Eur. Phys. J. H 42, 95–105 (2017)
DOI: 10.1140/epjh/e2017-80023-3

THE EUROPEAN PHYSICAL JOURNAL H

Editorial

Editorial introduction to the special issue "The Renaissance of Einstein's Theory of Gravitation"

Alexander Blum[1], Domenico Giulini[2], Roberto Lalli[1,a], and Jürgen Renn[1]

[1] Max Planck Institute for the History of Science, Berlin, Germany
[2] Institute for Theoretical Physics, Leibniz University of Hannover, Hannover, Germany

Received 18 April 2017 / Received in final form 4 May 2017
Published online 9 June 2017

On November 25th 1915, Albert Einstein submitted to the Royal Prussian Academy of Sciences the last of a series of papers that contained the final and fundamental equation of his theory of gravitation, which he called *General Relativity* (Einstein 1915, 1916). This equation contains the field-theoretic law according to which the energy-momentum distribution of matter sources acts on and reacts to the gravitational field. It was the final achievement of an "intellectual odyssey," which lasted more than eight years (Renn 2007; Gutfreund and Renn 2015).

The present issue is dedicated to the centenary anniversary of this momentous scientific achievement through a series of contributions that investigate the historical trajectory of Einstein's theory of gravitation. In spite of the celebratory character of the issue, we decided not to focus on the early history of General Relativity. Einstein's own path toward the theory, its reception in different national scientific communities and the further progress until the early 1950s have been discussed in an enormous amount of scholarly work during the past decades. Instead, we prefer to take this opportunity to explore in more detail the post-World War II developments of the theory, which only recently has become the subject of a lively debate among historians of science and physicists actively working on General Relativity and closely related fields.

This issue of EPJH aims to give new insights into the historical process through which Einstein's theory of gravitation came to turn into that fruitful and exciting branch of the physical research we know today. This process looked so splendid to some of the protagonists that physicist Will (1986, 1989) dubbed it the "Renaissance of General Relativity". But what is meant exactly by Renaissance? What kind of complex process does the term try to describe? Was it a mere consequence of the general growth of physics in the post-WW II period? Or did the phenomenon entail deeper epistemic transformations?

[a] e-mail: rlalli@mpiwg-berlin.mpg.de

From a superficial perspective, the history of Einstein's theory of gravitation might look like an inevitable success story. The recent detection of gravitational waves (Abbott et al. 2016), predicted by Einstein a century ago, have once more impressively underlined the central role that General Relativity will play in our understanding of fundamental interactions and cosmology. Today every physics student is told that General Relativity is one of the pillars of modern physics, together with quantum mechanics and quantum field theory.

But this has not always been the case. After an initial burst of excitement following the 1919 announcement that one of the few predictions of Einstein's theory–the gravitational deflection of light rays–had been confirmed, the theory underwent a period of stagnation, which lasted from the mid-1920s to the mid-1950s. Historian of physics Eisenstaedt (1986, 1989), who was the first to study the long-term history of General Relativity, called this phase the "low-water-mark" period. During this period only a few scientists worked on a theory that was seen by the majority of the physics community as mathematically extremely expensive with very little physical yield. And even physicists with undeniable strong mathematical inclination, like Pascual Jordan, were initially appalled by the "mismatch between the simplicity of the physical and epistemological foundations and the annoying complexity of the corresponding thicket of formulae" (Jordan 1952, p. 5).

Most of the invested work was seen to yield only either formal improvements or minor corrections to Newtonian predictions. As a result, the majority of theoretical physicists around the mid-1920s lost interest in the theory and preferred to focus on the far more exciting development of quantum mechanics, its plethora of applications to micro- and solid-state physics, which gave rise to much stronger and more fruitful connections with experimental activities and, last but not least, also promised much better career prospects. So, for a long time, a neo-Newtonian interpretation of General Relativity prevailed as the dominant attitude, where General Relativity was viewed merely as providing small corrections to Newtonian gravity, neglecting its fundamental aspects like the unification of the inertial and gravitational fields altogether.

An example of this attitude is that the physical meaning and domain of applicability of the full exterior Schwarzschild solution (including the horizon) remained unclear until the 1960s. There was a great amount of confusion as to whether the event horizon contained in this solution corresponded to a real spacetime singularity or whether its apparently singular nature was merely an artifact of an unsuitable choice of coordinates (Eisenstaedt 1987). This is not to say that during the low-water-mark period there was no important work on the Schwarzschild solution. Eisenstaedt himself and Luisa Bonolis in this volume show that researchers made significant progress and offered a number of insights on this issue, in some cases with direct connections to physical applications. The most important was certainly the work of Robert Oppenheimer and his co-authors on the application of General Relativity to stellar collapse in 1939 (Oppenheimer and Snyder 1939; Oppenheimer and Volkoff 1939). Nonetheless, these important advances did not become part of the shared knowledge of the experts on General Relativity. The criteria scientists used to evaluate the significance of specific advances and to define which were the important questions to be addressed varied considerably. What in hindsight could be considered important results were often ignored, and some of them remained controversial for decades. The relevance of General Relativity for the discipline of physics as a whole was also cause of disagreement. Oppenheimer himself strongly encouraged students and younger researchers to work on topics different from General Relativity. He did so in the mid-1950s, when another authoritative theoretical nuclear physicist, Wheeler, was instead making General Relativity the main focus of his research agenda. Such a fate of grossly diverging attitudes was also suffered by the theory of gravitational radiation. Here confusion reigned even as to whether gravitational waves were physically real, e.g., in the sense that they can carry energy from the source to the distant observer. Quite remarkably, Einstein himself came to doubt their physical existence in the 1930s (Kennefick 2007, pp. 79–104).

By the 1970s, the status of Einstein's theory of gravitation was completely different: the theory was perceived as an important, empirically well tested branch of theoretical physics, which had also produced a brand new and successful sub-discipline: relativistic astrophysics. It is important to note that the impact of General Relativity onto astrophysics was by far not exhausted by quantitative corrections, but also, and more essentially, by its addition of new qualitative features, e.g., concerning the structure and stability of stars, the formation of Black Holes, the emission of gravitational waves, and gravitational lensing as tools for mass detection. With at least one notable exception (Goenner 2017), most historians of science and physicists agree that sometime by the end of the 1960s a significant process had occurred, which might be described as a renaissance of the theory (see also Thorne 1994; Kaiser 2000; Kragh 2002; Kennefick 2007). In addition, the intimate connection of General Relativity with various mathematical branches with no previously established close connection with physics, like non-Riemannian differential-geometry and differential- and point-set topology, ceased be perceived as mere excess baggage. Rather, it turned into a positive aspect connected with the hope that new insights will emerge at the interface between mathematics and physics, eventually to the advantage of both sides. As an example we mention the 1967 *Battelle Recontres* lectures in mathematics and physics (DeWitt and Wheeler 1968), which, amongst others, brought together eminent mathematicians with no previous record in relativity, like Raoul Bott, Paul Federbush, Sigurdur Helgason, Stephen Smale, and Norman Steenrod, with physicists and mathematicians who had already worked in the field of relativity, including Brandon Carter, Yvonne Choquet-Bruhat, Bryce DeWitt, Cécile DeWitt-Morette, Robert Geroch, Stephen Hawking, André Lichnerowicz, Roger Penrose, Tullio Regge, John Wheeler and James York. The belief in a fruitful interaction was expressed succinctly on the cover of the proceedings volume, the title of which ends with $M \cap P \neq \emptyset$.

The descriptions of what this renaissance was, however, vary considerably, but most of them share a specific bias concerning the historical development of scientific theories: Next to Newton's Classical Mechanics, Einstein's General Relativity is often regarded as the prototypical example of a breakthrough in scientific theory associated with a framework created by a single ingenious scientist, on which all later developments are built, filling in the details without the need to revise the foundation. Consequently, the further development of the framework can only consist in integrating novel empirical evidence, working out the implications of the fundamental equations, and the introduction of new calculational techniques. Accordingly, the framework itself has no history of its own and fundamental progress can only occur through major upheavals. This view of theory (non-)development thus matches perfectly with the common reading of Kuhn's theory of scientific revolutions as radical paradigm changes, followed by long periods of normal science consisting of puzzle-solving. The latter are usually less important when trying to locate the decisive strategic moments in the history of theory formation. In reviewing, e.g., the history of gravitational waves on the occasion the recent discovery, the one constant is Einstein's "prediction" of 1916, with some puzzles and ambiguities resolved along the way. The debates concerning the existence of gravitational waves, which went on for at least 40 years, are then, also by historians of science, generally viewed as being a mere "comedy of errors" caused by lack of empirical evidence, lack of funding, disciplinary divides, lack of communication, and even personal idiosyncrasies.

This bias now severely restricts the scope of how one can interpret the evident turning point that is usually referred to as the Renaissance of General Relativity: It might be due to the influx of new empirical evidence, made possible by novel technologies, it might be an almost trivial consequence of the postwar political situation, with the unprecedented flow of money into virtually every branch of physics, or it might be due to the solution of a particularly persistent puzzle, which had been a bottleneck for further progress. Within this scope, it remains, however, difficult to explain how it came to (a) a burst of theoretical advances in several unrelated areas of General Relativity, which occurred (b) years before the major discoveries of 1960s radio-astronomy (the discovery of quasars in 1963, of the Cosmic Background Radiation in 1965, and of pulsars in 1967). And even if this sudden eruption were brushed aside as a mere coincidence, the question would remain why up to this turning point, so many major figures in General Relativity persistently mistrusted the qualitatively new implications of the theory, such as the possibility of unstoppable gravitational collapse or, to return to our example, the existence of gravitational waves.

These puzzles indicate, in our view, that the entire idea of scientific progress informing this narrative is seriously deficient. It seems more plausible, instead, that theories do have history and that their history does not consist merely in puzzle-solving. Indeed, if we admit that, in the Renaissance period, the conceptual foundation of General Relativity itself underwent a development, we can explain (a) why so many persistent problems suddenly became solvable within a relatively short time period; (b) why General Relativity was transformed from a marginal theory, primarily of mere philosophical and mathematical interest, into a vibrant field of research; and (c) how relativists were able to react so quickly to the unexpected experimental breakthroughs in astrophysics. This change did not consist in a *modification* of the foundation laid by Einstein: the Einstein Equation remains the cornerstone of General Relativity to this day. Rather, we are looking at an *extension* of the foundation: The theory of 1915 was insufficient to reach firm conclusions without being complemented by intuitions drawn from the resources of pre-relativistic physics or (for the case of cosmology) by philosophical considerations that were hardly generalizable to more mundane problems. Finding a general way to extract the physical content of the theory first became a major concern in the Renaissance years, with many papers opening with remarks concerning the difficulty of interpreting General Relativity. Only after the central issues had been resolved in the Renaissance was General Relativity applicable to any given physical problem, providing an interpretation in its own terms. The Renaissance was thus not a mere agglomeration of isolated results, but a global transformation in the character of the theory. Such a global transformation, which in Kuhnian terms might be described as a "paradigm shift," was hence not the premise, but rather the result coming at the end of a long period of problem-solving *within* General Relativity.

While historically-minded physicists have long been interested in the establishment of general relativity by Einstein in the 1910s, we hope that the papers in this special issue will demonstrate that the Renaissane period is of similar importance to those interested in the conceptual foundations of general relativity and its historical development (see also the programmatic articles by three of the editors, Blum et al. 2015, 2016). Given the focus of this journal, the authors in this volume have addressed from different perspectives one central aspect of the process of the Renaissance of Einstein's theory of gravitation: Its establishment as a field of study in its own right within the *discipline of physics*, rather than as an object of mere philosophical or mathematical analysis. It is the return of General Relativity to the mainstream of physics that the authors have discussed in different and pertinent cases and using a variety of approaches, some of which are somewhat different with respect to the style of papers that usually appear in the EPJH. The order of papers is more a conceptual than a chronological one, as the authors have discussed quite different aspects that cannot be easily considered as following a purely chronological progressive development.

The first paper in this special issue, by Alexander Blum and Thiago Hartz, focuses on the role of the program of constructing a quantum theory of gravity in the renaissance of General Relativity. They do this through a close reading and contextualization of a heretofore unpublished historical document, a report on a workshop on the quantization of the gravitational field, held in Copenhagen in the summer of 1957, published here for the first time. Held several months after the famous Chapel Hill conference, the Copenhagen workshop was arguably the first ever scientific meeting dealing solely with question of quantizing gravity and provides us with a unique glimpse at the role that this problem played in the physics of the time in general, and in the Renaissance of General Relativity in particular.

The notion of quantizing General Relativity carried with it an air of "domestication," whereby General Relativity would be brought from the realm of classical field theory (which the attempts by Einstein and others to construct a unified field theory did not transcend) into the domain of quantum theory, which formed the basis for most of the work in the physics mainstream, from solid state to high-energy nuclear physics. As the authors outline, it was Bryce DeWitt (the author of the report) who attempted to integrate these scattered attempts at domestication into the emerging renaissance community, by bringing together various approaches and finding a common agenda. As is well-known, the attempts at constructing a theory of quantum gravity have to this day not met with ultimate success, and there is still no universally accepted approach to the problem. The authors thus present DeWitt's attempt to find such a common agenda as a failed attempt to bring General Relativity back into the physics mainstream by strengthening its ties to high-energy physics and quantum field theory. This "physicalization" of General Relativity instead happened several years later solely via its connections to astrophysics and astronomy, while its relation to quantum theory remained elusive.

The theoretical discourse was in any case not the only way in which the General Relativity returned to the field of physics; there was also an exponential increase in experimental activities aimed at testing the predictions of the theory. Peebles' article in this volume reviews the early attempts to establish the field of the experimental study of gravity in the decade between the late 1950s and the late 1960s. Peebles argues that the growth was so impressive that one can well say that the field was actually born in that period. This is why he named the process the "naissance of experimental gravity physics". Building on his deep firsthand knowledge of the field, Peebles' review covers all the relevant scientific activities in experimental gravity physics of the period, although its main focus remains the pioneering work of Robert Dicke and of the group Dicke established in Princeton, where Peebles himself earned his PhD. The activity of this group, Peebles shows, had a relevant role in sparking this kind of research and strengthening its position as a relevant part of the physics endeavour. As for the historical factors underlying the process of the "naissance of experimental gravity physics," Peebles especially stresses the relevance of technological advances. New technologies were an essential component, for they allowed to draw unprecedented connections between the theory and the physical world. But by focusing on Dicke's own trajectory and his decision to change direction in mid-career, Peebles also shows that other factors, which cannot be reduced to the new possibilities opened by technological advances, played a similarly important role, such as the formulation of alternative theories of gravitation (such as the Brans-Dicke theory, also known as the Jordan-Thiry-Brans-Dicke theory) that provided both a theoretical background and the motivation to perform experiments designed to provide a crucial support to one of the competing theories, or the ability of individual experimenters to make use of the advancements in technology or in fields different from their own area of expertise. From the various different factors that shaped the early history of experimental gravity physics, Peebles draws some general lessons that are intended as food for thought for active experimental physicists.

Besides experiments designed to putting the theory to the test, the experimental-observational status of the general theory of relativity changed completely in the renaissance period. The two aspects of this change are in the area of gravitational-wave research and in the field of relativistic astrophysics. To these two subjects are devoted the last three articles of the volume. The experimental activity aimed at detecting gravitational waves exploded in the 1970s and there is good evidence that this activity was mostly a consequence of Weber's 1969 announcement that his attempts in this direction had been successful (Weber 1969, 1970). Up to this moment, Weber had pursued this activity alone or in collaboration with a few students and assistants. After the announcement, he had to face a number of controversies with his peers, who started to distrust his results after about 1972/73 (Collins 2004). To this dramatic, visionary figure, who may well be argued to have not obtained the credit he deserved during the time of his career, is dedicated a personal recollection by astronomer and historian of physics Virginia Trimble, who was also Weber's wife for the final twenty-eight years of his life.

The focus on the personal trajectory of a visionary and controversial scientist reminds us that the history of physics cannot simply be understood as the progressive accumulation of knowledge, but that several factors enter the development of science, some of them of non-scientific nature. And this is especially true in the case of controversies, where the debate is not only between individuals, but involves different social groups defined by disciplinary boundaries or different training. This is the perspective proposed by Daniel Kennefick in his paper "The Binary Pulsar and the Quadrupole Formula Controversy". In his book on the history of the theoretical quest on gravitational radiation, Kennefick (2007) had already shown that these theoretical developments were clouded by disagreements as to whether gravitational waves existed and which properties they had. The consensus on the fact that energy-carrying gravitational waves existed was in fact achieved only during the renaissance period. Even after the majority of theoreticians working on the theory of General Relativity came to accept the physical reality of gravitational waves, several aspects continued to remain matters of debate. Among them, the most pressing were whether binary systems decay because of gravitational damping and the related question as to whether the quadrupole formula first derived by Einstein (1918) in the linearized approximation scheme was the correct way to deal with this problem. Kennefick's contribution to this volume deals with the development of this theoretical controversy and the twofold role of the discovery of the binary pulsar system PSR 1913+16 in 1974 and the subsequent observation of its decay, firstly in sharping the theoretical controversy and then in closing it. In his narrative, Kennefick focuses in particular on the epistemic aspects related to the social separation of the involved scientists in different communities: that of application-/calculation-minded physicists on one side, and that of physicists with a strong commitment to mathematical rigor on the other. Kennefick stresses the relevance of this kind of social structure for the way in which such controversies evolve.

Finally, the birth of relativistic astrophysics, the crowning element of the return of General Relativity to the mainstream of physics, is addressed in the paper by Luisa Bonolis entitled "Stellar structure and compact objects before 1940: Towards relativistic astrophysics". The author looks at this process from the perspective of the continuity of research on astrophysical compact objects that became an active branch of research since the mid-1920s, after the development of quantum mechanics and its application to solid state physics. In this issue, Bonolis presents the first part of a two-part paper on the history of the astrophysics of highly dense objects up to its transformation into relativistic astrophysics during the 1960s. In this first part presented here, the early phases of these research activities up to the beginning of the Second World War are investigated. According to Bonolis, only by following the long-term development of the theoretical study of the stellar structure of compact objects and the connection of this research with contemporary developments in other branches of theoretical physics is one able to properly understand what is considered as an essential element of the"physicalization" of Einstein's gravitational theory; namely, the application of the theory to solve the physical problems related to the newly discovered astrophysical objects of quasars and pulsars, which since their discoveries required the application of a non-Newtonian theory of gravitation.

While the articles in this volume shed lights on important aspects of the process to which the volume is dedicated, the phenomenon was so huge and complex that many other aspects still remain to be investigated by historians of science and physicists. For starters, the list of topics that were at the focus of the renaissance is of course by no means exhausted by the papers presented in this issue. One such important issue is the problem of motion in General Relativity: Einstein's equations imply by way of integrability conditions a kind of local conservation concerning the energy-momentum exchange between the gravitational field and matter, thereby strongly restricting, or sometimes even determining, the dynamics of the latter. These integrability conditions may jeopardize consistency of approximation schemes if not properly taken into account. First advances were made already in the late 1930s through the work of Einstein et al. (1938, see also Havas 1989) but the problem remained largely neglected until it was unignorably put back onto the agenda of theoretical astrophysicists in the mid-1970s (Ehlers et al. 1976) where it remained as an active field of research ever since (Blanchet et al. 2011; Pützfeld et al. 2015). The problem of motion for extended bodies, in the course of the renaissance and beyond, also represents a promising field for historical research; see, e.g., Dixon's contribution "The New Mechanics of Myron Mathisson and Its Subsequent Development" in Pützfeld et al. (2015).

Another central conceptual development of the renaissance is a sharpening of the notion of singularity, culminating in the singularity theorems of Roger Penrose and Stephen Hawking in the mid-1960s (Penrose 1965; Hawking and Penrose 1970). The question of singularities provides an enticing case study for general questions of theory and concept development. The notion of a singularity had been around for a long while and had served as a criterion for excluding solutions of the Einstein equations as physically irrelevant. Such arguments had been used, e.g., in the 1930s to deny the existence of plane gravitational waves, and also to discredit simplified models of gravitational collapse or big-bang cosmology. One important renaissance development was the definitive disentangling of coordinate and genuine (non removable) singularities; this was just as much a sociological advance as it was a conceptual one, given the fact that many physicists and mathematicians appear to have been clear on this matter already in the 1930s, but that it only became generally accepted once there was an established community out of which such insights could spread. Here historians have the opportunity to study how a research field develops a memory and thus why, after the renaissance, John Stachel's dictum that everything worth discovering in General Relativity was discovered at least twice might no longer hold (Stachel 1992).

But going beyond this disambiguation, there was also the problem of finding a positive characterization of a proper (non coordinate) singularity that also made some intuitive physical sense. The community would ultimately agree on a definition as used by Penrose, Hawking and Geroch based on the notion of geodesic incompleteness, but this was still a compromise lacking many desirable features (Geroch 1968). Here the task for future historians could be to trace and understand its origins and the reasons for (and the conceptual impact of) its ultimate general acceptance, despite its flaws. This raises further more general questions on how the emerging community agreed upon matters such as formal definitions or standards of argument and mathematical rigour, and what impact formal-mathematical work had, e.g., on the inquiries into actual physical black holes.

Similarly probing epistemological questions may well be asked also concerning those topics that *are* addressed in the current volume. For the history of science progresses not just by broadening its scope, but also by delving ever deeper into the dynamics of the evolution of science. Detailed reviews, often written by physicists, such as Jim Peebles' contribution to this volume or Malcolm Longair's magisterial book *The Cosmic Century* on the history of astrophysics, often form the starting point for the kind of historico-critical analysis that historians of science (at least those with a strong interest in the actual content of the science under study) aspire to. Questions such as the ones we formulated for the case of singularities are still unanswered also for the case of one of the central topics of this issue, gravitational waves: How did formal-mathematical existence proofs for wave solutions of the Einstein equations (initiated by Felix Pirani and Andrzej Trautman in the late 1950s) actually relate to the more physical questions of gravitational wave emission and absorption discussed in this issue?

This question finally brings us to the third dimension in which the future historiography of General Relativity should and will progress: Besides the thematic broadening and the epistemological deepening, we also have simple chronological progression. The renaissance for all its import is certainly not the end of the exciting history of the singular theory that is General Relativity. All of the stories told and questions raised and answered in this issue can be extended towards the present. We have already outlined for the case of gravitational waves the challenge of pursuing and understanding the development that led from total skepticism about gravitational waves in the mid-1950s to their ultimate discovery 60 years later. Similarly, we see in the article by Blum and Hartz that in 1957 there was still considerable optimism that the construction of a quantum theory of gravity would be achieved within the next few years. When and for which reasons did this change and how did the different approaches to quantum gravity interact with and affect the General Relativity community at large in the following decades? In formulating these questions one can already see that we are here entering a territory that historians of science tend to eschew: Dealing with the history of scientific research programs that have not yet reached an accepted definite conclusion.

One argument often brought forth is that historians should not meddle in scientific matters that are still unresolved. This is hardly a tenable position: Political historians are very much engaging with the history of the 1970s, say, dealing with social and geopolitical issues that are by any standard unresolved to this day. And even if one would want to argue that history of science is somehow different in this regard, science can hardly be so neatly compartmentalized into solved and unsolved problems. Attempting to do so would lead to a highly incomplete picture of the scientific development in the second half of the twentieth century, or to an ultimately superficial understanding of merely the institutional and social aspects of modern science. The renaissance of General Relativity is, in fact, a perfect illustration of the fact that these institutional and social developments cannot be understood in isolation, but only through their intimate interaction with the content of science. For this reason, we conclude this introduction by stressing that there are plenty of open physical questions that the renaissance generation handed down to us in the hope that they may pique the interest of young physicists and historians of science alike.

Amongst the open issues is, of course, the overarching problem of how to reconcile the theoretical description of gravity with that of the remaining fundamental interaction. In absence of obvious phenomenological data the physical need for such a unification is usually seen in genuine predictions of General Relativity, like that of singularities and other spacetime features considered to be pathological. Currently it is rather unclear of whether this unification process will be more like a "quantization of gravity" or rather a "gravitization of Quantum Theory" (Penrose 2014a,b) or, perhaps most likely, both at the same time. We know that the impact of relaxing the spacetime symmetries of Special Relativity (Poincaré group) onto our familiar concepts of Quantum-Field Theory is dramatic: no particles, no vacuum, no scattering theory! There is so far no well founded intuition regarding proper replacements that might work beyond semi-classical situations. And even in semi-classical contexts the coupling of the classical gravitational field (the metric) to ordinary quantum mechanical matter is not understood in detail and from first principles. How does an atom react to the non-newtonian components of the Einsteinian gravitational field (i.e. gravitomagnetism and gravitational waves) and how does a non-classical delocalized state of a quantum system source a gravitational field? How do we formulate the equivalence principle in a way that applies to quantum matter (no point particles, no world lines, no clocks)? These are apparently mundane questions which have so far not received accepted answers.

Sill further down, on a purely classical level, hard technical problems remain. The full (non-linear) stability of even the simplest non-trivial solutions, like that of Schwarzschild and Kerr, are unknown. Sound approximation schemes in cosmology are lacking, which means that we do not know how to properly derive the Friedmann equations (which underlie our cosmological standard model) as controlled approximation to the full Einstein equations in situations with only approximate homogeneity. Calculations of gravitational radiation-reaction upon the sources are much more difficult than in the linear case of electrodynamics and certainly plagued with the same pathologies (runaway solutions). And even without radiation reaction, the analytical treatment of the equations of motion of structured (and necessarily spatially extended) bodies in strong gravitational fields is still not in a satisfying form, free of mathematical tricks and uncontrolled approximations (Blanchet et al. 2011; Pützfeld et al. 2015).

This list could easily be continued. But the fact that hard theoretical problems remain unsolved should not mislead the reader into thinking that this theory – General Relativity – lacks precision and foundation. Quite the opposite: The bigger the success the more ambitious our criteria for proper understanding become and develop. Modern precision tests in astrophysics and cosmology show that General Relativity can clearly bear comparison with the best predictive theories in all of physics. Moreover, the mathematical formulation of its physical and epistemological foundations is certainly no less adequate, and presumably even better, than that of any other fundamental theory in physics. We therefore feel entitled to predict that its future will be as astonishing and revealing as its past has already been.

References

Abbott, B.P. et al. 2016. Observation of Gravitational Waves from a Binary Black Hole Merger. *Phys. Rev. Lett.* **116**: 061102-1–061102-14

Blanchet, L., A. Spallici and B. Whiting. Eds. 2011. *Mass and Motion in General Relativity.* Springer, Dordrecht

Blum, A., R. Lalli and J. Renn. 2015. The Reinvention of General Relativity: A Historiographical Framework for Assessing One Hundred Years of Curved Space-time. *Isis* **106**: 598–620

Blum, A., R. Lalli and J. Renn. 2016. The Renaissance of General Relativity: How and Why It Happened. *Annalen der Physik* **528**: 344–349

Collins, H. 2004. *Gravity's Shadow.* Univ. of Chicago Press, Chicago

DeWitt, C. and J.A. Wheeler. Eds. 1968. *Battelle Recontres – 1967 Lectures in Mathematics and Physics* – $M \cap P \neq \emptyset$, W.A. Benjamin, Inc., New York

Ehlers, J., A. Rosenblum, J. Goldberg and P. Havas. 1976. Comments on gravitational radiation damping and energy loss in binary systems. *The Astrophysical Journal* **208**: L77–L81

Einstein, A. 1915. Feldgleichungen der Gravitation. *Preuss. Akad. Wiss. Berlin, Sitzungsberichte der Physikalisch-mathematischen Klasse*: 844–847

Einstein, A. 1916. Die Grundlage der Allgemeinen Relativitätstheorie. *Annalen der Physik* **354**: 769–822

Einstein, A. 1918. Über Gravitationswellen. *Preuss. Akad. Wiss. Berlin, Sitzungsberichte der Physikalisch-mathematischen Klasse*: 154–167

Einstein, A., L. Infeld and B. Hoffmann. 1938. The gravitational equations and the problem of motion. *Ann. Math.* **59**: 65–100

Eisenstaedt, J. 1986. La Relativité Générale à l'Étiage: 1925–1955. *Archive for History of Exact Sciences* **35**: 115–185

Eisenstaedt, J. 1987. Trajectoires et Impasses de la Solution de Schwarzschild. *Archive for History of Exact Sciences* **37**: 275–357

Eisenstaedt, J. 1989. The Low Water Mark of General Relativity. In *Einstein and the History of General Relativity, Einstein Studies, Vol. 1*. Ed. Don Howard and John Stachel. Birkhäuser, Boston, pp. 277–292

Geroch, R. 1968. What is a Singularity in General Relativity? *Annals of Physics* **48**: 526–540

Goenner, H. 2017. A golden age of general relativity? Some remarks on the history of general relativity. *General Relativity and Gravitation* **49**: 42. doi:10.1007/s10714-017-2203-1

Gutfreund, H. and J. Renn. 2015. *The Road to Relativity*. Princeton Univ. Press, Princeton, NJ

Havas, P. 1989. The Early History of the "Problem of Motion" in General Relativity. In *Einstein and the History of General Relativity, Einstein Studies, Vol. 1*. Ed. Don Howard and John Stachel. Birkhäuser, Boston, pp. 234–276

Hawking, S.W. and R. Penrose. 1970. The singularities of gravitational collapse and cosmology. *Proceedings of the Royal Society London Series A* **314**: 529–548

Jordan, P. 1952. *Schwerkraft und Weltall*, Vieweg & Sohn, Braunschweig

Kaiser, D. 2000. Roger Babson and the Rediscovery of General Relativity. In *Making Theory: Producing Theory and Theorists in Postwar America*. Ph. D. dissertation. Harvard Univ. Press, Cambridge, MA, pp. 567–595

Kennefick, D. 2007. *Traveling with the speed of thought: Einstein and the quest for gravitational waves*. Princeton Univ. Press, Princeton, NJ

Kragh, H. 2002. *Quantum generations: A history of physics in the twentieth century*. Princeton Univ. Press, Princeton, NJ

Oppenheimer, J.R. and H. Snyder. 1939. On continued gravitational contraction. *Physical Review* **56**: 455–459

Oppenheimer, J.R. and G.M. Volkoff. 1939. On massive neutron cores. *Physical Review* **55**: 374–381

Penrose, R. 1965. Gravitational collapse and space-time singularities. *Physical Review Letters* **14**: 57–59

Penrose, R. 2014a. On the Gravitization of Quantum Mechanics 1: Quantum State Reduction. *Foundations of Physics* **44**: 557–575

Penrose, R. 2014b. On the Gravitization of Quantum Mechanics 2: Conformal Cyclic Cosmology. *Foundations of Physics* **44**: 873–890

Pützfeld, D., C. Lämmerzahl and B. Schutz. Eds. 2015. *Equations of Motion in Relativistic Gravity*. Springer, Dordrecht

Renn, R. Ed. 2007. *The Genesis of General Relativity*. Springer, Dordrecht

Stachel, J. 1992. The Cauchy problem in general relativity-the early years. In *Studies in the History of General Relativity, Einstein Studies, Vol. 3*, edited by J. Eisenstaedt and A.J. Kox. Birkhäuser, Boston, pp. 407–418

Thorne, K. 1994. *Black Holes and Time Warps: Einstein's Outrageous Legacy*. Norton, New York

Weber, J. 1969. Evidence for the discovery of gravitational radiation. *Phys. Rev. Lett.* **22**: 1320–1324

Weber, J. 1970. Gravitational radiation experiments. *Phys. Rev. Lett.* **24**: 276–279

Will, C.M. 1986. *Was Einstein Right? Putting General Relativity to the Test*. Basic Books, New York

Will, C.M. 1989. The Renaissance of General Relativity. In *The New Physics*. Ed. Paul Davies. Cambridge University Press, Cambridge, pp. 7–33

Eur. Phys. J. H **42**, 107–157 (2017)
DOI: 10.1140/epjh/e2017-80015-8

THE EUROPEAN PHYSICAL JOURNAL H

The 1957 quantum gravity meeting in Copenhagen: An analysis of Bryce S. DeWitt's report

Alexander Blum[1,a] and Thiago Hartz[2,3]

[1] Max-Planck-Institut für Wissenschaftsgeschichte, Berlin, Germany
[2] Museu de Astronomia e Ciências Afins, Rio de Janeiro, Brazil
[3] Niels Bohr Archive, Copenhagen, Denmark

Received 9 March 2017 / Accepted 15 March 2017
Published online 1 June 2017

1 Introduction

Between June 15 and July 15, 1957, three physicists met at the Institute for Theoretical Physics – the renowned institute directed by Niels Bohr in Copenhagen, Denmark – in order to discuss quantum theories of the gravitational field. They were Bryce S. DeWitt, Stanley Deser, and Charles W. Misner. During the last eleven days of the meeting, they were joined by three other physicists, namely, Christian Møller, Oskar Klein, and Bertel Laurent.

As the meeting had been partially funded by the US Air Force, when DeWitt returned home, to North Carolina, he had to write a research report to the Wright Air Development Center – the US Air Force's research and development center located at the Wright-Patterson Air Force Base in Ohio. Such a report could go into quite some technical detail, as the Wright Air Development Center (or more specifically, the Aeronautical Research Laboratories) had recently hired a bona fide specialist, Joshua N. Goldberg, who was involved in funding and supporting many general-relativity-related activities from 1956 to 1963 (Goldberg 1992).

DeWitt finished the first version of the report, which has 9 pages, on 31 July 1957; the second and final version, which was considerably enlarged and incorporated comments from the other participants, was completed on 8 October 1957. As far as we know, there is only one remaining copy of the first version of the report and two remaining copies of the second version, one handwritten and one typeset. The first version of the document and the handwritten copy of the second version can be found in possession of Cécile DeWitt-Morette at the University of Texas at Austin. The typed copy of the second version of the document is deposited at the Aage Petersen Collection of Reprints and Manuscripts at the Niels Bohr Library and Archives of the American Institute of Physics, College Park, MD, United States. This last version

[a] e-mail: ablum@mpiwg-berlin.mpg.de

is the one that is reproduced in *Eur. Phys. J. H*, Doi: 10.1140/epjh/e2017-80016-0 and that we introduce and comment on in the next pages.

The relevance of DeWitt's report lies in the amount of background information it gives on the development of quantum gravity at this critical moment: Indeed, one can say that quantum gravity emerges as a research field in its own right at just this time. The Copenhagen workshop is the first scientific meeting dedicated solely to the problem of quantum gravity. And despite the very limited attendance, the report shows the interactions between various approaches and attitudes to the problem and can thereby also serve as an inspiration for today's highly segregated quantum gravity communities. In order to appreciate the full relevance of the report, it is important to have a clear picture of the development of quantum gravity until 1957. That is the goal of Section 2, where we give a broad panorama of the research on quantum gravity before 1957 and also introduce the actors and the tradition of larger conferences in general relativity in which the Copenhagen workshop is to be situated. In Section 3, we reconstruct the planning of the 1957 meeting. In Section 4 – which is the main part of the article – we present a careful analysis of DeWitt's 1957 report. Finally, in Section 5 we offer our conclusions.

2 General historical context

The broad historical context in which the Copenhagen workshop must be placed is the renaissance of general relativity in the 1950s. Recently, a growing number of historians are improving our understanding of this transition, from the marginalization of general relativity between the mid-1920s and the mid-1950s to the establishment of a consolidated area of research in the late 1950s. It is now clear that scientific, institutional, and generational aspects played a role in the establishment of the new community (Blum et al. 2015). In this section, we present those aspects that are most germane to quantum gravity in general and to our story in particular.

2.1 Research on quantum gravity before 1957

Before discussing the history of "quantum gravity", we need to briefly discuss our use of this term. It is, flat out, anachronistic. There is so far no historical analysis of the question when, how, and why the term arose, but it most certainly arose years after the Copenhagen meeting and is entirely absent from the report[1]. We use it merely for convenience – the authors' and the readers' – and take it to be synonymous with the "quantization of the gravitational field" of the report's title. This means, in particular, that it does not cover any kind of more general quantum theory that is supposed to reduce to a quantization of general relativity in some low-energy limit, such as modern-day string theory. Such theories were not yet on the horizon in 1957 and they would not have been deemed necessary: The use of terms such as "quantum *gravi*dynamics" clearly shows that the quantization of gravity was still thought to be essentially analogous to the quantization of electrodynamics, i.e., quantum *electro*dynamics, as the few scattered early works on quantum gravity had not yet revealed the full extent of the conceptual and formal difficulties involved. In the following we will briefly discuss these few scattered works[2].

[1] This is borne out by a simple Google Ngram search, which finds the first usage of the term "quantum gravity" in 1969. Interestingly, in the 1990s, the frequency of "quantum gravity" then already surpasses that of "quantum electrodynamics".

[2] The interested reader is referred to a volume currently in press by Dean Rickles and one of the authors which presents early sources in quantum gravity along with a historical contextualization (Blum and Rickles 2017).

When quantum mechanics and quantum field theory had been originally developed in the second half of the 1920s, there was only one quantization procedure, *the* quantization procedure, namely *canonical quantization*. When Léon Rosenfeld, then assistant of Wolfgang Pauli, one of the founders of the quantum theory of fields, made the first short-lived forays into the quantization of general relativity in 1930, he naturally applied this quantization procedure both when discussing the gauge invariance properties of the full theory (Rosenfeld 1930a) and when calculating actual toy problems in the linearized (or weak-field) approximation (Rosenfeld 1930b), where the metric is split up into $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$, where $\eta_{\mu\nu}$ is the Minkowski metric and $h_{\mu\nu}$ is taken to be a (small) perturbation[3].

In the ensuing decades, new quantization procedures developed, not necessarily, however, because of an interest in the foundations of quantum mechanics, but rather with a specific goal in mind, namely to cure the problems of quantum field theory, in the years before the development of renormalization procedures gave a handle on the problematic infinities. Almost invariably, such new procedures would be applied to the quantization of gravity within a few years. A case in point is a quantization procedure originally developed by Dirac (1932); it does not have a real name of its own, because in a sense it is simply equivalent to canonical quantization in the interaction picture (which Dirac introduced for the first time in the 1932 paper). It has, however, the advantage of being manifestly covariant and the disadvantage of *only* being applicable in the interaction picture, thereby necessitating a clean separation into a free field and perturbing interactions. We might thus call it *covariant perturbative quantization*. It consists of imposing covariant commutation relations on the free (i.e., interaction picture) field operators at two distinct space-time points, as opposed to the equal-time canonical commutation relations of canonical quantization.

The use of the interaction picture implied a severe restriction in the application of covariant perturbative quantization to gravity: It always had to rely on some sort of perturbation around a flat, Minkowski metric. Covariant perturbative quantization was first applied to linearized general relativity in the 1930s (Bronstein 1936)[4]. But this quantization procedure really only took off when it became the basis for the development of renormalized quantum electrodynamics in the late 1940s and thereby became the standard quantization technique within the high-energy nuclear/particle physics community[5]. In the wake of this great success, a number of physicists conjectured that a full theory of (covariant perturbative) quantum gravity could be constructed by, when expanding the metric in terms of the perturbation $h_{\mu\nu}$, keeping not only the linear term (as Bronstein had done), but also the infinite sum of higher-order, non-linear terms[6]. This was first proposed by Bryce DeWitt (1949) in his unpublished

[3] On Rosenfeld's pioneering work, see also (Salisbury and Sundermeyer 2016).

[4] A translation of this paper was printed as a "Golden Oldie" in General Relativity and Gravitation, see (Deser and Starobinsky 2012).

[5] See, e.g., (Tomonaga 1946) and (Schwinger 1948).

[6] This is closely connected to, though logically distinct from, the program of constructing the full theory of general relativity by starting from a free, massless spin 2 field in Minkowski space (first identified with a graviton by Fierz and Pauli (1939)) and then iteratively adding non-linear self-interaction terms through demands of self-consistency. Such a program was first pursued by Robert Kraichnan in his unpublished Bachelor's thesis (MIT, 1947). The thesis is not extant (it is not in possession of the MIT library nor of Robert Kraichnan's widow), and all we have to go by is the abstract of the thesis, which was kindly provided to the authors by Professor John Preskill of Caltech. But in a letter to his thesis advisor Hermann Feshbach (dated 11 July 1947, half a year after submitting his thesis), Kraichnan still could only assert: "I strongly suspect that the physical results of this formalism are identical or essentially similar to those of the general theory of relativity, but I haven't yet been able to get results one way or another". (The letter was in possession of Kraichnan's widow Judy

PhD thesis and then again by Suraj N. Gupta (1952a,b). This flat-space view was widely accepted in the high-energy community, as one can see from a remark made by Murray Gell-Mann – a leader of that community by the late 1950s – to Gregory Breit in June 1959 that he thought "very well of the Gupta-Feynman view regarding General Relativity with successive correction terms for a Spin 2 particle"[7]. This point of view only began to be called into question several years after the Copenhagen workshop, when Richard Feynman (1963) for the first time actually included the non-linear terms in an actual calculation, initiating a research program that culminated in the proof that perturbative quantum gravity is non-renormalizable and thus in the realization that the naive covariant perturbative quantization of general relativity does not work.

But already in the 1950s, the perturbative approach was far from uncontroversial and Gell-Mann's opinion was hardly universally accepted. There was the old guard who found the simple, formulaic unification of general relativity and quantum theory presented by Gupta to be a rather bland coda to the great conceptual upheavals of the early twentieth century[8]. But more important for our story is the opposition from the (at least somewhat) younger generation of "relativists" that led the renaissance of general relativity. They felt that the perturbative flat-space formulation was bound to lose essential aspects of the complex non-linear field theory that was general relativity. Indeed, the conviction that general relativity is in some sense special is probably one of the defining new beliefs of the renaissance period. The so-called low-water mark period of general relativity (from the mid-1920s to the mid-1950s) that preceded the renaissance was characterized by a "neo-Newtonian" interpretation (Eisenstaedt 1986, p. 148–158) of general relativity, which viewed the theory as only providing a small correction to Newtonian physics. In a similar vein, the perturbative approach to (quantum) gravity provided only small corrections to the flat Minkowski space-time of special relativity. This is to be contrasted with the point of view that gained traction in the renaissance of the 1950s, that especially the non-linearities of general relativity are the essential feature of the theory and lead to properties that are qualitatively different from both Newtonian gravity and special relativity.

One result that strongly promoted this point of view and greatly influenced some of the main proponents of quantum gravity in the renaissance had already been obtained in the late 1930s by Albert Einstein, Leopold Infeld, and Banesh Hoffmann (Einstein et al. 1938). They had shown that the gravitational field equations for empty space are sufficient to derive the equations of motion for a set of gravitationally interacting point-like masses (singularities). Eleven years later, Infeld and Schild (1949) were similarly able to prove the even more elementary result that the vacuum field equations determined the (geodesic) motion of a single point-like test particle in a given background field. These results were in stark contrast to the case of Maxwell

---

Moore Kraichnan, who kindly provided us with a copy. It is now with the Linda Hall Library for Science, Technology and Engineering, in Kansas City, MO.) His iterative derivation of general relativity was only published almost a decade later (Kraichnan 1955). By this time, the program had been proposed independently by Gupta (1954), who, however, only provided a heuristic sketch of a derivation in order to explain the analogies and disanalogies between his perturbative approach to quantum gravity and quantum electrodynamics. Nowadays, the constructive spin 2 program is mainly associated with Feynman, who based his lectures on gravitation on it (Feynman et al. 1995). Many years later, Deser (1987) argued that, in order to get the Einstein equations *with cosmological constant*, one could not simply start from a flat space background, but instead had to take an "*a priori* arbitrary background geometry".

[7] Gregory Breit Papers, Yale University, Notes on Yale Theoretical Physics Seminar.

[8] See, e.g., Wolfgang Pauli's letter to to Homi J. Bhabha, 26 July 1952, reprinted in (von Meyenn 1996, p. 682), where he describes Gupta's papers as "less interesting", despite the fact that they purported to solve a problem he had a longstanding interest in.

electrodynamics, where the Lorentz force cannot be derived from the vacuum Maxwell equations. This strongly suggested that the non-linear field theory such as general relativity could provide novel insights into the longstanding problem of integrating the dynamics of continuous fields with those of singular point particles which an essentially linear theory such as electrodynamics could never aspire to.

Following this lead, a new approach to quantum gravity was inaugurated by Peter Bergmann in the late 1940s. Dismissing the modern covariant quantization techniques, which treated the non-linearity of general relativity as mere perturbations, he proposed a return to a *canonical quantization* of the full non-linear theory. He envisaged a method for transforming the full non-linear Lagrangian into a Hamiltonian, which should then be (canonically) quantized, preserving in the quantum theory the essential features of the non-linearity of general relativity. The problem was to formulate a general method for obtaining a Hamiltonian from a Lagrangian with gauge symmetries, i.e., from the Lagrangian of a constrained system. In quantum electrodynamics and in the linearized gravitational theory, methods for dealing with this difficulty had been developed in the 1930s. It was, however, far from clear how to extend these methods to the non-linear case. The solution to this problem was developed independently by Peter Bergmann and his students (Bergmann et al. 1950), and by Pirani and Schild (1950), building on work by Paul Dirac (1950) and Paul Weiss (1938). That was the beginning of the canonical quantum gravity program. As opposed to the perturbative covariant approach, which in 1957 was viewed as either completed or intrinsically insufficient, canonical quantization of gravity was an active research program in 1957 and was extensively discussed at the Copenhagen workshop (Sect. 4.4).

While Bergmann proposed a return to the roots of quantum mechanics and the original canonical quantization procedure, new quantization techniques were proliferating. Misner, in an introductory passage – hailed by Carlo Rovelli (2004, Appendix B) as a moment of clarity in the history of quantum gravity – to his 1957 thesis on quantum gravity asserted that

> Four approaches have been suggested to discover the content of the quantum theory of general relativity... (Misner 1957, p. 497)

Besides the already mentioned canonical and perturbative covariant approaches, this list also included quantization using Schwinger's action principle and *Feynman* (or path integral, in modern parlance) *quantization*. And even this list was not exhaustive, omitting lesser known methods such as Peierls brackets.

The most prominent of the new quantization techniques was, however, certainly Feynman quantization, which was also the central focus of Misner's thesis. It had been developed by Feynman in the early 1940s, in an attempt to quantize classical theories which were not defined by a Hamiltonian or Lagrangian density, but only by an action. Such a classical theory had been devised by Wheeler and Feynman (1945) in order to eliminate both the infinite number of degrees associated with the electromagnetic field and the problematic self-interaction of charged matter, both of which were viewed as the classical origin of the divergence difficulties of quantum electrodynamics. Just like perturbative covariant quantization, Feynman quantization had thus been developed in the context of attempting to cure the woes of quantum field theory more generally. While Feynman never managed to quantize Wheeler-Feynman electrodynamics using his path-integral approach, his investigations did lead to the development of this new quantization method, which not only gave an interesting novel view of non-relativistic quantum mechanics, but also was of immense heuristic value in establishing the Feynman rules for regular quantum electrodynamics[9]. The idea to use Feynman quantization to quantize gravity was only a couple of years

[9] On the genesis of Feynman quantization, see, e.g., (Wüthrich 2010)

old in 1957 and had not progressed very far, but among the Copenhagen workshop participants it was probably the most popular approach overall (see Sect. 4.3).

2.2 Participants of the copenhagen meeting

The 1957 meeting took place in Copenhagen and thus it is no wonder that three of the participants (those three that only joined for the latter part) came from what might be called the Scandinavian tradition in general relativity. In Northern Europe, there had been persisting interest in general relativity from a few physicists since the late 1920s.

Christian Møller had become well-known due to his work on scattering theory – both electron-electron scattering and meson scattering – during the 1930s and 40s. He lectured at the Institute for Theoretical Physics in Copenhagen from the early 1930s until his death, in 1980 (Brevik 2011; Rozental 1985, p. 8). In the mid-1940s, he engaged himself in Heisenberg's S-matrix program, but put it aside to[10] write a highly regarded textbook (Møller 1952) based on his general relativity lectures during the 1930s and 40s, which became a classic text in general relativity (Wheeler and Ford 2000, p. 305).

One of the most important aspects in his teaching of general relativity was his attitude towards the four-dimensional formulation. He acknowledged the historical relevance of having recognized the similarities between space and time, but that was, according to him, not the best way of approaching the teaching of general relativity. Instead, he asserted one should "stress again the fundamental physical difference between space and time, which was somewhat concealed by the purely formal four-dimensional representation" (Møller 1952, p. v-vi). Such a program was envisaged for pedagogical reasons:

> The three dimensional point of view (...) leads to a reintroduction of dynamical concepts into the gravitational theory, which, I believe, makes it easier for the student fully to grasp the physical content of the general theory of relativity (Møller 1952, p. vi).

These pedagogical perspectives strongly influenced his research agenda, which began focusing on general relativity in the 1950s. He always emphasized the importance of studying how time was to be measured in general relativity and, therefore, dedicated most of his attention to analyzing the behavior of clocks in a gravitational field – as we will discuss in Section 4.1.

Oskar Klein had also been a member of the Copenhagen Institute for Theoretical Physics, from 1918 to 1923 and again from 1925 to 1930. In the latter period, he was Niels Bohr's assistant, as well as a leading figure in theoretical physics himself, obtaining important results such as the Klein-Gordon equation or the Klein-Nishina formula. In 1931 he became professor at Stockholm College (later Stockholm University), where he lectured on general relativity (Feyerabend 1995, p. 77; Fisher-Hjalmars and Laurent 1991). Stanley Deser (1977) observes that Klein was "one of the first who seriously considered general relativity in connection with quantum theory, an interest that remained with him from the time of the formulation of the Klein-Kaluza theory in 1926". After a long period of abstinence, he returned to the problem of quantum gravity in the mid-1950s, outlining the following research program:

> As a kind of programme we shall thus put forward the following claim: The operators to be used in quantum field theory should have a simple connection to a transformation group (so far insufficiently known) which contains

[10] Letter from Møller to Wolfgang Pauli, 13 November 1946, reprinted in (von Meyenn 1993).

> the general coordinate transformations in spacetime as a subgroup. The quantum conditions ought to characterize the group in question. In trying to develop a theory according to such a programme it should be kept in mind that the direct quantization according to the ordinary scheme of quantum mechanics of the Einstein equation meets with difficulties of the same type, but very much enhanced, as those met with in the quantization of the Maxwell equations. Also from this point of view it would seem preferable to start with the quantum conditions expressing group properties instead of starting with a Lagrangian density. This would probably make the theory still more symbolic and remote from direct observation than ordinary quantum field theory (Klein 1955, pp. 98–99).

This research program was further developed by **Bertel Laurent**. Laurent was Klein's last PhD student and was still working on his thesis on quantum gravity in 1957 (Deser 1995). The thesis, entitled "Studies in the Synthesis of Quantum Theory and General Relativity" was submitted on 5 December 1959.

Klein had also further explored the idea that the gravitational field might work as a regulator for quantum field theory:

> It is perhaps not unreasonable that the rigorous consideration of gravitational and perhaps other similar non-linear effects would do away with the remaining divergencies of electron theory. (Klein 1956, p. 61)

This went beyond the brute force regularization using a minimal length scale that had been explored already in the 1930s (Heisenberg 1938), and Klein's statement would have a deep influence on the 1950s generation – as we will discuss in Section 4.5.

The other three participants of the Copenhagen workshop came from the United States. There, general relativity had been practiced by a few physicists in the 1930s and 40s, to a large part due to European refugees. Until the mid-1950s, these researchers had virtually no interaction. This situation began to change, as several American physicists set up new centers of relativity research in the second half of the 1950s and, above all, due to the 1957 Chapel Hill conference (see Sect. 2.3).

Bryce DeWitt had been a PhD student of Julian Schwinger at Harvard. Schwinger himself was not interested in general relativity until the late 1950s; around 1950, lectures on general relativity at Harvard University were given by Philipp Frank, whose chief interests were the philosophical aspects of the theory (DeWitt 2011, p. 51). Nevertheless, Schwinger did supervise DeWitt's 1949 PhD thesis on self-energy problems related to the quantization of the gravitational field. DeWitt recalled:

> I had a strong feeling that Einstein's theory was in a sort of limbo, detached from the rest of physics, and that it was a shame that such a beautiful theory should be so ignored. I proposed to drag it forcibly into the then modern world by redoing Schwinger's QED calculations with the gravitational field added. I was very naive in those days (DeWitt 1996a).

DeWitt's thesis (which relied on the covariant perturbative quantization) exemplifies the initial expectation that a quantum theory of gravity could be constructed by applying methods from QED to the gravitational field. Of course several physicists had already anticipated that such an expectation was unfounded – for instance, Matvei Bronstein – but such warnings could at the time only be based on qualitative, heuristic reasoning[11]. It was only when the works of Bergmann, Dirac, Pirani, and Schild on constrained dynamics appeared, right after DeWitt completed his thesis, that it was widely acknowledged (outside the immediate high-energy/quantum field

[11] See (Blum and Rickles 2017).

theory community) that these results made the Lorentz-covariant perturbative approach to quantum gravity obsolete. In a paper based on his thesis, which DeWitt submitted to the Physical Review[12], he remarked:

> [A] perturbation philosophy is adopted. The linear (zero order) part of the gravitational Lagrangian function is subtracted from the full Lagrangian. The residue, which contains all the self-interaction effects [...] is treated as a perturbation but is never closely examined. This procedure results, of course, in a rather makeshift formulation of the theory [...]
> The use of a makeshift formulation has [...] been prompted by a more pressing consideration. At the time of initial writing of these papers no one had yet explicitly constructed a gravitational Hamiltonian for the general theory [of relativity]. [Footnote:] F. A. E. Pirani and A. E. Schild, and independently, P. G. Bergmann and his co-workers [...] have just recently succeeded in constructing the required Hamiltonian function. [...] It may eventually be of interest to reformulate some of the present calculations according to the rigorous Pirani-Schild-Dirac scheme.

That such a reformulation would not just "eventually be of interest", but was in fact the only thing of immediate interest, was the view taken by the paper's referee, Howard P. Robertson: "In view of the Pirani-Schild paper, which [DeWitt] has seen – and remarks on p. 3, footnote, that this should eventually be carried out in terms of their more rigorous theory – it would seem to me better to suggest he carry out his work in terms of their theory"[13].

Bit by bit, the perturbative approach began to be rejected by the general relativity community (in spite of Gupta's concise 1952 formulation of the perturbative theory). DeWitt then was faced with a dilemma. As a student of Schwinger, he had learned to give first priority to explicit covariance; but the Bergmann-Pirani-Schild canonical methods were moving away from explicit covariance and adopting an explicit separation of time and space. At the same time however, the successful covariant methods of QED – such as the Gupta-Bleuler method, the Schwinger action principle, the Peierls bracket, etc – were not directly applicable to the (non-linear) gravitational case. DeWitt dealt with this dilemma by adopting a many-pronged approach to quantum gravity, investigating several quantization methods and their interrelations. This included the canonical approach, but only in its original fully four-dimensionally covariant (so called parameter) form. We will return to this specific point in detail in Section 4.4. DeWitt's inclusive interest in the different quantization programs and his ambition to overcome their respective shortcomings by relating them to one another is a *leitmotif* of his entire 1957 Copenhagen report.

Between 1950 and 1953, DeWitt had postdoctoral stays (working on quantum gravity) at the Institute for Advanced Study (Princeton), the ETH Zurich, and at the Tata Institute (Bombay) (DeWitt 2011, p. 140). After working on hydrogen bomb development at Livermore Laboratory for a three-year period[14], he created in 1956, together with his wife Cécile DeWitt-Morette, the Institute of Field Physics at the University of North Carolina at Chapel Hill, North Carolina (DeWitt and Rickles 2011).

---

[12] Preprint of the article "On the application of quantum perturbation theory to gravitational interactions, part I", 1950. Bryce S. DeWitt Personal Files, in Cécile DeWitt-Morette's office, University of Texas at Austin, Austin, TX. Hereafter this collection will be referred to as DF.

[13] Referee report, 17 May 1950. Howard P. Robertson Papers, 1922–1980, box 7, folder 14. California Institute of Technology Archives.

[14] As outlined in "Cold War Curvature", a talk given by David Kaiser at "The 'Renaissance' of General Relativity in History: Assessing Einstein's Legacy in Post-World War II Physics", the General Relativity Centenary Conference held in Berlin from 2-5 December 2015.

Cécile DeWitt-Morette did some work on general relativity in the 1950s, but her main areas of expertise at that time were meson theory and Feynman's path integrals. While there were no other physicists actively working on general relativity when the DeWitts came to North Carolina, it was being taught by Eugene Merzbacher, who had also been a PhD student of Schwinger's. Around 1957, DeWitt started training his first graduate students, Robert Brehme (PhD '59), John J. Ging (PhD '60), and Hsin Yang Yeh (PhD '60).

Stanley Deser had also obtained his PhD with Schwinger. His thesis, completed in 1953, was, however, on meson scattering, i.e., on a topic entirely unrelated to general relativity. Deser then went to the Institute for Advanced Study in Princeton, where he attended a lecture by Albert Einstein and met Elsbeth Klein, the daughter of Oskar Klein, whom he would marry in 1956[15]. In 1955, Deser went to the Institute for Theoretical Physics in Copenhagen. Until his arrival in Copenhagen, Deser had not worked on general relativity. According to his recollections, it was after his meeting with Oskar Klein in Bern in 1955 that he started thinking more seriously about the subject:

> Our meeting and the [Bern] conference also affected me, and I began to think more seriously about the possibility that general relativity could be a universal regulator of the divergence problems of quantum field theory, and in particular of QED. This idea had also occurred to Landau, Pauli and Oskar himself (Deser 1995, p. 3).

We will return in detail to this research project of Deser's in Section 4.5.

The third American participant of the Copenhagen workshop, **Charles Misner**, had only recently obtained his PhD at Princeton under the supervision of John Archibald Wheeler. In 1957, Princeton was arguably the most prestigious center for general relativity in the United States, hosting not only Wheeler but also Eugene P. Wigner, who got involved in general relativity in 1955, when he began to work with Helmut Salecker on the quantum limitations of clocks in general relativity (Salecker and Wigner 1958) – a work deeply related to Møller's, as we will discuss in Section 4.1.

Wheeler had given his first series of lectures on general relativity in Princeton in 1952 (Wheeler and Ford 2000, p. 228) and soon took gravitation to be the cornerstone of his new research program that "everything is field". A paradigmatic example of that program was the *geon*, an object envisaged by Wheeler in which an electromagnetic wave is held together by its own gravitational field (Wheeler 1955; Hartz and Freire 2015). While a quantum theory of geons was (and is) still far off, Wheeler did invoke heuristic estimates of quantum effects, in order to make it plausible that quantum geons would eventually have the right dimensions for identifying them with elementary particles (as opposed to classical geons, which were much too large).

Wheeler, in his attempts to turn general relativity from an abstract philosophico-mathematical field into an integral part of physics, also supported experimental efforts in general relativity, starting in 1955 with Robert H. Dicke and Joseph Weber (see the article by Jim Peebles, Eur. Phys. J. H, `Doi: 10.1140/epjh/e2016-70034-0`). On the theoretical side, Wheeler's star PhD student of the 1950s was Misner. His 1957 thesis was specifically on the application of Feynman's path integral methods to the quantization of the gravitational field, but also discussed in detail the relation of this approach to other attempts at quantizing general relativity.

The three US physicists present at the 1957 Copenhagen meeting did not represent US-American quantum gravity research in its entirety. We have already mentioned the groups of Peter Bergmann and Alfred Schild. The German-born Bergmann had,

[15] See (Halpern 2004, p. 213-214) and a letter from Pauli to Gunnar Källén, 7 October 1956, reprinted in (von Meyenn 2001, p.687-688).

at age 21, become Albert Einstein's assistant in Princeton, a position that he occupied for five years. He later wrote an acclaimed textbook on both special and general relativity (Bergmann 1942). In 1947, he went to Syracuse University, where he created his research group on general relativity and quantum gravity. As two of Bergmann's collaborators observed:

> When Peter began his career at Syracuse in 1947, no US physics department had a center for research in general relativity. Indeed, very few physicists considered the area worthy of their time. Within the Syracuse physics department, Peter created one of the first groups specifically concerned with studying the general theory of relativity with the intent of reconciling that field with quantum theory. [...] Up to the mid-1950s, Peter and his students were the major contributors to the literature in general relativity (Goldberg and Schucking 2003).

Schild (also German-born) had obtained his PhD at the University of Toronto in 1946 under Leopold Infeld. He then went to the Carnegie Institute of Technology, in Pittsburgh, where he did his work on canonical quantization together with his PhD student Felix Pirani, who had followed Schild from Toronto. Schild moved to the University of Texas at Austin in 1957, where he created the Center for Relativity five years later (Schucking 1989).

Finally, at Purdue University, the Dutch physicist Frederik J. Belinfante set up a group working on quantum gravity in the 1950s. Belinfante pursued several approaches in parallel, considering both the quantization of alternatives to general relativity (the linear, Lorentz covariant Belinfante-Swihart theory) and the canonical quantization of general relativity itself, in both cases with a strong emphasis on the interaction of gravitation with other fields, especially spinorial matter.

Research explicitly on quantum gravity was, however, not the sole contemporary influence on the Copenhagen workshop. Early on, there was a very strong European mathematical tradition in general relativity (focusing on its relation with differential geometry), associated with the names of Tullio Levi-Civita, Cornelius Lanczos, Theóphile de Donder, Théophile Lepage, Élie Cartan, and Georges Darmois, among others. Some physicists even perceived general relativity as a branch of mathematics (Mercier 1992), and in some countries, in particular in France, this did not change in the 1950s. The doyen of French mathematical relativity in the 1950s was André Lichnerowicz[16]; his interest in quantum gravity began only in 1958[17], but the work of him and his school on the initial value problem (Lichnerowicz 1955) did have some impact on the discussions in Copenhagen, as we will see in Section 4.2.

After this brief overview of the participants of the Copenhagen workshop and the contemporary quantum gravity scene they were moving in, we arrive at the conclusion of this section. While there was a drastic increase in the research on general relativity in the 1950s, there were only a few groups working on quantum theories of the gravitational field. Even though the 1957 meeting in Copenhagen gathered only six physicists, it still represented a significant proportion of researchers interested in quantum gravity, and all the trends in quantum gravity research touched on in the last two sections were actually discussed at the Copenhagen meeting.

2.3 The first conferences on general relativity

The establishment of new research groups was not the only institutional advance of the 1950s, which also saw the organization of the first conferences devoted solely to general

[16] For more on this towering figure, see, e.g., (Kosmann-Schwarzbach 2009).

[17] See (Lichnerowicz 1964).

relativity and the creation of the International Committee on General Relativity and Gravitation, which aimed at improving the communication between the researches in the field (see Blum et al. 2015, p. 617).

In 1954, on the eve of the fiftieth anniversary of the creation of special relativity, André Mercier solicited the help of Wolfgang Pauli in co-organizing a conference on the occasion (Mercier 1992). The conference, which had ninety participating scientists and took place in Bern, Switzerland, in July 1955, also turned out to be memorial for the creator of the relativity theories: Albert Einstein had passed away in April, only three months before the conference (Mercier and Kervaire 1956).

The Bern conference has been hailed on various occasions as the most important moment in the renaissance of general relativity. Lichnerowicz (1992, p. 105) writes that "the International Congress at Bern [...] marked the true renaissance of interest in general relativity on the part of physicists, astronomers, and mathematicians". Mercier (1992, p. 119) claimed that there the quantum gravity problem "was asserted with steadfastness that did not leave much to be desired". Eisenstaedt (1986) chose the year 1955 as the end of the "low-water-mark" period.

For all its importance, the Bern conference was not the only central event of the renaissance and its importance is greater for Europe than it is for the United States. As Dean Rickles has pointed out, "the Bern conference consisted mostly of an older generation who had persistently thought about general relativity and quantum gravity for decades". (DeWitt and Rickles 2011, p. 19). In the US, the training of a new (arguably the first) generation of relativists had begun before the Bern conference and the event at which this generation established itself as a force to be reckoned with took place one and a half years after the Bern conference, on 18–23 January 1957, at Chapel Hill. It was organized by Bryce DeWitt and his wife Cécile DeWitt-Morette and was also the official inaugural event of the Institute of Field Physics, that had begun its activities in January 1956. The project was funded by Agnew H. Bahnson, a North Carolina industrialist and philanthropist.

The Bern and the Chapel Hill conferences were unrelated as far as their organizational structure was concerned. However, in order to create some sort of tradition – or even to increase the importance of their event –, Bryce and Cécile DeWitt began to refer to the Chapel Hill conference as the "Second International Conference on Gravity"[18]. In Chapel Hill, Lichnerowicz offered to organize the follow-up conference in Royaumont, France, in 1959 (see Lichnerowicz 1992, p. 91, and Lichnerowicz and Tonnelat 1962). In 1962, the next conference in the series was held in Warsaw, Poland, organized by Leopold Infeld (Ashtekar 2014; Infeld 1964). Nowadays, these conferences are viewed as having inaugurated the still ongoing tradition of GR conferences, counting Bern as GR0, Chapel Hill as GR1, Royaumont as GR2, and Warsaw as GR3.

Those conferences – the amount of new questions that were formulated, the intense discussions among its participants, the huge progress in general relativity from one conference to another – are the greatest historical evidence showing the renaissance of general relativity. Because of the Chapel Hill conference, the year 1957 is well recognized as a turning point in quantum gravity research. According to Rickles:

> The Chapel Hill conference on the Role of Gravitation in Physics [...] did for general relativity what Shelter Island did for quantum electrodynamics. The Chapel Hill conference was a genuine break from the Bern conference, both in terms of its organization, its content, but more so its *spirit* (DeWitt and Rickles 2011, p. 19-20).

The 1957 Copenhagen meeting happened just five months after the Chapel Hill conference and many of the participants in the Copenhagen conference had also been at

[18] Bahnson Memoranda #5 (16 October 1956) and #6 (1 November 1956). DF, box Institute of Field Physics.

Table 1. Who attended each conference?

Person/Conference	Bern '55	Chapel Hill '57	Royaumont '59	Warsaw '62
Stanley Deser (1931–)	Yes	Yes	Yes[19]	Yes
Bryce S. DeWitt (1923–2004)	No	Yes	Yes	Yes
Oskar Klein (1894–1977)	Yes	Yes	No	No
Bertel E. Laurent (1928–1993)	No	Yes	Yes	Yes
Charles W. Misner (1932–)	No	Yes	Yes	Yes
Christian Møller (1904–1980)	Yes	No	Yes	Yes

Chapel Hill. In Table 1, we present a list of the six participants of the 1957 Copenhagen meeting with the information on which of the first four major GR conferences they attended. Also thematically, several of the issues discussed in Chapel Hill reappeared in the Copenhagen discussions, as we will highlight systematically in Section 4. However, as opposed to the conferences listed above, the meeting in Copenhagen had a very specific focus: It was the very first gathering of researchers to discuss solely the quantization of gravitation. In the next section, we discuss how this specialized meeting came to be.

3 The history of the 1957 meeting in copenhagen

3.1 Christian Møller and the Research Group in Copenhagen

Since the establishment of the Institute of Field Physics, Bryce and Cécile DeWitt put a lot of effort into fund raising. Together with Agnew Bahnson, they approached a great number of industrialists and created a fair amount of publicity for their work, mainly through newspaper articles. A fundamental step in drawing general attention to the Institute was the 1957 Chapel Hill conference. It was therefore of central importance to attract well-known physicists of international renown to that event.

DeWitt made a preliminary list of eighteen "physicists whom we definitely want to come [to the Chapel Hill conference] and who would be expected to play an active role in the discussions"[20]. There were five physicists from Europe on the list: Niels Bohr, Christian Møller, André Lichnerowicz, Oscar Klein, and Marie-Antoinette Tonnelat. Apparently concerned with funding restrictions, DeWitt decided to choose just three of them. As far as we know, Bohr never got invited. He had already declined the invitation to attend the 1955 Bern conference, feeling that "he had not thought enough about the possible future of a generally relativistic cosmology that would have given him premonitions of a development to come" (Mercier 1992, p. 112).

Møller had been a key person in the development of the Institute for Theoretical Physics in Copenhagen. Denmark had been one of the candidates for hosting the CERN laboratory in a competition that was ultimately won by Switzerland in October 1952 (Krige 1987a). Møller took part in these discussions as one of the representatives of the Danish government (Pestre 1987, p. 161). As part of Denmark's involvement in the creation of CERN, the CERN Theoretical Study Group was established in Copenhagen in 1952. Møller became the Director of the group in September 1954, taking over a position that had previously been Niels Bohr's (Krige 1987b, p. 217). He kept his position as the director of the research group until it was replaced, in

[19] Wrongly referred to in the conference proceeding as F. Deser.

[20] Letter from Bryce S. DeWitt to Agnew H. Bahnson Jr., 7 March 1956. DF, box Institute of Field Physics, folder Bahnson Correspondence.

early September 1957, by a new Theoretical Study Group in Geneva, marking the end of CERN activities in Copenhagen (Krige 1987a, p. 246). Nevertheless, the staff of the Copenhagen group remained there, creating the Nordic Institute for Theoretical Physics (NORDITA), whose first director was Møller, from 1957 until 1971 (Rozental 1985, p. 117).

The Theoretical Study Group in Copenhagen was extremely successful. Each country that was a member of CERN was supposed to send one early career theoretical physicist every year (Iliopoulos 1996). In theory, they would work on theoretical subjects related to the problems posed by the experimentalists, but that original purpose of the group was hardly taken seriously. There was some nuclear physics work, for instance by Aage Bohr and Ben Mottelson, but also some very mathematical work, unrelated to experimental matters, by Rudolf Haag and Gunnar Källén, among others. Møller's own engagement with general relativity meant that he was open to not sticking to CERN experimental concerns. He warmly welcomed, for instance, Haag's idea of studying the Wigner classification of irreducible representations of the Poincaré group (Haag 2010, p. 267).

There were also young physicists associated with the Theoretical Study Group coming from outside the CERN member states, such as Steven Weinberg and Stanley Deser from the USA, who were funded by the National Science Foundation[21]. Deser stayed in Copenhagen for two years, from mid-1955 until the end of the group. Before going to Copenhagen, he was "working, with R. Arnowitt, on the question of renormalization, and non-perturbative solutions of 'non-renormalizable' theories such as ps(pv)"[22], that is, (ps=pseudoscalar) meson theories with (pv=pseudovector) derivative couplings, a subject that was related to the work done by Källén and others in the Copenhagen group (Arnowitt and Deser 1955). In Denmark, encouraged by Møller, Deser found a favorable environment to refocus his research on general relativity.

3.2 Christian Møller and the Chapel Hill project

DeWitt wrote to Møller in April 1956 telling him about the new research group at Chapel Hill and inviting him to join it as a visiting senior professor in the following academic year[23]. Møller's salary was going to be partially funded by a generous donation from the Research Corporation[24]. That was the very first invitation sent out by DeWitt as director of the Institute of Field Physics. DeWitt hardly knew Møller. They had probably met twice before, the first time at the Tata Institute in India in 1951 and the second time later that same year at the Institute for Theoretical Physics (Bhabha 1951). But Cécile DeWitt-Morette had worked under Møller's supervision in the late 1940s. In a report to the French Department of Education, Møller had the following to say on her stay in Copenhagen:

> After having finished her studies in Paris, Madame C. DeWitt, née Morette, spent the academic year 1947-48 at the Institute for Theoretical Physics, Copenhagen, as a Fellow of the Rask-Ørsted Foundation. [...] [Her] work showed her deep understanding of the physical problems and her mastery of the mathematical methods which also is so characteristic of her subsequent work. During

[21] Letter from Christian Møller to the National Science Foundation, 16 May 1955. Christian Møller Papers, Niels Bohr Archive, Copenhagen, Correspondence, box 2. Hereafter this collection will be referred to as MP.

[22] Letter from Stanley Deser to Christian Møller, 7 April 1955. MP, Correspondence, box 2.

[23] Letter from Bryce DeWitt to Christian Møller, 30 April 1956. MP, Correspondence, box 2.

[24] Bahnson Memorandum #6 (1 November 1956). DF, box Institute of Field Physics.

> her short visit to Copenhagen she made many friends, and due to her amiable personality it was easy for her to get into contact with other physicists at the Institute[25].

Thus, it is likely that inviting Møller to stay one year at Chapel Hill was her idea. The physics staff there consisted at that time of only six researchers, including Bryce DeWitt, Cécile DeWitt-Morette, Eugen Merzbacher, and Everett Palmatier. In the invitation letter, DeWitt referred to John Wheeler (who was then visiting Copenhagen) and Stanley Deser as those who could explain the project at Chapel Hill and its history. Wheeler, DeWitt stated as his greatest credential, was "to a great extent the godfather of the project" at Chapel Hill. The salary that they offered to Møller was quite attractive, $9000 for nine months plus the round trip fare to and from the United States[26]. The letter also mentioned that Eugen Merzbacher was teaching general relativity at Chapel Hill using Møller's textbook. In the same letter, DeWitt further informed Møller that he and Cécile DeWitt were planning "a small conference on gravitational theory" from February 4 to 8, 1957 – which ended up being the 1957 Chapel Hill conference. She added a handwritten postscript to the letter, "You do not know how much your visit would be appreciated", which she signed simply "Cécile". That was as much as Bryce and Cécile DeWitt could do in order to attract Møller to Chapel Hill[27]. He regretted not being able to accept the invitation:

> You may know that I am involved in the work with CERN. The Theoretical Study Group will stay in Copenhagen till the fall of 1957, and I have committed myself to leading the Group until that time. Thus, no other obligation can be considered in the year to come. [...] Also the conference in February would have been of much interest to me. However, you have certainly gathered from my above plans that, to my deep regret, I shall also be unable to come on this occasion[28].

Not having Møller, either as a visitor or at the conference, was for sure a great disappointment to the DeWitts. However it was more than simply a personal disappointment, since Møller would be the representative of the Institute for Theoretical Physics at the conference. And both Bryce and Cécile DeWitt had a clear idea of the relevance of having a well-known physicist from Copenhagen at the event. As the historian John Krige has shown, during the 1950s Copenhagen was perceived by the American funding agencies as the strategic center of the intellectual Cold War in Europe (Krige 2006, p. 169–172). DeWitt then invited to the 1957 conference the three other European physicists from his list, Lichnerowicz, Klein, and Tonnelat. A few months later, as the conference budget was finally confirmed in a favorable manner, that list of European physicists grew.

The Chapel Hill conference finally took place from 18–23 January 1957 and was a huge success. This helped secure further funds for the Institute of Field Physics, allowing Bryce and Cécile DeWitt to consider the possibility of inviting more physicists to join the staff on a temporary basis. Invitations were sent out to Bertel Laurent and to Felix Pirani.

In summer 1956, DeWitt had visited Copenhagen during a trip to Europe, in order to reassert in person the invitation to Møller, now for the academic year 1957–1958[29].

---

[25] Letter from Møller to DeWitt-Morette, 22 February 1965. MP, Correspondence, Supplement, box 1.

[26] For comparison, $1000 per month was the *combined* salary of Bryce and Cécile DeWitt.

[27] Letter from Bryce DeWitt to Christian Møller, 30 April 1956. MP, Correspondence, box 2.

[28] Letter from Christian Møller to Bryce DeWitt, 16 May 1956. MP, Correspondence, box 2.

[29] Bahnson Memorandum #5 (16 October 1956). DF, box Institute of Field Physics.

Møller had hesitated several months to confirm the acceptation. In February 1957, after the big conference, he finally answered that he wouldn't be able to stay for nine months at Chapel Hill, since he had already arranged a long stay at Pittsburgh[30]. DeWitt then got impatient and Cécile had to intervene[31]. She insisted that Møller should visit Chapel Hill as soon as his duties with CERN were finished, even if for just a short period:

> As we have not been able to "scare" (!...)[32] you into an extended period we would be very happy to have you on the faculty here for a two month period. [...] I hope there will be no unforeseen difficulty preventing you from coming to Chapel Hill[33].

Møller accepted this last invitation and arranged to stay at Chapel Hill for two months. Apparently, the visit to Pittsburgh had indeed been planned for a long time. In an interview given to the Danish newspaper Berlingske Aftenavis, Møller stated:

> That is an old invitation, which I now see myself in a position to accept [...]. The director of the [Carnegie] Institute [of Technology] has so to say repeated the invitation each year, but I have so far been compelled to refuse, among other things because of my work as head of the theoretical division of CERN, which, however, as of next autumn will no longer be located at Niels Bohr's institute. My lectures [in Pittsburgh] will focus on the general theory of relativity[34].

After his stay at the Carnegie Institute of Technology and a quick visit to Princeton, Møller finally went to Chapel Hill, however for just one month, beginning 25 January 1958. Despite the certain amount of frustration it brought to the DeWitts, having to wait so long for such a short visit, the attempt to get Møller to their institute for a longer period did strengthen the bonds between Copenhagen and Chapel Hill, and thereby made possible the 1957 meeting in Copenhagen.

3.3 Planning the 1957 meeting in Copenhagen

DeWitt had also made some efforts to attract to Chapel Hill Stanley Deser and Charles W. Misner as research assistants. Bahnson reported in one of his memoranda:

> On the way back from their stimulating summer in [the 1956] Les Houches [Summer School], Bryce DeWitt made a trip to Copenhagen and contacted Dr. Møller and Dr. Deser. It seems likely that Dr. Møller who is an outstanding European physicist, will join the project at Chapel Hill for several months next year and that Dr. Deser may come to Chapel Hill as a research assistant.

[30] Letter from Bryce DeWitt to Christian Møller, 12 February 1957. Letter from Christian Møller to Bryce DeWitt, 18 February 1957. MP, Correspondence, Supplement, box 1.

[31] Letter from Bryce DeWitt to Christian Møller, 12 February 1957. Letter from Christian Møller to Bryce DeWitt, 18 February 1957. MP, Correspondence, Supplement, box 1.

[32] Footnote in the letter: "Mille excuses – this shows only how much we wanted you for a long stay". She is referring to Bryce DeWitt's harsh letter from 12 February, insisting that Møller should stay for an extended period. This insistence was related to funding restrictions, as she explained to Møller in the letter from 22 February: The grant from the Research Corporation had been awarded on the supposition that Møller would stay at Chapel Hill for an extended period and could not be used for a shorter stay.

[33] Letter from Cécile DeWitt-Morette to Christian Møller, 22 February 1957. MP, Correspondence, Supplement, box 1.

[34] "Professor Chr. Møller til USA", Berlingske Aftenavis, 20 August 1957. Danish Newspaper Clippings Collection, volume 34, Niels Bohr Archive, Copenhagen.

> These are very outstanding people in the field of gravitation and contacting them personally was a worthwhile move in the next important steps of trying to attract the best personnel to Chapel Hill now that the problem of attack is becoming more well defined[35].

However, five months later, the situation had changed. Møller had already made clear that, due to his agreement with Pittsburgh, he could only consider a short visit to Chapel Hill; Deser had received a better offer from Julian Schwinger at Harvard; and Misner had to decline DeWitt's invitation because he needed a place where he could interact with mathematicians – something that was not possible at Chapel Hill. So DeWitt wrote to Bahnson:

> Seeing that it was unlikely that Deser and Misner could come to work with us next year and at the same time feeling very strongly that they have very interesting things to say and that we could do some very interesting things together, I got in touch with them to find out if they would be interested in getting together in Europe this summer. As you know, Deser is currently in Europe anyway and Misner was going there just for a vacation. Both of them responded very favorably and we also got in touch with Professor Schwinger of Harvard to find out if he would be interested in joining the group. He too, responded favorably since he was anyway going to be in Europe. According to present plans, a group of six of us will be in Copenhagen during the month of July working at the Institute for Theoretical Physics directed by Bohr. The six included myself, Deser, Misner, Laurent, Schwinger and Professor Klein of Stockholm[36].

We do not know the reason why Schwinger had to cancel his participation. It is understandable that Møller would not have been expected to join the meeting – even though he did ultimately join it – because the Theoretical Study Group was about to be dissolved, NORDITA was being created, and, in addition to all that, Møller was about to move to the United States for about half a year.

In any case, organizing such a meeting required a certain amount of funding. DeWitt applied to both the National Science Foundation and the US Air Force. As we already know, it was the latter who ended up funding the meeting. As Bahnson wrote in a memorandum:

> Bryce is leaving on MATS Air Force transportation [Military Air Transport Service] the end of May or the first of June [1957] to attend the conference of mathematics in Lille, France from June 3 to 8. He will spend the month of July in Copenhagen with five other outstanding physicists who will try to get a fresh point of view on the problem [of the quantization of gravity] in associating with one another, and I believe the next year will show exciting progress in the solution of a most profound and difficult problem[37].

Bahnson, a philanthropist with no formal training in physics, sometimes had rather fanciful expectations. In this case – that the following year was to "show exciting progress" in quantum gravity – they were not altogether unfounded. We will return to this question in the conclusions.

---

[35] Bahnson Memorandum #5, see footnote 29.

[36] Memorandum from Bryce S. DeWitt to Agnew H. Bahnson, 19 March 1957. DF, box Institute of Field Physics.

[37] Bahnson Memorandum #9 (7 May 1957). DF, box Institute of Field Physics.

4 An analysis of Bryce DeWitt's report

We now turn to the analysis of Bryce DeWitt's 1957 report. The division of Section 4 follows the report's structure, and the reader might be able to profit from it most by reading this section and the report in parallel.

4.1 The theory of measurement

The first section of the report discusses the measurement of a quantum gravitational field. It is remarkable that DeWitt decided to begin with such an epistemological issue. Since Møller took part in these discussions and since he joined the meeting only during the last eleven days, we can deduce that the presentation of the report did not follow the order of the discussions. Therefore, putting the measurement problem at the beginning was a choice of DeWitt's.

In the late 1960s, DeWitt would become the most well known supporter of Hugh Everett's interpretation to quantum mechanics, but in 1957 he was still quite close to Niels Bohr's views[38]. Indeed, the definition that DeWitt provides of a measurement is entirely Bohrian: measurements are consistency problems (Kalckar 1971, p. 127). He began the report with the claim that in the development of physical theories that extend previous concepts, at some point one must examine the self-consistency of the entire new framework – i.e., at some point one must develop an analysis of the possible measurements within that theory. DeWitt often started his texts with general methodological lessons learned from historical episodes. In this case he was referring to the early history of "electrodynamic theory", i.e., quantum electrodynamics, where the examination of the measurement problem in quantum theories came *after* the development of the mathematical formalism. Before returning to the 1957 report, we shall briefly discuss this historical episode, as it is essential for understanding DeWitt's position.

In 1929, Werner Heisenberg wrote a letter to Niels Bohr explaining some ideas that he had just sketched about the measurement of electromagnetic fields in quantum theory. In close analogy with the position-momentum uncertainty relations, Heisenberg expected that the electromagnetic field components – which were also non-commuting variables, according to the quantum electrodynamical formalism that he had just developed with Wolfgang Pauli (Heisenberg and Pauli 1929) – would be subjected to similar uncertainty relations. Heisenberg then designed two thought experiments with the purpose of visualizing the limitations on the simultaneous measurement of field components. These were just preliminary investigations, as he explained to Bohr:

> I believe that the essential features here are correct, but the utilization of the average values of E [the electric field] and H [the magnetic field] in the cube is still not quite solid. But would you in principle consider such a discussion reasonable?[39]

That letter was discussed in Copenhagen and two young physicists – Lev Landau and Rudolf Peierls – decided to tackle the problem themselves. There was at the time a widespread belief that the quantum theory of relativistic systems – in particular, quantum electrodynamics – was essentially wrong, due to the occurrence of several divergencies in the calculations (whose solution would appear only in the late 1940s

[38] This shift in DeWitt's opinion on the foundations of quantum mechanics – from a Bohrian perspective to a radical Everettian one – will be analyzed in a forthcoming article by one of the authors (TH).

[39] Letter from Werner Heisenberg to Niels Bohr, 16 June 1929. Translated in (Kalckar 1996, p. 7).

with the full development of renormalization methods). Landau and Peierls decided to search for the source of all the troubles. Following Bohr's ideas – "in the description of atomic phenomena, the quantum postulate presents us with the task of developing a 'complementarity' theory the consistency of which can be judged only by weighing the possibilities of definition and observation" (Bohr 1928, p. 580) – they decided to use a variant of Heisenberg's thought experiment to show that the very idea of a measurement was meaningless in relativistic quantum theory.

Their argument goes as follows. In order to measure an electromagnetic field one must observe its action on a test charge, whose movement provides all information one can obtain about the field. Landau and Peierls chose a point electron as the test charge. They demonstrated, through a thought experiment, that in a quantum context the radiation reaction of the test charge is uncontrollable, leading to the following fundamental uncertainty for each component of the electric field:

$$\Delta E_i \gtrsim \frac{\sqrt{\hbar c}}{(cT)^2}. \tag{1}$$

Since the mathematical formalism does not predict such single component uncertainties (only pair uncertainties), Landau and Peierls concluded – following Bohr's terminology – that the quantum theory of the electromagnetic field, in particular the relation between definition and observation, is not consistent. Generalizing that reasoning, they claimed that "in the correct relativistic quantum theory (which does not yet exist), there will therefore be no physical quantity and no measurement in the sense of wave mechanics" (Landau and Peierls 1931, p. 69, translation in ter Haar 1965, p. 50).

Niels Bohr was deeply bothered by these conclusions (Darrigol 1991; Jacobsen 2011). It took him almost three years to work out an adequate refutation, which came in the form of a long, difficult, and subtle paper written with Léon Rosenfeld (Bohr and Rosenfeld 1933). Its message, however, was quite simple: In order to make the best possible measurement, one should use not an electron, but a massive test charge (with no fixed charge/mass ratio), with macroscopic dimensions (so that its charge density is finite), and attach it to a series of springs (which act as a compensation device). Then the radiation reaction may be entirely eliminated. The resulting uncertainties for the electromagnetic field obtained from a thus improved thought experiment are exactly those expected from the mathematical formalism of quantum electrodynamics. Therefore, quantum theory may be consistently applied to relativistic systems after all, in the same way that it was applied to non-relativistic systems.

We now return to the 1957 report. DeWitt explains that the situation in the quantum theory of the gravitational field is different, since "the question of 'measurability' has been raised even before the advent of a formalism". The first discussions about the measurability of the quantum gravitational field dates back to the mid-1930 (Bronstein 1936; Solomon 1938; see also Blum and Rickles 2017; Hartz and Freire 2015, and references therein)[40]. However, DeWitt is not referring to those early efforts, of which apparently he was not aware, but actually to the 1957 Chapel Hill conference, where the measurement of a quantum gravitational field was discussed several times.

The Chapel Hill conference had two proceedings: One edited by Bryce DeWitt was published in the Reviews of Modern Physics in July 1957, and another written by Cécile DeWitt-Morette circulated as an internal report to the US Air Force – in the same style as the Copenhagen report, but considerably larger (167 pages). The latter proceedings were produced between the end of January and mid-March

[40] These discussions were still ongoing some fifty years later (von Borzeszkowski and Treder 1988, Chapter 3).

1957. Most of the talks and discussions were paraphrased, but several large sections of the discussions were reproduced verbatim (in quotation marks). This was possible due to major advances in portable tape recording technology in the 1950s, allowing Cécile DeWitt-Morette to go around during the conference and record virtually all the discussions[41]. Of course, the poor quality of the recording demanded complementary methods. As she explained:

> It can hardly be said that the report [of the Chapel Hill conference] gives a perfectly true picture of the conference. The report has been prepared from notes taken during the session, from material given by the authors, and from tape recordings. (The reporters have hoped to have a stenographic transcript available, but the cost of this transcript was beyond common sense.) Some contributions have been very appreciably abridged, some are reproduced practically verbatim, some are extended, and some have not been recorded, depending largely on the "communication" (both material and intellectual) between authors on the one hand and reporters and editors on the other (DeWitt and Rickles 2011, p. 34).

Due to these limitations, it cannot be deduced from the proceedings who was (were) the one(s) to raise the problem of measurement of the gravitational field in discussion. The proceedings indicate – in Cécile DeWitt-Morette's own words – that, after Peter Bergmann's talk, which initiated the second half of the conference, dealing with quantum gravity, the following discussion took place:

> Discussion then turned to the problems of measurement of the gravitational field. This item was placed first on the agenda in an attempt to keep physical concepts as much as possible in the foreground in a subject which can otherwise be quickly flooded by masses of detail and which suffers from lack of experimental guideposts. The question was asked: What are the limitations imposed by the quantum theory on the measurements of space-time distances and curvature? Since the curvature is supposed to be affected by the presence of gravitating matter, an equivalent question is to ask: What are the quantum limitations imposed on the measurement of the gravitational mass of a material body, and, in particular, can the principle of equivalence be extended to elementary particles? (DeWitt and Rickles 2011, p. 167)

Probably that was the first time that the 1950s generation expressed a concern about the measurement of a quantum gravitational field and the fact that it was placed at the beginning of the quantum gravity session, just as it was placed at the beginning of the Copenhagen report, makes it highly probably that it was Bryce DeWitt who put this problem on the agenda. In the proceedings, he appears, however, only as the (probable) author of an editorial note.

The main talk on measurements at Chapel Hill was given by Helmut Salecker, who presented his work with Eugene Wigner. He began by demonstrating that one can measure space-time distances using only a clock, with no need for rods. All quantum limitations in those measurements are therefore a consequence of the quantum limitations of clocks. Wigner had always been puzzled by Bohr and Rosenfeld's reasoning because the measurement device they proposed could hardly be reproduced in a real experiment (Wigner 2002, p. 9). According to Bohr, thought experiments should not be limited by what is actually possible; it should only be restricted by the rules of the formalism applied to idealized experimental situations. The tension between Bohr-Rosenfeld-type ideal clocks (with which the uncertainties implied by the formalism could be reproduced exactly) and physically feasible clocks (with which these

[41] Information obtained in a conversation of one of the authors (TH) with Cécile DeWitt-Morette, 4 August 2011, Austin, TX, USA.

uncertainties could only be approximately reproduced) was Wigner's main concern in his work with Salecker (Salecker and Wigner 1958).

After Salecker's talk at Chapel Hill, Rosenfeld took the floor to comment on the difficulties in measuring the quantum gravitational field. In the electromagnetic case discussed by him and Bohr in 1933, it had been necessary to use both positive and negative charges, in order to reduce the electromagnetic field of the test charge to a dipole field (with no monopole term). It had also been essential for the 1933 argument to be able to arbitrarily adjust the charge to mass ratio. In the gravitational case, the situation was quite different on both accounts. Rosenfeld thus agreed with Salecker's claim that the measurement of the gravitational field has further limitations than in the electromagnetic case, and probably would even produce single-component uncertainties, in the manner of Landau and Peierls (DeWitt and Rickles 2011, p. 178–179). The implications of such single-component uncertainties would have been dramatic: In essence, they would have made any straightforward quantization of the gravitational field impossible. At this point, Wheeler intervened. In Cécile DeWitt-Morette's words:

> Wheeler suggested that perhaps one should simply forget about the measurement problem and proceed with other aspects of theory. The history of [quantum] electrodynamics shows that it is always a ticklish business to conclude too early that there are certain limitations on a measurement. He would propose rather to emphasize the organic unity of nature, to develop the theory (i.e., quantum gravidynamics) first and then to return later to the measurement problem. He suggested that this was particularly appropriate when we don't even understand too much about the classical measurement process! (DeWitt and Rickles 2011, p. 179)

We see here the main antagonism of quantum gravity in 1957, that between Wheeler and DeWitt: The latter meticulous and diligent, trying to get the foundations straight, the former sanguine and visionary, interested only in the exciting things one might be able to do with quantum gravity. We will see this antagonism pop up throughout the report. At Chapel Hill, in any case, the measurement issue was not entirely forgotten. In particular, Rosenfeld expressed some second thoughts. He stated that "perhaps his original pessimism in regard to the measurability of gravitational fields, as compared to the electromagnetic case, might be unjustified" (DeWitt and Rickles 2011, p. 184–185). He argued that perhaps the uncertainties in the position and momentum of the test charge (i.e., in the gravitational case, a test mass) influence different components of the gravitational potential, so that any single component could be measured with arbitrary precision.

Despite Rosenfeld's cautious optimism, the discussions at Chapel Hill led DeWitt to state – in the Copenhagen report – that measurement issues in the quantum theory of the gravitational field "have been mainly directed toward proving the inapplicability of quantum mechanics in the gravitational domain". He then goes on, surprisingly at first glance, to suggest that perhaps the expectation of obtaining strong evidence against the quantum gravity program from measurement analyses may just be wishful thinking, because of the complexity of the problem. DeWitt thus seems to be subscribing here to Wheeler's opinion that one should not approach the problem of measurements in quantum gravity for the time being. We thus have here an interesting historical problem: Why then did DeWitt dedicate so much attention in Copenhagen to measurement matters? There are two possible answers.

First, DeWitt did not have at the time a good argument *for* the quantization of the gravitational field, the problem to which he was dedicating his career. His basic argument in 1957 was the following:

> No apology will be made for [the desire to attack the problem of developing a quantum theory of gravitation], although needless to say, recent experiments

> have nothing to do with it! In the author's opinion it is sufficient that the problem is there, like the alpinist's mountain (DeWitt 1957b, p. 377).

At the same time, there were theoretical arguments on measurability supporting the idea that perhaps the gravitational field should *not* be quantized. We could mention Salecker's and Rosenfeld's measurement arguments, as well as some other thought experiments formulated at the Chapel Hill conference that raised some doubts concerning the necessity of quantizing the gravitational field (DeWitt and Rickles 2011, p. 247–260; see also Zeh 2011). Thus, in order to protect his research project, DeWitt had to be able to argue against such doubts. That is the purpose, for instance, of the long calculation that DeWitt presented in the Copenhagen report. The great importance that he attributed to such a calculation can be seen from the fact that he had already included it in the Chapel Hill report, during the editing process (DeWitt and Rickles 2011, p. 167–169). As opposed to Wheeler, DeWitt did not want to (or did not feel able to) cavalierly brush such objections aside.

Second, DeWitt perhaps had some cautious hopes that the measurement analysis could be of some aid in the development of a quantum theory of gravitation. We will come back to this point in a few paragraphs.

DeWitt mentions that the discussions about measurement were raised by Christian Møller, who had discussed the matter with Wigner. Møller had also been interested in understanding the behavior of a clock in a gravitational field. It was well known (Møller 1952, p. 247) that the proper time τ of an observer moving with velocity v in a region subjected to a gravitational field with the scalar potential χ is given, in a system of coordinates that is time-orthogonal, by the expression

$$d\tau = dt\sqrt{1 + 2\chi/c^2 - v^2/c^2}, \tag{2}$$

where t is the proper time of an observer at rest in the absence of a gravitational field. Møller wanted to deduce expression (2) from a clock model – also for pedagogical reasons, as discussed in Section 2.2 – and to evaluate what would be the changes in expression (2) if one considered a real (Wignerian) clock (not an ideal Bohrian one).

Møller also expressed some concerns about approaching measurements at that stage of the development of the quantum theory of the gravitational field, since Bohr and Rosenfeld in 1933 had to systematically use the mathematical formalism of quantum electrodynamics in order to find the best possible way of measuring the electromagnetic field – as DeWitt mentions in the report:

> Professor Møller quoted a statement made by Professor Bohr after finishing his famous "measurement" paper with L. Rosenfeld, viz. "During the course of our study of the quantum limitations on the measurability of the electromagnetic field we made every possible mistake. In each case, in order to extricate ourselves, we had to go back and look at the formalism!"

That was meant to substantiate Wheeler's opinion that the measurement problem in quantum gravity was to be left alone for the time being. Instead of following Møller's advice, however, DeWitt ultimately decided to invert the argument and to attempt and define the formalism (i.e., the commutation relations of the theory) in terms of the measurement analysis. In this manner, the above quotation came to be the main source of inspiration for DeWitt's work in the five years following the Copenhagen meeting (Hartz and Freire 2015). In his contribution to the famous collection *Gravitation*, edited by Louis Witten in 1962, DeWitt stated:

> The author [DeWitt] is heavily indebted [to Bohr and Rosenfeld's 1933 paper] as will be immediately apparent in the sections to follow. This indebtedness may seem in one respect surprising, not, to be sure, because of any present-day

> diminution in the importance of this classic work, but because its content, as Bohr and Rosenfeld have themselves repeatedly indicated, was guided in every way by the existence of an already developed formalism, whereas here we are trying to "put the cart before the horse" – to develop the formalism itself with the aid of the ideas of the theory of measurability (DeWitt, 1962, p. 270).

This way of defining the mathematical formalism from the measurement analysis was central in much of DeWitt's later work on quantum field theory and quantum gravity (see, e.g., DeWitt 1971, 2003, p. 30–144; see also DeWitt 2011, p. 15).

4.2 Topological problems

This section at first glance appears somewhat out of place in this workshop on quantum gravity. DeWitt does offer a motivation for the study of topological problems in the context, stating that the classical behavior of the gravitational field is "not well understood" and implying (though only by historical analogy) that this is a necessary prerequisite for the construction of the corresponding quantum theory. There are, however, much more important reasons for the prominent place of a section on topology in this report and that is the role played by topological considerations in the research program of John Archibald Wheeler.

In 1953, after a decade of pursuing a program of reducing all of physics to direct particle interactions (as in Wheeler-Feynman electrodynamics), Wheeler had switched gears entirely. In the academic year 1952/53, he was teaching his famous first course on relativity, both special and general, with the aim of learning general relativity himself in order to, it appears from his notebooks, also cast the gravitational interaction in the form of a direct particle interaction. In the course of the term, however, Wheeler became enamored with Einstein's unified field theory program. This was, in a rather explicit sense, the exact opposite of Wheeler's program, namely to reduce everything to fields, as outlined by Einstein in the appendix to the fourth edition of his "Meaning of Relativity", which found an interested reader in Wheeler in late March 1953[42]:

> There is [...] the conviction that one cannot keep side by side the concepts of field and particles as elements of the physical description. [...] The field concept, however, seems inevitable, since it would be impossible to formulate general relativity without it. [...]
> For this reason I see in the present situation no possible way other than a pure field theory, which then however has before it the gigantic task of deriving the atomic character of energy. (Einstein 1953, p. 164-165)

How now to face this "gigantic task?" Wheeler at first focused on studying solutions of the vacuum Einstein-Maxwell equations with localized energy, the so-called geons (Wheeler 1955), but soon turned toward an approach that had first been suggested (though not pursued much further) by Einstein himself, already in the mid-1930s (Einstein and Rosen 1935), namely to model particles in a pure field theory through the topology of space(-time). Wheeler's paradigmatic example of topology as particles was the wormhole, a topological handle with closed electric field lines partially trapped inside, such that field lines appear to be diverging from the two points where the handle is attached to the rest of space (mimicking, in Wheeler's words, charge without charge).

[42] The above reconstruction of Wheeler's path is the preliminary result of a detailed study of Wheeler's notebooks and his intellectual trajectory in the 1950s, currently being pursued by one of the authors (AB) together with Dieter Brill.

While clearly inspired by Einstein, Wheeler's program differed from Einstein's in one central aspect. Einstein had hoped that the particle-like entities obtained in a unified field theory would (approximately) obey the laws of quantum mechanics, without these laws (with all their conceptual baggage, to which Einstein objected) having to be explicitly introduced into the foundations of the theory. For Wheeler, on the other hand, his classical considerations were merely a prelude to the eventual construction of the full quantum theory. This allowed him to bracket the obvious shortcomings of the classical topology as particles program, e.g., the fact that one could deduce an upper limit on the charge-to-mass ratio that was many orders of magnitude too small to reproduce even the most massive of elementary particles (Misner and Wheeler 1957). Wheeler (1957) was rather vague on how exactly these shortcomings of the program were to be cured in the quantum theory, but it centrally involved Feynman path integral quantization with integration over field configurations with non-trivial topologies and a special emphasis on the quantum fluctuations of the metric at the Planck scale. Topology was thus an essential element of Wheeler's approach to quantum gravity or, more properly, quantum gravity was an essential element of Wheeler's particles-from-topology program.

DeWitt did not mention these speculations of Wheeler's, instead framing the section on topology in a more conservative manner merely as an exploration of the classical theory. We will see in more detail in Section 4.5 that DeWitt was both highly skeptical of Wheeler's program and hesitant to criticize it outrightly, which explains his cautious formulation. This caution also most likely coincided with that of the representative of the Wheeler school in Copenhagen, Misner. He had worked both on the formal questions of the Feynman quantization of general relativity (see the next section) and had co-authored with Wheeler the paper on the classical theory of wormholes; but the highly speculative paper in which Wheeler sought to connect the two had been written by Wheeler alone. The discussion of topology at the workshop therefore appears to have focused entirely on classical issues, with Wheeler's "quantum foam" speculations looming as a distant motivation in the background.

One aspect discussed in Copenhagen was whether one might also be able to describe the sources of a scalar (Yukawa) meson field (believed at the time to be the adequate description of the interaction between nucleons) using wormhole-like structures ("topological tricks" as they are referred to in the report)[43]. In the Einsteinian tradition, Wheeler had focused exclusively on the gravitational and electromagnetic interactions. Indeed, he was highly skeptical of the meson theories of nuclear interactions of the time, as he stated, e.g., his 1954 Richtmyer lecture[44]:

> I find it impossible to regard the meson field hypothesis as more than a free invention – ingenious and suggestive, to be sure, but only a free invention –

[43] The report mentions that this problem was studied with the help of Arthur Komar who was a postdoc in Copenhagen at the time, apparently studying "the interaction of scalar and gravitational fields". There is no trace of this work in Komar's publications, but we do find the following remark in Wheeler's letter of reference for Komar to Aage Bohr in Copenhagen (24 April 1956, Wheeler papers, Box 15, Komar correspondence): "Komar has a strong background in mathematics, especially differential geometry [...] His background in nuclear physics and quantum electrodynamics is nowhere near so advanced, but I know he is particularly eager to do something in field theory and wants to come to Copenhagen for that reason". Komar does not appear to have pursued this any further after moving to Bergmann in Syracuse the following year.

[44] In contrast, Wheeler very much believed that the neutrino would remain an integral part of fundamental physical theory, despite the fact that it had not even been discovered at the time of the 1954 lecture. His attempts to construct wormholes that acted as sources of a classical neutrino Weyl spinor field failed, however (Klauder and Wheeler 1957).

> that distracts us from following out the implications of the theories we already have[45].

But for someone like Deser, coming from a high-energy nuclear physics background, it was a natural question to inquire whether Wheeler's methods could also be employed for the state-of-the-art theories of the nuclear interactions. As DeWitt writes, they found that this worked just fine for a massless[46] meson field, while there were some difficulties with a massive field[47].

While we have no details of these calculations, the difficulty is quite easily understood by looking at Misner and Wheeler's treatment of electromagnetic wormholes. Their study had consisted of two steps: First, ignoring the dynamics of the metric, they had studied the vacuum Maxwell equations in a doubly connected space[48] and verified the possible existence of wormholes. This study did not involve any reference to the Einstein equations or the dynamics of general relativity and worked just as well for the (vacuum) Yukawa (massive Klein-Gordon) field equations.

The second step in Misner and Wheeler's "Classical Physics as Geometry" was now to take into account also the dynamics of the metric. After having established the possibility of constructing charges from topology, they had then disregarded the topological aspects, treating a wormhole merely as two point charges, thereby neglecting the metric within the actual wormhole. Now, there was no exact solution of the Einstein equations for more than one massive particle, but, as Misner recalled:

> All through this time with Wheeler, I was pushing the idea that you could use the initial value problem in general relativity to get some results that were too complicated to get dynamically[49].

Misner had studied the work of André Lichnerowicz (1944) who had shown how to solve the initial value equations[50] for the metric of n massive bodies. For the joint paper with Wheeler, Misner had then generalized Lichnerowicz's solution of the initial

---

[45] John Archibald Wheeler papers, American Philosophical Society, Philadelphia, Box 182. In the remainder of the document, this collection will be referred to as JAWP.

[46] Of course the classical scalar field equation has no mass associated with it, only a characteristic length scale, which in the quantized field theory becomes the Compton wavelength of the field quantum. It has however, as witnessed by the report, been common for decades to call the classical field equation "massive" as well, and we do not wish to be particular.

[47] Tellingly, given Wheeler's objections to meson theory, when Misner reported back to Wheeler on his attempts to incorporate mesons in the their program, he did not mention the possible physical relevance, but only the calculational simplicity: "The scalar photon [i.e., the massless meson] makes a good example of how the process works, without nearly as much work as for electromagnetism". (Letter from Misner to Wheeler, 10 August 1957, JAWP, Box 18, Misner Correspondence)

[48] We use here Wheeler's terminology, where a "doubly connected space" is a space "with a handle", where two given points are "doubly connected" because there are two connections from one point to another, one through the bulk of the space and one through the handle. This is not to be confused with another possible usage of the term "doubly connected", implying a fundamental group of order 2. Spaces that are doubly connected in Wheeler's sense have an infinite fundamental group and would thus be categorized as infinitely connected in this latter terminology. A detailed historical study of the origins of this incompatible terminology is beyond the scope of this paper.

[49] Interview by Dean Rickles and Don Salisbury, 16 March 2011, Niels Bohr Library and Archives, American Institute of Physics, College Park, MD.

[50] In the report these are referred to as the Fourès equations, after Lichnerowicz's student Yvonne Fourès-Bruhat (later Yvonne Choquet-Bruhat) who had rigorously proven the well-posedness of the initial value problem in general relativity (Fourès-Bruhat 1952). See also (Lichnerowicz 1992).

to the case of n charged particles (wormhole mouths). Such a generalization was only possible for particles with a $1/r$ Coulomb potential, for which there existed an exact solution in the one-particle case, and thus could not be performed for a *massive* scalar field, with its exponentially decreasing Yukawa potential. This appears to have been the problem that Deser and Misner encountered in Copenhagen. It was, however, not a problem of principle, but one of solving the Einstein (or at least the initial value) equations for a particle creating a non-Coulomb potential.

It should be noted that the n-body (charged or uncharged) initial value solution not only makes no use of a wormhole topology, it also does not involve any singularities of the metric (as opposed to the Schwarzschild or Reissner-Nordström solutions) in order to ensure the mathematical well-posedness of the initial value problem. For this reason, in his adaptation of the Reissner-Nordström solution to an n-body solution of the initial value equations, Misner had first formulated the Reissner-Nordström metric in isotropic coordinates[51]. In these coordinates the singularity at the horizon occurs only in the time-time component g_{00} of the metric (i.e., in the lapse). Since g_{00} is unconstrained by the initial value conditions, the singularity could then be simply be smoothed over in the initial value reformulation, and also in its generalization to n bodies. The spatial Reissner-Nordström metric dl^2 in isotropic coordinates reads[52]:

$$dl^2 = \left[(1 + m/2r)^2 - (q/2r)^2\right]^2 \left(dr^2 + r^2 d\Omega\right). \tag{3}$$

Here, the mass m and the charge q are given in "natural units", where both Newton's constant and the speed of light are set equal to 1. For the case of several point charges, Misner replaced the spherical coordinates in the round brackets by Cartesian ones (the many-body problem is no longer spherically symmetric) and the conformal factor in square brackets by the more general expression $\left[\chi^2 - \phi^2\right]^2$, with χ a generalization of the Schwarzschild part

$$\chi = 1 + \sum_a \frac{\alpha_a}{|\mathbf{r} - \mathbf{r}_a|} \tag{4}$$

and ϕ a generalization of the Coulomb part

$$\phi = \sum_a \frac{\beta_a}{|\mathbf{r} - \mathbf{r}_a|} \tag{5}$$

where the sums go over all the charged point particles (wormhole mouths) located at the points with space coordinates $\mathbf{r}_a$ and the constants α_a and β_a are directly related to the masses and charges of those particles, albeit in a non-trivial manner. The topology of the initial three-dimensional hypersurface is then simply Euclidean with the set of points $\mathbf{r}_a$ removed.

In their paper, Wheeler and Misner had not commented on this fact. But the report gives some indication as to the reasons for the adaptation of the simplified topology: Apparently Misner *had* tried to construct a metric for an actual wormhole topology, already early on in his thesis work[53], but had failed. Revisiting this problem in Copenhagen, he finally succeeded. As he reported to Wheeler after the meeting:

> I looked again at the situation of a wormhole biting its tail ($S^2 \times S^1$, a torusy thing in 3-dim.) which was the first thing I did for you several years ago. Then

[51] Misner refers to these coordinates in the report as conformal coordinates, as the conformal flatness of the spatial Reissner-Nordström metric is manifest in these coordinates.

[52] There is a exponent of 2 missing on the square brackets in equation 243 of (Misner and Wheeler 1957).

[53] He had been working on topological issues as early as 1955, as witnessed by the abstract of his presentation at the APS meeting in April of that year (Misner 1955).

> I found no solution of the Einstein-Maxwell equations was possible, but this time I find them[54].

Misner would go on to work this out in detail after Copenhagen. His further work on the initial value problem for the wormhole would play an essential role in his contribution to the Arnowitt-Deser-Misner (ADM) collaboration, as we will discuss in Section 4.4. His renewed study of the wormhole initial conditions finally led to his 1960 paper, a two-page affair with no references at all to quantum gravity. We witness here an interesting feature of the renaissance of General Relativity, where a problem that first arises within a far-ranging research scheme, in this case referencing both unified field theory and the quantization of gravity, becomes divorced from these speculations and instead establishes itself as a research topic within the newly emerging field of general relativity proper. We saw in this section how topology divorced itself from the quantum problem, offering as it did interesting problems in its own right. In the next section, the first one to treat the actual quantization of gravity, we will see that the separation was in a sense mutual, with DeWitt trying to push back the Wheelerian speculations and the excessive complications introduced into the problem of quantizing gravity by dragging in topology, arguing instead for an isolation of the specific problem of quantum gravity.

4.3 Feynman quantization

Feynman quantization (in its application to field theory at least) was in 1957 still no more than a heuristic tool, far removed from any kind of formal rigor. The results that Feynman had obtained in quantum electrodynamics using his functional integral methods had really only been able to make an impact after Freeman Dyson had managed to re-derive them using more rigorous methods (the mainstream covariant techniques of relativistic quantum field theory)[55]. While drawing inspiration from Feynman's work for this re-derivation, Dyson recalls having strongly criticized Feynman's approach for its lack of rigor:

> I argued with him a lot, because I still had strong resistance to his way of doing things [...] I said to him: "Look, you've got to get the mathematics right, otherwise it doesn't make any sense. You've got to have some solid foundation in mathematics. It's alright to draw the pictures, but..."[56]

Still, Feynman quantization did not entirely fade from view in the 1950s, both because of its conceptual interest[57] and the promise it appeared to hold for providing non-perturbative methods in quantum field theory, so sought after in order to get a handle on the strong-coupling nuclear theories of the day.

---

[54] Letter from Misner to Wheeler, 10 August 1957, JAWP papers, Box 18, Misner Correspondence. The letter is written so long after the conference (and sent from Salzburg) as Misner went on a tour of Europe with his brother after the Copenhagen workshop (Email from Charles Misner to one of the authors (AB) of 1 April 2016).

[55] See (Schweber 1994).

[56] Web of Stories interview with Dyson, http://www.webofstories.com/play/freeman.dyson/71.

[57] As witnessed by the panegyrics in an early paper by Yoichiro Nambu (1950): "[Feynman's] ingenious method is indeed attractive, [...] because of its way of thinking which seems somewhat strange at first look and resists our minds that are accustomed to causal laws. According to the new standpoint, one looks upon the world in its four-dimensional entirety. A phenomenon that will come into play in this theatre is now laid out beforehand in full detail from immemorial past to ultimate future and one investigates the whole of it at glance".

As DeWitt outlines in the report, one of the main proponents of Feynman quantization and its application to *gravity* was John Wheeler, who had of course been Feynman's thesis advisor. He had suggested the Feynman quantization of general relativity as a thesis topic to Misner, indicating that this might be something that might be done within half a year. Misner had jumped at this chance, given that the other option was to write a 700-page thesis on axiomatic quantum field theory with Arthur Wightman[58]. Misner had then, however, soon switched gears and was mainly working on the so-called "already unified theory" (electromagnetism without electromagnetism) (Misner and Wheeler 1957), until it was pointed out to Wheeler and Misner by Bergmann that this theory had largely already been worked out by George Rainich thirty years earlier. In order to present original work for his PhD thesis, Misner had returned to the work on Feynman quantization in early 1957 and had, indeed within a few months, written a short thesis on the matter, which was published in the special issue of Reviews of Modern Physics dedicated to the Chapel Hill conference, despite the fact that Misner had not presented this work of his there[59].

The aspect of Feynman quantization mainly discussed in Copenhagen, however, had been done early on by Misner, before getting distracted by the already unified theory[60]. Wheeler had selected Misner for the problem of Feynman quantization because of his mathematical prowess, so it was an obvious candidate for the first problem to tackle: The measure involved in the functional integration over metric configurations. In his thesis work, Feynman had only treated integration over actual paths, i.e., over the time-dependent position of a quantum-mechanical particle. There were two ways to extend this to field theory: The first was to Fourier decompose the field and then treat each Fourier mode as a quantum mechanical oscillator to be treated according to Feynman's original quantum mechanical rules. This was what Feynman (1950) himself had done when he first applied his methods to QED proper. However, as Misner pointed out in his thesis, this again requires resorting to perturbations around a flat metric, since only then does one have a complete set of solutions (to the linearized vacuum Einstein equations) to do Fourier analysis with.

Wheeler had thus suggested a different way to do the functional integration for the case of gravity, namely to do *spatial* (rather than k-space, as in the Fourier decomposition approach) discretization and then integrate over the field values at the discrete space points x_i in the usual, quantum mechanical, Feynman manner. The most straightforward way to do this was simply to take as measure the product over all $dg_{\mu\nu}(x_i)$ and then integrate in a straight line from the values of the metric at the beginning of an infinitesimal time interval to their values at the end[61]. But as Misner had already pointed out in his thesis, the measure made no distinction between metrics with different signatures and thus there was no indication how it should return the Minkowski metric as the classical limit in the vacuum case. Misner had thus proposed a measure that also contained an inverse power of the negative metric determinant $(-g)$, which would become singular for processes in which the metric tried to change its signature (i.e, went through a point where the metric determinant was zero), concluding that

> The range of integration [...] is the connected region of $g_{\mu\nu}$-space containing $g_{\mu\nu} = \mathrm{diag}(-1, 1, 1, 1)$ and bounded by $\det g_{\mu\nu} = 0$.

[58] See interview of Misner by Chris Smeenk, 22 May 2001, https://www.aip.org/history-programs/niels-bohr-library/oral-histories/33697.

[59] On this story see (Wheeler and Ford 2000, p. 267–268) and interview with Misner mentioned in footnote 49.

[60] See (Misner 1957, p. 497).

[61] The same infinitesimal segmentation of the total time interval had been done by Feynman for quantum mechanics.

Misner had identified $(-g)^{-5/2}$ as the correct power of the negative metric determinant to appear in the measure, using his notion of homogeneity group – in essence demanding that the measure for the integration over the metric variables at some point x_i be invariant under the group of coordinate transformations at that point[62]. His argument had involved no recourse to the actual Lagrangian determining the dynamics of the metric, a feature he had explicitly highlighted in his thesis. From DeWitt's report we can gather that Laurent in Copenhagen had managed to derive Misner's measure from a different line of argumentation, based on the explicit structure of the Lagrangian; but no details are given and Laurent never published on the matter[63]. Misner's measure is obviously not the only measure to become singular for a change of signature, and DeWitt had apparently constructed a "preliminary measure" which would *also* be singular in case the initial or final hypersurface of the transition amplitude being calculated ceased to be space-like; he would eventually come to dismiss Misner's measure and the entire notion that the measure could be directly determined from the invariance group altogether (DeWitt 1962b)[64].

The main focus of the discussions in Copenhagen appears to have been the relation of these formal considerations to the more speculative hopes placed in the Feynman quantization approach, especially by Wheeler. Wheeler (1957, p. 607) had speculated that at scales on the order of the Planck length multiply connected spaces would "contribute importantly to the sum over histories". The question was now how this was to be reconciled with his PhD student's cautious formal attempts to construct a measure, which implied a "natural barrier" to the integration domain. This question was discussed extensively in Copenhagen, but with no definite outcome. Here, the division between DeWitt's and Wheeler's approaches clearly showed (see also Sect. 4.5): While Wheeler wanted to tease out the extreme and conceptually interesting features of a quantum theory of gravity, DeWitt repeatedly insisted that the formal questions needed to be settled first[65].

Given this formal bent, it is no wonder that DeWitt had found himself in Lille just a week before the Copenhagen workshop[66]. Here, from June 3-8, 1957, a conference on mathematical problems of quantum field theory had been held that is widely regarded as one of the founding events of axiomatic and algebraic field theory, and even of modern mathematical physics as a subdiscipline of physics[67]. Here, it appears, DeWitt had spoken with Julian Schwinger, his former PhD advisor, about another quantization procedure to be applied to general relativity, Schwinger's own action

[62] The power of 5/2 essentially arises as number of independent components of the metric (10) times rank of the metric tensor (2) over (dimension of spacetime (4) times weight of the metric determinant (2)).

[63] Misner, in a letter to Wheeler of 10 August 1957 (JAWP, Box 18, Misner correspondence), reporting on the workshop, stated that he "had some difficulty understanding" Laurent's arguments.

[64] This dismissal was accepted by Misner, as he outlines in his interview with Rickles and Salisbury, see footnote 49.

[65] DeWitt was far more explicit on the divide between his views and Wheeler's in the first version of the report, where he wrote: "[I]n contrast to Professor Wheeler's present views [...] we should prefer to work with a simple topology and a simple integration domain". (DF, Box "Institute of Field Physics", unnamed folder 7)

[66] DeWitt is not listed in the official list of participants of the conference reprinted in (Fredenhagen 2010), but the full report contains one comment by DeWitt in discussions (NN 1959). Schwinger is listed as a participant (though with the wrong initial), but his talk appears not to have gone well (see Mehra and Milton 2000, p. 381, and letter from Källén to Pauli of 10 July 1957, printed in von Meyenn 2005).

[67] See (Fredenhagen 2010), Wightman's remarks in (von Meyenn 2005, p. 1059, fn 4), and (Haag 2010, p. 274).

principle. As opposed to the covariant and Feynman quantization, Schwinger's method had not arisen from an attempt to solve the difficulties of quantum field theory. Rather, it had been developed by Schwinger in the early 1950s in response to the great success of Feynman's methods in quantum electrodynamics, as an attempt to put Feynman's heuristic quantization procedure on surer mathematical footing[68]. As Misner had pointed out in his article on Feynman quantization, it had not yet been applied to gravity.

It is hardly surprising that DeWitt, champion of a pluralist approach to quantum gravity, took the next step and quizzed the master himself on the potential of the Schwinger method in quantum gravity. But DeWitt appears not to have been particularly impressed with this potential: Schwinger's method just did not seem practical enough and indeed had not yet shown its merits by leading to new physical results (as opposed to the other three quantization methods discussed above and below). The main point discussed in Copenhagen regarding the Schwinger method was that it required[69] one to cast the field theory to be quantized into a form where the Lagrangian was only a first-order polynomial in the derivatives of the field variables. Such a reformulation existed for general relativity, using the connections as independent dynamical variables in addition to the metric, and goes by the name of Palatini formalism[70].

But this was not much discussed in Copenhagen, making the question of Schwinger quantization a mere appendix to the discussion of Feynman quantization. The only issue discussed in any detail was a "curious ambiguity", which probably refers to a difficulty first discussed by Weyl (1950): In the regular Palatini formalism, the usual relations between connection and metric appear as the Euler-Lagrange equations obtained from varying the connection variables. This is spoiled, however, by the inclusion of Dirac spinor fields, minimally coupled to the gravitational field using tetrads and infinitesimal rotations of the local tetrad frames in place of metric and connection. As Weyl had pointed out equivalence between the Palatini and the regular metric formulation could be re-established by adding to the Dirac-Einstein-Hilbert Lagrangian in the regular formulation an additional term quartic in the spinor field[71]. But even this discussion of an interesting aspect of Schwinger quantization and the use of the Palatini formalism it implied does not seem to have led to any novel conclusions in Copenhagen.

---

[68] See a letter from Freeman Dyson to Rudolf Peierls of 23 September 1950: "He [Schwinger] translates the Feynman Lagrangian formalism with "integration over histories" into a rigorous and conventional language. [...] You have to search carefully how [this] is hidden in the notes. Of course, the name of Feynman is never mentioned. Only I happen to know from other sources (as is also obvious when you see what Schwinger's method actually is) that Schwinger started the whole thing from making Feynman's method intelligible to himself". The letter is reproduced in (Lee 2009).

[69] This is not quite how DeWitt expresses it in his report, and indeed formally the Schwinger method could be applied to Lagrangians involving higher powers of the field derivatives. Indeed, several years before, the Syracuse group had attempted to apply Schwinger's method to general relativity in the usual formulation, only to run into the factor ordering complications discussed by DeWitt (Bergmann and Schiller 1953, section 5). Schwinger himself by the late 1950s regarded the construction of a first-order formulation as an essential prerequisite for his method. This is witnessed by references to the Schwinger method in papers by former Schwinger students, e.g., (Arnowitt and Deser 1959) or, in a non-quantum-gravity context, (Glashow 1959).

[70] Despite having been developed by Einstein himself, see (Ferraris et al. 1982).

[71] The highly plausible identification of the "curious ambiguity" with the difficulty discovered by Weyl was made by Stanley Deser in an email to one of the authors (AB) of 22 March 2016.

In the years immediately following the Copenhagen workshop, however, the Schwinger approach would deliver much more fruitful results than the Feynman one: Just one year later, Deser and another Schwinger student, Richard Arnowitt, began to tackle the quantization of gravity using the Palatini formalism and Schwinger quantization, soon joining forces with Misner (who had, after finishing his thesis, abandoned the idea of Feynman quantization) to construct, not a quantum theory of gravity, after all, but the ADM formalism[72].

4.4 Canonical quantization

As DeWitt writes in the report, the canonical quantization of the gravitational field was at the time a business primarily pursued by the group of Peter Bergmann in Syracuse. Bergmann's research program had been initiated in the late 1940s as a conservative alternative to Einstein's unified field theory program (on which Bergmann had worked in the 1930s and 1940s), a research program based on the established structures of both general relativity and quantum theory, in particular on the "tried and (so far) true method" of canonical quantization applied to the classical field theory of general relativity. The first order of business had been a reformulation of that classical theory in canonical (Hamiltonian) form, which had been achieved around 1950 in a two-step process: First, the Syracuse group had constructed a Hamiltonian formulation of general relativity with full four-dimensional symmetry (Bergmann et al. 1950), using the so-called parameter formalism. In a second step, they had then turned to (and successfully constructed) a 3+1-decomposed formulation, due to the difficulties of the four-dimensional formulation, both calculational and conceptual (Penfield 1951)[73].

Several years later, in 1957, Syracuse was still the hub of research in canonical quantum gravity, even though several of Bergmann's graduate students had turned away from the problem of quantum gravity and immersed themselves in problems within general relativity proper, such as gravitational waves or the problem of motion – a characteristic shift in focus of the renaissance period, which we have already observed in the case of Misner and the Wheeler school. Bergmann on several occasions had formulated what he took to be the central outstanding issue in the canonical quantization program, namely the identification of the "true observables". This question arose because of the general covariance of the Einstein field equations, which Bergmann had identified early on as the essential feature of general relativity that he believed would carry over into a quantum theory of gravity. General covariance implied that in the canonical formulation of the theory not all of the phase-space variables could be taken as independent dynamical quantities. Rather, there existed in the classical theory non-dynamical identities involving the phase-space variables, the so-called constraints, which had to be fulfilled in addition to the equations of motion. There were eight such constraints in total, four "primary" (arising directly from the defining relations for the canonical momenta), and four "secondary" (arising from the demand that the primary constraints be conserved in time). These constraints had been identified by the Syracuse group in 1950/51, along with the Hamiltonian, thereby establishing a full canonical formulation of general relativity.

The general setup had been quite familiar already in 1950 from quantum electrodynamics, where a similar, though less involved, situation occurred due to gauge

[72] See (Deser 2015).

[73] On this early history of the canonical approach to quantum gravity, see (Blum and Rickles 2017, section 5).

invariance. In QED, the next step had now traditionally[74] been to solve the constraints, leading to a modified Hamiltonian that now only contained independent physical degrees of freedom; in the case of QED, these were the transverse modes of the electromagnetic field (photons). This elimination of redundant degrees of freedom had, in QED, originally been performed only after quantization (Fermi 1932; Oppenheimer 1930), but it had later been established that this could also been done pre-quantization, identifying the elimination of the constraints as the choice of a certain gauge, the Coulomb or (in the absence of charged matter) radiation gauge (Heitler 1944).

After the establishment of the Hamiltonian formulation of general relativity in 1950/51, it was clear to Bergmann that the next step should be an elimination of the constraints, in analogy to (quantum) electrodynamics.

> [B]uild up the Hilbert space from only those states that satisfy all constraints and [...] make it, thus, deliberately, a small subspace of the functional space of all conceivable wave functionals without regard to constraints, [...] severely restricting the number of linear operations that may properly be called Hilbert operators. [...]
> Will the elimination of a large number of operations from consideration not embarass us as physicists, by eliminating the mathematical description of physically meaningful quantities?
> In answer, we find that the observables not ruled out are invariants, quantities that remain unchanged under all infinitesimal transformations with respect to which the theory is assumed to be invariant. For electrodynamic quantities, for instance, commutability with the subsidiary conditions of quantum electrodynamics implies gauge invariance. And truly, any quantity that can be given a well-defined numerical value must be an invariant.
> We have gone through a number of examples to convince ourselves that any physically observable quantity is an invariant, but this point is so obviously of major importance that it should be more fully investigated. Suffice it here to say that the point of view we have adopted is a generalization of and consistent with accepted practices in quantum electrodynamics and elsewhere. (Bergmann and Schiller 1953, pp. 12-13)

At the 1955 Pisa conference on elementary particle physics, Bergmann had introduced the name "true observables" for these invariant quantities, the operators corresponding to which should be the only ones operating in the reduced Hilbert space of constraint-obeying wave functionals. He further asserted that:

> It would appear that particularly in theories possessing general covariance the determination of the true observables is a necessary preliminary to their quantization. Unfortunately this determination remains so far an unsolved problem. (Bergmann 1956, p. 1178)

Bergmann thus believed that in quantum gravity one did not have the liberty of eliminating the constraints before or after quantization, as in quantum electrodynamics. Rather, one was forced to perform the reduction of the degrees of freedom already pre-quantization. In order to understand how Bergmann came to this conclusion, it is important to realize that "general covariance" for him was a somewhat wider notion that encompassed all those properties of general relativity which he considered essential and worthy of conserving in the transition to the quantum theory. And besides the invariance with respect to general coordinate transformations, this also meant

[74] This procedure had been replaced in the 1950s by more sophisticated, covariant techniques, such as the Gupta-Bleuler method.

the non-linearity of the theory. Such a theory would in its Hamiltonian and its constraints include terms non-linear in both the canonical coordinates and the canonical momenta and there was no unambiguous symmetrization procedure for transferring such an expression into the quantum theory. Indeed, for any given quantization rule for such expressions, it was to be expected that the commutator algebra of the constraints and the Hamiltonian would not close, destroying the self-consistency of the theory. This was the so-called factor ordering problem, and Bergmann hoped to avoid it by finding the true observables before quantizing:

> Once the observables have been determined, then presumably any Hermitian combination, which in the limit $\hbar \to 0$ goes over into the classical Hamiltonian represents a formally possible quantum theory. Then it will be necessary to develop new physical (rather than formal) criteria for the appropriate factor sequence. (Bergmann and Goldberg 1955, p. 538)

DeWitt had been interested in Bergmann's program early on. Initially, as a graduate student, he had approached the problem of quantum gravity from the perspective of the new covariant methods in quantum field theory, developed in the late 1940s by, among others, DeWitt's advisor Julian Schwinger. This approach – as discussed in Sections 2.1 and 2.2 – fundamentally relied both on perturbation theory and on Lorentz covariance, implying as a starting point not the full Einstein equations, but rather the linear approximation, supplemented as necessary with higher-order non-linear correction terms. As his thesis work was contemporaneous with the successful establishment of the Hamiltonian formulation of full general relativity by the Bergmann group (and the Toronto/Pittsburgh group consisting of Alfred Schild and his graduate students Felix Pirani and – later – Ray Skinner), the paper based on his thesis which he submitted to the Physical Review in 1950, had been effectively rejected as too long and detailed given its purely approximative approach. DeWitt had wholeheartedly accepted this criticism and delved into the canonical approach to quantum gravity, writing two papers on canonical quantum gravity, dealing, respectively, with the factor ordering problem (DeWitt 1952) and the inclusion of spinorial quantum fields (DeWitt and DeWitt 1952)[75].

DeWitt's work was based on the early Hamiltonian formulations, which displayed full four-dimensional symmetry, using the parameter formalism first developed by Paul Weiss in the 1930s[76]. As already mentioned, this formalism was soon abandoned by both the Syracuse and the Pittsburgh groups: Not only were the calculations difficult, it was also unclear physically how to interpret the fact that the space-time coordinates (as functions of the auxiliary parameters) had to be taken as non-commuting q-numbers, in addition to the metric tensor field. This step of abandoning the parameter formalism was, however, not followed by DeWitt, who briefly left the field in 1952, at about the time that the Pittsburgh group turned to the non-parameterized 3+1-formulation (see section 2.2). Returning to his work on quantum gravity in 1956, the interconnections between different approaches to the problem became a leitmotif for DeWitt. Initially focusing on the Feynman sum-over-histories (path integral) approach (see the preceding section), he returned to working on the canonical quantum gravity program after the 1957 Chapel Hill conference he had organized[77]. Here, DeWitt

[75] See (Blum and Rickles 2017).

[76] On the history and details of this formalism, see (Rickles and Blum 2015).

[77] The strong interconnection that DeWitt observed in his work on different quantum gravitational formalisms is also witnessed by his brief remark that his work on the canonical approach would shed light on the problem of the "true measure" arising in the path integral approach, by, it is assumed, identifying the true dynamical degrees of freedom.

picked up right where he had left things in 1952, using the parameterized formalism that had been abandoned elsewhere[78].

This anachronism was no problem in Copenhagen, where no representatives of the cutting edge of canonical quantization were present. Bergmann's (and Wheeler's) presence had originally been anticipated[79], but the final roster saw not a single (even former) member of the Syracuse group. This absence of experts in the field of constrained Hamiltonian dynamics, however, also implied that the details of DeWitt's work did not make much of a splash. In an informal report which Misner wrote for DeWitt, but which he only quoted in the first version of the report, not the final one, Misner stated that

> I learned to understand the basic ideas of canonical quantization of general relativity, which had escaped me on several previous attempts at reading the published research in this problem[80]...

But, as we will see later, it would take a second exposure for the relevance of these ideas to really sink in. Misner would later recall:

> I don't think I absorbed then the Dirac style primary/secondary constraint ideas Bryce gives in the report[81].

What then was this work of DeWitt's? There had already for some time been the hope that one might find a canonical transformation in the Hamiltonian formulation of general relativity that would turn (some of) the constraint equations into "pure momenta" that is into the form "one canonical momentum variable equal to zero". One would then quite easily be able to eliminate the constraint and the canonical momentum variable in question (along with its conjugate canonical coordinate) from the theory. Jim Anderson (who had been a PhD student of Bergmann's) recalls having thought along these lines already in the summer of 1954[82] and the Copenhagen report supports this view, naming Anderson as the one to have "conjectured the possibility of such a transformation". In Copenhagen, DeWitt now presented his derivation of just such a transformation which turned the primary constraints of the theory into pure momenta, thereby making a big step towards solving the central difficulties of the canonical quantization program outlined above.

The open question was whether such a transformation into pure momenta also existed for the secondary constraints, thereby completing the elimination of constraints in the theory. DeWitt could present no final results concerning this question, only a general theorem which stated that such a transformation would exist only if the Lie

[78] It should be noted that DeWitt was well aware of the later developments in the canonical theory. Indeed, he often uses terminology that really only makes sense in the unparameterized formalism. In particular, when he speaks of primary constraints, he means only those primary constraints that also show up in the unparameterized formalism. The additional primary constraints appearing in the parameterized formalism DeWitt does not attempt to eliminate. After all, this would simply be deparameterizing. We will adopt his terminology – hence, in the following, primary constraints refers only to the four primary constraints associated with canonical momenta for the metric variables.

[79] As witnessed by a letter from Klein to Wheeler of 20 April 1957, Relativity Notebook V, JAWP.

[80] First version of the 1957 Copenhagen report. DF, Box "Institute of Field Physics", unnamed folder 7.

[81] Email to one of the authors (AB) of 1 April 2016.

[82] Interview with Jim Anderson by Dean Rickles and Donald Salisbury, 19 March 2011.

group generated by the subset of constraints being eliminated[83] was Abelian[84]. But he was unable at the time to determine whether, in fact, the Lie group generated by the secondary constraints was Abelian, due to the underdeveloped state of the mathematical theory of "infinite Lie groups"[85].

But, as already mentioned, these exciting developments within the canonical quantization program hardly made an impression on DeWitt's audience. Rather, their main focus appears to have been on DeWitt's outdated mathematical tools, namely the parameter formalism. Whence this interest in DeWitt's parameters? On the one hand, the emphasis on the parameter formalism in the report appears to be due to DeWitt's own interests. He had apparently convinced himself that the abandonment of the parameter formalism by the other groups working on canonical quantization had been misguided. In particular, he believed that the use of the parameter formalism would help in finding the canonical transformation that turned the secondary constraints into pure momenta.

This claim of DeWitt's is somewhat hard to assess. He appears to indeed have eliminated the primary constraints in the parameterized formalism and to have found the Hamiltonian and the secondary constraints in the now reduced phase space[86]. DeWitt's "Hamiltonian" (Equation 3) looks somewhat curious, with its free Lorentz index – for any specific choice of parameterization, a specific linear combination of DeWitt's φ_μ, which are just are the constraints derived from the defining relations for the momenta canonically conjugate to the space-time coordinates, would be singled out as the actual Hamiltonian. This splitting up of the Hamiltonian makes the calculation of the secondary constraints somewhat cumbersome (16 Poisson brackets need to be calculated, for only four secondary constraints), but it does the job. The resultant secondary constraints of the parameterized formalism have some unusual features; in particular they are, as opposed to the unparameterized case, linear in the momenta, which gave DeWitt some hope, that they, too, could be transformed into pure momenta. And in a grant proposal to the Research corporation one year later[87] DeWitt would in fact claim that he was able to show quite generally that in the parameterized theory the Lie group generated by both the primary *and* the secondary

[83] Since the commutator/Poisson bracket of two constraints is a linear combination of constraints, the constraints can be viewed as elements of a Lie algebra.

[84] Apparently Anderson had conjectured that such a transformation would always exist for the primary constraints of the theory and that the only problem was to find it. DeWitt now stated that this only happened to be the case for general relativity, where the algebra generated by the primary constraints was indeed Abelian.

[85] That is, in particular, gauge symmetries. As DeWitt points out, the importance of such groups had also been stressed by Wolfgang Pauli in a series of lectures on group theory that he had given for the theory division of CERN in Copenhagen in September 1955 (Pauli 1965, p. 86). The lecture notes were first published as a CERN report in 1956, so they were certainly available in Copenhagen in Summer 1957. Pauli's emphasis was likely influenced both by Pauli's engagement with non-abelian gauge theories in 1953/54 (Straumann 2000) and with his work on updating his book on general relativity in 1955/56 (von Meyenn 2001, pp. 498ff).

[86] Both Bergmann (with Johanna Brunings) and Pirani and Schild had originally falsely believed that there were *only* primary constraints in the parameterized theory (Blum and Rickles 2017).

[87] Since Wheeler was one of DeWitt's references, this proposal can be found in the John Archibald Wheeler papers, Box 7, DeWitt correspondence. DeWitt's signature is dated 5 June 1958.

constraints was Abelian and that he had thereby managed to turn the secondary constraints into pure momenta as well[88]. But he went on:

> Having completed this stage of the work, however, DeWitt is now running into new difficulties. Because of the complexity of the equations, he is finding it very difficult to remove all of the non-observable quantities from the Hamiltonian of the system. Part of the trouble is that he now, at this point, has a lack of general principles which would guide him and enable him to see the forest instead of the trees. Stimulated by his modest success with the aid of group theoretical principles, DeWitt is now undertaking an investigation of the theory of infinite dimensional Lie groups.

DeWitt had thus finally, despite all his computational prowess, run into the same difficulties of complexity that had led others to abandon the parameter approach almost a decade earlier. A first publication on infinite dimensional Lie groups followed three years later (DeWitt 1961) and made no more reference to the parameterized formalism, leaving a test of the validity of DeWitt's claims as a challenge to contemporary physicists. Similarly, it is an open question how DeWitt's remarks on the nonlocality of the true observables (and especially the speculative footnote on the possibility of obtaining fermionic spinors) are to be interpreted in the usual, non-parameterized form of the theory[89].

In any case, as already mentioned several times, there was no-one at the Copenhagen workshop who could or wanted to delve into these details of constrained Hamiltonian dynamics. The main interest of the other participants appears to be summarized in the last paragraph of this section, the only place where a name other than DeWitt's is mentioned. The question being discussed was what was later coined the "problem of time", that is the vanishing of the Hamiltonian in General Relativity. In the parameterized formalism, the Hamiltonian always vanished, for the reasons that DeWitt discussed: It generates the dynamics based on the parameter T rather than on the physical time. This was well-known and was initially believed to be an artefact of the parameterized formalism which would vanish in the non-parameterized 3+1 formulation, as stated, e.g., by Bergmann and his PhD student Ralph Schiller:

> [T]he introduction of parameters [...] does not lead to serious modifications. Most important, the Hamiltonian becomes itself a constraint. There is, therefore, no "motion". Any state that obeys all constraints is a solution of the Schrödinger equation, provided we do not permit it to change in the course

[88] The authors have not been able to determine whether the proposal was accepted. Certainly Wheeler's lukewarm recommendation did not help. On 9 December 1958, he wrote to the Research Corporation (JAWP, Box 7, DeWitt correspondence): "I would not be honest if I tried to indicate that DeWitt is one of the top three investigators in the field of general relativity in the world; he would be more like number six in this country". This was not, it should be added, Wheeler's last word on the matter. Another two years later, on 29 November 1960, he would write to the Chairman of the Department of Physics at the University of North Carolina (ibid.): "I would say that no one in the world has at the same time a wider command of the mathematical techniques of relativity theory plus quantum theory – together with the drive to carry through more complex calculations to a definitive end – than does Bryce DeWitt".

[89] Though DeWitt was certainly not alone at the time with his conjecture concerning the non-locality, see, e.g., a letter from Bergmann to Pauli of 17 November 1957 (Bergmann Papers, Syracuse), where he states that he had also believed that the true observables in general relativity should be non-local. Note that this letter is not reprinted in the Pauli Correspondence volumes, despite the fact that Bergmann appears to have sent it to the editors.

> of "time". This apparent freezing is, however, a purely formal result of the introduction of parameters. (Bergmann and Schiller 1953, p. 13)

In the following years, however, the suspicion began to emerge that also in the non-parameterized Hamiltonian formulation of general relativity, where the Hamiltonian density gives the dynamics based on the coordinate time t, the Hamiltonian would vanish, as remarked again by Bergmann in the closing sentence of (Bergmann and Goldberg 1955) and discussed at the 1957 Chapel Hill conference (DeWitt and Rickles 2011, p. 191 and p. 231). The question also arose in Misner's PhD thesis on the path integral quantization of general relativity, where he had derived the Hamiltonian density operator $\mathcal{H}_{\rm op}$ appearing in the Schwinger-Tomonaga (relativistic Schrödinger) equation and found it to be zero. Misner went on to write that:

> Because of different methods of definition, our $\mathcal{H}_{\rm op}$ Hamiltonian is not necessarily the operator corresponding to the Hamiltonian defined in classical theory. However Professor J.L. Anderson at the Chapel Hill conference voiced suspicions that the classical Hamiltonian in general relativity would be zero. (Misner 1957, p. 508)

It appears now that at the Copenhagen workshop a consensus was reached that indeed Misner's Hamiltonian operator did not correspond to the classical Hamiltonian, but was in effect the Hamiltonian of the parameterized formalism, which vanished trivially. Misner's work thus had no bearing on the actual "problem of time", which then was not discussed any further in Copenhagen. This view clearly was adopted by Misner himself after Copenhagen: When Wolfgang Pauli wrote to him later in 1957, inquiring as to what the physical implications of the vanishing of Misner's Hamiltonian was, he replied that it was a mere triviality:

> It comes down to saying that [...] the Schwinger-Tomonaga equation considers variations that are as meaningless physically as coordinate transformations, and therefore produce no change in the state-vector. I think you, Bergmann and I all agree that a Hamiltonian density, if defined in this way, must vanish[90].

While this was the main upshot of DeWitt's presentations on canonical quantization in Copenhagen, his work would make a bigger splash several months later, when he presented it to a more receptive audience at the Stevens Institute of Technology in Hoboken, NJ. His host, Jim Anderson (1958), immediately picked up on DeWitt's work and presented the canonical transformation that allowed for the elimination of the primary constraints in the non-parameterized formalism just two months later. Anderson's work was, however, soon marginalized, as Paul Dirac (1958) had obtained the same results at exactly the same time (though, it appears, independently of both DeWitt and Anderson) and taken them quite a bit further. But another member of the audience at the Stevens Institute was now a lot more receptive to DeWitt's presentation than he had been in Copenhagen: Charles Misner.

According to Misner's recollections (which are based on the detailed notebooks he kept in the months after the Copenhagen workshop), he had turned in the fall of 1957 to the wormhole problem (see sect. 3.2):

> After the '57 summer I started a research notebook (Wheeler style) on 9 November 1957 when an ONR (Office of Naval Research) grant through the Math Department in Princeton [...] was supporting my research. There [...] I turned to trying to make Wheeler's "wormhole" sketch rigorous. One step was to find the initial conditions, done by early January 1958. The next step would

[90] Letter from Misner to Pauli, 19 November 1957, Bergmann Papers, Syracuse. Note that this letter is not reprinted in the Pauli Correspondence volumes.

> be to numerically evolve it in time. A page long note from an IBM session outlines on 31 December 1957 what this involves, including the question of how to tell the computer to lay out coordinates in the future as time progresses beyond the initial conditions. Both the coordinate choice problem, and how to take a step forward in time were studied[91].

Hearing DeWitt's presentation at the Stevens meeting, Misner now realized that this problem was closely related to DeWitt's work on eliminating constraints[92]. And several months later, Misner was really bringing his expertise on the matter to bear on what would become known as the ADM collaboration, culminating in the 1962 review article that is now primarily cited (Arnowitt et al. 1962).

4.5 Approximation methods

This section (we will return to the question of the idiosyncratic title in due time) deals mainly with the work of Deser (and Laurent) on rendering the quantum field theories of the non-gravitational (electromagnetic and nuclear) interactions finite by including the effects of the quantized gravitational field. While attributed to Klein by DeWitt, the origin of the idea is rather to be found in the work of Lev Landau and Wolfgang Pauli.

Landau's involvement in the measurability debate with Bohr and Rosenfeld in the early 1930s has already been discussed in Section 4.1. After this defeat, Landau essentially abandoned quantum electrodynamics for two decades, returning to it only in 1953, instigated by his young collaborators Alexei Abrikosov and Isaac Khalatnikov (Abrikosov 1973). Landau originally believed that they could, by using an approximation technique that went beyond the usual perturbation theory, prove that QED was in fact a perfectly finite theory, even as one let the momentum cutoff Λ go to infinity. But soon he realized a sign mistake and had to draw an entirely different conclusion: His approximation was in fact inconsistent for large cutoffs (and *a fortiori* for infinite cutoffs, i.e., for a fully Lorentz-invariant theory) (Ioffe 2012). But rather than viewing this as a mere defect of his approach, Landau revived the attack against relativistic quantum field theory that he had dropped twenty years before, and in a follow-up paper written together with Isaac Pomeranchuk – head of the theory division at the Moscow Institute for Theoretical and Experimental Physics and the leader of Soviet QFT research up until Landau's re-entering the field – drew much more radical conclusions from the anomaly (nowadays known as "Landau Pole") that he had identified, namely that the physical (renormalized) charge in QED would always be zero, independent of the value of the bare (unrenormalized) charge, that, in other words, QED was a non-interacting quantum field theory (Landau and Pomeranchuk 1965).

These views reached Wolfgang Pauli in Switzerland through several channels – indeed Landau had worked hard to communicate his ideas to Western physicists, as this slowly became possible in the years after Stalin's death. Landau had invited Pauli's Swedish collaborator Gunnar Källén to a Moscow conference on QED in the Spring of 1955 (Jarlskog 2014) and also wrote a contribution to the *Festschrift* celebrating Bohr's 70th birthday in 1955 (Pauli 1955), a contribution that the editors (which included Pauli) had solicited without really expecting it to materialize[93]. Pauli was quite happy to hear of Landau's views, which coincided with the conclusions he

[91] Email to one of the authors (AB) of 1 April 2016.

[92] Interview with Misner and Dieter Brill by Dean Rickles and Don Salisbury, 16 March 2011.

[93] Letter from Léon Rosenfeld to Pauli of 24 May 1954, published in (von Meyenn 1999).

himself had drawn from his recent investigations into the structure of renormalized quantum field theory with Källén:

> I am of course satisfied to see that he [Landau] has the same suspicions concerning quantum electrodynamics that I do. [...] From what I have read of the paper, it seems probable to me, however, that he has as little proof for this conjecture as I do. [...] So it appears to me: Landau's "nose" is still good, but mathematical proofs were never his strong suit[94].

But what intrigued Pauli most, was Landau's attempt to connect the failure of QED with gravitation. As Landau had written in paper for the 1955 Bohr *Festschrift*:

> Of course, no unambiguous physical conclusions can be drawn from the result obtained, that the point [i.e., infinite cutoff] interaction is zero in the case of electrodynamics. The energies Λ [...] are in every case very large. At these energies, the effects of gravitational interaction may exceed the electromagnetic effects, so that a discussion of electrodynamics as a closed system becomes physically incorrect. The idea is very attractive that this "crisis" in electrodynamics occurs for just those energies where the gravitational interaction is comparable with the electromagnetic. (Landau 1955, p. 60)

Landau gave no indication as to how gravitation might solve the zero charge crisis. Neither did Pauli. But he did have some ideas on how the inclusion of gravitation might affect quantum field theory in general. And just a few months after hearing of Landau's arguments, Pauli had the opportunity to present his ideas on the relevance of general relativity to field theory to the most interested audience imaginable, at the 1955 Bern relativity conference. This underlines one more time the catalytic role of the 1955 conference, both in bringing older physicists back to general relativity – Pauli of course having made his fame as a young man with a review article on general relativity – and introducing younger physicists to the theory. For in the audience was a young postdoc with a training in Schwingerian quantum field theory, but hardly any prior exposure to general relativity: Stanley Deser.

Pauli presented his ideas in the discussion following the talk by Oskar Klein. Klein was another older physicists, who had been prompted to return to research on relativity by the upcoming Bern meeting[95]. He had been working on the five-dimensional theory (that now of course bears his name, along with that of Theodor Kaluza) on and off, ever since his pioneering work in the mid-1920s. In his Bern lecture (Klein 1956), Klein, too, now spoke of the connection between quantum electrodynamics and general relativity. His early work on the five-dimensional theory had been concerned with establishing the connections between general relativity and the newly emerging wave mechanics, and he was now attempting to do a similar thing with the new renormalized quantum electrodynamics, a theory he was only now really starting to learn[96]. Klein's main argument for such a connection was that the discretization of the electric charge, obtained by making the wave functions of charged matter waves periodic in the fifth coordinate, should be regarded as "a quantum condition in classical disguise" (p. 59). He emphasized that the ratio of the Planck length (which appears in the "ordinary quantisation of gravitation theory", p. 61) and the periodicity length (which gave the "quantization" of the electric charge) was just the fine structure constant of QED and concluded that "[t]o have these two processes of quantization connected is the same as to determine the value of" the fine structure constant (Pauli's holy grail,

[94] Letter from Pauli to Källén, 24 April 1955, printed in (von Meyenn 2001).
[95] Pauli, as president of the conference, appears to actually have played a role in this, see letter from Klein to Pauli of 6 August 1954, published in (von Meyenn 1999).
[96] See letter from Klein to Pauli of 6 August 1954, (von Meyenn 1999).

incidentally). Klein went on to speculate that

> A near lying possibility of such a connection is that the relation between [the periodicity length] and [the Planck length] is determined by the renormalisation of the electric charge through vacuum polarisation, which in an adequate theory ought to be finite.

Klein's speculations prompted Pauli to the following remarks in discussion:

> [T]he connection [...] of the mathematical limitation of quantum electrodynamics with gravitation pointed out by LANDAU and KLEIN, seems to me to hint at the indeterminacy in space-time of the theory, invariant with respect to the wider group of general relativity. It is possible that this new situation so different from quantized theories, invariant with respect to the Lorentz group only, may help to overcome the divergence difficulties which are so intimately connected with a c-number equation for the light-cone in the latter theories (p. 69).

Pauli's vague remarks were discussed during the course of 1956 by Deser and Klein[97], now separated from their original Landau pole context and viewed as a general recipe for eliminating divergences (acting as a regulator) in quantum field theory. These discussions resulted in a paper by Deser (1957), first presented at the Chapel Hill Conference and then published in the Chapel Hill special issue of the Reviews of Modern Physics. In this paper, Deser (acknowledging his discussions with Klein) had managed to flesh out Pauli's ideas. Deser's idea was to use a path-integral approach, which in the case of a quantized gravitational field would also involve an integration over configurations of the metric field. This was supposed to ameliorate the divergences in the following manner: In the usual theory, the propagators (in position space) would exhibit their characteristic divergent behavior on the light cone by involving expressions of the form $1/(x_\mu - x'_\mu)^2$, a divergent behavior that lay at the heart of the divergences of quantum field theory in general. In a theory with a quantized gravitational field, Deser argued, such a singular expression should be replaced by an expression of the form

$$\int \mathcal{D}g e^{\frac{i}{\hbar}A_g} \frac{1}{\left(\int_x^x \sqrt{g_{\mu\nu}dx^\mu dx^\nu}\right)^2} \tag{6}$$

where $\mathcal{D}g$ is the functional integration over the gravitational field variables. The singularity of the usual propagator would thus be removed, since in integrating over the metric configurations there would only be a set of measure zero in which x and x' actually lie on each other's light cone.

Deser's ideas may have been a fleshing out of Pauli's brief remarks, but they, too, were still rather vague. At Chapel Hill, Deser consequently received a fair amount of criticism. He recalls that it was particularly the comments by Richard Feynman that convinced him of the "lack of control over the whole regularization scheme"[98] still present in his sketch. The conference proceedings also mention comments by Lichnerowicz and Fourès-Bruhat, remarking on the lacking definition of the measure (see Sect. 4.3), and an unattributed criticism regarding the order of integration: If one integrates over the spacetime coordinates first, before integrating over the gravitational ones, one encounters the usual divergences before gravity can step in to save the day. Deser had answered that this was "partly due to an unallowed interchange of limits", but the sketch he had provided was hardly sufficient to substantiate such an assertion.

[97] Email from Stanley Deser to one of the authors (AB), 12 May 2016.
[98] Email from Stanley Deser to one of the authors (AB), 24 March 2016.

The Copenhagen workshop then sees Deser and Laurent, several months after Chapel Hill, trying to gain control of the scheme through approximation methods. The approximations employed were rather radical, as they simply replaced the action A_g for the gravitational field, appearing in the phase of the path integrand, which includes both the usual Hilbert action and radiative corrections arising when integrating out the matter fields, by the most simple field action possible, namely the Klein-Gordon action, normally used for a scalar field (multiplied by the inverse of Newton's constant, to make up for the fact that the metric does not have the correct units to be a scalar field)[99]. But even with this radical approximation, they were unable to make progress on the idea, as the path integral was still not of Gaussian (the only calculable) form, due to the metric showing up in the denominator, which was of course central to the proposed elimination of the divergences. The work at Copenhagen can thus be considered the death throes of the gravity as regulator idea; Deser would not work on it again afterwards.

In DeWitt's report, Deser's work was packaged together with Misner and Wheeler's ideas on quantum foam, which they had developed during their time in Leiden one year earlier (see Wheeler and Ford 2000, p. 246-263). Not much work appears to have been done on quantum foam in Copenhagen, and DeWitt's packaging Deser's regulator work and quantum foam together seems somewhat curious at first glance. For sure, both of them relied on path integral heuristics. But, half a year earlier, after Deser's talk at Chapel Hill, Wheeler had rather pointed out the disanalogies, remarking that the effects of functional integration over metric configurations would be far more radical than Deser envisioned, since one would also have to integrate over configurations with different topologies (Wheeler 1957). DeWitt argued for lumping the two together because both of them assumed significant effects of high energy quantum gravity effects in low-energy theories, despite the extreme feebleness of the actual gravitational interaction. In particular, the inclusion of quantum gravitational effects precluded the use of a flat space Minkowski metric as a zeroth order approximation to the structure of space-time (hence, ultimately, the name of the section). For the case of quantum foam this was a vague, physically motivated conjecture, in Deser's case it clearly showed in his approximate expression for the matter propagator: If one here expands the denominator around the Minkowski metric, the leading term is again the usual propagator, light-cone singularities and all.

But concerning the expected effects of high-energy quantum gravity contributions on the low-energy theory, Deser and Wheeler were miles apart. Deser was merely arguing that the elimination of divergences he envisioned would not occur at any finite order in perturbations around a flat metric. The actual contributions to the low-energy propagator were expected to be minute. On the other hand, Wheeler was expecting to explain the actual existence of particles and charge in the low-energy theory as a result of high-energy quantum-gravitational fluctuations. At Chapel Hill, Deser indeed appears to have quite explicitly disagreed with Wheeler on this point. In a note that Wheeler took after Deser's presentation of the regulator work at Chapel Hill, he wrote:

> How important is high freq[uency] stuff? JAW [John Archibald Wheeler] says very; Deser says little[100].

[99] The fact that the metric is now effectively a scalar field in Deser and Laurent's approximation, is underscored by the fact that it always has lower indices, as a distinction between co- and contravariant is no longer possible.

[100] Relativity Notebook V, p. 99, JAWP.

So, it would appear that in this section, DeWitt is really just arguing against the view of the Wheeler school[101], especially where he states that the non-locality, which Wheeler expected to arise from topological fluctuations at the quantum level, should arise already at the classical level, and would thus directly appear in the quantum theory after the usual quantization and approximation procedures were applied to the canonical formulation of general relativity he was currently constructing (see the previous section). One can only speculate as to why DeWitt would dampen his criticism of Wheeler by formulating it primarily as a criticism of Deser. But certainly DeWitt was professionally very much dependent on Wheeler, who had been the "godfather" of DeWitt's Institute of Field Physics and even at this point had great influence on the funding DeWitt received there (See Footnote 88).

In any case, the question at stake here was really a methodological, rather than a physical one: Should one try to tease out the implications of quantum gravity without actually having such a theory at hand, or should one just press on and try to construct a full quantum gravitational formalism? DeWitt was here arguing for the latter. A very concise formulation of his view can be found in the above-mentioned grant application:

> [E]ither one must quantize the gravitational field or one must formulate a new and very basic principle in physics. As there is no indication at present of what this basic principle should be – or as one is not clever enough to bring it to light – it is worth attempting to quantize the gravitational field. Either one will succeed and thus understand the connection between two major theories of physics – the gravitational theory and the quantum theory – or one will fail and learn something from the failure.

4.6 Some special problems

The investigation of gravitational radiation had been from the beginning strongly influenced by a close analogy with electromagnetic theory. Einstein's original derivation of gravitational waves had relied on the weak field approximation, and thus involved none of the nonlinear aspects of the gravitational field, where most of its distinctive features were expected to appear. Moreover, gravitational waves – if they existed – were expected to be transverse waves traveling at the speed of light, just like their electromagnetic counterparts. Felix Pirani made, in 1962, a remark that summarizes well the resilience of analogies in the research on gravitational radiation:

> [O]ne cannot [...] expect all the familiar attributes of electromagnetic radiation to have analogues in the case of gravitational radiation. [Nevertheless, some] analogy has to be sought, because the concept of radiation is until now largely familiar through electromagnetic theory, and one cannot define gravitational radiation sensibly without some appeal to electromagnetic theory for guidance (Pirani 1962, p. 91, see also Rickles 2010).

But, bit by bit – as mentioned in section 2.1 – the particularities of general relativity began to move to the focus of attention and the analogy began to erode. DeWitt began section 6 of the report with a clear statement in that direction.

[101] There really does not appear to have been much of a difference of opinion on these matters between Deser and DeWitt. Deser, upon reading DeWitt's report for the first time in 2016, indeed remarked that the discussion hinted at here had made no impression at all and actually voiced his support for DeWitt's stance that "[w]e always have to expand" (Email to one of the authors (AB) of 22 March 2016). Conversely, only seven years after the Copenhagen meeting, DeWitt (1964) would actually bring forth an argument that Deser's speculations about gravity as a regulator were in fact correct.

The Copenhagen meeting was not the first time that DeWitt reflected on the differences between electromagnetic radiation and gravitational radiation. He first thought systematically about these matters when he was preparing a talk for the theoretical physics panel of the Third Annual Meeting of the American Astronautical Society, which took place in New York on 7 December 1956. The panel apparently gathered DeWitt, Peter Bergmann, Freeman Dyson, and Louis Witten, and was chaired by Agnew Bahnson[102]. That talk was DeWitt's first foray into gravitational radiation[103]. There, DeWitt stated:

> I would like to give you a glimpse into the subtlety of the necessary analysis [in general relativity] by considering how one would arrive at a unique definition of gravitational radiation. [...] Let's ask how we would, if we could, in principle, measure gravitational radiation. We have to work by analogy. How do we measure electromagnetic radiation? We measure electromagnetic radiation essentially by the jolt it gives us when it passes by us (DeWitt 1957a, p. 24).

The informal style was adapted to the purpose of the talk: it was aimed at publicizing the activities of the new Institute of Field Physics to a broad audience (DeWitt and Rickles 2011, p. 16). DeWitt explained the fundamental difference between electromagnetic and gravitational radiation in the following lines. In order to measure electromagnetic radiation, one must observe a varying force on a charged particle – i.e., a third time derivative of position. Since the force is the gradient of the potential, and supposing that one can transform time derivatives into space derivatives using the chain rule, the existence of an electromagnetic radiation is related to the *second* spatial derivative of the potential.

The measurement of a gravitational radiation is more subtle: "In the gravitational case, you don't feel any jolt because everything gets jolted by the same amount. [...] There is no net jolt. How are you going to measure it? Well, here is the conceptional solution" (DeWitt, 1957b, p. 25). DeWitt imagined the following thought experiment. Two objects are falling in free fall towards a planet. One only knows that there is a gravitational field because these objects approach each other, due to the convergence of the lines of force. Therefore, the existence of a gravitational field can be observed by a measurement of their mutual acceleration. Such an acceleration is the gradient of the gravitational force and, thus, the second spatial derivative of the potential. If one wants to observe a gravitational wave, one must have a variation in the mutual acceleration of those two objects. DeWitt concluded that the measurement of a gravitational wave depends on measuring the *third* derivative of the potential.

> This gives you an idea of how you have to go to a more subtle mathematical stage to reach an useful and invariant definition of gravitational radiation. This shows you an important difference between gravity and electromagnetism, and one which forces many analogies to break down (DeWitt 1957a, p. 25).

Probably DeWitt got the idea for the thought experiment from conversations with John Wheeler. The main experimental efforts in general relativity at that time were led, independently, by Joseph Weber and by Robert Dicke, both working at Princeton. Weber was particularly devoted to measuring gravitational waves, with no success. His research was in all respects guided by analogies with the electromagnetic radiation. At the 1961 Varenna summer school on experimental tests of general relativity, organized

[102] Bahnson Memoranda #5 (16 October 1956). DF, box Institute of Field Physics.

[103] In fact, he had discussed gravitational radiation in his PhD thesis, but in a quite naive way, without paying attention to the differences between the electromagnetic and the gravitational case (DeWitt 1949). Only in the 1956 talk did he reflect on the matter systematically.

by Christian Møller, Weber explained:

> Let us for a moment consider the [gravitational wave] problem from the standpoint of an experimentalist. He would really like to do laboratory experiments similar to those which Hertz did on electromagnetic waves almost a century ago, that is, he would like to generate and detect such waves in his laboratory (Weber 1962, p. 116).

The experiment carried out by Weber (1960) used a gravitational antenna (called "mass quadrupole detector") designed in close analogy to the electromagnetic case. While Weber's work initialized the entire field of gravitational wave detection, the analogy with the electromagnetic case was ultimately abandoned in favor of a detection method specific to gravitational waves (based on laser interferometry), culminating in the first detection of gravitational waves in 2015.

The relation between electromagnetic and gravitational radiation was mentioned several times at the Chapel Hill conference. Hermann Bondi expressed a critical perspective: "The analogy between electromagnetic and gravitational waves has often been made, but doesn't go very far, holding only to the very questionable extent to which the equations are similar" (DeWitt and Rickles 2011, p. 95). John Wheeler expressed a more optimistic view. According to him, the richness of Maxwell's theory should act as a source of enthusiasm for the relativists, providing questions and allowing them "to draw new richness out of the [general relativity] theory on the classical level" (DeWitt and Rickles 2011, p. 44). As an example, he mentioned: "[A] task ahead of us is the construction of the curve giving the spectrum of gravitational radiation incident on the earth, analogous to the known curve of the electromagnetic radiation spectrum; or at least the determination of upper limits on it" (DeWitt and Rickles 2011, p. 45). Therefore, Wheeler is not seeking formal analogies between both theories, but rather using electromagnetism merely a general source of inspiration to general relativity.

In section 6 of the Copenhagen report, DeWitt aimed at providing further illustration of the subtle differences between electromagnetic and gravitational radiation. Those examples show, in DeWitt's own words, "in a simple way how radically different the gravitational field is from other fields". Therefore, analogical methods should be used with due care, for gravitation was not just another field theory. This was one of the main conclusions of the Copenhagen meeting.

At the end of section 6, DeWitt refers to Dirac's 1938 article. It may be interesting to say here a little on DeWitt's work immediately after the Copenhagen meeting, since it was done under the influence of that classical article and was related to the radiation reaction issues. One of the questions raised at the Chapel Hill conference – as we mentioned in section 4.1 – was: "[C]an the principle of equivalence be extended to elementary particles?" (DeWitt and Rickles 2011, p. 167). DeWitt's second PhD student, Robert Brehme, asked him – probably in late 1957 – a similar question: Does a falling electric charge radiate? As DeWitt recalled:

> I was at that time [late 1957] trying to develop a canonical formalism for the gravitational field with the aim of creating a quantum theory of gravity, and I hoped that Brehme would assist me in this work. In fact, the work bogged down in the usual difficulties familiar to anyone who has tried to construct, and make sense of, a canonical quantum theory of gravity. So, in desperation, I agreed to let Brehme investigate the falling-charge problem; but I insisted that he do it properly. He was to begin by studying Dirac's famous 1938 paper on the classical radiation electron, in which all calculations are performed in a manifestly Lorentz invariant manner. He was then to translate this paper into the language of curved spacetime, keeping all the derivations manifestly

> generally covariant. He was not to introduce a special coordinate system at any stage (DeWitt 1996b, p. 34).

DeWitt and Brehme's research exposed for the first time several interesting aspects of radiation reaction in curved spacetimes, including a nice application of Synge's world function (DeWitt and Brehme 1960, see also Poisson et al. 2011, p. 19). In the course of their work, they studied Jacques Hadamard's book in great detail – which allowed DeWitt to master Green functions in curved spacetimes, a technique that turned out to be a fundamental ingredient in the covariant approach to quantum gravity that he would develop during the 1960s (DeWitt 1962a, 1965, 1967; Hadamard 1923).

5 Conclusions

As we hope to have shown through our contextualization and analysis, Bryce DeWitt's Copenhagen report not only provides a fascinating and invaluable overview of the state of quantum gravity in 1957, it also presents DeWitt's vision of establishing quantum gravity as a field of research within the emerging general relativity community of the renaissance. DeWitt's vision was an inclusive one: He had worked on perturbative approaches as a Ph.D. student, and in Copenhagen he sought to bring together canonical and Feynman quantization approaches, to open up a dialogue and to establish connections. He also sought to bring the older generation on board, which was more interested in epistemological questions of measurement and uncertainty relations. The question we might now ask is, given our observation in the introduction that to this date there is no unified research field of quantum gravity, whether the Copenhagen meeting was a success.

At first glance one may well answer in the affirmative: DeWitt made substantial progress on his parameterized, canonical approach; Møller, DeWitt, and Misner explored the role of clocks in establishing the consistency of a quantum theory of the gravitational field; Misner's and Laurent's different approaches to the measure in Feynman quantization began to converge; and Deser and Misner would soon begin to collaborate (together with Richard Arnowitt) on the Schwinger quantization of gravitation. But if we trace these developments only a little bit further, we realize that all of these hopeful beginnings soon turned into dead ends: DeWitt abandoned the canonical approach soon after, just as Misner abandoned his work on Feynman quantization. Laurent soldiered on, but without making much progress. DeWitt had promised, in the first version of the report, two papers by himself as an outcome of the Copenhagen workshop (one of them together with Laurent); they were never published. And while the ADM work provided important breakthroughs concerning the concept of energy in general relativity, or the most efficient formulation of the initial value problem (especially for numerical calculations), it was clear to the authors already in the early 1960s that it would not deliver the sought-after quantum theory of gravity. In the first version of the report, DeWitt had written:

> A formalism, which is something we do not even yet have in the case of quantum gravidynamics, must come first. Fortunately, it now seems as if we may have one in the not too distant future[104].

DeWitt had meant that the formalism "must come first" before the questions of measurability could be pursued in earnest. But his statement equally well applies to the question of community-building. The renaissance of general relativity had only been possible because there was a solid theoretical and formal basis on which the entire

[104] First version of the 1957 Copenhagen report. DF, Box "Institute of Field Physics", unnamed folder 7.

community could agree, the Einstein equations at the very least. There appeared to be in Copenhagen a convergence and an emerging understanding between the various approaches, but as each of them ran into their own individual difficulties, this budding consensus soon faded. DeWitt had envisaged a future in which a core formalism of quantum gravity would soon emerge (most probably in the best-defined approach, the canonical one). The different approaches would then be viewed as different perspectives on the same theory – as had happened in the history of quantum electrodynamics in the late 1940s. The various research agendas would then merge into a common field of quantum gravity research. So, e.g., even though he had dismissed Schwinger's action principle as a way of *constructing* the core formalism, he welcomed it as a future way of *analyzing* that formalism. As he wrote in the report:

> It seemed highly unlikely to the participants at Copenhagen that Schwinger's methods would really prove useful in the search for the hidden path to a quantum theory of gravitation. On the other hand, there is no doubt that when a quantum gravitational formalism is finally found Schwinger will be on hand to reformulate it.

But DeWitt's plans were frustrated. The uncontroversial core formalism (to be found at the end of the "hidden path") that had been DeWitt's objective in 1957 soon receded beyond the horizon and with it the memory of the Copenhagen workshop. Despite the fact that the origins of ADM can in many ways be traced to the Copenhagen meeting, Deser would later state that the Copenhagen meeting (as opposed to the conferences in Bern and Chapel Hill) had left "no traces" in his memory[105] and Misner also stated that he had "almost no physics recollections of the 1957 workshop"[106]. Rather than becoming a joint research field, quantum gravity gradually turned into a race, if it can be called that, with the various competitors running off in different directions, apparently indifferent to one another. As the different approaches moved apart, it became increasingly difficult to translate from one to the other.

DeWitt would make substantial progress in the decade after Copenhagen, culminating in his famous 1967 trilogy. He continued to contribute to the canonical approach (the Wheeler-DeWitt equation), but in the years after 1957 mainly focused on a new approach that essentially reversed the view he had held in Copenhagen. While he had then defended the position that discussions of measurement would come after the formalism, he now pursued the idea that the consistent measurability demands could actually define the formalism. But he pursued this path alone and without achieving ultimate success.

Essentially to this day, quantum gravity remains what DeWitt had fought so hard for it not to be: A high-energy phantom theory that could liberally be employed for heuristic investigations of measurability questions, the ultimate constitution of matter (Wheeler), or the consistency of low-energy quantum field theories (as in Deser's regulator attempts). We hope that the reader may draw some inspiration from the glimpse the Copenhagen report provides of the brief moment in time when there was a unified quantum gravity community.

Acknowledgements. We would like to thank Wolf Beiglböck, Stanley Deser, Domenico Giulini, Roberto Lalli, and Dean Rickles for valuable comments on the manuscript. We also thank Cécile DeWitt-Morette, the Niels Bohr Archive (Copenhagen), the Niels Bohr Library and Archives (College Park), and the American Philosophical Society (Philadelphia) for their authorization to consult their collections. T.H. thanks Olival Freire Jr., Reinaldo de

[105] Email to one of the authors (AB) of 22 March 2016.

[106] He *did* remember meeting his future wife in Copenhagen in the summer of 1957. Email to one of the authors (AB) of 23 March 2016.

Melo e Souza, and Finn Aaserud for conversations and advice; the funding agencies CNPq (Brazil) and Novo Nordisk Foundation (Denmark); and the Friends of the Center for History of Physics (American Institute of Physics) for a grant-in-aid which funded his research at the Niels Bohr Library and Archives. Open access funding provided by Max Planck Society.

References

Abrikosov, A.A. (1973). My years with Landau. *Physics Today* **26**: 56–60.

Anderson, J.L. (1958). Reduction of primary constraints in generally covariant field theories. *Physical Review* **111**: 965–966.

Arnowitt, R. and S. Deser. (1955). Renormalization of derivative coupling theories. *Phys. Rev.* 100(1):349–361.

Arnowitt, R. and S. Deser. (1959). Quantum theory of gravitation: General formulation and linearized theory. *Phys. Rev.* **113**: 745–750.

Arnowitt, R., S. Deser and C.W. Misner. (1962). The dynamics of general relativity. In Witten, L., editor, *Gravitation: An Introduction to Current Research*. Wiley, New York, Chap. 7, pp. 227–265.

Ashtekar, A. (2014). The last 50 years of general relativity and gravitation: from GR3 to GR20 Warsaw conferences. *General Relativity and Gravitation* **46**: 1706 (1–17).

Bergmann, P.G. (1942). *Introduction to the Theory of Relativity*. Prentice-Hall, New York.

Bergmann, P.G. (1956). Introduction of "true observables" into the quantum field equations. *Il Nuovo Cimento* **III**: 1177–1185.

Bergmann, P.G. and I. Goldberg. (1955). Dirac bracket transformations in phase space. *Phys. Rev.* **98**: 531–538.

Bergmann, P.G., R. Penfield, R. Schiller and H. Zatzkis. (1950). The Hamiltonian of the general theory of relativity with electromagnetic field. *Phys. Rev.* **80**: 81–88.

Bergmann, P.G. and R. Schiller. (1953). Classical and quantum field theories in the lagrangian formalism. *Phys. Rev.* **89**: 4–16.

Bhabha, H.J., editor (1951). *Report of an international conference on elementary particles*, Bombai. Tata Institute/UNESCO.

Blum, A., R. Lalli and J. Renn. (2015). The reinvention of general relativity: A historiographical framework for assessing one hundred years of curved space-time. *Isis* **106**: 598–620.

Blum, A.S. and D. Rickles, editors (2017). *Quantum Gravity in the First Half of the XXth Century: A Sourcebook*. Edition Open Access, Berlin.

Bohr, N. (1928). The Quantum Postulate and the Recent Development of Atomic Theory. *Nature* **121**: 580–590.

Bohr, N. and L. Rosenfeld. (1933). Zur Frage der Messbarkeit der elektromagnetischen Feldgrößen. *Det Kgl. Danske Videnskabernes Selskab Mathematisk-fysiske Meddelelser* **12** (8): 1–65.

Brevik, I.H. (2011). Christian Møller: The concepts of mass and energy in the general theory of relativity I-II (DKNVS Forhandlinger 1958). *Det Kongelige Norske Videnskabers Selskabs Skrifter*, pp. 93–102.

Bronstein, M. (1936). Quantentheorie schwacher Gravitationsfelder. *Physikalische Zeitschrift der Sowjetunion* **9**: 140–157.

Darrigol, O. (1991). Cohérence et complétude de la mécanique quantique: l'exemple de "Bohr-Rosenfeld". *Revue d'histoire des sciences* **44**: 137–179.

Deser, S. (1957). General relativity and the divergence problem in quantum field theory. *Rev. Mod. Phys.* **29**: 417–423.

Deser, S. (1977). Oskar Klein. *Phys. Today* **30**: 67–68.

Deser, S. (1987). Gravity from self-interaction in a curved background. *Classical and Quantum Gravity* **4**: L99–L105.

Deser, S. (1995). Oskar Klein: from his life and physics. Technical Report TH-95-09, CERN.

Deser, S. (2015). The legacy of ADM. *Physica Scripta* **90**: 068006 (4pp).

Deser, S. and A. Starobinsky. (2012). Editorial note to: Matvei P. Bronstein, Quantum theory of weak gravitational fields. *General Relativity and Gravitation* **44**: 263–265.

DeWitt, B. (1949). *I: The Theory of Gravitational Interactions. II: The Interaction of Gravitation with Light.* Ph.D. thesis, Harvard.

DeWitt, B. (1952). Point transformations in quantum mechanics. *Phys. Rev.* **85**: 653–661.

DeWitt, B. (1957a). Principal directions of current research activity in the theory of gravitation. *Journal of Astronautics* **4**: 23–28.

DeWitt, B. (1957b). Dynamical theory in curved spaces. I. A review of the classical and quantum action principles. *Rev. Mod. Phys.* **29**: 377–397.

DeWitt, B. (1961). Quantization of fields with infinite-dimensional invariance groups. *J. Math. Phys.* **2**: 151–162.

DeWitt, B. (1962a). The quantization of geometry. In Witten, L., editor, *Gravitation: An Introduction to Current Research*. Wiley, New York, Chap. 8, pp. 266–381.

DeWitt, B. (1962b). Quantization of fields with infinite-dimensional invariance groups. III. Generalized Schwinger-Feynman theory. *J. Math. Phys.* **3**: 1073–1093.

DeWitt, B. (1964). Gravity: A universal regulator? *Phys. Rev. Lett.* **13**: 114–118.

DeWitt, B. (1965). *Dynamical Theory of Groups and Fields*. Gordon and Breach, New York.

DeWitt, B. (1967). Quantum theory of gravity. II. The manifestly covariant theory. *Phys. Rev.* **162**: 1195–1239.

DeWitt, B. (1971). The Many-Universes Interpretation of Quantum Mechanics. In d'Espagnat, B., editor, *Fondamenti di meccanica quantistica: Rendiconti della scuola internazionale di fisica Enrico Fermi*. Academic Press, New York, pp. 211–262.

DeWitt, B. (1996a). Preliminary remarks before beginning his technical talk. In Ng, Y.J., editor, *Julian Schwinger: The Physicist, the Teacher, and the Man*. World Scientific, Singapore, pp. 29–31.

DeWitt, B. (1996b). The uses and implications of curved-spacetime propagators: A personal view. In Ng, Y.J., editor, *Julian Schwinger: The Physicist, the Teacher, and the Man*, pages 33–59. World Scientific, Singapore.

DeWitt, B. (2003). *The Global Approach to Quantum Field Theory*. Oxford University Press, New York.

DeWitt, B. and R.W. Brehme. (1960). Radiation damping in a gravitational field. *Ann. Phys.* **9**: 220–259.

DeWitt, B. and C.M. DeWitt. (1952). The quantum theory of interacting gravitational and spinor fields. *Phys. Rev.* **87**: 116–122.

DeWitt, C.M. (2011). *The Pursuit of Quantum Gravity: Memoirs of Bryce DeWitt from 1946 to 2004*. Springer, Berlin.

DeWitt, C.M. and Rickles, D., editors (2011). *The Role of Gravitation in Physics: Report from the 1957 Chapel Hill Conference*. Edition Open Access, Berlin.

Dirac, P.A.M. (1932). Relativistic quantum mechanics. *Proc. Roy. Soc. Lond.* **A136**: 453–464.

Dirac, P.A.M. (1950). Generalized Hamiltonian dynamics. *Canadian J. Math.* **2**: 129–148.

Dirac, P.A.M. (1958). The theory of gravitation in Hamiltonian form. *Proc. Roy. Soc. Lond.* **A246**: 333–343.

Einstein, A. (1953). *The Meaning of Relativity (Fourth Edition, including the Generalization of Gravitation Theory)*. Princeton University Press, Princeton.

Einstein, A., L. Infeld and B. Hoffmann. (1938). The gravitational equations and the problem of motion. *Ann. Math.* **39**: 65–100.

Einstein, A. and N. Rosen. (1935). The particle problem in the general theory of relativity. *Phys. Rev.* **48**: 73–77.

Eisenstaedt, J. (1986). La relativité générale à l'étiage: 1925-1955. *Archive for History of Exact Sciences* **35**: 115–185.

Fermi, E. (1932). Quantum theory of radiation. *Rev. Mod. Phys.* **4**: 87–132.

Ferraris, M., M. Francaviglia and C. Reina. (1982). Variational formulation of general relativity from 1915 to 1925 "Palatini's Method" discovered by Einstein in 1925. *General Relativity and Gravitation* **14**: 243–254.

Feyerabend, P. (1995). *Killing Time*. University of Chicago Press, Chicago.

Feynman, R.P. (1950). Mathematical formulation of the quantum theory of electromagnetic interaction. *Phys. Rev.* **80**: 440–457.

Feynman, R.P. (1963). Quantum theory of gravitation. *Acta Physica Polonica* **24**: 697–722.

Feynman, R.P., F.B. Moringo and W.G. Wagner. (1995). *Feynman Lectures on Gravitation.* Addison-Wesley, Reading, MA.

Fierz, M. and W. Pauli. (1939). On relativistic wave equations for particles of arbitrary spin in an electromagnetic field. *Proc. Roy. Soc. Lond.* **A173**: 211–232.

Fisher-Hjalmars, I. and B. Laurent. (1991). Oskar Klein. In Ekspong, G., ed. *The Oskar Klein Memorial Lectures.* World Scientific, Singapore, Vol. 1, pp. 1–9.

Fourès-Bruhat, Y. (1952). Théorème d'existence pour certains systèmes d'équations aux dérivées partielles non-linéaires. *Acta Mathematica* **88**: 141–225.

Fredenhagen, K. (2010). Lille 1957: The birth of the concept of local algebras of observables. *Eur. Phys. J. H* **35**: 239–241.

Glashow, S.L. (1959). The renormalizability of vector meson interactions. *Nucl. Phys.* **10**: 107–117.

Goldberg, J.N. (1992). US Air Force support of general relativity: 1956–1972. In Eisenstaedt, J. and Kox, A., editors, *Studies in the History of General Relativity*, volume 3 of *Einstein Studies*. Birkhäuser, Boston, pp. 89–102.

Goldberg, J.N. and E.L. Schucking. (2003). Peter Gabriel Bergmann. *Phys. Today* **56**: 64–66.

Gupta, S.N. (1952a). Quantization of Einstein's gravitational field: general treatment. *Proc. Phys. Soc.* **65**: 608–619.

Gupta, S.N. (1952b). Quantization of Einstein's gravitational field: linear approximation. *Proc. Phys. Soc. A* **65**: 161–169.

Gupta, S.N. (1954). Gravitation and electromagnetism. *Phys. Rev.* **96**: 1683–1685.

Haag, R. (2010). Some people and some problems met in half a century of commitment to mathematical physics. *Eur. Phys. J. H* **35**: 263–307.

Hadamard, J. (1923). *Lectures on Cauchy's problem in linear partial differential equations.* Yale University Press, New Haven.

Halpern, P. (2004). *The Great Beyond: Higher Dimensions, Parallel Universes, and the Extraordinary Search for a Theory of Everything.* Wiley.

Hartz, T. and O. Freire. (2015). Uses and appropriations of Niels Bohr's ideas about quantum field measurement, 1930–1965. In Aaserud, F. and Kragh, H., editors, *One hundred years of the Bohr atom: Proceedings from a conference*, volume 1 of *Scientia Danica. Series M, Mathematica et physica*, pages 397–418, Copenhagen. Det Kongelige Danske Videnskabernes Selskab.

Heisenberg, W. (1938). Die Grenzen der Anwendbarkeit der bisherigen Quantentheorie. *Zeitschrift für Physik* **110**: 251–266.

Heisenberg, W. and W. Pauli. (1929). Zur Quantendynamik der Wellenfelder. *Zeitschrift für Physik* **56**: 1–61.

Heitler, W. (1944). *The Quantum Theory of Radiation*, 2rd edn. Oxford University Press, Oxford.

Iliopoulos, J. (1996). Physics in the CERN theory division. In Krige, J., editor, *History of Cern - Volume 3.* North Holland, Amsterdam, Chap. 8, pp. 277–326.

Infeld, L., editor (1964). *Relativistic Theories of Gravitation: Proceedings of a conference held in Warsaw and Jablonna, July 1962*, London. Pergamon Press.

Infeld, L. and A. Schild. (1949). On the motion of test particles in general realtivity. *Rev. Mod. Phys.* **21**: 408–413.

Ioffe, B.L. (2012). The first dozen years of the history of ITEP theoretical physics laboratory. arXiv:1208.1386v1.

Jacobsen, A.S. (2011). Crisis, measurement problems, and controversy in early quantum electrodynamics: The failed appropriation of epistemology in the second quantum generation. In Kojevnikov, A., Carson, C., and Trischler, H., editors, *Quantum Mechanics and Weimar Culture: Revisiting the Forman Thesis.* Imperial College Press, London, pp. 375–396.

Jarlskog, C. (2014). *Portrait of Gunnar Källén*, chapter 52: At 1955 Moscow Meeting. Springer, Switzerland, pp. 223–225.

Kalckar, J. (1971). Measurability problems in the quantum theory of fields. In d'Espagnat, B., editor, *Proceedings of the International School of Physics "Enrico Fermi"*. Academic Press, New York, pp. 127–168.

Kalckar, J., editor (1996). *Niels Bohr Collected Works, Volume 7*. North-Holland, Amsterdam.

Klauder, J. and J.A. Wheeler. (1957). On the question of a neutrino analog to electric charge. *Rev. Mod. Phys.* **29**: 516–517.

Klein, O. (1955). Quantum theory and relativity. In Pauli, W., editor, *Niels Bohr and the Development of Physics*. Pergamon Press, London, pp. 96–117.

Klein, O. (1956). Generalisations of Einstein's theory of gravitation considered from the point of view of quantum field theory. In Mercier, A. and Kervaire, M., editors, *Jubilee of Relativity Theory – Proceedings*, pages 58–71, Basel. Birkhäuser.

Kosmann-Schwarzbach, Y. (2009). Tribute to André Lichnerowicz (1915–1998). *Notices of the AMS* **56**: 244–246.

Kraichnan, R.H. (1955). Special-relativistic derivation of generally covariant gravitation theory. *Phys. Rev.* **98**: 1118–1122.

Krige, J. (1987a). Case studies of some important decisions. In Hermann, A., Krige, J., Mersits, U., and Pestre, D., editors, *History of Cern - Volume 1*. North Holland, Amsterdam, Chap. 8, pp. 237–292.

Krige, J. (1987b). Survey of development. In Hermann, A., Krige, J., Mersits, U., and Pestre, D., editors, *History of Cern – Volume 1*. North Holland, Amsterdam, Chap. 7, pp. 209–236.

Krige, J. (2006). *American Hegemony and the Postwar Reconstruction of Science in Europe*. MIT Press, Cambridge, MA.

Landau, L. (1955). On the quantum theory of fields. In Pauli, W., editor, *Niels Bohr and the Development of Physics*. Pergamon Press, London.

Landau, L. and R. Peierls, (1931). Erweiterung des Unbestimmtheitsprinzips für die relativistische Quantentheorie. *Zeitschrift für Physik* **69**: 56–69.

Landau, L. and I. Pomeranchuk. (1965). On point interactions in quantum electrodynamics. In ter Haar, D., editor, *Collected Papers of L. D. Landau*. Clarendon Press, Oxford, pp. 654–658.

Lee, S., editor (2009). *Sir Rudolf Peierls: selected private and scientific correspondence*. World Scientific, New Jersey, Vol. 2.

Lichnerowicz, A. (1944). L'intégration des équations de la gravitations relativiste et le problème des n corps. *J. Math. Pures Appl.* **23**: 37–63.

Lichnerowicz, A. (1955). *Théories Relativistes de la Gravitation et de l'Électromagnétisme*. Masson, Paris.

Lichnerowicz, A. (1964). Propagateurs et quantification en relativité générale. In Infeld, L., editor, *Relativistic Theories of Gravitation: Proceedings of a conference held in Warsaw and Jablonna, July 1962*, pp. 177–188.

Lichnerowicz, A. (1992). Mathematics and general relativity: A recollection. In Eisenstaedt, J. and Kox, A., editors, *Studies in the History of General Relativity*, volume 3 of *Einstein Studies*. Birkhäuser, Boston, pp. 103–108.

Lichnerowicz, A. and Tonnelat, M., editors (1962). *Les Théories Relativistes de la Gravitation*, Paris. CNRS.

Mehra, J. and Milton, K.A. (2000). *Climbing the Mountain: The Scientific Biography of Julian Schwinger*. Oxford University Press, Oxford.

Mercier, A. (1992). General relativity at the turning point of its renewal. In Eisenstaedt, J. and Kox, A., editors, *Studies in the History of General Relativity*, volume 3 of *Einstein Studies*. Birkhäuser, Boston.

Mercier, A. and Kervaire, M., editors (1956). Fünfzig Jahre Relativitätstheorie. *Helvetica Physica Acta*, Supplementum IV.

Misner, C.W. (1955). Applications of topology to general relativity. *Phys. Rev.* **99**: 662.

Misner, C.W. (1957). Feynman quantization of general relativity. *Rev. Mod. Phys.* **29**: 497–509.

Misner, C.W. and Wheeler, J.A. (1957). Classical physics as geometry. *Ann. Phys.* **2**: 525–603.

Møller, C. (1952). *The Theory of Relativity*. Clarendon Press, Oxford.

Nambu, Y. (1950). The use of the proper time in quantum electrodynamics. I. *Progress of Theoretical Physics* **5**: 82–94.

NN, editor (1959). *Colloque International du Centre National de la Recherche Scientifique. Les problèmes mathématiques de la théorie quantique des champs.*, Paris.

Oppenheimer, J.R. (1930). Note on the theory of the interaction of field and matter. *Phys. Rev.* **35**: 461–477.

Pauli, W., editor (1955). *Niels Bohr and the Development of Physics*. Pergamon Press, London.

Pauli, W. (1965). Continuous groups in quantum mechanics. *Ergebnisse der exakten Naturwissenschaften* **37**: 85–104.

Penfield, R.H. (1951). Hamiltonians without parameterization. *Phys. Rev.* **84**: 737–743.

Pestre, D. (1987). The period of conflict, August-December 1951. In Hermann, A., Krige, J., Mersits, U., and Pestre, D., editors, *History of Cern - Volume 1*. North Holland, Amsterdam, Chap. 5, pp. 147–177.

Pirani, F. (1962). Survey of gravitational radiation theory. In n/a, editor, *Recent Developments in General Relativity*. Pergamon Press, New York, Chap. 6, pp. 89–105.

Pirani, F.A.E. and A. Schild. (1950). On the quantization of Einstein's gravitational field equations. *Phys. Rev.* **79**: 986–991.

Poisson, E., A. Pound and I. Vega. (2011). The motion of point particles in curved spacetime. *Living Reviews in Relativity*, 7.

Rickles, D. (2010). Quantum gravity meets &HPS. In Mauskopf, S. and Schmaltz, T., editors, *Integrating History and Philosophy of Science: Problems and Prospects*. Springer, Dordrecht, pp. 163–199

Rickles, D. and A.S. Blum. (2015). Paul Weiss and the genesis of canonical quantization. *Eur. Phys. J. H* **40**: 469–487.

Rosenfeld, L. (1930a). Zur Quantelung der Wellenfelder. *Ann. Phys.* **397**: 113–152.

Rosenfeld, L. (1930b). Über die Gravitationswirkungen des Lichtes. *Zeitschrift für Physik* **65**: 589–599.

Rovelli, C. (2004). *Quantum Gravity*. Cambridge University Press, Cambridge.

Rozental, S. (1985). *NB: Erindringer om Niels Bohr*. Gyldendal, Copenhagen.

Salecker, H. and Wigner, E. (1958). Quantum limitations of the measurement of space-time distances. *Phys. Rev.* **109**: 571–577.

Salisbury, D. and Sundermeyer, K. (2016). Léon Rosenfeld's invention of constrained Hamiltonian dynamics. arXiv:1606.06076.

Schucking, E.L. (1989). The first Texas symposium on relativistic astrophysics. *Phys. Today* **42**: 46–52.

Schweber, S.S. (1994). *QED and the Men Who Made It: Dyson, Feynman, Schwinger, and Tomonaga*. Princeton Univ Press, Princeton.

Schwinger, J. (1948). Quantum electrodynamics. I. A covariant formulation. *Phys. Rev.* **74**: 1439–1461.

Solomon, J. (1938). Gravitation et quanta. *Journal de Physique et Le Radium* **9**: 479–485.

Straumann, N. (2000). On Pauli's invention of non-Abelian Kaluza-Klein theory in 1953. arXiv:gr-qc/0012054v1.

ter Haar, D., editor (1965). *Collected Papers of L. D. Landau*. Pergamon Press, London.

Tomonaga, S. (1946). On a relativistically invariant formulation of the quantum theory of wave fields. *Progress of Theoretical Physics* **1**: 27–42.

von Borzeszkowski, H.-H. and H.-J. Treder. (1988). *The Meaning of Quantum Gravity*. Springer, Netherlandds.

von Meyenn, K., editor (1993). *Wolfgang Pauli: Wissenschaftlicher Briefwechsel mit Bohr, Einstein, Heisenberg u.a.*, volume III: 1940–1949. Springer, Berlin.

von Meyenn, K., editor (1996). *Wolfgang Pauli: Wissenschaftlicher Briefwechsel mit Bohr, Einstein, Heisenberg u.a.*, volume IV/Part I: 1950–1952. Springer, Berlin.

von Meyenn, K., editor (1999). *Wolfgang Pauli: Wissenschaftlicher Briefwechsel mit Bohr, Einstein, Heisenberg u.a.*, volume IV/Part II: 1953–1954. Springer, Berlin.

von Meyenn, K., editor (2001). *Wolfgang Pauli: Wissenschaftlicher Briefwechsel mit Bohr, Einstein, Heisenberg u.a.*, volume IV/Part III: 1955–1956. Springer, Berlin.

von Meyenn, K., editor (2005). *Wolfgang Pauli: Wissenschaftlicher Briefwechsel mit Bohr, Einstein, Heisenberg u.a.*, volume IV/Part IV: 1957-1958. Springer, Berlin.

Weber, J. (1960). Detection and generation of gravitational waves. *Phys. Rev.* **117**: 306–313.

Weber, J. (1962). Theory of methods for measurement and production of gravitational waves. In Møller, C., editor, *Evidence for Gravitational Theories (Proceedings of the International School of Physics "Enrico Fermi")*. Academic Press, New York, Chap. 4, pp. 116–140.

Weiss, P. (1938). On the Hamilton-Jacobi theory and quantization of a dynamical continuum. *Proc. Roy. Soc. Lond.* **A169**: 102–119.

Weyl, H. (1950). A remark on the coupling of gravitation and electron. *Phys. Rev.* **77**: 699–701.

Wheeler, J.A. (1955). Geons. *Phys. Rev.* **97**: 511–536.

Wheeler, J.A. (1957). On the nature of quantum geometrodynamics. *Ann. Phys.* **2**: 604–614.

Wheeler, J.A. and R.P. Feynman. (1945). Interaction with the absorber as the mechanism of radiation. *Rev. Mod. Phys.* **17**: 157–181.

Wheeler, J.A. and K. Ford. (2000). *Geons, Black Holes & Quantum Foam*. W. W. Norton, New York.

Wigner, E. (2002). Concept of observation in quantum mechanics. In Podolsky, B., Hart, J.B., and Werner, F.G., editors, *Conference Manuscript: Conference on the Foundations of Quantum Mechanics (1962). Book 1, Monday Evening Session*. Xavier University Exhibit.

Wüthrich, A. (2010). *The Genesis of Feynman Diagrams*. Springer, Dordrecht.

Zeh, H. (2011). Feynman's interpretation of quantum theory. *Eur. Phys. J. H* **36**: 63–74.

Eur. Phys. J. H **42**, 159–176 (2017)
DOI: 10.1140/epjh/e2017-80016-0

THE EUROPEAN PHYSICAL JOURNAL H

Historical document

Exploratory research session on the quantization of the gravitational field*

At the Institute for Theoretical Physics, Copenhagen, Denmark, June-July 1957

Bryce S. DeWitt[a]

Aeronautical Research Laboratory, Contract No. AF 33(616)-5367, Wright Air Development Center, Air Research and Development Command, United States Air Force, Wright-Patterson Air Force Base, Dayton, Ohio, USA

Received 9 March 2017 / Accepted 15 March 2017
Published online 1 June 2017

Abstract. During the period June–July 1957 six physicists met at the Institute for Theoretical Physics of the University of Copenhagen in Denmark to work together on problems connected with the quantization of the gravitational field. A large part of the discussion was devoted to exposition of the individual work of the various participants, but a number of new results were also obtained. The topics investigated by these physicists are outlined in this report and may be grouped under the following main headings: The theory of measurement. Topographical[1] problems in general relativity. Feynman quantization. Canonical quantization. Approximation methods. Special problems.

Foreword

An exploratory research session on problems connected with the quantization of the gravitational field was held from the 15th of June to the 15th of July 1957 at the

* Reproduced here from a copy of the original document available at the Aage Petersen Collection of Reprints and Manuscripts, Series I, Box 1, Folder 8, deposited at the Niels Bohr Library and Archives, American Institute of Physics, College Park, MD, United States. As far as the the editors know, that is the single remaining copy of the final, typed version of the document. The current version was edited and commented by Alexander Blum (Max-Planck-Institut für Wissenschaftsgeschichte, Berlin, e-mail: ablum@mpiwg-berlin.mpg.de) and Thiago Hartz (Museu de Astronomia e Ciências Afins, Rio de Janeiro, e-mail: thiagohartz@gmail.com). Their comments on the manuscript have been included in additional footnotes, marked as "Editors' comment". Thanks is due to Melanie J. Mueller, director of the Niels Bohr Library and Archives, for authorizing the reproduction of this document.

[a] deceased

[1] Editors' comment: This is a typo. DeWitt means "topological".

University Institute for Theoretical Physics, Copenhagen, Denmark. The following persons participated in this research session:

Professor Christian Møller, Director of CERN Theoretical Study Division, Institute for Theoretical Physics, Copenhagen, Denmark.[2]

Professor Oskar Klein, University of Stockholm, Stockholm, Sweden.[2]

Professor Bryce S. DeWitt, University of North Carolina, Chapel Hill, North Carolina.

Dr. Stanley Deser, Institute for Theoretical Physics, Copenhagen, Denmark.

Dr. Charles W. Misner, Princeton University, Princeton, New Jersey.

Dr. Bertel Laurent, University of Stockholm, Stockholm, Sweden.[2]

The session – or, rather, the sessions – were held in a very informal style on numerous mornings, afternoons, and evenings between the dates indicated. At the beginning there was some uncertainty as to how best to proceed. It was suggested, for example, that the group might attempt to tackle one specific problem with an eye toward eventual joint publication. It was soon realized, however, that this would be impractical, in view of the varied interests of the members of the group, the nature of the present outstanding problems in gravitational theory, and the fairly limited amount of time available. It was then proposed instead that each participant should endeavour to expound his own point of view and accomplishments to the others, and to proceed from there to further development, if possible. This, in the main, was the procedure adopted. It must be understood that the exposition itself was a lengthy process, involving many arguments at various temperature levels and frequent wandering into side issues. On the other hand, "plenary" sessions formed only a part of the activity. Often only two or three participants were involved in a single discussion, and much of the hard work was carried out in solitude.

The most noteworthy aspect of this mode of operation was a strong sense of luxury shared by the participants – luxury at having other experts close at hand, the luxury of not having to dig each other's ideas painfully out of the published literature, the luxury of being able to repeat a question several times (which is not possible in an ordinary conference) and of not having to understand the first time, and the luxury of having ideas repeated with different emphasis and in various settings and shadings.

There was not time during the session to prepare material for publication. However, it seems likely that results obtained in Copenhagen will be incorporated in two or three future papers. It is impossible at present to say when these articles will appear, as the results form only a part of several much larger research efforts currently underway.

The present report will serve to summarize the main findings at Copenhagen and some of the tentative conclusions reached. The report is subdivided according to topics covered. An attempt is made in each case to indicate briefly the status of the subject prior to the Copenhagen effort and then to discuss the subsequent progress made (if any) identifying the main contributors.

Thanks are due to the CERN Theoretical Study Division for its hospitality to the participants; to the students of Hagemann's Kollegium where three of the participants were housed; and to Professor Niels Bohr and the staff of the Institute for Theoretical Physics for providing the participants with a place to work. Special thanks are due to Wright Air Development Center for having pioneered this type of meeting and to the University of North Carolina for having undertaken the administration of it.

Bryce S. DeWitt

Chapel Hill, North Carolina,
October 8, 1957

[2] Participating during the last eleven days only.

Results

1 The theory of measurement

In every new development in physical theory which extends or generalizes previous concepts, there arises at a certain stage a need to examine the overall self-consistency of the innovation. In the development of the quantum theory this self-consistency problem has centered around the "theory of the measurement process", since it is a special view of what is meant by a "measurement" which forms the core of the quantum innovation in physics. Moreover, each new development in the quantum theory itself has eventually necessitated a reexamination of measurement theory. Normally this reexamination comes at the end of a process of formalistic development (e.g. in electrodynamic theory). In the application of quantum mechanics to the general theory of relativity, however, the question of "measurability" has been raised even before the advent of a formalism – which, as of this date, still does not exist. Such activity as has so far taken place on this problem seems to have been mainly directed toward proving the *inapplicability* of quantum mechanics in the gravitational domain. If such inapplicability were really true, theoretical physicists would, of course, be saved a great deal of labor. In view of the formidable nature of the task of uniting the quantum and general relativity theories, one cannot, however, avoid the suspicion that such efforts may stem partly from wishful thinking.

The question of measurability was raised in Copenhagen by Professor Møller, who had some lengthy discussions on it just a few weeks previously with Professor E. P. Wigner of Princeton. In approaching the problem of measurement, Wigner has begun by trying to construct the most efficient measuring devices which the principles of quantum mechanics allow. In general relativity theory the pertinent measurement consists of a determination of the invariant interval between space-time events, which Wigner showed could always be performed by means of clocks. Wigner's problem therefore reduced initially to the problem of constructing the most efficient possible clocks. Having done this Wigner then wishes to proceed to the study of the influence of such clocks themselves on the objects being measured. Preliminary reports by his student, Dr. H. Salecker, presented at the Chapel Hill Conference in January 1957[3], indicated that very serious limitations are imposed on the measurability of the gravitational effects of masses as small as the heavier elementary particles – a result of considerable significance for the validity of any attempt to quantize general relativity.

While Professor Møller was in general agreement with Wigner's analysis of clocks (aside from some purely technical modifications which he introduced in the work sessions and which it is unnecessary to reproduce in this report) he was utterly unable to follow Wigner's subsequent implications in regard to the limitations on the measurement of gravitational effects of small masses. These implications hinge on the determination of the perturb-effect of the clocks and their registering devices, but the picture which he has so far presented is by no means clear.

Professor Møller quoted a statement made by Professor Bohr after finishing his famous "measurement" paper with L. Rosenfeld, viz.

> During the course of our study of the quantum limitations on the measurability of the electromagnetic field we made every possible mistake. In each case, in order to extricate ourselves, we had to go back and look at the formalism!

In Professor Møller's view it will be equally necessary, in the case of quantum gravidynamics, to have a valid formalism available, before one can make any reliable statements about the measurement problem. And even when such a formalism is finally

[3] See "Conference on the Role of Gravitation in Physics" WADC Technical Report 57–216.

found, Møller does not believe Wigner's implications will be substantiated. On the contrary, Møller believes that the gravitational effect of small masses can always be measured in principle if one simply has an arbitrarily long time at one's disposal.

DeWitt and Misner adduced the following arguments in support of this stand:

Consider first a classical test particle of mass m which is initially at rest at the point x_0 in a gravitational potential φ. At the time t the position of the particle will be

$$x = x_0 - \frac{1}{2}\frac{\partial\varphi}{\partial x}t^2 \tag{1}$$

The gravitational field strength is given by $\partial\varphi/\partial x$ and can be immediately determined by a measurement of t and x. If, however, the particle is subject to quantum laws its initial position and velocity[4] are subject to uncertainties related by

$$\Delta v_0 = \frac{\hbar}{m\Delta x_0} \tag{2}$$

leading to a spread in the position uncertainty of amount

$$\Delta x = \Delta v_0 t = \frac{\hbar t}{m\Delta x_0} \tag{3}$$

The initial position measurement may be made by a photon of momentum uncertainty $\Delta p = m\Delta v_0$. The resulting uncertainty in the initial time may be ignored since it is given by $\Delta t_0 = \Delta x_0/c = \hbar/mc\Delta v_0 \ll t$ (assuming $\Delta v_0 \ll c$). The final position and time measurements may be made with arbitrarily high precision, using energetic photons, since the experiment is then over.

Suppose the gravitational field[5] φ is produced by another particle of equal mass m. Then $\varphi = -Gm/x$, $\partial\varphi/\partial x = Gm/x^2$, $\partial^2\varphi/\partial x^2 = -Gm/x^3$, and, apart from a factor 2 involved in transforming to center-of-mass coordinates and using the "reduced mass", equation (1) will hold with x interpreted as the separation distance.

Certain obvious conditions must now be satisfied in order that the gravitational field be determined from the classical equation (1), namely

$$\max\left(\Delta x_0, \Delta x\right) \ll |x - x_0| \ll x \tag{4}$$

or

$$\max\left(\Delta x_0, \frac{\hbar t}{m\Delta x_0}\right) \ll \left|\frac{\partial\varphi}{\partial x}\right| t^2 = \frac{Gmt^2}{x^2} \ll x \tag{5}$$

Another requirement, viz.

$$\left|\frac{\partial^2\varphi}{\partial x^2}\right| |x - x_0| \ll \left|\frac{\partial\varphi}{\partial x}\right| \tag{6}$$

or

$$x = \frac{\left|\frac{\partial\varphi}{\partial x}\right|}{\left|\frac{\partial^2\varphi}{\partial x^2}\right|} \gg |x - x_0| \tag{7}$$

which says that $|\partial\phi/\partial x|$ must not change appreciably during the course of the motion, is then automatically satisfied.

[4] Editors' comment: v_0

[5] Editors' comment: DeWitt means "gravitational potential".

Introducing dimensionless quantities $\Delta\xi_0$, ξ, T, by the relations

$$\Delta x_0 = \frac{\hbar^2}{Gm^3}\Delta\xi_0, \quad x = \frac{\hbar^2}{Gm^3}\xi, \quad t = \frac{\hbar^3}{G^2m^5}T \tag{8}$$

one may rewrite these conditions in the form

$$max\left(\Delta\xi_0, \frac{T}{\Delta\xi_0}\right) = \alpha\frac{T^2}{\xi^2} = \alpha\beta\xi \tag{9}$$

where α, $\beta \ll 1$. We may distinguish two cases.

$$I: \ \alpha\beta\xi = \Delta\xi_0 > \frac{T}{\Delta\xi_0} = \frac{1}{\alpha}\sqrt{\frac{\xi}{\beta}} \tag{10}$$

$$II: \ \alpha\beta\xi = \frac{T}{\Delta\xi_0} > \Delta\xi_0 = \frac{1}{\alpha}\sqrt{\frac{\xi}{\beta}} \tag{11}$$

Both cases lead to $\xi > \frac{1}{\alpha^4\beta^3}$. Writing

$$\xi = \frac{1}{\alpha^4\beta^3\gamma^2}, \quad T = \sqrt{\beta\xi^3} = \frac{1}{\alpha^6\beta^4\gamma^3} \tag{12}$$

where $\gamma < 1$, we have $\Delta\xi_0/\xi = \alpha\beta$ in case I and $\Delta\xi_0/\xi = \alpha\beta\gamma$ in case II. Equations (12) show that the measurement can be performed provided simply that x and t be chosen large enough. Since, however, t varies inversely as the fifth power of m, the required values quickly become fantastic for small m. The lowest values are obtained with the choice $\gamma = 1$ for which cases I and II coalesce.

There is one more condition which must be satisfied: the initial photon must have an energy $\ll mc^2$ in order that its own gravitational influence be negligible. This simply means that Δx_0 must be much larger than the Compton wavelength corresponding to m. Thus

$$\Delta x_0 \gg \frac{\hbar}{mc} \quad \text{or} \quad \Delta\xi_0 = \frac{1}{\delta}\frac{m^2}{\mu^2}, \quad \delta \ll 1 \tag{13}$$

where

$$\mu = \sqrt{\frac{\hbar c}{G}} \approx 10^{-5} \text{ g}. \tag{14}$$

Choosing $\gamma = 1$, so that

$$\frac{1}{\delta}\frac{m^2}{\mu^2} = \Delta\xi_0 = \frac{1}{\alpha^3\beta^2}, \quad T = \frac{1}{\alpha^6\beta^4} = \frac{1}{\delta^2}\frac{m^4}{\mu^4}, \tag{15}$$

we may distinguish two further cases:

$$A: \ m < \mu \tag{16}$$

Here we must take $\delta = \frac{m^2}{\mu^2}\alpha^3\beta^2 \lll 1$.

$$B: \ m > \mu \tag{17}$$

Here we may choose[6] freely (subjected to the restriction $\delta \ll 1$) and then take $\alpha^3\beta^2 = \frac{\mu^2}{m^2}$.

The absolute minimum values of x and t for measurability are given by $\alpha = \beta = 1$ in case A and $\alpha = \beta$, $\delta = 1$ in case B. Thus

$$x \gg \frac{\hbar^2}{Gm^3}, \quad t \gg \frac{\hbar^3}{G^2m^5} \quad \text{when } m < \mu \tag{18}$$

$$\left.\begin{aligned} x &\gg \tfrac{\hbar^2}{Gm^3}\left(\tfrac{m^2}{\mu^2}\right)^{\frac{7}{5}} = \left(\tfrac{\hbar^3 G^2}{mc^7}\right)^{\frac{1}{5}} \\ t &\gg \tfrac{\hbar^3}{G^2m^5}\tfrac{m^4}{\mu^4} = \tfrac{\hbar}{mc^2} \end{aligned}\right\} \text{ when } m > \mu \tag{19}$$

There is never any difficulty in satisfying the latter set of conditions. In fact, when $m \geq \mu$ the measurements can be performed entirely classically, without regard to quantum effects. To see how fantastic the figures can be, on the other hand, when $m < \mu$, consider the case of protonic mass: $m \approx 10^{-24}$ g. Then

$$\frac{\hbar^2}{Gm^3} \approx 10^{25} \text{ cm} \approx 10 \text{ million light years} \tag{20}$$

$$\frac{\hbar^3}{G^2m^5} \approx 10^{53} \text{ sec} \approx 10^{46} \text{ years} \tag{21}$$

Nevertheless there is no point of *principle* which makes such measurements impossible. There is no difficulty in constructing a clock accurate enough to measure such large time intervals, and massive enough so that its own position uncertainty does not mask the position measurement of the masses m during this time. Finally the perturbing effect of the clock's gravitational field may be eliminated by incorporating the clock in a huge spherical shell of[7] mass density completely surrounding the experiment.

2 Topological problems

Although no questions of principle seem to disallow the union of the quantum and general relativity theories (on the contrary, there are strong reasons for attempting such a union) there are nevertheless many difficulties to be faced before it can be successfully brought about. In all previous cases, the classical behavior of any physical system possessing a classical analog has been thoroughly understood before the quantum theory was applied to it. In the case of the gravitational field, however, this classical behavior itself is not well understood. The classical gravitational field is exceedingly complex and rich in possibilities as yet unexplored.

Among the more curious of these possibilities is that which arrises because the gravitational field determines the space in which it itself acts – and in particular, the topology of that space. The manifold of possible topologies is infinite, just as the manifold of possible fields within each topology is infinite. It is clearly of interest to examine at least some of the typical topological situations which can arise.

Dr. Misner, working with Professor Wheeler's group at Princeton, had already during the previous two or three years examined some of the simplest topological possibilities. At Copenhagen he reported in some detail on the results of this work, leaving a sharpened awareness of its importance and significance in the minds of the other participants. As the simplest special topology Misner called attention

[6] Editors' comment: δ

[7] Editors' comment: uniform

to the Schwarzschild metric. With the use of conformal coordinates it is easy to see that the simplest interpretation of the Schwarzschild metric is that of a two sheeted topology (cf. Einstein and Rosen) with no singularities. In emphasizing another aspect of this metric, Misner pointed out that the Schwarzschild metric is that which describes the gravitational field of a "mass", which is of course familiar to everyone. The special point which Misner made, however, was that from a certain point of view *there is no mass*; the appearance of a "source" of the gravitational field is obtained by a topological trick. By combining the Maxwell field with the Einstein field Misner had (as reported at the Chapel Hill conference) also been able to represent sources of the electromagnetic field (i.e. "charges") by means of topological tricks.

Dr. Deser raised the question as to whether the sources of other fields (e.g. nuclear fields) can also be represented by topological tricks. He and Misner investigated the scalar meson fields and, in the massless case, found that the topological representation is again possible. In the case of non-vanishing mass the Einstein equations are difficult to solve, but it again appears in principle possible to use topological tricks to represent sources. In the study of the latter problem Deser and Misner had the help of Dr. A. Komar who had been a member of the Institute for Theoretical Physics during the past year, was still in Copenhagen, and was studying precisely the problem of the interaction of scalar and gravitational fields.

(In view of recent work of Klauder and Wheeler it appears that similar possibilities do *not* occur for spinor fields (Klauder and Wheeler 1957). That is, topological tricks *cannot* be employed to represent sources of spinor fields.)

Misner also outlined the situation in regard to more complicated topologies, in which several "sources" are present. Conformal coordinates can again be used at the initial instant, for both the neutral mass and charged mass cases, and Misner has shown that fairly simple and intuitively reasonable metrics which satisfy the Fourès equations for the Cauchy initial value conditions can be introduced for these topologies. The interest of these topologies is that they are dynamically active, i.e. they *move*. With several sources one can have a many-sheeted topology, a two-sheeted topology, or a "wormhole" topology. (See report of the Chapel Hill Conference.) The initial value conditions are easiest to "mock up" for the wormhole topology, simple monopole terms being sufficient, whereas for the two-sheeted topology multipole terms of all orders are required. However, the theory of both types of topologies is now sufficiently well advanced so that one could probably put a two-source problem on a high-speed computing machine without excessive difficulty. This may be done in the next year or so and would be of very great interest, especially the determination of what happens in very close encounters between sources.

In view of the apparent tractability of the multi-source problem, DeWitt and Laurent were unable to see why Misner had seemed to have proved two years ago the impossibility of a particularly simple type of "wormhole" dynamics: the case of a simple toroidal structure dominated by a closed electric field. Misner therefore took the problem up again and found an error in his previous work. He found, contrary to his original finding, that the toroidal dynamics is indeed possible, and, moreover, he was able to solve the differential equation for the toroidal expansion in closed form.

Another aspect of the topological problem arises from the following considerations: Sources of fields may be represented by topological tricks. On the other hand, topologies may be completely described by the metric. To what extent, therefore, can fields quite generally be described purely by metric? G. Y. Rainich showed many years ago that the electromagnetic field could be completely described by metric quantities (Rainich 1925). Its stress tensor leaves a sufficiently characteristic imprint on the metric for the complete details of the field itself to be deduced. Misner has worked out the details of the necessary deciphering process more completely than can be found in Rainich's work. At Copenhagen he and Deser investigated the feasibility of a deciphering procedure for a scalar field as well, and readily found one. They were able to show, moreover, that one can also handle several fields at a time, i.e. that various field can be disentangled from one another and identified in detail merely through the combined stress tensor. One cannot, however, disentangle two fields of the same variety (i.e. identical), and the problem becomes difficult when there are interactions between the various fields. It is possible, owing to the work of Takabayasha (various articles in Prog. Theor. Phys. and Il Nuovo Cimento in recent years) who has arrived at a complete description of spinor fields by means of hydrodynamic quantities, that the work of Deser and Misner can also be extended to spinor fields.

3 Feynman quantization

Having a keen appreciation of the great amount of work which remains to be done in the purely classical aspects of gravitational theory, it was with considerable soberness that the participants approached the main problem of the work sessions – the quantization problem. It was recognized that the successful solution of this problem will have to overcome many obstacles. The most direct frontal attack on the problem is that suggested by Professor Wheeler, who proposes simply to write

$$\int \exp\left[\frac{i}{\hbar}(\text{Einstein action})\right] \, d\,(\text{field histories}) \tag{22}$$

in order to obtain the invariant transformation function for the gravitational field. Dr. Misner, working with Wheeler, has perhaps done more than anyone else in pushing the investigation along these lines, which for the cases of simpler field theories, were laid down originally by R. P. Feynman.

The Feynman approach, because of its directness, runs quickly into a number of difficulties, among the chief of which are: (1) the necessity to discover or define an appropriate invariant "measure" with respect to which to carry out the Feynman integration; (2) the necessity of defining invariant state functionals to serve as bounds in the Feynman integration; and (3) the necessity of finding appropriate operators in terms of which to define physically significant observables and boundary conditions. The latter two difficulties arise in any approach to quantization, but the first difficulty is of special immediacy and importance in the Feynman method.

By introducing the notion of "homogeneity group" Misner has succeeded in defining a possible invariant measure for the Feynman integration. In Misner's argument no use is made whatsoever of the dynamical aspect of the gravitational field as exemplified by its Lagrangian; the whole rests solely on arguments of general invariance. This is by no means an unreasonable point of view and certainly does not mean, for example, that the measure cannot also be determined in fact from the Lagrangian. For the precise form which the Lagrangian has depends on properties of general invariance. Nevertheless, since appeal to the Lagrangian is essential according to the *canonical* viewpoint (see next section) it is of importance to see whether or not the two viewpoints are in harmony.

Dr. Laurent, at Copenhagen, developed an argument based on local linearization of the Einstein equations and the use of "hyper delta functions" in functional space to show that the Lagrangian indeed leads to Misner's "measure". There are still some uncertainties in Laurent's argument, chiefly concerning the admissibility of certain types of state functionals, and in the view of DeWitt, who has been developing a rigorous canonical approach (see next section), the problem is still not settled. Misner's measure has the virtue that it becomes singular when the metric tries to change its signature, thereby imposing a natural barrier to the extension of the Feynman integration over nonphysical field histories. Misner's measure, however, allows physical conditions to be specified on surfaces which are not necessarily space-like. On the other hand, DeWitt has obtained a preliminary measure which not only becomes singular when the metric tries to change signature but also becomes singular when the metric attempts to change the character of a surface (on which physical specifications have been made) from space-like to time-like.

Considerable discussion developed on the question of the appropriate field-history domain for a Feynman integration. This problem has several aspects. First, it was asked whether the existence of a natural barrier in Misner's or DeWitt's measure necessarily excludes integration over nonphysical field histories. It seems that the propagation amplitude will damp out anyway very rapidly in such domains, owing to the classical action becoming imaginary instead of real, and integration over such regions may in fact be necessary in order to obtain correct analytic behavior of asymptotic wave functionals[8] just as quantization of simple systems necessitates penetration of the wave function into classically nonphysical regions (barrier penetration).

Secondly, the question of topologies was raised. If one be permitted to integrate over nonphysical domains must one not also consider the possibility of integrating over various topologies? Professor Wheeler has championed the answer "yes" to this question, although in view of the infinity of possible topologies the prospect is slightly terrifying. The question itself has special aspects. For example, suppose it has been decided that integration over nonphysical signatures is to be excluded. There still remains a question as to the existence of 4-dimensional domains of constant signature which are bounded by two or more positive definite 3-spaces having different topologies, or whether peculiar topologies can exist as intermediate states to simple topologies. Misner plans to carry out extensive researches on these questions with the aid of the Princeton topologists. The problem, however, is exceedingly difficult, most of its aspects being to date unsolved or unknown. This becomes particularly obvious when one recalls that the bounding 3-spaces may be topologically different merely in virtue of "knottedness", and even the "knot problem" for closed two dimensional domains has not been solved.

It was conjectured in Copenhagen that the topological integration problem may bear some slight analogy to a boundary condition problem. Consider the quantum theory of a simple particle in a box. In order to treat this system by the Feynman method it is necessary not only to integrate over all direct paths between two space time points but also to sum over the infinity of classically distinct types of paths corresponding to repeated reflections of the particle from the walls of the box. Only then will the wave function satisfy the necessary boundary condition of vanishing at the box wall. In some vaguely similar sense the summation over the infinity of topologically distinct gravitational field histories may correspond to a special boundary condition which may or may not be satisfied in Nature.

[8] i.e., "analytic" in the sense of the theory of functions of a nondenumerable infinity of complex variables. Another, more startling example of penetration into a nonphysical region is provided by the strict probabilistic interpretation of the scalar wave equation (Klein-Gordon). Using the relativistic position operator defined by Newton and Wigner (1949) one can show that the correct propagation function for the position operator is not the invariant delta function $\Delta(x - x')$ which vanishes outside the light cone, but rather the time derivative of the function $\Delta^+ = \frac{1}{2}(\Delta - i\Delta^{(1)})$ which does *not* vanish outside the light cone, although it does damp out rapidly whithin a Compton wavelength for the light cone. This, of course, is due to certain nonlocal features of relativistic particles when viewed from the traditional quantum standpoint. And here we have the first hint that vaguely similar, although profoundly deeper, nonlocal features will make this appearance in quantum gravitational theory.

Because of the formidable nature of the topological problems there is no doubt that any calculations which may eventually be made in quantum gravitational theory will initially assume both constant signature and constant topology. In virtue of the rich potentialities which variable topologies provide, however, one would be exceedingly rash at this stage to assert the necessity of such constancy.

Occupying a position complementary to that of the Feynman theory is the Schwinger quantization method. Just prior to the Copenhagen gathering Professor DeWitt had discussed with Professor Schwinger in Lille (Mathematical Congress) the question of applying his (Schwinger's) methods to the gravitational field. One major difficulty facing Schwinger's technique is that he deals with operators from the outset, and therefore must know a great deal about the commutation relations (i.e., dynamics) of the system he is dealing with already in advance. This becomes particularly troublesome when the theory if nonlinear as Einstein's is. Schwinger's suggestion was to work with the Palatini formalism which introduces the affine components along with the metric as fundamental field quantities. The appropriate Lagrangian is then linear in the field derivatives and the commutation difficulties arising from quadratic occurrences of the derivatives are thereby eliminated – according to Schwinger.

This possibility was considered at Copenhagen, but apart from a curious ambiguity which was noted in setting up gravitational interactions with other field within the Palatini formalism, work was not pushed very far in this direction. The trouble with using Palatini à la Schwinger is that the difficulties are simply postponed to a later stage, viz. the stage where one must begin to look at the constraints, which, in the Palatini formalism, are of the "second class" as well as the "first" (Dirac 1950). Within the Schwinger formalism constraints must be handled *ad hoc*. Each type of system must be given a special treatment, and the gravitational field will require the most special (i.e., difficult) treatment of all. It seemed highly unlikely to the participants at Copenhagen that Schwinger's methods would really prove useful in the search for the hidden path to a quantum theory of gravitation. On the other hand, there is no doubt that when a quantum gravitational formalism is finally found Schwinger will be on hand to reformulate it.

4 Canonical quantization

Although the Feynman viewpoint enables one to obtain rapid and far reaching insights into the quantum theory of a given system, it is less successful in separating and isolating the technical difficulties which arise when calculations are attempted on the system. There is a tried and (so far) true method of quantizing any system which possesses a classical analog. This is the so-called canonical quantization method laid down originally by Schrödinger, Heisenberg and Pauli. About ten years ago Professor P. G. Bergmann at Syracuse University took up the problem of applying this method to the gravitational field. It is he who first exposed in detail the main difficulties besetting the problem and who introduced the notion of "true observables" (coordinate-invariant quantities) and their relation to the primary and secondary constraints of the system (Dirac 1950).

At Copenhagen some new progress was made on Bergmann's problem by Professor DeWitt. During the preceding year DeWitt had been engaged in an effort to develop mathematical tools for the analysis of the constraints which arise in a canonical system as a result of the existence of invariance groups. His program consists of a study of the types of transformations to which the constraints themselves may be subjected, the types of Lie algebras generated by the constraints, the interrelation of the constraints, the Lagrangian and the "true measure" (i.e. for Feynman quantization of the system), and the elimination of the constraints from the quantum theory – implying a simultaneous discovery of the "true observables" and resolution of the factor-ordering ambiguity.

At Copenhagen DeWitt succeeded in rigorously isolating half of the non-observable variables, those corresponding to the primary constraints. The chief problem was to find a canonical transformation which transforms the primary constraints into pure momenta. Professor J. L. Anderson of Stevens Institute of Technology had previously conjectured the possibility of such a transformation, but it had not up to this time been carried out by anyone. In order to find the transformation it was necessary to solve a set of simultaneous variational differential equations of a type which will continue to characterize DeWitt's work for some time to come. DeWitt found the solution only after several days of seemingly fruitless effort – an indication of the difficulty of the whole technical problem. However, his solution is rigorous and he has not yet been forced to resort to a successive approximation scheme.

DeWitt is now able to turn his attention to the secondary constraints, by far the most interesting quantities in the theory and to the understanding of which the major effort is now being devoted. It is impossible at the present time to predict when this work will be finished or when it will have reached a form suitable for publication. DeWitt has incidentally shown that Anderson's conjecture about the transformability of constraints is generally wrong. It is correct only if the Lie group generated by the constraints is Abelian, and it happens that the group generated by the primary constraints can be shown to be necessarily Abelian. The important question raises itself as to whether the group generated by the secondary constraints is also Abelian or not. If it were, and if DeWitt could find the transformation which changes the secondary constraints into momenta, then the problem of formally quantizing general relativity would at last be solved, and one could begin to compute some physical quantities. Unfortunately, the question is not easily answered. Strong arguments can be adduced on one side to prove that the group is Abelian, but equally strong arguments can be found on the other side to prove that it is not. These contradictory conclusions stem from reasoning by analogy with finite Lie groups. The Lie groups of general relativity are, however, infinite – and very little is known of the mathematical properties of infinite Lie groups. Nevertheless, Professor W. Pauli of Zürich has conjectured that infinite Lie groups will eventually turn out to be of greater importance in physics than finite parameter groups have ever been (Pauli 1957). This is mentioned to indicate the significance of the problem which still lies ahead.

Of importance in DeWitt's work is the use of the parameter formalism of P. Weiss (Weiss 1938). The parameter formalism has the following principal advantages:

1. The parameters serve to label the points in the underlying space-time manifold once and for all. They are not subject to transformation. The so-called "coordinates" have then a much greater functional similarity to the metric components, than in the ordinary formalism and, like the metric components, can be imposed in an arbitrary way on the underlying manifold (subject only to the secondary constraints).
2. The change of viewpoint achieved by the use of parameters permits one to perform functional operations and transformations which would otherwise be impossible. This promises to greatly simplify the work of transforming the secondary constraints into momenta (if this is at all possible).

To convey something of the utility of the parameter formalism some of DeWitt's results will be recorded here:

One of the parameters is chosen as the independent parameter T of a canonical formulation. Since the choice is arbitrary no invariance is destroyed. Surfaces $T =$ constant are then regarded as space-like and a normal vector density ℓ_μ is introduced.

After isolation of the non-observable variables corresponding to the primary constraints, DeWitt finds a Schrödinger equation of the form

$$\varphi_\mu \equiv \lambda_\mu + \mathcal{H}_\mu = 0 \tag{23}$$

where

$$\begin{aligned}\mathcal{H}_\mu \equiv \frac{1}{2}\ell_\mu G_{\alpha\beta\gamma\delta}\Pi^{\alpha\beta}\Pi^{\gamma\delta} - 2\ell^{-4}\ell_\mu\ell^\gamma\Pi^{\alpha\beta}\nabla_{\gamma\alpha}\ell_\beta - \ell^{-2}\left(2\ell^\alpha\Pi^{\beta\gamma}\right.\\ \left.-\ell^\gamma\Pi^{\alpha\beta}\right)\nabla_{\mu\nu}g_{\alpha\beta} - \nabla_{\mu\alpha}\left(\Pi^{\alpha\beta}a_\beta\right) + Y_\mu\end{aligned} \tag{24}$$

$$G_{\alpha\beta\gamma\delta} \equiv g^{-\frac{1}{2}}\ell^{-2}\left(\bar{g}_{\alpha\gamma}\bar{g}_{\beta\delta} + \bar{g}_{\alpha\delta}\bar{g}_{\beta\gamma} - \bar{g}_{\alpha\beta}\bar{g}_{\gamma\delta}\right), \tag{25}$$

$$g \equiv |g_{\mu\nu}|, \quad \ell^2 = -g^{\mu\nu}\ell_\mu\ell_\nu, \quad \bar{g}_{\alpha\beta} = g_{\alpha\beta} + \ell^{-2}\ell_\alpha\ell_\beta \tag{26}$$

$$\ell^\alpha \equiv g^{\alpha\beta}\ell_\beta, \quad \nabla_{\alpha\beta} = \ell_\alpha\frac{\partial}{\partial x^\beta} - \ell_\beta\frac{\partial}{\partial x^\alpha}. \tag{27}$$

The λ_μ are momenta conjugate to the coordinate χ^μ and the $\Pi^{\alpha\beta}$ are momenta conjugate to the remaining variables $\psi_{\alpha\beta}$. Here $\psi_{\alpha\beta}$ is a symmetric tensor with ten components, of which, however, only six are independent, and $\Pi^{\alpha\beta}$ satisfies $\Pi^{\alpha\beta}\ell_\beta \equiv 0$. In the quantum theory λ_μ becomes an operator $-i\hbar\delta/\delta x^\mu$ acting on the state functional $\psi[x]$ in the Schrödinger equation $\varphi_\mu\psi[x] = 0$. The relation between $\psi_{\alpha\beta}$, $g_{\alpha\beta}$ and the variables a_α conjugate to the primary constraints ϕ^α is: $g_{\alpha\beta} = \frac{1}{2}(\ell_\alpha a_\beta + \ell_\beta a_\alpha) + \psi_{\alpha\beta}$. The Y_μ are rather complicated functions of the a_α and $\psi_{\alpha\beta}$, but involving no momenta.

The Schrödinger equation is, as one would expect, seen to involve the momenta $\Pi^{\alpha\beta}$ quadratically, trough the "Hamiltonian vector" $\mathcal{H}_\mu$. However, only the component of $\mathcal{H}_\mu$ parallel to ℓ_μ is thus quadratic. The remaining three components are linear in the momenta and correspond to the redundancy in the description of space-like surfaces by means of parameters.

The secondary constraints are obtained by taking the Poisson bracket of φ_μ with the primary constraints ϕ^ν. DeWitt finds

$$\begin{aligned}(\varphi_\mu, \phi'^\nu) \equiv \Big[\frac{1}{2}\ell^\nu\mathcal{H}_\mu + \ell^{-2}\ell^\alpha\Pi^{\nu\gamma}\nabla_{\gamma\alpha}\ell_\mu - \frac{1}{2}\left(2g^{\alpha\nu}\Pi^{\beta\gamma} - g^{\nu\gamma}\Pi^{\alpha\beta}\right)\nabla_{\mu\gamma}g_{\alpha\beta}\\ - \nabla_{\mu\alpha}\Pi^{\alpha\beta} + \frac{1}{2}\ell^\nu\nabla_{\mu\alpha}\left(\Pi^{\alpha\beta}a_\beta\right) - \frac{1}{2}\ell^\nu Y_\mu\Big]\delta(u-u') + \frac{\delta Y_\mu}{\delta a'_\nu}\end{aligned} \tag{28}$$

where u and u' are the labels of points on a space-like surface. The above expression involves the momenta quadratically trough the occurrence of $\mathcal{H}_\mu$. However, because of the Schrödinger equation, $\mathcal{H}_\mu$ may be replaced by $-\lambda_\mu$, leaving an expression involving sixteen components which are linear in the momenta. Twelve of these components are mere repetitions of the three components of the Schrödinger equation corresponding to the aforementioned redundancy in the description of space-like surfaces. The remaining four components are the secondary constraints:

$$\begin{aligned}\chi_\mu \equiv -\frac{1}{2}\ell_\mu\ell^\nu\lambda_\nu + \ell^{-2}\ell^\alpha\ell^\nu g_{\mu\beta}\Pi^{\beta\gamma}\nabla_{\gamma\alpha}\ell_\nu - \ell^\nu\Pi^{\beta\gamma}\nabla_{\nu\gamma}g_{\mu\beta}\\ + \frac{1}{2}\ell^\nu\Pi^{\alpha\beta}\nabla_{\nu\mu}g_{\alpha\beta} - \ell^\nu g_{\mu\beta}\nabla_{\nu\alpha}\Pi^{\alpha\beta} + \frac{1}{2}\ell_\mu\ell^\nu\nabla_{\nu\alpha}\left(\Pi^{\alpha\beta}a_\beta\right) + Z_\mu = 0\end{aligned} \tag{29}$$

where

$$g_{\mu\alpha}\ell^\nu\frac{\delta Y_\nu}{\delta a'_\alpha} - \frac{1}{2}\ell_\mu\ell^\nu Y_\nu\,\delta(u-u') \equiv Z_\mu\,\delta(u-u'). \tag{30}$$

By means of the parameter formalism the secondary constraints have been rendered *linear* in the momenta. This will greatly simplify the procedure of transforming these constraints themselves into pure momenta, if this is at all possible. However, the most startling feature is the occurrence of $\ell^\nu \lambda_\nu$ in the fourth secondary constraint:

$$\ell^{-2}\ell^\mu \chi_\mu \equiv \frac{1}{2}\ell^\nu \lambda_\nu - \frac{1}{2}\ell^{-2}\ell^\mu \ell^\nu \Pi^{\beta\gamma} a_\beta \nabla_{\nu\gamma}\ell_\mu - \ell^{-2}\ell^\mu \ell^\nu \nabla_{\nu\gamma}\left(\Pi^{\beta\gamma}\psi_{\mu\beta}\right) + \ell^{-2}\ell^\mu Z_\mu \quad (31)$$

A transformation which changes this component into a pure momentum will *of necessity* produce a *functional mixing up* of the coordinates χ^μ with the metric components. After the "true observables" are then extracted the resulting theory will have a *nonlocal* character which is unique among field theories[9]. DeWitt is currently devoting his efforts to the study of expression (31), optimistic in his expectation that the secondary Lie group will prove to be Abelian like the primary group. However, even if this group proves to be non-Abelian all is not lost. It simply means that a (doubtless difficult) study of the classification and representations of infinite non-Abelian Lie groups will have to be undertaken.

Another advantage of the parameter formalism, noted in Copenhagen, was that it served in good measure as an aid to the participants in understanding one another. It served to fix ideas and to separate problems of general covariance from problems of constraints. By providing an explicit representation it served as an aid in translating the mathematical abstractions of Misner's work into concrete terms. For example, it enabled the other participants to understand Misner's result that *the Hamiltonian operator in any "topologically invariant" theory vanishes* (Misner 1957). It is not hard to see that Misner's use of the term "topologically invariant" is equivalent to using a dynamics based on the parameter T. But the state vector ψ does not depend on T, and the "Hamiltonian" corresponding to T consequently vanishes. True dynamics is obtained only through the constraint equation $\varphi_\mu \psi[x] = 0$.

5 Approximation methods

Since a definitive formulation of quantum gravitational theory is yet to be achieved, it may seem somewhat premature to consider the technical problems of computation at this stage. Nevertheless, attention was given to this matter at Copenhagen, partly with an eye toward seeing what difficulties one may expect to encounter in the future.

In the view of Dr. Deser, ordinary expansion methods (in hoped-for future calculations) will be worthless. He points out that in general relativity even a matter Lagrangian will have, strictly speaking, no "free-particle" part, because the "coupling" to the gravitational field enters already in the kinetic energy in a multiplicative fashion. He maintains that the special character of this fundamentally new mode of coupling will make itself felt most critically at high energies and forbid any approximations from "free-particles" just at the point where the divergence difficulties

[9] Here the nonlocality of the quantized gravitational field, hinted at in the preceding section, becomes explicit. A curious possibility arises at this point: The only momenta quadratically involved in the Schrödinger equation (23) are the $\Pi^{\alpha\beta}$. However, because the extraction of the true observables will involve the coordinates x^μ themselves in a peculiar way it may happen that the Schrödinger equation, after extraction, will involve also the λ_μ quadratically. The physical meaning of such a situation is not at all clear, just as the original interpretation of the relativistic wave equation, involving second time derivatives, was not clear. It might happen that a "square root" procedure analogous to that used in the development of the Dirac equation would have to be adopted, thus allowing both spinors and Fermi statistics to come into physics from Einstein's theory from a curious back door.

of modern field theories now exist. Dr. Misner also expresses a similar view, calling attention to the fact that if gravitation is to occupy a significant place in modern physics, it can do so only by being *qualitatively* different from other fields. As soon as we assume the gravitational field to behave qualitatively like other fields we find that it is quantitatively insignificant. It is in its qualitative difference that its very special importance lies.

It is fair to state that these points of view stem largely from the philosophies of Professors Klein and Wheeler respectively. In the qualitative uniqueness of the gravitational field Klein hopes to find a mechanism for "smearing out the light cone" and thus softening the singularities of relativistic quantum field theory. Wheeler on the other hand looks to the richness of topological possibilities in general relativity to characterize the uniqueness of the gravitational field. For Wheeler the vacuum is in a state of turmoil – a "foam like" structure with an exceedingly complex topology. This view of the vacuum prevents one from basing an approximation procedure on an assumption of Euclidean topology in zeroth order.

The computational suggestions of Wheeler and Misner have not progressed beyond the most rudimentary level of dimensional considerations. Deser's suggestions are hardly any further advanced. However, Deser is at least able to write down an explicit expression which make some of the difficulties explicit. Employing Feynman's functional integration technique he obtains the following approximate formal expression for the propagation function of a relativistic particle modified by its gravitational interactions:

$$\Delta(x,x') \sim N^{-1} \int \Delta(x,x',g_{\mu\nu})\, e^{\frac{i}{\hbar}A}\,[\delta g_{\mu\nu}] \tag{32}$$

where N is a normalization factor, A is the gravitational action and

$$\Delta(x,x',g_{\mu\nu}) \sim \left[\int_x^{x'} (g_{\mu\nu}dx^\mu dx^\nu)^{\frac{1}{2}}\right]^{-2} \tag{33}$$

the latter integration being taken along a geodesic. Making still further approximations one may write $A = \frac{1}{G}\int g_{\mu\nu}\Box^2 g_{\mu\nu} d^4x$ and, for x near x',

$$\Delta(x,x',g_{\mu\nu}) = \left[g_{\mu\nu}\,(x_\mu - x'_\mu)\,(x_\nu - x'_\nu)\right]^{-1} \tag{34}$$

so that

$$\Delta(x,x') \sim N^{-1} \int \frac{e^{\frac{i}{\hbar G}\int g_{\mu\nu}\Box^2 g_{\mu\nu}d^4x}}{g_{\mu\nu}(x)\,(x_\mu - x'_\mu)(x_\nu - x'_\nu)}\,[\delta g_{\mu\nu}] \tag{35}$$

One might now hope that by studying this formal functional integral one would gain certain insights as to valid approximation schemes entirely different from the usual expansion methods. For certain systems, in fact, this is occasionally possible, by treating the Gaussian forms into which the functional integral can be cast, as ordinary integrals. Unfortunately the integral above is not Gaussian as it stands and any attempt to obtain an asymptotic series by expanding the denominator fails to remove the singularity of the propagation function, which Deser shows by other arguments must be absent. It would seem that here is an example where the special features of the Feynman quantization technique prove to be of little advantage – at least as far as practical calculations are concerned.

At Copenhagen Deser and Laurent devoted considerable effort, but without much success, toward developing an approximation scheme, or at least the framework of one, based on Deser's functional integrals combined with the observation that in any quantum gravidynamical calculation $\hbar$ and G always occur together in the product $\hbar G$, and hence that the classical limit ($\hbar \to 0$) and the Euclidean limit ($G \to 0$) occur together.

To DeWitt this observation was strong evidence that, contrary to the opinions of the other participants, ordinary expansion procedures should be excellent. In his view the unique features – in particular, the non-locality – of the gravitational field must exist already in the classical theory. Once a canonical formalism is completely developed quantization should be straightforward. No agreement was reached on these points, and there the matter rests.

During the course of the Copenhagen gathering DeWitt obtained a result which, if it had been just the opposite, *would* have had a direct bearing on the matter. The investigation leading to this result actually arose initially from quite another source and not out of an effort to refute Deser's stand. It occurred to DeWitt that one might be able to find an appropriate functional transformation from the x^μ and $g_{\mu\nu}$ over a space-like surface[10] to another 14-fold, three dimensional continuum of quantities which would render the Einstein equation *linear*! Such a possibility caused DeWitt to fear that all his previous work might prove to be a great joke – that his carefully developed procedures might, in fact, be leading him steadily (however slowly) *to* such a linearizing transformation, which would then imply an isomorphism between Einstein's theory and a linearized theory. If this were true DeWitt could have saved himself a lot of work by looking for such a transformation at the outset. The basis for DeWitt's notion was not entirely idle or trivial. There *are* certain classes of nonlinear differential equations which can be rendered linear by transforming to new independent variables which are functions of both the original dependent and independent variables. Certain of the equations arising in the derivation of the well-known exact solutions of Einstein's equations are of this form. However, when DeWitt set up the variational differential equations which would have to be satisfied by such a functional transformation he found, by an expansion procedure, that they were mutually incompatible. Hence no such transformation exists.

Had such a transformation existed then Deser would have had no ground to stand on. The appropriate procedure in calculations with the gravitational field would obviously be to proceed just as with any other field. This would *not* mean that gravitation would fail to retain its unique nonlocal features, for the functional transformation would still mix up the x^μ and the $g_{\mu\nu}$ and hence introduce non-locality into the matter interactions. Even though the transformation does not exist, and Deser's stand remains therefore unrefuted, DeWitt's stand is not thereby imperiled. DeWitt maintains that the non-locality is still quite real – even classically – and that the impossibility of linearization simply means that the true observables of the gravitational field really do suffer self-interactions, trough a coupling which is also, presumably, nonlocal.

6 Some special problems

Very frequently during the Copenhagen work sessions the participants found themselves confronted by a constantly recurring difficulty, namely the inadequacy of the ready-at-hand concepts of other field theories when dealing with the gravitational field. Often an argument which seemed to be proceeding smoothly would suddenly find itself enmeshed in a contradiction, which on further analysis generally proved to be the result of the inapplicability of some idea or other which had always proved valid in previous contexts. This situation can be illustrated by a simple example, viz. the idea of radiation damping.

Consider first an electrically charged particle scattering electromagnetic radiation. An incoming electromagnetic wave causes the particle to being oscillating.

[10] For the consideration of such a possibility, use of the parameter formalism is essential. Without the parameter formalism the functional transformations would have to be restricted to the $g_{\mu\nu}$ alone, and it is easy to show that no linearization is possible within this framework.

The oscillation of the particle causes it to emit radiation of its own. But this emitted radiation reacts back on the particle, causing it to lag slightly behind the motion it would have if it did not radiate. This effect becomes more pronounced the higher the frequency of oscillation, causing the scattering cross section which is constant at low frequencies to diminish somewhat at very hight frequencies. Accompanying the lag in the particle's motion is also a net acceleration of the particle in the direction of the incoming radiation, representing the effect of radiation pressure. The scattering itself may be regarded as an interference effect between the incoming primary radiation and the emitted secondary radiation. As the lag of the particle becomes more pronounced (at higher frequencies) the interference effect becomes less efficient and the scattering cross section (and, in fact, the oscillatory amplitude of the particle) is "damped".

At first sight one would suppose that a similar phenomenon should occur in general relativity. One simply replaces electromagnetic waves by gravitational waves and the electric charge by the particle mass. For this type of a scattering computation the weak field approximation should be eminently suitable, since cross sections are normally amplitude independent. If one proceeds with the calculation, however, one finds that, in the weak field approximation, there is no scattering at all. The geodetic motion of a particle in a weak gravitational radiation field is simply uniform rectilinear. From a more sophisticated viewpoint, of course, this is not surprising. The left hand side of the Einstein weak-field equation has identically vanishing divergence (linearized Bianchi identities). On the right hand side stands the particle stress tensor, and this must have identically vanishing divergency too. But this happens only when the particle motion is uniform rectilinear. In order to obtain real scattering effects one must proceed to at least one stage beyond the weak field approximation. But here one runs into great complications because of the nonlinearity of the field equations.

The nonlinearity of the gravitational field equations is, of course, the mechanism by which Einstein's theory manages to conserve all forms of energy, including gravitational. The last point is important. In the informal discussions, replete with ready-at-hand concepts, which characterized the Copenhagen work sessions, it was very easy to run into a contradiction if one did not take care to keep close track of the energy contained in the gravitational field itself. But the gravitational field energy is a very elusive quantity. In the first place, it cannot be localized. In the second place, there are perfectly well behaved systems (e.g., closed universes; cylindrical gravitational waves) for which it seems curiously indefinable, even in the large. These problems, which remain unsolved today, were encountered at a very early phase in nearly every discussion.

Returning to the discussion of radiation damping, in order to complete the picture, we may next point out that there is another method of approach to the problem which is completely unavailable in other field theories. Since the choice of coordinates in general relativity is entirely arbitrary one may, if one likes, view the particle as being always at rest instead of oscillating. This suggests the natural approach of regarding the scattering particle as a Schwarzschild singularity – or, better, as the connecting link in a two-sheeted topology – and of treating the incoming radiation as a small perturbation on the Schwarzschild metric. The appropriate mathematics has already been worked out by Professor Wheeler and Dr. T. Regge (to be published). However, this completely transforms the problem. It becomes like a problem of scattering from a fixed potential, and it is very difficult to see how a phenomenon like radiation damping can make its appearance even in an exact solution. Or, rather, it is quite clear that although a Heitler integral equation can be set up for the scattering problem, the behavior of the phase shifts will have little relation to that in the electromagnetic case, and represent much more subtle effects.

Another aspect of the problem may be pointed out. First, let us note that the above discussion of the electromagnetic case was a bit hasty. If one tries to solve the problem rigorously, using a point particle so as to maintain relativistic invariance, one runs into a self-energy divergence difficulty. This can be circumvented by the method of Dirac which is basically a classical mass renormalization prescription (Dirac 1938). In the renormalized theory, however, the oscillating particle does not suffer a phase *lag* at hight energies, but a phase *advance* instead (so-called "pre-acceleration")[11]. In the gravitational case, with the use of the Schwarzschild metric, the renormalization problem does not arise. The mass of the particle is already determined by the metric at large distances. A new flexibility in boundary conditions appears, however. Boundary conditions must be specified on both topological sheets, and one has the possibility of allowing energy to leak from one sheet to the other.

7 Conclusion

The example of the preceding section shows in a simple way how radically different the gravitational field is from other fields. Although the participants repeatedly stressed this fact, they were nearly as often shocked or tricked by it, almost as if they didn't really believe it. In some sense the heart of the matter lies in the curious position of the concept of energy in general relativity. On the one hand the invariance of the theory leads to strong conservation laws, among which one expects the law of conservation of energy. On the other hand the very concept of energy somehow seems to dissolve. The participants agreed that this concept needs a thorough study and review. One is tempted to say that once the concept of energy is understood, everything will be understood. It seems to play such a central role in all the bizarre and specially interesting manifestations of the theory, from the nonlinearity of the equations to the occurrence of the operator $\ell^\mu \mathcal{H}_\mu$ in the secondary constraints of the canonical formalism.

It is quite possible, and even likely, that a major corollary of a successful development of a canonical formalism will be a clear cut definition of energy in the gravitational field. This is because a canonical formalism must of necessity define a quantity with the properties of energy, namely the operator which generates displacements in time. Professor Klein, who has laid great emphasis on the displacement properties of the energy operator in relativistic field theories, was impressed by the progress that DeWitt has achieved in the canonical approach, having been previously very pessimistic about the outlook for traditional methods. He was particularly pleased at DeWitt's demonstration of how easy it is to pass from the canonical formalism to the theory of Feynman and how the Dirac theory of constraints, which underlies DeWitt's work, contains Laurent's theorems on Feynman quantization as corollaries.

It is appropriate to conclude this report by repeating Klein's emphasis on the fact that the theory of gravitation is a theory of space and time. Whether this space-time has real nonlocal properties, what form these properties (if any) take, and whether the energy concept can be given a sharp definition, are still unknown, and it is impossible to say when the answers will be found. However, many new and interesting sides to these problems were learned and discovered by the participants at Copenhagen. This report has attempted to present the most important of these.

References

Dirac, P.A.M. 1938. Classical Theory of Radiating Electrons. *Proc. Roy. Soc. A* **167**: 148−169.

Dirac, P.A.M. 1950. Generalized Hamiltonian Dynamics. *Can. J. Math.* **2**: 129−148.

[11] The cross section still damps out at high frequencies.

Klauder, J. and Wheeler, J.A. 1957. On the Question of a Neutrino Analog to Electric Charge. *Rev. Mod. Phys.* **29**: 516−517.

Misner, C.W. 1957. Feynman Quantization of General Relativity. *Rev. Mod. Phys.* **29**: 497−509.

Newton, T.D. and Wigner, E.P. 1949. Localized States for Elementary Systems. *Rev. Mod. Phys.* **21**: 400−406.

Pauli, W. 1957. *Lectures Delivered to the CERN Theoretical Study Division in 1956-1957.*

Rainich, G.Y. 1925. Electrodynamics in the General Relativity Theory. *Trans. Am. Math. Soc.* **27**: 106−136.

Weiss, P. 1938. On the Hamilton-Jacobi Theory and Quantization of a Dynamical Continuum. *Proc. Roy. Soc. A* **169**: 102−119.

Open access funding provided by Max Planck Society.

Eur. Phys. J. H **42**, 177–259 (2017)
DOI: 10.1140/epjh/e2016-70034-0

THE EUROPEAN
PHYSICAL JOURNAL H

Robert Dicke and the naissance of experimental gravity physics, 1957–1967

Phillip James Edwin Peebles[a]

Joseph Henry Laboratories, Princeton University, Princeton NJ, USA

Received 27 May 2016 / Received in final form 22 June 2016
Published online 6 October 2016

Abstract. The experimental study of gravity became much more active in the late 1950s, a change pronounced enough be termed the birth, or naissance, of experimental gravity physics. I present a review of developments in this subject since 1915, through the broad range of new approaches that commenced in the late 1950s, and up to the transition of experimental gravity physics to what might be termed a normal and accepted part of physical science in the late 1960s. This review shows the importance of advances in technology, here as in all branches of natural science. The role of contingency is illustrated by Robert Dicke's decision in the mid-1950s to change directions in mid-career, to lead a research group dedicated to the experimental study of gravity. The review also shows the power of nonempirical evidence. Some in the 1950s felt that general relativity theory is so logically sound as to be scarcely worth the testing. But Dicke and others argued that a poorly tested theory is only that, and that other nonempirical arguments, based on Mach's Principle and Dirac's Large Numbers hypothesis, suggested it would be worth looking for a better theory of gravity. I conclude by offering lessons from this history, some peculiar to the study of gravity physics during the naissance, some of more general relevance. The central lesson, which is familiar but not always well advertised, is that physical theories can be empirically established, sometimes with surprising results.

1 Introduction

This is an examination of how the experimental study of gravity grew in the late 1950s and through the 1960s. The subject was a small science then that can be examined in some detail in the space of this paper. It offers a particularly clear illustration of the importance of ideas as well as technology in the origins of lines of research, which in this case have grown into Big Science.

In the mid-1950s the experimental exploration of gravity physics was generally considered uninteresting. This was in part because there seemed to be little that could be

[a] e-mail: pjep@princeton.edu

done, apart from incremental improvements of the three classical tests of general relativity. But a serious contributing factor was that influential scientists accepted general relativity theory as a compellingly logical extension from classical electromagnetism to the theory of gravity (an excellent example is the presentation in Landau and Lifshitz 1951, in the first English translation of the 1948 edition of *The Theory of Fields*). I take this broad acceptance of Einstein's general relativity, at a time when it had little empirical basis, to be an argument from the "nonempirical evidence" that respected scientists found the theory to be a logically compelling extension of what had come before. The term is borrowed from Dawid (2015), who argued for the merits of "non-empirical theory assessment" in present-day string theory, where the prospects for empirical assessment seem to be even more remote than it seemed to be for gravity physics in the 1950s. In Dawid's (2016) words,

> By the term non-empirical confirmation I denote confirmation of a theory by observations that lie beyond the theory's intended domain: unlike in the case of empirical confirmation, the confirming observations are not predicted by the theory they confirm. Non-empirical confirmation resembles empirical confirmation, however, in being based on observations about the world beyond the theory and its endorsers. Main examples of non-empirical confirmation are based on the observations that scientists have not succeeded in finding serious alternatives to a given theory, that in some sense comparable theories in the research field have turned out predictively successful once tested, and that a theory provides explanations that had not been aimed at during the theory's construction.

My use of the term "nonempirical evidence", in a second theme of this paper, is meant to be in line with this statement, but I have ventured to include considerations of the vague but commonly applied criterion of elegance, or simplicity. I take as prototype for this criterion Einstein's (Einstein 1945, p. 127) comment about the cosmological constant, Λ, in the appendix for the second edition of *The Meaning of Relativity*:

> The introduction of the "cosmologic member" into the equations of gravity, though possible from the point of view of relativity, is to be rejected from the point of view of logical economy.

The term Λ in Einstein's field equation remained unpopular among influential scientists in the 1950s, and increasingly so through the 1990s (as reviewed in Sect. 7.1). But despite its inelegance Λ was eventually added to the standard and accepted theory under the pressure of experimental advances. We may of course counter this example of the force of empirical evidence with a prime illustration of the successful application of nonempirical evidence: General relativity theory now passes demanding empirical tests.

The change in thinking about the possibilities of probes of gravity physics that led to the experimental establishment of general relativity is part of what Will (1986) termed the renaissance of general relativity, and Blum et al. (2015) termed its reinvention. Both names are appropriate for the subject as a whole, but on the empirical side the connotation of revival is inappropriate, because not a lot had happened earlier. There was a paradigm shift in community opinion, in the sense of Kuhn's (1962) *Structure of Scientific Revolutions*: In the 1950s it was generally accepted that there is little of interest to do in experimental gravity physics, as one sees in the heavy ratio of theory to experiment in the international conferences in the 1950s reviewed in Section 2. In the 1960s new directions in experimental programs were becoming a familiar and accepted part of science. But since there was not a shift in standard and accepted fundamental physics I avoid the word "revolution" and use instead the term, "the naissance of experimental gravity".

I take this naissance to have started in 1957, at the Chapel Hill Conference on *The Role of Gravitation in Physics* (DeWitt 1957), where Dicke (1957a) emphasized the sparse experimental exploration of gravity physics, the promise of new technology that could help improve the situation, and experiments to this end in progress in his group. Others may have been thinking along similar lines, but Dicke was alone in making a clear statement of the situation in print and starting a long-term broad-based program of empirical investigations of gravity. I take the naissance to have lasted about a decade before maturing into a part of normal science.

Ideas are important to the general advance of science, but they may play a particularly big role in the origins of a research activity. A clear illustration is the ideas Dicke and others found for new lines of research from old arguments associated with Ernst Mach and Paul A.M. Dirac. This situation was important enough to the development of modern gravity physics to be reviewed in some detail, in Section 4 of this paper.

Section 2 reviews the state of thinking about general relativity as a physical science at the start of the naissance, as revealed by the proceedings of international conferences in the late 1950s and early 1960s. Two of the leading actors in this search for "The Role of Gravitation in Physics" were John Archibald Wheeler and Robert Henry Dicke; their thinking is discussed in Section 3. Section 4 reviews ideas that motivated Dicke and others, and Section 5 presents a consideration of the style of exploration of these and other ideas in what became known as the Gravity Research Group. Activities in the group are exemplified by accounts of the two experiments in progress that Dicke mentioned at the Chapel Hill Conference. Section 6 presents more brief reviews of other significant advances in experimental gravity physics, from 1915 through to the nominal end of the naissance. The great lines of research that have grown out of this early work have shown us that the theory Einstein completed a century ago fits an abundance of experimental and observational evidence on scales ranging from the laboratory to the Solar system and on out to the observable universe. This is a striking success for assessment from nonempirical evidence, but I hope not to be considered an anticlimax. As noted, we have been forced to add Einstein's (1917) cosmological constant, despite its general lack of appeal. Empirical evidence can be surprising. Lessons to be drawn from this and other aspects of the history are discussed in Section 7[1].

[1] This paper was written in the culture of physics, and certainly could be made more complete by broader considerations, perhaps examined in the culture of the histroy of science. For example, in Section 7.5 I discuss support for research during the naissance by agencies that are more normally associated with the military. My thoughts in this section on why the agencies were doing this are only schematic; much more is known, as discussed for example in Wilson and Kaiser (2014), and I expect they and other historians are better equipped to follow this interesting story. Industries also supported postwar research in gravity physics. Howard Forward is listed as a Hughes Aircraft Company Staff Doctoral Fellow in the paper by Forward et al. (1961) on modes of oscillation of the Earth, which figured in the search for gravitational waves (Sect. 6.6). George Gamow's tour as a consultant to General Dynamics in the 1950s is celebrated for the missed chance for Gamow and Hoyle to hit on the thermal sea of radiation left from the hot early universe (as recalled by Hoyle 1981). I do not know why there were such connections between industry and gravity physics; as graduate students we used to joke that the aircraft companies were hoping to find antigravity. The Hughes Fellowships still exist, but I understand now focus on more practical training. During the naissance, nontenured faculty in physics at Princeton University typically had appointments halftime teaching and halftime research, the latter supported by a funding agency, often military. It allowed many more junior faculty than positions available for tenure. The junior faculty were supposed to benefit from experience and contacts that could lead to jobs elsewhere, usually successfully. And research greatly benefitted from many active young

2 General relativity and experimental gravity physics in the 1950s

The modest state of experimental research in gravity physics in the mid-1950s is illustrated by the proceedings of the July 1955 Berne Conference, *Jubilee of Relativity Theory* (Mercier and Kervaire 1956), on the occasion of the 50th anniversary of special relativity theory, the 40th for general relativity. Of the 34 papers in the proceedings there is just one on the fundamental empirical basis for general relativity: Trumpler's (1956) review of the modest advances in two of the original tests of general relativity, measurements of the deflection of light by the Sun and of the gravitational redshift of light from stars. Baade presented an important development, his correction to the extragalactic distance scale. He did not contribute to the proceedings, but the record of discussions of his report includes Robertson's (1956) comment on the possible need for "the disreputable Λ". Einstein's Λ became part of established gravity physics, but this happened a half-century later.

The experimental situation was improving, though slowly at first. At the March 1957 Chapel Hill *Conference on The Role of Gravitation in Physics* (DeWitt 1957; DeWitt and Rickles 2011), Dicke (1957a) stressed the contrast between the scant experimental work in gravity physics and the dense tests and experimental applications of quantum physics, and he outlined a research program he had commenced a few years earlier aimed at improving the situation. The only other commentary in the proceedings on empirical advances was Lilley's (1957) review of radio astronomy. Of immediate interest was the possibility of distinguishing between the Steady State and relativistic cosmological models by counts of radio sources as a function of flux density. The counts proved to be faulty for this purpose, but they have proved to be of lasting interest for their near isotropy, an early hint to the large-scale homogeneity of the observable universe (Sect. 6.13).

At the June 1959 Royaumont Conference on *Les Théories Relativistes de la Gravitation* (Lichnerowicz and Tonnelat 1962), Weber (1962) presented an analysis of how to build a gravitational wave detector. This was the only experimental paper among the 46 in the proceedings, and one may wonder how edifying the more technical aspects were to an audience that likely was almost entirely theorists. But Weber was introducing a new direction in the experimental investigation of general relativity theory.

We see a strikingly abrupt change of emphasis to the search for empirical probes of gravity in the July 1961 NASA *Conference on Experimental Tests of Theories of Relativity* (Roman 1961). Discussions of projects to which NASA could contribute included measurements of relativistic timekeeping in artificial satellites; tracking of artificial satellite orbits, including a satellite that shields a test mass from atmospheric drag and light pressure by jets that keep the enclosed test mass centered; the design conditions for a test of the relativistic Lense-Thirring inertial frame-dragging effect; and a search for detection of gravitational waves on a quieter site, the Moon.

The general lack of interest in experimental general relativity and gravity physics in the 1950s was at least in part a result of competition from many interesting things to do in other branches of physics, from elementary particle physics to biophysics. But

people. This comfortable and productive arrangement at elite universities ended during the Vietnam War, possibly a casualty of resentment of the particularly loud protests of the draft at elite universities that perhaps were least seriously afflicted by the draft. Perhaps there was something to our joke that Senator Mansfield sought to prevent the military from corrupting our young minds. But I am not capable of judging the truth of this matter, or the actual effect on curiosity-driven research. In Section 3, I mention Wheeler and Dicke's peaceful coexistence with their quite different philosophies of research in gravitation. Historians might see room for closer examination of this situation and, I expect, many other aspects of how empirical gravity physics grew.

Fig. 1. Members of the senior faculty in the Department of Physics, Palmer Physical Laboratory, Princeton University, in about 1950: from the left Rubby Sherr, Allen Shenstone, Donald Hamilton, Eric Rogers, Robert Dicke, Walker Bleakney, John Wheeler, Rudolf Ladenburg, and Eugene Wigner.

ongoing advances of technology were offering new possibilities for better probes of relativity. In some cases a particular technical advance motivated a specific experiment. For example, by 1955 Townes's group had a working ammonia beam maser (Gordon et al. 1955); Møller (1957) acknowledged "stimulating discussions" with Townes "on problems of general relativity in connection with the maser;" and a year after that Townes's group published a variant of the Kennedy-Thorndike aether drift experiment (Cedarholm et al. 1958; as discussed below in Sect. 6.2). And a year after publication of the Mössbauer (1958) effect, Pound and Rebka (1959) announced their plan to use it for an attempt at a laboratory detection of the gravitational redshift, and Pound and Rebka (1960) announced the detection a year after that (as reviewed in Sect. 6.4). Dicke followed this pattern, but with a particular difference: he was systematically casting about for experiments that may help improve the empirical basis for gravity physics.

3 Wheeler and Dicke on the role of gravitation in physics

At the 1957 Chapel Hill "Conference on The Role of Gravitation in Physics" John Archibald Wheeler spoke on the need to better understand the physical meaning of general relativity, which he accepted as the unquantized theory of gravity. Robert Henry Dicke spoke on the need to better establish the experimental basis for gravity physics, and perhaps find a unquantized theory that is even better than general relativity. The two are shown in Figure 1 with other members of the senior faculty of the

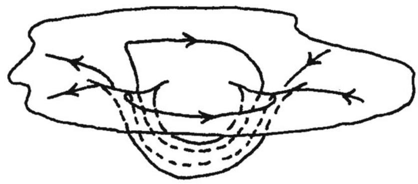


Fig. 2. Wheeler's sketch of a wormhole that is charged without charge, from the 1957 Chapel Hill Conference.

Department of Physics at Princeton University in about 1950. The photograph was taken a few years before both decided to turn to research in relativity and gravitation, when they were in mid-career: Wheeler was 44 and Dicke 39 in 1955.

Wheeler's research interests had been in theoretical nuclear, particle, and atomic physics. In his autobiography (Wheeler and Ford 1998, p. 228) Wheeler recalled that his notes from 1952 revealed that he had learned "from Shenstone 1/2 hour ago that I can teach relativity next year" which was "my first step into territory that would grip my imagination and command my research attention for the rest of my life" (Allen Shenstone, then chair of physics, is second from the left in Fig. 1). By the end of the 1950s Wheeler was leading an active research group in theoretical general relativity, with participation by students, postdocs, faculty, and a steady stream of visitors.

Dicke recalled (in Lightman and Brawer 1990, p. 204) that during sabbatical leave at Harvard in 1954–1955 he was thinking about the Eötvös experiment, which tests whether the acceleration of gravity may depend on the nature of the test particle. The constraint was remarkably tight, but Dicke saw that it could be done even better with the much better technology he could use. Bill Hoffmann (2016), who was a graduate student then, recalled that Dicke returned from Harvard "all fired up about gravity experiments". The quite abrupt change in direction of his active research career to gravity from what might be summarily termed quantum optics is illustrated by the list in Appendix A of the research topics of his graduate students before and after Dicke turned to gravity physics[2].

We have samples of what Wheeler and Dicke were thinking as they turned to investigations of relativity and gravity physics from the proceedings of the Chapel Hill Conference. The title of Wheeler's (1957) paper, *The Present Position of Classical Relativity Theory and Some of its Problems*, reflects the separation of topics at the conference to "unquantized general relativity" and "quantized general relativity". The latter was as fascinating, challenging and widely debated then as it is now. Figure 2 from Wheeler's paper shows an example of his thinking in classical relativity: a wormhole in space-time threaded by electric flux, giving us "charge without charge." He also spoke of "mass without mass," in his concept of a geon produced by the nonlinear interaction of electromagnetic and spacetime curvature fields. Wheeler's thinking about quantized general relativity is illustrated by a report of his contribution to a discussion at the Chapel Hill Conference:

> WHEELER envisions "foam-like structure" for the vacuum, arising from these fluctuations of the metric. He compared our observation of the vacuum with the view of an aviator flying over the ocean. At high altitudes the ocean looks smooth, but begins to show roughness as the aviator descends. In the ease of the vacuum, WHEELER believes that if we look at it on a sufficiently small scale it may even change its topological connectedness, thus (in the illustration

[2] Dicke's early life, the story of how he joined the faculty at Princeton University, and what he did when he got there, are reviewed in Happer et al. 1999.

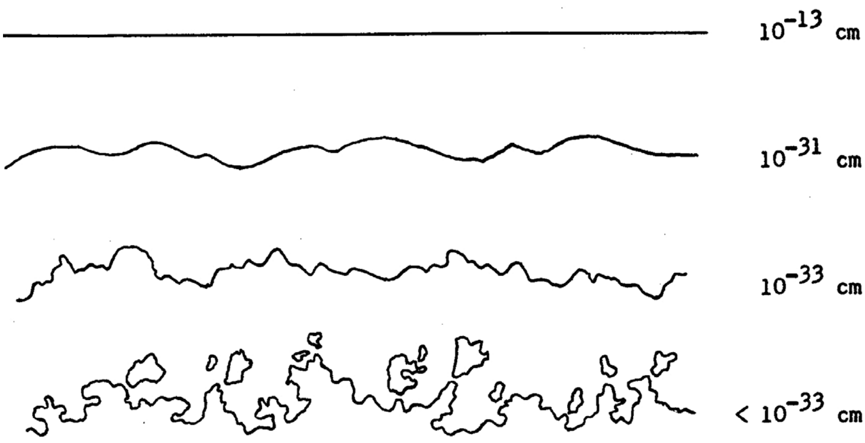


Fig. 3. Wheeler's sketch of spacetime foam, from discussion at the 1957 Chapel Hill Conference.

in Fig. 3):

Wheeler's imaginative approach to physics, and his modes of instruction of graduate students, produced great science and great generations of scientists. His inspiring style is seen in his opening paragraph in the Chapel Hill proceedings:

> We are here to consider an extraordinary topic, one that ranges from the infinitely large to the infinitely small. We want to find what general relativity and gravitation physics have to do with the description of nature. This task imposes a heavy burden of judgement and courage on us, for never before has theoretical physics had to face such wide subject matter, assisted by so comprehensive a theory but so little tested by experiment.

Wheeler's part in the renaissance of the theoretical side of general relativity deserves a study that would expand on his recollections in Wheeler and Ford (1998), but only a few brief points may be noted here. The presence of singularly active research groups in gravity theory and experiment in the same physics department, commencing in the 1950s, cannot have been planned: Wheeler and Dicke turned to relativity and gravity from other research directions well after arriving at Princeton. The presence of Einstein at the Princeton Institute for Advanced Study, and the relativist Howard Percy Robertson, who was based at Princeton University from 1929 to 1947, may have set a tradition that had some influence. Perhaps that is exemplified by Wheeler's feeling, which I heard expressed on occasion, that a philosophically satisfying universe has closed space sections, as Einstein had argued in the 1920s (and is reviewed in Sect. 4). Wheeler took an active interest in what was happening in Dicke's group. But Wheeler spent far more time studying the physical significance of general relativity theory as Einstein had written it down, and the possible approaches to its quantization.

Dicke's thinking about his change of direction of research is illustrated by these quotes from his 1957 Chapel Hill paper, *The Experimental Basis of Einstein's Theory* (Dicke 1957a, p. 5):

> It is unfortunate to note that the situation with respect to the experimental checks of general relativity theory is not much better than it was a few years after the theory was discovered – say in 1920. This is in striking contrast to the situation with respect to quantum theory, where we have literally thousands of experimental checks.

> ...
>
> Professor Wheeler has already discussed the three famous checks of general relativity; this is really very flimsy evidence on which to hang a theory.
>
> ...
>
> It is a great challenge to the experimental physicist to try to improve this situation; to try to devise new experiments and refine old ones to give new checks on the theory. We have been accustomed to thinking that gravity can play no role in laboratory-scale experiments; that the gradients are too small, and that all gravitational effects are equivalent to a change of frame of reference. Recently I have been changing my views about this.

In the second of these quotes Dicke was referring to Wheeler's summary comments on the classical three tests of general relativity: the orbit of the planet Mercury, the gravitational deflection of light passing near the Sun, and the gravitational redshift of light from stars. As it happens, the redshift test Wheeler mentioned, the measured shifts in the spectra of two white dwarf stars in binary systems, proved to be accurate in one case but quite erroneous in the other (as discussed in Sect. 6.4). Wheeler and Dicke might instead have referred to the measured wavelength shifts of solar absorption lines (St. John 1928), but this test was problematic because the measured shifts varied across the face of the Sun and varied with the depth of origin of the lines in the solar atmosphere, largely results of turbulence. The measured solar redshifts were roughly in line with general relativity, however, tending to scatter around the predicted value by no more than about 25%, so one might say that in 1957 general relativity had passed about two and a half tests.

At the Chapel Hill Conference Dicke mentioned work in progress in his group, largely in the discussion, beginning with this exchange:

> BERGMANN: What is the status of the experiments which it is rumored are being done at Princeton?
>
> DICKE: There are two experiments being started now. One is an improved measurement of "g" to detect possible annual variations. This is coming nicely, and I think we can improve earlier work by a factor of ten. This is done by using a very short pendulum, without knife edges, just suspended by a quartz fiber, oscillating at a high rate of around 30 cycles/sec. instead of the long slow pendulum. The other experiment is a repetition of the Eötvös experiment. We put the whole system in a vacuum to get rid of Brownian motion disturbances; we use better geometry than Eötvös used; and instead of looking for deflections, the apparatus would be in an automatic feed-back loop such that the position is held fixed by feeding in external torque to balance the gravitational torque. This leads to rapid damping, and allows you to divide time up so that you don't need to average over long time intervals, but can look at each separate interval of time. This is being instrumented; we are worrying about such questions as temperature control of the room right now, because we'd like stability of the temperature to a thousandth of a degree, which is a bit difficult for the whole room.
>
> ...
>
> We have been working on an atomic clock, with which we will be able to measure variations in the moon's rotation rate. Astronomical observations are accurate enough so that, with a good atomic clock, it should be possible in three years' time to detect variations in "g" of the size of the effects we have been considering. We are working on a rubidium clock, which we hope may be good to one part in 10^{10}.

PHYSICAL AND ASTROPHYSICAL CONSTANTS
(Units of $\hbar$, c, m)

10^0	10^{20}	10^{40}	10^{80}
Masses Elementary Particles	Reciprocal "weak Coupling Const.^{2} β decay, μ decay π → μ decay, etc., etc., etc.	Reciprocal Gravitational Coupling Const.^{2} $\left(\text{Ratio } \frac{\text{Electrical force}}{\text{Gravitational}}\right)$	Number of Particles in Universe to Hubble radius
e^2 $\left(\alpha = \frac{e^2}{\hbar c}\right)$		Age of universe	
Other "Strong" Coupling Const.$^{+2}$		Hubble radius of universe	

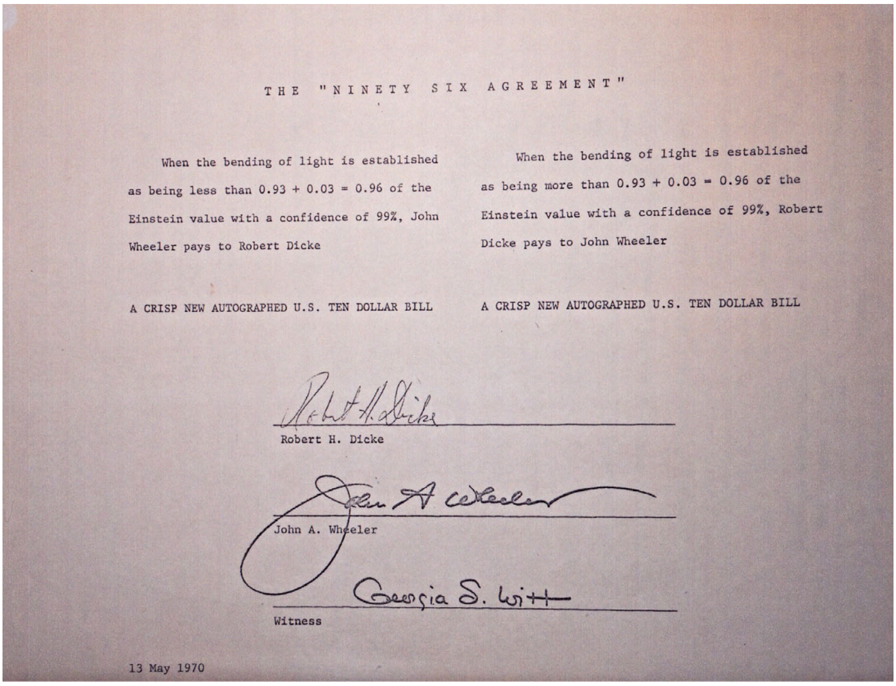
Fig. 4. Dicke's table of orders of magnitude of physical parameters, from his paper at the 1957 Chapel Hill Conference.

In these comments Dicke mentioned three experimental projects. Tracking the motion of the Moon is reviewed in Section 6.10.2. The pendulum and Eötvös experiments are discussed in Section 5, for the purpose of illustrating how Dicke's group operated.

In the third of the above quotes Dicke mentioned that "I have been changing my views". Though not stated at this point in the discussion, it seems likely that his new views included the comments earlier in the paper about the table in Figure 4 (copied from his Chapel Hill paper). The table was meant to motivate the idea that the strengths of the gravitational and weak interactions may evolve as the universe expands. This idea, and a more detailed discussion of the table, was presented in Dicke (1957b). We may suppose that Dicke's change of views also included his interest in Mach's Principle, for although he did not mention Mach in the Chapel Hill proceedings he advanced arguments along Machian ideas in Dicke (1957c). The next section reviews these ideas, their influence on Dicke and others, and their role in shaping the progress of experimental gravity physics.

While Dicke was actively casting about for evidence that might lead to a better gravity theory than general relativity, Wheeler was not at all inclined to question the validity of Einstein's relativity. Their differences are illustrated by the 1970 document[3] shown in Figure 5. At the time Dicke felt reasonably sure that the scalar-tensor gravity theory discussed in Section 4.4 fits the evidence better than general relativity. The details are reviewed in footnote 12 in Section 6.12; here we need only note a few points. The scalar-tensor theory predicts that the precession of the orbit of Mercury is smaller than in general relativity. However, Dicke (1964) had pointed out that if the interior of the Sun were rotating about as rapidly as the gas giant planets Jupiter and Saturn, the rotation of the solar surface being slowed by drag by the wind of plasma blowing away from the Sun, then the oblate solar mass distribution would contribute to the Newtonian precession, leaving a smaller non-Newtonian residual, consistent with the scalar-tensor theory. The Dicke and Goldenberg (1967) measurement of the shape of the Sun agreed with this idea. This interpretation required that the gravitational deflection of light by the Sun is no more than about 0.93 times the general relativity prediction. The added 0.03 in the document is not explained; it may be an indication of caution on Dicke's part. But my impression was that Dicke was confident of the

[3] The witness, Georgia Witt, was Wheeler's secretary. The document is discolored, as if it had been tacked to a wall for a long time. But I was only recently made aware of it by Martin McHugh, who found it in the Robert Henry Dicke Papers, Box 15, Folder W, Department of Rare Books and Special Collections, Princeton University Library.



THE "NINETY SIX AGREEMENT"

When the bending of light is established as being less than 0.93 + 0.03 = 0.96 of the Einstein value with a confidence of 99%, John Wheeler pays to Robert Dicke

A CRISP NEW AUTOGRAPHED U.S. TEN DOLLAR BILL

When the bending of light is established as being more than 0.93 + 0.03 = 0.96 of the Einstein value with a confidence of 99%, Robert Dicke pays to John Wheeler

A CRISP NEW AUTOGRAPHED U.S. TEN DOLLAR BILL

Robert H. Dicke

Robert H. Dicke

John A. Wheeler

John A. Wheeler

Georgia S. Witt

Witness

13 May 1970

Fig. 5. A wager on the tests of gravity theories.

argument because he had a new measurement – the shape of the Sun – that seemed to agree with two old considerations – Mach's principle and Dirac's large numbers in the context of the scalar-tensor gravity theory. Wheeler rested his case on the elegant simplicity of Einstein's general relativity theory, and he certainly gave the impression of confidence that the theory would continue to pass the improving tests. The network of tests is much tighter now, and it continues to agree with Wheeler's confidence in Einstein's theory. It might be noted, however, the evidence does require close to flat space sections, contrary to Einstein's and Wheeler's preference for a closed universe.

In my recollection Wheeler and Dicke were comfortable with their differences, but I do not know of any examples of joint research or more than casual exchanges of ideas. (An exception is that Wheeler guided Dicke's work on his 1939 Princeton undergraduate senior thesis, "A Logical Development of Quantum Mechanics and the Raman effect in the Atom.")

I recall that other senior members of the Princeton physics department in the 1960s felt that the path to a deeper fundamental theory would be through quantum particle physics. They respected Dicke and Wheeler as excellent physicists who had made curious career choices.

4 Mach's principle, Dirac's large numbers, and scalar-tensor gravity theory

Considerations along lines discussed in this section motivated Einstein's development of general relativity and his introduction of basic elements of modern cosmology. They also motivated Pascual Jordan in his creation of the scalar-tensor gravity theory, and Dicke in his development of empirical gravity physics. These are loosely specified ideas that have never been part of the broadly accepted belief system in physics and astronomy, but they are important for an understanding of how empirical gravity physics grew, and they continue to attract interest, for evolving reasons.

4.1 Mach's principle

Just as motion may be considered meaningful only relative to the rest of the matter in the universe, it is logical to some to conjecture that inertial motion is meaningful only relative to what the rest of the matter is doing. The idea has a long history; it is most familiar now from the discussion in Ernst Mach's book, *The Science of Mechanics* (Mach 1893, p. 284 in the 1960 edition of the English translation). Mach commented on Newton's point, that if the surface of the water in a bucket is curved then the bucket is observed to be rotating relative to the distant stars. This could be taken to mean that rotation has an absolute meaning, but Mach and others questioned that. In Mach's words (translated from the German),

> No one is competent to say how the experiment would turn out if the sides of the vessel increased in thickness and mass till they were ultimately several leagues thick. The one experiment only lies before us, and our business is, to bring it into accord with the other facts known to us, and not with the arbitrary fictions of our imagination.

Mach's admonition is in line with the theme of this paper: experiments matter (though Mach's positivist-empiricist philosophy was overly strict about imagination: Einstein's great imagination led to general relativity theory, which now proves to fit many facts later derived from experiments). The quote could be read to argue that the rotation of a very massive bucket might be expected to drag the motion of an inertial frame defined by local measurements, including the behavior of the surface of the water in the bucket. This would be an elegant anticipation of the relativistic Lense-Thirring inertial frame dragging near a rotating mass concentration such as the Earth, as follows. It is simplest to imagine a spherical shell of mass M and radius R rotating at angular velocity Ω. The general relativity prediction, in lowest approximation, is that an inertial frame inside the shell precesses relative to distant matter at angular velocity, ω, which is in order of magnitude

$$\omega \sim \frac{GM}{Rc^2}\Omega. \tag{1}$$

The successful test of this relativistic prediction applied outside the solid Earth is discussed in Section 6.8.

If in equation (1) we replace M by the observable mass of the universe in the relativistic Friedman-Lemaître cosmology that is expanding at about escape velocity, and we replace R by the Hubble length (the distance at which the linear relation between the distances of galaxies and their redshifts extrapolates to apparent recession velocity equal to the velocity of light), then the factor GM/Rc^2 is of order unity, so $\omega \sim \Omega$. Brill and Cohen (1966) examined this situation. It might invite one to imagine that, if the observable universe were said to be rotating, then local inertial frames would rotate with it, which is to say that the rotation would be meaningless. We may consider this qualitative argument to be one way to express Mach's Principle.

The broader physical implication of Mach's massive rotating bucket argument are still debated. Einstein, in his 1921 lectures on *The Meaning of Relativity* (Einstein 1923), noted that in general relativity a galaxy in otherwise empty asymptotically flat space-time could rotate, with all the usual effects of rotation, but it would be rotation relative to an otherwise empty universe. Einstein argued that, if this situation were allowed, it would mean that "Mach was wholly wrong in his thought that inertia, as well as gravitation, depends upon a kind of mutual action between bodies" and that "from the standpoint of epistemology it is more satisfying to have the mechanical properties of space completely determined by matter, and this is the case only in a space-bounded universe" (Einstein 1923, p. 119). In subsequent editions the sentence ends "only in a closed universe". Wheeler continued to argue for the philosophical appeal of this argument for a closed near homogeneous relativistic universe (as one sees in Misner et al. 1973, §21.12). Following Einstein (1923), some take Mach's Principle to be that a philosophically acceptable universe is described by a solution of Einstein's field equation in which inertial motion everywhere is determined solely by the distribution and motion of matter everywhere.

The considerations presented in the 1921 lectures may account for Einstein's (1917) bold proposal that, apart from local fluctuations, the universe is homogenous. There was no empirical evidence of this at the time. His 1917 argument for homogeneity is difficult to follow, but the 1921 reasoning seems clear: homogeneity would prohibit the phenomenon of non-Machian inertia in an asymptotically flat space that contains only a single concentration of matter. The concept of large-scale homogeneity (which might be stated to be that the universe is a spatially stationary and isotropic random process) has come to be known as Einstein's Cosmological Principle. It was very influential in the development of cosmology well before there was any observational evidence in support of homogeneity. The network of well-checked cosmological tests we have now make a persuasive case for Einstein's Cosmological Principle from Mach's Principle (Sect. 6.13.3), which may be counted as a notable example of the power of nonempirical evidence – when it is right.

One may debate whether the application of the Cosmological Principle really makes general relativity theory Machian. The issue does not seem to have long interested Einstein, but Machian ideas continued to fascinate others. For example, general relativity theory requires the same consistency of inertial motion defined by local experiments and observations of distant matter in a universe that satisfies the Cosmological Principle but has an arbitrarily small mean mass density. Dicke's thinking, as I recall it, was that this means Einstein's theory is unsatisfactory, that we need a better one that would make the distinction between inertial and noninertial motion meaningless in a universe that is empty apart from some test particles with arbitrarily small masses.

Examples of thoughts along such directions are to be found in Sciama (1953, 1964), Brans and Dicke (1961), Lynden-Bell (2010), and articles in the book, *Mach's Principle: From Newton's Bucket to Quantum Gravity* (Barbour and Pfister 1995). The sense of the thinking may be captured in Sciama's (1953) proposal that "if local phenomena are strongly coupled to the universe as a whole, then local observations can give us information about the universe as a whole". Such ideas, perhaps drawn from interpretations of Mach's principle, perhaps drawn from some other holistic concept of physical reality, inspired many of the gravity experiments discussed in Sections 5 and 6.

4.2 Dirac's large numbers hypothesis

Another line of thought that was influential during the naissance started with the observation that the ratio of electrostatic to gravitational forces of attraction of a proton and electron is an exceedingly large number. The ratio of the expansion time t in the relativistic Friedman-Lemaître cosmology to the characteristic time e^2/m_ec^3 defined by atomic parameters also is an exceedingly large number. And the two numbers have the same order of magnitude:

$$\frac{e^2}{Gm_pm_e} \sim t\frac{m_ec^3}{e^2} \sim 10^{40} \sim N. \tag{2}$$

Here m_e and m_p are the masses of the electron and proton, e is the magnitude of their charge, and c is the velocity of light. The number n_p of protons in the observable part of the universe, in the relativistic model, is really large, on the order of $n_p \sim N^2$. Dirac (1937) mentioned Eddington's discussion of these large numbers (without citation; he may have meant Eddington 1936, p. 272). Dirac's Large Numbers Hypothesis (LNH) was that these numbers are very large because N has been increasing for a very long time, and that the parameters in equation (2) have been changing in such a way as to preserve rough equality of the two ratios. If physical units may be chosen such that the atomic parameters in equation (2) are constant, or close to it, then preservation of the approximate numerical agreement of the ratios would require that as the universe expands, and the expansion time t increases, the strength of the gravitational interaction decreases, as:

$$G \sim t^{-1}. \tag{3}$$

This would mean that the number of protons is increasing as $n_p \sim N^2 \sim t^2$. Dirac was not clear about the manner of increase of n_p; he mentioned particle production in stellar interiors. The thought turned instead to the idea that n_p is the number of protons in the observable part of the universe, which has been increasing (in the standard cosmology and variants).

4.3 The weak and strong equivalence principles

Dicke argued that Dirac likely was on the right, Machian, track: As the universe evolves physical "constants" evolve under the influence of the evolving concentrations of matter. But my observation was that Dicke was even more attracted to ideas that suggested interesting experiments from which something of value might be uncovered. At the Chapel Hill Conference Dicke (1957a) mentioned experiments in progress motivated by the idea that Dirac's LNH

> would imply that the gravitational coupling constant varies with time. Hence it might also well vary with position; hence gravitational energy might contribute to weight in a different way from other energy, and the principle of equivalence might be violated, or at least be only approximately true. However, it is just at this point that the Eötvös experiment is not accurate enough to say anything; it says the strong interactions are all right (as regards the principle of equivalence), but it is the weak interactions we are questioning.
>
> ...
>
> Assuming that the gravitational binding energy of a body contributes anomalously to its weight (e.g., does not contribute or contributes too much), a large body would have a gravitational acceleration different from that of a small one. A first possible effect is the slight difference between the effective weight of an object when it is on the side of the earth toward the sun and when it is on the side away from the sun.

The comment about a difference of gravitational accelerations of a large body and a small one is worth noting. It may be related to Dicke's (1962a) later remark: if the strength of the gravitational interaction were a function of position, then the gravitational binding energy of a massive body, such as the planet Jupiter, would be a function of position, and the gradient of the energy would be a force that would cause the orbit to differ from that of a low mass particle with negligible gravitational binding energy. Finzi (1962) discussed the same effect, in connection with the orbits of white dwarf stars. Nordtvedt (1968) introduced the formal analysis of the effect.

The principle of equivalence Dicke mentioned at the Chapel Hill Conference has come to be termed the Strong Equivalence Principle: the prediction in general relativity theory that what is happening on Earth is quite unaffected by the disposition of all exterior mass, in our Solar System, galaxy, and the rest of the universe, apart from tidal fields and the determination of local inertial motion. The Weak Equivalence Principle is that the gravitational acceleration of a test particle is independent of its nature. It is tested by the Eötvös experiment, as Dicke noted. Early discussions of the distinction are in Dicke (1957a; 1959b, pp. 3–4; 1962a, pp. 15–31).

The pendulum experiment Dicke mentioned at the Chapel Hill Conference, to check for variation of the gravitational acceleration g at a fixed point on Earth, was a search for a possible violation of the Strong Principle. Maybe, Dicke suggested, the strength G of gravity varies as the Earth moves around the Sun and moves relative to the rest of the mass of the universe, producing an annual variation of g, or maybe G is decreasing as the universe expands, producing a secular decrease of g. At Chapel Hill Dicke also mentioned precision tracking of the motion of the Moon. He did not explain, but in later papers (notably Hoffmann et al. 1960) stated the purpose to be to check for Dirac's LNH expressed in equation (3), again in violation of the Strong Principle. Section 6.10.2 reviews how tracking the Moon grew into a demanding test of general relativity with a tight constraint on the LNH, including the variation of g that the pendulum experiments were meant to probe.

4.4 The scalar-tensor gravity theory

Pascual Jordan took Dirac's LNH seriously (as Schucking 1999 described). Jordan (1937, 1949) reviewed Eddington's (1936) and Dirac's (1937) considerations of the large numbers and, with Dirac, contemplated the idea that n_p is growing because matter is being created, maybe in stars. Jordan and Möller (1947) and Jordan (1948) took the LNH as motivation for replacing Newton's gravitational constant G by a scalar field in a scalar-tensor gravity theory in which the evolution of the scalar field could agree with the conjecture that the strength of the gravitational interaction is decreasing as the universe expands. In his book, *Schwerkraft und Weltall*, Jordan (1952) took note of Teller's (1948) point, that a larger G in the past would imply a hotter Sun, which if too hot would violate evidence of early life on Earth, a serious constraint. Also in *Schwerkraft und Weltall*, the setup of the scalar-tensor theory required local violation of energy-momentum conservation if the strength of the gravitational interaction were evolving, again leading Jordan to contemplate generation of matter (p. 143), perhaps in stars, perhaps causing the masses of stars to increase as G decreases. In *Schwerkraft und Weltall* and *Die Expansion der Erde* (Jordan 1966) Jordan considered the growing evidence for continental drift, which he pointed out might be caused by a decreasing value of G that relieved stresses that allowed the continents to move. In the preface to *Die Expansion der Erde*, Jordan wrote that "I must report that R. Dicke has independently arrived at similar hypothetical consequences" (as expressed in the English translation in Jordan 1971). Jordan (1966) may have been referring to Dicke (1961a) or (1962b), where he discussed issues of continental drift.

Fierz (1956) wrote down the special case of Jordan's approach to a scalar-tensor theory that preserves standard local physics, eliminating the violation of local energy conservation that Jordan was thinking might be relevant for stellar evolution. In the notation of Brans and Dicke (1961), Fierz's action (with units chosen so $c = 1$) is

$$S = \int \sqrt{-g}\, d^4x \left[\phi R + 16\pi L + w\, \phi_{,i}\phi^{,i}/\phi\right]. \tag{4}$$

Newton's gravitational constant G is replaced by ϕ^{-1}, where ϕ is a scalar field. Since ϕ does not enter the action L for matter and radiation, local physics is standard, local energy is conserved, and the gravitational acceleration of a test particle is independent of its nature, consistent with the Eötvös experiment. The presence of ϕ in the denominator of the gradient energy density term makes the units consistent. The source term for ϕ in equation (4) is $w^{-1}R$, where the Ricci tensor R is a measure of spacetime curvature. Brans (1961) showed that this can be reduced to

$$\Box\,\phi = 8\pi T/(3+2w), \tag{5}$$

where T is the trace of the matter stress-energy tensor. Thus we see that the strength of the coupling of the scalar field to the rest of physics scales as about w^{-1}, where the constant w is a free parameter. The larger the choice of w the smaller the source of variation of ϕ. That is, at large w equation (4) may approach the Einstein-Hilbert action of general relativity with constant ϕ and $G = 1/\phi$.

Carl Brans (Ph.D. Princeton 1961) arrived in Princeton in 1957 as a graduate student intending to work on mathematical analyses of spacetime structure. He recalled that Charles Misner (whose Ph.D. in 1957 was directed by Wheeler) suggested that he talk to Dicke, who was looking for someone to turn Mach's Principle and Dirac's LNH into a gravity theory. Dicke told Brans about Sciama's (1953) schematic model for how this might be done; Brans then independently hit on the scalar-tensor action in the form of equation (4), with preservation of local energy-momentum conservation (Brans 2008; 2016). Brans later learned that Jordan and Fierz had done it first.

In his earliest publications on gravity physics Dicke (1957b, 1957c) referred to Jordan (1952) and Fierz (1956). If at the time Dicke had recognized the significance of the theory in these papers it would have been in character for him to have told Brans about them rather than Sciama, but we can only speculate about this. In any case, we see that Dicke's references to Jordan's research on scalar-tensor gravity theory and its physical implications were consistently brief though complete. Jordan's theory was cited in Dicke (1957b, 1957c); Brans and Dicke (1961) and Dicke (1962c), on their considerations of the scalar-tensor theory; and in Dicke (1964) and Dicke and Goldenberg (1967), on the search for a solar oblateness that might allow more room for the scalar-tensor parameter w in equation (4). Jordan does not seem to have been overly troubled by this manner of acknowledgement of his work, as indicated by their correspondence[4]. In a letter to Dicke dated 2.7.1966 Jordan wrote

> I tried to study comprehensively the whole field of empirical facts which might be suited to allow a test of Dirac's hypothesis; and I have now the impression (or I hope to be justified to think so) that all relevant empirical facts are in best accord with theoretical expectation, and the whole picture is quite convincing. I shall be very anxious, as soon as the book is published, to learn what you think of it. (Some points are a little deviating from what you preferred to assume; I come to the result of a rather great value of $\dot\kappa/\kappa \sim 10^{-9}/$ year ... I intend to be in USA for about a month ... Naturally I should be extremely glad to have the occasion to visit you.

Here $\dot\kappa/\kappa$ is Jordan's notation for the fractional rate of change of the strength of gravity; the notation used here is $\dot G/G$. Dicke's reply, dated 7 July 1966, after a paragraph welcoming Jordan's plan to visit the USA, was

> I was pleased to learn that you are publishing a new book. I think that perhaps you would agree with me that the implications for geophysics and astrophysics of a time rate of change of the gravitational interaction is one of the most fascinating questions that one could consider. I always have my mind open

[4] Quotations from the Robert Henry Dicke Papers, Box 4, Folder 4, Department of Rare Books and Special Collections, Princeton University Library.

> looking for some new fragment of information that could have a bearing on this question. I am curious to know how you could have a time rate of change of gravitation as great as 10^{-9} per year and am looking forward to reading about it in your book.

At the time we felt reasonably sure that the evolution of G could not be faster than about a tenth of Jordan's value (Sect. 6.10). Jordan and Dicke disagreed about aspects of the science, as here, but I know of no indication of disagreements between the two beyond this normal course of events in research.

For a more complete review of the historical development of scalar-tensor theories see Goenner (2012). For other reviews of Jordan's and Dicke's thinking about an evolving G and its possible effects in geophysics and cosmology see Kragh (2015a, 2015b; 2016).

5 The gravity research group

The experimental advances in gravity physics in the 1950s and 1960s grew out of independent work in many laboratories. But since Dicke was the central actor it is appropriate to give particular attention to his approach. The examples discussed here are drawn from the projects he mentioned at the Chapel Hill Conference: a repetition of the classical Eötvös experiment and a search for possible variation of the gravitational acceleration by a suitably designed pendulum experiment. This discussion is meant to illustrate some of Dicke's characteristic methods, including recollections of how he assembled the Gravity Research Group.

5.1 The Princeton static and dynamic Eötvös experiments

The Eötvös experiment tests a starting idea of general relativity, that the gravitational acceleration of a test particle is independent of the nature of the particle. Dicke's (1957a) comments about the considerable advances in technology from what was available to Eötvös are illustrated in Figure 6. The photograph on the upper left shows Eötvös and colleagues, likely measuring gravitational field gradients. He later turned this methodology to the comparison of gravitational accelerations of a variety of materials (Eötvös et al. 1922, Eötvös posthumously). To avoid disturbing the instrument Eötvös had to let the balance come to rest in isolation, then approach it and promptly use the telescope for a visual observation of the orientation of the balance before it could respond to the gravitational field gradient produced by the mass of his body. The line drawing on the right, which shows the setup of the Princeton dynamic version, illustrates three of the improvements that Dicke mentioned in response to Bergmann's question: the balance was placed in a vacuum, to reduce dissipation and the attendant Brownian noise; the behavior of the balance was measured and recorded remotely, removing disturbances by the observer; and the system was buried to reduce the disturbing effects of temperature variations and the wind and the noise. Dicke also chose to compare the gravitational accelerations of test masses toward the Sun. The advantage was that a difference of accelerations toward the Sun would be manifest as an effect on the balance with a 24-hour period. That removed the need to rotate the apparatus, apart from checks of reproducibility. But this strategy of course required careful attention to diurnal disturbances.

The static version of the Princeton Eötvös experiment employed a feedback loop to hold a torsion balance in place; the dynamic version tracked the oscillation frequency of a torsion pendulum. Dicke mentioned the static approach at the March 1957 Chapel

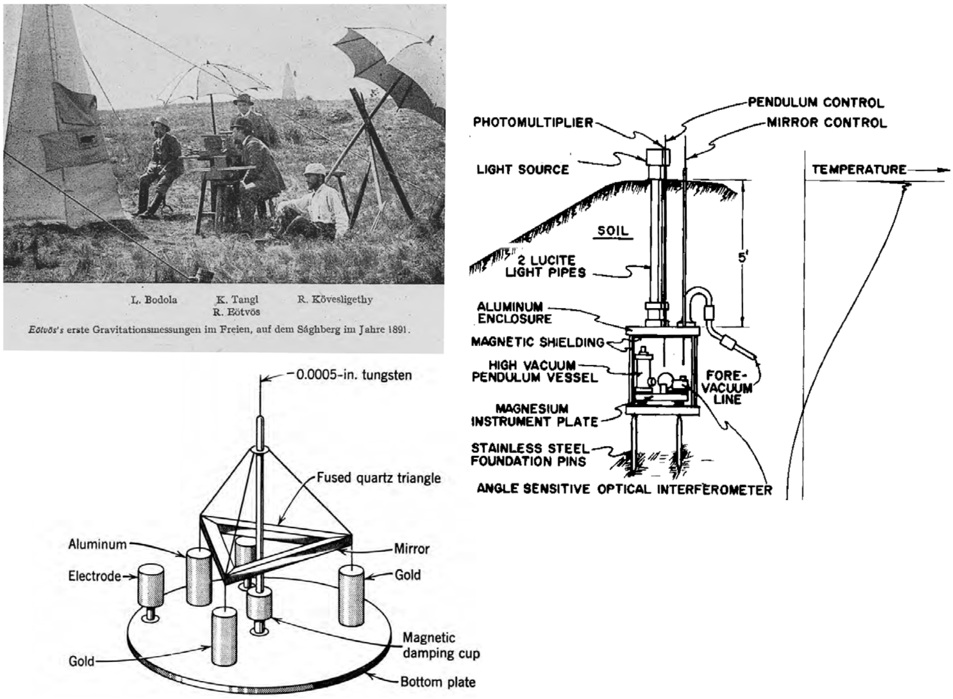


Fig. 6. The photograph in the upper left shows Eötvös and colleagues, likely measuring gravitational field gradients. On the right is an illustration of features common to the two repetitions of the Eötvös experiment in Dicke's group: the balance placed in a vacuum, the torque detected remotely, and the instrument buried. On the lower left is a sketch of the torsion balance used in the static version (in an early design).

Hill conference. In the fall of 1957 Sidney Liebes joined the Gravity Research Group and Dicke suggested to him the dynamic version and the optical interferometer method of precision timing of the pendulum. He left it to Liebes to design, construct, and operate the experiment.

In the dynamic version the torsion balance was a 2.5-cm long tubular bar of fused silica with an aluminum test mass inserted in one end and a platinum mass inserted in the other. The equilibrium position of the bar was oriented east-west. The bar was suspended by a thin fused-silica fiber (which Liebes drew by the traditional crossbow technique: fix one end of a silica bar, attach the other end to an arrow without a point, melt the center of the silica bar, shoot the arrow across the room, and then try to find the fiber). The amplitude of oscillation approached 90°, the period was approximately 8 minutes, and the decay constant in the high vacuum was several months, which is about 10^4 oscillations. The apparatus was buried five feet beneath the bleachers of the Princeton football stadium, to suppress diurnal temperature variations. The upper end of the torsion fiber was attached to a gimbal mechanism that enabled eddy current damping of pendulum vibrations induced by football games and other disturbances.

The left-hand part of Figure 7 is a top view of the angle-sensitive optical interferometer Dicke proposed to Liebes. One element is an optical flat on one side of the pendulum. The beam-splitter (semi-reflective mirror) caused the nearly monochromatic light from a low-pressure sodium lamp to follow the opposing paths through the interferometer marked by the solid and dashed lines. The emerging beams combined

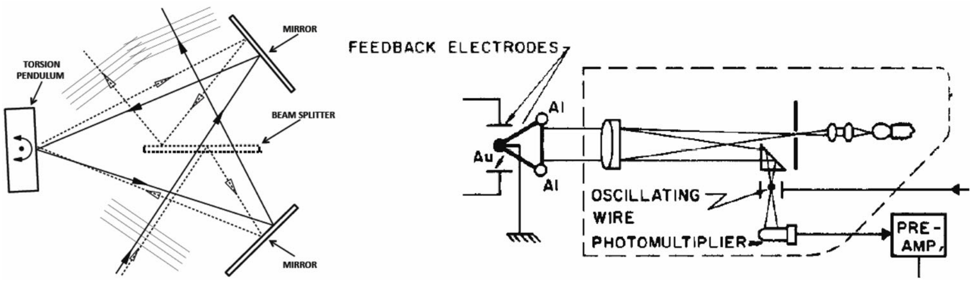


Fig. 7. The angle-sensitive detectors in the Princeton Eötvös experiments in the dynamic version, on the left (Liebes 1963, 2016), and the static version, on the right (Roll et al. 1964).

to form an interference pattern, the output of which was integrated by a photomultiplier. The bar is shown slightly displaced from the interferometric null-fringe position. In passage through optical null the fringes race apart and the photomultiplier detects a momentary characteristic pulse of light. The pulse timings were measured with a stabilized high-frequency oscillator and frequency counter system. A violation of the Weak Equivalence Principle would be manifest as a 24-hour variation of the gravitational torque on the pendulum by the Sun, which would produce a diurnal variation in the times of optical null as the torsion pendulum crossed in one direction, and an opposite shift in times of optical null when crossing in the other direction. This experiment was reported at a conference (Liebes 1963), but not published.

The static version was buried in another then more remote part of the campus. The sketch of the torsion balance on the lower left side of Figure 6 shows the fused silica triangle that held the test masses, gold and aluminum, in an arrangement that suppresses tidal torques. This triangle is drawn in heavy lines in the right-hand sketch in Figure 7, showing the two feedback electrodes that straddled the gold test mass (but straddled the aluminum mass in the earlier design in Fig. 6). Roll et al. (1964) described this part of the experiment as follows:

> At the heart of the experiment is the instrumentation for measuring very small rotations of the torsion balance ... The light is focused through a 25 μ slit, reflected from the aluminized flat on the quartz frame of the torsion balance, deflected off the telescope axis by a small prism, and the image of the slit focused on a 25 μ-diam tungsten wire. By locating this wire in the field of a small magnet and connecting it in a balanced bridge oscillator circuit, it was made to oscillate at its mechanical resonance frequency of about 3000 cps and with an amplitude of 25 to 50 μ ... When the diffraction pattern of the 25 μ slit produced by the 40 mm diameter telescope lens is centered exactly on the equilibrium position of the oscillating wire ... the photomultiplier will detect only the even harmonics of the 3000 cps fundamental frequency. As the torsion balance rotates slightly and shifts the diffraction pattern off center, the fundamental frequency will begin to appear in the photomultiplier output. The phase of the fundamental (0° or 180° relative to the oscillator signal driving the wire) indicates the direction of rotation of the pendulum, and its amplitude is proportional to the magnitude of the rotation for sufficiently small angular displacements ... The full width at half maximum of this curve (the "line width" which must be split by the detection apparatus) is about 30 μ or 3×10^{-5} rad ... these processes are all performed by a lock-in amplifier.

The 25 μ tungsten wire was informally known as the wiggle-wire. The experimental result is discussed in Section 6.7. Peter Roll (2016) recalled that, in writing the paper on this experiment (Roll et al. 1964),

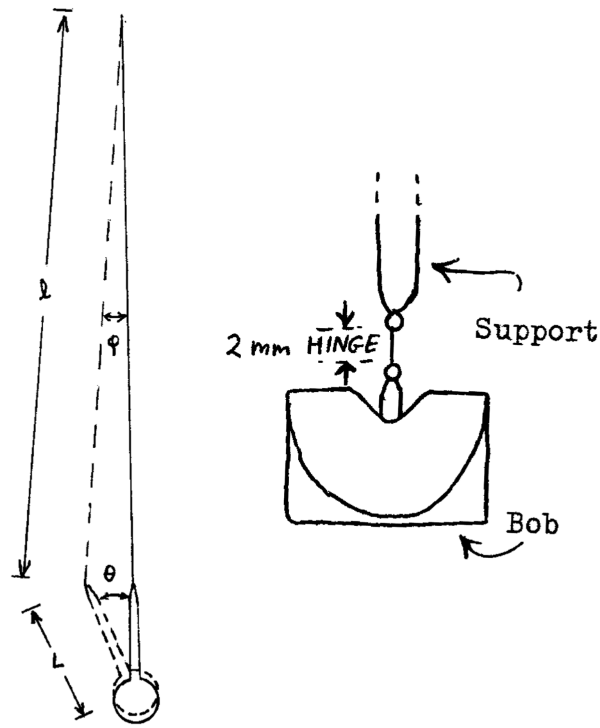


Fig. 8. Schematics of the Hoffmann (left) and Curott (right) pendulum designs.

it was important for the final paper on the Eötvös-Dicke experiment to document it in enough detail so that subsequent generations would not have to do as much legwork to understand exactly what was done and how. Dicke's contributions to the paper were the rationale for doing it in the first place and the basic design of the apparatus. Bob Krotkov got the first versions of the equipment designed in detail, built, and working. I came along at the end, contributing some of details of the Au-Al-Al apparatus and procedures, the final analyses of data and error sources, and analysis of Eötvös and Renner results. Bob Dicke was the master of both the theory behind the experiments and the experimental and apparatus design. His introduction to the 1964 paper explains the years he spent looking for a scalar gravitational field.

5.2 Searching for variation of the gravitational acceleration

The pendulum experiment Dicke mentioned at Chapel Hill was meant to test whether the gravitational acceleration at a fixed spot on Earth may vary as the Earth moves around the Sun, or as the universe expands. The ideas motivating this test are reviewed in Section 4. Features that Dicke mentioned at Chapel Hill are illustrated in Figure 8, in two different pendulum designs, the left-hand sketch from Bill Hoffmann's (1962) dissertation and the right-hand sketch from David Curott's (1965) dissertation. The rapid oscillations (22 Hz in Hoffmann's experiment, 5 Hz for Curott) made these pendulums less sensitive to external disturbances and much more sensitive to a slow variation of g, because more oscillations were observed in a given time. Each was drawn from a single piece of silica, with no knife edges that may wear, and very small energy loss in the vacuum. The pendulums were electrostatically driven. The dissipation times were about 6 hours for Hoffmann, 20 hours for Curott, or about 4×10^5 oscillations in each experiment. The period measurement during free decay used a source of light that passed through a slit, was reflected off an optical flat on the bottom of the pendulum, and focused on the vertex of a prism. As the pendulum swung the light passed from primary detection by a photomultiplier on one side of the prism to detection by a photomultiplier on the other side. A frequency standard with good short-term stability was phase-locked to the pendulum to follow the average pendulum crossing time.

Dicke chose the designs for the two versions of the Eötvös experiment, static and dynamic, and the two versions of the pendulum experiment. We may suppose he meant to explore possibilities for optimum methods. Was the time right for these exploratory experiments? Might they have been done much earlier? Hoffmann's (2016) response was

> A small fused silica inverted pendulum could have been fabricated and operated in a vacuum before the 1950's. But the accuracy of the measurement of the pendulum frequency depended on recent technological advances. These included sensitive phototubes used to detect the light beam reflected from the pendulum; low noise, stable, battery-operated vacuum tubes developed for undersea communication cables and used with the phototubes; lock-in amplifiers pioneered by Bob Dicke and used for precision phase locking and monitoring weak signals; fast counters and digital printers; a commercial ultra-stable crystal oscillator and a commercial atomic clock (General Radio Atomichron) for precise timing; and a programable digital computer (IBM 650) for Fourier analysis of the measurements to identify an annual variation. The experiment was built at the time transistors were coming into use and benefited by use of this new technology.
>
> ...
>
> The IBM 650 was purchased by Princeton and installed in the Gauss house, a Victorian home on Nassau Street, around 1957–1958. Prior to that Princeton had limited use of the Institute for Advanced Study MANIAC computer. During 1960 I was given the keys to the building for a night and spent the night there alone with with punched cards containing my data and a machine language Fourier transform program written by Bob Krotkov. The result is in my thesis. This would not have been possible two years earlier.

Curott (2015) expressed the same opinion, and added that his experiment

> couldn't have been done much earlier since it depended upon current (1960's) ultra-vacuum technology and electronic timing techniques. The emerging computer availability on campus also played an important role since daily corrections had to be made for minute by minute positions of moon and sun (tidal corrections). The data analysis would have been daunting, if not impossible, before electronic computers.

5.3 Dicke's synchronous detection

Dicke was the leading exponent of the method of synchronous lock-in amplifier detection employed in the Princeton Eötvös and pendulum experiments discussed above. He used it earlier in the Dicke (1946) microwave radiometer for suppression of the effect of receiver noise by switching between the source to be measured and a stable reference source, with detection of the output synchronized to the switching frequency. Lock-in amplifiers were also used in the Gravity Group measurement of the solar gravitational redshift (Sect. 6.4); the development of static gravimeters (Sect. 6.6); the measurement of the shape of the Sun (Sec 6.12); and the Princeton search for the CMB (Sect. 6.13.2). The first commercial lock-in amplifiers were produced in 1962 by the then privately held company, Princeton Applied Research, founded by Dicke and colleagues largely at Princeton University and the Princeton Plasma Physics Laboratory. Curott (1965) acknowledged use of a "Prototype of a model commercially available from Princeton Applied Research Corp". In connection with the measurements of the shape of the Sun (Kuhn et al. 1988 and earlier references therein), Kuhn (2016) recalled that

> Dicke's solar oblateness measurements both with Goldenberg and Hill, and later with Ken and I had, at their heart, at least one analog lock-in amplifier. Even in the 80's when we had early microcomputers to do synchronous demodulation, Bob had a very clever scheme to measure the position of the solar image in reference to the occulting disk using a 'JB8' lock-in amplifier. It made the Mt. Wilson oblateness measurements possible[5].

Weiss (2016) emphasized

> the critical idea in the Dicke Eötvös experiment in using quiet (low noise) sensors in a feedback loop to damp and position a mechanical instrument. The idea of holding an instrument at a fixed position and then reading it out by recording the force to hold it in that fixed position is critical to many precision mechanical experiments that have followed the Dicke Eötvös experiment. The feedback system is used to keep the response of the system linear and if done cleverly can be used to suppress normal modes of structures that through non-linearities in the system cause noise in the mode that carries the physical information of the measurement being made. LIGO uses many thousands of such feedback systems.

The reference is to the first successful detector of gravitational waves, the Laser Interferometer Gravitational-Wave Observatory (LIGO et al. 2016). Weiss (2016) added that

> the technique of modulating a physical effect to make it measurable above the $1/f$ noise used in the wiggle-wire telescope of the angle detector in the Dicke Eötvös experiment (Sect. 5.1) and in the Brault technique for finding the center of a spectral line (Sect. 6.4) was the other important technique developed by Dicke for precision experiments. The idea did not originate with Dicke but it was developed to an art by him and is now part of the stable of tricks used in virtually all precision measurements. The technique has a name – suppressed carrier modulation detection – and uses the lock-in amplifier.

The judgement of experimental colleagues at Princeton University was that, with its successors, the lock-in amplifier "probably has contributed as much to experimental Ph.D. theses as any device of the past generation" (Happer et al. 1999).

5.4 Assembling the gravity research group

How did Dicke assemble the Gravity Research Group group? Recollections of the five main contributors to the pendulum and Eötvös experiments discussed in this section offer a fair sample.

[5] The initials JB indicate Jim Braults' early design.

David Curott: When I entered the graduate program I intended to go into Controlled Fusion Research. I assumed there was no reason to question General Relativity. Without my knowing about Dicke's Group, in my first summer I was assigned a summer research assistantship in Dicke's Gravity Group, and that opened my eyes to the need for gravity research and the opportunities offered by new technologies. Dicke and his group excited my imagination and I stayed in the Group.

Bill Hoffmann: I entered graduate school at Princeton in the fall of 1954, after graduating from a small college (Bowdoin), intending to be a theoretical physicist. I soon found that this was not for me, but I was not interested in mainstream experimental work in atomic, nuclear, or high energy physics. When I first met Bob Dicke on his return from Harvard sabbatical in 1955 he was all fired up about gravity experiments. His enthusiasm and ideas about gravitation experimentation captivated me. I knew that this was the person I wanted to work with. I was attracted by the challenges to be inventive and the stimulation from others in his group.

Robert Krotkov: I joined Dicke's group because he invited me. I don't remember just what research I did at first – but I do remember where we first talked. It was in Palmer lab. As one came in the front door there was a hallway to the right, a hallway to the left, and straight ahead stairs leading up to a landing. It was on that landing that we happened to meet each other. How one remembers the really important things! Eugene Wigner directed my Ph.D. With Wigner, I would carry out some calculations, come to him, and he would tell me what to do next. With Dicke, I got a big picture and saw where what we were doing would fit in.

Sidney Liebes: I've had a lifelong fascination with relativity, first kindled, in my youth, by reading George Gamow's *Mr. Tomkins in Wonderland*. Nearing completion of my Ph.D. experiment at Stanford, testing a prediction of quantum electrodynamics, I asked George Pake whether he knew anyone doing experimental gravity physics. George's response: "Bob Dicke at Princeton". That prompted my phone call to Dicke, which resulted in an offer to join the faculty as an instructor and work in the Gravity Research Group. Dicke created an environment where I felt totally uninhibited in how I spent my time, never prompted, directed or monitored, beyond periodic Group meeting update contributions. I believe that environment to have been a critical factor in my independent rediscovery of gravitational lensing.

Peter Roll: My first introduction to general relativity and gravitation was a summer reading assignment of Ernst Mach's book on Mechanics, and several "wasted" hours in the Yale Physics Library digesting a small textbook on tensor calculus and GR, when I should have been working on physics problems and papers. In the spring of 1960 I was completing my Ph.D. dissertation in experimental nuclear physics at Yale and looking for a job. Bob Beringer, a senior faculty member at Yale and a colleague of Bob Dicke at the MIT Radiation Lab during the war, had just learned that Dicke was looking for a newly-minted Ph.D. to join his group at Princeton, finishing the Eötvös experiment that Bob Krotkov had started. Beringer encouraged me to look into the position and told me a bit about Dicke's background and accomplishments. After a telephone call and a trip to Princeton, my mind was made up – I wasn't going to find another position or another place that was anywhere near as interesting and challenging.

I might add that in the early 1960s, while I was attending Gravity Research Group meetings, I still had in mind writing a dissertation in theoretical elementary particle physics. But Bob had suggested that I look into constraints on possible evolution of the fine-structure constant from the degree of consistency of radioactive decay ages based on different long-lived atomic nuclei. That got me interested in the rhenium-osmium decay,

$$^{187}\mathrm{Re} \rightarrow ^{187}\mathrm{Os} + e^- + \bar{\nu}, \tag{6}$$

because I had noticed that the quite small decay energy made the decay rate quite sensitive to a change of the fine-structure constant. Since the decay energy was not well known I started looking into how I might better measure it. But Dicke had come to know me well enough to instruct me, first that I ought to stick with theory, and second that I am better suited to the kind of theory we had been doing in his group.

To exemplify the diverse directions of research in the group I refer to the list of reports of work in the Gravity Research Group at the January 1963 meeting of the American Physical Society[6]: Dicke on *Mach's Principle and Laboratory Physics*; Brault on *Gravitational Red Shift of Solar Lines*; Liebes on *Test of the Principle of Equivalence*; Turner and Hill on *New Experimental Limit on Velocity-Dependent Interactions of Clocks and Distant Matter*; Peebles on *Experimental Restrictions on Generally Covariant Gravity Theories*; Faller on *Absolute Determination of the Gravitational Acceleration*; Hoffmann on *Pendulum Gravimeter for Monitoring the Gravitational Acceleration as a Function of Time*; and Roll and Dicke on *Equivalence of Inertial and Passive Gravitational Mass*. By the time of this meeting other groups were making important contributions to experimental gravity physics, as reviewed in the next section, but none rivaled this searching range of investigations.

Worth recording also is Dicke's style of operation with his group. He tended to explain in some detail the motivation for a proposed project, outline possible methods, sometimes in detail, and then stand back to let us get to work. I recall David Wilkinson remarking that Dicke followed with great interest the construction of the Roll and Wilkinson microwave radiometer, built at Dicke's suggestion to look for radiation left from a hot Big Bang. But Wilkinson recalled that Dicke ventured very few suggestions about how they were doing it. I also recall Dicke encouraging me to keep thinking of ideas about possible implications of the Roll-Wilkinson experiment, whether a detection or upper limit. But I think the only idea he offered was the possible connection of the baryon Jeans mass in the hot Big Bang model to globular star clusters (Peebles and Dicke 1968).

6 A review of experimental gravity physics through the naissance

This is a review of highlights of experiments in gravity physics from 1915 through the naissance to its nominal completion in about 1968, as experimental research in this subject became part of normal science. I mean to include what astronomers term observations, and also probes of gravity physics derived from measurements made for other purposes. I mention Ph.D. theses completed in the course of research aimed at probing gravity, because students made considerable contributions, particularly so in Dicke's Gravity Research Group (It's worth recalling that Dicke's group often met on Friday evenings; we complained but attended because the discussions were too interesting to miss). This review of how the subject grew is largely assembled from the published literature, but I refer to personal recollections from some of the actors, many of them Dicke's students and colleagues, as well as my own experience since arriving at Princeton as a graduate student in 1958 and joining the group soon

[6] Bulletin of the American Physical Society, Volume 8, pp. 27–29.

after that. It should be understood that the sparse references to what happened after 1968 are meant only to illustrate what grew out of the naissance. Also, as a theorist, I am aware that a better-informed examination of experimental methods, and how they evolved during the naissance, would be a valuable addition to these summary comments.

6.1 A timeline for experimental gravity physics

The timeline in Figure 9 summarizes experimental developments in gravity physics from 1915 to the nominal completion of the naissance. The choice of developments to be marked required some creative accounting. The orbit of Mercury was the only demanding test of general relativity for a long time, so it ought to be marked; I put it at Einstein's 1915 relativistic interpretation of the excess precession (line 1 in the figure). I stretch the decade of the naissance to admit the Shapiro et al. (1968) measurement of the relativistic time delay of planetary radar pulses that pass near the Sun (line 38 in the figure). I mark a few developments that were initiated during the naissance and later became important tests. Thus line 22 marks Schiff's (1960) argument for a gyroscope experiment to test the relativistic dragging of inertial frames (Sect. 6.8), which led to Gravity Probe B (Everitt et al. 2015). The 1965 argument by the Dicke Group for a Lunar Laser Ranging Experiment (Alley et al. 1965; line 32 in the figure) is marked because it was an early step toward an extraordinarily demanding program of tests of gravity physics (Sect. 6.10.2). Zel'dovich's (1968, line 37) discussion of the quantum field zero-point contribution to the vacuum energy density is marked because, though little recognized then, it was and remains a major challenge for gravity and quantum physics.

Zwicky (1933) showed that, under conventional physics, the stability of the Coma cluster of galaxies requires "dunkle Materie" considerably in excess of what is seen in the stars (line 9 in Fig. 9). The problem was noted by others; an example is Babcock's (1939) demonstration that the gaseous nebulae on the outskirts of the galaxy M31 are moving relative to the galaxy much more rapidly than would be expected from the observed mass of the stars. This and other evidence for dark matter was neglected during the 1960s, but it was so important to the aftermath (as briefly reviewed in Sect. 6.13.3) that it is marked in the figure, at Zwicky's paper.

Line 11 marks another of Zwicky's (1937) contributions, his discussion of how a galaxy could act as a gravitational lens. It is entered because it was the first informed recognition of this possibly observable effect, though Zwicky's paper was little noticed until a reference to it appeared in broader analyses of gravitational lensing effects by Klimov (1963), Refsdal (1964), who cited Einstein's (1936) paper, and Liebes (1964), who cited Einstein (1936) and Zwicky (1937), while Zwicky in turn cited Einstein (1936). This is discussed in Section 6.5.

Jocelyn Bell Burnell found the pulsar phenomenon in 1967. This produced the first observational evidence for the existence of neutron stars, as pulsars (Gold 1968 and references therein). Pulsars are not marked in the timeline because their potential for demanding tests of gravity physics was seen only with the discovery of the Hulse and Taylor (1975) binary pulsar.

The paper on relativistic collapse by Oppenheimer and Snyder (1939) is important, but not marked because this theory had little to do with the evolving empirical situation in the 1960s. The evidence for the presence of supermassive black holes (or compact objects that act like them) in galactic nuclei was developing at this time. An early step was Burbidge's (1959) estimate of the energy in plasma in the radio sources in the most luminous radio galaxies, about 10^{60} ergs. This is equivalent to the annihilation of about 10^6 $M_\odot$, which seemed excessive for energy production in

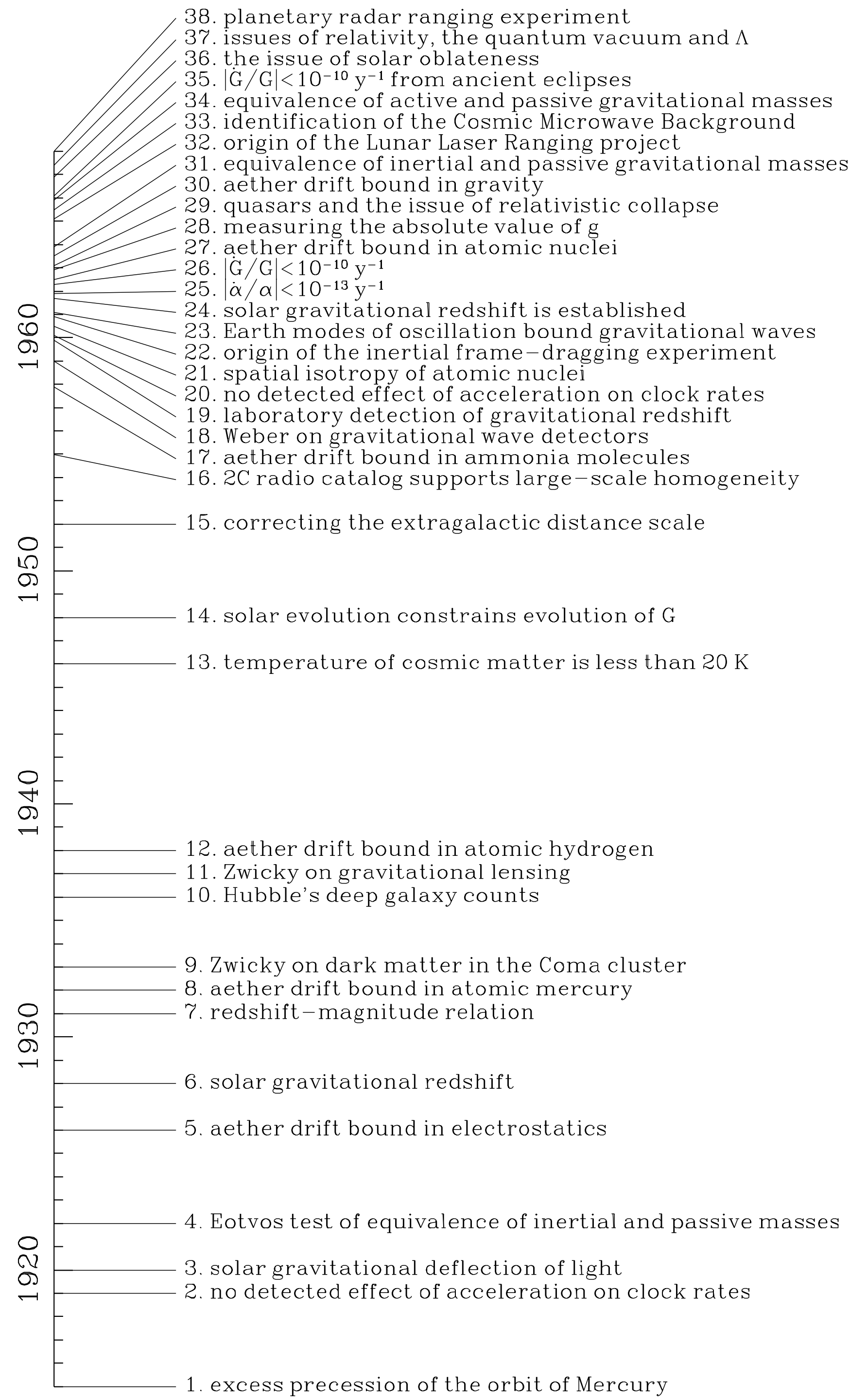


Fig. 9. A timeline of notable advances in experimental gravity physics, from the completion of general relativity theory in 1915 to the transition to a normal empirical science in the late 1960s. The exceptions to the rule that experiments are marked at completion are explained in the text.

stars. Schmidt's (1963) discovery of quasars added another energy problem, the source of the large optical luminosities of these compact objects. This forced community attention to the issues of gravitational collapse and black hole physics, as seen in the title of the 1963 Texas Symposium on Relativistic Astrophysics: *Quasistellar Sources and Gravitational Collapse* (Robinson et al. 1965). Lynden-Bell (1969) reviewed the proposal that the energy source for quasars and luminous radio galaxies is relativistic collapse to a massive black hole, and he argued that black holes remnant from quasar activity may be present in the nuclei of normal galaxies. This is now the paradigm. The discovery of quasars is entered in the timeline (line 29 in Fig. 9) to mark its effect on thinking about gravitational collapse.

The evidence for stellar mass black holes traces from the discovery of Cygnus X-1 as an X-ray source (Bowyer et al. 1965), and the evidence from optical identification that it is a binary system with a main sequence primary star and a companion X-ray source with mass at least 3 $M_\odot$ (Webster and Murdin 1972; Bolton 1972). That plausibly meant that the companion is too massive to be anything but a black hole. But the companion mass depends on an estimate of the mass of the primary, which depends on its distance, which was uncertain. The evidence now is that Cygnus X-1 contains a stellar remnant black hole with mass 15 ± 1 $M_\odot$ (Orosz et al. 2011). Although the stellar mass black hole phenomenon is now well established, and Cygnus X-1 was discovered in 1965, the evidence for stellar mass black holes developed late enough not to be marked in the timeline.

Incremental experimental advances in gravity physics are not marked. Thus the first measurement of the gravitational deflection of light passing close to the Sun (Dyson et al. 1920), is shown (line 3), but not the several repetitions, though as Trumpler (1956) emphasized they were very important in making the case that the solar deflection likely had been detected and is within about 10% of the predicted value. Trumpler reported that two white dwarf stars have measured redshifts consistent with predicted gravitational redshifts. The theory and observation for one of the white dwarfs are within 15% of present measurements, and they are consistent with general relativity. But the measured and predicted redshifts of the other white dwarf, although consistent with each other, both are one quarter of present results. Section 6.4 reviews this situation. For the timeline it seems to be more appropriate to mark St. John's (1928) case for detection of the gravitational redshift of light from the Sun (line 6 in Fig. 9). Although turbulent motions caused the measured redshift to vary across the face of the Sun, and to vary with atomic weight, the measurements generally were within about 30% of the prediction. I also mark Brault's (1962) solar redshift measurement, in line 24, because his strategy of solar line choice, and his much improved ability to measure the line shape and line center, successfully suppressed the systematic variation across the face of the Sun, which considerably improved the case for detection of the gravitational effect. The Pound and Rebka (1960) laboratory detection of the gravitational redshift is marked too (in line 19), because although its formal uncertainty was about twice Brault's the hazards of systematic errors in the two measurements were quite different. The consistency makes a very good case that the gravitational redshift effect really was detected by 1962, and determined to be within about 10% of the relativistic prediction.

The following review of gravity experiments offers still more explanations of this timeline. Will's (1993) analysis of experimental developments presented a timeline running from 1960 to 1980, in a useful complement to Figure 9. Another's selection of entries for Figure 9 certainly could differ, but we may be sure it would show the strikingly abrupt change in density of entries in the late 1950s, at the naissance, which is the point of this figure.

Table 1. Aether drift constraints.

Timeline[a]	Experiment	Bound[b]
–	isotropy of the velocity of light (1887)	$\epsilon \lesssim 0.01$
5.	static electromagnetic field (1926)	$\epsilon \lesssim 0.01$
8.	5461 Å atomic mercury line (1932)	$\epsilon \lesssim 0.03$
12.	molecular hydrogen ion Hβ line (1938)	$\epsilon \lesssim 0.1$
17.	24 GHz ammonia molecule inversion transition (1958)	$\epsilon \lesssim 10^{-3.5}$
21.	Zeeman splitting in the Li^7 atomic nucleus (1960)	$\epsilon \lesssim 10^{-16}$
27.	14.4 keV nuclear gamma-ray line (1962)	$\epsilon \lesssim 10^{-4.5}$
30.	annual variation of gravitational acceleration (1962)	$\epsilon \lesssim 0.1$

[a]Line number in Figure 9. [b]Bound on the coupling parameter in the models in equations (8) to (11) for the aether drift velocity in equation (7).

6.2 Tests for preferred motion

The experiments reviewed in this subsection test whether the time kept by a local clock, such as transitions in a molecule, or atom, or atomic nucleus, or the behavior of electromagnetic fields, might be affected by position or motion, in a violation of the Strong Equivalence Principle. The violation would of course have to be more subtle than the Galilean transformation that was ruled out by the demonstration that the velocity of light is close to isotropic. Some of these experiments were motivated by thoughts of an aether. Others might be attributed to the Machian thoughts discussed in Section 4, that the disposition and motion of all of the matter in the universe might define a preferred position or motion manifest in local physics.

The results of measurements summarized in Table 1 are expressed as bounds on a parameter ϵ that is meant to be a measure of coupling to an effective aether. For definiteness I take the velocity v of our motion through the aether to be comparable to our velocity relative to the thermal 3K Cosmic Microwave Background (the CMB discussed in Sect. 6.13.2),

$$v \simeq 300 \text{ km s}^{-1}. \tag{7}$$

The definition of ϵ has to be *ad hoc*, of course, because our motion through an effective aether need not be comparable to our motion through the CMB, and, more important, because we do not have a viable theory of ϵ.

The first entry in the table, from before the start of the timeline, is the Michelson-Morley bound on the anisotropy of the velocity of light[7], which in a Galilean transformation bounds the effective motion relative to the aether to $v_e \lesssim 4$ km s^{-1}. The coupling parameter here is defined by

$$v_e = \epsilon v, \tag{8}$$

with v in equation (7). This implies the Michelson-Morley bound, $\epsilon \lesssim 0.01$, entered in the last column of the first line of the table.

Jaseja et al. (1964) compared frequencies of two neon-helium masers (in which helium pumps population inversion for neon emission lines at wavelength $\sim 10^4$ Å) placed in orthogonal directions on a platform that can rotate. The stimulated emission frequency is defined by the mirror separation, so the Galilean argument for the Michelson-Morley experiment applies here: the relative frequency difference is expected to be $\delta\nu/\nu = \frac{1}{2}(v_e/c)^2 \cos\theta$, by the Galilean argument. The bound,

[7] Michelson 1903, in his book *Light waves and their uses*, reviewed the situation, with the memorable conclusion that "The theory may still be said to be in an unsatisfactory condition".

$\delta\nu/\nu \lesssim 10^{-11}$, on the effect of rotation of the system, so as to point it in different directions relative to an aether drift, indicates $v_e \lesssim 1$ km s^{-1}, a factor of three better than Michelson-Morley. The older method did so well with a much more poorly defined frequency because the same wave, or photon, sampled the path difference.

The second entry in the table is Chase's (1926) revisit of a pre-relativity expression for the magnetic field $\boldsymbol{B}$ induced by uniform motion of a charge distribution at velocity $\boldsymbol{v}$ relative to the aether. A Galilean transformation of a static charge distribution and electric field produces magnetic field

$$\boldsymbol{B} = \epsilon\, \boldsymbol{v} \times \boldsymbol{E}/c. \tag{9}$$

With $\epsilon = 1$ these electric and magnetic fields are a solution to Maxwell's equations[8]. We know now that this solution does not represent a static situation, of course, but the constant ϵ in equation (9) allows a measure of the possible coupling of the local electromagnetic field to Chase's effective "stationary ether" determined by the rest of the matter in the universe. Equation (9) would indicate that a charged capacitor moving at velocity $\boldsymbol{v}$ relative to the aether induces a magnetic field whose energy depends on the angle between $\boldsymbol{E}$ and $\boldsymbol{v}$, producing a torque to minimize the energy. Electromagnetic energies are large enough that Chase's torsion balance yielded an impressively tight constraint: the velocity of the experimental apparatus relative to the local stationary ether has to be $v \lesssim 4$ km s^{-1} if $\epsilon = 1$. Motion relative the stationary ether defined by the CMB (Eq. (7)) gives the bound on ϵ in the last column of the table. The first column is the line number in the timeline in Figure 9.

Kennedy and Thorndike (1932) sought to complete the argument for special relativity from the Michelson-Morley experiment by testing whether the time kept by a source of light for an optical interferometer might vary with the motion of the source relative to some preferred frame (while the interferometer arm lengths might be unchanged, or perhaps change in some other way). Their interferometer with unequal arm lengths could then show a variation of the interference pattern as the motion of the light source changed. The frequency of the light source as a function of velocity $\boldsymbol{w}$ of the mercury atoms relative to an effective aether may be modeled as

$$\nu = \nu_o(1 \pm \epsilon\, w^2/c^2), \qquad \boldsymbol{w} = \boldsymbol{u} + \boldsymbol{v}, \tag{10}$$

where $\boldsymbol{u}$ is the circumferential velocity of the interferometer due to the rotation of the Earth, $\boldsymbol{v}$ is the velocity of the Earth relative to the aether, and the parameter ϵ replaces the factor 1/2 in the Kennedy-Thorndike model (their Eq. (5)). The diurnal fractional frequency shift Kennedy and Thorndike looked for is then

$$\frac{\Delta\nu}{\nu} = \frac{2\epsilon\, u\, v}{c^2} \cos\, \phi, \tag{11}$$

where ϕ is the angle between the two vectors (with the sign in Eq. (10) absorbed in ϕ), and the terms that are nearly isotropic for purpose of the diurnal measurement are ignored. It is worth pausing to consider the numbers entering their result. The light was the spectral line of atomic mercury at wavelength $\lambda = 5461$ Å. The difference of path lengths in the two arms was Δs ~ 30 cm, which translates to $n = \Delta\mathrm{s}/\lambda \simeq 6\times10^5$ wavelengths. The measured bound on the diurnal fringe shift was $\delta n < 6\times10^{-5}$, which gives $\delta n/n \lesssim 10^{-10}$. We are considering the model $\delta n/n = \delta\nu/\nu$. The circumferential (Earth rotation) speed is $u \sim 0.3\,\mathrm{km\,s^{-1}}$. These numbers in equation (11), more

[8] Consider a static charge distribution ρ with electric field that satisfies $\nabla \cdot \boldsymbol{E} = 4\pi\rho$ and $\nabla \times \boldsymbol{E} = 0$. To represent uniform motion through the aether at velocity $\boldsymbol{v}$ let the charge density be $\rho = \rho(\boldsymbol{r} - \boldsymbol{v}t)$, with current density $\boldsymbol{j} = \rho\boldsymbol{v}$ and $\boldsymbol{E} = \boldsymbol{E}(\boldsymbol{r} - \boldsymbol{v}t)$. Then equation (9) with $\epsilon = 1$ satisfies Maxwell's equations.

carefully used, give the Kennedy-Thorndike bound on the velocity of the Earth relative to the aether, $v < (24 \pm 19)/\epsilon$ km s^{-1}. With the adopted Earth velocity (Eq. [7]) this translates to the bound on ϵ in the third line in the table. It depends on the impressively tight bound on the fringe shift $\delta n/n$.

Ives and Stilwell (1938) tested the second-order transverse Doppler shift in a beam of hydrogen consisting largely of the molecular hydrogen ion H_2^+ and the heavier ion with three protons bound by two electrons. They measured the longitudinal and transverse shifts in the analog of the Hβ line in the molecules at $\lambda \sim 5000$ Å. The measured transverse Doppler shifts in beams moving in the North, South, East or West directions, were found to differ by no more than about $\Delta\lambda \sim 0.003$ Å. This bounds the fractional shifts in times kept by the ions to $\Delta\nu/\nu \lesssim 10^{-6}$, for ions moving at beam velocities $u \sim 10^3$ km s^{-1}. The bound on ϵ from equation (11) is entered in the fourth row in Table 1. Recent precision tests of the relativistic Doppler effect by variants of the Ives and Stilwell experiment considerably improved this bound. Botermann et al. (2014) reported that measurements of frequency shifts of hyperfine structure lines of Li^+ ions moving at speed $w/c = 0.3$ agree with the relativistic prediction to a few parts in 10^9. Since they do not report problems with reproduceability as the Earth rotates and moves around the Sun we may take it that their bound in equation (10) bounds the coupling parameter to $\epsilon \lesssim 10^{-8}$ for the hyperfine structure lines.

Cedarholm et al. (1958) used ammonia beam masers for an aether drift test based on Møller's (1956, 1957) considerations of how atomic clocks might test relativity. The experiment compared the frequencies of two masers with ammonia beams that moved in opposite (antiparallel) directions at beam velocity $u \simeq 0.6$ km s^{-1}. The fractional change of frequency when the system was rotated by 180° was bounded to $\Delta\nu/\nu \lesssim 10^{-12}$, in trials repeated for a year (Cedarholm and Townes 1959). Møller (1957) and Cedarholm et al. (1958) argued for the expression in equation (11), without the parameter ϵ, from a consideration of the Doppler effect on radiation reflected by the walls of a resonant cavity moving through an effective aether. But it seems best to follow Turner and Hill (1964) in going directly to the model in equation (11), including ϵ, as a fitting function. Here the bound on $\Delta\nu/\nu$ translates to the bound on ϵ entered in the fifth row of the table.

Cocconi and Salpeter (1958) wrote that, "If Mach's Principle holds, we might then expect that the slight asymmetries in the distribution of matter at large would result in slight deviations from at least some of the laws of mechanics and gravitation which are commonly assumed to be exact", and could produce a "diurnal variation in the period of a quartz crystal (or a pendulum) clock". They did not refer to similar arguments by Sciama (1953) and Dicke (1957b), or the pendulum experiment Dicke (1957a) mentioned at Chapel Hill, but it is easy to imagine Cocconi and Salpeter were thinking along the lines of one of the many other trails of thought tracing back to Mach's arguments. They went on to consider the possibility that the electron inertial mass is slightly different for motions transverse and parallel to a preferred direction set by the large-scale distribution of matter, and pointed out that an inertial mass anisotropy would cause the Zeeman splittings between atomic levels with neighboring magnetic quantum numbers, m, to differ, depending on the orientation of the magnetic field relative to some preferred direction, because the patterns of electron motions relative to the magnetic field differ for different values of m. Cocconi and Salpeter (1960) suggested that the test could be even more sensitive if the Zeeman splittings were probed by the Mössbauer (1958) effect. Hughes et al. (1960), Drever (1960, 1961), and Virgilio Beltran-Lopez (1962, Ph.D. Yale) turned to a still more sensitive test by nuclear magnetic resonance measurements of Zeeman splittings of the four levels of the $I = 3/2$ spin of the atomic nucleus of Li^7. In a simple model of a $P_{3/2}$ neutron in the potential well of the six inner nucleons, the constraint on anisotropy of the nucleon inertial mass was found to be $\delta m/m < 10^{-20}$. This impressed Dicke; it is the subject of my first paper with him (Peebles and Dicke 1962a). In the model for $\delta m/m$ expressed in the form of equation (10), with the nucleon velocity approaching a few percent of the velocity of light, the bound on coupling to an effective aether is in the fifth row in Table 1.

Two groups independently tested for preferred motion using Mössbauer's (1958) discovery and explanation of the very narrow nuclear γ-ray absorption line spectrum produced when the recoil momenta of the nucleus emitter and absorber are taken up by the crystal lattices rather than by the atomic nuclei. In the report of their experiment, Champeney et al. (1963) referred to Ruderfer (1960), who proposed placing the γ-ray source at the center of a turntable and the Mössbauer absorber near the edge. The origin of the other experiment, at Princeton, may be traced back to a November 1959 letter from Dicke to Robert Pound at Harvard. Pound (2000) recalled that Dicke wrote

> I note from your recent note in the Physical Review Letters that we have been inadvertently treading on each others research. For the past couple of months Ken Turner has been working full time on the very problem you discuss.

The problem Dicke mentioned in this letter was a laboratory detection of the gravitational redshift by means of the Mössbauer effect. Pound recalled that Dicke had considered using a silver isotope that Pound and colleagues, with more experience in condensed matter, felt was not likely to produce a usefully strong absorption line, and that Dicke sent his graduate student, Turner, to Pound to learn the technology. Dicke (1963, p. 187) recalled that

> It started out about September 1959 as an attempt to measure the gravitational red shift using the Mössbauer effect which had just been discovered, but it soon became apparent that there were two other groups working on this problem, and to avoid a horse race it was dropped about November in favor of the one to be described.

The experiment "to be described" was by Kenneth Turner (1962, Ph.D. Princeton). Champeney et al. and Turner both used a Co^{57} γ-ray source and Fe^{57} absorber. Turner placed the source near the rim of a standard centrifuge wheel and the absorber and detector near the axis. The conclusions from the two experiments are that our effective velocity in Earth's equatorial plane relative to the effective aether is limited to $\epsilon v = 160 \pm 280$ cm s^{-1} (Champeney et al. 1963) and $\epsilon v = 220 \pm 840$ cm s^{-1} (Turner and Hill 1964). These results in the model in equation (8) are summarized in the seventh row in Table 1.

Dicke's interest in the idea that gravity may be affected by motion relative to a preferred frame led to the experiment he mentioned at the 1957 Chapel Hill Conference (and is discussed in Sect. 5), to check whether the gravitational acceleration g at a fixed position on Earth might vary as the Earth moves around the Sun or as the universe expands. This was the subject of the doctoral dissertation experiments by William Hoffmann (1962, Ph.D. Princeton) and David Curott (1965, Ph.D. Princeton). Hoffmann concluded "that the amplitude of any annual variation of the gravitational constant, is less than 4 parts in 10^8", which is comparable to what may be inferred from other g measurements and "has considerable promise for accurate g measurements". Curott reported "a frequency increase of $1.7 \pm .4$ parts in 10^9 per day", but this tentative indication of a detection did not pass later tests. The fitting function in equation (11), expressed as the fractional variation of G as the Earth moves around the Sun with speed $u = 30$ km s^{-1}, with a conservative constraint from these two experiments, $\delta G/G \lesssim 10^{-7.5}$, limits the parameter ϵ in equation (11) to the value in the eighth row in Table 1.

The intended point of Table 1, which is meant to be a fair sample of what experimentalists were doing and the span of dates of the experiments, is the following.

The search for tests of special relativity, or the idea of some sort of effective aether, or the Machian considerations in Section 4, or the elegance of some other holistic world picture that would relate local to global physics, has been persistently interesting enough to enough people to motivate many experimental explorations. All this work has not revealed any departure from the Strong Equivalence Principle. This consideration continues in Section 7.1, following discussions of other probes of the Strong Principle.

6.3 Time kept by accelerating clocks

The time kept by an accelerated molecule, atom, or atomic nucleus may be affected by mechanical stresses that distort local wave functions and electromagnetic fields, but within broadly accepted ideas these mechanical effects may be computed, or estimated, using standard atomic or condensed matter physics, even when one is considering the possibility that the physical parameters in the computation may depend on position or on velocity relative to some aether. But tests for an intrinsic effect of acceleration on timekeeping must be considered. Rutherford and Compton (1919) briefly reported the test entered in line 2 of the timeline in Figure 9: they found no effect on the rate of decay of radioactive material fixed to the edge of a spinning disk. Ageno and Amaldi (1966) presented an edifying review of improvements of this experiment, including their own version. The timeline marks (in line 20) the great advance in sensitivity afforded by the Mössbauer (1958) effect in the centrifuge experiment by Hay et al. (1960). The setup was similar to the anisotropy tests (Turner 1962; Champeney et al. 1963), except that Hay et al. tested the mean transverse Doppler effect. The measured fractional second-order Doppler shift, $\delta\nu/\nu \sim 10^{-13}$, was found to agree reasonably well with the relativistic $v^2/2c^2$ prediction. This means the fractional shift in the intrinsic atomic nucleus clock rate due to its acceleration must be well below a few parts in 10^{13} at the largest experimental acceleration, 6×10^7 cm s^{-2}. The characteristic relativistic acceleration, $c^2/r \sim 10^{33}$ cm s^{-2}, defined by the radius r of the nucleus is much larger, however.

6.4 Gravitational redshift, relativistic timing, and tired light

Trumpler (1956) mentioned the test of gravitational redshift in two white dwarf stars (in multiple systems, so that the radial velocities of the companions yield the correction for the motions of the systems). He reported that the observed and predicted redshifts of Sirius B are $v_{\rm obs} = 19$ km s^{-1} and $v_{\rm pred} = 20$ km s^{-1}. This was without attribution, but Adams (1925) and Moore (1928) both measured $v_{\rm obs} = 21$ km s^{-1}, close to Trumpler's number. These two measurements were meant to be independent: Adams used the 100-inch reflector at Mount Wilson, Moore the 36-inch refractor at Mount Hamilton, which might be expected to be differently affected by light scattered from the main sequence companion star Sirius A. Both referred to Eddington for the prediction, $v_{\rm pred} = 20$ km s^{-1}, which apparently was satisfactorily close to the two measurements. But Greenstein et al. (1971) argued that the Adams and Moore

> spectra were badly contaminated by Sirius A light, and the results depended on measurements of metallic lines, such as the Mg II line λ4481, which are now known not to occur in white dwarfs. Consequently, these redshifts are of historical interest only.

Greenstein, Oke, and Shipman, in observations when Sirius A had moved further away from the white dwarf Sirius B, reducing the problem with scattered light, found

$v_{\rm obs} = 89 \pm 16$ km s^{-1} and $v_{\rm pred} = 83 \pm 3$ km s^{-1} (the measurements for Sirius B have not changed much since then: Barstow et al. 2005). We see that in the 1950s the two measurements of the redshift of Sirius B were consistent, but both were wrong by a factor of four. They apparently confirmed the relativistic prediction, but it too was wrong, by a like factor. Hetherington (1980) and Greenstein et al. (1985) debated the meaning of this interesting situation.

The second white dwarf star Trumpler mentioned is 40 Eridani B, for which Popper (1954) found $v_{\rm obs} = 21 \pm 4$ km s^{-1} and $v_{\rm pred} = 17 \pm 3$ km s^{-1}. Greenstein and Trimble (1972), making use of the developing art of image intensifiers, found $v_{\rm obs} = 23 \pm 5$ km s^{-1} and $v_{\rm pred} = 20 \pm 9$ km s^{-1} for this white dwarf, consistent with Popper. These numbers agree with the relativistic prediction, and have not changed much since then.

Trumpler (1956) remarked that "For more than 30 years I have been working on a program of measuring the radial velocities (Doppler shifts) of stars in galactic star clusters". This well-experienced observer did not pause to consider whether the observations of the radial velocities of the two white dwarf stars might be questionable. This is not a criticism of Trumpler, but rather a serious cautionary example of the hazards of empirical science. This discussion continues in Section 7.7.

The other early test of the gravitational redshift, St. John's (1928) measurements of the redshift of light from the Sun, was vexed by turbulence and outflows manifest as distortions of absorption line shapes, variations of the measured line shifts from center to limb of the Sun, and systematic variations of the line shifts with the binding energy of the ion, which correlates with the depth of formation of the line in the Solar atmosphere. But the values of the line shifts were roughly in accord with relativity, usually to about 30%, arguably good enough to be entered in Figure 9 (line 6), but not a very convincing detection.

James Brault (1962, Ph.D. Princeton), at Dicke's suggestion, improved the situation. Brault measured the shift of the solar sodium D_1 absorption line at 5896 Å, which is strong enough to allow a tight measurement of the line shape. And the sodium ionization potential is small enough that the line is thought to largely originate above the turbulence in the photosphere but below the chromosphere, in a region where non-gravitational perturbations might be expected to be minimal. Brault used a wavelength modulation technique that he showed stably defined the line center, as follows. The position of the output slit of the spectrometer oscillated at frequency ω, so that the narrow band of wavelengths admitted to the photodetector varied with time as

$$\lambda(t) = \lambda_o + \delta\lambda \sin \omega t. \tag{12}$$

The first term on the right-hand side, λ_o, was adjusted until the photodetector output showed no component at the slit oscillation frequency ω; only the harmonics were detected. The value of λ_o at this point defined a measure of the line center. This strategy will be recognized as similar to the phase-sensitive lock-in amplifier technique used in the Roll et al. (1964) Eötvös experiment reviewed in Section 5. Brault probed the line shape by measuring how λ_o depended on the scan amplitude $\delta\lambda$ in equation (12). Brault demonstrated that, in his chosen line and range of scan amplitudes, λ_o is insensitive to $\delta\lambda$. This is the wanted signature of a satisfactorily symmetric line shape. He also showed that the line center defined this way is insensitive to position on the Sun, scanning from center to limb. These two results make a reasonable case that the measurements were not seriously affected by nongravitational disturbances. Brault's conclusion was that "The ratio of the observed red shift to the theoretical value is found to be 1.05 ± 0.05".

This beautiful experiment was fully published only in Brault's thesis. Dicke (1963, pp. 189–191) outlined the experiment, and I take the liberty of showing in Figure 10

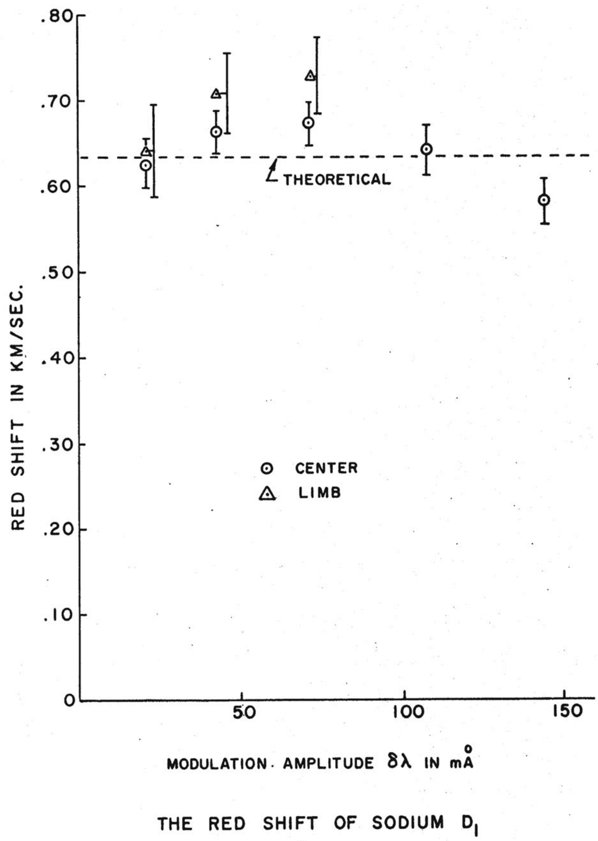


Fig. 10. Brault's (1962) test of the solar gravitational redshift.

Brault's (1962) summary figure. The horizontal axis is the amplitude $\delta\lambda$ of the spectrum scan. The vertical axis shows the measured line shifts at solar center and limb; other figures in the thesis showed samples across the full face of the Sun. The figure shows the insensitivity to scan amplitude and position on the Sun that make the case for a reliable test of the relativistic prediction marked by the dashed line.

Pound and Rebka (1960) used Mössbauer's (1958) effect to obtain a 10% laboratory detection of the gravitational redshift of γ-rays falling 23 m through helium in a tower at Harvard. The precision was comparable to Brault's, but the situation was very different, which is important in the search for systematic errors. Also, the laboratory experiment was under much better control and capable of improvement. Pound and Snider (1964) brought the laboratory precision to 1%.

An even more demanding test of the gravitational redshift and relativistic timekeeping grew out of another great advance in technology, atomic clocks. Møller's (1957) early recognition of their promise is discussed in Section 2. At the Chapel Hill Conference Dicke (1957a) mentioned work at Princeton on rubidium atomic clocks, in collaboration with Carroll Alley (1962, Ph.D. Princeton). After considerably more development (Mark Goldenberg 1960, Ph.D. Harvard), Vessot et al. (1980) reported a test of gravitational timekeeping by the Gravity Probe A rocket flight of an atomic hydrogen maser to 10 000 km altitude. The timing measurements agreed with the prediction of general relativity theory to about a part in 10^4.

Irwin Shapiro's (1964) "fourth test of general relativity", which probes another aspect of relativistic timing, used planetary radar astronomy (as reviewed in Pettengill and Shapiro 1965) to check the relativistic prediction of an increase in return times of radar pulses reflected by Mercury or Venus when the line of sight passes close to the Sun. Shapiro (2015) recalled that

> My "entry" into testing GR was not influenced as far as I could tell by Bob's broad and well broadcast – in scientific meetings and the like – approach to (re)start this experimental field. I didn't attend any of these meetings, nor did I read any of the proceedings. I had the idea to enter it from the prospects for radar astronomy and my knowledge of the Mercury test; my first thought, in the late 1950s, was to check on, and improve, the measurements of the relativistic advances of the perihelia of the inner planets. From there I went on to think of more than a half dozen different tests, almost all not original with me, save for the radar/radio approach. All but two of them were eventually carried out by my colleagues, students, and me, and also by others.

Shapiro et al. (1968) termed their radar timing measurements "preliminary", but they presented a clear detection of the relativistic prediction of the time delay. This experiment was an elegant addition to the tests of general relativity. It is the last entry in the timeline in Figure 9, line 38; it is taken to mark the end of the naissance of experimental gravity physics. More recent measurements of microwave signals transmitted from the ground to the Cassini spacecraft, retransmitted by the spacecraft, and detected on the ground, as the line of sight to the spacecraft passed near the Sun, established the relativistic effect of the Sun on the timing of radiation to a few parts in 10^5 (Bertotti et al. 2003).

Yet another aspect of relativistic timing is the shift to the red in the spectra of distant galaxies of stars. In the 1950s and earlier it was reasonable to ask, with Zwicky (1929), whether the starlight might have been shifted to the red by some physical process operating along the line of sight, rather than by the expansion of the universe. This question helped inspire Kennedy and Thorndike (1931) to check one conceivable physical effect on the frequency of propagating light: they measured the effect on the frequency of the 5641 Å mercury line after moving through 50 000 volts potential difference, "because it has required only a modification of apparatus devised for another purpose", with a null result.

Tolman (1930, Eq. 30) pointed out that Zwicky's "tired light" model for the cosmological redshift is tested by measuring how the surface brightnesses of galaxies vary with the redshift. Under standard local physics, and assuming the light propagates freely through a spacetime described by a metric tensor, Liouville's theorem tells us that the integrated surface brightness (energy received per unit time, area, and solid angle) varies with the ratio of the observed wavelength $\lambda_{\rm obs}$ of a spectral feature to the laboratory wavelength $\lambda_{\rm em}$ at the source as $i \propto (\lambda_{\rm em}/\lambda_{\rm obs})^4$. One of the four powers of $\lambda_{\rm em}/\lambda_{\rm obs}$ may be attributed to the loss of energy of each photon as its frequency decreases, one to time dilation of the rate of detection of photons, and two powers to aberration of the solid angle of a bundle of radiation. In a simple tired light model in a static universe only the first factor would operate: $i \propto (\lambda_{\rm em}/\lambda_{\rm obs})$. The test is important but the precision is limited by the variable properties of galaxies (Geller and Peebles 1972; Sandage 2010). A much tighter test follows from Tolman's demonstration that the four factors serve to preserve the form of a thermal radiation spectrum as a homogeneous universe expands and freely propagating radiation cools. The spectrum of the Cosmic Microwave Background radiation discussed in Section 6.13.2 is quite close to thermal, as shown in Figure 13. Since the universe is observed to be optically thin to radiation at these wavelengths, the tired light model predicts that the CMB spectrum cannot remain thermal as the radiation is redshifted, contrary

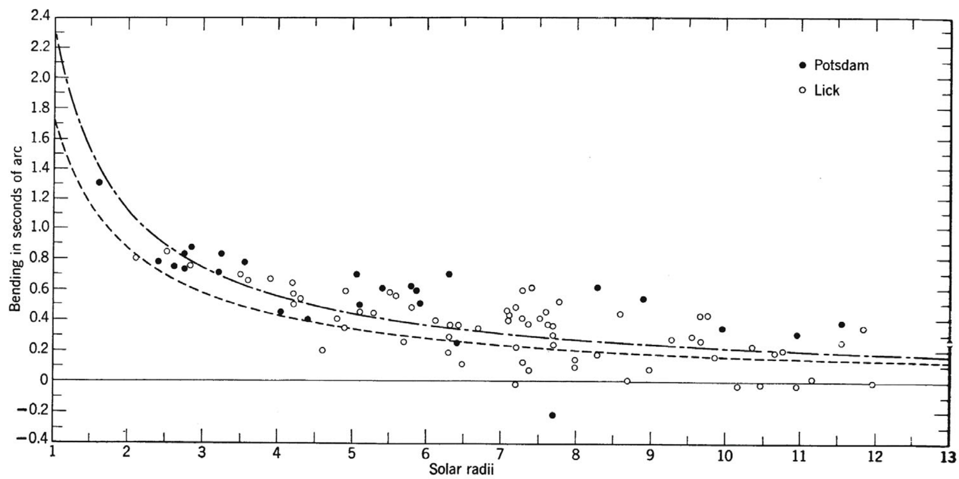


Fig. 11. Radial star displacements measured during the 1922 (open circles) and 1929 (filled circles) solar eclipses, assembled and reviewed by Bertotti et al. (1962). The dashed curve is the relativistic prediction, the dot-dashed curve a fit to the Potsdam data.

to the measurements. The tired light model clearly is wrong. The measured spectrum is consistent with freely propagating radiation in a very close to homogeneous expanding universe with standard local physics.

6.5 Gravitational lensing and deflection of light

Bertotti et al. (1962) presented a careful review of the state of the tests of general relativity. Figure 11 shows their summary of the test of the gravitational deflection of light by the Sun for two sets of observations in favorable star fields. In their assessment, the most important uncertainty was the conversion of distances of stars from the Sun on the plates to angular distances from the Sun in the sky, in the eclipse plates and in the comparison dark sky exposures. The conversions differed because of temperature differences and uncertainties in the positions of plates in the telescope. Bertotti et al. concluded that the observed pattern of radial shifts of angular distances during the eclipse "does not contradict the $1/r$ law predicted by general relativity, but neither does it give much support to such a dependence on distance". Indeed, one sees in Figure 11 that the case for the $1/r$ law rests on the one star closest to the Sun in one of the eclipses.

At the 1955 Bern Conference Trumpler (1956) listed ten measurements of the solar gravitational deflection of light, all with probable errors of about 10%, and concluded that "If one considers the various instruments and methods used and the many observers involved, the conclusion seems justified that the observations on the whole confirm the theory". This of course assumes the $1/r$ law. We may conclude that, at the level of these somewhat tepid assessments of the weight of the evidence, general relativity in 1955 had passed two critical tests, from the orbit of the planet Mercury and the observations of deflection of starlight by the Sun. As discussed in Section 6.4, the third classical test, gravitational redshift, was more doubtful.

Shapiro (1967) pointed out that radio interferometer observations of radio-loud quasars could detect the predicted gravitational deflection. The first results (Seielstad et al. 1970; Muhleman et al. 1970) were about as precise as the optical observations but capable of considerable improvement. Indeed, the solar gravitational deflection is now detected in directions over a large part of the celestial sphere (Shapiro et al. 2004). Will (2015) reviewed the history of this test.

Gravitational deflection causes a sufficiently compact mass to act as a gravitational lens. Renn et al. (1997) reviewed the early history of papers on this effect. Notable among them was Einstein's (1936) publication of his earlier thoughts about the double images and increase of apparent magnitude (intensity enhancement when lensing increases the solid angle of the image while preserving the surface brightness) to be expected when one star passes nearly in front of another. Renn et al. reported that Einstein was not optimistic about the possibility of an observation of these lensing effects, but the intensity enhancement produced by lensing by stars and planets is now well observed, and termed microlensing. Zwicky (1937) assessed the prospects for detecting Einstein's (1936) effects when the gravitational lens is a galaxy, with emphasis on the observationally important effect of intensity enhancement. Zwicky's vision of gravitational lensing by the mass observed to be concentrated around galaxies is now a valuable probe of the mass distribution on larger scales, in the dark matter outside the concentration of stars in galaxies. Zwicky's informed discussion is marked as line 11 in Figure 9.

Klimov (1963), Liebes (1964), and Refsdal (1964) independently presented more detailed analyses of the prospects of observing effects of gravitational lensing. Klimov discussed lensing of a galaxy by the mass concentration in a foreground galaxy close to the line of sight, and took note of the Einstein ring produced at close alignment if the lens is close enough to axially symmetric. Liebes and Refsdal mainly discussed lensing by "point-like" mass concentrations such as stars, and they emphasized the phenomenon of intensity enhancement that has proved to be so observationally important. Liebes considered the lensing signature of planets around stars, and the possibility of observing lensing of more remote stars in our galaxy by intervening visible stars or dark or faint objects that might contribute to the mass of the Milky Way. This is the line of ideas taken up by the MACHO search for massive dark objects (e.g. Alcock et al. 2000). Liebes cut a plastic lens that simulated the properties of a gravitational lens (Liebes 1969). A disk source viewed through the lens exhibited dual and ring images; a target placed in the lensed beam of a bright source exhibited the intensity amplification. Refsdal anticipated lensing measures of Hubble's constant H_o and lensing measures of galaxy masses.

Walsh et al. (1979) were the first to identify an observation of lensing: a double image of a quasar. Wheeler pointed out in remarks at the 1959 Royaumont Conference (Lichnerowicz and Tonnelat 1962, pp. 269–271) that lensing produces an odd number of images, but the odd image in the Walsh et al. observation is demagnified, perhaps obscured by dust in the lensing galaxy, and certainly hard to observe. Gravitational lensing now probes the mass distributions in clusters of galaxies (Hoekstra et al. 2013) and the mean mass distribution around galaxies (Bahcall and Kulier 2014).

6.6 Gravitational waves

Joseph Weber's interest in gravitational waves traces back at least to the analysis by Weber and Wheeler (1957) of the *Reality of the Cylindrical Gravitational Waves of Einstein and Rosen* (in a physical system that is translationally invariant in one direction, but not axisymmetric). Although the paper by Einstein and Rosen (1937) presented an argument for the reality of gravitational waves that carry energy, in analogy to electromagnetic waves, the idea was still controversial. One sees this in Einstein's note at the end of the paper (after a skeptical reception of an earlier version by *The Physical Review*), that he had corrected the conclusion "after the departure of Rosen". In his contribution to the proceedings of the 1955 Bern Conference Rosen concluded by endorsing the "conjecture ... that a physical system cannot radiate gravitational energy". The point that gravitational waves certainly carry energy, which can be deposited as work by a gravitational wave acting on a viscous body, was made by Weber and Wheeler (1957), by Feynman (in DeWitt 1957, p. 143) at the Chapel Hill Conference, and by Bondi (1962), at the Royaumont Conference. Bondi also argued that gravitational waves can be produced by a nongravitational explosion, but that, for a purely gravitational binary star system, "I am somewhat doubtful whether such a system will radiate at all." (I suspect that this was because Bondi did not consider the effect of radiation reaction on the equation of motion of the point-like stars.) Despite this quite confusing theoretical situation, Weber (1962) presented to the Royaumont conference his practical examination of how to build a gravitational wave detector. This was one of the earliest steps toward the development of experimental gravity physics (and accordingly is entered in line 18 in Fig. 9).

Nancy Roman, who was the first Chief of the Astronomy, Solar Physics, and Geophysics Programs at the NASA Office of Space Sciences, saw the growing possibility and interest in better probes of relativity afforded by space science, and organized the July 1961 NASA *Conference on Experimental Tests of Theories of Relativity.* The proceedings (Roman 1961) record Weber's report of progress in building a gravitational wave detector (Weber 1960), and his discussion of the measurements of modes of acoustic oscillation of the Earth, including the interesting quadrupole modes that would be excited by the long wavelength gravitational waves of general relativity (Forward et al. 1961; marked as line 23 in Fig. 9). Weber discussed the possibility of placing gravimeters on the Moon, which he expected would be a quieter place to look for the excitation of quadrupole modes of acoustic oscillation by gravitational waves. A decade later the Apollo 17 astronauts placed a Lacoste-type spring gravimeter on the Moon (Giganti et al. 1973). An unfortunate 2% miscalculation of the masses needed to trim the balance to the gravitational acceleration at the position where the gravimeter was placed on the Moon prevented operation at design sensitivity. Interest in this approach continues: Lopes and Silk (2014) analyzed the possibility of detecting gravitational wave excitation of quadrupole acoustic oscillations of the Sun.

Weber (1969, 1970) reported evidence that his bar detectors had found gravitational waves, based on coincident detection of events: unusual departures from the mean noise fluctuations in bar detectors separated by 1000 km, at dimensionless strain estimated to be $h \sim 10^{-16}$. This attracted considerable interest from theorists and experimentalists. The complaint that the indicated gravitational wave strain corresponds to an unreasonably large energy density certainly was worth noting, but the central issue was whether the events were real, perhaps signaling the effect of gravitational waves in some better theory, or perhaps signatures of some other new phenomenon. Tyson and Giffard (1978) reviewed several independent experiments that did not confirm Weber's event rates. But the neutrinos detected from supernova 1987A showed an interesting correlation with Weber Bar events in detectors in Maryland and Rome (Aglietta et al. 1989). It is difficult to find a community consensus of what this might mean. In a review of Weber Bar detectors, Aguiar (2011) wrote: "Did the bars detect gravitons from SN1987A or some other particles that excited the bars by thermoelastic processes? In any case, we hope that another supernova will solve this problem".

In the scalar-tensor gravity theory there could be observable effects of temporal or spatial variations of the scalar field value that determines the local strength of gravity. Morgan et al. (1961) searched for an annual periodicity of earthquakes that might have been triggered by an annual variation in the scalar field value as the Earth moves around the Sun. Their conclusion was that "The occurrence of this periodicity would be understandable if the gravitational constant were to vary as the earth-sun distance changes or as Earth's velocity relative to a preferred coordinate frame changes; however, the observed periodicity cannot be interpreted as conclusive support for such a hypothesis". A scalar wave could excite the Earth's 20-minute breathing mode (the high-Q nearly spherically symmetric mode with no nodes in the radial function). Jason Morgan (1964, Ph.D. Princeton) looked for geophysical, lunar and planetary indications of this effect, without significant detection.

Dicke's interest in the possibility of a laboratory detection of the low-frequency modes of oscillation of the Earth that might be driven by long wavelength tensor or scalar gravitational waves (as discussed by Forward et al. 1961) led to projects in his group to build or modify LaCoste-type (spring) gravimeters, with sensitivity increased by making use of the precision detection of the test mass deflection afforded by the phase-sensitive lock-in amplifier techniques discussed in Section 5 (Robert Moore 1966, Ph.D. Princeton; Weiss and Block 1965; Block and Moore 1966). This work contributed to the creation of networks for low-frequency seismology that produced detailed detections of Earth's low-frequency modes of oscillation. Jonathan Berger (2016) recalled that, when he was a graduate student,

> Barry Block came to IGPP (Walter Munk's institution) at UCSD in 1965 (or 1966) followed shortly thereafter by Bob Moore. Bob brought with him 2 (I think) modified LaCoste gravimeters which he had developed for the thesis work at Princeton and afterwords at U. Maryland. At the same time, Freeman Gilbert and George Backus (also at IGPP) were developing the theory and mathematical methods for inverting normal mode data to resolve details of Earth structure. These instruments soon produced some spectacular observations of the Earth's normal modes from relatively frequent earthquakes that whetted appetites for more such data.

This grew into the project, International Deployment of Accelerometers, with Berger as director, that is now part of the Global Seismographic Network. It yields probes of the internal structure of the Earth and measures of earthquakes, storm surges, tsunamis, and underground explosions[9].

Let us pause to consider how the analysis by Forward et al. (1961) of the possible effect of low frequency gravitational waves on Earth's modes of oscillation helped interest the experimental gravity community in gravimeters, which aided the development of global seismology, which detected low frequency Earth oscillations, which Boughn et al. (1990) turned into further exploration of the bounds on the energy density in long wavelength gravitational waves. And let us consider also that the search for detection of gravitational waves may be dated to have begun with Weber's paper in the proceedings of the 1959 Royaumont Conference. A quarter of a century later precision timing showed that the Hulse-Taylor binary pulsar system is losing energy at the rate expected from radiation of gravitational waves (Taylor and Weisberg 1982). A quarter of a century after that LIGO detected gravitational waves (LIGO Scientific Collaboration and Virgo Collaboration 2016), completing Weber's vision.

My impression is that the very qualities that Weber needed to pioneer this difficult science made it exceedingly difficult for him to deal with critical reactions to his early results. Weber in the early 1960s impressed me for his great determination, his indifference to experts who were not sure these waves even exist, his love of the chase for a wonderful phenomenon, and his energetic accounts in seminars at Princeton on how he was building and instrumenting his detectors. To be noted also is his checks of significance of signals by coincidences in detectors first separated by a few kilometers, then 1000 km. The near coincidence of events in the well-separated LIGO interferometers was a key element in the first convincing gravitational wave detection (LIGO et al. 2016).

---

[9] As discussed in the web sites for Project IDA, at http://ida.ucsd.edu/, and the Global Seismographic Network, at https://www.iris.edu/hq/programs/gsn.

6.7 Masses: active, passive, inertial, and annihilation

The equations

$$F = m_i a, \qquad F = m_i g = \frac{GM_a m_p}{r^2}, \qquad E = m_e c^2, \tag{13}$$

define four masses to be assigned to an object. In the first equation F is the force, mechanical or gravitational, on an object with inertial mass m_i moving with acceleration a. In the second equation m_p is the passive gravitational mass of the object that is moving with acceleration g due to the gravitational force of attraction by a second body with active gravitational mass M_a at distance r. In the last equation m_e is the mass defined by the annihilation energy E. These definitions follow Bondi (1957), who was largely concerned with the possibility of negative mass (at the time some wondered whether antimatter falls up, as in Schiff 1958). To be discussed here is the empirical situation in nonrelativistic physics, where in the standard model the four masses are equal (the situation is more complicated in relativistic situations. In general relativity an ideal fluid with mass density ρ and pressure p, with $c = 1$, has active gravitational mass density $\rho + 3p$, inertial and passive gravitational mass densities $\rho + p$, and energy density ρ).

The Eötvös experiment discussed in Section 5 tests whether the ratios m_i/m_p of inertial to passive gravitational masses, and hence gravitational accelerations, may depend on the natures of compact nearly massless test particles. Eötvös et al. (1922) bounded differences of gravitational accelerations among a considerable variety of materials to about 1 part in 10^8 (line 4 in Fig. 9). The static experiment in Dicke's group, with the synchronous detection scheme illustrated in Figure 7 at the heart of the instrument, found that the gravitational accelerations of gold and aluminum toward the Sun differ by no more than 3 parts in 10^{11} (Roll et al. 1964; line 31 in Fig. 9). Braginskiĭ and Panov (1971) found that the fractional difference of gravitational accelerations of aluminum and platinum is less than about 1 part in 10^{12}. The present bound reaches parts in 10^{13} on fractional differences of gravitational accelerations toward the Earth and Sun (Adelberger et al. 2009).

A difference in the ratio m_a/m_p of active to passive gravitational masses for different materials is in principle detectable as a difference in the measured values of Newton's constant G in Cavendish balance experiments using different materials, but precision measurements of G are difficult. In Kreuzer's (1968) survey of the literature measurements of G, using materials ranging from glass to mercury, scatter by a few parts in 10^3, indicating a similar bound on fractional differences of m_a/m_p among the elements. To improve the constraint Lloyd Kreuzer (1966, Ph.D. Princeton; Kreuzer 1968; line 34 in Fig. 9), floated a solid body in a fluid at near neutral buoyancy, meaning the passive masses of the body and the fluid it displaced were nearly the same. If the active masses of the body and the displaced fluid were sensibly different it would be detected as a change in the gravitational attraction of a Cavendish-type balance as the body was moved through the fluid toward and away from the balance. The inevitable departure from exact neutral buoyancy was measured by the force needed to support the body in the fluid as the fluid mass density was adjusted, with the balance response extrapolated to zero support. Kreuzer concluded that the ratios of active to passive gravitational masses of bromine (which dominated the mass of the fluid) and fluorine (which dominated in the solid) differ by less than about 5 parts in 10^5 (the reanalysis of Kreuzer's data by Morrison and Hill 1973 confirms Kreuzer's conclusion).

If active and passive gravitational masses were not equivalent then, as Kreuzer (1966) discussed, a dumbbell consisting of two different massive bodies held at fixed separation by a rod would accelerate in the absence of any external force. This is a situation we are conditioned to reject but must consider, on the principle that unexamined assumptions in science ought to be challenged. Following Kreuzer's experiment, Bartlett and Van Buren (1986) pointed out that, if the composition of the Moon were not spherically symmetric, as suggested by a difference between the centers of figure and mass of the Moon, then the departure from equivalence of active and passive mass would contribute to the acceleration of the Moon. Lunar Laser Ranging measurements place a tight bound on this effect (Sect. 6.10.2).

As a graduate student, at Dicke's invitation, I pored over measurements of binding energies of atomic nuclei and nuclear spectroscopy mass measurements, to test the equivalence of mass and binding energy, but without publishing. Much more recently Rainville et al. (2005) and Jentschel et al. (2009) examined budgets of energy and mass in radiative nuclear transitions. The γ-ray wavelength was measured by crystal Bragg spectroscopy, and converted to annihilation mass by standard values of $\hbar$ and c. The inertial mass difference was measured by "cyclotron frequencies (inversely proportional to the mass) of ions of the initial and final isotopes confined over a period of weeks in a Penning trap" (from Rainville et al. 2005). The conclusion from these impressively precise measurements is that the inertial and annihilation energy masses agree to a few parts in 10^7.

6.8 Gravitational frame-dragging

Michelson and Gale's (1925) paper on *The Effect of the Earth's Rotation on the Velocity of Light* is notable for the heroic size of the Sagnac interferometer, 2010 $\times$ 1113 feet (640 by 320 meters). Earth's rotation was detected, but that is a long way from detecting the relativistic prediction of frame-dragging (Eq. (1)). Pugh (1959), and Schiff (1960) in collaboration with experimentalists at Stanford, realized that advances in artificial satellite technology offered a realistic possibility of detection of frame-dragging. The demanding condition for detection is that the drag on the satellite by light pressure and the last traces of the atmosphere be canceled well enough to suppress torques on the gyroscopes. This was discussed at length at the NASA Conference (Roman 1961). This prompted marking the origin of Gravity Probe B in Figure 9 (line 22) at 1961. The successful completion a half century later, in a satellite in polar orbit, detected the predicted geodetic precession in the plane of the orbit, and it detected the inertial frame-dragging normal to the orbit to be expected from the rotation of the Earth. The frame-dragging was measured to be 37.2 ± 7.2 milli arc seconds per year, consistent with the relativistic prediction (Everitt 2015 and the following 20 papers on Gravity Probe B). Ciufolini and Pavlis (2004) found a roughly matching result by analysis of the precision tracking of two high-density LAGEOS Earth satellites, which do not have provision for cancellation of atmospheric drag, but are dense enough to allow a reasonably secure correction.

6.9 Absolute measurements of gravitational acceleration and the gravitational constant

Precision measurements of the absolute value of the gravitational acceleration g have practical applications, as in monitoring ground water level changes, and some that are purely curiosity-driven, as in probes of the Machian ideas reviewed in Section 4. In the 1950s measurements of the absolute value of g (expressed in established standards of length and time, with surveys that transferred the value of g from one place to another) used carefully designed pendulums or freely falling objects timed by photoelectric detection of intersections of light beams. In his Ph.D. thesis Faller (1963,

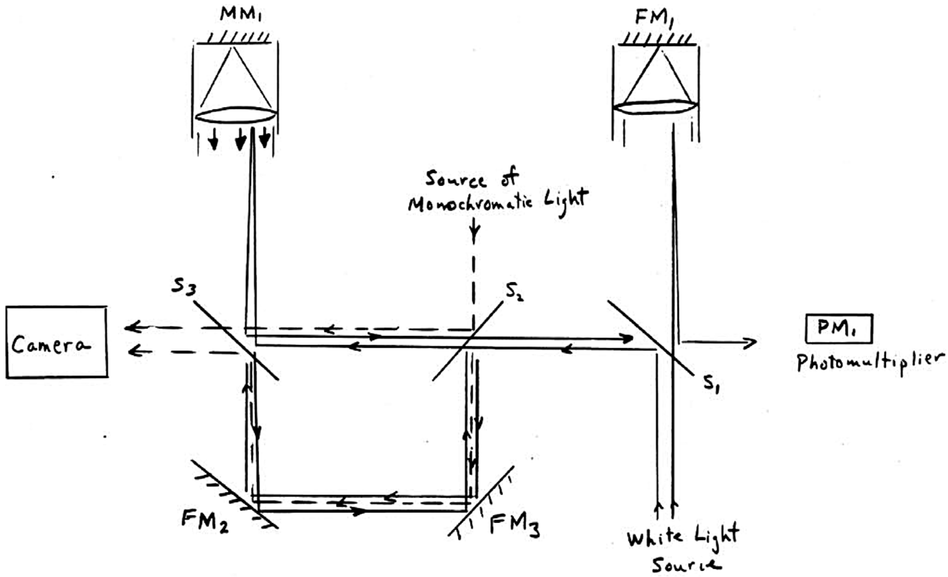


Fig. 12. A preliminary sketch of Faller's thesis measurement of the absolute value of the gravitational acceleration, from Jim Wittke's Princeton Graduate Alumni Newsletter, 1962.

Ph.D. Princeton) described origins of his interferometer measurement:

> About six years ago, it was suggested that an interferometric method might also provide a better approach to the problem of an absolute "g" determination[12]. In particular, the falling object might be one plate of an interferometer. The work, probably due to the advent of inertial navigation, was never brought to a conclusion as the particular interest at that time was concerned with the possibility of navigating submarines gravitationally. It is, however, the suggestion of employing an interferometer in order to make an absolute "g" determination that has been taken up and made use of in the experiment described here.

Faller's reference 12 in this quote reads "J.G. King and J.R. Zacharias of M.I.T". King[10] recalled that

> We wanted to take two mirrors and make a Fabry-Perot interferometer out of them, and this is pre-laser, so you do it with collimated light, filtered. And now the upper plate is held up by electrostatic field, so you shut the field off and it drops so suddenly that it can't wiggle sideways, and it falls down, and the interference fringes go swittt, and you measure that, and that gives you g. Interestingly, this was going on I think at Hycon Eastern, a small local company.

In his thesis Faller emphasized the great problem with this plan: the interference pattern in a Fabry-Pérot interferometer is exceedingly sensitive to the relative orientation of the plates, and if one of the plates is falling its orientation is exceedingly difficult to control. The solution in Faller's thesis and the many later versions of his experiment was to replace the plates by corner reflectors.

The arrangement in Faller's experiment is sketched in Figure 12 (in the experiment marked as line 28 in Fig. 9). Light from the source at the lower right partially passed

---

[10] Interview of John G. King by George O. Zimmerman on 2009 November 18, Niels Bohr Library & Archives, American Institute of Physics, College Park, MD USA

through the half-silvered mirror S_1, returned from the fixed corner reflector FM1 on the upper right, and was reflected by S_1 to the photomultiplier. Part of the light from the source was reflected by S_1 and S_2, returned by the freely falling corner reflector MM1 on the upper left, and arrived at the photomultiplier to be combined with the light from the other path. The fixed mirrors FM2 and FM3 served as calibration. The bandwidth of the light source was broad enough that the position of the freely falling corner reflector at a fringe, relatively bright or dark, was defined by the condition that at the photomultiplier the phase as a function of wavelength at fixed falling corner reflector position was at an extremum. The precision this allowed in Faller's thesis measurement of g was comparable to what was achieved by other methods, but could be improved. Faller and Hammond (1967) showed that the newly available stabilized lasers with sharply defined wavelength allowed sharply defined fringes and a much better measurement. The present standard (advertised in http://www.microglacoste.com/fg5x.php), with applications such as long-term monitoring of water table levels and continental uplift following the last Ice Age, in Faller's (2015) words "pretty much is the legacy of what, with Bob's help, I started in Palmer Lab".

One would also like to have an accurate absolute measurement of Newton's gravitational constant, G, the better to weigh the Earth and the rest of the solar system, as well as to establish a tight system of fundamental physical parameters. An indication of the difficulty is that Heyl (1930) measured $G = 6.670 \times 10^{-8}$ cm^3 g^{-1} s^{-2}, with a scatter among gold, platinum, and glass weights at about a part in 10^4. That is to be compared to the present CODATA standard, $G = 6.6741 \times 10^{-8}$ cm^3 g^{-1} s^{-2}, with fractional uncertainty reduced from Heyl by about a factor of two. Faller (2014a) reviewed this situation.

6.10 Evolution of Newton's gravitational constant

Section 4 reviewed Dirac's Large Numbers Hypothesis (LNH), that the strength of the gravitational interaction (as measured by the value of G in standard units) may be very weak because it has been decreasing for a very long time. The idea has been probed in several ways.

6.10.1 Meteorites, eclipses, and geophysics

Teller (1948) pointed out that a larger G in the past would imply a hotter Sun, which if too hot would violate evidence of early life on Earth. Teller made an important point (which is marked as line 14 in Fig. 9). But the Hubble age then was underestimated by a factor of seven, meaning the constraint on the evolution of G is significantly less tight than Teller estimated. Dicke (1962b) added the consideration that Earth's climate is seriously affected by the greenhouse effect of water vapor and the opposite effect of clouds. He argued that the net effect is quite uncertain, but that this consideration further weakens Teller's constraint. Dicke (1962b) also noted that stellar evolution models with standard gravity physics predict that the luminosity of the Sun has been increasing: The conversion from hydrogen to helium that keeps the Sun shining has been increasing the mean molecular weight in the core, which increases the temperature required to support it, which increases the luminosity. Dicke's colleague at Princeton University, Martin Schwarzschild, estimated that the present luminosity of the Sun, relative to the luminosity minimum at formation 4.5 Gyr ago, is $L_{\mathrm{present}}/L_{\mathrm{initial}} = 1.6$. This estimate is in Schwarzschild 1958, equations (8) and (23). A more recent estimate in Ribas (2010) is $L_{\mathrm{present}}/L_{\mathrm{initial}} = 1.3$. A larger G in the past might have helped keep Earth warm enough for early life. But, as Dicke argued, the relation between the history of the surface temperature of the Earth and the luminosity of the Sun was quite uncertain. The problem still with us.

A quantitative bound on the past luminosity of the Sun (in Peebles and Dicke 1962b; the title, *The temperature of meteorites and Dirac's cosmology and Mach's Principle*, makes manifest Dicke's motivation) considered that if G were larger in the past it would mean that parent meteorite bodies would have been warmer, the diffusive loss of radiogenic argon faster, and the apparent potassium-argon radioactive decay ages of meteorites smaller than their true ages. We were following Goles et al. (1960), who used this consideration to constrain the environments of parent meteorite bodies under the assumption of conventional physics. We concluded that the strength of the gravitational interaction could not have been decreasing more rapidly than

$$-\frac{1}{G}\frac{dG}{dt} \lesssim 1 \times 10^{-10}\ \mathrm{yr}^{-1}. \tag{14}$$

And we assured the reader that "This limit does not seem to rule out any of the cosmologies in which the strength of the gravitational interaction is variable". That included the estimate in equation (3) based on the LNH. This result is marked at line 26 in Figure 9.

Ideas about how the effect of a decreasing value of G might be manifest in geophysical phenomena were mentioned in Section 4. In particular, a secular change of G would affect the Earth's moment of inertia, its distribution of angular momentum in the solid Earth and in atmospheric and oceanic currents, and its orbit around the Sun, all of which could have a detectable effect on timings of solar eclipses, which have been recorded back to the time of Babylon some 3000 yr ago. Curott (1966) examined the eclipse evidence, and Dicke (1966) analyzed the many geophysical considerations that may help determine the effect of a secular evolution of G on the evolution of the angular velocity of the Earth's surface and hence the predicted eclipse timings. Dicke concluded that the evidence bounds the evolution of G at about the level of equation (14). This was by an approach that is so different that it merits a separate entry as line 35 in Figure 9.

6.10.2 Lunar Laser Ranging

At the Chapel Hill Conference, Dicke (1957a) expressed his interest in better measurements of the motion of the Moon. Though not explained in the proceedings, he was at the time seeking to test whether the strength of gravity has been decreasing as the universe expands (Sect. 4.2 Eq. (3)). This is the stated goal in the papers on tracking artificial satellites or the Moon by Hoffmann et al. (1960) and Dicke et al. (1961) (with, in the first of these papers, acknowledgement of contributions by all 13 members of the Gravity Research Group at that time, some much more important for the project than others). These two papers discussed the prospects for precision optical tracking of the angular position of a satellite by a "searchlight illuminating a corner reflector", or else "sunlight illuminating a sphere", or a "flashing light on the satellite". The proposal had become more specific in the paper by Alley et al. (1965) with the title, and the argument for, *Optical Radar Using a Corner Reflector on the Moon*. The author list included Dicke, with Alley, Bender, and Faller, who were his former graduate students, Wilkinson, who was then an assistant professor in Dicke's group, and two who were not in the group: Henry Plotkin at NASA Goddard Space Flight Center and Peter Franken at the University of Michigan, Ann Arbor. The paper mentioned the possibility of a "few interesting measurements", including "an accurate check on lunar orbit theory", "Simultaneous ranging by several stations, with known geographical positions, could be used to measure the size and shape of the earth", a "connection could be established between the American and European geodesic networks", and "disturbances or periods which are not explained by the perturbations contained in the lunar theory ... if seen, might be identified with gravitational wave effects".

Recollections of how this line of thought grew into the Lunar Laser Ranging Experiment are in the proceedings of Session 2 of the 2014 19th International Workshop on Laser Ranging, Annapolis. In these proceedings Plotkin (2014) recalled that in 1960 he moved to the Goddard Spaceflight Center, where his "job was supposed to be developing optical systems for photographing satellites and stars in order to calibrate Microwave and Radio tracking systems", and that he stopped at Princeton along the way to discuss the task with Dicke, Carroll Alley, and others. Dicke "suggested that when I get to NASA I consider putting cube corners on the moon or launching cube corners into space orbits and photographing reflections from powerful searchlight illumination". In an undated document[11] Plotkin reported that by 1964 "NASA launched the first of the satellites with arrays of fused-quartz cube corner retroreflectors to act as cooperative targets for laser radar stations" which "achieved ranging precision of about 1 meter". This is the methodology for the Lunar Laser Ranging Experiment (and for the precision tracking of the high mass density LAGEOS satellites used to map Earth's gravitational field, and, as a byproduct, detect relativistic inertial frame dragging; Sect. 6.8). Also in the Annapolis proceedings are Faller's (2014b) recollections of the corner reflectors he used in his absolute measurement of the acceleration of gravity (Sect. 6.9), his work on the design of the arrays of corner reflectors to be placed on the Moon, and the serious challenge he and colleagues faced in finding the first photons to be received by reflection from the corner reflectors, made even more difficult because the corner reflector arrays were not placed at the planned position on the Moon. Faller (2014b) also presented Dicke's previously unpublished essay on *Lunar Laser Ranging Reminiscences*. Alley (1972) published his recollections of the project. Bender (2015) recalled that

> in Sept., 1968 when we first heard that the Apollo 11 astronauts would only be allowed to spend 2 hours on the lunar surface, compared with the 4 hours originally planned, Bob Dicke being the one who asked if a reflector package could be considered as a contingency experiment for the Apollo 11 landing may well have carried more weight with the NASA people than if he hadn't been involved.

The NASA Apollo 11 astronauts placed this first package of corner reflectors on the Moon in July 1969. The USSR Lunokhod 1 unmanned lunar rover left the second package on the Moon after launch in November 1970.

The origin of the Lunar Laser Ranging Experiment is marked at line 32 in the timeline in Figure 9, at publication of the paper by Alley et al. (1965) that argued for this program. Faller et al. (1969) reported that measurements of the distance to the corner reflectors "can now be made with an accuracy approaching 15 cm". Distance measurements to the lunar corner reflectors now approach a precision of 1 mm, and they tightly constrain possible departures from general relativity theory, including the Strong Equivalence Principle mentioned in Section 4.3. The inferred relative gravitational accelerations in the Sun-Moon-Earth system bound the fractional differences of the ratios of inertial to passive gravitational masses of the Earth and Moon to no more than about a part in 10^{13} (Williams et al. 2012). It is particularly relevant for the thread of this history that the evolution of the strength of gravity is bounded to

$$\frac{1}{G}\left|\frac{dG}{dt}\right| \lesssim 10^{-12}\ \mathrm{yr}^{-1}. \tag{15}$$

[11] http://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/19680025204.pdf.

Dicke's dream of tests of gravity physics from precision measurements of the motion of the Moon, which he first mentioned in print in 1957, has been abundantly realized. But Hoffmann et al. (1960) had proposed tracking satellites to look for evolution of G "that might be expected to amount to about one part in 10^{10} per year". This is two orders of magnitude larger than the bound.

6.10.3 Other constraints

The strength of the gravitational interaction in the early universe is constrained by the pattern of variation of the Cosmic Microwave Background radiation (the CMB) across the sky (Sect. 6.13.2). The pattern is largely determined by the evolving dynamical interaction of the spatial distributions of matter and radiation through to redshift $z \sim 1000$, when in the standard model the primeval plasma combined to largely neutral atomic hydrogen, eliminating the non-gravitational interaction between matter and radiation. Bai et al. (2015) found that the standard six-parameter ΛCDM cosmology (defined in Sect. 6.13.3), with G as a seventh free parameter, fitted to the Planck CMB satellite data (as in Planck Collaboration 2015a), with other cosmological tests, constrains the value of G (in standard units for the rest of physics) at redshift $z \sim 1000$ to be within about 8% of the laboratory value. The precision is modest but the reach is impressive.

If, in still earlier stages of expansion of the universe, the strength of gravity had been different from now, making the rate of expansion and cooling different from the prediction in general relativity, it would have affected the formation of the light elements at redshift $z \sim 10^9$ (Sect. 6.13.2). The effect of a change of expansion rate was illustrated in Figure 2 in Peebles (1966). Dicke (1968) explored this effect in the scalar-tensor gravity theory. But the light element abundance measurements have been found to be in line with the constant strength of gravity in the ΛCDM cosmology.

Also to be noted is the impressively tight and direct bound on the present rate of change of the strength of gravity from timing measurements of the orbit of a binary pulsar (Zhu et al. 2015), at precision comparable to equation (15). This measurement and the many other constraints reviewed here on the possible evolution of the strength of gravity conflict with an elegant idea, Dicke's readings of Mach's Principle and Dirac's Large Numbers (Sect. 4). The discussion of how empirical evidence can, on occasion, conflict with elegant ideas, continues in Sections 6.12 and 7.1.

6.11 Evolution of the weak and electromagnetic interactions

Dicke (1957a,b) proposed extending Dirac's LNH, that the strength of gravity is very small because it has been decreasing for a very long time, to a decreasing strength of the weak interaction, but more slowly because the weak interaction is not as weak as gravity. This would mean β-decay rates are decreasing. Dicke (1959a) tested for this by comparing radioactive decay ages of meteorites from α- and β-transitions, with inconclusive results.

The electromagnetic interaction is even less weak, but Dicke (1959b) pointed to the long-standing idea that the Planck length $\sqrt{\hbar G/c^3}$ may play some role in determining an effective momentum cutoff for quantum field theory (e.g. Landau 1955; Deser 1957; Arnowitt et al. 1960). If G were decreasing, and the Planck length decreasing with it, maybe the fine-structure constant $\alpha = e^2/\hbar c$ that measures the strength of the electromagnetic interaction is decreasing, perhaps as the absolute value of the logarithm of the Planck length (Landau 1955). Dicke set me the dissertation topic of an evolving electromagnetic interaction (James Peebles 1961, Ph.D. Princeton; Peebles and Dicke 1962c). A dynamical value of α implies a long-range fifth force of interaction that could violate the constraint from the Eötvös experiment. The thesis finessed that by introducing two metric tensors as well as the scalar field to replace α (since developed in much more detail in Lightman and Lee 1973). I found empirical limits on the rate of evolution of α from estimates of how a changing fine-structure constant would affect relative rates of radioactive decay by α-particle emission, nuclear fission, positron emission, and electron capture and emission. Estimates of the consistency of published radioactive decay ages of meteorites and terrestrial samples led to the conclusion that "the data could not be used to eliminate a change in α of 0.1% in the past 4.4 billion years" (this is entered in line 25 in Fig. 9).

Recent bounds are better. Examples are

$$\begin{aligned}
&\frac{1}{\alpha}\left|\frac{d\alpha}{dt}\right| \lesssim 10^{-16}\ \mathrm{yr}^{-1} \text{ at } z = 0,\\
&\frac{1}{t}\frac{|\Delta\alpha|}{\alpha} \lesssim 10^{-16}\ \mathrm{yr}^{-1} \text{ at } z = 0.14,\\
&\frac{1}{t}\frac{|\Delta\alpha|}{\alpha} \lesssim 10^{-14}\ \mathrm{yr}^{-1} \text{ at } 0.2 \lesssim z \lesssim 1,\\
&\frac{1}{t}\frac{|\Delta\alpha|}{\alpha} \lesssim 10^{-12.5}\ \mathrm{yr}^{-1} \text{ at } z = 1200.
\end{aligned} \tag{16}$$

In the second through fourth lines $\Delta\alpha/\alpha$ is the allowed fractional change in the value of the fine-structure constant and t is the time to the present from the time at the redshift z of observation. The first line is a "preliminary constraint" from comparisons of optical atomic clock rates of aluminum and mercury ions (Rosenband et al. 2008). The second line is from bounds on shifts of the resonant energies for slow neutron capture by fission products in the Oklo natural nuclear reactor. The geophysical considerations are complicated, but the many reanalyses (Shlyakhter 1976; Damour and Dyson 1996) argue for reliability of the bound. The third line is an example of constraints from line spacings in quasar spectra (Albareti et al. 2015). The last line is the constraint from the pattern of variation of the Cosmic Microwave Background radiation across the sky, fitted to the ΛCDM cosmological model (Planck Collaboration 2015a; Sect. 6.13.3). For another review of the history of ideas about the value and possible evolution of the fine-structure constant see Kragh (2003).

The length of the list in equation (16) is another illustration of the persistent fascination with the idea that local physics may be influenced by the global nature of the universe in which local physics operates. Motivations for this idea have evolved since the Machian arguments that so fascinated Dicke and others, but interest in the holistic concept continues.

6.12 Tests of the scalar-tensor theory

Brans and Dicke (1961) found that the most demanding constraint on the scalar-tensor gravity theory was the measured precession of the orbit of the planet Mercury. They concluded that this measurement required that the parameter w in the theory (Eq. (4)) satisfy $w > 6$. This of course assumes a correct accounting of the masses that determine the orbit. But Dicke (1964) reminded the reader that, before relativity, the excess precession in Newtonian gravity theory was imagined to be caused by mass not taken into account in the standard computations, and that

> One old suggestion seems not to have had the attention that it deserved. If the Sun were very slightly oblate, the implied distortion of the sun's gravitational

> field would result in a rotation of the perihelia of the planets. To produce a discrepancy in Mercury's orbit as great as 8 per cent of the Einstein value would require an excessively small visual oblateness, only 5 parts in 10^5 amounting to a difference between the solar equatorial and polar radii of only 0.05" arc.

One might imagine that the Sun is slightly oblate because the solar interior is rotating more rapidly than the surface, maybe a result of spinup as matter was drawn into the growing Sun, in line with the far shorter rotation periods of the gas giant planets Jupiter and Saturn. The rotation of the solar surface might have been slowed by the transfer of angular momentum to the solar wind, as Dicke inspired me to analyze in some detail.

Dicke (1964) reported that he, Henry Hill, and Mark Goldenberg had "built a special instrument to observe photoelectrically the Sun's oblateness, and preliminary measurements were made during a few weeks toward the close of the summer of 1963". Dicke and Goldenberg (1967) later reported that the measured fractional difference of equatorial and polar radii is 5.0 ± 0.7 parts in 10^5, about what Dicke (1964) felt would allow an interesting value of the scalar-tensor parameter, w, in equation (4)[12]. The result was not widely welcomed, in part because it challenged general relativity theory, and certainly in part because identifying and measuring a level surface on the Sun is seriously challenging. Subsequent measurements by Hill et al. (1974), and Hill and Stebbins (1975), used techniques that Hill, then in the Gravity Research Group, and his Princeton graduate student Carl Zanoni, originally meant to be used for measurement of the gravitational deflection of starlight passing near the Sun without the aid of a Solar eclipse. This required design for the strong rejection of diffracted sunlight (Zanoni 1967, Ph.D. Princeton; Zanoni and Hill 1965), which benefitted the design of the Hill et al. oblateness measurements. The conclusion was that the solar oblateness is too small for a serious effect on the orbit of Merury. This proved to be consistent with the probe of the internal rotation of the Sun afforded by the splitting of frequencies of modes of solar oscillation with different azimuthal numbers m. This revealed that the solar interior is rotating nearly as a solid body at close to the mean of the rotation rates seen at different latitudes on the solar surface (Thompson et al. 2003). Rozelot and Damiani (2011) reviewed the history and status of this subject. We may conclude that the mass distribution in the Sun is not likely to have disturbed the test of gravity theory from the orbit of Mercury. The Dicke and Goldenberg (1967) measurement is marked at line 36 in the timeline (Fig. 9), as a step toward closing the case for this important test of gravity physics.

[12] The situation in the late 1960s was summarized by Dicke (1969). He reported that the contribution to the precession of the obit of Mercury by the Dicke-Goldenberg measurement of the oblateness of the solar mass distribution is 3.4 ± 0.5 seconds of arc per century. This contribution subtracted from the measured precession, 43.10 ± 0.44, is the precession to be attributed to the departure from Newtonian gravity physics. And this precession divided by the relativistic prediction, 43.03, is 0.920 ± 0.015. This would be a serious challenge for general relativity. In the scalar-tensor theory with $w = 5$ this ratio is $(4+3w)/(6+3w) = 0.905$, or one standard deviation below the measurement. At about this time the first radio interferometer measurements (Seielstad et al. 1970; Muhleman et al. 1970) showed that the relativistic deflection of light by the sun is 1.02 ± 0.12 times the relativistic prediction. In the scalar-tensor theory with $w = 5$ the deflection is $(3 + 2w)/(4 + 2w) = 0.93$ times the relativistic prediction. This again is about one standard deviation below the measurement. And one standard deviation is a quite acceptable degree of consistency. The gravitational redshift is the same in general relativity and the scalar-tensor theory. That is, in 1969 Dicke's evidence was that the scalar-tensor theory with $w = 5$ is consistent with the classical tests, provided the solar oblateness measurement was correct. The scalar-tensor solar gravitational deflection of light, 0.93 times the relativistic prediction at $w = 5$, is the gravitational deflection in the Wheeler-Dicke wager shown in Figure 5.

The remarkably tight bound on the scalar-tensor gravity theory from the Cassini satellite radar timing experiment is $w > 10^{4.5}$ (Bertotti et al. 2003). In simple solutions for the expansion of the universe in the scalar-tensor theory this bound implies that the rate of change of the strength of gravity is limited to $-\dot{G}/G \lesssim 10^{-14}$ yr^{-1}, tighter even than the more direct bounds from the Lunar Laser Ranging Experiment (Eq. (15)) and pulsar timing (Zhu et al. 2015). Mchugh (2016) reviewed the history of tests that have progressively tightened the bound on w. The original vision of the scalar-tensor theory certainly must be scaled back to at most a small perturbation to general relativity. But the scalar-tensor theory still fascinates, figuring in the search for a deeper gravity physics, as one sees by its frequent mention in the review, *Beyond the cosmological standard model* (Joyce et al. 2015). And an excellent experimental program similarly can lead in unexpected directions. The solar distortion telescope Dicke first designed in the early 1960s was turned to detection of variation of the solar surface temperature as a function of solar latitude during five years of the solar cycle, a serious contribution to the continuing attempts to interpret the observed variation of the solar constant with the sunspot cycle (Kuhn et al. 1988). Kuhn et al. (2012) reviewed the history and present status of precision space-based measurements of the shape of the Sun.

6.13 Cosmology and the great extension of tests of gravity physics

General relativity theory inspired modern cosmology, and it was recognized in the mid-1950s that cosmological tests might in turn test relativity. Thus Walter Baade spoke at the 1955 Berne Conference, *Jubilee of Relativity Theory*, on progress in the measurement of a fundamental datum for cosmology, the extragalactic distance scale. This was an important step toward what has grown to be the rich science of cosmology.

6.13.1 The expanding universe

Hubble (1929) found the first reasonably clear evidence of the linear relation between galaxy redshifts and distances,

$$v = H_o r, \tag{17}$$

which Lemaître (1927) had shown is to be expected if the universe is expanding in a homogeneous and isotropic way.[13] The cosmological redshift, z, is defined by the ratio of wavelengths of features in the observed spectrum and the laboratory wavelengths at emission, as $1 + z = \lambda_{\mathrm{obs}}/\lambda_{\mathrm{em}}$. At small z the redshift may be considered to be a Doppler shift, where the recession velocity is $v = cz$. The distance r may be inferred from the inverse square law using estimates of the intrinsic luminosities of observed objects, expressed as absolute magnitudes, using the observed energy flux densities, expressed as apparent magnitudes. A commonly used term thus is the redshift-magnitude relation (though one also uses the relation between intrinsic linear sizes and observed angular sizes). This redshift-magnitude relation also describes observations of objects at greater distances, where the redshift z is large, and one looks for the relativistic effects of spacetime curvature.

Hubble and Humason (1931) considerably improved the case for equation (17), reaching redshift $z \sim 0.07$ for giant galaxies, and Hubble (1936) showed still better evidence in a larger sample that reached $z \simeq 0.15$. These results were very influential

[13] Lemaître derived this linear relation from a solution to Einstein's field equation, but it follows more generally from standard local physics in a near homogeneous spacetime described by a metric tensor.

in the development of cosmology. The Hubble and Humanson paper on measurements of the redshift-magnitude relation accordingly is marked at line 7 in Figure 9.

Another important early advance in cosmology was Hubble's (1936) deep counts of galaxies as a function of limiting apparent magnitude (line 10 in Fig. 9). The counts were not inconsistent with Einstein's Cosmological Principle, that the distribution of the galaxies is homogeneous in the large-scale average. The counts reached an impressively large redshift (estimated at $z \sim 0.4$ in Peebles 1971). Systematic errors in Hubble's counts were problematic enough that they could not rule out Mandelbrot's (1975) argument discussed in Section 7.1 for a fractal galaxy distribution with fractal dimension D well below three. But one could conclude that, if there were an observable edge to the universe of galaxies, the most distant ones would be flying away at near relativistic speeds.

It was understood early on (e.g., Tolman 1934, p. 185) that the relativistic Friedman-Lemaître cosmological models predict that at high redshift there may be a departure from the redshift-magnitude relation implied by equation (17), depending on the model parameters (including the mean mass density and the mean curvature of space sections at constant world time). Interest in detecting a departure from equation (17) increased with the introduction of the Steady State cosmology, which makes a definite prediction for the form of the redshift-magnitude relation (Bondi and Gold 1948; Hoyle 1948). Humason et al. (1956) reported progress in measurements of this relation, which had reached redshift $z \sim 0.2$, a modest advance over Hubble (1936) two decades earlier. Sandage (1961) presented a detailed analysis of the prospects for further advances in the application of this and other cosmological tests. We might take as prophetic Sandage's remark that (in his italics) "*If observations show q_o to be -1, we cannot decide between a steady-state universe and a Lemaitre-Eddington universe.*" The parameter q_o is a measure of the departure from equation (17), and the value $q_o = -1$ is predicted by the Steady State cosmology and by the general relativity Friedman-Lemaître model if the expansion rate is dominated by Einstein's cosmological constant, Λ. As it happens, well-checked measurements of the redshift-magnitude relation for supernovae (of type Ia, the explosions of white dwarf stars) are close to $q_o = -1$ (Riess et al. 1998; Perlmutter et al. 1999). This is the degeneracy Sandage noted. The degeneracy is broken, and the classical Steady State cosmology convincingly ruled out, by other cosmological tests. The evidence briefly reviewed in Section 6.13.3 is that we live in an evolving universe now dominated by Einstein's Λ.

Establishing the value of Hubble's constant, H_o, in equation (17) by astronomical methods requires determination of the intrinsic luminosities of extragalactic objects, a difficult task. Hubble and Humason (1931) slightly increased Hubble's (1929) estimate to $H_o \simeq 560$ km s^{-1} Mpc^{-1}, a result that was generally accepted until the 1950s. For example, in the monograph *Relativity, Thermodynamics, and Cosmology*, Tolman's (1934, p. 177) comment about the measured value of H_o was that "It is believed that the uncertainty in the final result is definitely less than 20 per cent". But this meant that the characteristic expansion time is $1/H_o \sim 2 \times 10^9$ yr, which was uncomfortably short, roughly half the largest radioactive decay ages of terrestrial minerals and estimates of stellar evolution ages. It was not widely discussed then, but Lemaître's 1931 model, with the values of Λ, the cosmic mean mass density, and space curvature chosen so the universe passed through a time of slow expansion, could reconcile the large estimate of H_o with a long expansion time. It did require a very special adjustment of parameters, however.

Bondi and Gold (1948) pointed out that the Steady State cosmology allows galaxies of arbitrarily great age, albeit with great scarcity, thus allowing a large age for our particular galaxy. But the Steady State cosmology predicts that the mean age of the galaxies is $1/(3H_o) \sim 6 \times 10^8$ yr (for the estimate of H_o accepted then). Gamow (1954) remarked that this would mean that "we should find ourselves surrounded by a bunch of mere youngsters, as the galactic ages go". He remarked that this seemed inconsistent with the observation that our Milky Way galaxy and neighboring galaxies have similar mean stellar spectra, indicating similar stellar evolution ages.

Walter Baade improved the situation with his announcement at the 1952 Rome Meeting of the International Astronomical Union that he had found a correction to the extragalactic distance scale: he found that cepheid variable stars of type I (younger, with larger heavy element abundances) are considerably more luminous than previous estimates. This increased estimates of extragalactic distances, and reduced Hubble's constant to about $H_o = 180$ km s^{-1} Mpc^{-1} according to Baade[14]. Sandage (1958) improved the situation still more by his correction for misidentification of gaseous nebulae as luminous stars. This brought his estimate to $H_o \simeq 75$ km s^{-1} Mpc^{-1}, "with a possible uncertainty of a factor of 2". For further commentary on these and related developments see Trimble (1996). The Planck Collaboration (2015b) found $H_o = 68 \pm 1$ km s^{-1} Mpc^{-1} indirectly derived from precision CMB anisotropy measurements. This is close to Sandage's central value, and still large enough to require the "disreputable Λ", as Robertson (1956) put it at the Bern Conference.

The revision of the distance scale was influential enough to be marked in the timeline (line 15, at Baade's announcement of his correction at the 1952 IAU conference). Its importance is illustrated by Bondi's comment added to the second edition of his book, *Cosmology* (Bondi 1960, p. 165): "It is not easy to appreciate now the extent to which for more than fifteen years all work in cosmology was affected and indeed oppressed by the short value of T (1.8×10^9 years) so confidently claimed to have been established observationally". The remark is fair, but one might ask why the theorists did not challenge the observers.

At the Bern Conference Klein (1956) proposed an alternative to the large-scale homogeneity assumed in the Friedman-Lemaître and Steady State models. Perhaps the galaxies are drifting apart into empty space after an explosion of a local concentration of matter. Klein may not have been aware that Hubble's (1936) deep galaxy counts require near relativistic recession velocities, hence a relativistic explosion. But still the idea was well worth considering, because Einstein's picture of a homogeneous universe was not apparent in the surveys of the galaxy distribution available at that time. Oort (1958) emphasized this in his commentary on the observational situation in cosmology, in the proceedings of the Solvay Conference on *La structure et l' évolution de l'univers* (Stoops 1958). Oort commenced his paper with the statement that "One of the most striking aspects of the universe is its inhomogeneity". He went on to review the observations of clustering of galaxies on the largest scales that could be surveyed then. But Oort was willing to estimate the cosmic mean mass density, because the assumption of homogeneity in the average over still larger scales was not inconsistent with Hubble's (1936) deep galaxy counts.

Measures of the large-scale distribution of extragalactic objects were starting to improve in the late 1950s. At the Chapel Hill Conference, Lilley (1957) discussed the Second Cambridge Catalogue of Radio Sources (Shakeshaft et al. 1955). The main point of interest then was the count of sources as a function of flux density, for which the Steady State cosmology makes a definite prediction. This test was spoiled by side-lobe confusion of source identifications. But the map of angular positions of the radio sources was not so seriously afflicted, and the strikingly uniform distribution of sources across the sky was not what one would expect in a fractal universe (this is further discussed in Sect. 7.1). The radio source map, with Hubble's deep counts and the

[14] Baade's presentation to the IAU is reported without details in the Transactions of the IAU, Volume 8, p. 397, in a summary of the talks presented. We can be sure Baade also reported his distance scale correction at the 1955 Bern Conference, as discussed in Section 2, though he did not contribute a paper to the proceedings.

observed linearity of the redshift-distance relation, was among the first observational indications that Einstein's argument from Mach's Principle for the Cosmological Principle might be right (and for this reason the 2C catalog is entered in line 16 in Fig. 9). More evidence for the Cosmological Principle came with the discovery that space is filled with the near uniform sea of microwave radiation discussed next.

6.13.2 The sea of cosmic microwave radiation

In the late 1940s George Gamow, with his student Ralph Alpher and their colleague Robert Herman, developed a theory of element formation in the early stages of expansion of the universe in an initially hot dense Friedman-Lemaître cosmological model (Ralph Alpher 1948, Ph.D. George Washington University; Gamow 1948; Alpher et al. 1953). In this theory space is filled with a sea of thermal radiation, now known as the Cosmic Microwave Background, or CMB. The properties of this radiation have been read now in considerable detail that yields demanding tests of general relativity theory. The story of how the thermal radiation in this hot Big Bang model was predicted, and the community came to recognize its existence, is complicated. My analysis of how Gamow and colleagues arrived at their cosmology with its thermal radiation is in Peebles (2014). The book, *Finding the Big Bang* (Peebles et al. 2009), recalls how the radiation was discovered, interpreted, and its properties explored. Peebles (2012) adds personal recollections. A few aspects of this CMB story are entered here as part of the history of how experimental gravity physics grew.

The CMB research program began in the summer of 1964 when Dicke gathered Peter Roll, David Wilkinson, and me, junior faculty and a postdoc in his Gravity Research Group, to discuss a proposal for what the universe might have been doing before it was expanding. Dicke suggested that the universe may have been collapsing, maybe in a cycle of expansion and contraction. The bounce would have been seriously irreversible, producing entropy largely in the form of a sea of thermal radiation. This idea of irreversibility of a bouncing universe traces back at least to Tolman (1934, p. 443), who pointed out that if the bounce conserved entropy then in a cyclic universe each cycle would last longer. Dicke proposed a specific physical model for entropy production in the bounce. And it is very typical of Dicke that his idea suggested an experiment: a search for the radiation.

Ideas about early universe physics have changed, but Dicke's thought about entropy production is worth recording here. Stars in the previous cycle would have been converting hydrogen into heavier elements, releasing nuclear binding energy of several million electron volts for each nucleon that became bound in a heavier atomic nucleus. This binding energy would have been radiated as some 10^6 starlight photons per nucleon, because the starlight photons have energy on the order of an electron volt. If the contraction before the bounce to the next expansion were deep enough, then during the contraction these starlight photons would have been blueshifted to energies above several million electron volts, enough to photodissociate the elements heavier than hydrogen that were formed in stars during the previous expansion and collapse. That would yield fresh hydrogen for the generations of stars in the next expansion. Just a few of the blueshifted starlight photons would have been needed to release each nucleon from a heavier atomic nucleus, and the remaining 10^6 photons per nucleon would relax to a sea of thermal radiation, in a seriously irreversible process[15].

[15] The estimate in Dicke and Peebles (1979) of the number of cycles needed to produce the present entropy per nucleon can be improved. For a more direct approach, let the energy density in the Cosmic Infrared Background (the CIB) produced by stars and active galactic nuclei during the cycle of expansion and contraction be f times the energy in the CMB.

Dicke suggested that Roll and Wilkinson build a microwave radiometer that might detect this thermal CMB radiation. This was timely for Roll, because the Roll et al. (1964) Eötvös experiment was completed, and it was timely for Wilkinson, because he had not yet settled on a long-term project after arrival from completion of his Ph.D. at the University of Michigan. Dicke suggested that I look into possible theoretical implications. At the time this was to me just another of Dicke's many ideas, most of which I found interesting to explore. This one was very interesting. It was speculative, to be sure, but that was not out of the ordinary in the Gravity Research Group. And Dicke's idea has proved to be wonderfully productive.

While at the MIT Radiation Laboratory during the Second World War Dicke invented much of the microwave radiometer technology Roll and Wilkinson used. We had to remind him that Dicke et al. (1946) had already used a Dicke radiometer to place a limit on the CMB temperature. They reporting a bound of 20 K on "radiation from cosmic matter at the radiometer wave-lengths", 1.0 to 1.5 cm (line 13 in Fig. 9). The radiometer Roll and Wilkinson built used Dicke's technique of suppression of receiver noise by switching between the horn antenna and a reference source of thermal radiation at known temperature, for which Roll and Wilkinson used liquid helium.

A half decade before Dicke's suggestion to Roll and Wilkinson, an experimental microwave receiving system at the Bell Telephone Laboratories detected more noise than expected from accounting of noise sources in the instrument (DeGrasse et al. 1959). This curious result was repeatable (as Hogg 2009 recalled), but not widely discussed. The noise originating in the receivers in the Bell experiments was so small that Dicke's switching technique was not needed, but Penzias and Wilson (1965) used it to remove any chance that the "excess antenna temperature" could be attributed to some error in the noise accounting, again using a liquid helium reference source (they could switch much more slowly than in the Roll and Wilkinson radiometer, because their system noise was so small). Penzias and Wilson also carefully tested and eliminated possible terrestrial sources of the excess radiation. They had an important measurement. They learned of a possible interpretation when they learned of the Roll and Wilkinson search for a sea of radiation left from the early universe.

It was clear that this excess radiation, if extraterrestrial, must be close to isotropic, because the signal did not vary appreciably as the Earth and antennas rotated. Partridge and Wilkinson (1967) improved this to a demonstration that the large-scale variation of the CMB intensity across the sky is less than about 3 parts in 10^3. This was another early argument for Einstein's Cosmological Principle. Space was known to be close to transparent at these wavelengths, because radio sources were observed at high redshift, so the isotropy had to mean either that the universe is nearly uniformly filled with this radiation, or else that the distribution of the radiation is significantly inhomogeneous but close to spherically symmetric, with us near

---

Suppose the bounce produces no entropy apart from the thermalization of the CIB, and suppose nucleons and radiation are conserved in the bounce. Then an easy exercise shows that the entropy per nucleon in the CMB after the bounce is $(1+f)^{3/4}$ times the entropy per nucleon before the bounce. The observed energy density in the CIB is about $f \sim 0.1$ times the energy in the CMB (Hauser et al. 1998). If the CIB is not going to receive much more energy in the rest of the present cycle, then in this model the entropy per nucleon in the next cycle will be $1.1^{3/4}$ times the entropy per nucleon in this cycle, roughy a 10% increase. Not addressed, of course, is what might have caused the bounce, and how the quite clumpy distribution of matter, including black holes, just prior to the bounce could have become so very close to the near homogeneous condition we may expect to obtain at the start of the next cycle, as it did at the start of our cycle according to the established cosmology. The Steinhardt and Turok (2007) model offers a way to deal with these issues, but in a manner that eliminates the entropy Dicke was thinking about.

the center of symmetry. The latter seemed less likely then, and it is now convincingly ruled out by the cosmological tests.

In Gamow's (1948) hot Big Bang picture, and the Dicke et al. (1965) interpretation of the Bell Laboratories excess noise, the radiation intensity spectrum would be expected to be close to thermal. This is because the spectrum would have to have been very close to thermal in the early universe, when the high density and temperature would have produced statistical equilibrium, and the CMB thermal spectrum would be expected to have been preserved as stars and galaxies formed, because the CMB heat capacity is large compared to what is expected to be readily available in matter. In the Steady State cosmology one is free to postulate continual creation of radiation as well as matter, and one is free to postulate the spectrum of the radiation at creation. But the radiation spectrum would not have relaxed to thermal because, as has been noted, the universe is observed to be transparent at these wavelengths. And since the detected radiation would be the spectrum at creation convolved over redshift it would require an exceedingly contrived spectrum at creation to make the detected spectrum close to thermal. One also had to consider that the Bell excess noise might only be the sum of local sources of microwave radiation, making it consistent with the Steady State cosmology or a cold Big Bang model. But again, in this local source picture the spectrum would not likely be close to thermal. The spectrum measurements thus were critical to the interpretation of this radiation.

The Penzias and Wilson (1965) and Roll and Wilkinson (1966) experiments operated at 7.4 cm and 3.2 cm wavelength, respectively. Their measurements at the two wavelengths were consistent with the same effective radiation temperature, which argued for a close to thermal spectrum in the long wavelength, Rayleigh-Jeans, part of the spectrum. An important early advance was the recognition that the observation of absorption of starlight by interstellar cyanogen (CN) molecules in the first excited level as well as the ground level offered another measure of the spectrum, at the wavelength 0.26 cm of transition between the ground and excited levels. McKellar (1941) had translated the observed ratio of populations in the excited and ground levels to an effective temperature $T = 2.3$ K. George Field, Patrick Thaddeus, and Neville Woolf (whose recollections are in Peebles et al 2009, pp. 75–78, 78–85, and 74–75 respectively) recalled their prompt recognition that the Bell radiation, if close to thermal, could account for the observed CN excitation. Iosif Samuilovich Shklovsky (1966) in the Soviet Union independently made the same point. New measurements of the CN excitation temperature were found to be consistent with the Bell and Princeton radiometer measurements (Field et al. 1966; Thaddeus and Clauser 1966). That meant the CMB intensity at 0.26 cm and the measurements at 7.4 cm and 3.2 cm wavelength are reasonably consistent with a thermal spectrum, including the departure from the Rayleigh-Jeans power law about at the CN wavelength. The departure from the Rayleigh-Jeans power law was demonstrated also by a radiometer intensity measurement at 0.33 cm (Boynton et al. 1968). These results are marked at line 33 in the timeline in Figure 9.

In the 1960s discussions of how cosmology and gravity physics might be tested by CMB measurements were beclouded by indications that there is a considerable excess of energy relative to a thermal intensity spectrum at wavelengths near and shorter than the thermal peak. In the hot Big Bang picture this excess would require the postulate that explosions of some sort in the early universe added considerable radiation energy. Such explosions surely would have obscured any primeval patterns imprinted on the radiation by whatever imprinted the departures from exact homogeneity in the primeval mass distribution that grew into galaxies. This uncomfortable situation was resolved a quarter of a century after identification of the CMB, by two independent groups (Mather et al. 1990; Gush et al. 1990). Both showed that the spectrum is very

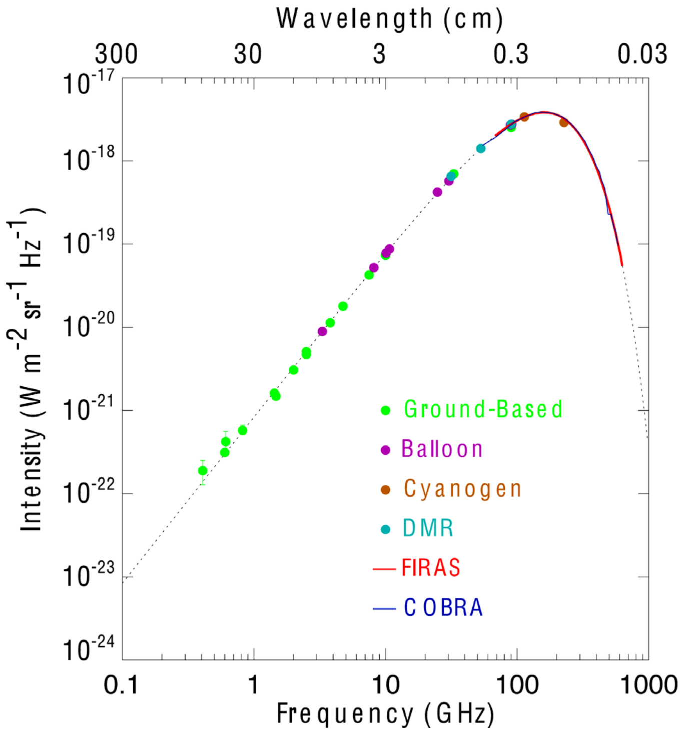


Fig. 13. The energy (intensity) spectrum of the Cosmic Microwave Background radiation that nearly uniformly fills space (Kogut 2012). The dotted curve is the theoretical Planck blackbody spectrum; the solid curve near the peak shows measurements.

close to thermal over the radiation intensity peak. Kogut (2012) compiled these and the other spectrum measurements shown in Figure 13[16].

The strikingly close agreement with a thermal spectrum has two important implications. First, we have noted that space in the universe as it is now is observed to be close to transparent. This means that the CMB would not have relaxed to the thermal spectrum shown in Figure 13 in the universe as it is now. Formation of the thermal spectrum requires that the universe expanded from a state dense and hot enough to have forced thermal relaxation. This serious evidence that our universe is evolving, drawn from such a simple figure (albeit one based on many far from simple measurements), is to be ranked with the memorable advances in the exploration of the world around us. The spectrum does not offer a serious constraint on gravity physics, however, because preservation of the thermal spectrum as the universe expands follows if spacetime is well described by a close to homogeneous and isotropic line element with standard local physics. The thermal spectrum in Figure 13 argues for cosmic evolution, but it does not argue for general relativity theory.

Second, the demonstration in 1990 that the CMB spectrum is close to thermal meant that the CMB need not have been seriously disturbed from its primeval condition. This meant that there was a chance that measurements of the CMB could be

[16] It is to be noted that most of these data are from measurements of the differences between the CMB intensity and the intensity of radiation from black thermal sources at well-calibrated temperatures. Since the theoretical Planck blackbody spectrum is not as well tested as these comparison measurements, the accurate conclusion is that the CMB spectrum is very close to thermal.

mined for comparison to what is predicted by theories of cosmic evolution in the early universe. That is, we might not have to deal with the complexities of nongravitational disturbances of the CMB by the processes of galaxy formation. Peebles and Yu (1970) introduced the radiative transfer computation that predicts the primeval oscillations in the power spectra of the distributions of the CMB and of the matter, under the assumptions of general relativity theory and adiabatic initial conditions (meaning the primeval spatial distribution $n_\gamma(\boldsymbol{x})$ of the CMB photons and the distribution $n(\boldsymbol{x})$ of the matter particles have the same small fractional departures from exact homogeneity, $\delta n_\gamma/n_\gamma = \delta n/n$). The first pieces of evidence that these assumptions are on the right track were found some three decades after Peebles and Yu, from measurements of the angular distribution of the CMB and the spatial distribution of the galaxies. This history was reviewed in Peebles et al. (2009).

6.13.3 The cosmic microwave background and general relativity

To add to the picture of how tests of gravity physics grew out of research in the 1960s, including the discovery and early steps in exploration of the Cosmic Microwave Background, I offer this overview of the CMB-based cosmological tests that reached fruition well after the naissance. These tests treat gravity physics in linear perturbation from the relativistic Friedman-Lemaître solution. The linearity greatly simplifies theoretical predictions, but of course it means the tests are limited to this approximation. The linear treatment of gravity and the effect of departures from homogeneity on the rate of expansion of the universe remains a good approximation even at low redshifts, when the mass distribution has grown strongly nonlinear, because the gravitational potentials are still small, on the order of $(v/c)^2 \sim 10^{-6}$, everywhere except close to neutron stars and massive compact objects, presumably black holes. On the scale of things these compact objects act as dark matter particles.

The standard and accepted six-parameter ΛCDM cosmology starts with the cosmologically flat Friedman-Lemaître solution to Einstein's field equation. This solution is perturbed by departures from homogeneity that are assumed to be initially adiabatic, Gaussian, and near scale-invariant. The stress-energy tensor in Einstein's equation is dominated by the CMB, nucleons, neutrinos, the hypothetical nearly collisionless initially cold nonbaryonic dark matter, or CDM, and Einstein's cosmological constant, Λ. That is, the term ΛCDM signifies a considerable number of assumptions in addition to the presence of Λ and CDM.

Gravity in general relativity theory in linear approximation is represented by two potentials, sourced by the active gravitational mass density and the energy flux density. At redshifts well above $z_d \simeq 1200$ baryonic matter was fully ionized and the plasma and radiation acted as a fluid made slightly viscous by diffusion of the radiation. This matter-radiation fluid has mass density ρ, pressure p, active gravitational mass density $\rho+3p$, and inertial and passive mass densities $\rho+p$, where ρ is the energy density and p is the pressure. Near redshift z_d the CMB radiation is to be described by its photon distribution function in single-particle phase space, with the CMB evolution in phase space described by the Boltzmann collision equation for the photons propagating according to the equation of motion in the gravitational potentials and scattering off the free electrons. The dark matter is modeled as an initially cold ideal gas that behaves as Newton would expect, while the baryons are treated as a fluid that suffers radiation drag. The CMB and neutrino sources for the gravitational potentials are to be computed as integrals over the distribution functions in their single-particle phase spaces.

The two points of this perhaps unduly schematic accounting are that the measurements of the CMB probe a considerable variety of elements of general relativity theory, and that the measurements are interpreted in a model that depends on a considerable number of assumptions in addition to general relativity. It is important therefore that there is a considerable variety of measurements. The variety is large enough that an adequate review is too long for this paper, so I refer instead to the Planck Collaboration (2015b) discussion of the empirical situation from the CMB measurements.

6.14 On the empirical case for general relativity theory

An assessment of the weight of empirical evidence for an idea or theory in natural science is of course informed by the degree of precision of the supporting measurements, but at least equally important is the variety of tests that probe the situation in independent ways. For consider that a measurement interpreted by the wrong theory may be precise but quite wrong. The situation for tests of gravity physics and general relativity is illustrated in Figure 14, where the probes are ordered by characteristic length scale, from the laboratory to the edge of the observable universe. The section numbers in parentheses indicate discussions of lines of research that originated in the 1960s and earlier. The picture has been made more complete by adding the results from observations of binary pulsars (Hulse and Taylor 1975), as noted in Section 6.12, and the LIGO et al. (2016) gravitational wave detection. The picture could be made even more complete by marking the tests of the gravitational interaction on scales down to 0.01 cm (Adelberger et al. 2009). Still other probes of gravity physics are discussed by the Planck Collaboration (2015b) and references therein. The point of the figure is that gravity has been looked at from many sides now, in phenomena that are observed and analyzed by a variety of methods, by measurements that span a broad range of length scales. The case for general relativity theory as a good approximation to what actually has been happening rests on the abundance and variety of these tests. In my opinion they make the case about as good as it gets in natural science.

Might there be a still better approximation, perhaps one that eliminates the need for the hypothetical dark matter? The case for its existence is indirect, from the consistency of the network of cosmological tests. A precedent might be noted: Before the Reines-Cowan detection there was compelling indirect evidence for the existence of neutrinos from measured decay energy spectra and transition rates (as reviewed in Blatt and Weisskopf 1952, though there still was uncertainty about the Fermi and Gamow-Teller interactions). There is the big practical difference that the expected detection rate for neutrinos was known, while the parameters for direct detection of the dark matter are far less well characterized. But in the philosophy of this paper the weight of the indirect evidence for neutrinos in the early 1950s was about as good as the weight of the present indirect evidence for the dark matter of the ΛCDM theory (at the time of writing). We may be surprised, of course, and general relativity with its dark matter and cosmological constant certainly might be disestablished. That seems quite unlikely, but an arguably likely, and welcome, development would be the establishment of a still better theory. It may show us how these hypothetical components, Λ and CDM, fit into the rest of physics, perhaps within an extension of the standard model for particle physics, perhaps in a generalization of general relativity. Some such eventuality seems likely, simply because it has been the normal course of events in the history of physics.

7 Lessons from the naissance of experimental gravity physics

This history of the empirical exploration of gravity in the 1960s offers some lessons that are largely of historical interest because there no longer is much left of the

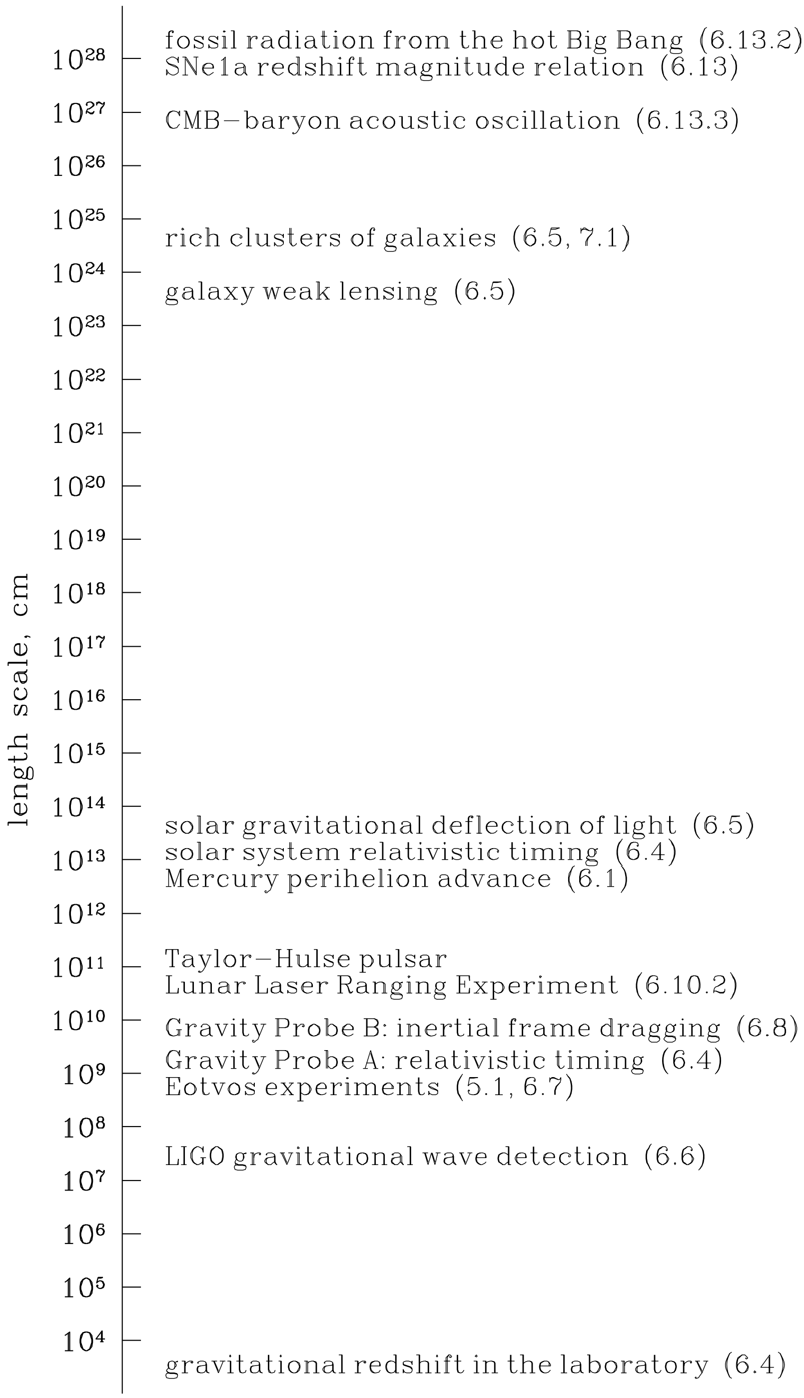


Fig. 14. Length scales of probes of gravity physics.

low-hanging fruit of the 1960s. But some lessons are of broader significance, and are conveniently examined in some detail in the context of experimental gravity physics because the subject was such a small science in the 1960s.

7.1 Nonempirical evidence and empirical surprises

The story of how gravity physics grew during the naissance shows the power of concepts of logic and elegance advocated by influential scientists informed by previously successful advances in science. This may be termed the importance of standard and accepted belief systems. The story also shows the power of empirical evidence, which can surprise us, and on occasion force revisions of the belief systems.

In 1915 there was only one seriously demanding test of general relativity theory, from the motion of the planet Mercury, which orbits the Sun at roughly 10^{13} cm. The empirical situation was little improved in 1955. But a half century after that Einstein's theory has proved to pass a demanding network of tests on scales ranging from the laboratory to the edge of the observable universe, at the Hubble length $H_o^{-1} \sim 10^{28}$ cm. The latter is an extrapolation of fifteen orders of magnitude in length scale from the best evidence Einstein had, from Mercury's orbit, and the former, down to the laboratory, is a similarly enormous extrapolation. The successes of these extrapolations of a theory discovered a century ago is a deeply impressive example of the power of nonempirical evidence. But not all elegant ideas can be productive, of course. I offer three examples drawn from this history of the naissance of experimental gravity physics.

The distribution and motion of the mass around us has been shown to affect local inertial motion, in the Lense-Thirring effect. It has seemed natural to some to imagine that other aspects of how local physics operates might also be related to what is happening in the rest of the universe. We have an excellent empirical case that our universe is expanding and cooling from a very hot dense early state, and it has seemed natural to look for evolution of how local physics operates, perhaps to be described by evolution of the values of its dimensionless parameters. Perhaps proper clock rates defined by molecules, atoms, or atomic nuclei, which may be small enough that tidal stresses may be ignored, and may be adequately shielded from electromagnetic fields, cosmic rays, and all that, may still be affected by position or motion or cosmic evolution. As discussed in Section 4, this holistic concept has inspired some to think of an effective aether; some to think of Mach's Principle; Jordan to introduce a scalar-tensor theory that would express Dirac's Large Numbers Hypothesis; and Dicke to think of this scalar-tensor theory as expressing both the LNH and Mach's Principle. An holistic concept led Einstein to his prediction of the large-scale homogeneity of the universe. If Einstein had consulted astronomers he would have received no encouragement, but the observable universe proves to agree with his thinking. But although this holistic thinking, which Einstein so successfully applied, has continued to seem beautiful to some, it has not led to anything of substance since Einstein. The searching tests reviewed in Section 6 agree with the idea that local physics, including gravitation, is quite unaffected by what the rest of the universe is doing.

For a second example, consider the opinion of influential physicists in the 1950s that the cosmological constant Λ is inelegant if not absurd. The situation then was reviewed in Peebles and Ratra (2003, Sect. III). The power of this thinking is seen in the 1980s through the 1990s in the general preference for the relativistic Einstein-de Sitter cosmological model, with $\Lambda = 0$ and negligible mean space curvature. The first clear evidence that the observed galaxy peculiar velocities are smaller than expected in the Einstein-de Sitter model, if the space distribution of matter is fairly traced by the observed distribution of the galaxies, was presented by Davis and Peebles (1983). It was countered by the hypothesis that the galaxy space distribution is significantly more clumpy than the mass. The consequences of this biasing picture for the evolution of cosmic structure were first examined by Davis et al. (1985). Biasing agrees with the observed greater clustering of the rare most massive galaxies, but it is not naturally consistent with the observed quite similar distributions of the far more abundant normal and dwarf galaxies. This was first seen in the CfA galaxy redshift survey maps (Davis et al. 1982, Figs. 2a and 2d). By the arguments summarized in Peebles (1986), the biasing picture was a nonempirical postulate meant to preserve $\Lambda = 0$ and avoid curvature of space sections. Some (including de Vaucouleurs 1982; Peebles 1986; Brown and Peebles 1987; Peebles et al. 1989; Maddox et al. 1990; and Bahcall and Cen 1992) took seriously the growing empirical evidence that, assuming general relativity theory, either Λ or space curvature differs from zero. But the title of a paper adding to this evidence, *The baryon content of galaxy clusters: a challenge to cosmological orthodoxy* (White et al. 1993), illustrates the general reluctance to abandon the orthodox Einstein-de Sitter cosmology with $\Lambda = 0$.

The preference for negligible large-scale space curvature was largely inspired by the inflation picture of the early universe, which in some scenarios was meant to account for the observed large-scale homogeneity by exceedingly rapid expansion of the early universe that might have stretched out and thus suppressed observable departures from homogeneity. The great expansion also would have stretched out and suppressed space curvature. This preference for suppressed space curvature proves to agree with stringent bounds obtained much later (Planck Collaboration 2015b, Sect. 6.2.4, and references therein). It is another impressive success for nonempirical evidence. But the empirical evidence is that we must learn to live with Λ (or something in the stress-energy tensor that acts like it), even though its numerical value is nonzero but absurdly small compared to natural estimates from quantum physics.

For the third example, consider Mandelbrot's (1975) argument that the galaxy space distribution may be a fractal. In a simple fractal distribution, particles are placed in clusters, the clusters placed in clusters of clusters, the clusters of clusters in clusters of clusters of clusters, and so on up. The mathematics of fractals is elegant, and there are interesting applications. Thus Mandelbrot explained that the measured length of the coastline of Brittany depends on the length resolution used to make the measurement. And Mandelbrot took note of de Vaucouleurs' (1970) point, that the galaxy distribution in surveys then available certainly resembled a fractal. This line of argument, if more widely pursued earlier, could have had the beneficial effect of driving observational tests. But as it happened, interest in the elegance of a fractal universe started to grow just as empirical evidence for Einstein's large-scale homogeneity was at last emerging (as discussed in Sect. 6.13, and reviewed by Jones et al. 2004).

Elegant ideas tend to be lasting. The Steady State cosmology is elegant, as is the fractal concept, and the two arguably find their places in the multiverse of eternal inflation. The role of the scalar field that Jordan and Dicke envisioned is now seen to be at most an exceedingly subdominant effect at low redshifts, but the possible role of scalar fields in gravity physics still is discussed. Examples are superstring theories (Uzan 2003), models for inflation, and the search for a deeper gravity physics (Joyce et al. 2015). Also to be considered is the notion that the cosmological constant, Λ, in the present standard cosmology may be an approximation to the energy of a slowly evolving scalar dark energy field (Peebles and Ratra 1988, who referred back to Dicke's thinking about the slow evolution of the strength of gravity). The very weak strength of the gravitational interaction is still discussed, now under the name of the gauge hierarchy issue. The idea that the gravitational interaction has been growing weaker as the universe expands, which so interested Dirac, Jordan, and Dicke (Sects. 4 and 6.10), still is explored (Sects. 6.10.2 and 6.12). Also still explored is the possible evolution of the fine-structure constant, as exemplified by equation (16), and evolution of the other dimensionless physical parameters, though now inspired largely by issues of completion of quantum physics (as reviewed by Uzan 2003).

Another old issue of completeness is the considerable difference between the relativistic constraint on the vacuum energy density, to be represented by the value of Einstein's cosmological constant if the vacuum is invariant under Lorentz transformations, and simple estimates of the sum of energies of the fields of particle physics. Early thinking was reviewed in Peebles and Ratra (2003, Sect. B.3). In particular, Rugh and Zinkernagel (2002) recalled Wolfgang Pauli's recognition in the 1930s that the zero-point energies of the modes of the electromagnetic field presented a problem. In the English translation (Pauli 1980) of Pauli's (1933) *Die allgemeinen Prinzipien der Wellenmechanik*, Pauli wrote that, in the quantization of the electromagnetic field,

> a zero-point energy of $\frac{1}{2}\hbar\omega_r$ per degree of freedom need not be introduced here, in contrast to the material oscillator. For …this would lead to an infinitely large energy per unit volume because of the infinite number of degrees of freedom of the system …Also, as is obvious from experience, it does not produce any gravitational field.

The zero-point energy of each mode of a matter field, which is computed by the same quantum physics, must be real to account for measured binding energies, but these zero-point energies also add up to a absurdly large – negative for fermions – mean vacuum energy density. This curious situation was discussed in passing in Gravity Research Group meetings in the 1960s, but not to the point of publishing, because we did not know what to make of it. Zel'dovich (1968) published, and pointed out that the vacuum energy likely appears in the form of Einstein's cosmological constant, though with an apparently absurd value. Zel'dovich's paper was visible enough, and the issue for gravity physics and quantum physics is deep enough, to be entered in the timeline as line 37.

The Anthropic Principle, for which one might claim nonempirical evidence, offers the thought that we live in an element of a multiverse, or in a part of a really extended universe, in which the sum of zero-point energies, latent heats, and whatever else contributes to the vacuum energy density, happens to be small enough to allow life as we know it (Weinberg 1987). The earlier version of this line of thinking that we heard in the Gravity Research Group was that the age of the expanding universe has to be on the order of 10^{10} years, to allow time for a few generations of stars to produce the heavy elements of which we are made, but not so much time that most of the stars suitable for hosting life on a planet like Earth around a star like the Sun have exhausted their supply of nuclear fuel and died (Dicke 1961b). The resolution of the vacuum energy puzzle by the nonempirical evidence of the Anthropic Principle is considered elegant by some, ugly by others, a healthy situation in the exploration of a serious problem.

7.2 Prepared and unprepared minds

It is observable in this history of experimental gravity physics that "chance favors the prepared mind" (in an English translation of the thought attributed to Pasteur). Consider, for example, how quickly Pound and Rebka turned the announcement of Mössbauer's effect into a laboratory detection of the gravitational redshift (Sect. 2). In his recollections Pound (2000) mentioned Singer's (1956) discussion of how an atomic clock in an artificial satellite might be used to detect the gravitational redshift. Pound recalled that he and Rebka recognized that the Mössbauer effect presented an opportunity and "a challenge; namely, to find a way to use it to measure relativistic phenomena, as I had wanted to do with atomic clocks; however, the clocks had not proved sufficiently stable". Dicke's letter to Pound (Sect. 6.2) showed that, when Pound and Rebka (1959) announced their intention to do this experiment, at least two other groups also were working on it. We see that minds were well prepared for this experiment. Wilson and Kaiser (2014) considered how Shapiro's thinking about planetary radar ranging experiments helped prepare his mind for the measurement of the relativistic effect on the time delay of planetary radar pulses that pass near the Sun (Sect. 6.4). Dicke was prepared and searching for probes of gravity physics. For example, as NASA was learning to fly rockets Dicke and his group were exploring how to use this new space technology for precision measurements of the orbits of satellites, perhaps by tracking angular positions of corner reflectors illuminated by searchlights, and then, when the technology allowed it, turning to pulsed or continuous wave lasers for precision ranging to corner reflectors on the Moon. The results of Lunar Laser Ranging, after considerably more work, had an intended consequence, the production of demanding tests of gravity physics. For Dicke, of course, an unintended consequence was a much tighter bound on evolution of the strength of gravity than he had expected.

We may consider also a situation in which minds were not prepared. The research by Gamow and colleagues in the late 1940s, on element formation in the early stages of expansion of a hot Big Bang cosmology, has grown into a demanding set of tests of gravity physics (as reviewed in Sect. 6.13.3). Gamow (1953) outlined this research in lectures at the 1953 Ann Arbor *Symposium on Astrophysics*. But I have not found any mention of the research by Gamow's group in the proceedings of the four international conferences on general relativity and gravitation that I have discussed in this history: Bern in 1955, Chapel Hill in 1957, Royaumont in 1959, and NASA in 1961. There is no mention of it in the proceedings of the next conference in this series, the 1962 Warszawa and Jablonna *Conférence internationale sur les théories relativiste de la gravitation* (Infeld 1964), there is no mention in two IAU Symposia where it might have figured: the *Paris Symposium on Radio Astronomy*, 1958 (Bracewell 1959), and *Problems of Extra-galactic Research*, Berkeley California, 1961 (McVittie 1962), and no mention in the *Solvay Conference on Cosmology*, Brussels 1958 (Stoops 1958). Bondi's (1952, 1960) book, *Cosmology*, which was the best review of research in cosmology during the naissance of experimental gravity, gave references to papers on element formation by Gamow's group, in a list of papers for further reading, but there was no discussion of these ideas in the text. We may conclude that, until 1965, the leading figures in relativity and cosmology were not prepared for the sea of thermal microwave radiation, the CMB. They did not know about Gamow's (1948) ideas or else they did not consider them promising.

Donald Osterbrock attended Gamow's Ann Arbor lectures, and remembered Gamow's idea of helium production in a hot Big Bang (as Osterbrock 2009 recalled). Geoffrey Burbidge also was at the Ann Arbor conference. Burbidge (1958) later commented on evidence that the helium abundance in our galaxy is larger than might be expected from production by stars. He did not mention Gamow's thinking, however. If Burbidge attended Gamow's lectures he had forgotten them, or considered them unpromising. Osterbrock and Rogerson (1961) added to the evidence for a large helium abundance in the Milky Way, even in the apparently older stars that have lower abundances of heavier elements. They pointed out that "the build-up of elements to helium can be understood without difficulty in the explosive formation picture". Their reference was to Gamow (1949). This is the first published announcement of a possible detection of a fossil from a hot early stage of expansion of the universe. It received no significant notice. Hoyle and Tayler (1964), apparently independently, also announced the possible significance for cosmology of the large cosmic helium abundance. We cannot know how this paper would have been received if another candidate remnant from a hot early universe, the sea of microwave radiation, had not been announced at close to the same time. This radiation was first detected, as anomalous noise in microwave telecommunications experiments, at the Bell Telephone laboratories (Sect. 6.13.2). Engineers were aware of the anomaly in 1959 (DeGrasse, Hogg, Ohm, and Scovil 1959). Five years later Penzias and Wilson had made a good case that the anomalous noise could not be instrumental or terrestrial. But their minds were not prepared for cosmology until they learned of the search for fossil microwave radiation by Dicke's Gravity Research Group.

What delayed general recognition of Gamow's hot early universe theory? Weiss (2016) pointed to two factors that in his experience might have contributed to the delay. First, Gamow's plan had been to account for the origin of all the elements (apart from possibly subdominant contributions by nuclear burning in stars). By 1950 Enrico Fermi and Anthony Turkevich, at the University of Chicago, had concluded that element buildup in the hot early universe would very likely end at the isotopes of hydrogen and helium. The failure of Gamow's plan may have tended to obscure the still viable – and now established – idea of formation of isotopes of the lightest elements. The second factor was the absence of experimentalists in Gamow's group[17]. This is to be contrasted with the active interactions of theory and practice in Dicke's group. In 1964 Peter Roll and David Wilkinson had little prior experience in microwave technology. The prior feeling I had for cosmology was that the homogeneous expanding universe solution to Einstein's equation seemed to be grossly oversimplified so as to offer workable exam problems; I thought it was to be compared to the acceleration of a frictionless elephant on an inclined plane. But we encouraged each other in our willingness to learn experiment and theory from Dicke's challenges.

Another factor that must be mentioned was Gamow's supreme lack of interest in details. In the 1950s he continued to write important papers; a notable example was the demonstration of the significant time-scale challenge to the Steady State cosmology (Gamow 1954, as discussed in Sect. 6.13.1). But Gamow (1953a, 1953b, 1956) turned from his perceptive 1948 physical argument for a hot Big Bang to a far less persuasive picture. His new picture assumed that the expansion rate predicted by the Friedman-Lemître equation became dominated by space curvature just when the mass densities in nonrelativistic matter and in the sea of thermal radiation passed through equality (with cosmological constant $\Lambda = 0$). Knowing the present matter density and Hubble's constant (for which Gamow took $\rho_m(t_o) = 10^{-30}$ g cm^{-3} and $H_o^{-1} = 10^{17}$ s or a little longer), the model predicts present temperature $T_o = 6$ to 7 K (Gamow 1953a, 1956). This specific prediction, of a not very low temperature, could have encouraged an experimental program aimed at detecting Gamow's radiation[18]. But a closer look could have discouraged the project, because it would have revealed that there was no basis for the starting assumption. The approach only places an upper bound of about 50 K on the present temperature, and no lower bound[19]. An interested

[17] Fermi and Turkevich certainly understood experiments. But their contribution to Gamow's program was a theoretical analysis, which they did not bother to publish, apart from giving their results to Alpher and Gamow.

[18] Bernard Burke (2009, p. 182) concluded that the resources were available to detect the CMB at $T_o = 2.725$ K in the 1950s. But I do not know of any evidence that anyone actually considered looking for Gamow's 6 to 7 K thermal radiation.

[19] In his popular book, *Creation of the Universe*, Gamow (1952) arrived at present temperature $T_o = 50$ K (p. 52), which indeed follows from his adopted age of the universe, $t_o = 10^{17}$ s, under the assumption that the expansion rate was dominated by radiation up to the present. This is a physically reasonable upper bound on the present radiation temperature in a relativistic cosmology given the present age. Gamow may not have meant it to be a serious estimate of T_o, however. On page 78 he wrote that at the time of equality of mass densities in matter and radiation the radiation temperature was $T_{\mathrm{eq}} = 300$ K. With his

experimentalist would have had to go back to Gamow (1948) to see a well-motivated physical argument for a present temperature of this order. Chernin (1994) argued for the elegance of Gamow's (1953a, 1953b, 1956) simplifying assumption that allowed a succinct and reasonably accurate account of cosmic evolution with present radiation temperature close to the Alpher and Herman (1948) estimate[20], about 5 K. This is a fair point. But the considerations in Gamow (1948) are elegant too, they are about as simple, and they are based on a specific physical condition, that thermonuclear reactions in the early universe produce a significant but not excessive mass fraction in atomic weights greater than the proton. It is difficult to understand why Gamow turned to the much less well motivated theory he presented in the 1950s.

One may also wonder why there were such different community reactions to two announcements of evidence of fossils from a hot early universe. The first, by Osterbrock and Rogerson (1961), that the unexpectedly large abundance of helium may be a remnant from the hot early universe, received little notice. The second, by Dicke et al. (1965), that the unexpectedly large noise in Bell communications detectors may be a remnant from the hot early universe, received abundant attention. The latter was the more interesting to physicists, because they could set about measuring the radiation spectrum and angular distribution. But the former certainly could interest astronomers, who could have been motivated to more closely examine helium abundances in stars and nebulae, and the processes of stellar helium production and dispersal. Perhaps those who study the sociology of science are best positioned to explore why the presence of more microwave radiation than expected from known radio sources received so much more attention than the presence of more helium than expected from production in known stars. The rest of us might bear in mind that nature is quite capable of surprising us.

7.3 Speculative and programmatic experiments

This history presents us with examples of experiments that were purely speculative, done simply because they were possible; experiments that were inspired by ideas such as those just discussed that were speculative according to accepted ideas but attractive in other philosophies; and experiments that may be termed programmatic: designed to find what was expected from standard and accepted ideas.

A good illustration of the first, speculative, class is the experimental test of equivalence of active and passive gravitational masses (Sect. 6.7). Discovery of a violation would be shocking, but one must consider that during the naissance little was known about empirical gravity physics, and such purely speculative experiments were well worth doing to help improve the situation. It would be difficult now to find support for a more precise laboratory test, which is regrettable; physics should be challenged.

The second class includes the tests for sensitivity of local physics to what the rest of the universe is doing. Some of these experiments were inspired by thoughts of a luminiferous aether of some sort, others by thoughts of Mach's Principle and Dirac's Large Numbers Hypothesis. Explorations of these ideas motivated experiments that have informed us about the nature of gravity. Consider, for example, Dicke's fascination with the idea that the strength of the gravitational interaction may be

standard value for the present matter density this scales to present temperature $T_o = 7$ K, the value in Gamow (1953b).

[20] The situation in 1948 was confused (Peebles 2014). The Alpher and Herman CMB temperature estimate was based on their fit to the broad range of observed cosmic element abundances. This approach was soon found to fail. But Gamow's (1948) estimate of the conditions for light element formation, which are close to the now established cosmology, extrapolate to $T_o = 8$ K for Gamow's assumed present baryon density (Peebles 2014, Eq. (24)).

evolving, and that precision tracking of the orbit of the Moon might reveal the effect. Without Dicke's persistence and influence, would NASA, and its counterpart in the Soviet Union, have gone to the trouble of placing corner reflectors on the Moon? We cannot answer, of course, but we do see how elegant ideas may lead to great results, as in the demanding tests of gravity physics from the Lunar Laser Ranging Experiment.

I place in the programmatic class of experiments during the naissance the measurements of the gravitational redshift and deflection of light, and the search for tensor gravitational waves. This is the delicate art of finding what one expects to find (Hetherington 1980). One might argue that when these experiments were done during the naissance they belonged to the second class, because they were inspired by a theory that was not empirically well supported. The distinction is that the theory was broadly accepted, on nonempirical grounds. But empirical evidence is far better, in the philosophy of science as exploration of the world around us. The programmatic searches for what general relativity theory predicts have been deeply important to this empirical establishment.

As gravity physics grew the distribution of experiments among the three classes evolved, from considerable activity on the purely speculative side during the naissance to the present emphasis on the programmatic side, now that we have a well-established theory that tells us what to look for. This programmatic side is essential to the experimental gravity physics program, but it is not minor pedantry now to pursue experiments designed to be sensitive to departures from the standard model, such as checks of equivalence of the four masses defined in equation (13), or tests of the gravitational inverse square law (e.g. Adelberger et al. 2009), or searches for cosmic evolution of physical parameters such as ratios of elementary particle masses and the strengths of the fundamental interactions. All our physical theories are incomplete, the world is large, and it has ample opportunities to surprise us yet again.

7.4 General purpose and purpose-built instruments

The pendulums usually seen in teaching laboratories for measurements of the acceleration, g, of gravity look very different from the ones illustrated in Figure 8 and used in the Hoffmann (1962) and Curott (1965) experiments to probe possible evolution of g. They look very different again from the sketch in Figure 12 of Faller's (1963) falling corner reflector experiment to measure the absolute value of g. These Princeton experiments, and the two versions of the Eötvös experiment (Liebes 1963; Roll et al. 1964), were designed, or we may say purpose-built, to be optimum for a specific measurement. The Pound and Rebka (1959) laboratory measurement of the gravitational redshift was a purpose-built experiment too, but it was inspired by the appearance of a new tool, the Mössbauer effect, which made the experiment possible. The Princeton experiments certainly made heavy use of new tools as they became available, but many could have been done a decade earlier, though, I am informed, distinctly less well, or might have been done instead a decade later, and better. These are examples of experiments that awaited someone to act on the idea that a closer examination of a particular issue is worth doing, perhaps by design of a purpose-built experiment.

At Gravity Research Group meetings Dicke told us about his preference for purpose-built experiments, and his dislike of sharing raw experimental data that others could reanalyze in foolish ways. The closest he came to Big Science was in the Lunar Laser Ranging Experiment, but it is to be observed that as that project grew he withdrew. The community has adapted to the working conditions of Big Science; shining examples are the precision measurements of the CMB, the precision statistical measures of the natures and distributions of the galaxies, and the LIGO detection of gravitational waves. Dicke perhaps could not have anticipated the power of "data mining" in modern observational surveys that can stimulate thinking about unexamined ideas. But there still is a good case for his preference for an experiment designed for optimum examination of a particular issue. For example, the Sloan Digital Sky Survey (Alam, Albareti, Allende, et al. 2015 and references therein) has made wonderfully broad contributions to our knowledge of the statistical properties of galaxies, but of course it could not have been designed for optimum exploration of all issues. An example is the spiral galaxies in which only a small fraction of the observed stars rise well above the disk (Kormendy, Drory, Bender, and Cornell 2010). Such galaxies are common nearby, and they are a fascinating challenge for theories of galaxy formation based on the ΛCDM cosmology discussed in Section 6.13.3, because the predicted tendency of galaxies to grow by merging would tend to place stars in orbits that rise well above the disk. Further exploration of this interesting phenomenon might best be served by an observational program designed for optimum examination of the properties of this particular class of galaxies.

7.5 Support for curiosity-driven research

The financial support for speculative curiosity-driven research during the naissance in experimental gravity physics is worth recalling, because science and society have changed since then. Research by Dicke and his Gravity Research Group was supported in part by the United States National Science Foundation, and in part also by the United States Army Signal Corps in the late 1950s, by the Office of Naval Research of the United States Navy through the 1960s, and in some papers one also sees acknowledgement of support by the U.S. Atomic Energy Commission. The Hoffmann, Krotkov, and Dicke (1960) proposal for precision tracking of satellites was published in the journal *IRE Transactions on Military Electronics*. The first paper in the issue, by Rear Admiral Rawson Bennett, USN, opened with the sentence "This issue . . . is devoted to the United States Navy's interest and effort in space electronics". The military certainly had reason to be interested in space electronics, but surely had little interest in the Hoffmann et al. proposal to test the idea that the strength of gravity may be changing by about a part in 10^{10} per year. But the military seemed to be comfortable supporting what Dicke and his group wanted to investigate.

This situation was not unusual. High energy physics papers often acknowledged support from military agencies. In experimental gravity physics the development of the maser used in the Cedarholm et al. (1958) aether drift test (Sect. 6.2) was supported "jointly by the Signal Corps, the Office of Naval Research, and the Air Research and Development Command". The Pound and Rebka (1960) laboratory detection of the gravitational redshift (Sect. 6.4) acknowledged support "in part by the joint program of the office of Naval Research and the U.S. Atomic Energy Commission and by a grant from the Higgins Scientific Trust". Irwin Shapiro, then at the MIT Lincoln Laboratory, recalled (Shapiro 2015) that application of his new test of general relativity by planetary radar ranging (Sect. 6.4) required a more powerful transmitter, and that after due deliberation the director of Lincoln Laboratory "called up an Air Force general, who provided Lincoln Laboratory with funding, and asked him for \$500,000, which he then granted, for a new transmitter to enable Lincoln to carry out the experiment". This substantial financial contribution allowed a substantial advance in gravity physics. Wilson and Kaiser (2014) analyzed how the planetary radar experiments by Shapiro and colleagues may have been related to the US military research effort to detect and perhaps somehow learn to deal with an incoming USSR intercontinental ballistic missile. But Shapiro (2015) recalled that the military did not make any attempt to guide directions of the research by him and his colleagues on testing general relativity. That also was the experience in Dicke's Gravity Research Group.

The thinking by military funding agencies at the time may have combined the thoughts that this curiosity-driven research did not cost much, perhaps apart from exceptional cases such as Shapiro's transmitter, could do no harm, might lead to something of eventual value to society and the military, and perhaps also might help the military stay in contact with people whose expertise and advice they might seek on occasion. Weiss added that

> The military was absolutely the most wonderful way to get money. Their mission at that time – and that's something that's grossly misunderstood by all the people that got into trouble with Vietnam and everything else – the military was in the business of training scientists. They wanted not to get caught again the next time there was a [need for a] Manhattan Project or a Rad Lab [MIT Radiation Laboratory].

This permissive attitude of military funding agencies to speculative ideas had the result that in Dicke's group new directions of research often were pursued and the results then reported to the funding agencies, without prior approval of a well-reasoned motivation. This has changed. Compare Dicke's invitation to his graduate student Jim Faller to make an absolute measurement of "little g" (Sect. 6.9) to the recent US National Science Foundation invitation to propose an absolute measurement of "Big G"[21]. This evolution of thinking, from bottom-up toward top-down, is natural in a maturing science, though disturbing from the standpoint of innovation. During the naissance the scant empirical basis for gravity physics made it very appropriate to pursue the speculative curiosity-driven experiments that the military agencies were inclined to support. But the 1970 Mansfield Amendment ended this by prohibiting the Defense Department from funding "any research project or study unless such project or study has a direct and apparent relationship to a specific military function". The last of the Gravity Research Group papers to acknowledge support from the ONR was published in 1972. And the arteries of natural science hardened a little.

7.6 Unintended consequences of curiosity-driven research

Pure curiosity-driven research has had such great unplanned consequences that custom may obscure recognition. So let us remember Dicke's invitation to Jim Faller to measure the acceleration of gravity by dropping a corner reflector (Sect. 6.9), which grew into technology to measure changes in water table levels and the continental rebound from the last Ice Age. Remember also Dicke's interest in probing the physics of temporal and spatial variation of the strength of gravity, and probes for gravitational waves, which led him to invite Bill Hoffmann and David Curott to design pendulums suited for the purpose, and Barry Block, Bob Moore, and Rai Weiss to do the same with spring-type (LaCoste) gravimeters. Section 6.6 recalls how this contributed to the Global Seismographic Network that monitors phenomena of practical interest to us all: earthquakes, tsunamis, storm surges, and underground explosions.

Of course, a good deal of curiosity-driven research serves only to satisfy curiosity. The Global Seismographic Network offers measures of the internal structure of the Earth, which certainly is interesting, though perhaps most satisfying to the curiosity of specialists. John Wheeler was fascinated by "thought experiments" that illustrate the curious apparently acausal nature of quantum physics. He discussed this at an Einstein Centennial Address delivered at the University of Maryland in 1979, on the

[21] http://www.nsf.gov/pubs/2016/nsf16520/nsf16520.htm

centenary of Einstein's birth. Bill Wickes and Carroll Alley were in the audience. Wickes (1972) had completed his thesis with Dicke toward the end of the naissance; Alley's research with Dicke is reviewed in Sections 6.4 and 6.10.2. Wickes (2016) recalled that

> Carroll and I independently realized that we could actually do Wheeler's double-slit delayed choice experiment, using some of the lasers and fast-switching methods that Carroll had pioneered for the lunar ranging work. So we joined forces and recruited Oleg Jakubowicz to do the work for his Ph.D. thesis . . . I like to think of the whole project as an elegant intersection of Wheeler's imagination and Dicke's practical tutelage.

This story of how a Wheeler "thought experiment" became a real experiment is told in Wheeler and Ford (1998, pp. 336–338). Quantum physics is not really acausal: It does not allow you to foresee movements of the stock market, even in principle. But it is satisfying to see real experimental demonstrations of the well-advertised non-intuitive nature of quantum physics, even to non-specialists who take an interest in the world around us. And it is to be observed that many who are not involved in research in natural science find it satisfying to know that Einstein's general relativity theory of gravity, which he discovered a century ago, has been experimentally shown to be a good approximation to how gravity actually operates.

7.7 Establishing and disestablishing elements of natural science

I offer some concluding thoughts about the empirical establishment of general relativity theory discussed in Section 6.14. It is challenging if not impossible to define our working scientific philosophy. This attempt is meant to explain why the empirical case for general relativity theory is to me about as persuasive as it gets in natural science, even though the theory is manifestly incomplete.

Common thinking in physical science, and adopted here, is that a theory is empirically established if it produces substantially more successful predictions than it allows adjustable parameters. The latter must take account of the fact that a sensible theorist will choose to work on the ideas that look most promising from the evidence at hand. Thus tests of gravity physics on the scale of the Hubble length depend on the ΛCDM theory of cosmic structure formation, which grew out of several under discussion in the 1990s (Peebles and Silk 1990). The choice, initially CDM, was informed by the phenomena as well as simplicity. In effect, this was a free parameter to be added to the other adjustments made to fit the theory to the evidence.

Also to be considered is the hazard Hetherington (1980) termed "finding too facilely what they expected to find". Hetherington's example is the incorrect observational confirmation of an incorrect prediction of the gravitational redshift of the white dwarf star Sirius B (reviewed in Sect. 6.4). Another example is Dawid (2015) comment that Einstein's general relativity "had been confirmed by Eddington's measurement of starlight bending in 1919". The measurement was greeted by the media as a great triumph for Einstein's theory, which perhaps is what Dawid had in mind. But the measurement was dubious, and if correct it only added "very flimsy evidence on which to hang a theory" (Dicke 1957a, as discussed in Sect. 3). A third example is the BICEP2 measurement of the pattern of polarization of the CMB. The measurement was rightly welcomed as an important experimental advance. The initial interpretation of the BICEP2 measurement was that it had detected the effect of gravitational waves to be expected if inflation in the very early universe set our initial conditions, and if the expansion rate during inflation had been rapid enough to have produced waves of the proposed amplitude. This was greeted initially as a great triumph, a demonstration that inflation really happened. But it was soon understood that the experiment was not a credible detection, but rather an upper bound on the possible effect of inflation, if inflation actually happened.

There could be undetected examples of misleading "finding the expected" in the tests of gravity physics summarized in Section 6.13.3. Consider for example the SNeIa redshift-magnitude relation marked in Figure 14. It was a brilliant completion of Sandage's (1961) great goal for a test of cosmology. But if this had been the only evidence for the detection of Einstein's cosmological constant, Λ, the community would have had a choice: accept Λ in the standard model, or argue that that the evidence of its detection is misleading because, despite very careful tests that argue to the contrary, SNeIa at high redshift had lower luminosities than apparently identical ones nearby. In this hypothetical situation, with no other relevant evidence, I expect the community would have argued for the latter, because a nonzero value of Λ is awkward from a theoretical point of view, and it destroys the elegant simplicity of the Einstein-de Sitter model that was so influential in the 1990s. This would have been an example of willful disregard of the unexpected. In fact this evidence for Λ was not disregarded, perhaps in part because the community was too sensible, but certainly also because the CMB anisotropy measurements, with the astronomers' value for the extragalactic distance scale, H_o^{-1}, also were seen to be pointing to nonzero Λ (the situation was reviewed in Peebles et al. 2009, pp. 447–477). And, at the time, measurements of stellar evolution ages with the astronomers H_o^{-1} were found to be difficult to understand in the absence of a nonzero Λ. This litany of crosschecks for other tests may be continued at considerable length. And as discussed in Section 6.13.3, it is this abundance of crosschecks that argues against excessive willful misinterpretation of the evidence, and empirically establishes general relativity theory as a good approximation to reality.

The result is not what Dicke anticipated in the mid-1950s when he decided to remedy the scant attention to experimental gravity physics. I do not recall his mentioning the tightening evidence against the scalar-tensor theory; Parkinson's disease might have slowed his recognition of these developments. But his big questions remain. Is local physics really not related to the rest of the universe, apart from the tantalizing effect on inertial motion? Is the enormous difference of strengths of gravity and electromagnetism really only a result of anthropic selection? Is the classical theory of gravity really in a satisfactory state?

7.8 Future developments

Among open issues in gravity physics, the one most immediately relevant to the empirical theme of this paper may be the question of what happened in the very early universe, before classical general relativity theory could have been a good approximation. It may be instructive to compare thinking now and in the 1950s. General relativity theory then was widely accepted as logically compelling. Now the inflation picture for the very early universe is accepted by many to be promising, even logically compelling. There is the serious difference that general relativity is a theory, while inflation is a framework on which to hang a theory. But there is a serious similarity. In the 1950s the empirical basis for general relativity was generally considered to be necessarily schematic, because better experiments were not feasible. Now the empirical basis for inflation, or other ideas about the very early universe, is considered to be necessarily schematic, because better experiments are not feasible. The community was surprised by the abundance of evidence that has grown out of the naissance of experimental gravity physics. If the community will not be surprised by another harvest of tests it will have to be content with evidence from elegance and logical completeness. But the communities of natural science have been surprised by many great

Table A.1. Graduate student research with Dicke, pre-gravity.

Name	Research
Alexander Pond	A Experimental Investigation of Positronium (1952)
George Newell	A Method for Reducing the Doppler Width of Microwave Spectrum Lines (1953)
Bruce Hawkins	The Orientation and Alignment of Sodium Atoms by Means of Polarized Resonance Radiation (1954)
Robert Romer	A Method for the Reduction of the Doppler Width of Microwave Spectral Lines (1955)
James Wittke	A Redetermination of the Hyperfine Splitting in the Ground State of Atomic Hydrogen (1955)
Christopher Sherman	Nuclear Induction with Separate Regions of Excitation and Detection (1955)
Lowell White	The Gyromagnetic Ratio of the Electron in the Metastable State of Hydrogen (1956)
Peter Bender	The Effect of a Buffer Gas on the Optical Orientation Process in Sodium Vapor (1956)
Edward Lambe	A Measurement of the g-Value of the Electron in the Ground State of the Hydrogen Atom (1959)

empirical advances. This experience suggests there may be more surprising empirical developments to come, perhaps even some that postpone the need for nonempirical assessment of some future deeper theory of the nature of gravity and how the world began.

Acknowledgements. I have greatly benefitted from recollections, guidance to the literature, and advice on the preparation of this paper, from Eric Adelberger, Pete Bender, Jon Berger, Bart Bernstein, Steve Boughn, Paul Boynton, Richard Dawid, Dieter Brill, David Curott, Jim Faller, Masataka Fukugita, Richard Garwin, Henry Hill, Bill Hoffmann, David Kaiser, Helge Kragh, Bob Krotkov, Jeff Kuhn, Adele La Rana, Sid Liebes, Ed McDonald, Martin Mchugh, Charlie Misner, Lyman Page, Bruce Partridge, Peter Roll, Irwin Shapiro, Joe Taylor, Scott Tremaine, Virginia Trimble, Rainer Weiss, Bill Wickes, and Clifford Will. I must make special mention of Masataka Fukugita, Bill Hoffmann, Bruce Partridge, Irwin Shapiro, Virginia Trimble, Rainer Weiss, and the editors of this journal, for their careful readings and annotated commentaries on how to improve the paper; Dieter Brill, for his notes of what was said at Gravity Group meetings; Sid Liebes, who made the drawing of the interferometer on the left side of Figure 7; and Bill Hoffmann, who had in his possession the drawing of the interferometer in Figure 12, and made the first draft of the tables in Appendix A. I must also recognize with gratitude Linda Chamberlin's help in guiding me through the vast but not always well organized resources of the Princeton University Library. This research was supported through the good offices of the Department of Physics, Princeton University.

Appendix A: Dicke's gravity research group

Examples of research in Dicke's Gravity Research Group were discussed in Sections 5 and 6. This Appendix is meant to give a broader picture by listing all of Dicke's Ph.D. graduate students and all the post-Ph.D. members of his Gravity Research Group. The records I have may be incomplete; I would be grateful for information about anyone I have overlooked or misidentified. The dissertation titles serve to illustrate Dicke's quite abrupt mid-career change of direction of research.

Table A.1 lists the pre-gravity Ph.D. theses Dicke directed, with title and year of acceptance of the thesis by the Princeton Department of Physics. The compilation of Table A.2, on research in the Gravity Research Group, required somewhat more

Table A.2. Research with Dicke in the Gravity Group.

Name	Status	Research
Robert Krotkov	post-Ph.D.	Comparison between theory and observation for the outer planets (Krotkov and Dicke 1959)
Carl Brans	Ph.D.	Mach's Principle & Varying Gravitational Constant (1961)
James Peebles	Ph.D.	Observational Tests and Theoretical Problems with Variable Strength of the Electromagnetic Interaction (1961)
Carroll Alley	Ph.D.	Optical Pumping and Optical Detection Involving Microwave & Radio Frequency Coherence Effects (1962)
William Hoffmann	Ph.D.	A Pendulum Gravimeter for Measurement of Periodic Annual Variations in the Gravitational Constant (1962)
Kenneth Turner	Ph.D.	New Limit on Velocity Dependent Interaction Between Natural Clocks and Distant Matter (1962)
James Brault	Ph.D.	The Gravitational Red Shift in the Solar Spectrum (1962)
Dieter Brill	post-Ph.D.	Experiments on Gravitation (Bertotti, Brill, Krotkov 1962)
James Faller	Ph.D.	An Absolute Interferometric Determination of the Acceleration of Gravity (1963)
John Stoner	Ph.D.	Production of narrow balmer spectrum lines in an electron-bombarded atomic hydrogen beam (1963)
Henry Hill	post-Ph.D.	Experimental Limit on Velocity-Dependent Interactions of Clocks and Distant Matter (Turner and Hill 1964)
Sidney Liebes	post-Ph.D.	Gravitational Lenses (Liebes 1964)
Curtis Callan	Ph.D.	Sphericaly Symmetric Cosmological models (1964)
Lawrence Jordan	Ph.D.	The velocities of 4 BeV/c pions and 8 BeV/c pions, kaons, and protons (1964)
Jason Morgan	Ph.D.	An Astronomical and Geophysical Search for Scalar Gravitational Waves (1964)
William Hildreth	Ph.D.	The interaction of scalar gravitational waves with the Schwarzschild Metric (1964)
Peter Roll	post-Ph.D.	The equivalence of inertial and passive gravitational mass (Roll, Krotkov, Dicke 1964)
Rainer Weiss	post-Ph.D.	A Gravimeter to Monitor the ${}_0S_0$ Dilational Mode of the Earth (Weiss and Block 1965)
David Curott	Ph.D.	A Pendulum Gravimeter for Precision Detection of Scalar Gravitational Radiation (1965)
David Wilkinson	post-Ph.D.	Cosmic Black-Body Radiation (Dicke, Peebles, Roll, and Wilkinson 1965)
Robert Moore	Ph.D.	Study of Low Frequency Earth Noise and New Upper Limit to the Intensity of Scalar Gravitational Waves (1966)
Lloyd Kreuzer	Ph.D.	The Equivalence of Active and Gravitational Mass (1966)
Barry Block	post-Ph.D.	Measurements in Earth mode frequency, electrostatic sensing & feedback gravimeter (Block and Moore 1966)
Mark Goldenberg	post-Ph.D.	Solar Oblateness and General Relativity (Dicke and Goldenberg 1967)
Carl Zanoni	Ph.D.	Development of Daytime Astrometry to Measure the Gravitational Deflection of Light (1967)
Dennis Heygi	Ph.D.	The Primordial Helium Abundance as Determined from the Binary Star System μ Cassiopeiae (1968)
B. Edward McDonald	Ph.D.	Meridian Circulation in Rotating Stars (1970)
Lawrence Cathles	Ph.D.	The viscosity of the Earth's mantle (1971)
William Wickes	Ph.D.	Primordial Helium Abundance and Population-II Binary Stars: Measurement Technique (1972)
Jeffrey Kuhn	Ph.D.	Global scale photospheric velocity fields: Probes of the solar interior (1980)
Ken Libbrecht	Ph.D.	The shape of the Sun (1984)

creative rules. A graduate student who completed a Ph.D. under Dicke's direction is marked as Ph.D. in the second column, and the third column gives the thesis title and date of acceptance. Some graduate students published papers in addition to their dissertations while they were group members, and some stayed in the group as postdocs after completion of the Ph.D., but none of this information is entered. Those who joined the group as a postdoc, instructor, or assistant professor are marked as post-Ph.D.s in the second column. The third column in this case lists the title and reference for the first publication reporting research done in the group (but excluding multiple-author papers that are already in the table). Some titles are shortened to fit. A more complete picture would list visitors, some of whom stayed for a day or two, perhaps to present a seminar, and others stayed longer, but I do know of reasonably complete records. Dicke's publications with colleagues outside the group, and his single-author papers, are not tabulated, but almost all of his papers on gravity are discussed in the text. Apart from special cases to be noted the undergraduate papers and theses Dicke directed are not entered.

Some entries in Table A.2 require special explanation. Dicke advised Lawrence Cathles's 1965 Princeton undergraduate paper on *The physics of glacial uplift*. Cathles wrote his Ph.D. thesis on deeper exploration of the same subject, in the Princeton Department of Geological and Geophysical Sciences, under the direction of Jason Morgan, who wrote his Ph.D. thesis under Dicke's direction in the Department of Physics. While Henry Hill was in the gravity group he directed Carl Zanoni's 1957 Ph.D. research, which is entered. Curtis Callan was most closely associated with Sam Treiman's research in elementary particle theory, but Dicke proposed and supervised research on his dissertation. Wheeler supervised Dieter Brill's 1959 Ph.D. thesis, but Brill often attended Gravity Group meetings, and he wrote a 1962 review of experimental tests of relativity with Bob Krotkov, who was in the Gravity Group, and Bruno Bertotti, who was visiting the group.

References

Adams, W.S. 1925. The relativity displacement of the spectral lines in the companion of Sirius. *The Observatory* **48**: 337-342

Adelberger, E.G. 2015. Private communication

Adelberger, E.G., J.H. Gundlach, B.R. Heckel, S. Hoedl and S. Schlamminger. 2009. Torsion balance experiments: A low-energy frontier of particle physics. *Progress Particle Nucl. Phys.* **62**: 102-134

Ageno, M. and E. Alamdi. 1966. Experimental Search for a Possible Change of the β Decay Constant with Centrifugal Force. *Atti Della Accademia Nazionale Dei Lincei* Serie 8 **8**: 35pp

Aglietta, M., G. Badino, G. Bologna, et al. 1989. Analysis of the data recorded by the Mont Blanc neutrino detector and by the Maryland and Rome gravitational-wave detectors during SN 1987 A. *Nuovo Cimento C Geophysics Space Physics C* **12**: 75-101

Aguiar, O.D. 2011. Past, present and future of the Resonant-Mass gravitational wave detectors. *Res. Astron. Astrophys.* **11**: 1-42

Alam, S., F.D. Albareti, C. Allende Prieto, et al. 2015. The Eleventh and Twelfth Data Releases of the Sloan Digital Sky Survey: Final Data from SDSS-III. *Astrophys. J. Suppl.* **219**: 12, 27 pp.

Albareti, F.D., J. Comparat, C.M. Gutiérrez, et al. 2015. Constraint on the time variation of the fine-structure constant with the SDSS-III/BOSS DR12 quasar sample. *Month. Not. Roy. Astron. Soc.* **452**: 4153-4168

Alcock, C., R.A. Allsman, D.R. Alves, et al. 2000. The MACHO Project: Microlensing Results from 5.7 Years of Large Magellanic Cloud Observations. *Astrophys. J.* **542**: 281-307

Alley, C.O., Jr. 1962. *Optical Pumping and Optical Detection Involving Microwave and Radio Frequency Coherence Effects*. Ph.D. Thesis, Princeton University

Alley, C.O. 1972. Story of the Development of the Apollo 11 Laser Ranging Retro-Reflector Experiment: One Researcher's Personal Account. *Adventures in Experimental Physics* Ed. B. Maglich, World Science Communications, Princeton N.J. pp. 132-156

Alley, C.O., P.L. Bender, R.H. Dicke, et al. 1965. Optical Radar Using a Corner Reflector on the Moon. *J. Geophys. Res.* **70**: 2267-2269

Alpher, R.A. 1948. *On the Origin and Relative Abundance of the Elements*. Ph.D. thesis, The George Washington University

Alpher, R.A. and R. Herman. 1948. Evolution of the Universe. *Nature* **162**: 774-775

Alpher, R.A., J.W. Follin and R.C. Herman. 1953. Physical Conditions in the Initial Stages of the Expanding Universe. *Phys. Rev.* **92**: 1347-1361

Arnowitt, R., S. Deser and C.W. Misner. 1960. Finite Self-Energy of Classical Point Particles. *Phys. Rev. Lett.* **4**: 375-377

Baade, W. 1956. The Period-Luminosity Relation of the Cepheids. *Publications of the Astronomical Society of the Pacific* **68**: 5-16

Babcock, H.W. 1939. The rotation of the Andromeda Nebula. *Lick Observatory Bulletin* **19**: 41-51

Bahcall, N.A. and R. Cen. 1992. *Astrophys. J. Lett.* **398**: L81-L84

Bahcall, N.A. and A. Kulier. 2014. Tracing mass and light in the Universe: where is the dark matter? *Month. Not. Roy. Astron. Soc.* **439**: 2505-2514

Bai, Y., J. Salvado and B.A. Stefanek. 2015. Cosmological constraints on the gravitational interactions of matter and dark matter. *J. Cosmology Astroparticle Phys.* issue 10, article id. 029, 22 p.

Barbour, J.B. and H. Pfister. 1995. *Mach's Principle: From Newton's Bucket to Quantum Gravity*. Birkhäuser, Boston

Barstow, M.A., H.E. Bond, J.B. Holberg, et al. 2005. Hubble Space Telescope spectroscopy of the Balmer lines in Sirius B. *Month. Not. Roy. Astron. Soc.* **362**: 1134-1142

Bartlett, D.F. and D. van Buren. 1986. Equivalence of active and passive gravitational mass using the moon. *Phys. Rev. Lett.* **57**: 21-24

Beltran-Lopez, V. 1962. *Part I. Microwave Zeeman Spectrum of Atomic Chlorine. Part II. Measurements on Anisotropy of Inertial Mass*. Ph.D. Thesis, Yale University

Bender, P.L. 2015. Private communication

Berger, J. 2016. Private communication

Bertotti, B., D. Brill and R. Krotkov. 1962. Experiments on Gravitation. In *Gravitation: An Introduction to Current Research*, edited by L. Witten, Wiley, New York, pp. 1-48

Bertotti, B., L. Iess and P. Tortora. 2003. A test of general relativity using radio links with the Cassini spacecraft. *Nature* **425**: 374-376

BICEP2 Collaboration, P.A.R. Ade, R.W. Aikin, et al. 2014. Detection of B-Mode Polarization at Degree Angular Scales by BICEP2. *Phys. Rev. Lett.* **112**: 241101, 25pp.

Blatt, J.M. and V.F. Weisskopf. 1952. *Theoretical Nuclear Physics*. Wiley, New York

Block, B. and R.D. Moore. 1966. Measurements in the Earth mode frequency range by an electrostatic sensing and feedback gravimeter. *J. Geophys. Res.* **71**: 4361-4375

Blum, A., R. Lalli and J. Renn. 2015. The Reinvention of General Relativity: A Historiographical Framework for Assessing One Hundred Years of Curved Space-time. *Isis* **106**: 598-620

Bolton, C.T. 1972. Identification of Cygnus X-1 with HDE 226868. *Nature* **235**: 271-273

Bondi, H. 1952. *Cosmology*. Cambridge, Cambridge University Press

Bondi, H. 1957. Negative Mass in General Relativity. *Rev. Mod. Phys.* **29**: 423-428

Bondi, H. 1960. *Cosmology*. Cambridge, Cambridge University Press, second edition

Bondi, H. 1962. On the physical characteristics of gravitational waves. In Lichnerowicz and Tonnelat (1962), pp. 129-125

Bondi, H. and T. Gold. 1948. The Steady-State Theory of the Expanding Universe. *Month. Not. Roy. Astron. Soc.* **108**: 252-270

Botermann, B., D. Bing, C. Geppert, et al. 2014. Test of Time Dilation Using Stored Li^+ Ions as Clocks at Relativistic Speed. *Phys. Rev. Lett.* **113**: 120405, 5 p.

Boughn, S.P., S.J. Vanhook and C.M. O'Neill. 1990. Observational limits on a millihertz stochastic background of gravitational radiation. *Astrophys. J.* **354**: 406-410

Bowyer, S., E.T. Byram, T.A. Chubb and H. Friedman. 1965. Cosmic X-ray Sources. *Science* **147**: 394-398

Boynton, P.E., R.A. Stokes and D.T. Wilkinson. 1968. Primeval fireball intensity at $\lambda =$ 3.3 mm. *Phys. Rev. Lett.* **21**: 462-465

Bracewell, R.N. 1959. Ed. *Paris Symposium on Radio Astronomy*. Stanford University Press, Stanford, USA

Braginskiĭ, V.B. and V.I. Panov. 1971. Verification of the Equivalence of Inertial and Gravitational Mass. *Zhurnal Eksperimental'noi i Teoreticheskoi Fizik* **61**: 873-879; English translation in *Soviet J. Exp. Theor. Phys.* **34**: 463-466

Brans, C.H. 1961. *Mach's Principle and a Varying Gravitational Constant*, Ph.D. Thesis, Princeton University

Brans, C.H. 2008. Scalar-tensor Theories of Gravity: Some personal history. *Am. Inst. Phys. Conf. Ser.* **1083**: 34-46

Brans, C.H. 2016. 65 Years in and Around Relativity. In *At the Frontier of Spacetime: Scalar-Tensor Theory, Bell's Inequality, Machs Principle, Exotic Smoothness*, edited by T. Asselmeyer-Maluga. Springer

Brans, C., and R.H. Dicke. 1961. Mach's Principle and a Relativistic Theory of Gravitation. *Phys. Rev.* **124**: 925-935

Brault, J.W. 1962. *The Gravitational Red Shift in the Solar Spectrum*. Ph.D. Thesis, Princeton University

Brill, D.R. and J.M. Cohen. 1966. Rotating Masses and Their Effect on Inertial Frames. *Phys. Rev.* **143**: 1011-1015

Brown, M.E. and P.J.E. Peebles. 1987. The local extragalactic velocity field, the local mean mass density, and biased galaxy formation. *Astrophys. J.* **317**: 588-592

Burbidge, G.R. 1958. Nuclear Energy Generation and Dissipation in Galaxies. *Publ. Astron. Soc. Pacific* **70**: 83-89

Burbidge, G.R. 1959. The theoretical explanation of radio emission, in Bracewell (1959), pp. 541-551

Burke, B.F. 2009. Radio astronomy from first contacts to the CMBR. In Peebles, Page, and Partridge (2009), pp. 176-183

Cedarholm, J.P. and C.H. Townes. 1959. A New Experimental Test of Special Relativity. *Nature* **184**: 1350-1351

Cedarholm, J.P., G.F. Bland, B.L. Havens and C.H. Townes. 1958. New Experimental Test of Special Relativity. *Phys. Rev. Lett.* **1**: 342-343

Champeney, D.C., G.R. Isaak and A.M. Khan. 1963. An 'aether drift' experiment based on the Mössbauer effect. *Phys. Lett.* **7**: 241-243

Chase, C.T. 1926. A Repetition of the Trouton-Noble Ether Drift Experiment. *Phys. Rev.* **28**: 378-383

Chernin, A.D. 1994. FROM THE HISTORY OF PHYSICS: How Gamow calculated the temperature of the background radiation or a few words about the fine art of theoretical physics. *Phys. Usp.* **37**: 813-820

Ciufolini, I. and E.C. Pavlis. 2004. A confirmation of the general relativistic prediction of the Lense-Thirring effect. *Nature* **431**: 958-960

Cocconi, G. and E.E. Salpeter. 1958. A Search for Anisotropy of Inertia. *Il Nuovo cimento* **10**: 646-651

Cocconi, G. and E.E. Salpeter. 1960. Upper Limit for the Anisotropy of Inertia from the Mössbauer Effect. *Phys. Rev. Lett.* **4**: 176-177

Curott, D.R.F. 1965. *A Pendulum Gravimeter for Precision Detection of Scalar Gravitational Radiation*. Ph.D. Thesis, Princeton University

Curott, D.R. 1966. Earth deceleration from ancient solar eclipses. *Astron. J.* **71**: 264-269

Curott, D.R. 2015. Private communication

Damour, T. and F. Dyson. 1996. The Oklo bound on the time variation of the fine-structure constant revisited. *Nucl. Phys. B* **480**: 37-54

Davis, M., J. Huchra, D.W. Latham and J. Tonry. 1982. A survey of galaxy redshifts. II – The large scale space distribution. *Astrophys. J.* **253**: 423-445

Davis, M., G. Efstathiou, C.S. Frenk and White, S.D.M. 1985. The evolution of large-scale structure in a universe dominated by cold dark matter. *Astrophys. J.* **292**: 371-394

Davis, M. and P.J.E. Peebles. 1983. A survey of galaxy redshifts. V – The two-point position and velocity correlations. *Astrophys. J.* **267**: 465-482

Dawid, R. 2015. *String Theory and the Scientific Method*. Cambridge University Press, Cambridge

Dawid, R. 2016. Private communication

DeGrasse, R.W., D.C. Hogg, E.A. Ohm and H.E.D. Scovil. 1959. Ultra-Low-Noise Measurements Using a Horn Reflector Antenna and a Traveling-Wave Maser. *J. Appl. Phys. Lett.* **30**: 2013

Deser, S. 1957. General Relativity and the Divergence Problem in Quantum Field Theory. *Rev. Mod. Phys.* **29**: 417-423

de Vaucouleurs, G. 1970. The Case for a Hierarchical Cosmology. *Science* **167**: 1203-1213

de Vaucouleurs, G. 1982. Five crucial tests of the cosmic distance scale using the Galaxy as fundamental standard. *Nature* **299**: 303-307

DeWitt, C.M. 1957, Ed. *Conference on the Role of Gravitation in Physics*. Wright Air Development Center Technical Report 57-216; Springfield, Carpenter Lithography and Printing

DeWitt, C.M. and D. Rickles. 2011, Eds. *The Role of Gravitation in Physics: Report from the 1957 Chapel Hill Conference.* Berlin: Edition Open Access, 2011). See http://www.edition-open-sources.org/media/sources/5/Sources5.pdf

Dicke, R.H. 1946. The Measurement of Thermal Radiation at Microwave Frequencies. *Rev. Sci. Instrum.* **17**: 268-275

Dicke, R.H. 1957a. The Experimental Basis of Einstein's Theory. In DeWitt (1957), pp. 5-12

Dicke, R.H. 1957b. Principle of Equivalence and the Weak Interactions. *Rev. Mod. Phys.* **29**: 355-362

Dicke, R.H. 1957c. Gravitation without a Principle of Equivalence. *Rev. Mod. Phys.* **29**: 363-376

Dicke, R.H. 1959a. Dirac's Cosmology and the Dating of Meteorites. *Nature* **183**: 170-171

Dicke, R.H. 1959b. New Research on Old Gravitation. *Science* **129**: 621-624

Dicke, R.H. 1961a. The Nature of Gravitation. In *Science in Space*, edited by Lloyd V. Berkner and H. Odishaw. McGraw-Hill, New York, pp. 91–118

Dicke, R.H. 1961b. Dirac's Cosmology and Mach's Principle. *Nature* **192**: 440-441

Dicke, R.H. 1962a. Machs principle and equivalence. In *Evidence for Gravitational Theories: Proceedings of Course 20 of the International School of Physics "Enrico Fermi"*, edited by C. Møller, Academic, New York, pp. 1-49

Dicke, R.H. 1962b. The Earth and Cosmology. *Science* **138**: 653-664

Dicke, R.H. 1962c. Mach's Principle and Invariance under Transformation of Units. *Phys. Rev.* **125**: 2163-2167

Dicke, R.H. 1963. Experimental Relativity. In *Relativity, Groups, and Cosmology*, edited by C.M. DeWitt and B. DeWitt. Gordon and Breach, New York, pp. 164-315

Dicke, R.H. 1964. The Sun's Rotation and Relativity. *Nature* **202**: 432-435

Dicke, R.H. 1966. The Secular Acceleration of the Earth's Rotation and Cosmology. In *The Earth-Moon System*, edited by B.G. Marsden and A.G.W. Cameron. Plenum Press, New York, pp. 98-164

Dicke, R.H. 1968. Scalar-Tensor Gravitation and the Cosmic Fireball. *Astrophys. J.* **152**: 1-24

Dicke, R.H. 1969. General relativity: survey and experimental tests. *Contemporary Physics* **1**: 515-531, Proceedings of the International Symposium held at the International Centre for Theoretical Physics, Trieste, 7–28 June, 1968

Dicke, R.H. and H.M. Goldenberg. 1967. Solar Oblateness and General Relativity. *Phys. Rev. Lett.* **18**: 313-316

Dicke, R.H. and P.J.E. Peebles. 1979. The big bang cosmology – enigmas and nostrums. In *General Relativity: An Einstein Centenary Survey*, edited by S.W. Hawking and W.I. Israel. Cambridge University Press, Cambridge, pp. 504-517

Dicke, R.H., R. Beringer, R.L. Kyhl and A.B. Vane. 1946. Atmospheric Absorption Measurements with a Microwave Radiometer. *Phys. Rev.* **70**: 340-348

Dicke, R.H., W.F. Hoffmann and R. Krotkov. 1961. Tracking and Orbit Requirements for Experiment to Test Variations in Gravitational Constant. In *Space Research II*, edited by H.C. van de Hulst, C. de Jager and A.F. Moore. North-Holland, Amsterdam, pp. 287-291

Dicke, R.H., P.J.E. Peebles, P.G. Roll and D.T. Wilkinson. 1965. Cosmic Black-Body Radiation. *Astrophys. J.* **142**: 414-419

Dirac, P.A.M. 1937. The Cosmological Constants. *Nature* **139**: 323

Drever, R.W.P. 1960. Upper limit to anisotropy of inertial mass from nuclear resonance. *Philos. Mag.* **5**: 409-411

Drever, R.W.P. 1961. A search for anisotropy of inertial mass using a free precession technique. *Philos. Mag.* **6**: 683-687

Dyson, F.W., A.S. Eddington and C. Davidson. 1920. A Determination of the Deflection of Light by the Sun's Gravitational Field, from Observations Made at the Total Eclipse of May 29, 1919. *Philos. Trans. Roy. Soc. London Ser. A* **220**: 291-333

Eddington, A.S. 1936. *Relativity Theory of Protons and Electrons.* Cambridge University Press, Cambridge

Einstein, A. 1917. Cosmological Considerations on the General Theory of Relativity. *S.-B. Preuss. Akad. Wiss.* 142-152

Einstein, A. 1923. *The Meaning of Relativity.* Princeton University Press, Princeton

Einstein, A. 1936. Lens-Like Action of a Star by the Deviation of Light in the Gravitational Field. *Science* **84**: 506-507

Einstein, A. 1945. *The Meaning of Relativity.* Princeton University Press, Princeton, second edition

Einstein, A. and N. Rosen. 1937. On Gravitational Waves. *J. Franklin Institute* **223**: 43-53

Eötvös, R.V., D. Pekár and E. Fekete. 1922. Beiträge zum Gesetze der Proportionalität von Trägheit und Gravität. *Ann. Phys.* **373**: 11-66

Everitt, C.W.F., B. Muhlfelder, D.B. DeBra, et al. 2015. The Gravity Probe B test of general relativity. *Classical and Quantum Gravity* **32**: 224001, 29 p.

Faller, J.E. 1963. *An Absolute Interferometric Determination of the Acceleration of Gravity.* Ph.D. Thesis, Princeton University

Faller, J.E. 2014a. Precision measurement, scientific personalities and error budgets: the sine quibus non for big G determinations. *Philos. Trans. Roy. Soc. A* **372**, 20140023, 18 p.

Faller, J.E. 2014b. In http://cddis.gsfc.nasa.gov/lw19/Program/index.html#sess2, 15 p.

Faller, J.E. 2015. Private communication.

Faller, J.E. and J. Hammond. 1967. Laser-interferometer determination of the acceleration of gravity. *IEEE J. Quantum Electron.* **3**: 266-267

Faller, J.E., I. Winer, W. Carrion, et al. 1969. Laser Beam Directed at the Lunar Retro-Reflector Array: Observations of the First Returns. *Science* **166**: 99-102

Field, G.B., G.H. Herbig and J. Hitchcock. 1966. Radiation Temperature of Space at λ2.6 mm. *Astron. J.* **71**: 161

Fierz, M. 1956. Über die physikalische Deutung der erweiterten Gravitationstheorie P. Jordan's. *Helvetica Physica Acta* **29**: 128-134

Finzi, A. 1962. Test of Possible Variations of the Gravitational Constant by the Observation of White Dwarfs within Galactic Clusters. *Phys. Rev.* **128**: 2012-2015

Forward, R.L., D. Zipoy, J. Weber, et al. 1961. Upper Limit for Interstellar Millicycle Gravitational Radiation. *Nature* **189**: 473

Gamow, G. 1948. The Origin of Elements and the Separation of Galaxies. *Phys. Rev.* **74**: 505-506

Gamow, G. 1949. On Relativistic Cosmogony. *Rev. Mod. Phys.* **21**: 367-373

Gamow, G. 1952. The creation of the universe. Viking Press, New York, 147 p.

Gamow, G. 1953a. Lectures. In *Symposium on Astrophysics.* University of Michigan, Ann Arbor, June 29 to July 24, pp. 1–30

Gamow, G. 1953b. Expanding Universe and the Origin of Galaxies. *Dan. Mat. Gys. Medd* **27**, 10, 15pp

Gamow, G. 1954. On the steady-state theory of the universe. *Astron. J.* **59**: 200

Gamow, G. 1956. The physics of the expanding universe. *Vistas in Astronomy* **2**: 1726-1732
Geller, M.J. and P.J.E. Peebles, 1972. Test of the Expanding Universe Postulate. *Astrophys. J.* **174**: 1-5
Giganti, J.J., J.V. Larson, J.P. Richard and J. Weber. 1973. Apollo 17: Preliminary Science Report SP-330: pp. 12.1–12.4
Goenner, H. 2012. Some remarks on the genesis of scalar-tensor theories. *General Relativity and Gravitation* **44**: 2077-2097
Gold, T. 1968. Rotating Neutron Stars as the Origin of the Pulsating Radio Sources. *Nature* **218**: 731-732
Goldenberg, H.M. 1961. *The Atomic Hydrogen Maser*. Ph.D. Thesis, Harvard University
Goles, G.G., R.A. Fish and E. Anders. 1960. The record in the meteorites - I. The former environment of stone meteorites as deduced from K 40-Ar 40 ages. *Geochimica et Cosmochimica Acta* **19**: 177-195
Gordon, J.P., H.J. Zeiger and C.H. Townes. 1955. The Maser-New Type of Microwave Amplifier, Frequency Standard, and Spectrometer. *Phys. Rev.* **99**: 1264-1274
Greenstein, J.L. and V.L. Trimble. 1967. The Einstein Redshift in White Dwarfs. *Astrophys. J.* **149**: 283-298
Greenstein, J.L., J.B. Oke and H.L. Shipman. 1971. Effective Temperature, Radius, and Gravitational Redshift of Sirius B. *Astrophys. J.* **169**: 563-566
Greenstein, J.L., J.B. Oke and H. Shipman. 1985. On the redshift of Sirius B. *Roy. Astronom. Soc. Quarterly J.* **26**: 279-288
Gush, H.P., M. Halpern and E.H. Wishnow. 1990. Rocket measurement of the cosmic-background-radiation mm-wave spectrum. *Phys. Rev. Lett.* **65**: 537-540
Happer, W., P.J.E. Peebles and D.T. Wilkinson. 1999. Robert Henry Dicke. *Biographical Memoirs, National Academy of Sciences* **77**: 1-18
Hauser, M.G., R.G. Arendt, T. Kelsall, et al. 1998. The COBE Diffuse Infrared Background Experiment Search for the Cosmic Infrared Background. I. Limits and Detections. *Astrophys. J.* **508**: 25-43
Hay, H.J., J.P. Schiffer, T.E. Cranshaw and P.A. Egelstaff. 1960. Measurement of the Red Shift in an Accelerated System Using the Mössbauer Effect in Fe^{57}. *Phys. Rev. Lett.* **4**: 165-166
Hetherington, N.S. 1980. Sirius B and the Gravitational Redshift: An Historical Review. *Roy. Astron. Soc. Quarterly J.* **21**: 246-252
Heyl, P.R. 1930. A Redetermination of the Constant of Gravitation. *Bur. Stand. J. Res.* **5**: 1243-1290
Hill, H.A. and R.T. Stebbins. 1975. The intrinsic visual oblateness of the sun. *Astrophys. J.* **200**: 471-475
Hill, H.A., P.D. Clayton, D.L. Patz, et al. 1974. Solar Oblateness, Excess Brightness, and Relativity. *Phys. Rev. Lett.* **33**: 1497-500
Hoekstra, H., M. Bartelmann, H. Dahle, et al. 2013. Masses of Galaxy Clusters from Gravitational Lensing. *Space Sci. Rev.* **177**: 75-118
Hoffmann, W.F. 1962. *A Pendulum Gravimeter for Measurement of Periodic Annual Variations in the Gravitational Constant*. Ph.D. Thesis, Princeton University
Hoffmann, W.F. 2016. Private communication
Hoffmann, W.F., R. Krotkov and R.H. Dicke. 1960. Precision Optical Tracking of Artificial Satellites. *IRE Transactions on Military Electronics* **4**: 28-37
Hogg, D.C. 2009. Early Low-Noise and Related Studies at Bell Laboratories, Holmdel, N.J. In Peebles, Page, and Partridge (2009), pp. 70–73
Hoyle, F. 1981. The Big Bang in Astronomy. *New Scientist* **92**: 521-524
Hoyle, F. 1948. A New Model for the Expanding Universe. *Month. Not. Royal Astron. Soc.* **108**: 372-382
Hoyle, F., and R.J. Tayler. 1964. The Mystery of the Cosmic Helium Abundance. *Nature* **203**: 1108-1110
Hubble, E. 1929. A Relation between Distance and Radial Velocity among Extra-Galactic Nebulae. *Proc. Natl. Acad. Sci.* **15**: 168-173
Hubble, E. 1936. *The Realm of the Nebulae*. Yale University Press, New Haven

Hubble, E. and M.L. Humason. 1931. The Velocity-Distance Relation among Extra-Galactic Nebulae. *Astrophys. J.* **74**: 43-80

Hughes, V.W., H.G. Robinson and V. Beltran-Lopez. 1960. Upper Limit for the Anisotropy of Inertial Mass from Nuclear Resonance Experiments. *Phys. Rev. Lett.* **4**: 342-344

Hulse, R.A. and J.H. Taylor. 1975. Discovery of a pulsar in a binary system. *Astrophys. J. Lett.* **195**: L51-L53

Humason, M.L., N.U. Mayall and A.R. Sandage. 1956. Redshifts and magnitudes of extragalactic nebulae. *Astron. J.* **61**: 97-162

Infeld, L. 1964. Ed. *Conférence internationale sur les theéories relativiste de la gravitation*. Pergamon Press, Oxford

Ives, H.E. and G.R. Stilwell. 1938. An Experimental study of the rate of a moving atomic clock. *J. Opt. Soc. Am.* **28**: 215-226

Jaseja, T.S., A. Javan, J. Murray and C.H. Townes. 1964. Test of Special Relativity or of the Isotropy of Space by Use of Infrared Masers. *Phys. Rev.* **133**: 1221-1225

Jentschel, M., J. Krempel and P. Mutti. 2009. A validity test of $E = mc^2$. *Eur. Phys. J. Special Topics* **172**, 353-362

Jones, B.J.T., V.J. Martínez, E. Saar and V. Trimble. 2004. Scaling laws in the distribution of galaxies. *Rev. Mod. Phys.* **76**: 1211-1266

Jordan, P. 1937. Die physikalischen Weltkonstanten. *Naturwissenschaften* **25**: 513-517

Jordan, P. 1948. Fünfdimensionale Kosmologie. *Astron. Nachr.* **276**: 193-208

Jordan, P. 1949. Formation of the Stars and Development of the Universe. *Nature* **164**: 637-640

Jordan, P. 1952. *Schwerkraft und Weltall*. Braunschweig, Vieweg

Jordan, P. 1966. *Die Expansion der Erde*. Braunschweig, Vieweg

Jordan, P. 1971. *The Expanding Earth*. Pergammon Press, New York

Jordan, P. and C. Möller. 1947. Über die Feldgleichungen der Gravitation bei variabler "Gravitationslonstante". *Zeitschrift für Naturforshung* **2a**: 1-2

Joyce, A., B. Jain, J. Khoury and M. Trodden. 2015. Beyond the cosmological standard model. *Phys. Rep.* **568**: 1-98

Kennedy, R.J. and E.M. Thorndike. 1931. A Search for an Electrostatic Analog to the Gravitational Red Shift. *Proc. Natl. Acad. Sci.* **17**: 620-622

Kennedy, R.J. and E.M. Thorndike. 1932. Experimental Establishment of the Relativity of Time. *Phys. Rev.* **42**: 400-418

Klein, O. 1956. On the Eddington Relations and their Possible Bearing on an Early State of the System of Galaxies. In Mercier and Kervaire (1956), pp. 147–149

Klimov, Y.G. 1963. Occulted Galaxies and an Experimental Verification of the General Theory of Relativity. *Astronomicheskii Zhurnal* **40**: 874-881 (English translation: in *Soviet Astronomy* **7**: 664-669 (1964))

Kogut, A. 2012. Private communication

Kormendy, J., N. Drory, R. Bender and M.E. Cornell. 2010. Bulgeless Giant Galaxies Challenge Our Picture of Galaxy Formation by Hierarchical Clustering. *Astrophys. J.* **723**: 54-80

Kragh, H. 2003. Magic Number: A Partial History of the Fine-Structure Constant. *Archive for History of Exact Sciences* **57**: 395-431

Kragh, H. 2015a. Pascual Jordan, Varying Gravity, and the Expanding Earth. *Phys. Perspective* **17**: 107-134

Kragh, H. 2015b. Gravitation and the earth sciences: the contributions of Robert Dicke. arXiv:1501.04293, 25 p.

Kragh, H. 2016. *Varying Gravity: Dirac's Legacy in Cosmology and Geophysics*. Birkhäuser Verlag, Basel, in press

Krauss, L.M. and B. Chaboyer. 2003. Age Estimates of Globular Clusters in the Milky Way: Constraints on Cosmology. *Science* **299**: 65-70

Kreuzer, L. 1966. *The Equivalence of Active and Passive Gravitational Mass*. Ph.D. Thesis, Princeton University

Kreuzer, L.B. 1968. Experimental Measurement of the Equivalence of Active and Passive Gravitational Mass. *Phys. Rev.* **169**: 1007-1012

Kuhn, J.R. 2016. Private communication

Kuhn, J.R., K.G. Libbrecht and R.H. Dicke. 1988. The surface temperature of the sun and changes in the solar constant. *Science* **242**: 908-911

Kuhn, J.R., R. Bush, M. Emilio and I.F. Scholl. 2012. The Precise Solar Shape and Its Variability. *Science* **337**: 1638-1640

Kuhn, T.S. 1962. *The Structure of Scientific Revolutions.* University of Chicago Press, Chicago

Landau, L. 1955. On the quantum theory of fields. In *Niels Bhor and the Development of Physics*, edited by W. Pauli, L. Rosenfeld and V. Weisskopf. McGraw-Hill Book Company, New York, pp. 52–69

Landau, L. and E. Lifshitz, 1951, *The Classical Theory of Fields*, translated from the Russian by M. Hamermesh. Addison-Wesley, Reading

Lemaître, G. 1927. Un univers homogène de masse constante et de rayon croissant, rendant compte de la vitesse radiale des nébuleuses extra-galactiques. *Annales de la Société Scientifique de Bruxelles* **47**: 49-59

Lemaître, G. 1931. The expanding universe. *Month. Not. Roy. Astron. Soc.* **91**: 490-501

Lichnerowicz, M.A. and M.A. Tonnelat. 1962. Eds. *Les Théories Relativistes de la Gravitation.* Centre national de la recherche scientifique, Paris, 475 p.

Liebes, S. 1963. Test of the Principle of Equivalence. *Bull. Am. Phys. Soc.* January 1963, p. 28

Liebes, S. 1964. Gravitational Lenses. *Phys. Rev.* **133**: 835-844

Liebes, S. 1969. Gravitational Lens Simulator. *Am. J. Phys.* **37**: 103-104

Liebes, S. 2016. Private communication

Lightman, A. and R. Brawer. 1990. *Origins: The Lives and Worlds of Modern Cosmologists.* Harvard University Press, Cambridge Mass.

Lightman, A.P. and D.L. Lee. 1973. New Two-Metric Theory of Gravity with Prior Geometry. *Phys. Rev. D* **8**: 3293-3302

LIGO Scientific Collaboration and Virgo Collaboration 2016. Observation of Gravitational Waves from a Binary Black Hole Merger. *Phys. Rev. Lett.* **116**, 061102, 16 p.

Lilley, A.E. 1957. Radio Astronomical Measurements of Interest to Cosmology. In DeWitt (1957), pp. 130-136

Lopes, I. and J. Silk. 2014. Helioseismology and Asteroseismology: Looking for Gravitational Waves in Acoustic Oscillations. *Astrophys. J.* **794**: article id. 32, 7 p.

Lynden-Bell, D. 2010. Searching for Insight. *Ann. Rev. Astron. Astrophys.* **48**: 1-19

Mach, E. 1893. *Die Mechanik in Ihrer Entwickerung Historisch-Kritisch Dargestellt*; English translation *The Science of Mechanics* 1960, Chicago, The Open Court Publishing Company

Maddox, S.J., G. Efstathiou, W.J. Sutherland and J. Loveday. 1990. Galaxy correlations on large scales. *Month. Not. Roy. Astron. Soc.* **242**: 43P-47P

Mandelbrot, B. 1975. *Les Objects Fractals.* Flammarion, Paris

Mather, J.C., E.S. Cheng, R.E. Eplee Jr., et al. 1990. A preliminary measurement of the cosmic microwave background spectrum by the Cosmic Background Explorer (COBE) satellite. *Astrophys. J. Lett.* **354**: L37-40

Mchugh, M.P. 2016. The Brans-Dicke theory and its experimental tests. In *At the Frontier of Spacetime: Scalar-Tensor Theory, Bell's Inequality, Machs Principle, Exotic Smoothness*, edited by T. Asselmeyer-Maluga. Springer

McKellar, A. 1941. Molecular Lines from the Lowest States of Diatomic Molecules Composed of Atoms Probably Present in Interstellar Space. *Publications of the Dominion Astrophysical Observatory Victoria* **7**: 251-272

McVittie, G.C. 1962. Ed. *Problems of Extra-Galactic Research.* Macmillan, New York

Mercier, A. and M. Kervaire. 1956. Eds. *Jubilee of Relativity Theory. Helvetica Physica Acta, Suppl.* **IV**

Michelson, A.A. 1903. *Light waves and their uses.* University of Chicago Press, Chicago

Michelson, A.A., and H.G. Gale. 1925. The Effect of the Earth's Rotation on the Velocity of Light, II. *Astrophys. J.* **61**: 140-145

Misner, C.W., K.S. Thorne and J.A. Wheeler. 1973. *Gravitation*. W.H. Freeman and Co., San Francisco

Møller, C. 1956. The Ideal Standard Clocks in the General Theory of Relativity. In Mercier and Kervaire 1956, pp. 54–57

Møller, C. 1957. On the Possibility of Terrestrial Tests of the General Theory of Relativity. *Nuovo Cimento, Suppl.* **6**: 381-398

Moore, J.H. 1928. Recent Spectrographic Observations of the Companion of Sirius. *Publ. Astron. Soc. Pacific* **40**: 229-233

Moore, R.D. 1966. *A Study of Low Frequency Earth Noise and a New Upper Limit to the Intensity of Scalar Gravitational Waves*. Ph.D. Thesis, Princeton University

Morgan, W.J. 1964. *An Astronomical and Geophysical Search for Scalar Gravitational Waves*. Ph.D. Thesis, Princeton University

Morgan, W.J., J.O. Stoner and R.H. Dicke. 1961. Periodicity of Earthquakes and the Invariance of the Gravitational Constant. *J. Geophys. Res.* **66**: 3831-3843

Morrison, D. and H.A. Hill. 1973. Current Uncertainty in the Ratio of Active-to-Passive Gravitational Mass. *Phys. Rev. D* **8**: 2731-2733

Mössbauer, R.L. 1958. Kernresonanzfluoreszenz von Gammastrahlung in Ir^{191}. *Zeit. Phys.* **151**: 124-143

Muhleman, D.O., R.D. Ekers and E.B. Fomalont. 1970. Radio Interferometric Test of the General Relativistic Light Bending Near the Sun. *Phys. Rev. Lett.* **24**: 1377-1380

Nordtvedt, K. 1968. Equivalence Principle for Massive Bodies. I. Phenomenology. *Phys. Rev.* **169**: 1014-1016

Oort, J. 1958. Distribution of Galaxies and the Density of the Universe. In Stoops (1958), pp. 183–203

Oppenheimer, J.R. and H. Snyder. 1939. On Continued Gravitational Contraction. *Phys. Rev.* **56**: 455-459

Orosz, J.A., J.E. McClintock, J.P. Aufdenberg, et al. 2011. *Astrophys. J.* **742**: 84, 10pp.

Osterbrock, D.E. 2009. The Helium Content of the Universe. In Peebles, Page, and Partridge (2009), pp. 86–92

Osterbrock, D.E. and J.B. Rogerson, Jr. 1961. The Helium and Heavy-Element Content of Gaseous-Nebulae and the Sun. *Publ. Astron. Soc. Pacific* **73**: 129-134

Partridge, R.B. and D.T. Wilkinson. 1967. Isotropy and Homogeneity of the Universe from Measurements of the Cosmic Microwave Background. *Phys. Rev. Lett.* **18**: 557-559

Pauli, W. 1933. Die allgemeinen Prinzipien der Wellenmechanik. *Handbuch der Physik, Quantentheorie*. Springer, Berlin

Pauli, W. 1980. *General principles of quantum mechanics*. Springer, Heidelberg, 1980

Peebles, P.J.E. 1961. *Observational Tests and Theoretical Problems Relating to the Conjecture that the Strength of the Electromagnetic Interaction may be Variable*. Ph.D. Thesis, Princeton University

Peebles, P.J.E. 1966. Primordial Helium Abundance and the Primordial Fireball. II. *Astrophys. J.* **146**: 542-552

Peebles, P.J.E. 1971. *Physical Cosmology*. Princeton University Press, Princeton, N.J.

Peebles, P.J.E. 1986. The mean mass density of the Universe. *Nature* **321**: 27-32

Peebles, P.J.E. 2012. Seeing Cosmology Grow. *Ann. Rev. Astron. Astrophys.* **50**: 1-28

Peebles, P.J.E. 2014. Discovery of the hot Big Bang: What happened in 1948. *Eur. Phys. J. H* **39**: 205-223

Peebles, P.J. and R.H. Dicke. 1962a. Significance of Spatial Isotropy. *Phys. Rev.* **127**: 629-631

Peebles, J. and R.H. Dicke. 1962b. The Temperature of Meteorites and Dirac's Cosmology and Mach's Principle. *J. Geophys. Res.* **67**: 4063-4070

Peebles, P.J. and R.H. Dicke. 1962c. Cosmology and the Radioactive Decay Ages of Terrestrial Rocks and Meteorites. *Phys. Rev.* **128**: 2006-2011

Peebles, P.J.E. and R.H. Dicke. 1968. Origin of the Globular Star Clusters. *Astrophys. J.* **154**: 891-908

Peebles, P.J.E. and B. Ratra. 1988. Cosmology with a time-variable cosmological 'constant'. *Astrophys. J. Lett.* **325**: L17-L20

Peebles, P.J. and B. Ratra. 2003. The cosmological constant and dark energy. *Rev. Mod. Phys.* **75**: 559-606

Peebles, P.J.E. and J. Silk. 1990. A cosmic book of phenomena. *Nature* **346**: 233-239

Peebles, P.J.E. and J.T. Yu. 1970. Primeval Adiabatic Perturbation in an Expanding Universe. *Astrophys. J.* **162**: 815-836

Peebles, P.J.E., R.A. Daly and R. Juszkiewicz. 1989. Masses of rich clusters of galaxies as a test of the biased cold dark matter theory. *Astrophys. J.* **347**: 563-574

Peebles, P.J.E., L.A. Page, Jr. and R.B. Partridge. 2009. *Finding the Big Bang*. Cambridge University Press, Cambridge, UK

Penzias, A.A. and R.W. Wilson. 1965. A Measurement of Excess Antenna Temperature at 4080 Mc/s. *Astrophys. J.* **142**: 419-421

Perlmutter, S., G. Aldering, G. Goldhaber, et al. 1999. Measurements of Ω and Λ from 41 high-redshift supernovae. *Astrophys. J.* **517**: 565-586

Pettengill, G.H. and I.I. Shapiro. 1965. Radar Astronomy. *Ann. Rev. Astron. Astrophys.* **3**: 377-410

Planck Collaboration 2015a. Planck intermediate results. XXIV. Constraints on variations in fundamental constants. *Astron. Astrophys.* **580**: A22, 25 p.

Planck Collaboration 2015b. Planck 2015 results. XIII. Cosmological Parameters. arXiv:1502.01589v2, 67 pp.

Plotkin, H. 2014. In http://cddis.gsfc.nasa.gov/lw19/Program/index.html#sess2, 3 p.

Popper, D.M. 1954. Red Shift in the Spectrum of 40 Eridani B. *Astrophys. J.* **120**: 316-321

Pound, R.V. 2000. Weighing photons. *Classical and Quantum Gravity* **17**: 2303-2311

Pound, R.V. and G.A. Rebka. 1959. Gravitational Red-Shift in Nuclear Resonance. *Phys. Rev. Lett.* **3**: 439-441

Pound, R.V. and G.A. Rebka. 1960. Apparent Weight of Photons. *Phys. Rev. Lett.* **4**: 337-341

Pound, R.V. and J.L. Snider. 1964. Effect of Gravity on Nuclear Resonance. *Phys. Rev. Lett.* **13**: 539-540

Pugh, G.E. 1959. *Weapons Systems Evaluation Group Research Memorandum No. 11, The Pentagon, Washington 25, D.C.*; https://einstein.stanford.edu/content/sci_papers/papers/Pugh_G_1959_109.pdf

Rainville, S., J.K. Thompson, E.G. Myers, et al. 2005. *Nature* **438**: 1096-1097

Refsdal, S. 1964. The gravitational lens effect. *Month. Not. Roy. Astron. Soc.* **128**: 295-306 and 307-310

Renn, J., T. Sauer and J. Stachel. 1997. The origin of gravitational lensing: a postscript to Einstein's 1936 Science paper. *Science* **275**: 184-186

Ribas, I. 2010. The Sun and stars as the primary energy input in planetary atmospheres. *IAU Symposium* **264**: 3-18

Riess, A.G., A.V. Filippenko, P. Challis, et al. 1998. Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant. *Astronomical J.* **116**: 1009-1038

Robertson, H.P. 1956. Cosmological Theory. In Mercier and Kervaire (1956) pp. 128–146

Robinson, I., A. Schild and E.L. Schucking. 1965. Eds. *Quasistellar Sources and Gravitational Collapse*. University of Chicago Press, Chicago

Roll, P.G. 2016. Private communication

Roll, P.G., R. Krotkov and R.H. Dicke. 1964. The equivalence of inertial and passive gravitational mass. *Ann. Phys.* **26**: 442-517

Roll, P.G. and D.T. Wilkinson. 1966. Cosmic Background Radiation at 3.2 cm-Support for Cosmic Black-Body Radiation. *Phys. Rev. Lett.* **16**: 405-407

Roman, N.G. 1961. Ed. *Conference on Experimental Tests of Theories of Relativity*. Available at https://einstein.stanford.edu/content/sci_papers/papers/1961_SU_Relativity_Conf.pdf

Rosenband, T., D.B. Hume, P.O. Schmidt, et al. 2008. Frequency Ratio of Al^+ and Hgl^+ Single-Ion Optical Clocks; Metrology at the 17th Decimal Place. *Science* **319**: 1808-1812

Rozelot, J.-P. and C. Damiani. 2011. History of solar oblateness measurements and interpretation. *Eur. Phys. J. H* **36**: 407-436

Ruderfer, M. 1960. First-Order Terrestrial Ether Drift Experiment Using the Mössbauer Radiation. *Phys. Rev. Lett.* **5**: 191-192

Rugh, S.E. and H. Zinkernagel. 2002. The quantum vacuum and the cosmological constant problem. *Studies in History and Philosophy of Modern Physics* **33**: 663-705

Rutherford, E. and A.H. Compton. 1919. Radio-activity and Gravitation. *Nature* **104**: 412

Sandage, A. 1958. Current Problems in the Extragalactic Distance Scale. *Astrophys. J.* **127**: 513-526

Sandage, A. 1961. The Ability of the 200-INCH Telescope to Discriminate Between Selected World Models. *Astrophys. J.* **133**: 355-392

Sandage, A. 2010. The Tolman Surface Brightness Test for the Reality of the Expansion. V. Provenance of the Test and a New Representation of the Data for Three Remote Hubble Space Telescope Galaxy Clusters. *Astronomical J.* **139**: 728-742

Schiff, L.I. 1958. Sign of the Gravitational Mass of a Positron. *Phys. Rev. Lett.* **1**: 254-255

Schiff, L.I. 1960. Possible New Experimental Test of General Relativity Theory. *Phys. Rev. Lett.* **4**: 215-217

Schmidt, M. 1963. 3C 273: A Star-Like Object with Large Red-Shift. *Nature* **197**: 1040

Schucking, E.L. 1999. Jordan, Pauli, politics, Brecht, and a variable gravitational constant. *Phys. Today* **52**: 26-31

Schwarzschild, M. 1958. *Structure and Evolution of the Stars*. Princeton University Press, Princeton

Sciama, D.W. 1953. On the origin of inertia. *Month. Not. Roy. Astron. Soc.* **113**: 34-42

Sciama, D.W. 1964. The Physical Structure of General Relativity. *Rev. Mod. Phys.* **36**: 463-469

Seielstad, G.A., R.A. Sramek and K.W. Weiler. 1970. Measurement of the Deflection of 9.602-GHz Radiation from 3C279 in the Solar Gravitational Field. *Phys. Rev. Lett.* **24**: 1373-1376

Shakeshaft, J.R., M. Ryle, J.E. Baldwin, et al. 1955. A survey of radio sources between declinations -38° and 83°. *Memoirs of the Royal Astronomical Society* **67**: 106-152

Shapiro, I.I. 1964. Fourth Test of General Relativity. *Phys. Rev. Lett.* **13**: 789-791

Shapiro, I.I. 1967. New Method for the Detection of Light Deflection by Solar Gravity. *Science* **157**: 806-808

Shapiro, I.I. 2015. Private communication

Shapiro, I.I., G.H. Pettengill, M.E. Ash, et al. 1968. Fourth test of General Relativity: Preliminary Results. *Phys. Rev. Lett.* **20**: 1265-1269

Shapiro, S.S., J.L. Davis, D.E. Lebach and Gregory, J.S. 2004. Measurement of the Solar Gravitational Deflection of Radio Waves using Geodetic Very-Long-Baseline Interferometry Data, 1979 1999. *Phys. Rev. Lett.* **92**: 121101, 4 p.

Shklovsky, I.S. 1966. Relict Radiation in the Universe and Population of Rotation Levels of Interstellar Molecules. *Astronomical Circular, Soviet Academy of Science* **364**: 1-3.

Shlyakhter, A.I. 1976. Direct test of the constancy of fundamental nuclear constants. *Nature* **264**: 340

Singer, S.F. 1956. Application of an Artificial Satellite to the Measurement of the General Relativistic "Red Shift". *Phys. Rev.* **104**: 11-14

Steinhardt, P.J. and N. Turok. 2007. *Endless Universe Beyond the Big Bang*. Doubleday, New York

St. John, C.E. 1928. Evidence for the Gravitational Displacement of Lines in the Solar Spectrum Predicted by Einstein's Theory. *Astrophys. J.* **67**: 195-239

Stoops, R. 1958. *La structure et L'évolution de l'Univers*. Brussels, Institut International de Physique Solvay

Taylor, J.H. and J.M. Weisberg. 1982. A new test of general relativity - Gravitational radiation and the binary pulsar PSR 1913+16. *Astrophys. J.* **253**: 908-920

Teller, E. 1948. On the Change of Physical Constants. *Phys. Rev.* **73**: 801-802

Thaddeus, P. and J.F. Clauser. 1966. Cosmic Microwave Radiation at 2.63 mm from Observations of Interstellar CN. *Phys. Rev. Lett.* **16**: 819-822

Thompson, M.J., J. Christensen-Dalsgaard, M.S. Miesch and J. Toomre. 2003. The Internal Rotation of the Sun. *Ann. Rev. Astron. Astrophys.* **41**: 599-643

Tolman, R.C. 1930. On the Estimation of Distances in a Curved Universe with a Non-Static Line Element. *Proc. Natl. Acad. Sci.* **16**: 511-520

Tolman, R.C. 1934. *Relativity, Thermodynamics, and Cosmology.* Clarendon Press, Oxford

Trimble, V. 1996. H_0: The Incredible Shrinking Constant, 1925–1975. *Publ. Astron. Soc. Pacific* **108**: 1925-1975

Trumpler, R.J. 1956. Observational Results on the Light Deflection and on Red-shift in Star Spectra. In Mercier and Kervaire (1956), pp. 106–113

Turner, K.C. 1962. *A New Experimental Limit on the Velocity Dependent Interaction Between Natural Clocks and Distant Matter*. Ph.D. Thesis, Princeton University

Turner, K.C. and H.A. Hill. 1964. New Experimental Limit on Velocity-Dependent Interactions of Clocks and Distant Matter. *Phys. Rev.* **134**: 252-256

Tyson, J.A. and R.P. Giffard. 1978. Gravitational-wave astronomy. *Ann. Rev. Astron. Astrophys.* **16**: 521-554

Uzan, J.-P. 2003. The fundamental constants and their variation: observational and theoretical status. *Rev. Mod. Phys.* **75**: 403-455

Vessot, R.F.C., M.W. Levine, E.M. Mattison, et al. 1980. Test of relativistic gravitation with a space-borne hydrogen maser. *Phys. Rev. Lett.* **45**: 2081-2084

Walsh, D., R.F. Carswell and R.J. Weymann. 1979. 0957 + 561 A, B – Twin quasistellar objects or gravitational lens. *Nature* **279**: 381-384

Weber, J. 1960. Detection and Generation of Gravitational Waves. *Phys. Rev.* **117**: 306-313

Weber, J. 1961. Discussion in Roman (1961), pp. 104–115

Weber, J. 1962. On the Possibility of Detection and Generation of Gravitational Waves. In Lichnerowicz and Tonnelat (1962), pp. 441–450

Weber, J. 1969. Evidence for Discovery of Gravitational Radiation. *Phys. Rev. Lett.* **22**: 1320-1324

Weber, J. 1970. Gravitational Radiation Experiments. *Phys. Rev. Lett.* **24**: 276-279

Weber, J. and J.A. Wheeler. 1957. Reality of the Cylindrical Gravitational Waves of Einstein and Rosen. *Rev. Mod. Phys.* **29**: 509-515

Webster, B.L. and P. Murdin. 1972. Cygnus X-1 – a Spectroscopic Binary with a Heavy Companion? *Nature* **235**: 37-38

Weinberg, S. 1987. Anthropic bound on the cosmological constant. *Phys. Rev. Lett.* **59**: 2607-2610

Weiss, R. 2000. Interview by Shirley K. Cohen. Pasadena, California, May 10, 2000. Oral History Project, California Institute of Technology Archives.

Weiss, R. 2016. Private communication

Weiss, R. and B. Block. 1965. A Gravimeter to Monitor the ${}_0S_0$ Dilational Mode of the Earth. *J. Geophys. Res.* **70**: 5615-5627

Wheeler, J.A. 1957. The Present Position of Classical Relativity Theory and Some of its Problems. In DeWitt (1957), pp. 1–5

Wheeler, J.A. and K. Ford. 1998. *Geons, Black Holes and Quantum Foam.* New York, Norton

White, S.D.M., J.F. Navarro, A.E. Evrard and C.S. Frenk. 1993. The baryon content of galaxy clusters: a challenge to cosmological orthodoxy. *Nature* **366**: 429-433

Wickes, W.C. 1972. *Primordial Helium Abundance and Population-II Binary Stars: a New Measurement Technique*. Ph.D. Thesis, Princeton University

Wickes, W.C. 2016. Private communication

Will, C.M. 1986. *Was Einstein Right? Putting General Relativity to the Test.* Basic Books, New York

Will, C.M. 1993. *Theory and Experiment in Gravitational Physics*, second edition. Cambridge University Press, Cambridge

Will, C.M. 2015. The 1919 measurement of the deflection of light. *Classical and Quantum Gravity* **32**: 124001, 14 p.

Williams, J.G., S.G. Turyshev and D.H. Boggs. 2012. Lunar laser ranging tests of the equivalence principle. *Classical and Quantum Gravity* **29**: 184004, 11 p.

Wilson, W. and D. Kaiser. 2014. Calculating Times: Radar, Ballistic Missiles, and Einstein's Relativity. In *Science and Technology in the Global Cold War*, edited by N. Oreskes and J. Krige. MIT Press, Cambridge Mass., pp. 273–316

Zanoni, C.A. 1967. *Development of Daytime Astrometry to Measure the Gravitational Deflection of Light*. Ph.D. Thesis, Princeton University

Zanoni, C.A. and H.A. Hill. 1965. Reduction of Diffracted Light for Astrometry Near the Sun. *Journal of the Optical Society of America* **55**: 1608-1611

Zel'dovich, Ya.B. 1968. The Cosmological Constant and the Theory of Elementary Particles. *Usp. Fiz. Nauk* **95**: 209-230 [English translation in Sov. Phys. Usp. **11**: 381-393]

Zhu, W.W., I.H. Stairs, P.B. Demorest, et al. 2015. Testing Theories of Gravitation Using 21-Year Timing of Pulsar Binary J1713+0747. *Astrophys. J.* **809**: 41, 14 p.

Zwicky, F. 1929. On the Red Shift of Spectral Lines through Interstellar Space. *Proc. Natl. Acad. Sci.* **1**5: 773-779

Zwicky, F. 1933. Die Rotverschiebung von extragalaktischen Nebeln. *Helvetica Physica Acta* **6**: 110-127

Zwicky, F. 1937. Nebulae as Gravitational Lenses. *Phys. Rev.* **51**: 290-290

Open access funding provided by the Max Planck Institute for the History of Science.

Eur. Phys. J. H **42**, 261–291 (2017)
DOI: 10.1140/epjh/e2016-70060-5

THE EUROPEAN PHYSICAL JOURNAL H

Personal recollection

Wired by Weber

The story of the first searcher and searches for gravitational waves

Virginia Trimble[1,2,a]

[1] Department of Physics and Astronomy, University of California, Irvine CA, 92697-4575, USA
[2] Queen Jadwiga Observatory, Rzepiennik, Poland

Received 28 October 2016 / Received in final form 5 November 2016
Published online 16 February 2017

Abstract. Joseph Weber started thinking about possibilities for detecting gravitational waves or radiation in about 1955. He designed, built, and operated the first detectors, from 1965 until his death in 2000. This paper includes discussions of his life, earlier work on chemical kinetics and what is now called quantum electronics, his published papers, pioneering work on gravitational waves, and its aftermath, both scientific and personal.

1 Introduction

I was married to Joseph Weber for the last $28\frac{1}{2}$ years of his life[1], from 16 March 1972 until his death on 30 September 2000. My knowledge of his work during this period is largely firsthand. For earlier times, I have relied on published papers, narratives he wrote for various organizations, and the stories he told me. Items it occurred to me to check always turned out to be true, which is perhaps not entirely irrelevant to the rest of this story.

First we must decide what to call the phenomenon. "Gravity waves" sounds simplest and is sometimes heard, but the phrase was already in use for a process in the earth's atmosphere for which gravity is the restoring force. Exact parallel with the electromagnetic case would give us gravitational radiation, which has the advantage of definitely conveying the idea that energy is being carried, a point once in dispute. Standing waves exist, but no standing radiation! Weber used waves and radiation interchangeably over the years, but the current community has fastened

[a] e-mail: `vtrimble@uci.edu; vtrimble@astro.umd.edu`

[1] This first sentence is a paraphrase of what Jennifer Homans (2012) wrote about her late husband, Tony Judt, in a very memorable piece in the *New York Review of Books*. I suppose it also counts as a conflict of interest statement. That the arrival date of the first LIGO event, 14 September 2015, was Joe's 15th Jahrzeit – anniversary of death, on the Jewish calendar – was called to my attention by his granddaughter, Elizabeth Weber Handwerker.

onto "gravitational waves" at least partly because the word "radiation" tends to frighten or worry the uninitiated (like the word nuclear, most often when it is mispronounced as "noocooler", though I have never heard anybody say "radashiun"). I will use whichever words appear on original documents.

Curiously, the first generally-recognized pulse of gravitational waves (Abbott et al. 2016) reached Earth just 100 years (minus about 7 weeks) after Einstein submitted the first of his four famous November 1915 papers, "On the General Theory of Relativity" (Gutfreund and Renn 2015, pp. 161–162). All were submitted on Thursdays; perhaps this was just the day when the editor of the *Sitzungsbericht* of the Prussian Academy of Sciences opened his mail. "Cosmological Consequences in the General Theory of Relativity", which introduced the cosmological constant (Einstein 1917) was also a Thursday child (8 February 1917).

The normally-cited Einstein (1916, 1918) papers on gravitational waves arrived on 22 June and 31 January respectively, also Thursdays (Gutfreund and Renn 2015). In due course, he had second thoughts, of which more later, but I don't think he was ever quite so wishy-washy on the subject as indicated in *The Economist's* report of the LIGO event (Anon 2016).

2 Background and early work

Joseph (Yonah ben Yakov) Weber was born in Paterson, New Jersey on 17 May, 1919, the second son and last of four American-born children of Leah (Lena on some documents) Stein and Jacob Weber. The family name was originally Gerber and was changed to match a passport that was available quickly and cheaply when, in 1909, they decided to leave the part of the Russian Empire now called Lithuania. Joe's first name arose from a misunderstanding when his mother went to enroll him in school, and his near-lack of the standard regional/ethnic accent was due to his being knocked down by a bus at age five and having to be re-taught to speak by a therapist from Philadelphia (at public school expense, incidentally).

By age 10, Weber had assembled his first crystal set and joined the Passaic County Amateur Radio Club. Not long after, he contributed to the family excheqeur by working in a radio store, which he found both more interesting and more remunerative than a paper route or caddying. He graduated from the Paterson Talmud Torah and Eastside High School already in love with a class mate, Anita Straus, who went on to Smith College, receiving a BA in physics in 1940. Joe's high school annual describes his activities as Mathematics Club and Orchestra; his hobbies as Amateur radio, chemistry, and astronomy; and characterizes *Ambition* as "????". Classmates had mentioned money, travel, reading, and cooking as "ambitions" in the year book.

Joe looks younger than most of his classmates, as indeed he was. Thus, although he had received a congressional appointment to the US Navy Academy (USNA) in 1935, he was initially too young to be admitted, and spent a year at Cooper Union, though a good deal of his education also came from the Danforth Memorial Library in Paterson. He described his favorite book as Maxwell's *Relative and Absolute Motion*, which apparently does not exist. The editors suggest that the book might actually have been the 1876 *Matter and Motion*.

Weber graduated from the USNA with the class of 1940. Among his accomplishments as an Annapolis cadet was wiring the mess hall for sound, so that one fine evening the tones of Schubert's Great C Major symphony drowned out the clatter of cutlery and the chatter of cadets. His first assignment was to the aircraft carrier Lexington, and he was above decks when she was sunk in the battle of the Coral Sea on 8 May 1942 (having steamed out of Pearl Harbor on 5 December 1941). Part of his compensation for "articles lost in a marine disaster" purchased an engagement ring for Anita, and they married not long after Navy rules permitted this. She continued teaching high school physics for several more years.

His next task was skippering the submarine chaser SC 690, which had six inch guns and a 10 cm radar, neither standard equipment, but the result of his scrounging and ability to maintain things. After participating in the first wave of the Sicilian landing, he spent a year studying electronics at the Naval Postgraduate School and was then assigned to run electronic countermeasures for the Bureau of Ships.

When Weber resigned his Navy commission (as lieutenant-commander), he was offered jobs by several of the companies that had received grants from his office, but instead he accepted a full professorship of electrical engineering at the University of Maryland in 1948. They requested (firmly!) that he earn a Ph.D. in something, somewhere, soon. Peebles et al. (2009) describe his interaction with George Gamow at George Washington University, which might be described as one of at least half a dozen near misses at pre-1965 discovery of the cosmic microwave background radiation. In the event, Joe became a graduate student at the Catholic University of America while both a Maryland professor and a consultant at the Naval Ordnance Laboratory, as was his thesis advisor, Keith Laidler (1916–2003), a native of Liverpool, who returned to Canada not long after Weber's 1951 Ph.D. (not, as far as I know, causal).

2.1 Chemical kinetics and the inversion spectrum of ammonia

Ordinary ammonia, NH_3, has a strong microwave (K-band) absorption feature at 1.25 cm (24 GHz), a splitting that results because the N can be on either side of the triangle of H atoms. Deuterated ammonia (NH_2D) has a similar feature, but at a different frequency. Weber's thesis experiment used a ten-foot microwave waveguide filled initially with NH_3 and a supply of HD deposited on the walls. The changing strength of the two absorption features when broad-band microwaves propagated down the waveguide revealed the rate and amount of the conversion of NH_3 to NH_2D (Weber and Laidler 1950, 1951a,b; Weber 1951). In the course of the experiment, they measured the wavelengths of the transitions more accurately than anyone before or for sometime after.

Related work included Kurt Shuler and Weber (1954) on ionization in flames, and the construction and use of a Stark effect microwave spectrometer (Marshall and Weber 1957a,b) applied to OCS, previously studied by the Townes group. This was Sam Marshall's Ph.D. thesis, with Weber as the effective advisor, though their affiliations were listed as Naval Ordnance Lab and CUA (Marshall) and NOL and U Md (Weber). Weber and Laidler both listed NOL and CUA. I have not seen a copy of Weber's actual thesis and am not sure that he even had one by 1972. Joe had several other students during his days in the electrical engineering department, but his name did not appear on their papers, they are not listed on his CV (nor are his physics students), but some remained close friends.

2.2 Inverted populations as amplifiers

Graduate physics at Catholic University around 1950 was largely taught in weekly 7–10 PM classes (good for returning veterans with day jobs), and largely taught by Karl Herzfeld, who had served for Austria in WWI, coming to the United States in time to escape the horrors of being Jewish there before and during World War II and also in time to be John A. Wheeler's thesis advisor. Joe recalled that, the moment he heard about the Einstein A and B coefficients, he realized that a population of molecules with more of them in an upper than in a lower energy state could be used as an amplifier.

So, in March, 1951, Weber went into the lab, constructed what we would now call an ammonia molecular beam maser, and learned that it worked as a very high resolution spectrometer, but that a solid state device would be needed to make an amplifier with a useful gain-band-width product (Weber 1959a,b). Bloembergen (1956) had made the same point. According to the review (Weber 1959b), the first observation of weak maser-type amplification was in 1950 by Purcell and Pound (1951), who looked at resonance of the Li^7 nucleus in a LiF crystal when it was moved very quickly from a strong magnetic field to a field pointing the opposite direction, resulting in an inverted level population and so a negative temperature in a Boltzmann equation.

During 1951–52, Weber discussed amplification by ammonia at seminars at Princeton and in the Washington area and at a 1952 IRE meeting in Ottawa Canada, under the title "Amplification of microwave radiation by substance not in thermal equilibrium" (Weber 1953a). Later that same year, there came a letter dated 23 November 1953 from Charles H. Townes, requesting a reprint of the paper, because, he wrote, he had a student named J.P. Gordon, who was working "on a related topic". The authors of the eventual Columbia paper (Gordon et al. 1954) were J.P. Gordon (called Jim), a graduate student who never finished his Ph.D. and H.J. Zeiger (called Herb), a postdoc who had completed a Ph.D. under I.I. Rabi (Johnson 2016). And then there were Basov and Prokhorov.

This requires a bit of expansion. That meeting was the Electron Tubes Conference of the Institute of Radio Engineers, and the publication was in the Proceedings of the Institute of Radio Engineers, Professional Group on Electron Physics. Weber had joined the IRE in 1944, became a full member in 1946, a senior member in 1953, a fellow in 1958, and a life fellow in 1985, for which the requirement was that the sum of a member's age and his years of fellowship exceed 100, and no, the arithmetic doesn't quite work out. In any case, in the interim, the IRE had merged with the Institute of Electrical Engineers (IEE) to become the IEEE, the Institute of Electrical and Electronic Engineers. A recently deceased German-Israeli-American engineer whose memberships followed a similar course once assured me that the IRE was the more prestigious of the merging Institutes.

Secondary sources (Campbell 1960; Kastler 1985; Thorne 1994; Glanz 2000) sometimes mention Weber's talk and paper as the first "open" description of what became masers and later lasers (the word was a Townes coinage). The explanation is a May, 1952 presentation by Basov and Prokhorov who "pointed out the theoretical possibility of a device producing microwaves by using stimulated emission at an All-Union conference on radiospectroscopy". (Kastler 1985, who should not be blamed for the misspaced prepositional phrase because he was quoting from a textbook by Bertolotti 1983). The Soviet paper appeared after the Gordon et al. (1954) one (Basov and Prokhorov 1954).

Some other items worth noting: first, Townes's first, actual, physical maser used ammonia gas. Second, Donald H. Menzel (1937) had earlier remarked that radiation passing through interstellar gas could, in principle, be amplified rather than absorbed, though he expected the effect to be small; several interstellar molecules, including CH_4 and H_2O, do in fact mase. Third, when Weber was elected to the University of Maryland Engineering Hall of Fame in 1988 the citation was for the earliest publication of quantum electronics principles. Fourth, when he first visited UC Irvine in February 1972, most of the then-members of the department (from Fred Reines and John Pelham on down in age) already knew or knew of him for that work and were the first to inform me that he should have shared the Nobel Prize.

Weber taught microwave engineering and related subjects at the University of Maryland from 1948 until about 1961, when he gradually moved from the engineering department into physics and began to teach quantum mechanics and all the other ills to which the physics flesh is heir. A baker's dozen papers (Weber 1953a, 1954a,b,c, 1955, 1956a,b, 1957, 1960b,c,d, 1961b; Weber and Hinds 1962) deal with a wide range of topics concerning various forms of electromagnetic radiation, many theoretical, and few with "maser" in the title or abstract. In 1959, when he was asked to review the topic for *Reviews of Modern Physics* (Weber 1959b), only four of the many references were to his own work – the primordial IRE paper, one each on vacuum fluctuation noise and maser noise considerations, and the Gravity Research Foundation prize essay that already looked forward to the use of maser amplifiers in possible designs for gravitational wave detectors (Weber 1959a).

The review also has a discussion of the energy levels of ruby and its properties as a potential amplifier. The second ring Joe ever gave me was a glorious emerald-cut laboratory grown ruby (set in yellow gold, with diamonds on either side, and accompanied by a biblical passage generally quoted as "a woman of valor", with a mention that his life would have played out very differently if he had known about rubies in 1952, but that he had no regrets). The first ring? That was the Tiffany solitaire engagement ring and wedding band, a few weeks after we met.

Some additional background material, from my point of view, appears in Trimble (2000, 2014, and 2016). As late as 1969, Weber was asked by Gordon & Breach Publishers to edit a pair of volumes of reprints of critical papers from the history of masers (Vol. I) and lasers (Vol. II) with commentary.

The actual affiliation listed on Weber (1953a) and other early papers is the Glenn L. Martin College of Engineering and Aeronautical Sciences, University of Maryland, College Park, Maryland. Apparently the then university president, "Curly" Bird had tried to recruit von Kármán to head up the new school and got the response, "Mr. Bird, where is Maryland?" From 1954–55 onward, progress in understanding and constructing masers and their ilk, as spectrometers, frequency standards, and amplifiers, proceeded quite rapidly among a fairly compact set of institutions including Columbia, Bell Telephone Labs, Princeton, and MIT. The best guide to who did what, when, and how is probably the articles reprinted and the commentary given by Weber (1969a,b), unless you prefer a whole book like Bertolotti (1983). One micro-factoid: while Purcell and Pound's (1951) experiment used magnetic fields to invert a thermal population, Weber's first operating device, a spectrometer in effect, used a reversing electric field.

A couple of more sidelights: (1) Good (1946) is a wonderful introduction to the inversion spectrum of ammonia, well written even by the higher standards of the time. (2) A reasonable question is "why did he publish there?" The answer is on page 51 of the reprint volume *Masers* (Weber 1969a).

"As noted earlier, I had presented a discussion of this principle at the 1952 Ottawa electron tube research conference. It had been my intention to publish these results in a widely read journal. Early in 1953 Professor H.J. Reich of Yale University wrote to say that he had been chairman of the 1952 electron tube conference program committee, and was also editor of a (not so widely read) journal. As a result the conference summary report was published in the June 1953 Transactions of the Institute of Radio Engineers Professional Group on Electron Devices". Near the end of his career, when most physicists had no use for results from room-temperature bar detectors for gravitational waves, Weber again published largely in low-prestige journals and conference proceedings.

But, when Weber became entitled to his first sabbatical, for 1955–56, he chose to go to Princeton with a pair of fellowships, on which J. Robert Oppenheimer and John A. Wheeler would be his advisors. In other words, he had started to think about gravitation, and he actually spent the second half of that academic year in Leiden with Wheeler, where the widow of Paul Ehrenfest (Tatiana Afanassieva, who published at least one paper with her husband) gave him the photograph of Einstein that appears as the frontispiece of his textbook (Weber 1961a).

Joe's love of tinkering survived the grim years of "Weber never did anything right" in the physics community, and our dining room featured a traditional oil-burning Shabbas lamp fitted with flickering tiny lightbulbs. He always picked the simplest solution to a mechanical problem that would work: a piece of string to lengthen a pull chain on a lamp, duct tape when I sawed through the long cord on the electric saw trying to get ivy off the house (some of which, unfortunately, was poison ivy), tightening the screws on the rattling dash board of our 1968 Camaro. And Weber picking the correct capacitor or resistor from a box of miscellaneous components was a joy to watch.

Collins (2004) recorded that he had been driven by Weber from the University to the gravity building in what he described as an old car. It was not the 1968 Camaro, which was one of our California cars. In fact, since the year was 1975, it was either Tweedledee or Tweedledum, the near-twin 1964 and 1965 Dodges that served perfectly well until a tree falling from the grounds of the Chevy Chase Country Club destroyed the garage and Tweedledee inside. The house, which Collins wrote that he had not seen, had five bedrooms, typical for Chevy Chase, and when the sad moment came that I had to sell it after Joe died, it went on the market as what a colleague called "a million dollar fixer-upper???" because neither of us was very good about taking time away from science to worry about interior paint and carpets, or new appliances, when the old ones worked just fine. Some of the money endowed the Joseph Weber Award in astronomical instrumentation for the American Astronomical Society.

3 The paper trail

No complete list of Weber's publications exists. This is not uncommon; I found the same thing while writing an entry for the *Biographical Encylopedia of Astronomers* (Hockey et al. 2014) on Thomas Gold[2]. What happens is that the author gives a talk at the Tierra del Fuego conference or submits a paper to the journal *Cosmologica Acta et Retracta* and enters the preprint into his[3] CV, but, when the proceedings finally appear or the paper is accepted and published a year or two later, he never goes back to fill in the details.

This phenomenon is particularly true here because so many of Weber's later papers were in conference proceedings, especially after the American community had decided to disbelieve his work. Colleagues in China, India, France, Italy, Pakistan, and a few other places (there was Malaysian currency in a drawer with all the rest when he died) continued to invite him to give talks at their meetings. But, as a lower limit, he was the author or co-author of at least 130 papers, dated 1950 to 2001, with at least 35 co-authors, never in large groups. He presented research results at a minimum of 50-some conferences, not counting meetings of the American Physical Society, from 1950 to about 1998.

The early papers were in chemical kinetics, as mentioned in Section 2.1. Then came "Amplification of Microwave Radiation by Substance not in Thermal Equilibrium",

[2] Gold, by the way, taught general relativity at the Cavendish Lab in Cambridge UK in 1949–53, so the subject was not completely neglected even in this period of relative disinterest. Peebles (2016) elegantly presents the American recovery of experimental gravity physics elsewhere in this volume. Blum et al. (2015) have recently discussed this "low water mark of GR" and its "Renaissance" in the context of an international historiographical framework and with references to a number of other authors and papers who earlier perceived a similar structure to the history. I am reluctant to copy out their references, lest it give the impression that I have read them.

[3] Well, yes, women (including me) also make this mistake, and I occasionally have to ask an editor or organizer, "What ever became of. . . ?" Don't you?

from that 1952 talk (Weber 1953a). The missing article was characteristic; Weber's speech was always slightly laconic and hesitant, a lingering relic of the speechless year, and he said that social conversation should not be a competitive sport. Quantum electronics publications ended with the edited volumes on masers and lasers (Weber 1969a,b).

Meanwhile, publications on gravitational waves began with Weber and Wheeler (1957), which addressed their reality (next section) and continued to the end (Weber 2000). This last paper is a bit of an embarrassment. I have a copy; I know it was accepted because the proofs came shortly after Joe died; I corrected the proofs (about which he had been very concerned in his last days) and returned them along with a couple of paragraphs of biographical material requested by the editor. But I cannot remember the name of the journal (or the editor) concerned, and it was a sufficiently non-prestigious one that the paper is not to be found by the Astrophysics Data Service.

The textbook (Weber 1961a) *General Relativity and Gravitational Waves* was, he later said, part of his effort to learn the subject thoroughly. I read it in about 1963, long before I met the author, because it was the thinnest GR text on the library shelves at UCLA when I thought I should learn something about the field, which was not then taught there. His attention to possible coherent detectors for neutrinos began with Weber (1981a) on "Exchange of Energy with Large Numbers of Particles" and also continued to the end.

I have copies of very few of these papers, though the archives at the University of Maryland Library took, and I assume still has, one of each item that was in his files in fall, 2000, including proposals, referees' reports both written and received, letters to and some copies of letters from Joe, and so forth. Most of the conference volumes from his Maryland office, the Chevy Chase house, his UCI office, and our Irvine apartment were donated to the Niels Bohr Library of the American Center of Physics in Maryland, along with hundreds of others of his books, and a good many of mine from Maryland home and office. The library disposed of conference proceedings only when they were duplicates, and most were not.

The sole-author papers in the reference list from 1951 to 2001 are intended to provide a summary of Weber's work as presented in the (relatively) more accessible publications. Not all are explicitly cited in the present text.

4 Does gravitational radiation exist? What others thought

Wave-like solutions to the equations of general relativity were never in doubt. The critical point is whether they would carry energy away from accelerated mass quadrupoles. This is more difficult. Indeed even in the electromagnetic case, there are both advanced and retarded potentials, and it is in some sense left for observations to decide that we see only the latter. There was, once upon a time, the Wheeler-Feynman (1948) absorber theory of radiation, supposed to provide an explanation, if the universe were either closed or in steady state. Curiously, that was what Feynman chose to talk about in a seminar organized by Caltech students soon after the October 1965 announcement of his Nobel Prize (I was there, part of a little coven of astronomy grad students somewhere near the middle of the auditorium).

Even before general relativity, Poincaré (1905) was an early ("first" is always dangerous) proponent of the idea that gravitational information must travel in wave form to convey to the surroundings of some system that the system has changed and that masses in the surroundings must respond accordingly. Poincaré is also supposed to have been the person Einstein said would have been most likely to discover a general theory of relativity if he himself had not (this is strictly third-hand rumor). Poincaré's view was predicated at least partly on analogies with electromagnetic radiation, also made use of by Feynman (1962–63–63), Weber (1961b), and many others. Levi-Civita (1917) denied the physical reality of gravitational waves because the covariant stress-momentum-energy tensor is equal to zero. It is conceivably not a coincidence that he was the president of the Commission on Relativity of the International Astronomical Union when it voted itself out of existence in 1925.

Different discussions of gravitational waves give rather different sets of physicists credit and blame for doubting existence and establishing it, Weber's own (1961a) version counted on the side of the angels Einstein (1916, 1918), Eddington (1923), Fierz and Pauli (1939, on what sort of wave would go with a spin-two graviton), Landau and Lifshitz (1951, there are many other editions in multiple languages), Bonnor (1954), Lichnerowicz (1955), Foures-Bruhat (1956), Bondi (1957), Pirani (1957), Brill (1959), Robinson and Trautman (1960) and naturally Weber and Wheeler (1957). You will not find "Bonnor (1954)" in the references here. I think it must be a ghost of Bonnor (1959) on "Spherical Gravitational Waves", in which he calculates, using retarded potentials, that the loss of gravitational mass from a system is equal to the energy carried by the waves, when you go beyond the linear approximation, and for the case that the changing quadrupole moment is driven by a non-gravitational force, for instance a spring connecting two masses. He states that he has been unable to "find an answer for the case of an isolated system with only gravitational forces".

The waves are of Petrov (1954, cited by Petrov 1962) type II. Feynman (1962–63), like Fierz and Pauli (1939) was primarily interested in a particle approach, showing that you would get GR as what was carried by a spin-two, massless particle. He writes "graviton" as if it were an old word, which it was, having (according to the *Oxford English Dictionary*) appeared as far back as 1942 in *Chemical Abstracts*, which spoke of "plane waves of a particle of spin 2 (graviton)".

Of standard texts, Møller (1952) has no interest in the topic; Adler et al. (1975) regard the linear, weak-field case as suitable for an introductory treatment; Hartle (2003) gives us two chapters, going from the weak, linear case up to strong field calculations. And the one you've all be waiting for (because you weren't strong enough to carry it with you), Misner, Thorne, and Wheeler (1973) ask the student to follow Bondi (1957, 1965) and Bondi and McCrea (1960) in showing that you can extract energy from a gravitational wave and use it to heat a stick, a real effect though a gedanken experiment. They provide the standard quadrupole formula for energy radiated by point masses in a binary system (tiresomely in $c = G = 1$ units), which made an early appearance in a book on interstellar communication (Dyson 1963). The exercise for the student is 18.5, and, in due course, they devote chapters 35, 36, and 37 to a mathematically intense treatment of gravitational waves.

Peebles (2016, elsewhere in this volume) says that the issue of existence was never in doubt after 1957, relying on the argument given by Weber and Wheeler (1957).

How could existence ever have been questioned, and who dared oppose this team of heavyweights? Well, it all started with Einstein and Rosen (1937). The story has been well told by Kennefick (2007), though I first heard it from Martin Blume, former editor in chief for the American Physical Society. He had looked back in the files of the *Physical Review*, to which that paper was first submitted, and refereed (anonymously) by H.P. Robertson, who gave the same advice directly to Einstein. Results were the published paper, with many changes, and Einstein's resolve never again to send anything to *Physical Review*. Items that it was possible to get wrong included whether the part of the stress-energy tensor representing a traveling wave could be transformed away by repeated jiggling of reference frames and whether a particle hit by a wave would be moved and so absorb energy. The answer to that is no, if you forget to include the radiation reaction as it re-emits. Similar things can be said about electromagnetic radiation. Another way of describing the problem is to say that the equation of motion of a particle is not damped by the radiation term (Infeld 1938).

Indeed, my impression is that Leopold Infeld was the key player on the opposing team. The standard place to start is Einstein et al. (1938), which dealt with the problem of motion in Einsteinian gravity (not yet fully solved, in the view of Damour 1987) and which is said to have distressed Banesh Hoffmann by the non-alphabetically ordering of the authors. At any rate, Infeld produced a steady stream of anti-gravitational radiation papers (Infeld 1936, 1937, 1938, 1954, 1956, 1957, 1959), in due course involving his students (Infeld and Wallace 1940, Infeld and Schild 1949) and others (Infeld and Plebanski 1960, Infeld and Scheidegger 1951, Infeld and Michalska-Trautman 1960). He appears to have handed over to Scheidegger (1953 and elsewhere), whose review of existence vs. non-existence also mentions points about energy content of the waves being zero and the transforming away of the wave part of the stress-energy tensor. Even Eddington (1924) worried a bit about the choice of coordinate systems, since it seemed to him that the velocity of propagation of gravitational information would be equal to the speed of light only for a properly chosen set of coordinates. At various moments, Weyl (1944) and Rosen (1956) were also non-believers.

Let's stop for a moment with Bonnor's (1959) calculations, using retarded potentials. He states clearly that, in second-order approximation, the mass lost by a system with an oscillating quadrupole moment is exactly equal to the energy carried by waves calculated in the first order approximation… for a system driven by non-gravitational forces, like a pair of masses joined by a spring, set into oscillation. But he is unable to reach the same conclusion if the system is isolated and only gravitational forces are at work.

By the time of the 1959 Royaumont conference, essentially all the participants agreed that gravitational waves could carry energy, if properly calculated and not considered in infinite space(-time), for which energy is not well defined. This is not the same as saying energy is zero! In the concluding remarks, however, Bergmann (1962) said that it had probably been unfair of them to reach this conclusion in the absence of the strongest opponent. He expressed hopes that Infeld would be at the next such meeting in a few years. Four other Polish physicists appear on the list of participants and Fock, Ivanenko and Petrov from the USSR.

Bergmann also expressed hopes for major contributions to cosmology from radio astronomy and for improved neutrino detection devices that could reach sources beyond earth. His precise words near the end were "If one of Weber's schemes to observe gravitational radiation should become realistic, that, too, would provide us with a completely new channel of information". Bergmann's English was largely British, so "scheme" did not sound quite so pejorative to him as it does to Americans now. His French text speaks of "dispositifs imagines" as the equivalent of schemes. In any case, by the 1962 GR3, his view was that there was no point in Weber's work and that nothing would come of it for 100 years. The remark may well appear in some report of the meeting, but I remember it as a Weber quote. Peter was anyhow approximately half right on the time scale!

Surely by now opposition has died out (in the sense claimed by Planck)? Not entirely. An arXiv posting by Loinger (2003) is described as a history of the discovery of the non-existence of gravitational waves. Rather impressively, only two days after the public announcement of the LIGO burst, Loinger and Marsico (2016) had produced a single page called "On LIGO's Detection of a Gravitational Wave", concluding that no gravitational wave can be emitted, because the gravitational trajectories of the interacting bodies of an ensemble (with no other forces) are geodesic lines (compare Bonnor's conclusion above). A large number of Italian institutions are represented on the LIGO team (Abbott et al. 2016) but apparently not the University of Milan, Loinger's home institution.

Perhaps nothing to do with the story, but Infeld's disciple Scheidegger spent most of his career at Imperial Oil Ltd, Calgary, and Infeld himself seems to have had particularly bad luck, even by the norm for Jews from Poland born in 1898. Canada had no Joseph McCarthy, but they did send him back to Poland as a suspected Communist after he had spent more than a decade there. Historian and philosopher of physics Allen Janis (personal communication, May 2016), who knew most of the people involved in this issue, tells me that Infeld's Canadian-born second wife never learned to speak Polish, making their 18 years back in Warsaw less pleasant than might otherwise have been the case.

5 The pioneer

The version of the story Joe told me began with one of his young sons banging his head against the wall at night, keeping his father awake and with extra time to start thinking about how he might bring what he called the beautiful theory of general relativity into contact with laboratory experiments. At any rate, he took his fellowships from the Guggenheim Foundation and the National Research Council in 1955–56 up to Princeton (the University and the Institute for Advanced Study) with J.R. Oppenheimer and J.A. Wheeler to be his advisors. Weber had some amusing Oppenheimer stories, but it was with Wheeler that he headed to the Lorentz Institute for Theoretical Physics in spring, 1956. Anita, who had to cope with the shopping and child-care, learned some Dutch; Joe did not. Lest we not pass this way again, he later held another Guggenheim (1962–63) again at Princeton and a Fulbright (summer 1963). A Wheeler/Princeton custom which followed him the rest of his life was the bound lab notebook, with numbered pages, into which went notes from colloquia and meetings, calculations of anything he thought should be calculated, equipment designs, and updates on the status of experiments, for instance the equivalent noise temperatures of the bar detectors each time he rang them down from input electrical signals. He is writing in one of those notebooks in the "Gravity Building" in one of the photos on file at the Emilio Segrè Visual Archives at the American Center for History of Physics in College Park. Another of those photographs shows, higher in the room, the lovely Japanese figure of a woman in white, the only mistress, Joe said, his first wife would allow him.

Weber had been working on various aspects of electromagnetic radiation, so he naturally thought of ways that energy from gravitational waves might drive changes in the Maxwell tensor (Weber 1980). The first designs that made it into print (Weber 1959a, 1961a, 1962a) show a pair of masses connected by a spring (energy to be extracted from the wire) and then two very massive piezoelectric crystals, connected to amplifiers and a receiver to cross-correlate the outcoming electric currents and look for relative motion between the crystals. A minor aspect of "national culture" is that Weber nearly always represented his detectors with equivalent circuits, with a driving voltage, resistance, inductance, and capacitances (Figs. 1 and 2), while Bondi (1957, 1962) and others trained in Britain tended to think of "springs and dashpots".

The next attempt, with Robert L. Forward and David M. Zipoy, was to design a high frequency pumped parametric capacitor, coupled to the end of a massive cylinder. Later, Hirakawa at the University of Tokyo explored the capacitor strategy (Hirakawa and Narihara 1975), as did Jean-Paul Richard (1976), who had joined the University of Maryland physics department after Weber moved over from engineering[4].

Weber also looked briefly at the possibility of a free-mass interferometer (Forward 1971, 1978; Moss et al. 1971). I enter the story peripherally at this point, because

[4] Hirakawa remained on the list of people with whom we exchanged holiday cards at year's end for several years, and I believe he died fairly young of tuberculosis. Richard's decision to go his own way obviously did the Maryland gravitational wave group no good in community eyes.

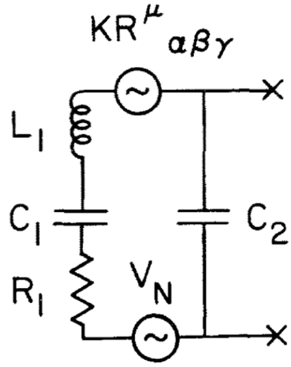


Fig. 1. Equivalent circuit for a bar antenna. $KR^{\mu}_{\alpha\beta\gamma}$ is the Riemann tensor that stretches and compresses the bar, creating electric currents in the crystals. V_N is all the noise sources put together. From Weber (1984), on p. 1186.

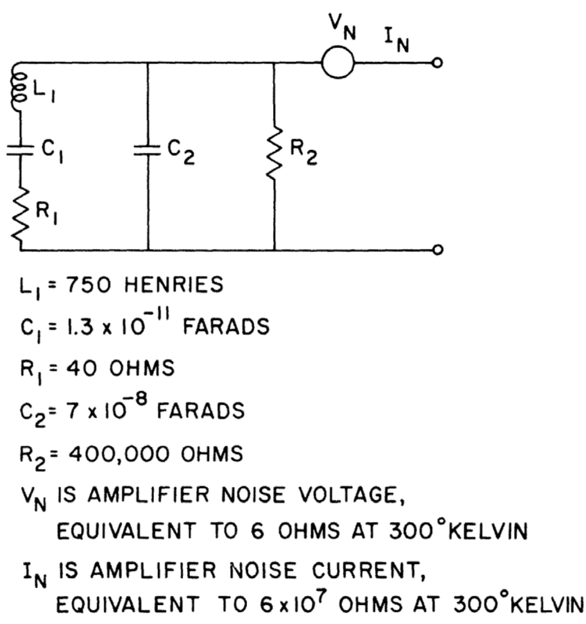


Fig. 2. Equivalent circuit for the "noise" part of the bar detector in the Gravity Building in June 1974, with numerical values of the inductances, capacitances, resistors, and voltages. From Lee et al. (1976), on p. 897.

I knew Bob Forward slightly from his occasional attendance at general relativity seminars at Caltech in 1966–68, when I was a graduate student there. He invited me back home to southern California from a postdoc in Cambridge, England, to spend a couple of weeks in December 1969 at Hughes Research Lab (where he had returned after completing his Ph.D. and a short postdoctoral term with Weber in Maryland). Forward said his goal was to teach me to be a systems engineer, but what I actually calculated for him were (a) the amount of gravitational radiation to be expected from a pulsar if its quadrupole moment was what was implied by glitches being starquakes in which mountains smoothed down and (b) what astronomical sources could be reached optically with some new form of intensity interferometry. The answers were

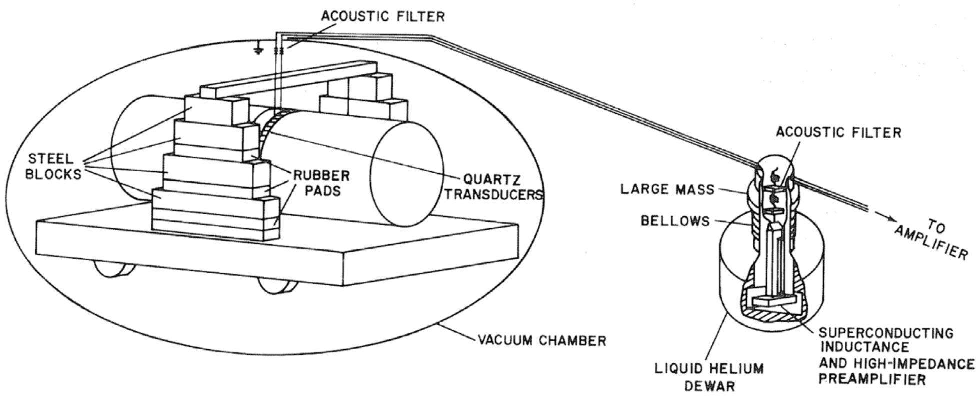


Fig. 3. The bar detector design as of 1964. Filtering, preamplifiers, mid amplifiers all became more complex over the years, and the quartz transducers were replaced by PZT. From Weber (1980), on p. 453.

(a) not much and (b) the sun. But while there I saw his first (Forward 1971) free-mass interferometer intended as a 3-meter long detector for gravity waves. The substrate was a granite slab from a supplier of gravestones. Forward later operated a somewhat larger free-mass interferometer that reached a sensitivity in displacement of the masses of about one part in 10^{15} (Forward 1978). Rainer Weiss (1972) at MIT had begun thinking along the same lines after completing a balloon experiment to study the microwave background radiation.

Forward at the time spoke very highly of Weber, and he was one of Joe's two best men when we married a second time in a synagogue in Orange County.

Meanwhile, back at the College Park ranch, Joe had realized that enormous piezoelectric crystals probably didn't exist (advertizing literature from the period talks about sizes measured in centimeters not meters) and that they weren't necessary. Instead, one could take a many-kg mass of something cheaper and stiffer, use it as the detector and take the energy out via hand-size quartz or PZT (lead-zirconate-titanate) crystals attached firmly to the bar. One of the group members later remarked that chunks of tungsten went glump when you hit them, but aluminum rang. The first published plans (Weber 1959a,b, 1961b) show single massive crystals, but the first thing actually built was a modest, 8″ diameter aluminum bar with quartz crystals bonded to it (Fig. 3).

Other bar detectors were built through the 1960's, with graduate student Joel Sinsky (interviewed and celebrated by Collins 2004) traveling to superintend the construction of both the aluminum bars and the crystals, the latter coming from Gulton Industries and Clevite. A good deal of experimenting was also needed to identify the right adhesive to attach crystals to bar, and the most-reproduced photograph from this period shows Weber bending over a bar, gluing on crystals or attaching wires to them to bring the signal out of the giant vacuum chambers that held the bar and disk detectors (Fig. 4). I remember Eastman 910 (pronounced nine-ten) and Araldite cement.

Aluminum bars of various sizes and shapes, at various temperatures have Q-values (ring-down time in units of the resonant periods) up to 100 000 to 500 000 (Weber 1980). They must be isolated acoustically and electromagnetically from everything you can think of that is not a gravity wave, and all of the longer Weber papers have discussions of some aspects of the problems. At best, they will have noise temperatures near room temperature, unless you go to cryogenic systems, and no one expected

Fig. 4. One of the several standard images of Weber posed as if working on a bar antenna, either gluing on crystals or soldering them to wires that would connect to the outside. Probably 1969.

signals anything like that powerful. Thus the goal was to operate two or more antennas further apart than the reach of earth tremors, power failures, cosmic ray showers, noisy traffic, and so forth.

One could then look for instants of coincident high output from two or more bars. First these were at different places on the University of Maryland campus, then, with increased funding, in a specially-built "Gravity Building" near the university golf course, and, for the data runs that led to the first announcements of evidence for the detection of gravitational radiation (Weber 1968c, 1969c), there was a large bar each in the Gravity Building and at Argonne National Laboratory near Chicago, where Roy Ringo kept an eye on things between Weber's frequent visits.

Initially, the electrical signal was transmitted back to College Park from Argonne via dedicated phone lines and a microwave link, and the two signals traced out by Easterline-Angus strip chart recorders, frequently in red ink. Weber, secretary Alessandra Esposito, and others (including me at one stage) then examined the charts, looking for times when both outputs had been above some pre-chosen threshold for a fraction of a second. The numbers varied over the years. The first bar was resonant at about 1400 Hz, the later ones and the disk (in its "breathing" mode) at 1660 or 1661 Hz.

Over the next few years, groups at Stanford, Saskatchewan, Glasgow, Louisiana State, Rochester, Rome, Moscow State, the Max Planck Institute in Munich, Caltech, Bell Telephone Labs, the University of Tokyo, and IBM built resonant bar detectors, generally just one per site and not run together for coincidences. None reported results consistent with the Maryland experiments, and all at some stage announced that Weber's data were just noise of some sort. It is not true that none of the Maryland results were ever confirmed (next section), but that was the impression left in most physicists' minds.

Numerous changes in the Maryland installations also occurred over the years, many generated in-house, some in response to assorted criticism. First the data were recorded separately at the two locations, and data tapes flown back for computer analysis by programmer Brian Reid (whose name appears as Reed in some secondary sources), lest the phone lines somehow be the cause of the coincident power pulses. The crystals were replaced with brass mushroom-shaped transducers fastened to the far ends. Super-conducting quantum interference devices (SQUIDs) were tried as coupling. Preferred and non-preferred algorithms for deciding what constituted a coincidence were implemented by several students as data experts. And so forth, as we all slipped gradually into the situation described in Section 8.

One part of the procedure the current interferometer community agrees Joe got right. The rate of accidental, noise-based, simultaneous increases in power output from two bars was measured by inserting time-delays into one of the data streams. A plot of the number of coincidences vs. time delay was expected to show a peak at zero, as indeed such plots (e.g. Fig. 5) very frequently did. Much later, the same technique was used to decide whether there were really time coincidences between gamma ray bursts and power peaks from the bars. Again the answer was frequently yes (Fig. 6). I entered the picture briefly at that point by suggesting that the programmer might also try running the program with a deliberately wrong time for the GRB, in case something in the time-shifting process accidentally enhanced count rates. If this was ever done, I did not see the result.

The largest mass detector then and now was the earth, and it was instrumented to look for excitation of its normal modes by passing gravitational waves (Forward et al. 1961). The upper limit was actually low enough to be able to say that gravitational radiation at periods of 5–20 min did not close the universe. The detectors did, however, see the 54 min "gravest normal mode" of the earth and a few others after a devastating Richter 9.5 May 1960 earthquake in Chile.

One of the referees has asked for my "take" on the process by which Joe Weber was "voted off the island". This has been addressed by science writers (Bartusiak 2000), historians of science (especially Franklin 1994, 2010), a sociologist (Collins 2004, 2011), speakers at conferences on relativity and gravitational radiations (e.g. Shaviv and Rosen 1975), and websites (Aufmuth, accessed spring 2016, is particularly unpleasant, partly because Weber is represented by a single, unflattering photograph,

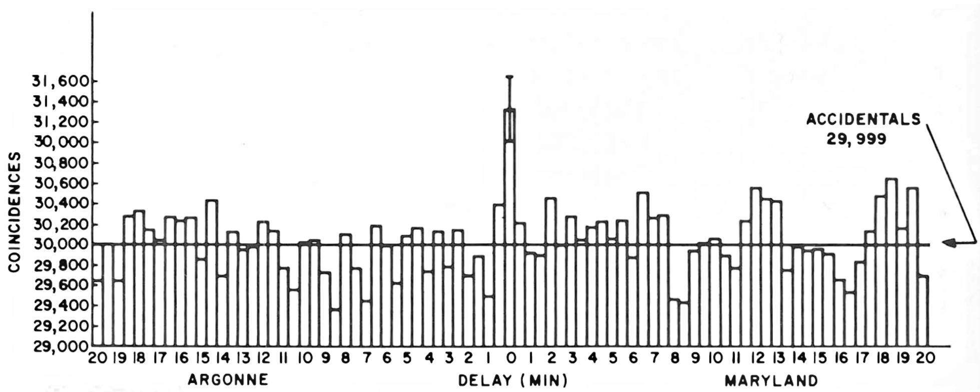


Fig. 5. Time delay histogram, giving coincidences vs. time delay for continuous operation of the Argonne and Maryland bars, August 3–October 17, 1974. Bins are 0.1 seconds wide (despite the coordinate label). From Weber (1980), on p. 458.

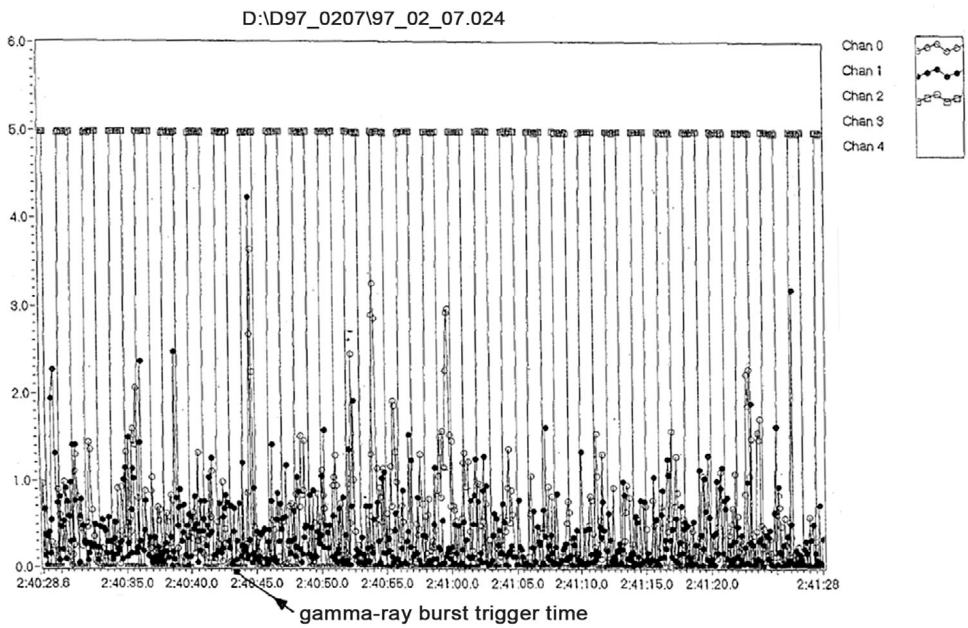


Fig. 6. Plot of output of two bars (shown as filled and open circles) for about one minute around the trigger time of a gamma ray burst. From an overhead transparency in Weber's files at the time of his death. Data analysis related to Weber and Radak (1996).

while all the other groups move forward with flattering ones). I am obviously not the right person to ask. No one wants "a critic on the hearth" (Joe built beautiful fires, by the way, and could also drive large trucks as well as ships and carrier-based aircraft). I have three specific memories of that 1974 GR7 conference in Israel: (1) the sight of a dead camel by the road; (2) incoming GRG president Nathan Rosen getting off the tour bus and helping the driver to change a flat tire; and (3) an incident outside the church of the nativity in Bethlehem that led us to say that the outgoing president of GRG was "the sort of person who would give you the coat off his back, but reluctantly". That we were not all friends and would soon be even less so was not my dominant thought, but rather how lucky I was to be married to someone whose work had spawned so much interest and who could still communicate freely with older Israelis in Yiddish. Our next trip to Israel was for the 1975 Texas symposium, where we met conductor Zubin Mehta and the mayor of Jerusalem, Teddy Kollek.

6 The lunar surface gravimeter

The highpoint, at least financially, of the 1960s and early 1970s was the Lunar Surface Gravimeter, NASA and other funding for which took care not only of equipment purchase and construction, Weber's summer salary and travel, his student/postdoc group and technicians, but also some theoretical students working with other advisors at Maryland.

NASA had wanted a scientific component for the Apollo program from the beginning. Rock samples were an obvious part of this, and their composition contributed to the current "best buy" model of lunar formation. The most clearly productive aspect was the installation and later upgrade of a lunar corner reflector, which has permitted more than 45 years of nearly continuous monitoring of the moon's distance, during which that distance has increased at 3–4 centimeters per year, owing to tidal drag (which also slows the earth's rotation). The proposal for the LSG (whose design was somewhat inspired by the device used to look for triggering of the earth's normal modes) was a response to a request for proposals for Apollo science. The intent was to turn the moon into a gravitational wave detector, with the LSG responding to triggering of the moon's normal modes by passing gravitational waves.

What is the difference between a seismometer and a gravimeter? A seismometer (earthquake detector) is supposed to be driven by solid-body (electromagnetic) forces from its supporting structure. An early Chinese one had a ring of dragon heads around a vase holding jade balls in their mouths. When the device was shaken by a quake wave, a dragon dropped his ball, and his location in the ring indicated the direction from which the wave had come. A gravimeter, in contrast, is isolated as much as possible, by acoustic and electromagnetic filters, from the underlying substrate, and is supposed to respond to changes in local g (which is 9.81 m/s^2 on the surface of the earth).

The earth-based device had used a La Coste – Romberg sensor with the instrument package assembled by Jerome V. Larson (of the EE department) and Weber. The sensor was a mass, spring, and lever system with a period of 20 seconds. It was critically damped and temperature controlled near 50 °C, at which the first derivative of the force constant with respect to frequency vanished. A capacitor plate attached to the mass served as an element of an AC bridge. If local g changed, the mass was accelerated and the bridge became unbalanced, resulting in an error voltage. This voltage was amplified and used to restore the mass to its equilibrium position by means of a closed-loop servosystem and electrical forces. The measured output voltage was recorded and computer-analyzed to look for the frequencies of the earth's normal modes. This device worked as designed (Weber and Larson 1966).

Under NASA rules, however, the LSG had to be built by American industry. Bendix corporation was the contractor selected, and Weber, as PI, was not allowed to touch the construction process. Sadly, the LaCoste – Romberg sensor was assembled with a misunderstanding of the value of g on the moon (it is NOT 9.81 m/s^2, which is why Apollo astronauts could leap tall buildings at a single bound). The LSG was emplaced by Apollo 17 astronauts (the last team). Data were multiplexed with other Apollo 17 ALSEP instruments and sent to Earth stations, recorded on "range tapes", from which NASA employees extracted the data for each separate experiment and turned it over to the PIs.

It was almost instantly obvious that something was wrong, which rendered the channels with tidal data and free mode data almost useless. But the high frequency seismological data could still be used after the beam was rebalanced by a command radioed from earth to the instrument, though this changed the resonance characteristics of the detection mechanism.

At this point Mr. Russell L. Tobias (from whose account, Tobias 2013, this material is largely taken) joined the team, with the task of analyzing what was coming from the LSG and looking for coincidences with output from the College Park bars. The 7-track tapes were difficult to read, and Tobias with a representative from Lockheed Electric Company managed to improve NASA's tape drive maintenance procedures. Other people involved were John Gigante, an electrical engineer, Bruce Weber, a data technician, and several electronic technicians contracted through Pulse, Inc, including Ms. Pota Fitzgerald, and the senior programmer, Brian Reid, who left to earn a graduate degree in computer sciences elsewhere.

The university of Maryland's central Univac 1108 computer system was inadequate for the data processing, and NASA provided a dedicated DEC PDP-11. Joe regarded that computer as dubious, and described NASA's reclaiming it as analogous to the story "The Ransom of Red Chief"[5]. Russell, however, spent a final, successful all-night session with it and managed to accumulate enough processed data for his thesis, which focussed on the comparison between lunar acceleration and the aluminum bar events.

Any number of footnotes could be added to this tale, of which I pick out two. First, Russell was already reasonably certain he was not aiming for a career in academe. He has been a successful member of private industry throughout the interim. The second is that his father and Joe Weber had been lab partners during their freshman year at Cooper Union (1935–36), though the event apparently did not leave very happy memories on either side. The senior Tobias completed a degree in chemical engineering in 1939 and a masters at Brooklyn Polytechnic, but did not complete a doctorate at George Washington University. He invented an early form of artificial rubber used by soldiers during WWII but Joe, of course, was in the Navy.

I suppose there is at least one other lesson from the LSG 1969-present tale. With modern (2013 not 2023!) search techniques, someone with an unusual name is much easier to find that someone with a common one. I located Russell within 24 hours of deciding to try to find him for the NASA studying-old-moonquakes project, but there are several other Joseph Webers to be found trolling compilations of physics papers, including one who works on gravitational radiation. Idle browsing of my own name brought up six different obituaries of women named Virginia Trimble, none of them, fortunately, mine, though one had lived in Kissimmee Florida, where I was about to go that day.

Incidentally, although Joe Weber is generally now perceived as having been a rather solitary person, and a very large fraction of his papers are single-author, the 1980 overview thanks 22 people for contributions to the design and construction of the experiments, data analysis, and helpful discussions. Scanning other sets of acknowledgements, co-authors, and my own memory brings that total up to something like 45 or 50. Of the pioneering groups, Robert Forward died in 2002; David Zipoy is apparently living in retirement in Florida; and I have no idea what became of most of those involved in the project, but the most faithful was Darrell Gretz, the technician who was the last person to drive Joe from the Chevy Chase house to the Gravity Building and back (except for me) as he was dying. It is perhaps not totally irrelevant

[5] "Red Chief" was a small boy whose parents had found him so tiresome that, when he was kidnapped, they declined to pay the ransom. The kidnappers found him so tiresome they eventually paid the parents to take him back.

that our annual parties (1973–91) on each coast tended to attract about 100 people each, a large fraction of them scientists, and, of course, significant others.

7 The aftermath

For a period of about ten years, the relativity, physics, and other scientific communities expressed very considerable interest in Weber's work. This was manifest in an enormous number of invitations to give talks at conferences and in physics departments and in invitations to write review and scientific articles. The award of the First Prize in the essay competition of the Gravity Research Foundation (Weber 1959a) came at the beginning of this period, and the Babson Award (1970) from the same Foundation[6] and the Boris Pregel Prize of the New York Academy of Sciences for research in physics and/or astronomy (1973) near the end of it. In between came his 1971 election to the International Committee on General Relativity and Gravitation, which would become the governing body of the International Society on General Relativity and Gravitation in 1974. A highlight of the 1971 (Copenhagen) meeting in my mind came when the Russians stood up together and walked out of the business meeting, because they had received their visas as scientists not as voters.

I realize now that I actually witnessed the transition from general interest to widespread distrust. Weber gave an invited talk on his results from the operation of a widely-separated pair of "Weber bars" at that 1971 Copenhagen meeting (GR6). After his presentation, a group of young postdocs from the Institute of Theoretical Astronomy in Cambridge tried to figure out what might be going on. He had reported the events as having come from the general direction of the galactic center, but bar detectors have a front-back symmetry, so that the pulses could have been coming from the opposite direction, which is very close to the direction to the Crab Nebula with its active pulsar. We had intended to ask him the next day for some details of the frequency response of the bar, on the grounds that the Crab pulsar, as it slowed, might be passing through a submultiple of the bar frequency. The question never got asked, because Anita Straus Weber had died that day, and Joe took off immediately from Copenhagen to return to the US.

At the 1970 Texas Symposium on Relativistic Astrophysics in Dallas, he described an additional detector, a massive aluminum disk (Fig. 7) whose radial "breathing" mode would be excited if gravitational waves had a dipole component. This is zero in general relativity, but might have been 7% or so of the quadrupole power if the Brans-Dicke (1961) scalar-tensor theory of gravitation had been correct[7].

Weber's invitation to give an endowed lecture at the University of Southern California in early February 1972 was also near the transition point. This was the occasion for our first real meeting (which led to marriage on 16 March 1972). I participated in a small fraction of the data processing over the next couple of years.

By the time of the 6th Texas Symposium (December 1972 in New York), it was clear that portions of the community were no longer supportive. Joe thought it might moderate the hostility of his critics if I gave the presentation for the group (Trimble and Weber 1973). This, to put it mildly, did not happen (Weber et al. 1973,

[6] Roger Babson regarded gravitation as an obstacle to be overcome, and his foundation was originally aimed at anti-gravity. This hasn't happened so far, though the annual prize essay competitions continue, and the cash prizes have come at useful times for some of the winners, like the young Stephen Hawking (Kaiser 1987).

[7] The situation was typified by a then-young, very bright (now distinguished and retired) strong supporter of GR who remarked to me that he could see where Dicke was wrong, but Weber had him worried.

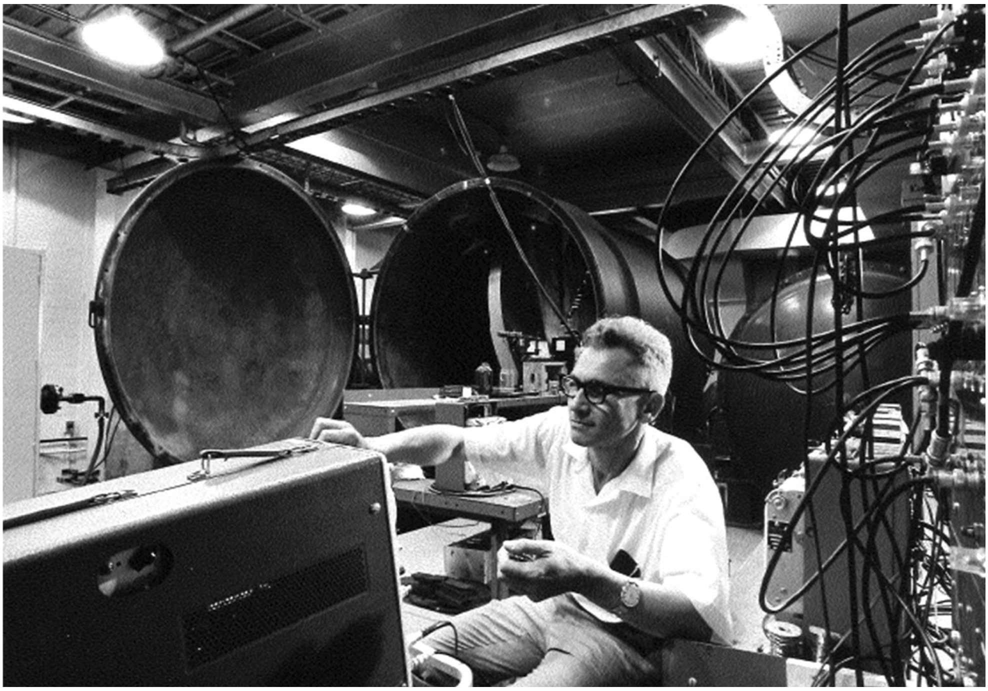


Fig. 7. Foreground: Weber with archaic imaging device; Background: the disk antenna in an open vacuum chamber.

Tyson 1973). It was then 30 years before I gave another invited "Texas" talk, and before the topic of gravitational radiation made it back into the plenary program (Schutz 2003), though there had been parallel sessions on the topic in the interim. Peter Bergmann had just died (October 19, 2002) at the time of that 21st Texas Symposium. Schutz did not cite Weber, and his name got a laugh when I responded to a questioner, who had said that something in my talk about X-rays from supernovae was "controversial", "You don't know what controversial means unless you've been married to Joe Weber for 28 years" (Trimble 2003, also my first time back on the Texas program in 30 years).

Joe was, of course, not unaware of the changing intellectual climate. It was probably fall, 1974 when he walked into the dining room, saying "Poor Sweetheart! Her husband thinks he's discovered gravity waves and sold the idea to Howard Hughes for a lot of money!" Grin. Check held up. It was $15,000, which was then a good deal of money. It was divided equally, at Joe's decision, among him, Bob Forward, and Dave Zipoy, then in the Maryland astronomy program. Forward was at Hughes, and had arranged the sale of the patents with the idea that a rotating quadrupole, operated in the near-field mode, might detect underlying oil deposits of lower density than the surrounding rock.

In the period 1965–75 or thereabouts, when Joe was asked to give very large numbers of colloquia, seminars, and conference talks, he developed a "tour" version. I first heard it at USC in February 1972 and again at UCI in early March that year. Anonymous referee II remembers a 1975 version at Louisiana State University as "one of the most memorable seminars I ever attended" (Well, we were all a good deal younger then). Weber would start with the Einstein metric and go very quickly from memory, through the Christoffel symbols and on to the Riemann and Ricci tensors and then the quadrupole formula for emission by a pair of orbiting masses, M, in a form like

$$L = \frac{32G}{5c^5}\omega^6 \mu a^2 \tag{1}$$

where μ = reduced mass, $(M_1M_2)/(M_1 + M_2), a$ = separation of the pair, and ω = their angular frequency. He called attention to the G upstairs and c^5 downstairs, suggesting that the resultant emission must be very weak, but then, again quickly and from memory, replaced the separation, a, by two Schwarzschild radii and the ω by the angular frequency of two star-sized masses at that separation. Magically, c came upstairs and G downstairs, and, the speaker suggested with a broad grin, the power emitted might not after all be so very small.

The referee supposed that the idea of using Schwarzschild radii (that is, a binary black hole) came from John Wheeler, during one of Weber's visits to Princeton. Impossible, of course, to prove a negative, but I think perhaps not, Weber (1961a) has roughly the Landau-Lifshitz (1951) version of equation (1). And in a review (Weber 1980) he thanks J.R. Oppenheimer and F.J. Dyson for encouragement during his 1962–63 stay and cites Dyson (1963) for the formula. It is necessary to distinguish "black hole" in the sense of something whose size is essentially its Schwarzschild radius and the singularity or whatever else might be going on at the center. Weber said frequently that he did not think singularities occurred in nature and that Einstein, if he had realized that GR predicted such singularities, would have abandoned the theory.

But the first search for "frozen stars", meaning ones at their Schwarzschild radii, was the work of Zeldovich and Guseinov (1966). That normal binary evolution could produce such systems was well known by 1971 (Paczyński 1971). The first observed black hole accretor, Cygnus X-1, dates from 1972, and Joe was actually a co-author on a 1973 paper (Trimble et al. 1973) that tried to push the accretor mass down into the neutron star range. Weber and Zeldovich had a warm relationship that dated back at least to the Warsaw GRG.

As community doubts grew, the funding climate, of course, also changed. NSF support dropped to \$50 000 per year in 1975, the same time frame in which the agency began supporting the project that became LIGO. Some years, increasingly, it was zero. Late dollar pulses came for an attempt at cryogenic bar detectors, whose noise should, of course, be much smaller. The one at He^4 temperature (near 4 K) ran briefly, but experienced noise from the boiling helium. Other groups attempted indium-plated aluminum, which crackled, and pure indium when struck goes thud, not ring. A helium dilution refrigerator, intended to bring the operating temperature down to milli-Kelvin, never worked properly. The manufacturer eventually took it back. And He^3 is one of the most expensive substances on earth. A presentation on "Development of cryogenic gravitational wave antennas at the University of Maryland" was given at the 8th International Conference on General Relativity and Gravitation in Waterloo Canada on August 10, 1977. The authors were listed as W. Davis, D. Gretz, J.P. Richard, and J. Weber (who actually gave the talk). Proceedings were never published.

The morning we were to leave for the meeting, the airline phoned to say that Canadian air services were on strike and we would be dropped at Rochester NY. Joe instantly phoned the airport there, reserved nearly the last rental car available. We flew, landed, drove across the border, and were in time for the opening reception (always the best part of a conference, I think). GRG's host there was Werner Israel, who under the previous GRG rubric would have become president of the society in succession to 1974 host Nathan Rosen, for the next three years, but declined the office, which has since been an elective one. We even saw Niagara Falls.

Still later, NASA-Goddard provided some funding for Weber to look for coincidences between bar events (the two large ones were by then long since both in the gravity building) and gamma ray bursts, whose extra-galactic character had not yet quite been established. The funding was primarily for the support of a postdoc, Bronislav Radak, borrowed from the high-energy group, to process the data, with the understanding that he knew very little about general relativity or the purposes of the experiment, and so could do the processing in a truly blind fashion (S. James Gates 2016 email personal communication). Coincidences between bar pulses and GRBs were reported by Weber and Radak (1996) covering the period 1990–91 (Fig. 6).

8 Neutrino detectors and a new cross section for bars?

Meanwhile, however, a (probably) well-meaning colleague set Weber off in a rather different direction. It was spring (about 1980) because we were in California, visiting Caltech for a colloquium or something and lunching in the mostly-student cafeteria variously called Chandler Dining Hall and "The Greasy" (much less so than when I was a grad student there in the 1960s). At the same table were a couple of physics graduate students and Richard Feynman. Joe was trying to explain something about how he thought the bar detectors worked. Feynman, characteristically impatient, said something along the lines of "oh, why don't you give up on gravitational waves; go look for neutrinos or something".

Joe took this as serious advice (I don't know that it wasn't), and started thinking back to single, large, perfect crystals, like pink ruby, quartz, sapphire, and silicon for bars of inches rather than feet. The idea was that they might scatter in an analogy to Mössbauer scattering of gamma rays, that is, recoiling as a whole, but with an interaction cross section that was proportional to the number of dipoles among the atoms or molecules. Weber (1981a, 1984, 1985b) are discussions of the theory. An early referee reported that he couldn't prove the analysis was wrong and that Weber seemed to have invented s-matrix theory by a non-standard method. Joe then went back and did the calculation in standard notation, getting the same answer.

Neutrino sources used included tritium in titanium tritide (12 keV), reactors at both UC Irvine and the National Bureau of Standards (now NIST, 1.6 Mev), and the sun (0–430 keV). Results appear in many conference proceedings, but Weber (1984) is the easiest to access. It describes the detectors in detail. All are torsion balances such that a deflection unbalances a radio frequency bridge. That signal is amplified, used to restore the balance to rest position, and recorded as the signal (Fig. 8). The tritium and solar experiments took place inside a wood building, with extra seismic and EM shielding, and extra temperature control, inside the Gravity Building. The tritium case used two 13-gram (65 carat) colorless, clear sapphire crystals, and the source was cycled back and forth in front of the balance. In the reactor case, a larger sapphire was used as a shield some of the time, and it indeed seemed to block the neutrinos. The solar case was in effect an Eötvös experiment, because the torsion balance had one mass of sapphire and one of lead, and was seen to twist in time with the direction to the sun being aligned perpendicular to the face of the crystals.

The paper cites standard texts (Compton and Allison on Nuclear Physics, Jackson on E&M, Yang and Lee on weak interaction physics, Feynman and Hibbs) and also specific contributions by Lamb, Mössbauer, Eötvös, Dicke, and Braginsky (who was also one of those who built a bar detector and remained friendly). Individuals thanked include Gregory Wilmot, the programmer, Larry Spruch and Syd Bludman for theoretical advice, Ray Davis of the original solar neutrino experiment for inspiration, Frank Desrosier, John Giganti, and Jay Kimbell for constructing the apparatus and electronics. Funding had come from the NSF, DARPA, the Defense Nuclear Agency, and the Strategic Defense Initiative (Star Wars) Office of Innovative Science and Engineering. The defense connection was the possibility of using small, portable neutrino (or anti-neutrino, they were the same for this experiement, as were all three flavors)

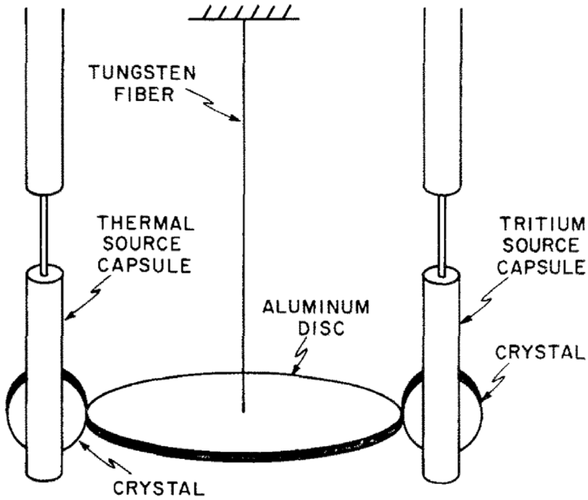


Fig. 8. Design of the neutrino detector used with a titanium tritide source. Because the expected force on the right hand crystal from the neutrinos pushing out would be nearly equal and opposite to the gravitational force, the dummy on the left was made to have equal mass and to radiate as much heat as would the decaying tritium. The two source capsules were lifted up and down together in cycles of several minutes, halting at the top and opposite the crystals. A radiofrequency bridge became unbalanced whenever the torsion balance twisted. The unbalance voltage was amplified and both recorded as the signal and used to restore to balance to its equilibrium position. The whole thing was, of course, enclosed in a vacuum chamber, maintained at about 10^{-6} Torr. From Weber (1984), on p. 1197.

detectors to locate nuclear submarines from above the ocean. The tritium sources proved to have half-lives much less than 12 years (counted in the lab and as neutrino sources). The electrons from a few decays (which did not get out of the shielding) heated the H^3 gas which expanded until the capsule leaked (Weber 1984).

More proposals went in, some with my name suggested as co-investigator. Joe had by then passed 70 and, under the rules of the time, been forced to retire. But, once again, funding and interest from the community dropped more or less to zero, though in 2002, a couple of scientists from Pirelli Tire Company came to Maryland and took away many pounds of neutrino detector equipment. I secretly kept one small pink ruby and one PZT crystal from the bar detector era as souvenirs. Other sourvenirs? Well there was an even larger clear, but not colorless crystal that Joe said had been intended for an engagement ring for an elephant, but ended up as a pendent for me. And as a result of the "Star Wars" connection, I got to dance a Viennese waltz with Eugene Wigner, a claim very few full professors of physics can make.

Weber (1989 and many other conferences) came to think that some similar coherent scattering process might enormously enhance the cross section for the bar detectors for gravitational waves, accounting for his positive results and the negative ones from groups that tried different arrangements. The various neutrino experiments had yielded signals very much like the calculated ones, so why shouldn't the bar detectors?

And then along came Supernova 1987A (A because it was the first found that year, not because of its importance). At the moment the neutrino and electromagnetic signals started reaching earth on February 23rd, none of the better-supported cryogenic or other bar detectors were up and running, so there were records only from the two at Maryland and one at Rome. Results appeared mostly in conference proceedings (Weber 1988b, 1990, 1994 is only a subset), and, worse luck, the gravitational radiation pulses were time-coincident not with the universally-accepted burst of neutrinos recorded at Kamioka and at IMB (and probably also at Baksan), but with the 5 neutrino-like events recorded by the Mt. Blanc detector about 5 hours earlier. A few theorists at the time attempted two-stage super-nova models, for instance collapse first to a neutron star and then to a black hole (see Trimble 1988 for many early references), but once again interest quickly waned. I cite here only Amaldi et al. (1988) to clarify that the Italian involvement had originally included a very high profile physicist.

After 30 September 2000, I received very many expressions of sympathy from physics and astronomy colleagues; one from a member of the Rome group (who would I imagine now prefer to remain anonymous) said that they were again finding a correlation with the galactic center and would be announcing this soon. I don't think they ever did, but see Galeotti and Pizzella (2016) on SN 1987A results.

What became of everything? The two largest bars remained on-line in the Gravity Building until Joe died; the last data tape ran out a day or two later. One of the bars is now in an exhibit at the Hanford site of the LIGO detector. I gave a colloquium talk there at the time it was dedicated. The original baby bar was given to the Smithsonian for an Einstein centenary exhibit in 1978. It is probably still somewhere in their storage room. Three other bars are still in storage at the University of Maryland, thanks to the kind offices of Lorraine DeSalvo, who is considering that they might be made into some sort of large art object. The disk antenna, hundreds of data tapes, and decades of bound notebooks presumably found their way into recycling. The last notebook was in the Gravity Building, about half full, and recorded the last year or so of measurements of the equivalent noise temperatures of the bars, which had gradually climbed as the cement holding the crystals on aged and cracked.

Coming down almost to the present, in May, 2014 Akira Banchi, associated with the Japanese TV organization NHK, wrote asking my permission to use some film footage of Weber working on one of the bars, taken in the early 1970s and used in a 2005 NSF documentary. I have not seen either the documentary or what NHK did with the footage, but of course I gave permission.

Logically last, though not quite chronologically, the Lunar Surface Gravimeter had a sort of afterlife. In 2013 and 2014 I heard from an Apollo ALSEP Missing Data Focus Group, involving people at NASA, Univ. of Colorado, Univ. of Maryland, Rice, UCLA, the University of Arizona Lunar and Planetary Lab, Texas Tech, MIT, and elsewhere. They were attempting to recover as many as possible of the original data tapes from experiments in the Apollo Lunar Science Experiment Packages, of which the LSG had been one in the last, December 1972, Apollo 17 flight. A happy outcome of this for me was reconnection with Dr. Russell Tobias, who, as a Maryland graduate student, had taken primary charge of analyzing the LSG tapes up until the time funding was withdrawn and who also prepared the final report on the project. Some of the information he provided appears in Section 6.

9 Gravitational waves today

The official announcement of the first event detected by LIGO came while I was writing this and is to be found as Abbott et al. (2016). A number of other papers are in press, on arXiv, or in preparation; a second event is being discussed by official LIGO speakers at conferences and colloquia; and a third is rumored. I was present at the official NSF 11 February LIGO press conference, at the kind invitation of NSF director Dr. France Cordova, a friend since she was a graduate student at Caltech. Press coverage was, of course, widespread, in *Nature*, *Science*, and all the rest. My favorite discussion was that of Bartusiak (2016) because she mentions Weber's work with reasonable charity, as she did in her book (Bartusiak 2000).

All the relevant conferences (and some not so relevant ones) and very many department colloquium series are arranging LIGO talks. This time, there are 1004 collaborators to share the task, well perhaps only 1001, because the initial paper includes three authors who died before publication. In fact, even I have been asked to give a few talks, and I was asked on very short notice to provide a nomination of LIGO folks for a foundation prize a few days before the prize committee was to meet. They won, with the three leaders I had suggested to the fore; and have since won at least three additional major prizes with the same three people on top. Some of the information needed to make Ronald Drever part of the lead trio came from Collins (2004).

Many groups are tooling up or activating earlier plans for "multiwavelength astrophysics", that is, attempts to locate electromagnetic, neutrino, or cosmic ray counterparts for GW bursts. Indeed I am part of one of these groups, the "transient and variable source working group" for the Large Synoptic Survey Telescope. But I am still a widow and still miss beyond words our daily "after breakfast hug" and the voice caroling out as steps ascended the stairs, "I wish to announce my safe return!"

Acknowledgements. I am indebted to Drs. Robert L. Forward and Vera Cooper Rubin for, respectively, professional and personal introductions to Joseph Weber, to Dr. Russell Tobias for his account of the Lunar Surface Gravimeter, to Dr. Fred M. Johnson (a Townes student at Columbia) for some details of the maser project there, and to the editorial board of EPJH for the invitation to compile this well-timed history. The editors of the special issue, referee Allan Franklin, and anonymous referee II (hi David!) contributed some important missing references (or requests for them!) and other useful information. MPIWG student Bendix Düker very bravely and very expertly turned a scruffy typescript into the well-formatted article you see before you. Joe and Anita's granddaughter Dr. Elizabeth Weber Handwerker located school annual pages and other information about her grandparents for which I am grateful. My deepest debt is, of course, to Joseph Weber, truly as the Navy says in bestowing commissions, an officer and a gentleman, and my best friend for more than 28 years.

Note added in proof. On 1 November 2016, Prof. Jayanth R. Banavar, Dean of the University of Maryland College of Mathematical, Computational, and Natural Sciences, organized there a gravitational waves festivity. Of the old gang, at least Darrell Gretz and John Giganti were there, alive and well. My talk (the previous day) was an abbreviated version of this paper. Darrell has recently written up his memories of the years working with Weber and is of the opinion that the bar detectors were responding to some real physical phenomenon.

References

Abbott, B.P. et al. 2016. Observation of Gravitational Waves from a Binary Black Hole Merger. *Phys. Rev. Lett.* **116**: 061102-1–061102-14.

Adler, R.J., H.J. Bazin and H. Shifter. 1975. *Introduction to General Relativity*, 2nd Ed. McGraw Hill.

Amaldi, E., P. Bonifazi, S. Frasca, M. Gabellieri, D. Gretz, G.V. Pallottino, G. Pizzella, J. Weber and G. Wilmot. 1988. Analysis of the data recorded by the Maryland and Rome room temperature gravitational wave antennas in the period of the SN 1987A. In M. Kafatos and A.G. Michalitsianos, eds. *Supernova 1987A in the Large Magellanic Cloud*, Cambridge University Press, Cambridge, pp. 453–462.

Anonymous. 2016. *The Economist*, 13 February, p. 77.

Aufmuth, P. http://www.geo600.uni-hannover.de/~aufmuth/JoeWeber.pdf ("Joseph Weber 1919–2000 Offizier & Gentleman") accessed October, 2016.

Bartusiak, M. 2000. *Einstein's Unfinished Symphony*. Joseph Henry Press, Washington, DC.

Bartusiak, M. 2016. The long road to detecting gravity waves. *Science News* **189**: 24–27.

Basov, N.G. and A.H. Prokhorov. 1954. Application of molecular beams for radiospectroscopic study of molecular rotational spectra. *Journal of Experimental and Theoretical Physics* **27**: 431–438 (the page numbers are different in the English translation).

Bergmann, P.G. 1962. Allocution de Cloture/Summary of the Colloque International de Royaumont. In Lichnerowicz and Tonnelat (1962), pp. 463–471 in English, 451–462 in French.

Bertolotti, M. 1983. *Masers and Lasers: An Historical Approach.* Adam Hilger, Bristol.

Blum, A.S., R. Lalli and J. Renn. 2015. The Re-invention of General Relativity: A Historiographical Framework for Assessing One Hundred Years of Curved Space-time. *ISIS* **106**: 598–620.

Bloembergen, N. 1956. Proposal for a new type solid state maser. *Phys. Rev.* **104**: 324–327.

Bondi, H. 1957. Plane gravitational waves in general relativity. *Nature* **179**: 1072–1073.

Bondi, H. 1962. On the physical characteristics of gravitational waves. In Lichnerowicz and Tonnelat (1962), pp. 129–135.

Bondi, H. 1965. Some special solutions of the Einstein equations. In A. Trautman, F.A.E. Pirani, and H. Bondi, eds. *Lectures on General Relativity.* Prentice Hall, Englewood Cliffs NS, pp. 375–489.

Bondi, H. and W.H. McCrea. 1960. Energy transfer by gravitation in Newtonian Theory. *Proceedings of the Cambridge Philosophical Society* **56**: 410–413.

Bonnor, W.B. 1957. Non-singular fields in general relativity. *J. Math & Mech.* **6**: 213.

Bonnor, W.B. 1959. Spherical gravitational waves. *Phil. Trans. Roy. Soc. A* **251**: 233–271.

Brans, C. and R.H. Dicke. 1961. Mach's principle and a relativistic theory of gravitation. *Physical Review* **124**: 925–935.

Brill, D. 1959. On the positive definite mass of the Bondi-Weber-Wheeler time-symmetric gravitational waves. *Annals of Physics* **7**: 466–483.

Cameron, A.G.W., ed. 1963. *Interstellar Communication.* Benjamin Press, New York.

Campbell, C. 1960. *The design of a two-level solid state maser.* Ph.D. Thesis, Univ. of St. Andrews.

Collins, H. 2004. *Gravity's Shadow.* Univ. of Chicago Press, Chicago.

Collins, H. 2011. *Gravity's Ghost.* Univ. of Chicago Press, Chicago.

Damour, T. 1987. The problem of motion in Newtonian and Einsteinian gravity. In. S.W. Hawking and W. Israel, eds. *300 Years of Gravitation.* Cambridge University Press, Cambridge, pp. 128–198.

Davis, H., D. Gretz, J.P. Richard and J. Weber. 1977. Development of cryogenic gravitational wave antennas at the University of Maryland. *8th International Conference on General Relativity and Gravitation, Waterloo, Canada, August 10, 1977.*

Dyson, F.J. 1963. In Cameron (1963), p. 115.

Eddington, A.S. 1923. The Propagation of gravitational waves. *Proc. Roy. Soc. A* **102**: 268–282.

Eddington, A.S. 1924. *The Mathematical Theory of Relativity*, 2nd Ed., Cambridge University Press, Cambridge, sect. 57.

Einstein, A. 1916. Näherungsweise Integration der Feldgleichungen der Gravitation. *Preuss. Akad. Wiss. Berlin, Sitzungsberichte der Physikalisch-mathematischen Klasse*: 688–696.

Einstein, A. 1917. Kosmologische Betrachtungen zur allgemeinen Relativitätstheorie. *Preuss. Akad. Wiss. Berlin, Sitzungsberichte der Physikalisch-mathematischen Klasse*: 142–152.

Einstein, A. 1918. Über Gravitationswellen. *Preuss. Akad. Wiss. Berlin, Sitzungsberichte der Physikalisch-mathematischen Klasse*: 154–167.

Einstein, A. and N. Rosen. 1937. On Gravitational Waves. *Journal of the Franklin Institute* **223**: 43–54.

Einstein, A., L. Infeld and B. Hoffmann. 1938. The gravitational equations and the problem of motion. *Ann. Math.* **59**: 65–100.

Feynman, R.P. 1962–63. Lectures on Gravitation, notes taken, duplicated and distributed by Fernando B. Moringigo & William G. Wagner.

Fierz, M. and W. Pauli. 1939. Relativistic wave equations for particles of arbitrary spin in an electromagnetic field. *Proc. Roy. Soc. A* **173**: 211–232.

Forward, R.L. 1971. Multidirectional multipolarization antennas for scalar and tensor gravitational radiation. *General Relativity and Gravitation* **2**: 149–159.
Forward, R.L. 1978. Wideband laser-interferometer gravitational-radiation experiment. *Physical Review D* **17**: 379–390.
Forward, R.L., D. Zipoy, J. Weber, S. Smith and H. Benioff. 1961. Upper limit for interstellar millicycle gravitational radiation. *Nature* **189**: 473.
Foures-Bruhat, Y. 1956. Sur l'intégration des équations de la relativité générale. *J. Rat. Mech. Anal.* **5**: 951–966.
Franklin, A. 1994. How to avoid the experimenters' regress. *Studies in the History and Philosophy of Modern Physics* **25**: 463–491.
Franklin, A. 2010. Gravity Waves and Neutrinos: The Later Work of Joseph Weber. *Perspectives in Science* **18**: 119–151 and references therein.
Galeotti, P. and G. Pizzella. 2016. New analysis for the correlation between gravitational wave and neutrino detectors during SN1987A. *Eur. Phys. J. C* **76**: 426.
Glanz, J. 2000. Obituary of Joseph Weber. *New York Times*, 9 October, p. A19.
Good, W.E. 1946. The Inversion Spectrum of Ammonia. *Phys. Rev.* **70**: 213–218.
Gordon, J.P., H.J. Zeiger and C.H. Townes. 1954. The maser – New type of microwave amplifier, frequency standard, and spectrometer. *Phys. Rev.* **95**: 282–290.
Gutfreund, H. and J. Renn. 2015. *The Road to Relativity.* Princeton Univ. Press, Princeton, NJ.
Hartle, J.P. 2003. *Gravity.* Addison Wesley, San Francisco, CA.
Hirakawa, H. and K. Narihara. 1975. Search for Gravitational Radiation at 145 Hz. *Phys. Rev. Lett.* **35**: 330–334.
Hockey, T. et al., eds. 2014. *Biographical Encyclopedia of Astronomers*, 2nd Ed. Springer, New York, NY, pp. 823–825.
Homans, J. 2012. Tony Judt: A Final Victory. *New York Review of Books*, 22 May.
Infeld, L. 1936. The New Action Function and the Unitary Field Theory. *Proc. Cambridge Philos. Soc.* **32**: 127–137.
Infeld, L. 1937. A new group of action functions in the unitary field theory. II. *Proc. Cam. Phil. Soc.* **33**: 70–78.
Infeld, L. 1938. Electromagnetic and gravitational radiation. *Phys. Rev.* **53**: 836–841.
Infeld, L. 1954. On the motion of bodies in general relativity theory. *Acta Physica Polonica* **13**: 187–204.
Infeld, L. 1956. On equations of motion in general Relativity Theory. *Helvetica Physica Acta* **29**: 206–209.
Infeld, L. 1957. Equations of motion in general relativity and the action principle. *Rev. Mod. Phys.* **29**: 398–411.
Infeld, L. 1959. Equations of motion and gravitational radiation. *Ann. Phys.* **6**: 341–367.
Infeld, L. and R. Michalska-Trautman. 1960. The two-body problem and gravitational radiation. *Ann. Phys.* **55**: 561–575.
Infeld, L. and J. Plebanski. 1960. *Motion and Relativity.* Pergamon Press, Oxford UK.
Infeld, L. and A.E. Scheidegger. 1951. Radiation and gravitational equations of motion. *Canadian J. Math.* **3**: 195–207.
Infeld, L. and A. Schild. 1949. On the motion of test particles in general relativity. *Rev. Mod. Phys.* **21**: 408–413.
Infeld, L. and P.R. Wallace 1940. The equations of motion in electrodynamics. *Phys. Rev.* **57**: 797–806.
Johnson, F.M. 2016. Personal communication.
Kaiser, D. 1987. Roger Babson and the Rediscovery of General Relativity. In *Making Theory and Theorists in Postwar America.* Ph.D. Dissertation, Harvard University, pp. 567–595.
Kastler, A. 1985. Birth of the maser and laser. *Nature* **316**: 307–309.
Kennefick, D. 2007. *Traveling with the speed of thought: Einstein and the quest for gravitational waves.* Princeton Univ. Press, Princeton, NJ, esp. pp. 61–65.
Landau, L. and E.M. Lifshitz. 1951. *The Classical Theory of Fields.* Addison Wesley, San Francisco, CA, Ch. 11 (and many other editions in many languages).

Lee, M.D., D. Gretz, S. Steppel and J. Weber. 1976. Gravitational Radiation Detector Observations in 1973 and 1974. *Phys. Rev. D* **14**: 893–906.

Levi-Civita, T. 1917. Realtà fisica di alcuni spazi normali del Bianchi. *Rendiconti della Reale Accademia del Lincei* **26**: 519–531.

Lichnerowicz, A. 1955. *Théories Relativistes de la Gravitation et de l'Electromagnétisme.* Masson, Paris.

Lichnerowicz, A. and M. Tonnelat, eds. 1962. *Theories Relativistes de la Gravitation* (Proceedings of the 1959 Royaumont conference) CNRS, Paris.

Loinger, A. 2003. Non-existence of gravitational waves. The stages of the theoretical discovery (1917–2003). arXiv:physics/0312149.

Loinger, A. and T. Marsico. 2016. Email, 6 March 2016.

Marshall, S.A. and J. Weber. 1957a. Plane parallel plate transmission line Stark microwave spectrograph. *Rev. Sci. Instrum.* **28**: 134–137.

Marshall, S.A. and J. Weber. 1957b. Microwave Stark effect measurement of the dipole moment and polarizability of carbonyl sulfide. *Phys. Rev.* **105**: 1502–1506.

Menzel, D.H. 1937. Physical Processes in Gaseous Nebulae I. *Astrophys. J.* **85**: 330–339.

Misner, C.W., K.S. Thorne, and J.A. Wheeler. 1973. *Gravitation.* W.H. Freeman, San Francisco, CA.

Møller, C. 1952. *The Theory of Relativity.* Oxford University Press, Oxford UK.

Moss, G.E., L.R. Miller and R.L. Forward. 1971. Photon-noise-limited laser transducer for gravitational antenna. *Appl. Opt.* **10**: 2495–2498.

Paczyński, B. 1971. Evolutionary Processes in Close Binary Systems. *Ann. Rev. Astron. Astrophys.* **9**: 183–208.

Peebles, P.J.E. 2016. Robert Dicke and the naissance of experimental gravity physics, 1957–1967. *Eur. Phys. J. H*, Doi:10.1140/epjh/e2016-70034-0.

Peebles, P.J.E., L.A. Page and R.B. Partridge. 2009. *Finding the Big Bang.* Cambridge University Press, Cambridge, pp. 6 & 181.

Petrov, A.Z. 1954. Classification of spaces defining gravitational fields *Sci. Notes Kazan State Univ.* **114**: 55–69.

Petrov, A.Z. 1962. Classification invariante des champs de gravitation. In Lichnerowicz and Tonnelat (1962), pp. 107–112.

Pirani, F.A.E. 1957. Invariant formulation of gravitational radiation theory. *Phys. Rev.* **105**: 1089–1099.

Poincaré, H. 1905. Sur la dynamique de l'électron. *Comptes Rendus Hebdomadaires de l'Académie des Sciences de Paris* **140**: 1504–1508.

Purcell, E.M. and R.V. Pound. 1951. A nuclear spin system at negative temperature. *Phys. Rev.* **81**: 279–280.

Richard, J.P. 1976. Sensor and suspensions for a low-temperature gravitational wave antenna. *Rev. Sci. Instrum.* **47**: 423–426.

Robinson, I. and A. Trautman. 1960. Spherical Gravitational Waves. *Phys. Rev. Lett.* **4**: 431–432.

Rosen, N. 1956. Gravitational waves. In A. Mercier and M. Kervaire, eds. *Jubilee of Relativity Theory.* Helvetica Physica Acta, Supplementum IV, Birkhausen Verlag, Basel, pp. 171–175.

Scheidegger, A.E. 1953. Gravitational motion. *Rev. Mod. Phys.* **25**: 451–468.

Schutz, B.F. 2003. LISA and the gravitational wave universe. In R. Bandiera et al., eds. *Texas in Tuscany, XXI Symposium on Relativistic Astrophysics*, World Scientific, Singapore, pp. 91–102.

Shaviv, G. and N. Rosen, eds. 1975. *General Relativity and Gravitation: Proceedings of the Seventh International Conference* (GR7), John Wiley, New York, NY.

Shuler, K.E. and J. Weber. 1954. A microwave investigation of the ionization of hydrogen-oxygen and acetylene-oxygen flames. *J. Chem. Phys.* **22**: 491–502.

Thorne, K.S. 1994. *Black Holes and Time Warps.* W. W. Norton, New York, NY, p. 366.

Tobias, R.L. 2013. Email dated 7 March.

Trimble, V. 1988. 1987A: The greatest supernova since Kepler. *Rev. Mod. Phys.* **60**: 859–871.

Trimble, V. 2000. Obituary of Joseph Weber. *Bulletin of the American Astronomical Society* **32**: 1691–1693.

Trimble, V. 2003. Supernovae: Ground zero and the aftermath. In R. Bandiera et al., eds. *Texas in Tuscany, XXI Symposium on Relativistic Astrophysics*. World Scientific, Singapore, pp. 269–284.

Trimble, V. 2014. Joseph Weber. In Hockey et al. (2014), pp. 2301–2303.

Trimble, V. 2016. Joseph Weber. In B. Wszolek and A. Kuzmica, eds. *Czestochowski Kalendarz Astronomiczny 2016*. Astronomica Nova, Chestochowska, Poland, pp. 171–175.

Trimble, V. and J. Weber. 1973. Gravitational radiation detection experiments with Disk-shaped and cylindrical antennae and the lunar surface gravimeter. In D.J. Hegyi, ed. *Sixth Texas Symposium on Relativistic Astrophysics*, *Ann. NY Acad. Sci.* **224**: 93–100.

Trimble, V., W.K. Rose and J. Weber. 1973. A low-mass primary for Cygnus X-1? *Monthly Notices of the Royal Astron. Soc.* **162**: pink pages 1–4.

Tyson, J.A. 1973. Gravitational radiation. In D.J. Hegyi, ed. *Sixth Texas Symposium on Relativistic Astrophysics*, *Ann. NY Acad. Sci.* **224**: 74–92.

Weber, J. 1951. Pressure broadening of an ammonia inversion line for foreign gases. *Phys. Rev.* **83**: 1058–1059.

Weber, J. 1953a. Amplification of microwave radiation by substance not in thermal equilibrium. *Transactions of the IRE, PGED* **3**: 1–4.

Weber, J. 1953b. Quantum theory of a damped electrical oscillator and noise. *Phys. Rev.* **90**: 977–982.

Weber, J. 1954a. Quantum theory of a damped electrical oscillator and noise II. The radiation resistance. *Phys. Rev.* **94**: 211–215.

Weber, J. 1954b. Vacuum fluctuation noise. *Phys. Rev.* **94**: 215–217.

Weber, J. 1954c. Vacuum fluctuation noise and dissipations. *Phys. Rev.* **96**: 556–559.

Weber, J. 1955. Scattering of electromagnetic waves by wires and plates. *Proc. IRE* **43**: 82.

Weber, J. 1956a. Exact quantum theory solution for the damped harmonic oscillator. *Phys. Rev.* **101**: 1619–1620.

Weber, J. 1956b. Fluctuation dissipation theorem. *Phys. Rev.* **101**: 1620–1626.

Weber, J. 1957. Maser noise considerations. *Phys. Rev.* **106**: 537–541.

Weber, J. 1959a. Gravitational Waves. First Prize Essay, Gravity Research Foundation, New Boston, NH.

Weber, J. 1959b. Masers. *Rev. Mod. Phys.* **31**: 681–710.

Weber, J. 1960a. Detection and generation of gravitational waves. *Phys. Rev.* **117**: 306–313.

Weber, J. 1960b. Phase as a dynamical variable in quantum mechanics. In *Proc. of Rochester Conference on Coherence of Electromagnetic Radiation*.

Weber, J. 1960c. Coherence properties of electromagnetic radiation. In *Proceedings of Johns Hopkins University Conf. on Electronic Countermeasure*.

Weber, J. 1960d. Some aspects on noise in low noise receivers. In *Proc. of MIT Symposium on Low noise Receivers*.

Weber, J. 1961a. *General Relativity and Gravitational Waves*. Interscience Publ. NY.

Weber, J. 1961b. Quantum electronics and new gravitation experiments. In *Proc. of 2nd International Conf. on Quantum Electronics*.

Weber, J. 1962a. On the possibility of detection and generation of gravitational waves. In Lichnerowicz and Tonnelat (1962), pp. 441–450.

Weber, J. 1962b. Theory of methods for measurement and production of gravitational waves. In C. Møller, ed. *Evidence for Gravitational Theories* (Varenna 1961). Academic Press, pp. 116–140.

Weber, J. 1963a. Remarks on gravitational experiments. *Nuovo Cimento* **29**: 930–934.

Weber, J. 1963b. Gravitation and light. In H.Y. Chiu and W.F. Hoffman, eds. *Gravitation*, W.A. Benjamin Inc. NY.

Weber, J. 1963c. Gravitational Waves. In H.Y. Chiu and W.F. Hoffman, eds. *Gravitation*, W.A. Benjamin Inc. NY.

Weber, J. 1964a. Noise considerations in gravitational experiments. *Nuovo Cimento* **30**: 462–464.

Weber, J. 1964b. Gravitational radiation experiments. In C. DeWitt and B. DeWitt, eds. *Relativity Groups and Topology* (Les Houches 1963) Gordon & Breach, New York, NY, pp. 865–882.

Weber, J. 1964c. Gravitation and Light. In H.Y. Chiu and W.F. Hoffmann, *Gravitation and Relativity*, W.A. Benjamin, pp. 90–105.

Weber, J. 1965a. Introductory Remarks: Lasers and free electron amplifiers. *Annals of the NY Acad. of Sciences* **122**: 571–578.

Weber, J. 1965b. Some notes on masers and lasers. *Proc. NY Acad. Sci.* **22**: 832.

Weber, J. 1966a. Gravitational shielding and absorption. *Phys. Rev.* **146**: 935–937.

Weber, J. 1966b. Observation of the thermal fluctuations of a gravitational wave detector. *Phys. Rev. Lett.* **17**: 1228–1230.

Weber, J. 1966c. Gravitational experiments on the lunar surface, Conference Document published as Weber 1967b.

Weber, J. 1967a. Gravitational radiation. *Phys. Rev. Lett.* **18**: 498–501.

Weber, J. 1967b. Lunar gravity investigations. In. E. Burgess, ed., *Physics of the Moon*, Advances in Astronautical Sciences 13, p. 199

Weber, J. 1968a. Gravitational waves, *Physics Today* **21**: 34–39.

Weber, J. 1968b. Gravitational radiation from the pulsars. *Phys. Rev. Lett.* **21**: 295–296.

Weber, J. 1968c. Gravitational-wave-detector events. *Phys. Rev. Lett.* **20**: 1307–1308.

Weber, J. 1969a, ed. *Masers: A Collection of Reprints with Commentary*, Vol. 9. Gordon and Breach, New York, NY.

Weber, J. 1969b, ed. *Lasers: A Collection of Reprints with Commentary*, Vol. 10A. Gordon and Breach, New York, NY.

Weber, J. 1969c. Evidence for the discovery of gravitational radiation. *Phys. Rev. Lett.* **22**: 1320–1324.

Weber, J. 1970a. Gravitational radiation experiments. *Phys. Rev. Lett.* **24**: 276–279.

Weber, J. 1970b. Anistropy and polarization in the gravitational radiation experiments. *Phys. Rev. Lett.* **25**: 180–184.

Weber, J. 1970c. The new gravitational radiation detectors. *Lettere al Nuovo Cimento* **4**: 653–658.

Weber, J. 1971a. Gravitational Radiation Experiments. In C.G. Kuper and A. Peres, eds. *Relativity and Gravitation*, Gordon and Breach, New York, NY, pp. 309–322 (proceedings of a seminar held in Haifa honoring Nathan Rosen).

Weber, J. 1971b. The detection of gravitational waves. *Scientific American* **224**: 22–29.

Weber, J. 1971c. Disc-cylinder Argonne-Maryland gravitational radiation experiments. *Il Nuovo Cimento* **4B**: 197.

Weber, J. 1971d. Experimental test of symmetry of gravitational radiation. *Phys. Lett. A* **34**: 271–273.

Weber, J. 1972a. Advances in gravitational radiation detection. *General Relativity and Gravitation* **3**: 59.

Weber, J. 1972b. Computer analyses of gravitational radiation detection coincidences. *Nature* **240**: 28.

Weber, J. 1977. Gravitational radiation detector observations in 1973 and 1974. *Nature* **266**: 243.

Weber, J. 1980. The Search for Gravitational Radiation. In A. Held, ed. *General Relativity and Gravitation*, Vol. 2, Plenum Publishing Co., New York, pp. 435–467.

Weber, J. 1981a. Exchange of Energy with Large Numbers of Particles. *Phys. Rev. A* **23**: 761–762.

Weber, J. 1981b. New method for increase of interaction of gravitational radiation with an antenna. *Phys. Lett. A* **81**: 542–544.

Weber, J. 1984. Gravitons, neutrinos and antineutrinos. *Foundations of Physics* **14**: 1185–1209.

Weber, J. 1985a. Gravitational wave experiments. In B. Korsonoglu et al. eds. *High Energy Physics* (in honor of P.A.M. Dirac in his Eightieth Year), Plenum Publishing Co., New York, pp. 199–210.

Weber, J. 1985b. Method for observation of neutrinos and antineutrinos. *Phys. Rev. C* **31**: 1468–1475.
Weber, J. 1986a. Gravitational antennas and the search for gravitational radiation. In J. Weber and T. M. Karade, eds. *Gravitational Radiation and Relativity: Proceeding of the Sir Arthur Eddington Centenary Symposium*, Vol. 3, World Scientific, Singapore, pp. 1–77.
Weber, J. 1986b. Coherent scattering of neutrinos and antineutrinos by quarks in a crystal. *American Inst. Phys. Proc.* **150**: 1038.
Weber, J. 1988a. Apparent observation of abnormally large coherent scattering cross section using KeV and MeV energy antineutrinos and solar neutrinos. *Phys. Rev. D* **38**: 32–39.
Weber, J. 1988b. Neutrinos, gravitons, metrology and gravitational radiation. In V. De Sabbata and V.N. Melnikov, eds. *Gravitational Measurements, Fundamental Metrology, and Constants*. NATO ASI Series C230, Kluwer, Dordrecht, pp. 467–500.
Weber, J. 1989. Gravitational antenna bandwidths and cross sections. In B.F. Schutz, ed. *Gravitational Wave Data Analysis*, NATO ASI C253, Kluwer, Dordrecht, pp. 195–200.
Weber, J. 1990. Gravitational radiation antennas: history, observations, and lunar surface operations. In A.E. Potter and T.L. Wilson, eds. *Physics and Astrophysics from a Lunar Base*, *AIP Conf. Proc.* **202**: 159–202.
Weber, J. 1991a. Velocity of propagation of gravitational radiation, mass of the graviton, range of the gravitational force, and the cosmological constant. In A. Zichichi et al., eds. *Gravitation and Modern Cosmology: the Cosmological Constant Problem* (in honor of 75th birthday of Peter Bergmann) Plenum Publishing Co., New York, pp. 17–20.
Weber, J. 1991b. New methods for neutrino detection, and solar neutrino interactions with a single-crystal earth core. In S. Flodmark, ed. *Proc. of Conf. on New Approaches in Geomagnetism and the Earth's Rotation*. World Scientific, Singapore, pp. 199–220.
Weber, J. 1992a. Supernova 1987A; gravitational-wave antenna observations, cross-sections, correlations with six elementary particle detectors, and resolution of past controversies. In A.I. Janis and J.R. Porter, eds. *Recent Advances in General Relativity*, Birkhauser, Boston, pp. 230–240.
Weber, J. 1992b. Gravitational radiation antenna observations, theory of sensitivity of bar and interferometer systems and resolution of past controversies. In N. Sanchez and A. Zichichi, eds. *Current Topics in Astrofundamental Physics*, World Scientific, Singapore, pp. 508–534.
Weber, J. 1992c. Neutrinos and antineutrinos in astronomy and astrophysics. In N. Sanchez and A. Zichichi as above, pp. 560–578.
Weber, J. 1993. Gravitational experiments at supercolliders. In W. Schroeder, ed. *The Earth and the Universe*. Festschrift in honour of Hans-Juergen Trader. Science Editions, Bremen, pp. 439–451.
Weber, J. 1994. Supercollider gravitational experiments. In V. de Sabbata and Ho Tso-Hsuiu, eds. *Cosmology and Particle Physics*, NATO ASI C427, Kluwer, Dordrecht, pp. 271–278.
Weber, J. 1998. Gravitational radiation antenna backgrounds and cross sections. *Physics Essays* **11**: 593–599.
Weber, J. 1999. Correlated gamma ray trigger times with gravitational radiation detector pulses from the bursting pulsar J1744-28. *Physics Essays* **12**: 781–784.
Weber, J. 2000. Gravitational radiation-antenna observations. Submitted, accepted, proofread. If published, in a journal not covered by ADS.
Weber, J. and G. Hinds. 1962. Interaction of Photons and Gravitons. *Phys. Rev.* **128**: 2414–2421.
Weber, J. and K.J. Laidler. 1950. Variations of rate of desorption with extent of surface coverage. *J. Chem. Phys.* **18**: 1416–1418.
Weber, J. and K.J. Laidler. 1951a. Kinetics of the ammonia-deuterium exchange by a microwave method. *J. Chem. Phys.* **19**: 381–382.
Weber, J. and K.J. Laidler. 1951b. Microwave spectroscopic investigations of the kinetics of the heterogeneous ammonia deuterium exchange. *J. Chem. Phys.* **19**: 1089–1096.
Weber, J. and J. Larson. 1966. Operation of LaCoste Romberg Gravimeter at sensitivity approaching the thermal fluctuation limits. *J. Geophys. Res.* **71**: 6005–6009.

Weber, J. and K.E. Shuler. 1954. A microwave investigation of the ionization of hydrogen-oxygen and acetylene-oxygen flames. *J. Chem. Phys.* **22**: 491.

Weber, J. and J.A. Wheeler. 1957. Reality of the cylindrical gravitational waves of Einstein and Rosen. *Rev. Mod. Phys.* **29**: 509–515.

Weber, J. and V. Trimble. 1973. On the response of a gravitational radiation detector to magnetic field fluctuations. *Phys. Lett.* **45A**: 353–354.

Weber, J., V.H. Hughes, P. Kafka, R.W.P. Drever, C.W. Misner and J.A. Tyson. 1973. General discussion on gravitational waves. *Ann. NY Acad. Sci.* **224**: 100–107.

Weber, J., M. Lee, D.J. Gretz, G. Rydbeck, V.L. Trimble and S. Steppel. 1977. New gravitational radiation experiments. *Phys. Rev. Lett.* **31**: 779–783.

Weber, J., V. Ferrari, G. Pizzella and M. Lee. 1982. Search for correlations between the University of Maryland and the University of Rome gravitational radiation antennas. *Phys. Rev. D* **25**: 2471.

Weber, J. and B. Radak. 1996. Search for correlations of gamma ray bursts with gravitational radiation antenna pulses. *Nuovo Cimento B* **111**: 687–692.

Weiss, R. 1972. Lincoln Research Laboratory of Electronics (MIT) Quarterly Report No. 105, 54076.

Weyl, H. 1922. *Space-Time Matter*. Methuen, London.

Weyl, H. 1944. How far can one get with a linear field theory of gravitation in flat space-time? *Am. J. Math.* **66**: 591–604.

Wheeler, J.A. and R.P. Feynman. 1948. Interaction with the absorber as the mechanism of radiation. *Rev. Mod. Phys.* **17**: 157–181.

Zeldovich, Ya.B. and O.H. Guseinov. 1966. Collapsed stars in binaries. *Astrophysical Journal* **144**: 840–842.

Eur. Phys. J. H 42, 293–310 (2017)
DOI: 10.1140/epjh/e2016-70059-2

THE EUROPEAN PHYSICAL JOURNAL H

The binary pulsar and the quadrupole formula controversy

Daniel Kennefick[a]

University of Arkansas, Physics Department, Fayetteville, AR 72701, USA

Received 26 October 2016
Published online 20 December 2016

1 Introduction

The recent detection of gravitational waves by the advanced LIGO instruments (Abbott et al. 2016) has brought the phenomenon to public attention in a way never seen before. Given the long and enormous effort to detect this elusive form of radiation, it is interesting to look back at the reasons why scientists became sufficiently convinced of their reality to pursue this difficult experiment. While theoretical work based on Einstein's theory of General Relativity predicted the existence of the phenomenon, it is hardly surprising to learn that a previous observational result played a central role in convincing scientists and funders that the goal of detection was achievable. In this paper we look at the way in which this observational result, the measurement of orbital decay in the binary pulsar PSR 1913+16, interacted with an ongoing debate amongst theorists known as the quadrupole formula controversy (Kennefick 1999 and 2007). We shall see that the observational result at first sharpened and intensified the theoretical debate, before it became one of the reasons why the debate came to a close.

2 Controversy

The background to the story can be sketched relatively briefly (for a fuller account, see Kennefick 2007). The theory of gravitational waves dates to 1916 with Einstein's first paper on the subject, only half a year after his publication of the final form of his general relativity theory. In 1918 Einstein published a paper correcting a certain error from the paper of 1916, and presenting, for the first time, the quadrupole formula, expressing the rate of emission of gravitational wave energy by a system of accelerating masses. The formula gets its name because, as Einstein showed, the power radiated in gravitational waves by a system of accelerating masses is proportional to the square of the third time derivative of the system's mass quadrupole moment.

When Einstein derived the quadrupole formula it was on the basis of the linearized approximation of general relativity. This permitted him to make the calculation relatively straightforward, because in the coordinate system adopted by him

[a] e-mail: danielk@uark.edu

the linearized equations of gravity take on a form which is directly analogous to the Maxwell equations for electromagnetism, a theory in which the role of radiation was, and is, reasonably well understood. But, since general relativity is a non-linear theory, this linearized approximation can hold only for very weak fields, which specifically excludes systems, such as a binary star system, which are held together by their own gravitational interaction. Since it is only this type of system which (as far as we know today) might be capable of producing detectable gravitational waves, this approximation leaves something to be desired as far as sources go (keep in mind that we are still awaiting the first evidence of gravitational waves from anything other than a binary system). It is thought to be ideal for the study of gravitational wave detectors however. The question then is, does the quadrupole formula give a reasonable approximation of the source strength of possible astrophysical sources of gravitational waves, especially binary stars?

Aspects of this question were debated vigorously from the mid-1950s to the mid-1970s during the renaissance of General Relativity (a period defined by Cliff Will 1986). By the mid seventies most theorists accepted that binary star systems did generate gravitational waves, but whether the quadrupole formula could be correctly applied to them was still doubted by some experts. This quadrupole formula controversy, by that time, appeared to be showing signs of coming to a close, in that the remaining skeptics were obliged to object, from time to time, to comments made by other theorists which treated the problem as largely settled (Havas 1973).

What is interesting about the role of the binary pulsar in this story is that there are good grounds for believing that its primary role was to stimulate the controversy into new life. It is usually thought of as the agency by which the controversy was settled (and this is certainly a role which is of interest to this paper), but another possible reading is that it actually made the controversy more prominent and more contentious and that this served, with time, to bring it to a conclusion by focusing the attention of theorists upon it. One might speculate that we are dealing with a controversy downsizing principle, in analogy with the problem of cosmic downsizing in extragalactic astronomy, which revolves around the observation that over time quasars come to have smaller and smaller black holes. Since black holes should only ever grow in size, it is claimed that this observational effect arises because the big ones have already used up all their fuel and "turned off". The situation is thought to be similar to that which obtains for stars, where the larger stars, which paradoxically contain more fuel, burn the fuel at a far faster rate and live a much shorter life than do less massive stars.

In the case of scientific controversies we may similarly expect, at any given moment, to find many more small and almost moribund controversies than strident ones, because the former will be more long-lived. The fuel which is only slowly consumed in a small controversy is not the number of issues to be debated. I agree with those who think such points are all but inexhaustible. The fuel is the number of potential participants in the controversy. Where the number of participants is low, each of them may feel comfortable conceding a long period of debate to what is a manageable number of colleagues. As the number involved in the controversy rises, the ability to mediate the controversy by direct personal relations between all participants is strained. The consequences of remaining on the fence become less predictable as they become potentially more serious, since more people involved means potentially more influential people having a vested interest in the outcome. The participants come under pressure to take a definitive position and tend to do so more quickly. To continue with the analogy, the fuel is more quickly processed through the various stages, from open minded participant, to committed protagonist, to close-minded ideologue, at the end of which no further debate is possible. In essence, the controversy which burns most brightly extinguishes itself most quickly. To be sure, I am merely taking a long-established piece of folk wisdom and dressing it up in academic clothes. The phrase "slow-burning controversy", already nicely encapsulates the image I am trying to convey.

So let us examine briefly the course of the quadrupole formula controversy in the 1970s. We have already summarized the debate over whether binary stars could emit gravitational waves, a debate which flourished in the late fifties and early sixties. There then followed a period in which it was regarded as settled, by a large majority, that binary stars did undergo radiation damping as a result of gravitational wave emission. The detail of how this occurred was perhaps not regarded as a terribly pressing problem, given that no one was familiar with any known astronomical systems which, according to the quadrupole formula itself, would undergo a measurable decay in their orbits. The state of affairs bore a close approximation to the situation in controversies which have passed the point of crystallization, which is to say that even though there remained some who doubted the consensus opinion that the quadrupole formula was approximately correct, their views did not receive much public airing. In fact, however, it was still possible for their views to be aired, the problem was simply not important enough for major notice to be taken of anyone's views on the matter.

A thorough account of the views of the skeptics as to why confidence in the wide applicability of the quadrupole formula was misplaced is given in chapters 9 and 11 of Kennefick 2007 (for a more concise account see Kennefick 1999). Briefly, by the early 1970s quite a bit of work had been done by theorists to try to show that binary stars did radiate (to leading order) according to this formula. But sceptics objected that the calculations which had been done lacked mathematical rigor. To quote the abstract (in its entirety) of what might be called the manifesto of the sceptics, a 1976 paper by Jürgen Ehlers, Arnold Rosenblum, Joshua Goldberg and Peter Havas

> It is argued that a formula for the energy loss due to gravitational radiation of bound systems such as binaries has not yet been derived either exactly or by means of a consistent approximation method within general relativity, a view which contradicts some widely accepted claims in the literature. The main approaches used to obtain such a formula are critically reviewed, and it is pointed out that the derivations presented so far either contain inconsistencies or are incomplete.

Very few exact results have been published in the history of gravitational wave research, so the most relevant part of this abstract is that decrying the lack of consistency in the approximation schemes used. These schemes typically involved expansions of quantities in powers of v/c where v is a velocity (for instance of one of the stars in the binary) and c is the speed of light. As long as v/c is small, higher order terms in the expansion are ignored. But, objected the sceptics, no effort was made to check whether coefficients in these neglected higher order terms might not be large enough to offset the small size of v/c. In general there was no attempt at error control at all. Physicists objected that, for most stellar binaries, v/c was a truly tiny number, much smaller than in many other calculations of physical interest where they were accustomed to be just as cavalier in their approach. Furthermore mathematicians like consistency in their approximation schemes (as Ehlers et al. advertise above) and this means truncating all terms at the same order in v/c. But physicists are prone to occasionally truncating some quantities at higher orders than others, if they feel that these quantities are more physically relevant. In short, at least part of the argument is whether you feel that the calculation should be conducted according to the reliability standards of physics or mathematics. Of course Ehlers, Rosenblum, Havas and Goldberg were all physicists, but they came from a branch of physics, General Relativity, which was much closer to mathematics than other branches were. Indeed in some Universities the relativists were housed in the mathematics department. Ehlers and company might have preferred to say that they were upholding the standards of BOTH math and physics before they were confident enough to quote a result, but this irked some of their colleagues, who observed that none of the four were prepared to give a definitive answer to the problem themselves, but were willing to critique the efforts of others!

To give, briefly, two examples of approximations which alarmed the skeptics, one would be the use of point masses to represent the two stars in the binary system. Of course the calculation is greatly simplified by pretending that all of the mass in the system is concentrated at two points in space, but it is known that tidal effects (amongst other things) are missed when doing so. Most astrophysicists would argue that these tidal effects would be very small unless the two bodies were very close, but relativists responded that such confidence came largely from experience derived from Newtonian calculations not General Relativity itself. Another issue concerned the need to use different kinds of approximation to describe the motion of the two stars and to describe the behavior of the wave far from the star. Some way of matching between these two solutions was needed, in order that boundary conditions on the waves could be unambiguously and correctly applied to the motion of the stars. Indeed, as argued in Kennefick 2007, discrepancies between many early calculations of gravitational wave emission from binary stars are probably traceable to just this failure to impose proper boundary conditions. The fact that these were just two of many issues which were debated means that calculations which the sceptics would have applauded on one ground they would criticise on another. Thus what seemed like an impartial attempt to move the field forward to the sceptics, may have seemed like incessant and insatiable nit-picking to others.

One important bone of contention (which both the issues already mentioned fed into) can be described in the following way. Since it is the accelerated motion of the two stars in the binary which is expected to generate the gravitational waves, many physicists felt that it was appropriate, since the motion of binary star systems was well studied, to describe such motion and then calculate the waves which would be thereby produced. The sceptics pointed out, however, that such motion schemes were not demonstrated to be actual solutions of the Einstein equations (Ehlers et al. 1976). It was possible, for instance, to calculate the motion of some binaries by assuming the absence of gravitational radiation and then put that motion into the equations and find out what radiation would be produced by it. But since the motion would be modified by including the radiation, there was a logical inconsistency involved in this approach. Essentially the history of the theoretical endeavor was of a long slow process of painstakingly altering calculational schemes to address various objections, with different researchers insisting on different levels of rigor before they were satisfied that the answer was known to some reasonable level of approximation. From the point of view of the controversy, a key question is, when does the debate end? When everyone is satisfied? Or is it legitimate to cut off this discussion when some participants would rather it be continued? In analyzing the history of this debate, I introduced the concept of the Theoreticians' Regress (Kennefick 2007; modelled on earlier work on the Experimenters' Regress by Harry Collins). This describes the dilemma confronted by theorists whose calculations fail to agree. Because the calculations are complex, finding errors in them is an open-ended process which can lead to debates of arbitrary length. Since the most reliable method of testing a calculation is to see whether it yields the correct result, it is difficult to evaluate the claims of competing calculations when the result is itself at issue. In such a situation one must often fall back on one's expert assessment of the abilities of the theorists themselves. Since you may naturally be more apt to trust a theorist whom you know better, it follows that one's social network may have a bearing on one's view of which calculation is yielding the correct scientific answer.

A good example of the status of the debate on the eve of the discovery of the binary pulsar is the June, 1973 Paris meeting on gravitational waves at which Peter Havas (a European emigré to the United States who specialized in the radiation problem) gave a talk outlining his view that the question whether binary stars did emit gravitational waves at all was still unsettled, and advancing his critique of the main calculations which agreed with the quadrupole formula result (Havas 1973). In the conference proceedings, two of the remarks in response to Havas' talk can be regarded as sharing his skepticism, two as disagreeing with it, and two as neutral (at least phrased in a neutral way). This certainly suggests not only that Havas had leave to raise such issues with his peers, but also that he had an audience part of which, at least, was sympathetic. At the same time, the problem was not at the forefront of theoretical concerns at that moment. It was not considered irrelevant or uninteresting, after all the very fact of the conference being held at all suggests otherwise, but the fact that no astrophysical applications had been discovered certainly lessened its urgency.

Within little over a year the situation was transformed completely.

3 Discovery

Pulsars were discovered in 1967 by Jocelyn Bell and Tony Hewish using the Interplanetary Scintillation Array at the Mullard Radio Astronomy Observatory near Cambridge, England. It quickly became apparent that pulsars were a real-life instance of a long standing theoretical entity, the neutron star, which had been first proposed by Walter Baade and Fritz Zwicky decades previously, in 1933 (see Haensel et al. 2007, pp. 2–4 for a brief history). The problem of gravitationally collapsed objects became of greater theoretical interest following the discovery of quasars by radio astronomers in the fifties and was further stimulated by the pulsar discovery. By the early seventies only a few dozen pulsars were known, and Joe Taylor of the University of Massachusetts, together with his graduate student Russell Hulse, proposed to do a computerized search for them with the large Arecibo dish in Puerto Rico to provide a much larger ensemble of discovered objects. It was a specific aim of Taylor's proposal that such a large number of pulsars might feature one which was part of a binary system (Hulse 1997). This would permit the measurement of the mass of the pulsar, a topic of immense astrophysical interest, since the very idea of neutron stars had arisen following the work of Subramanian Chandrasekhar on the limiting mass of white dwarf stars. That a close binary neutron star system had been suggested as a possible source of detectable gravitational waves as early as 1963 by Freeman Dyson was almost certainly not on Taylor's mind as he began his pulsar search. This was all the more true since Dyson's suggestion had been made in the context of a proposal that arbitrarily advanced alien civilizations might construct such systems for the purpose of interstellar navigation.

In early July 1974 Hulse, down at Arecibo, recorded a pulsar, just barely strong enough to be detected by the system, unusually sensitive for its day as it was, whose position on the sky automatically baptized it with the name PSR 1913+16. After confirmation that this was indeed a pulsar, including measuring its period, Hulse recorded the word "fantastic" on his observing record, referring to the fact that the pulsar had the second shortest period known at that time. At this point he had no notion that it was in a binary system, only the rotational period of the neutron star itself had been measured, not its orbital period. The only foretaste of what was to come was that subsequent attempts to confirm that rapid pulse in these first observations did not agree, to Hulse's frustration. He even went so far as to cross out and erase these subsequent attempts from his log (Hulse 1997).

In late August Hulse returned to this object, in a routine way, to try to confirm its period. As before he found that its period kept changing with each measurement. Indeed, by a curious coincidence, he found that he almost repeated the same set of measurements each time the pulsar came overhead at Arecibo (the dish at Arecibo is so large it is built into a small valley, and thus cannot observe very far from the zenith of the sky). This would turn out to be due to the fact that the pulsar binary has an orbital period of just under 8 hours, and thus completes a little over 3 orbits with every rotation of the Earth. It did not take Hulse long to convince himself that he had discovered a pulsar in a binary system, and it was immediately clear to him and to his advisor Taylor that they were dealing with an extraordinary system. An eight hour orbital period represented an orbiting system involving massive objects with an unprecedently small physical separation from each other. Indeed word got around quickly about the new discovery, to the extent that the first theoretical paper commenting on the binary pulsar appeared in late 1974 (Damour and Ruffini 1974), while the discovery paper itself appeared only in 1975 (Hulse and Taylor 1975).

There can be little doubt that interest in the radiation problem from binary stars was reinvigorated by the binary pulsar discovery. Here was a real world example of a system where radiation damping might actually be measurable. Of course there were doubts expressed, on the theoretical side (Damour and Ruffini 1974) that the effect really would be measurable, but the experimenters were nevertheless not ruling it out. In an interview Joe Taylor recalls his own view at the time (interview conducted by the author by phone on 2nd May, 2008) . . .

> The person who put us onto that was Bob Wagoner. It happened that once the news was out and it became public that this thing was there and that we were observing it, I responded to a number of invitations to go and give talks about it and ended up making a grand tour around North America where I made five or six stops and one of them was at Stanford and Bob Wagoner there actually gave me his paper predicting the orbital period decay to carry back with me since he knew I was going to be at Harvard a couple of days later and I handed it to Alex Dalgarno the editor of ApJ Letters. So it was Bob's paper (Wagoner 1975) that I first began to take seriously and to recognize that with the current state of the art then, in October 1974 of doing pulsar timing, it was clear that, if his numbers were right, and I assumed they were, it would take us a number of years to see any effect, but not an unreasonable number and if we could improve the timing accuracy a little bit it might happen even sooner and that's more or less what happened.

While relativists were excited about a number of tests of general relativity which could be made for this system whose components were moving under the influence of unprecedently strong gravitational forces, it seems that the measurement of the binary pulsar orbital decay came significantly earlier than most people expected, as Taylor agrees (interview, 2nd May, 2008):

> I think that's right and that's largely because at that time it wasn't yet recognized that doing really high precision timing of pulsar signals was a very important goal.

Nevertheless the possibility was in the air from late 1974 onwards, and the fact that it would take a significant amount of time gave the theorists ample time in which to apply new techniques and increased effort to the problem of analyzing the orbital evolution of such a system as it responded to its own gravitational wave emission.

To what extent was this activity on the theoretical side visible to the experimenters? Given that their result, when available, was likely to have a decisive effect on the controversy, it is remarkable that they went totally unaware of it until they finally had a result to announce. This announcement was made, in its earliest version, at the ninth Texas Symposium on Relativistic Astrophysics in Munich in 1978 (Taylor and McCulloch 1980). The Texas series of meetings had a tradition of announcements of important observational results. The first Texas meeting had been held in response to the growing interest in quasars as new objects discovered by radio astronomers in the late fifties (Robinson et al. 1965). Taylor's talk in Munich is one of the more celebrated of the announcements made at this series of meetings (interview, 2nd May, 2008).

> Well, I'll tell you when I first even knew that there was any debate, was at the Texas Symposium in Munich.[1] And so somebody asked me a question, well let me back up just a little bit. I was scheduled to give a paper there on something like the second or third day of the conference, and Jürgen Ehlers, who was one of the conference organizers, recognized that somehow not getting to this until nearly the last day of the conference was not a good idea. So he asked me to get up and say just a few words about it in a session on the first day so that at least people would know what I looked like and we could talk in the halls, and so forth, afterwards. So I did that and I basically gave the result[2] and said I'll give all the details at the scheduled time the day after tomorrow, or something like that. Somebody then in the audience asked a question, I don't remember who it was, 'when you say that you have seen the period decay and it agrees with the prediction, what prediction are you using?' And I sort of was blind-sided by that. I just thought that everyone knew how to calculate this, except maybe me. And so I think I must have stood there wondering how to answer for a minute and Tommy Gold, who happened to be the session chairman, whispered in my ear, 'Landau and Lifshitz', so I said it's given in Landau and Lifshitz. So that more or less is what transpired. I mean, I remember having conversations later with people about it and I began to realize that, of course, that was just sort of an heuristic formula and the calculation wasn't even derived, I guess, in Landau and Lifshitz, it was given as an exercise for the student to do.

It is humorous to note that Gold, the session chairman, had been, with his collaborator Bondi, one of the early skeptics concerning whether binary stars could emit gravitational radiation. Although Gold would certainly have been very familiar with Landau and Lifshitz' treatment, he might also have been inclined to agree with Bondi's comment (to the present author, quoted in Kennefick 2007), that it was very "glib".

So once Taylor was apprised of the existence of the controversy, what was his reaction (interview, 2nd May, 2008)?

> So ok, so I was aware then that there was a controversy about it. Whenever I quizzed theorists, that I knew pretty well, about it, they tended to be people like Kip Thorne, for example. Kip always said, 'oh yes, you know, we're still worrying about the mathematical details, but we know it's right.' And my impression was that, I think pretty much I gained the impression that you convey to a large extent in your book as well[3], that the more mathematically oriented physicists, and particularly those who had been doing relativity in mathematics

[1] At this point on the interview recording, the author can hear himself say 'Really.'

[2] As quoted in Weisberg and Taylor 1981, the binary pulsar decayed at a rate of $(-2.5 \pm 0.3) \times 10^{-12}$, compared to a value predicted from the quadrupole formula of $(-2.38 \pm 0.02) \times 10^{-12}$.

[3] A reference to Kennefick 2007, illustrating one of the problems faced by an oral historian who wishes to write books and continue doing oral histories!

> departments, were still concerned about the lack of rigor and the full mathematical beauty, but the physicists like Thorne and Feynman and others just had little patience with that kind of concern and wanted to get on with it and see what you could do with it. And they more or less told me 'don't worry about it.'

So communication between theorists and experimenters contained this interesting feature, that a reasonably lively controversy amongst the theorists could be completely invisible to the experimenters. Obviously the controversy was not one which consumed the total energy of theorists in the field, but it still involved a good deal of back and forth and even a dedicated workshop, during the period in question, and yet no mention was made of its existence within Taylor's hearing. Partly, as Taylor says, this was because of the kind of theorists he was talking to. In the field of relativistic astrophysics, there were people close to the astrophysics end of the spectrum, and people closer to the relativity end, and Taylor, as an astrophysicist, was naturally more likely to talk to those on the astrophysics end. Since those theorists were less likely to be skeptical of the quadrupole formula, they naturally chose not to bring up any caveats about the derivations which they felt were unlikely ever to have a bearing on the observations underway. Furthermore, and this bears on a point I will try to bring out at the end of the paper, they may have felt some slight embarrassment that there existed theorists in their field who still doubted the canonical understanding of gravitational radiation in general relativity.

4 Trading zones and pidgins

In his book *Image and Logic* Peter Galison (1997), one of the pioneers of the careful micro-study of physicists in action, argues that different groups of scientists, in particular experimental and theoretical physicists often speak different technical languages and encounter difficulty in communicating with each other. He argues that, in such situations, physicists find it useful to develop a pidgin, a term used to describe a secondary language, formed usually from a mishmash of other languages, used to facilitate trade between different peoples. Galison describes the conceptual space between different groups of physicists as a trading zone and discusses the use of pidgins, which in his usage may refer to particular mathematical constructs designed to permit experimenters and theoreticians (let's say) to discuss and compare the predictions of the latter with the results of the former.

The binary pulsar is an interesting case to observe the possible need for trading zones, since it was a discovery by radio astronomers who had, otherwise, relatively little contact with relativists interested in gravitational waves. At the same time their field had arisen alongside the broader culture of relativistic astrophysics, which was formed by a first contact between radio astronomers and relativists after the discovery of quasars. To what extent do we observe the need for a trading zone between experimenters and theorists in our particular story? Certainly there seem to be areas of physics in which theorists and experimenters talk to each other regularly and apparently freely, and it is certainly also true that when physicists, even from very different subject areas, converse, they speak a recognizable technical language which seems to be quite unconscious of boundaries. Indeed, for the physicist, the international, inter subject quality of physics speech is one of the defining experiences of being a physicist (no doubt the same may be true for scholars in other disciplines). Nevertheless there is some evidence, in the case of the binary pulsar story, supporting the model put forward by Galison. One promising way to understand how scientists deal with trading zones, when and if they occur, is through the notion of *interactional expertise*, a concept which describes the ability of someone to talk intelligibly and usefully to an expert about their field, even if they are not (yet) capable of working in that field, which would be full expertise (Collins et al. 2007). It may be that, even where physicists lack direct expertise to work in a neighboring field, they at least possess interactional expertise to talk with their fellow physicists in that field.

Let us begin with the discovery of the binary pulsar in 1974. The two astronomers involved, Joseph Taylor and Russell Hulse, both received educations fairly typical of astronomers of their generation in that they were educated primarily in physics (in fact Hulse was still a graduate student when he discovered the binary pulsar). In this context, particularly as the two men were working in radio astronomy, astronomy is conceived of as being more or less a sub-discipline of physics, albeit an unusually ancient one which still maintained a certain level of institutional independence. As such they took courses in general relativity, a subject within physics which was typically considered an optional higher level course, but one which might be especially relevant to those planning to specialize in astronomy. As radio astronomers interested in pulsars, relativity theory was clearly relevant to an understanding of the source of the signals they planned to study, but not nearly as relevant and routine as the physics of the electromagnetically based detectors and instruments they operated.

Accordingly Joe Taylor describes one of his first actions on discovering that he had a binary pulsar with a uniquely close orbit involving unprecedently intense gravitational interaction between the two components (interview, 2nd May, 2008).

> We'd both taken the obligatory, or almost obligatory, relativity course in University, as part of our physics training, but neither one of us was very deeply into relativity. My wife was much amused when one day, this was when I was at the University of Massachusetts, of course, I said I don't have to teach today, I'm going to drive into Boston and visit the Tech Coop. And I spent the day in the MIT bookstore and came back with a pile of books, Weinberg, and Misner, Thorne and Wheeler and all the other ones that you would imagine. She was much amused that I spent the next few months deeply engrossed in these books.

So certainly the astronomers felt a need to get up to speed with the elements of relativistic orbital motion (the books referenced are Weinberg 1972 and Misner et al. 1973). To what extent was there a language gap between them and the practitioners of this discipline? Partly the gap was a social gap. Neither Taylor nor Hulse habituated amongst relativists and therefore did not partake in their discourse. So Taylor went unaware of the ongoing quadrupole formula controversy, throughout the time when, as we would be tempted to say today, he was determining the outcome of this controversy.

But leaving aside this question of discourse, when Taylor and his collaborators did speak to relativists, could they make themselves understood and be understood? Clearly they could, for the most part. But some obstacles were encountered. By the time Taylor and company were dealing with the orbital decay of the binary pulsar, Hulse had finished his doctorate and moved on. A collaborator with whom Taylor published many of the early papers announcing and discussing the orbital decay was Joel Weisberg. Weisberg does recall language difficulty playing some modest role in talking to theorists, before they found a long term collaborator in a talented young French relativist, Thibault Damour (interview conducted by the author, by phone, on 24th February, 2000).

> It's interesting, we had a failed attempt to work with one person. And I think the problem was he couldn't talk well enough to experimentalists. He couldn't give us results that were easily interpretable by us, whereas Thibault could. It was quite interesting.

Weisberg describes the kind of theorist that would be helpful in the process of theory testing using the binary pulsar data, saying "it had to be people who could talk a language I could understand." Regarding the one failed effort mentioned above, the problem had a very practical aspect, "he [the theorist] couldn't give us specific things to test." At the same time he emphasizes that their eventual collaborator, Damour was "brilliant" and "made fundamental progress", so "it wasn't just a language thing." He adds (in a private communication) that the "theorist 'speaking the right language' was not, by itself, enough for a successful collaboration."

Nevertheless, to examine the "language thing," I suspect it is fair to say that, in the absence of a relativity community, Taylor and Weisberg would have been capable of performing calculations to establish the predictions of certain theories (though their case was a particularly difficult one, given the strong fields associated with neutron stars, so whether they could have carried on the calculations while pursuing their observational program is certainly open to doubt). In fact, as we shall see, they did contribute original work on the theory side. The problem seems to me to be legitimately a question of language and society, in the sense that Taylor and Weisberg's problem was not primarily that they lacked the expertise to do the calculations. That much they could have acquired, and did acquire, with time and effort (but again, the kinds of calculations which would have satisfied skeptics like Ehlers would have been especially challenging). What they lacked was fluency in the language spoken by theorists, and social standing within the discourse of theory. The existence of theories to test is inextricably linked with the existence of theorists who developed them, who have a vested interest in the testing. Since the theorists are the experts, it is understandable that the astronomers, like Taylor and Weisberg, would feel distinctly hesitant about publicly putting forth calculations in an area that was not their own realm of expertise. We get a sense of this in their 1981 paper announcing the orbital decay result, where they cagily refuse to be drawn into the controversy over the validity of the quadrupole formula.

> We are also aware that some relativists hold Einstein's quadrupole formula, which underlies the calculation of energy loss rates in [this paper], to be invalid for gravitationally free-falling systems [such as binary stars]. Obviously the dispute about what the theory actually predicts must be resolved, but the present experimental situation does not by itself seem to demand any changes.

It is also worth noting that there was one other advantageous aspect to Taylor and Weisberg's eventual collaborator Damour, in addition to his ability to speak their language and his "brilliance." Damour's views on the quadrupole formula controversy were similar to those of skeptics like Ehlers, and therefore his calculations went to great lengths to address many of their stated concerns. To quote from his 1983 review paper

> In 1979 Taylor, Fowler and McCulloch reported the observations of a secular acceleration of the mean orbital longitude of the binary pulsar PSR 1913+16: i.e. in other words, a secular diminution of the time of return of the periastron ... While this effect had been qualitatively and quantitatively predicted on the basis of the above-mentioned heuristic argument, it had not been validly demonstrated to be a consequence of Einstein's theory; on one hand because the detailed calculations were not complete enough to control all the terms of the equations of motion and were plagued by mathematical inconsistencies, and, on the other hand, because the methods of calculation did not apply to a system, like the binary pulsar, containing "compact" objects ... with very strong self gravitational fields.

Since, as we have seen, experimenters have better things to do with their time than to be drawn into arcane theoretical debates, it is important that the calculations which are done by theorists are not black boxes whose inner workings are totally opaque to the experimenters. It is important that the results of these calculations can be couched in a form which deals with observables pertinent to the actual measurements being made. The need for what Galison would describe as a pidgin helped to produce the parametrized post-Newtonian (PPN) framework as a tool to mediate the theory testing process. The PPN framework is a way of expressing results from the Newtonian theory with correction terms based upon the post-Newtonian approximation of General Relativity, with the addition to each term of a parameter. Each parameter can be defined differently if one is using a different modern theory of gravity. Thus a theory can be expressed in terms of these ten parameters which will permit anyone engaged in a weak-field (solar-system) test of gravity to quickly determine which theory makes what prediction about that quantity.

The PPN formalism is important because the theory-testing process requires an alliance of theorists and experimenters. Theorists made predictions based on their calculations. Experimenters made measurements which were then compared to the results of the calculations. But some theories have very few published results which experimenters can test. The PPN framework demands only that the parameters have been worked out for a given theory for experimenters to be able to determine what the prediction for that theory would be in the case of the particular test they have in mind. This PPN framework had been widely used during solar system tests of general relativity, but was ill-adapted to the binary pulsar case because it presumed that the gravitational fields involved were very weak. Nevertheless a somewhat similar, but much less general (focusing as it did upon the case of gravitational radiation emission) parametrization was established which facilitated the theory testing aspect of Weisberg and Taylor's 1981 paper. To quote from Clifford Will's paper on the subject (1977)

> Because of the complexity of many alternative theories of gravitation beyond the post-Newtonian approximation, we have not attempted to devise a general formulation analogous to the PPN framework beyond writing equation (2) with arbitrary parameters. However, we can provide a general description of the method used to arrive at equation (2), emphasizing those features that are common to the theories being studied.

So given the existence of a pidgin to create a trading zone between astronomers (and others) interested in doing theory testing and gravitational theorists, why did the astronomers shrink from commenting directly on the quadrupole formula itself? One obvious answer is that the pidgin was not designed to facilitate such a conversation. It permitted comparisons between calculations derived from different theories. It was not designed for the more complex and open-ended task of critiquing subtle details of such calculations. Another answer is that the barriers were as much social as linguistic (the two must obviously be linked). The astronomers felt they lacked the social standing to weigh in on a question which obviously fell within the purview of the theorists. Because the controversy over which calculation within a given theory was the correct one depended on subtle judgments, it naturally required the expertise of the practicing theorists. This is precisely the meaning of the Theoreticians' Regress, that it depends on subtleties of expert judgment and not on some closed algorithmic model of how to perform a calculation.

5 Skeptics' dilemma

I have argued that the closing of debate in the quadrupole formula controversy occurred at least partly because of the quickening effect caused by the binary pulsar increasing the importance of the controversy. At the same time, the lifetime of the controversy, once the binary pulsar data became available, was greatly constrained by the existence of experimental data which bore directly on the topic at issue. For the theoretical controversy to continue indefinitely, there would have to have been a significant effort to contest either the experimental evidence or the interpretation of it. The fact that there was no such significant attack on the ruling interpretation of the binary pulsar data certainly limited the lifetime of the controversy, so it is interesting to look at the reaction of the skeptics to the work of Taylor and his collaborators.

In any problem of orbital mechanics there are many mechanisms which might account for all or part of an observed change in orbital period. That even the most famous agreements between theory and observation can be challenged in this way is shown by the saga of Robert Dicke's efforts to measure the oblateness of the Sun (the degree to which its shape departs from a perfect sphere). Dicke had pointed out that if the Solar oblateness turned out to be significantly different from zero, its gravitational influence on the orbit of Mercury would throw out the close agreement between the prediction of General Relativity and the observed perihelion advance of the planet Mercury (Dicke and Goldenberg 1967). As with the case of the Mercury Perihelion, the binary pulsar data seemed particularly impressive because it agreed with the prediction of the quadrupole formula with little or no need to take into account of other factors. The interpretation was that the system was very "clean." The corollary to this, naturally, is that any evidence that the system was not so clean would throw out the agreement. Given this opening to challenge the *interpretation* of the binary pulsar data, it is interesting that the gravitational wave skeptics were not involved in proposing alternative mechanisms.

Certainly there were those who considered it, amongst them Peter Havas and, very likely, his former student Arnold Rosenblum. They were to the fore in demanding that the observations not be accounted a successful test of general relativity given that (in their opinion) the quadrupole formula had not been shown to be a valid prediction of that theory. Joe Taylor recalls that certain people were particular about this question of terminology (interview, 2nd May, 2008).

> Well, let me think, the people who kept bugging me about it, so to speak, were Peter Havas, Fred Cooperstock and Arnold Rosenblum. Arnold bugged me about it a lot. Anyway, they just kept saying 'Look, even though you have an experimental number now, we're not even sure what the theoretical number is and you can't go around saying that you've confirmed something.' So I tried to remain outside of the argument, letting the theorists fight it out until they all ... persuaded one another. So that seemed to be the best thing for me to do and we were simply concerned with getting an experimental result that we were happy with.

The alternative scenarios to the gravitational wave interpretation were actually put forward in print, but generally not by the skeptics. This may have been because the skeptics found themselves in a similar position to the experimenters. They had a vested interest in the debate, but lacked the special expertise which would have permitted them to comment. Likely dissipative mechanisms (or even non-dissipative ones) fell within the purview of astrophysics rather than relativity, and were explored and commented upon by astrophysicists rather than relativists.

The most important issues which had to be dealt with in demonstrating that the observed decay agreed with the quadrupole formula prediction was the nature of the unseen companion in the system, and the relative acceleration of the binary pulsar to our solar system. If the unseen companion was a sufficiently compact object, like another neutron star (which is now firmly believed to be the case) then it would undergo little deformation as a result of the visible pulsar's tidal effect. But if it was a normal star, it would develop a marked oblateness which would in turn create a perturbation in the orbit of the pulsar (a tidal friction-like effect) which would be difficult, except over longer timescales, to distinguish from the orbital decay due to radiation damping. Effects of this type would, however, have affected other measurements made in the system, and with time the experimenters became convinced that the system was extraordinarily clean. As Taylor and McCulloch (1980) stated in their paper from the Texas Symposium

> If one were given the task of designing an ideal machine for testing gravitation theories, the result might be a system rather similar to PSR1913+16; an accurate clock of large mass and small size, moving at high speed in an eccentric orbit around a similar object located in otherwise empty space. To be sure, one would place the system somewhat closer to the Earth than ∼5 kpc, or would arrange for a more powerful transmitter to convey the clock pulses to terrestrial telescopes; but we cannot expect Nature to be concerned with the inadequacies of our instrumentation!

This sense of wonder at the sheer serendipity of coming across such a system (many relativity theorists had sworn for decades that no system would ever be found in which gravitational wave effects would be measurable) was brought into focus for me after the more recent discovery of the "double pulsar" a system with an even closer orbit than the original binary pulsar, in which both pulsars are visible from Earth. I have heard this system referred to as "a relativistic astrophysicist's wet dream."

Taylor and McCulloch's comment illustrates the three main technical challenges in creating a match between theory and experiment for this system. First, the system must be in empty space. The presence of interstellar gas, for instance, would certainly alter the orbit of the system with time, as a result of dynamical friction. A related issue would be if the pulsars themselves were blowing off material at a significant rate, in which case the mass loss would affect the orbital motion. Secondly, as we have seen, both objects must be compact objects, such as neutron stars, so that perturbations due to the failure of the bodies to behave as point sources can be ignored. As a corollary to this, if the system contained a third massive object, this would obviously also affect the orbit of the two known components. Finally, the object should be close to us, not only for reasons of detection, but because a more distant object is in a more different orbit around the center of the galaxy and would be accelerating more strongly with respect to us here on Earth (for a list of references and discussion of a number of these issues, see Damour and Taylor 1991).

It is a well known result of special relativity that systems which are in inertial motion with respect to each other have clocks which run at different rates. If the systems are accelerating with respect to each other, then their respective clocks will alter, with time, in their relative rates of running. Since the solar system and the binary pulsar system are in different orbits around the galactic center they are not in the same inertial frame with each other. Accordingly the sensitive timing which is required to measure the orbital damping effect is also capable of measuring the relative accelerations of these two systems. In so far as doubt persisted about the validity of the quadrupole formula, this was a bad thing. Indeed, at one point during the 1980s, it did happen that the analysis of measurements of the binary pulsar did fall out of agreement with the quadrupole formula, by a much smaller amount than had been at issue in the earlier theoretical debate (in so far as that debate had ever been completely quantified). A close analysis of the relativistic theory of timing between the two systems, carried out by Taylor in collaboration with Thibault Damour, showed that the discrepancy could be explained on the basis of fully accounting for the timing issues (Damour and Taylor 1991).

Ultimately, as Taylor recalls, the situation reached the point where, if one *assumed* the validity of the quadrupole formula, one could make an accurate determination of the position of the binary pulsar in the galaxy, based on its relative acceleration. This measurement was more accurate than was possible by other methods at that time. This makes as good a moment as any to mark the end of the quadrupole formula controversy. When a prediction turns from a thing to be tested to a tool to be used, the debate is surely closed (and this, of course, goes some way to explain the impatience of non-skeptics to achieve that moment of closure). It is a mark of the importance of the controversy that the measurement of the distance to the galactic center which could have been provided by the binary pulsar data never became a canonical one, though it is in agreement with subsequent measurements using other techniques.

As Damour and Taylor put it in 1991

> If we assume that the standard general relativistic framework ... is valid we see that, in a few years, the measurement of $\dot{P}_b^{\mathrm{obs}}$ (the rate of decay of the binary pulsar's orbit) can be turned into a measurement of ... the galactic constants R_o (the distance form the Solar System to the Galactic center) and v_o (the speed of galactic rotation at about the center at the position of the solar system) (especially v_o, which presently contributes the biggest uncertainty). Such a "pulsar timing" measurement of v_o would be free from many of the astrophysical uncertainties that have plagued other determinations.

Since the Taylor-Hulse discovery, subsequent binary pulsars have been found where the relative acceleration of the two systems does not permit a particularly accurate determination of the rate of orbital damping. Had the controversy persisted so far this might have provided some opening for skeptics. However the discovery of the double pulsar in 2003, a system in which both pulsars are oriented so that both their radio beams are visible from the Earth, has provided a system with even stronger orbital damping than the original binary pulsars, whose results are in agreement with it.

How much interpretive flexibility was there for skeptics to continue the controversy? This has been a bone of contention in the field of science studies, where gravitational wave physics has been the subject of long term sociological analysis by Harry Collins (Collins 1994, 2004), some of whose conclusions have been challenged by the philosopher Allan Franklin (Franklin 1994). Did the skeptics largely abandon the fight because, as Franklin would have it, they were rational actors or, as Collins would have it, they had run out of sociological space in which to continue the argument? I suspect both considerations played a role. A rational actor will certainly take sociological considerations into account when determining whether to continue a debate. Most physicists do not wish to face social ostracism, even in a cause they believe to be right. At the same time any social constructivist will agree that the ruling out of certain arguments as work in the field progresses, the limitations placed on interpretative flexibility in the ebb and flow of debate, can tax the ingenuity of even the most stubborn skeptics to the point at which they give up the struggle. The social struggle can become unequal in a double sense, in that sceptics are both outnumbered and outmaneuvered by their opponents. Whether the maneuvering was all in vain, given the inevitable verdict of nature is, of course, an interesting question, but not one that is trivial to answer by the historian's method.

That skeptics *considered* continuing the battle is clear enough. Although Fred Cooperstock did retire from the fray for a decade or so after the mid-eighties, he subsequently put forward a new argument that gravitational waves would not propagate energy through empty space. The failure, for a numbers of years, of the new generation of gravitational wave detectors like LIGO, to detect gravitational waves passing by the Earth, provided some opening for skeptics like Cooperstock. He and others put forward arguments that the existing theory is correct for *sources* like the binary pulsar, but fails for *detectors* like LIGO, thus explaining why evidence existed for gravitational waves binary neutron star systems, even though as we have not, as yet, detected them[4]. The specifics of these new skeptical arguments vary widely[5]. It is important to distinguish between the scepticism of professional physicists like Cooperstock, and the irreconcilable objections of amateurs who focus on the sheer expense of the detectors which, they claim, can never succeed in detecting anything. These amateurs are prone to claim that the recent detections must be fraudulent in some way. Their arguments are not engaged in any way by the professionals in the field.

Peter Havas, when I interviewed him in 1995, certainly spoke of the openings he believed had existed, at least for a time, for an attack on the standard interpretation of the pulsar timing results. He still entertained significant doubts about the consensus which had emerged at that time. Joe Taylor reports that Havas, and his student Arnold Rosenblum, did ask to see some of the data and that he sent them a magnetic tape containing some (private communication). When he asked them a year later whether they had made progress they indicated that they had been distracted by other problems. Nevertheless, a search for Arnold Rosenblum's papers on the SAO/NASA Astrophysics Data System server shows that, from the mid-eighties, after several years spent on his calculations of gravitational wave emission that did not agree with the quadrupole formula, he then devoted a number of papers to the problems of relativistic timing in orbital and binary systems. Although none of this series of papers referred directly to the binary pulsar, they are strongly suggestive that he had spent a considerable amount of time thinking about this issue, leading him into that field[6].

Therefore we can say that the skeptics considered a foray against the conventional interpretation of the binary pulsar data, but decided against it. One can say that the physics of the situation obliged them to react this way, in that they felt they could not overturn the hard empirical evidence provided by the binary pulsar data. But one can also say there were sociological reasons. They were not in a position to do their own experiment to challenge the data, because they lacked the standing in that field which would have permitted them to enter it with any hope of success. For starters they would never have been granted time on a radio telescope to do their own measurements of this system (one group of astronomers did do some independent timing measurements of the binary pulsar, guided by data supplied by Taylor, and concluded that Taylor and his collaborators were correct in their results on the orbital decay, see Boriakoff et al. 1982). Even worse, in so far as the interpretation of the data could be challenged by theorists, it was by astrophysicists with experience in the study of stellar binaries and pulsars, not by relativists experienced in gravitational waves. Thus from a professional point of view the skeptics were in a double bind which, combined with their increasing isolation within their own community, as the debate moved towards a final resolution, prevented any kind of continuation of the public debate. Whatever private doubts were held by a few theorists about the reliability of the existing calculations, the empirical result was regarded as beyond dispute. The final option open to the skeptics, arguing that Taylor had simply got it wrong, was undoubtedly not entertained because of the outstanding reputation which Taylor enjoyed within the astrophysics community for his careful and painstaking work.

In the case of the binary pulsar replication demanded access to radio telescope time to look at the same system or, better, the discovery of an independent system.

[4] We cannot hope, with current technology, to detect the gravitational waves emitted by the known binary pulsar systems. It is only when such systems reach their terminal point and spiral into each other and merge that Earth-based detectors can hope to observe them.

[5] A sample of modern gravitational wave skepticism is given by the following references: Cooperstock 1992, Bel 1996 and Aldrovandi et al. 2008.

[6] Arnold Rosenblum died tragically young in 1991 (Cohen et al. 1991).

But, as we have seen, subsequent systems were often not as ideal for this experiment as the original. Not until the discovery of the double pulsar can we be said to have a fully comparable replication of the original, so one can certainly speculate that there may have been some scope for further controversy in the decades between 1980 and the early years of the twentieth century, had there been sufficient sociological space to support such a debate. But while logical space for disputation may have remained, the skeptics had run out of sociological space. Indeed, there is every reason to believe that the field of gravitational wave physics could ill afford to permit such a controversy to linger for that amount of time, lest it put its own disciplinary standing at risk.

6 Conclusions

We have seen how the measurement of orbital decay in the binary pulsar helped to convince physicists and others that gravitational waves were real and possibly detectable. The fact that the results vindicated long-standing predictions of the theory was also vital in establishing confidence in the theory underpinning planned detectors such as LIGO. Even though social boundaries and possible differences of language prevented the observers, Taylor and his collaborators, from directly declaring that the skeptics were wrong, their experimental work did have the effect of successively reducing the social space in which they might have continued their objections to the use of the quadrupole formula. However, as I show elsewhere (Kennefick 2007), this closing down of the contested terrain was accomplished not just by the increasingly accurate experimental work but by increasingly sophisticated theoretical work. Additionally close collaboration between the observers (Taylor and Weisberg) and one of the theorists (Damour) played a role in bringing the controversy to a close. As we have seen, a good boundary point marking the shift from controversy to post-controversy was the moment when one might make important discoveries (the distance from Earth to the center of the galaxy) by assuming that the quadrupole formula was true. When a contested result becomes, in its turn, a tool for research, then any lingering skepticism receives short shrift from most workers in the field.

Acknowledgements. I would like to thank Joseph Taylor, Clifford Will, Thibault Damour, Joel Weisberg and the late Peter Havas all of whom permitted me to interview them for the research which gave rise to this paper. All of the interviews, except the one with Peter Havas, were recorded. Both Harry Collins and Allan Franklin discussed some of the issues bearing on this paper with me many times, and aspects of it are based on an unpublished draft of a paper written by Collins and I. I would like to thank both of them for their help and inspiration on this work. Diana Buchwald and Kip Thorne both helped me far more than I can recall in the early stages of this work, and I would also like to thank David Rowe for giving me the chance to finally turn it into a paper.

References

Abbott, B.P. et al. 2016. Observation of Gravitational Waves from a Binary Black Hole Merger. *Physical Review Letters* **116**: 061102.

Aldrovandi, R., J.G. Pereira, R. da Rocha and K.H. Vu. 2008. Nonlinear Gravitational Waves: Their Form and Effects. arXiv:0809.2911v1.

Baade, W. and F. Zwicky. 1934. Remarks on Super-Novae and Cosmic Rays. *Physical Review* **46**: 76-77.

Bel, L. 1996. Static Elastic Deformations in General Relativity. Available at arXiv:gr-qc/9609045.

Boriakoff, Valentin, D.C. Ferguson, M.P. Haugan, Y. Terzian and S.A. Teukolsky. 1982. Timing Observations of the Binary Pulsar PSR 1913+16. *The Astrophysical Journal* **261**: L97-L101.

Cohen, J.M., Peter Havas, and V. Gordon Lind. 1991. Arnold Rosenblum. *Physics Today* 45: 81. Another obituary of Rosenblum appeared in the *New York Times* of January 7, 1991.

Collins, H.M. 1994. A Strong Confirmation of the Experimenters' Regress. *Studies in History and Philosophy of Science Part A* **25**: 493-503.

Collins, H.M. 2004. *Gravity's Shadow*. Univ. of Chicago Press, Chicago.

Collins, H.M. 2009. We cannot live by scepticism alone. *Nature* **458**: 30-31.

Collins, H.M., R. Evans and M. Gorman. 2007. Trading Zones and Interactional Expertise. *Studies in the History and Philosophy of Science A* **38**: 657-666.

Cooperstock, F.I. 1992. Energy Localization in General Relativity: A New Hypothesis. *Foundations of Physics* **22**: 1011-1024.

Damour, T. 1983. Gravitational radiation and the motion of compact bodies. In *Gravitational Radiation*, edited by N. Deruelle and T. Piran. North-Holland, Amsterdam, pp. 59-144.

Damour, T. and R. Ruffini. 1974. Sur certaines vérifications nouvelles de la Relativité générale rendues possibles par la découverte d'un pulsar membre d'un système binaire. *Comptes Rendu de l'Académie des Sciences de Paris, séries A* **279**: 971-973.

Damour, T. and J.H. Taylor. 1991. On the Orbital Period Change of the Binary Pulsar PSR 1913+16. *The Astrophysical Journal* **366**: 501-511.

De Witt, C.M. 1957. *Conference on the Role of Gravitation in Physics*, proceedings of conference at Chapel Hill, North Carolina, January 18–23, 1957. (Wright Air Development Center (WADC) technical report 57–216, United States Air Force, Wright-Patterson Air Force Base, Ohio). A supplement with an expanded synopsis of Feynman's remarks was also distributed to participants (a copy can be found, for example, in the Feynman papers at Caltech).

Dicke, R.H. and H.M. Goldenberg. 1967. Solar Oblateness and General Relativity. *Physical Review Letters* **18**: 313-316.

Dyson, F. 1963. Gravitational Machines. In *Interstellar Communications*, edited by A.G.W. Cameron. Benjamin Press, New York, pp. 115-120.

Ehlers, J., A. Rosenblum, J.N. Goldberg and P. Havas. 1976. Comments on Gravitational Radiation Damping and Energy Loss in Binary Systems. *The Astrophysical Journal* **208**: L77-L81.

Einstein, A. 1916. Näherungsweise Integration der Feldgleichungen der Gravitation. *Königlich Preussische Akademie der Wissenschaften Berlin, Sitzungsberichte*: 688-696.

Einstein, A. 1918. Über Gravitationswellen. *Königlich Preussische Akademie der Wissenschaften Berlin, Sitzungsberichte*: 154-167.

Einstein, A. and N. Rosen. 1937. On Gravitational Waves. *Journal of the Franklin Institute* **223**: 43-54.

Feynman, R.P. and R. Leighton. 1988. *What do you care what other people think? Further adventures of a curious character*. Norton, New York. Remark quoted appears on p. 91 of the Bantam paperback edition (New York, 1989).

Franklin, A. 1994. How to Avoid the Experimenters' Regress. *Studies in History and Philosophy of Science Part A* **25**: 463-491.

Galison, P. 1997. *Image and Logic: A Material Culture of Microphysics*. Univ. of Chicago Press, Chicago.

Haensel, P., A.Y. Potekhin, and D.G. Yakovlev. 2007. *Neutron Stars 1: Equation of State and Structure*. Springer, New York.

Havas, P. 1973. Equations of Motion, Radiation Reaction, and Gravitational Radiation. In *Ondes et Radiation Gravitationelles* proceedings of meeting, Paris, June, 1973. Editions du Centre National de la recherche scientifique, Paris, pp. 383-392.

Hulse, R. 1997. The Discovery of the Binary Pulsar. In *Nobel Lectures in Physics 1991–1995*, edited by G. Ekspong. World Scientific, Singapore.

Hulse, R.A. and J.H. Taylor. 1975. Discovery of a Pulsar in a Binary System. *Astrophysical Journal* **195**: L51-L53.

Kaiser, D. 2009. Birth Cry of Image and Logic. *Centaurus* **50**: 166-167.

Kennefick, D. 1999. Controversies in the History of the radiation reaction problem in General Relativity. In *The Expanding Worlds of General Relativity, Einstein Studies, volume 7*, edited by H. Goenner, J. Renn, J. Ritter and T. Sauer, Birkhauser Verlag, Boston, pp. 207-234. Also available at arXiv:gr-qc/9704002.

Kennefick, D. 2007. *Traveling at the Speed of Thought: Einstein and the Quest for Gravitational Waves*. Princeton University Press, Princeton, NJ.

Misner, C., K.S. Thorne and J.A. Wheeler. 1973. *Gravitation*. Freeman, San Francisco.

Robinson, I., A. Schild and E.L. Schucking. 1965. *Quasi-stellar sources and gravitational collapse, including the proceedings of the First Texas Symposium on Relativistic Astrophysics*. Univ. of Chicago Press, Chicago.

Rosen, N. 1940. General Relativity and Flat Space I. *Physical Review* **57**: 147-150.

Royal Swedish Academy of Sciences. 1993. Press Release announcing the Nobel prize winners in Physics for 1993, issued 13 October, 1993 and retrieved on the web at http://nobelprize.org/nobel_prizes/physics/laureates/1993/press.html on Apr 21, 1993.

Taylor, J.H. and P.M. McCulloch. 1980. Evidence for the Existence of Gravitational Radiation from Measurements of the Binary Pulsar 1913+16. In *Proceedings of the Ninth Texas Symposium on Relativistic Astrophysics* edited by Jürgen Ehlers, Judith Perry and Martin Walker. New York Academy of Sciences, New York, pp. 442-446.

Wagoner, R.V. 1975. Test for the Existence of Gravitational Radiation. *Astrophysical Journal* **196**: L63-L65.

Weinberg, S. 1972. *Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity*. Wiley, New York.

Weisberg, J.M. and J.H. Taylor. 1981. Gravitational Radiation from an Orbiting Pulsar. *General Relativity and Gravitation* **13**: 1-6.

Will, C.M. 1977. Gravitational Radiation from Binary Systems in Alternative Metric Theories of Gravity: Dipole Radiation and the Binary Pulsar. *The Astrophysics Journal* **214**: 826–839.

Will, C.M. 1986. *Was Einstein Right? Putting General Relativity to the Test*. Basic Books, New York.

Will, C.M. and D.M. Eardley. 1977. Dipole Gravitational Radiation in Rosen's theory of gravity: Observable effects in the binary system PSR 1913+16. *The Astrophysical Journal* **212**: L91-L94.

Open access funding provided by the Max Planck Institute for the History of Science.

Eur. Phys. J. H **42**, 311–393 (2017)
DOI: 10.1140/epjh/e2017-80014-4

THE EUROPEAN PHYSICAL JOURNAL H

Stellar structure and compact objects before 1940: Towards relativistic astrophysics

Luisa Bonolis[a]

Max Planck Institut für Wissenschaftsgeschichte, Boltzmannstraße 22, 14195 Berlin, Germany

Received 8 March 2017 / Accepted 8 March 2017
Published online 28 April 2017

Abstract. Since the mid-1920s, different strands of research used stars as "physics laboratories" for investigating the nature of matter under extreme densities and pressures, impossible to realize on Earth. To trace this process this paper is following the evolution of the concept of a dense core in stars, which was important both for an understanding of stellar evolution and as a testing ground for the fast-evolving field of nuclear physics. In spite of the divide between physicists and astrophysicists, some key actors working in the cross-fertilized soil of overlapping but different scientific cultures formulated models and tentative theories that gradually evolved into more realistic and structured astrophysical objects. These investigations culminated in the first contact with general relativity in 1939, when J. Robert Oppenheimer and his students George Volkoff and Hartland Snyder systematically applied the theory to the dense core of a collapsing neutron star. This pioneering application of Einstein's theory to an astrophysical compact object can be regarded as a milestone in the path eventually leading to the emergence of relativistic astrophysics in the early 1960s.

1 Introduction

Despite its enormous influence on scientific thought in its early years, general relativity experienced a so-called 'low-watermark period', going roughly from the mid-1920s to the mid-1950s (Eisenstaedt 1986, 1987a, 2006), during which it remained cut off from the mainstream of physics and was perceived as a sterile, highly formalistic subject. Accompanied by a series of major astrophysical discoveries, the status of General Relativity definitely changed in the 1960s, when it became an extremely vital research stream of theoretical physics. Quasars, the cosmic microwave background radiation, and pulsars – soon identified as rotating neutron stars – led to the recognition that physical processes and astrophysical objects exist in the universe that are understandable only in terms of the general theory of relativity. In providing definitive proof of the existence of neutron stars, the discovery of pulsars and binary X-ray sources, made

[a] e-mail: lbonolis@mpiwg-berlin.mpg.de

even plausible the possibility of black holes, entities that had previously existed only in the minds of a few theorists. In raising new challenges to the emerging relativity community, these had of course an important role in strengthening the process which turned general relativity into a "subdiscipline of physics" (Blum et al. 2015, 2016).

However, the view of a community of relativists magically awakened from its slumber by the new astrophysical discoveries is too one-dimensional. As Alexander Blum, Roberto Lalli, and Jürgen Renn have outlined in their historiographical framework exploring the main factors underlying the return of general relativity into the mainstream of physics, a complex series of elements underlying such process must be taken into account: intellectual developments, epistemological problems, technological advances, the characteristics of post-World War II and Cold-War science, as well as the newly emerging institutional settings. Starting from the mid 1950s, further implications began to be explored and general relativity gradually came into focus as a physical theory. This framework, in which they propose to speak of a *reinvention of general relativity*, rather than a *renewal*, is leading to an understanding of the reinvention as a result of two main factors: the recognition of the untapped potential of general relativity and an explicit effort at community-building. These two factors allowed this formerly dispersed field to benefit from the postwar changes in the science landscape.

The dynamics underlined in (Blum et al. 2015) is actually independent from – and prior to – the major astrophysical discoveries of the 1960s. Up to that time, the view prevailed that general relativistic effects were significant only for cosmology. However, the violent events that seemed to occur in the core of radio galaxies involving enormous energies corresponding to a rest-mass energy of 10^6 solar masses ($M_\odot$) (Burbidge 1959), the growing field of nuclear astrophysics (Burbidge 1962; Burbidge et al. 1957; Cameron 1958), and the eventful discovery of quasars, had prepared the stage for the emerging awareness at the beginning of the 1960s of physical processes in which general relativistic effects are dominant and that could release much larger fractions of the rest mass as energy than the small fraction provided by the binding energies of nuclei. Such processes that did seem possible in the framework of general relativity suggested the actual existence of astrophysical objects in the universe satisfying requirements that appeared to be beyond the scope of nuclear physics.

The problem of finding the source of the tremendous energy stored in cosmic rays and magnetic fields of some powerful radio galaxies, led to a theory put forward by William Fowler and Fred Hoyle in January 1963. They suggested that exceedingly massive star-like objects probably could exist with masses up to 10^8 times that of the sun at the center of those galaxies. The gravitational collapse of such supermassive stars could be the driving force behind the great amount of energy emitted by those strong radio sources (Fowler and Hoyle 1963a). Their opinion was that in the process of contraction of a mass of $10^7 - 10^8 M_\odot$ "general relativity must be used" in order to obtain the energies of the strongest "stellar-type" sources (Fowler and Hoyle 1963b, p. 535).

A few months after this proposal, new objects were discovered, having apparently masses of this order of magnitude, dimensions of about a light week, and having a luminosity two orders of magnitude larger than the luminosity of a large galaxy having dimensions a million times larger and containing something like 10^{11} stars. In particular, the crucial identification of the high redshift of the already known radio source 3C273 (Hazard et al. 1963; Oke 1963; Schmidt 1963) and of the source 3C48 (Greenstein and Matthews 1963), made now even more pressing the problem to explain the mechanism whereby these and other sources that were masquerading as a star and were thus identified as "quasi-stellar" objects, managed to radiate away the energy equivalent of five hundred thousand suns at a very fast rate.

The "supermassive stars" suggested by Fowler and Hoyle immediately became an attractive explanation for these new peculiar astrophysical objects, that appeared to be farther away than most known galaxies but were luminous enough to be observed by optical telescopes. Their enormous luminosity could also sharply change in the course of one week, as analysis of historical plate material of Harvard Observatory showed (Smith and Hoffleit 1963). As such enormous energies must be emitted by regions less than one light-week across, collapsed objects became candidates for the engine of quasi-stellar radio sources.

The intriguing discovery of quasi-stellar radio sources – soon renamed *quasars* (Chiu 1964, p. 21) – with their large red-shifts and corresponding unprecedented-large radio and optical luminosities, opened up the discussion on a series of exciting questions. Among the problems raised were the following: Were these objects the debris of a gravitational implosion? By what machinery could gravitational energy be converted into radio waves? Would gravitational collapse lead to indefinite contraction and a singularity in space time? If so, how should theoretical assumptions be changed to avoid this catastrophe? (Robinson et al. 1965, Preface).

"The topic was just right for reporting and sorting out observations as well as for theoretical analysis" (Schucking 1989, p. 51): during the summer 1963, three relativists in Dallas, Ivor Robinson, Alfred Schild, Engelbert Schucking, realized that a conference bridging the gap between the still exotic world of general relativity and the realm of astrophysics, might be well timed, and it would be a perfect occasion to make known the recently created Southwest Center for Advanced Studies. They immediately involved Peter Bergmann, an influential relativist who had been associated with Einstein at the Institute for Advanced Study in Princeton since 1936, and sent out letters of invitation. Three hundred relativists, optical and radio astronomers, and theoretical astrophysicists attended the *International Symposium on Quasi-Stellar Sources and Gravitational Collapse* (Robinson et al. 1965), the first of the long series of Texas Symposia, which set up the stage merging two seemingly distant fields: general relativity and astrophysics, so distant that the organizers had to invent a new label for this brand new field: "The suspicion existed that quasars might have something to do with relativity and thus might fit into an imaginary discipline combining astronomy with relativity. One of us – Alfred, Ivor or I? – invented a catch phrase for this new field of science: *relativistic astrophysics* [emphasis added]" (Schucking 1989, p. 50).

Robert Oppenheimer was asked to chair the first session, a most natural choice, because of his involvement in the first systematic application of Einstein's general theory of relativity to a compact astrophysical object. Oppenheimer's three papers published between 1938 and 1939, each with a different collaborator (Oppenheimer and Serber 1938; Oppenheimer and Volkoff 1939; Oppenheimer and Snyder 1939), are regarded as a milestone both in his scientific production and in the path eventually leading to the emergence of relativistic astrophysics in the early 1960s. In speaking of observations of "incredible grandeur" (Schucking 1989, p. 50), Oppenheimer officially opened the discussion on topics such as neutron stars or the possibility of gravitational collapse to a singularity in space-time, topics investigated within the context of the considerable revival of interest in the properties of matter at high densities and compact stars going on since the end of the 1950s (Ambartsumyan and Saakyan 1960, 1962a,b; Beckedorff 1962; Cameron 1959; Hamada and Salpeter 1961; Harrison et al. 1958; Migdal 1959; Salpeter 1960, 1961; Tauber and Weinberg 1961; Zeldovich 1962a,b). During the conference Bergmann remarked that in the past "general relativistic effects had been observed only in the weak field limit. Now new developments of astrophysics have made relativity a more physical theory" (Chiu 1964, p. 34). Following the 1963 Texas Symposium (Harrison et al. 1965; Robinson et al. 1965), many important advances in understanding black holes developed from new astrophysical observations and theoretical developments.

Cosmology, with its strong connections with general relativity since its early days, provided a continuity through the low-water-mark period, up to the post-war years, even if it was generally considered an "esoteric" field without any real connection with physics and having a scant observational basis. In cosmology, general relativity directed the course of the observational researches in the realm of the galaxies, once the paradigm of an expanding universe became firmly established. However, while in the past it was the geometry, kinematics, and dynamics of the universe which were in the foreground, in the post-war development of cosmology, physical processes in the universe, involving elementary particles, electromagnetic radiation, and nuclear reactions, became a dominant interest, establishing a new and wider interaction with other fields.

From a different perspective, studies on dense matter and compact astrophysical objects, merging interdisciplinary fields like nuclear physics and astrophysics – both having many intersecting topics especially with post-war cosmology – provided since the 1920s another form of continuity, through the 1930s and the 1940s. During the post- and Cold War period, implosion and explosion problems, related to the design of thermonuclear weapons, brought about renewed interest in investigations on highly dense stellar matter and on the abandoned problem of gravitational collapse within Einstein's theory. New tools, typical of post-war science, were now available: the impressive advances in nuclear science combined with the first powerful computers, designed to perform the complex calculations for thermonuclear weapons, were now used to calculate the equation of state of condensed stellar matter up to the endpoint of thermonuclear evolution. While a new community of researchers in general relativity was achieving novel fundamental theoretical insights into Einstein's equations, the interaction between different scientific communities tackling interconnected astrophysical problems led to a resurgent awareness of processes in which general relativistic effects might play a dominant role.

A reconstruction of how the emergence of relativistic astrophysics in the early 1960s can be understood as the culmination of a complex process including the longstanding tradition of the astrophysical study of compact objects and its connections with general relativity, will be the subject of a forthcoming article.

The present contribution is examining how fundamental premises were laid during the period going from the mid 1920s to the the end of the 1930s by theoretical investigations on such a basic topic, officially inaugurated in 1926 by Ralph Fowler's pioneering paper examining the problem of degenerate dense matter in white dwarf stars (Fowler 1926a). These studies were accomplished at the intersection of different theoretical frameworks involving several disciplines and sub-disciplines and developed into a knowledge network involving some leading actors whose multidisciplinary competences were instrumental in catalyzing the flourishing of this process. Such developments led at the end of the 1930s to Oppenheimer's contributions on the relativistic gravitational collapse of a neutron star. These works were rediscovered after the war and became a starting point for further investigations on the connection between compact objects and general relativity, eventually leading to what Kip Thorne called "the *golden age* of black hole research" (Thorne 2003, pp. 74–80), the decade going from 1964 to the mid-1970s, "an era that revolutionized our understanding of general relativity's predictions" (Thorne 1994, p. 258).

In the renaissance of general relativity and cosmology, central themes have been of course the study of relativistic gravitational collapse, black holes and neutron stars (Miller 1993). Fascinating reconstructions of the evolution of ideas about black holes and other 'dark astrophysical objects' were offered by physicists who have been protagonists in the quest to understand Einstein's legacy and its 'predictions about the

Universe'. I am especially referring to Kip Thorne (Thorne 1994) and Werner Israel (Israel 1987), whose valuable efforts have contributed to outline fundamental steps along the 'meandering paths' of this history, providing an important basis for reflecting on the evolution of scientific ideas and the formulation of new concepts, that together with astronomical observations fuelled the actual merging of astrophysics with general relativity. Other excellent essays have addressed the related evolution of the concept of neutron stars (Baym 1983) or more specifically have discussed the contributions to this story by main actors in this narrative, like Georges Lemaître (Eisenstaedt 1993; Kragh 1987), Robert J. Oppenheimer and Lev Landau (Hufbauer 2006, 2007; Yakovlev 1994), or Subrahmanyan Chandrasekhar (Wali 1990) and George Gamow (Hufbauer 2009; Kragh 2005). Other relevant references will be cited in due course.

The present attempt has instead adopted the perspective to follow the path of the evolving concept of a dense core in stars using it as a guiding key to reconstruct in detail the tapestry of interrelated ideas and changing models related to a series of fundamental questions: on one side the theoretical problem of the structure and evolution of stars up to their endpoint states, and on the other the role of such core as a virtual laboratory to investigate the behaviour of matter under extreme conditions of densities and pressures prevailing in stars that are impossible to realize on Earth. Its evolution as a challenging physical object – constantly connected to the problem of the origin of stellar energy – transformed the core into a testing ground for the emerging field of nuclear physics, also testifying the quickly changing relationship between physics and astrophysics during the 1930s. Such investigations, resulting from the interaction between different material and intellectual cultural practices, provided the multifaceted context and the theoretical framework within which Oppenheimer and his collaborators were able to work out the final fate of a collapsing neutron star at the end of the 1930s, which in retrospect was considered "the greatest of his discoveries: the black hole" (Schucking 1989, p. 50).

2 Prologue: The 'nonsensical message' of white-dwarf stars

In 1925, while Einstein was generalizing Bose's distribution function for the case of a fixed number of particles, Wolfgang Pauli, stimulated by Edmund C. Stoner's analysis of the quantum states of the electrons in complex atoms (Stoner 1924), proposed the exclusion principle as a general phenomenological rule governing the behavior of electrons in multi-electron systems (Pauli 1925).

Pauli's proposal triggered the development, independently by Enrico Fermi (Fermi 1926a,b) and P.A.M. Dirac (Dirac 1926), of a quantum statistics applicable to a gas of particles that obey the exclusion principle. As a fundamental physical principle rooted in quantum theory, the new quantum statistics provided the tool to treat an assembly of identical particles like a gas of electrons, and in turn, it immediately prompted Pauli's quick reaction (Pauli 1927). In order to prove that the Fermi-Dirac statistics – and not Bose-Einstein – was the right statistics to be applied to the degenerate electron gas ("beim materiellen Gas die Fermische und nicht die Einstein-Bosesche Statistik die zutreffende ist"), Pauli derived a physical consequence that could be experimentally verified: he pointed out that the weak temperature-independent paramagnetism of the metals, might be interpreted semi-quantitatively by representing the conduction electrons – free to move inside the metal – as a 'Fermi gas' of free particles and demonstrated that the electron gas in a typical metal is highly degenerate[1].

[1] Dirac derived the general theory of the behavior of quantum particles including both Fermi's result (which he apparently did not know about) and the Bose-Einstein result as

In December of that same 1926, while Pauli was submitting his pivotal contribution applying the Fermi-Dirac statistics to metals, the only form of dense matter known on Earth, Ralph Howard Fowler's paper 'On Dense matter', actually the *very first* application of the new statistics, discussing a degenerate gas of electrons in white dwarf stars, had already appeared in the 10 December issue of the *Monthly Notices of the Royal Astronomical Society* (Fowler 1926a).

Fowler's interest in the quantum theory and in the applications of physical ideas to the theory of valence, made him especially enthusiastic of the new quantum mechanics and its application to various areas of mathematical physics[2]. Fowler's early experience in problems of the behavior of solutions of second-order differential equations was at the root of his investigations with Edward A. Milne on stellar structure and the application of kinetic theory and statistical mechanics to stellar atmospheres (Fowler and Milne 1923). The statistical-mechanical investigations continued with further papers on the absorption lines in stellar spectra and on the ionization in stellar interiors written alone or in collaboration, like the seminal studies of gases in stars (Fowler and Guggenheim 1925a,b).

By this time, Fowler's studies of gases in stars, matched with his deep knowledge of statistical mechanics (Milne 1945), had fully set the stage for his interest in what appeared to be very peculiar stellar objects that had puzzled astronomers since many years (Holberg 2009).

Since 1862, the astronomer Alvan Clark, Jr. had been able to see the companion of Sirius, a very faint star, almost exactly ten thousand times fainter than Sirius itself, whose existence had already been discovered many years earlier, only through its gravitational influence. By 1910, reliable data showed that the faint companion, Sirius B, had a mass equal to 0.96 of that of the sun ($0.96M_\odot$), but was 400 times less luminous. "Nothing unusual thus far, but then came the bombshell" recalled Willem J. Luyten many years later (Luyten 1960, p. 30). From 1921 Luyten began a systematic general survey of the whole sky to search for white dwarfs, and it appears that he was the first to use the term that was subsequently popularized by Arthur Eddington (Holberg 2009).

In 1915 the American astronomer Walter Adams, an expert in stellar spectra, was able to secure the spectrum of the faint companion of Sirius A: "The great mass of the star, equal to that of the Sun and about one-half that of Sirius, and its low luminosity, one-hundredth part of that of the Sun and one ten-thousandth part of that of Sirius, make the character of its spectrum a matter of exceptional interest" (Adams 1915). The spectrum of Sirius B was quite puzzling: contrary to every expectation, Sirius B was white, in spite of its very low intrinsic brightness. Its spectrum was not very different from that of Sirius A[3].

In 1924 Arthur S. Eddington, the most influential astrophysicist of his time, brought these remarkable properties to the attention of the astronomical world. At the time, the conventional wisdom was that equilibrium against gravitational collapse was maintained in all stars by the internal pressure of the matter composing the star which had been heated into a gas, presumably by 'subatomic energy', as Eddington

special cases of his general theory. But Fermi's more 'physical' approach, discussing the problem of the quantization of a monoatomic ideal gas in a harmonic trap, probably explains why the expression 'Fermi gas' – and not Fermi-Dirac gas – is since then generally used in referring to an ensemble of a large number of fermions.

[2] After the great war he had started working on quantum theory, the kinetic theory of gases and in particular statistical mechanics, a field in which he made remarkable contributions. During his collaboration with Charles G. Darwin between 1922 and 1923, Fowler developed new methods in statistical mechanics that were later also applied to deal with the equilibrium states of ionized gases at high temperature.

[3] At that time, many dwarf, faint stars were already known, but they were all red.

pointed out. He had actually been one of the first to put forward such hypothesis. In discussing the relation between the masses and luminosities of the stars (Eddington 1924a), Eddington dedicated a specific section to white dwarfs, that "have long presented a difficult problem". Eddington then synthetized his views in an article sent in parallel to *Nature* (Eddington 1924b) in which he emphasized the importance of giant stars and white dwarfs as objects apparently escaping the standard laws that at that time allowed the construction of the Hertzsprung-Russell diagram, which gave the relationship between luminosity and surface temperature of a star, and according to which all stars appeared to be arranged in a practically continuous sequence.

An ordinary gas becomes comparatively incompressible at high density because of the finite volume occupied by its atoms or molecules. However, argued Eddington, atoms are mainly empty space (Eddington 1924b, p. 787): "at the high temperature within a star these sphere are completely destroyed, and this limit to the compression disappears. The stellar atom is highly ionized, and the peripheral electrons which determine its effective size have been detached [...] the ions, or broken atoms, can be packed much more tightly [...] There might thus exist stars far more dense than any material yet known to us. This may be the key to a puzzle presented by the companion of Sirius and a few other stars known as 'white dwarfs'." As he himself mentioned in (Eddington 1926, p. 10), "it had been suggested to him independently by Newall, Jeans and Lindemann that in stellar conditions the atoms themselves would break up to a considerable degree, many of the satellite electrons being detached"[4]. The conclusion was that the deduced very high density, according to the views he had presented, should not be accepted "as absurd".

Eddington, who was already well known for his commitment to Einstein's general theory of relativity, immediately added that it seemed unnecessary to debate the proposed alternatives at length, because, as several writers had pointed out, "the question could probably be settled by measuring the Einstein shift of the spectrum" for which Eddington proposed a value of about 20 km per second, "if the high density is correct".

Within a year, Adams, following Eddington's request, had carried out careful spectroscopic observations with the 100-inch telescope and measured the redshift. He found that, after allowance was made for the relative orbital motion of the two stars, the observed displacement was 19 km/s (Adams 1925a) (Adams 1925b, p. 387): "The results may be considered, therefore, as affording direct evidence from stellar spectra for the validity of the third test of the theory of general relativity, and for the remarkable densities predicted by Eddington for the dwarf stars of early type of spectrum". Eddington commented (Eddington 1926, p. 173): "This observation is so important that I do not like to accept it too hastily until the spectroscopic experts have had full time to criticize or challenge it; but so far as I know it seems entirely dependable. If so, Prof. Adams has killed two birds with one stone; he has carried out a new test of Einstein's general theory of relativity and he has confirmed our suspicion that matter 2000 times denser than platinum is not only possible, but is actually present in the universe". According to the astronomer Henry Norris Russell, this remarkable result was marking "a very definite advance in our knowledge of both the foundations of science and the constitution of matter" (Russell 1925a)[5], and Hans Thirring considered this effect a new, useful tool for the astrophysicist (Thirring 1926).

Adam's measurement, which would be strongly revised at the end of the 1960s (Greenstein et al. 1971; Holberg 2010), at the moment provided an evidence for the extreme compression of stellar matter, as emphasized by Eddington, and made clear that the existence of stars having the extraordinary qualities of what were by 1924

[4] The problem was thoroughly discussed in (Eddington 1926, pp. 165–172).

[5] See also Russell article 'Remarkable new tests favor the Einstein theory' (Russell 1925b).

cited in the literature as *white dwarfs*, not only was removing the necessity of confining the Einstein test to the sun, but was establishing for the first time a connection between general relativity and compact objects lying light years away, well beyond the solar system. But this fundamental thread of our story remained suspended and isolated for a long time.

Eddington's major monograph *The Internal Constitution of the Stars*, which concluded and summarized the results obtained during the first quarter of our century, was published in 1926 (Eddington 1926). Great progress had been made in the preceding years in the study of stellar interiors. The fundamental equations governing the structure of a star in radiative equilibrium had been established, and the role of ionization in determining the properties of interior stellar matter had been clearly recognized. Eddington's 'standard model' of stellar structure based on stars for which the perfect-gas law held and energy transport via radiation prevailed, yielded information on temperature and density in the interior of main-sequence stars and it was realized that the ideal-gas equation of state was a good approximation for all these stars.

Eddington dedicated a large discussion to white dwarfs. The extremely high density of the companion of Sirius A had been confirmed by Adams – but the puzzle remained. He was in fact uneasy as to what would ultimately happen to these superdense stars: "I do not see how a star which has once got into this compressed condition is ever going to get out of it. So far as we know, the close packing of matter is only possible so long as the temperature is great enough to ionise the material. When the star cools down and regains the normal density ordinarily associated with solids, it must expand and do work against gravity. *The star will need energy in order to cool*". At zero temperature all random motion should cease, according to ideas generally accepted up to 1926. In a cold star, nothing should prevent electrons and nuclei from recombining. The star would need to expand 10,000-fold to accommodate the volume of its neutral atoms. Where would it find the energy to do this? Its available fuel has been exhausted and it has no other resources: "We can scarcely credit the star with sufficient foresight to retain more than 90 per cent. in reserve for the difficulty awaiting it. It would seem that the star will be in an awkward predicament when its supply of sub-atomic energy ultimately fails. Imagine a body continually losing heat but with insufficient energy to grow cold!" concluded Eddington (Eddington 1926, p. 172).

Eddington had already remarked in 1922, during his Royal Astronomical Society Centenary address, that "Strange objects which persist in showing a type of spectrum entirely out of keeping with their luminosity, may ultimately teach us more than a host which radiate according to rule" (Eddington 1922). But what was the meaning of the apparently 'nonsensical message' (Eddington 1927, p. 48) coming from the companion of Sirius? Eddington had actually materialized a veil, whose corner was lifted by Ralph Fowler, who promptly responded, taking up the challenge and addressing the brand new astrophysical problem related to the nature of such dense stars.

Eddington's book *The internal constitution of stars* had been written between May 1924 and November 1925. As explained in the Preface dated July 1926, Eddington worked on the proofs up to March, and Fowler himself probably read at least parts of the volume in a preliminary stage. At the end of the preface Eddington is acknowledging him as his "referee in difficulties over points of theoretical physics" while Milne had read the proof sheets. Eddington, Milne and Fowler, were member of a small circle of very influential scientists, with strong common interests, all working in Cambridge. Their scientific discussions stimulated Fowler to apply quantum mechanics to the white dwarf problem raised by Eddington, whose enigma perfectly matched the advent of quantum mechanics and the Pauli exclusion principle.

Fowler's own interests and publications of 1926 are self explanatory in this sense. In April 1926 he had published 'The statistical mechanics of assemblies of ionized atoms and electrons', a detailed theoretical analysis in terms of electrons and positive nuclei that allowed him to tackle the properties of matter at the temperatures and densities occurring in stars (Fowler 1926b). On August 26, 1926, Dirac's paper containing the Fermi-Dirac distribution was communicated by Fowler to the Royal Society (Dirac 1926)[6]. On November 3, Fowler presented his own work to the Royal Society in which he systematically worked out the quantum statistics of identical particles, exploring the relationship between statistical mechanics and the new quantum mechanics, especially in connection with the Fermi-Dirac statistics (Fowler 1926c). After having thoroughly delved into the question, Fowler could thus devote his attention to Eddington's paradox, "*A star will need energy to cool*" and he tackled the problem in a most general perspective. Fowler reformulated the paradox posed by Eddington in clearer physical terms and resolved it introducing the notion of electron degeneracy for the first time. At the temperatures and densities that may be expected to prevail in the interiors of the white-dwarf stars, the electrons will be highly degenerate and all the available parts of the phase space with momenta less than the Fermi threshold are occupied, consistently with the Pauli exclusion principle. They fill all the energy levels, exactly like the electrons in an atom on the Earth. Therefore, the total kinetic energy evaluated according to such distribution will be about two to four times the negative potential-energy and Eddington's paradox does not arise (Fowler 1926a, p. 115): "The apparent difficulty was due to the use of a wrong correlation between energy and temperature, suggested by classical statistical mechanics".

In his classical monumental volume *Statistical Mechanics* Fowler well described the "absolute final state" – which he named *black-dwarf stage* – in which there is only one possible configuration left, when temperature ceases to have any meaning, and the pressure of the fully degenerate electron gas is large enough to balance the weight of the stellar layers attempting to collapse inward due to the gravitational pull (Fowler 1929, p 552): "As these stars go on radiating they will if anything condense still further and ultimately may well lose all their superfluous energy and fall to zero temperature. We may perhaps venture to refer to their probable final state as the black dwarf stages [...] The black-dwarf material is best likened to a single gigantic molecule in its lowest quantum state. On the Fermi-Dirac statistics, its high density can be achieved in one and only one way, in virtue of a correspondingly great energy content. But this energy can no more be expended in radiation than the energy of a normal atom or molecule. The only difference between black-dwarf matter and a normal molecule is that the molecule can exist in a free state while the black-dwarf matter can only so exist under very high external pressure".

Fowler's 1926 paper constituted a major breakthrough in astrophysical theory and would become one of the great landmark works in the realm of stellar structure. It was the first demonstration that the new quantum statistics could explain an important property of bulk matter and at the same time, in accounting in a general way for the observed characteristics of white-dwarf stars, it was a clear-cut example of the solution

[6] At that time, P.A.M. Dirac was Fowler's research student. Working under his influence, Dirac wrote papers on quantum theory and they also collaborated writing several articles in the period 1924-1926. Fowler's great commitment to the new quantum mechanics, testified by his scientific production of those years, was instrumental also in the early formation of Robert J. Oppenheimer, who was in Cambridge during the period 1925-1926. See for example his first paper related to his sojourn in Cambridge, entitled 'On the Quantum Theory of the Problem of the Two Bodies' presented in July 1926 by Fowler at the Cambridge Philosophical Society (Oppenheimer 1926). Apart from thanking Fowler and Dirac, Oppenheimer collaborated with Fowler in two papers published in that same 1926.

of an astrophysical problem depending upon features in which quantum mechanics differs essentially from any previous theory.

In mentioning white dwarfs and difficulties in the problem of stellar evolution, Eddington concluded his small volume *Stars and Atoms* expressing in the preface the feeling that the whole difficulty seemed to have been removed by R.H. Fowler's investigations, but cautiously adding that "there is something of fundamental importance that remains undiscovered". The very last words of the volume were dedicated to what was believed to be the final state of the white dwarf and perhaps therefore of every star: "If any stars have reached state No. 1 they are invisible; like atoms in the normal (lowest) state they give no light. The binding of the atom which defies the classical conception of forces has extended to cover the star. I little imagined when this survey of Stars and Atoms was begun that it would end with a glimpse of a *Star-Atom* [emphasis added]". Eddington could not imagine how prescient he was in saying that "white dwarfs appeared to be a happy hunting ground for the most revolutionary developments of theoretical physics" (Eddington 1927, p. 125). It was only the beginning. Further developments in the study of dense matter, the emergence of modern nuclear age as well as a new generation of scientists, would set the stage for modern challenging theoretical questions stemming from this new state of matter.

3 Metals and star interiors: Yacov Ilich Frenkel

Up to 1924, no one had given serious thought to abnormally dense matter. It was a remarkable coincidence that just at the time when matter of exceedingly great density was discovered in astronomy, physicists were developing the tools to tackle this subject. The idea of electrons free to move, but subject as an ensemble to the laws of quantum statistics, contained such a basic concept that was independently applied to electrons in metals. In seeing the proofs of Pauli's paper in spring 1927 (Pauli 1927), Arnold Sommerfeld, who well knew the problems of the classical electron theory of metals, was so impressed that he said that "one should make further application to other parts of metal theory"[7]. Introducing the idea that the free electrons in a metal constitute a Fermi gas, he was in fact able to explain the heat capacity catastrophe within the framework of Fermi-Dirac statistics.

Sommerfeld presented his theory at the International Volta Congress, held in September 1927 in Como, Italy, (Sommerfeld 1928a) (Sommerfeld 1928b). As Pauli later recalled: "I met Fermi personally the first time at the Volta-congress in Como, 1927 [...] Heisenberg introduced us with the words 'May I introduce the applications of the exclusion principle to each other', or with some similar joke"[8]. Following his talk Hendrik A. Lorentz, Enrico Fermi, Edwin Hall, and in particular Yacov Ilich Frenkel, participated in the discussion. Having worked since 1924 on the theory of metals, on which Frenkel was considered an authority, he had been invited to the Volta Memorial Conference, where he delivered a paper on the theory of metals in which he first formulated the main premises of quantum theory of electric conductiv-

[7] W. Pauli to F. Rasetti, October 30, 1956. The Pauli Letter Collection, CERN Archives. As recalled in (Hoddeson et al 1992, p. 102): "Sommerfeld liked the electron theory of metals well enough to make it a theme at his institute. In the summer of 1927, he lectured in his special-topics course on the 'structure of matter' to a small circle of advanced students, showing basic consequences of the application of Fermi-Dirac statistics to the electron gas [...] Electrons in metals were the main concern in Sommerfeld's theoretical research seminar in the winter 1927/1928".

[8] Pauli to Rasetti, see footnote 7.

ity (Frenkel 1928)[9]. In March 1928, Frenkel continued by letter the discussion with Sommerfeld begun at the Volta Congress (Frenkel 1966, p. 130) and in mid June he submitted his article in which he applied the 'Pauli-Fermi' electron gas theory to the problem of the cohesive forces (Frenkel 1928b)[10].

The atoms in metals lose their last one or two electrons, and these are free to move inside the metal. This problem was quite similar to atoms in white dwarf matter: electrons stripped from atoms were free to move between the compressed nuclei over the entire star. Sommerfeld had now extended the classical electron theory of metals developed by Paul Drude and H.A. Lorentz including quantum statistics[11]. The great interest arisen around the quantum properties of the electron theory of metals led in a natural way many people – some of whom were already working in the field – to discuss dense matter in white dwarfs, which became a virtual laboratory to test theories on degenerate matter in more extreme realms. Frenkel, described by Peierls as "a man of great versatility and originality" (Peierls 1997, p. 318), wrote in this regard an article on the application of the Pauli-Fermi electron gas theory to the problem of the cohesive forces, whose importance for the theory of white dwarfs remained almost unknown to astrophysicists. The fourth, and last section of the paper is entitled 'Superdense stars' (Frenkel 1928b, p. 244). Frenkel never uses the term 'white dwarf', probably because he wants to present the theory in a more general 'physical' sense, not necessarily connected with specific astrophysical objects. Moreover, he does not cite (Fowler 1926a). In his third section, specifically dedicated to the Thomas-Fermi atom, Frenkel had attempted to transfer the statistical atom model to the nucleus regarded as a sphere filled with *protons and electrons* (following the spread general view on nuclear models of that time) and reasoned that the electrons inside the nuclear volume were a strongly compressed gas. As a second step Frenkel transferred this analogy to white dwarfs, assuming them to be essentially homogeneous spheres consisting of a "mixture of electrons and ions gases", that he called "Kerngas" [nuclear gas]. It is then remarkable that, basing on such a "bad" model of the nucleus, he was nevertheless able to get a good representation of white dwarf matter calculating the relationship among the mass, radius, and density of the star.

As others would do in the following months, Frenkel remarked that the strongly compressed electron gas becomes relativistic. He stressed how from condensation of a metal vapour, metallic matter is obtained that can be considered as a 'single gigantic molecule' in which outer electrons are no more bound to single atoms, but are forming a 'gas system'. A further compression would set all electrons free. It is not possible to get such conditions in an Earth laboratory, hence Frenkel is naturally led to reason on dense stars, with density $\rho = 10^6$ g/cm^3: "Such pressure can actually exist only inside stars with a sufficiently large mass".

---

[9] Frenkel worked at the Leningrad Physico-Technical Institute on topics connected with problems of the structure of matter – especially solid and liquid bodies. Between 1922 and 1924 he had published *The Structure of Matter*, a complete theoretical analysis of the field, and in 1924 the *Electron Theory of Solids*, which later served as a basis for further original work. In particular, he also published papers on the theory of electric conductivity of metals and on the electron theory of solids, being considered one of the outstanding physicists of the Soviet Union. He spent in the period 1925-1926 in Germany, at a time coincident with the foundation of quantum mechanics. During 1927-1929 he published mostly on the theory of metals. See for example (Frenkel 1928a,c).

[10] The details of this pioneering work, in which Frenkel cites his previous research on metal theory, are discussed in (Yakovlev 1994).

[11] In this regard, see Sommerfeld and Bethe's review in the 1933 *Handbuch der Physik* (Sommerfeld and Bethe 1933) and (Hoddeson et al. 1987) for a reconstruction of the development of the quantum-mecanical electron theory of metals during the period 1928-1933.

Frenkel's investigations not only show how quickly stars were becoming a testing ground for theoretical physicists, but testify the ongoing transition of interest from the outer atomic layers to the nuclear realm. The method of analogies, as a search for connection among different physical contexts, was typical for Frenkel. He applied these considerations to reason on the superdense matter inside heavy nuclei treating them as solid bodies, a model having in turn a strong analogy with Gamow's coeval treatment of nuclei as an assembly of α-particles "treated somewhat as a small drops of water in which the particles are held together by surface tension" (Gamow 1929, 1930, p. 386).

Frenkel's remarkable article also discusses the existence of two types of superdense stars, consisting of non relativistic and ultrarelativistic electron gas and correctly estimated that the mass of a stable star, which is in a relativistic degenerate state, cannot exceed a definite maximum, $M \geq M_{\odot}$, somewhat larger than the mass of the sun. This really unexpected result went completely unnoticed at the moment.

4 First debate on dense matter in stars

Fowler's work had already led to a first qualitative understanding of the structure of white dwarfs, but a quantitative theory was still needed. Like Frenkel, others used the Fermi gas model to calculate the relationship among the mass, radius, and density of the white dwarf stars, assuming them to be essentially homogeneous spheres of electron gas.

The fundamental question of the degeneracy of electrons inside stars was also discussed by the German astrophysicist Wilhelm Anderson, working at Tartu University in Estonia, in an article submitted to the *Zeischrift für Physik* in July 1928 (Anderson 1928a). In the last part he mentioned an hypothesis put forward by the Australian physicist Kerr Grant, who had quickly reacted to Eddington's discussion on the unusual density of white dwarfs proposing that the mean density at the centre of the star could even be "fifty' million" times the mean density, instead of only "fifty" times according to Eddington's guess (Grant 1926). Based on the assumption that the properties of stellar material do not vary in a continuous manner from the star's surface to its centre, Grant also suggested a *central core* in which formation of heavy elements could take place with conversion of matter into radiation.

Anderson then submitted a second paper at the end of December (Anderson 1928b) following an article by the Soviet physicist Georgii Pokrowski (Pokrowski 1928), who had put forward a theory according to which "the mass of a star must have a maximum value" that would be obtained when the nuclei of completely ionized atoms touched each other and had estimated this density to be $4 \times 10^{13\pm1}$ g/cm^3. Provided that the nuclei could not be compressed, this should be the maximum density that matter could be in, that is the state of nuclear matter[12].

In a subsequent article, Anderson continued the analysis of Pokrowski's theory extending these considerations to the whole mass of the star (Anderson 1929a). His opinion was that the highest possible density, of the order of 10^{13} g/cm^3 could be only reached only under a very high pressure, that is in a central "core" of a star where the main part of the mass should be concentrated. This meant that it could happen not only within white dwarfs, but also in giant stars. In this case the idea of gravitational contraction as a source of energy for the star could not be a surprising possibility.

[12] According to Pokrowski, when a star having the limiting mass is also reaching its maximum density the gravitational potential at surface must have the critical value $\phi = c^2$. "In this case, recalled Anderson, no energy can leave the surface of the star, because to remove the mass of a quantum of energy $\frac{h\nu}{c^2}$ a work of exactly $\frac{h\nu\phi}{c^2} = \frac{h\nu c^2}{c^2} = h\nu$ would be required (Anderson 1929a, p. 389).

Anderson did not mention Frenkel's contribution that had already appeared (Frenkel 1928b), but an added note at the end cited an article by Edmund Clifton Stoner (Stoner 1929) discussing the limiting density in white dwarf stars. Anderson correctly remarked that in Stoner's formula from which the maximum possible density in a star could be calculated based on its mass, the latter had "ignored the variability of the mass of the electron" attaining velocities of the order of the velocity of light for stars having masses of the order of the sun and announced a new article meant to discuss Stoner's theory.

Stoner, who had been working at the Cavendish Laboratory since the early 1920s, developed theoretical interests encouraged by Fowler, and in 1924 published the already mentioned work on the distribution of electrons among atomic levels, a problem of topical interest to chemists as well as physicists. It attracted much attention thanks to Sommerfeld, who mentioned it as "a great advancement" in the preface of the fourth edition of his classic book, *Atomic Structure and Spectral Lines*[13].

Later Stoner developed a strong interest in magnetism and made a relevant contribution by introducing quantum ideas in the elucidation of the magnetic behaviour of matter, and did pioneering work on magnetism and the application of Fermi-Dirac statistics to the theory of para-and ferromagnetic phenomena. But his years in Cambridge – where eminent astrophysicists like Eddington and Milne worked – and of course Fowler's relevant paper, sparkled his interest on astrophysical topics, which became his lifelong interest.

Stoner wrote a paper to investigate "the question as to whether there is a limit to electron 'congestion' [...] under the gravitational conditions in the stars" (Stoner 1929, pp. 64–65). He was inspired by Eddington and Fowler, but especially by James Jeans's ideas about the departure from the ideal gas laws in some stellar interiors behaving "as if in a 'quasi-liquid' condition owing to the congestion of the atoms" (Jeans 1927),

He started from Fowler's idea that white dwarfs are supported by electron degeneracy pressure but went further, discussing under simplifying assumption whether there might be a limiting density "due to the 'jamming' of the electrons (owing to the exclusion principle which forms the basis of the Fermi statistics)". He modeled the star as a sphere of uniform density of material composed of completely ionized atoms: "the density increases as the sphere shrinks, and the limit will be reached when the gravitational energy released just supplies the energy required to squeeze the electrons closer together. The limiting case of high density occurs when the effective temperature is zero". In these calculations he followed Fowler in neglecting the electrostatic potential energy and considered only the kinetic energy in the degenerate electron gas and the gravitational energy of the star as given by Eddington (Eddington 1926, p. 87). He also neglected the kinetic energy of the nuclei, which is small, and obtained a limiting density of electrons, $n = 9.46 \times 10^{29}\ (M/M_\odot)^2\ \mathrm{cm}^{-3}$, in which M is the mass of the star in question and $M_\odot$ is the mass of the sun. He found an expression for the maximum density of a star of mass M, consisting of a mixture of fully-ionized atoms, approximately given by $\rho = 3.85 \times 10^6\ (M/M_\odot)^2\ \mathrm{kg\ m}^{-3}$, giving a value for the mean density of Sirius B in fairly good agreement with the modern value. The concept of a superdense 'core' within stars already materialized by Grant and Anderson,

[13] See Stoner's biography (Bates 1969), especially p. 214, where it is reported how the scheme came to Stoner's mind and how he wrote a note that he submitted to Rutherford, who in turn passed it to Fowler, with whom Stoner had at that time "several most helpful discussion on theoretical points". Fowler was so impressed that he asked him to write a full and detailed paper about it. As already mentioned, Stoner's article contained a scheme for the treatment of electron distribution derived from experimental data, that actually had a strong impact on Pauli, and influenced his formulation of the explicit statement that later became known as the Pauli exclusion principle.

was taken up by Stoner who concluded suggesting that "white dwarfs contain a core of material approaching the limiting density" being in an "almost incompressible or 'quasi-liquid' state, due to the 'congestion' of the electrons".

Such limiting density, observed Yûsuke Hagihara, a Japanese astronomer who had also studied in Cambridge under Eddington and Henry F. Baker during the 1920s, was considerably lower than a density of the order of 10^{17}, corresponding to the density reached by a star like the sun, whose volume had been reduced to a radius "equal to its Schwarzschild singularity $\left(\frac{2GM_{\odot}}{c^2}\right)$", that is "a few kilometers". "The most reasonable explanation" for this reassuring value, confirmed by the observations, would be "that this is the limit of the relativistically possible density". Hagihara also cited (Pokrowski 1928), who had given a value of the order of 10^{13} (Hagihara 1931, p. 107). In his 'Theory of the Relativistic Trajectories in a Gravitational Field of Schwarzschild', he thus emphasized that the solutions to the motion inside the circle corresponding to the Schwarzschild radius "is inadmissible from the principle of relativity", being "quite improbable that in any star the distance $r = \alpha$ or $2m$ from the center lies outside its radius" and that "the statement that a very massive star can entirely absorb the light emitted from its surface and never be seen from outside, is quite fallacious" (Hagihara 1931, pp. 173–174).

It seemed for a while that the white-dwarf stage – or rather the 'black-dwarf' stage as Fowler described it – represented the last stage of stellar evolution for all stars and thus their density appeared both from a theoretical and a physical point of view a limiting density. Moreover, since a finite state seemed possible for any assigned mass, one could rest with the comfortable assurance that all stars would have the 'necessary energy to cool', according to Eddington's expression. But this assurance was soon broken when it was realized that the electrons in the centers of degenerate masses begin to have momenta comparable to $m_e c$ and the electron gas must thus be treated relativistically.

As already remarked, Anderson immediately reacted to Stoner's paper criticizing calculations in which Stoner had used the rest-mass for the mass of electrons, and demonstrated that as the density increases the degenerate electrons in the centers of white dwarf stars comparable to or higher than the mass of the Sun, begin to attain velocities on the order of the velocity of light and that in this case the variation of the electron mass with velocity must be taken into account by using the equations of special relativity. He thus concluded that Stoner's assumptions led to "gröblich falschen Resultaten" [gross false results] in the case of white dwarfs having a mass comparable to the mass of our sun (Anderson 1929b, p. 852). His attempt to extend the equation of state of a degenerate electron gas to the relativistic domain was not correct, but it made the conceptual coupling of relativity and quantum statistical mechanics and indicated that Stoner's treatment implied a maximum value for the white dwarf mass.

Stoner's response to Anderson arrived in a paper submitted in December 1929 (Stoner 1930), where he worked out with more rigor the effect of the relativistic change of mass, still for the idealized case for a sphere of uniform density and formulating the correct relativistic equation of state (Nauenberg 2008; Thomas 2011). Stoner calculated that "For spheres of increasing mass the limiting density varies at first as the square of the mass, and then more rapidly, there being a limiting mass (2.19×10^{33} grams) [i.e. of the order of the sun's mass] above which the gravitational kinetic equilibrium considered will not occur" (Stoner 1930, p. 963), thus confirming Anderson's unexpected result of a critical mass for white dwarfs. On p. 952 of Stoner's article a figure with curves showing variation of limiting electron concentration with mass appears, comparing Anderson's and Stoner's results with the straight line which is obtained when special relativity is neglected. This figure is clearly showing that

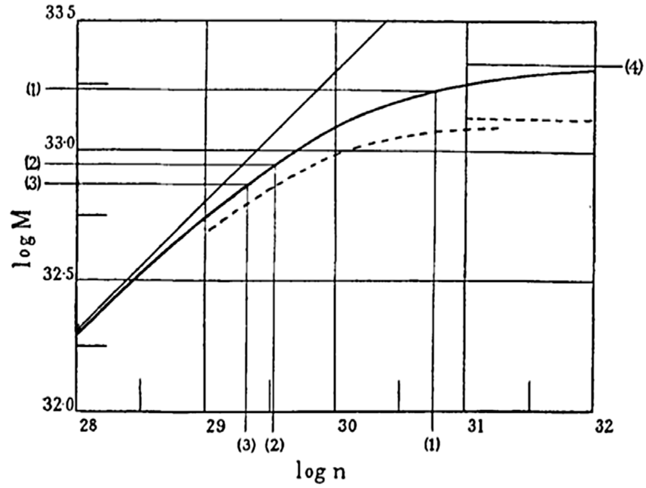


Fig. 1. Variation of limiting electron concentration (n) with mass (M) in a sphere of uniform density. The points (1), (2), (3), (4) correspond to Sirius B, o_2 Eridani B, Procyon B, and the limiting mass M (2.19×10^{33}) (Stoner 1930, p. 952). (Figure reproduced with the permission of Taylor and Francis Ltd, www.tandfonline.com. This figure is subject to copyright protection and is not covered by a Creative Commons license.)

a limiting mass is obtained when the crucial role of special relativity is considered. In the following page he commented that "The number of stars known to be of the white dwarf type is small, but this does not necessarily indicate that stars of very high density are uncommon. Dense stars of ordinary mass will have a small radius, and so will be faint objects [...] 'Black dwarfs' (to use Fowler's term) would not be observed".

Neither Stoner, nor Anderson speculated in these papers about what might be the fate of more massive stars. Stoner simply noted that "gravitational kinetic equilibrium will not occur" (Stoner 1930, p. 963). In 1936 Anderson then published his habilitation thesis: *Existiert eine obere Grenze für die Dichte der Materie und Energie?* [Does it exist an upper limit for the density of matter and energy?] (Anderson 1936).

5 Edward Milne and the idea of a condensed core in stars

The Stoner-Anderson debate on the structure of white dwarfs was someway embedded in a hot controversy between Milne and Eddington about the problem of stellar structure and the source of energy in stars, in which several other people became involved. At the end of 1929, Milne presented his investigations on the relation between the masses, luminosities, and effective temperatures of the stars from a standpoint which was "philosophically different" from that adopted by Eddington (Milne 1929, p. 17). Milne criticized Eddington's theory on the ground that stars with a point source of energy, and a point concentration of mass, at the center (or a reasonable physical approximation to this arrangement) would be more stable than Eddington's models; in such stars the central temperatures would be very high indeed[14].

On the other hand, underscored Milne (Milne 1930a, p. 16), "The theory of Sir Arthur Eddington does not claim to account for the observed division of stars into dense stars and stars of ordinary density, nor does it establish the division of ordinary stars into giants and dwarfs. On the other hand, it claims to establish what

[14] For an excellent detailed discussion on the Milne-Eddington controversy and for a comprehensive history of theories of stellar structure see (Shaviv 2009).

is known as the mass-luminosity law from considerations of equilibrium only, that is, without introducing anything connected with the physics of the generation of energy". Inspired by dense matter in white dwarfs – and by the Grant-Anderson-Stoner discussions on a dense core – Milne's investigations materialized in attempts to construct stellar models by using the properties of degenerate matter and by introducing the idea of collapsed models, too massive for their gas pressure to support their mass if perfect-gas conditions prevailed, and of centrally-condensed models, whose density and temperature rose to enormous values at the centre (Milne 1930a) (Milne 1930b).

Previous more vague ideas of condensed cores within stars already mentioned were now acquiring a definite status in Milne's theory and were meant to play a role in the constitution of *all* stars.

In the opening lines of his 'Analysis of stellar structure' (Milne 1930a), Milne stressed how, according to the then current theory (by which he meant that of which the researches of Eddington were the basis), the ordinary stars (giants and dwarfs) were considered masses of perfect gas with central temperatures of the order of 10^7 degrees and with central densities of the order of 50 times the mean density. However, stressed Milne, the current theory failed entirely to account for 'white dwarfs', so that he wanted to show in the paper that "a perfect-gas star in a steady state is in nature an impossibility, and that actual stars must either possess a small but massive core of exceedingly high density and temperature, or else must be almost wholly (that is, save for a gaseous fringe) at a very high density" (Milne 1930a, p. 4). According to Milne, this appeared to be a fundamental property of steady-state configurations, and it corresponded to the observed division of the stars into "ordinary stars" (giants and dwarfs) and "white dwarfs". "This result, added Milne, was not a consequence of any special hypothesis, but was flowing naturally from the method of analysis: "The division of 'ordinary stars' into 'giants' and 'dwarfs' would appear to be less fundamental and not to indicate any special difference of structure".

According to their luminosity they must be either 'centrally-condensed' or 'collapsed'. In such centrally condensed models, with super high temperatures and density, the perfect-gas law could not hold. Actually, for the first time, the term 'collapse' in the context of 'collapsed configurations' that would become of common usage in the future, was first introduced in astrophysics by Milne in this theory.

Milne's 'Analysis' paper had been preceded by another on the origin of stellar energy and the mechanism of its evolution (Milne 1930c) in which he located the source of stellar energy "In the intensely hot, intensely dense nucleus where the temperatures and densities are high enough for the transformation of matter into radiation to take place with ease", very provocatively stressing how "The consequences [of his investigations] amount to a complete revolution in our picture of the internal constitution of the stars". His ideas about the source of energy in stars were mixing on one side the hypothesis of matter annihilation, and on the other the possibility of a "synthesis or radioactive elements" thanks to the extremely high temperatures "of the order of 10^{10} degrees or higher" that were "to be assigned to the central condensation"[15].

Milne's papers on the analysis of stellar structure aroused a large and vigorous debate. These issues, that had for long already been a subject of disagreement between James Jeans and Arthur Eddington, entered a new phase through the work of Milne, whose views aroused a great interest because of the novelty underlying his proposal. The whole of the meeting of the Royal Astronomical Society of January 9, 1931, was dedicated to a general discussion on such fundamental topics, but Milne's ideas on the composite model for stars were not generally accepted. Milne's model was challenged especially by Norris Russell, Thomas Cowling and Bengt Strömgren. The latter reported the main critics in (Strömgren 1931).

[15] Milne went back to this problem in a new article with a quite similar title (Milne 1931a).

"The irony of all this", as emphasized in (Shaviv 2009, p. 238), "is that Eddington thought at the beginning that his theory explained the gaseous giant stars and not the dwarf (main sequence) stars. As it turned out, Eddington's theory explains the dwarfs and Milne's theory explains the giant stars".

However, at the threshold of the nuclear era, these animated discussions among astrophysicists contributed to the spreading of ideas about the "core" as a brand new physical object made of 'nuclear' dense matter within stars, ready for more specific "physical" investigations.

6 Chandrasekhar enters the lions' den

While Anderson and Stoner were publishing the results of their work about the relativistic effects on electron degeneracy in white dwarfs, the 19-years-old Subrahmanyan Chandrasekhar, known throughout his life as Chandra in the scientific world, was traveling on a ship from Bombay to Europe, determined to study and carry on research under Fowler at Cambridge. As a student at Presidency College from 1925, Chandra found a growing liking for physics and mathematics and an ongoing attraction to English literature[16]. In the autumn of 1928, during his trip around the world, Sommerfeld visited India and lectured at Presidency College in Madras. Chandra made it a point to meet Sommerfeld, from whose book *Atombau und Spektrallinien* [*Atomic Structure and Spectral Lines*] he had learned quantum theory: "I saw in the newspapers that he was going to be there, and so I went and saw him in the hotel [...] He gave me the copy of his papers on the electron theory of metals which were then in press, and his papers were clear enough for me to understand the Fermi statistics". Chandra was taken aback to learn that the old Bohr quantum theory, on which Sommerfeld's book was based, was superseded "by the discovery of wave mechanics by Schrödinger, and the new developments due to Heisenberg, Dirac, Pauli and others", and that the Pauli exclusion principle replaced Boltzmann statistics with Fermi-Dirac statistics (Wali 1990, p. 62). The young student was someway shocked by such revolutionary news, but Sommerfeld offered him the galley proofs of his as-yet-unpublished paper on the new Fermi-Dirac quantum statistics and its application to the electron theory of metals.

"At about the same time, added Chandra, I read Eddington's *Internal constitution of the stars*. It's quite readable. And it was the simultaneous knowledge of Eddington's *Internal constitution of the stars*, together with modern statistics, at least modern as of then, through Sommerfeld, that turned my interest into the theory of white dwarfs and related matters". (Chandrasekhar 1977). Glancing through the *Monthly Notices of the Royal Astronomical Society* he found Fowler's paper containing still another application of the Fermi-Dirac statistics to the dense stellar matter in the form of degenerate electrons in white dwarfs: "So it seemed to me that there was an area in which one could go right in. I could understand the Fermi statistics; I knew the theory of polytropes; I had read Fowler's paper; I could understand it. Right there, there was something which I could do. So that is how I started" (Wali 1990, p. 62). Within a few months he had enough mathematical preparation to understand the new statistics and was able to apply his knowledge to the problem of Compton scattering (Chandrasekhar 1929). In January 1929 Chandra sent it to Ralph Fowler, whose monumental book *Statistical Mechanics* had just come out. So Chandra's scientific career began with a series of papers on 'the new statistics' published between 1929 and 1930, when he was only an eighteen-year-old undergraduate student. During Chandra's final year at Presidency College, Werner Heisenberg went on a lecture tour

[16] For biographical portraits of Chandra see (Wali 1990) and (Parker 1997).

and Chandra was in charge of Heisenberg's visit to the college and was entrusted with the responsibility of showing him around Madras. He had a wonderful occasion of discussing with him his papers and of increasing his expectations towards the idea of perfecting his studies in Europe. Even before completing his final examinations, Chandra was awarded the Government of India scholarship and on 31 July 1930, when he was only 19-years-old, he left Bombay on the steamer Pilsna, a liner of Lloyd Triestino travelling from Bombay to Venice across the Arabic Sea, the Channel of Suez and the Mediterranean.

During his travel towards Venice, from where he would travel by rail to London, he continued to work on a paper he had completed just before his departure. In it he had developed Fowler's theory of white dwarfs further, combining it with Eddington's mathematical model for an isolated mass of gaseous stellar material in equilibrium under its own gravitational forces, the so-called polytropic gas sphere, a crude approximation to more realistic stellar models with a simple relationship between pressure p and density ρ: $p = K\rho^{1+1/n}$, where n, the polytropic index, and K are constants that depend on the properties of the particles making up the gas. In working out the statistical mechanics of the degenerate high density electron gas at the center of the white dwarf, he realized, as Fowler had not, that the upper levels of the degenerate electron gas (which are those affected by changes in density and temperature) are relativistic. This meant that the pressure supporting the star against gravity grows no faster than the increasing gravitational force as the star contracts, in contrast with the familiar nonrelativistic situation where the pressure increases more rapidly than the gravitational forces ultimately providing a sufficient pressure to block further contraction. Chandra thus found that this limiting form of the equation of state had a dramatic effect on the predicted mass-radius relation: instead of predicting a finite radius for all masses, the theory was now predicting that the radius must tend to zero as a certain limiting mass is reached, above which the internal pressure of the white dwarf cannot support the star against collapse.

The value of the limiting mass found by Chandra was $5.76\mu_e^{-2}M_\odot$ where μ_e denotes the mean molecular weight per electron. For the expected value $\mu_e = 2$, the limit is 1.44 solar masses[17]. The existence of this limiting mass meant that a white-dwarf state does not exist for stars that are more massive.

This paper on the limiting mass was rather puzzling also for Fowler, and for this reason its publication was rather delayed, as recalled by Chandra (Chandrasekhar 1977): "I had written this paper in July; and I gave it to Fowler in September and he never did anything with it, whereas he sent my other paper to the *Philosophical Magazine*. And fundamentally it is because neither Milne nor Fowler wanted to accept the fact that there was a maximum mass...". It eventually appeared in July 1931 in the *Astrophysical Journal*, published by the University of Chicago, at the time much less important than the *Monthly Notices of the Royal Astronomical Society* where Eddington and Milne generally published their work.

When he was writing his first paper, during the summer 1930, Chandra did not understand what the mass limit meant: "I didn't know how it would end, and how it related to the 3/2 low-mass polytropes [...] I knew it must be significant, because Milne was working on the 3/2 polytropes at that time. He thought that every star must have a white dwarf core. And I couldn't see how that could be true". But Chandra also recalled: "I would say that I fully understood its implications by the end of 1930". That is after having worked at his second and especially at his third

[17] Both Stoner and Chandra assumed that the mean molecular weight of the gas is 2.5. This value that would soon have to be substantially reduced in the light of the evidence presented by H.N. Russell and B. Strömgren that stars contain large amounts of hydrogen, even if in the case of the white dwarf stage the hydrogen has been largely consumed.

paper on the theory of white dwarfs communicated by Milne himself at the Royal Astronomical Society (Chandrasekhar 1931c)[18].

Chandra entered the field of stellar structure, the central thread of theoretical astrophysics, at a time when the Eddington-Milne controversy was on the verge of exploding. In this doing, he had to confront towering figures like the so-called 'triumvirate' represented by Jeans, Eddington and Milne, that dominated the scene in the British area, and not only. Especially Eddington and Milne were invariably cited by anybody entering the field of the internal constitution of stars.

On November 14, 1930, Fowler invited Chandra to attend the meeting of the Royal Astronomical Society, because he wanted to introduce him to Milne, who that same day was presenting his paper 'destroying' Eddington's view of the interior of stars: "The paper is supposed to cause a lot of sensation," Chandra wrote to his father (Wali 1990, p. 85). He also added: "Milne's results are in a sense a generalization of my own on the density of dwarf stars [...] Mine is one of the limiting cases of Milne's formulae. I think he will refer to my papers". But this was not the case, Milne did not mention Chandra's contribution, even if he followed his work and encouraged him, later even suggesting a collaboration. However he did not really understand the importance of Stoner's and Chandra's results, which were clearly disturbing his theory according to which *all* stars must have a white dwarf core. At the same time, his interest in Chandra's work was justified. He hoped that Chandra's work might support him in his rivalry with Eddington. One of the main consequences of Milne's analysis was the explanation not only of the existence of white dwarfs – his collapsed configurations – but also of the principal characteristics of these configurations. In his third contribution to the subject of white dwarfs (Chandrasekhar 1931c), following Milne's analysis of stellar structure (Milne 1930a,b), Chandra wanted to develop the theory of collapsed configurations a stage further and with this aim he performed a very detailed analysis introducing the relativistically degenerate core.

In the meantime, the same idea of using the Lane-Emden equations for polytropes (Chandrasekhar 1939, pp. 84–182), taking into account the special relativistic effects in the equilibrium of stellar matter for a degenerate system of fermions, came independently to Lev Davidovitch Landau, who, like Chandra, was quite explicit in pointing out the existence of the critical mass. His paper (Landau 1932), appearing in 1932, had the roots in Milne's proposal of a composite model for stars, and was meant as a critical contribution to the Eddington-Milne debate. But it went much beyond and had a key role in exporting Milne's idea of dense cores into the realm of physics. This migration to a different cultural context determined the starting of a new career for superdense cores within stars, which in Landau's theoretical investigations transformed into a well defined and promising physical object, whose potentialities would later be fully revealed with the advent of the neutron era. But at the moment Landau was not able to grasp the deep implications of his work, since he wrote his paper a year before Chadwick announced the discovery of the neutron.

[18] His second paper on the subject, 'The density of white dwarf stars', actually appeared in the Supplement of February 1931 to the *Philosophical Magazine* (Chandrasekhar 1931b). Now Chandra, who had become aware of Stoner's work (Stoner 1929) (Stoner 1930), reconsiders the problem of the density from the point of view of the theory of polytropic gas spheres, deriving a formula for the mean density "on considerations which are a much nearer approximation to the conditions *actually existent* in a white dwarf". He thus avoided Stoner's assumption that the density is uniform throughout the star, but recognized that "At any rate...the *order of magnitude* of the density which one can on purely theoretical considerations attribute to a white dwarf is the same" (Chandrasekhar 1931b, p. 595).

7 Landau 1932: a transition paper

Since the mid-nineteenth century, spectral analysis as applied to the study of stars had established a new relationship between physics and astronomy. When quantum theory decoded the enigma of spectral lines, astrophysics became to grow as a branch of physics. The lively debate about the structure of stars and their source of energy which saw dense stars at the crossroad of different discussions implying the quantum behavior at microscopic level, gradually began to shift towards the nuclear realm. Already in 1920, in investigating models for the nucleus, Rutherford had put forward the idea of a neutral particle formed by a bound state of proton and electron that he named 'neutron'. This idea resurfaced in particular towards the end of the 1920s, when attention of physicists was more and more shifting from the outer electron layers of the atom (whose theory had been definitely settled by quantum mechanics) to the nucleus, that was now considered as the new frontier. Within this general trend, dense stars were being once more rediscovered as physical laboratories for speculating on nuclear processes.

In May 1931, well in advance with Chadwick's breakthrough short note to *Nature* (Chadwick 1932) announcing the 'Possible existence of the neutron', the particle postulated by Rutherford in his Bakerian Lecture to the Royal Society in 1920 (Rutherford 1920), Langer and Nathan Rosen of MIT proposed that the combination of an electron and a proton "would be very useful in explaining a number of atomic and cosmic phenomena" (Langer and Rosen 1931). Their main aim was to "offer a way of describing the process of building up of the heavier elements" but they also proposed that a part of the packing energy released in the formation of neutrons from hydrogen could be "radiated in a single quantum" thus explaining the production of cosmic radiation as observed by Millikan and others[19]. As a third application, in the section 'High density matter in stars', they proposed the neutron to be at the origin of the formation of very dense cores in stars: "The usual explanation of the white dwarfs involving a high degree of ionization of the atoms is not the only one. There are in fact great advantages from this point of view in favor of our neutron. Being small it has a great mean free path and is comparatively insensitive to light pressure. It therefore goes easily to the center of a gravitating mass. Being neutral and having an extremely small external field, it permits high densities to build up before it deviates appreciably from perfect gas behavior".

A growing interest in the nuclear dimension as a realm of relevant processes in stars is also testified to by an article by Seitarô Suzuki, 'Constitution of the white dwarf stars', in which he mentions white dwarf degenerate matter, assuming that all atoms of various elements are stripped of their extranuclear electrons and that all kinds of nuclei of atoms are formed entirely from protons and electrons. He then concluded that the heavy radio-elements exist abundantly in the white dwarfs (Suzuki 1931).

Again white dwarfs' dense matter is used to reason at a nuclear level, where still protons and electrons are the protagonists of nuclear processes. It is in this context that a much celebrated paper by Landau, 'On the theory of stars', dated February 1931, but appearing only in February 1932, was conceived (Landau 1932). Landau had begun his scientific activity in 1926 at the Leningrad Physical Technical Institute (now the Ioffe Physical Technical Institute, St. Petersburg) and graduated from Leningrad State University in January 1927, at the age of 19, having as supervisor Yakov I. Frenkel, head of the Theoretical Physics Department (Kapitza and Lifshitz 1969). Landau's early entourage included George Gamow and Dimitri Iwanenko. Their group

[19] In 1931, the theory of cosmic rays as charged particles had not yet been established, so they referred to the spread theory of cosmic radiation as very high energy gamma rays.

was known as the Three Musketeers at the University of Leningrad. Between 1926 and 1928 Landau worked with Iwanenko on quantum theory, but during this period he became particularly close to the brilliant Matvey Bronstein, who had in the meantime joined the so called jazz-band group. Gorelik and Frenkel (Gorelik and Frenkel 1994, pp. 23–24) outline how Bronstein, fascinated by astronomy, introduced his physicist friends to astronomers like Viktor Ambartsumyan, who would later become a first class astrophysicist[20].

In 1929, on an assignment from the People's Commissariat of Education, and later thanks to a Rockefeller grant, Landau travelled abroad and for one and a half years worked in Denmark, Great Britain and Switzerland. The most important experience was his stay in Copenhagen where, at the Institute of Theoretical Physics, theoreticians from all Europe were attracted by Niels Bohr. Like many others, and in particular like his friend Gamow, Landau was strongly influenced by Bohr, whom he always considered his only teacher.

Rudolf Peierls met Landau in Zurich for the first time during the autumn of 1929: "[...] we discussed things a lot [...] I cannot remember all the things, we discussed, but certainly he was then already very interested in astrophysics" (Peierls 1977). In Zurich, at that time, young physicists around Pauli, like Peierls, Bloch, Leon Rosenfeld, were struggling mostly on problems of metals and, of course, quantum electrodynamics, the most debated subject at that time (Rosenfeld 1963). According to Arkadii Migdal (Migdal 1977), "Landau's idea was that theoreticians should not be devoted to one special part of physics".

In the early spring of 1931, Landau shared his time between Copenhagen and Zurich[21]. Gamow, too, was there from September 1930 to May 1931, and during that last month he completed his book on the constitution of atomic nuclei and radioactivity, the first one ever on theoretical nuclear physics (Gamow 1931, 1968): "I remember that Landau was helping me with the mathematics, with calculating the perturbation and so on. And these formulas were all derived by Landau"[22].

It happened that, in that same period, on August 19, 1930, Fowler wrote to Bohr that he had "very exciting news from Milne. He is convinced now that he has found exactly, where Eddington is wrong in his astrophysical theories" and that he would tell more in detail on his arrival in Copenhagen in September. Milne had just published in *Nature* his article on stellar structure and the origin of stellar energy (Milne 1930c), where he put forward the idea that a core of very dense material would form within stars, "a kind of 'white-dwarf' at its centre, surrounded by a gaseous distribution of more familiar type; the star is like a yolk in an egg [...] It is to this nucleus that we must look for the origin of stellar energy, a nucleus the existence of which has previously been unsuspected". His detailed analysis of stellar structure outlining his theory of a composite model for stars would appear in the following months

[20] Still during his student years, Bronstein wrote his first astrophysical papers, which contained an important contribution to the theory of stellar atmospheres in the form of the so-called Hopf-Bronstein relation (Bronstein 1929, 1930). Milne himself, one of the founders of the field, recommended Bronstein's second paper for publication in the *Monthly Notices of the Royal Astronomical Society*. In 1931, *Uspekhi Fizicheskikh Nauk* published a detailed survey by Bronstein entitled 'The Modern State of Relativistic Cosmology' (Bronstein 1931). It was the first review of cosmology in the USSR. Landau and Bronstein continued to collaborate even after the former had moved to Kharkov in 1932.

[21] In particular, he visited Bohr from 8 April to 3 May 1930, 20 September–20 November 1930, and 25 February–19 March 1931 (Pais 1993, p. 359).

[22] As remarked by Pais, (Pais 1993, p. 325): "Most of the people Gamow thanked for valuable advice belonged to the Copenhagen circle: Bohr himself and also Gamow's friends and contemporary fellows at the institute, Hendrik Casimir from Leiden, Lev Davidovich Landau from Leningrad, and Nevill Francis Mott from Cambridge".

(Milne 1930a), but it was clearly the involved question of the stellar energy problem that definitely triggered the physicists' interest. On August 26 Bohr answered Fowler's letter: "It shall be a great pleasure to discuss the many actual problems and not least the interpretation of the astrophysical evidence in which I am very interested. He added that they were expecting in September various visits, in particular Rosenfeld and Gamow. He did not mention Landau, who actually had written him on August 23 from Cambridge about his intention of visiting Copenhagen some time in the middle of September[23].

Fowler arrived in Copenhagen on September 11. All this meant that already during the summer and early fall of 1930 Milne's theory was largely discussed in Copenhagen[24].

Landau's background easily explains the motivations behind his interest in these questions. Under the spell of Bronstein's passion for astrophysics, and having studied the unusual magnetic properties of the degenerate electron gas in a metal, like others at that time, Landau extended his investigations on the behaviour of a relativistic degenerate electron gas in a more extreme and challenging realm: the interior of stars[25]. As will be clarified in the following, Bohr himself was especially interested in the problem of stellar energy, so that the whole matter certainly became a hot topic during Landau's stay in Copenhagen. The first relevant paper applying quantum mechanics to stellar element synthesis by nuclear reactions – that can be actually regarded "as one of the pioneering contributions to nuclear astrophysics" (Kragh 1996, p. 85) – had been written in 1929 by Fritz Houtermans and Robert d'Escourt Atkinson based on Gamow's theory of α-decay (d'E. Atkinson and Houtermans 1929b). As a first step they had actually sent a short note to *Nature* 'Transmutation of the lighter elements in stars' (d'E. Atkinson and Houtermans 1929a). Theoretical nuclear physics was entering stellar interiors: solving the problem of energy generation in stars might also account for the abundances of the various chemical elements. As acknowledged by the authors, Gamow had been deeply involved in "numerous discussions", and it is easy to imagine that in turn he must have abundantly discussed these topics with Landau going back to Russia that same 1929, after having travelled through Europe.

Landau's paper is dated February 1931, but soon after he returned to Leningrad and only on January 7, 1932, he submitted it to the brand new *Physikalische Zeitschrift der Sowjetunion* – the first Soviet physical journal published in languages other than Russian – that is more than one month before Chadwick's announcement of having detected the neutron (Landau 1932). That same 1932 Landau moved to Kharkov, where he became head of the Theoretical Division of the newly organized Ukrainian Physicotechnical Institute, an offshoot of the Leningrad Institute[26].

[23] Archives for the History of Quantum Physics, Bd. AHQP/BSC 19: Niels Bohr. Scientific correspondence, 1930–1945.

[24] Milne's long paper appeared in the November number of the *Monthly Notices of the Royal Astronomical Society* (Milne 1930a), together with related papers by R.H. Fowler, N. Fairclough, and T.G. Cowling, presented during the meeting held at the Royal Astronomical Society on January 9 1931, completely devoted to a debate on the subject.

[25] In 1930, Landau collaborated with Frenkel in an article on the quantization of free electrons in a magnetic field. During his stay in Cambridge he tried to explain Pyotr Kapitza's results concerning the dependence of the electrical conductivity of metals on an external magnetic field applying the methods of quantum mechanics to the problem of the anomalous properties of the electric conductivity of bismuth in strong magnetic field. This work resulted in his theory of diamagnetism, that became a basis of research in solid-state physics.

[26] The establishment of the *Physikalische Zeitschrift der Sowjetunion* had been promoted by Iwanenko, who was in the Kharkov Institute of Physics from 1929 to 1931 as first director of its theoretical division (Sardanashvily 2014).

Landau's paper has generally attracted attention as one of the milestone's – actually as a starting point – in the path towards the idea of compact collapsed objects (Yakovlev et al. 2013). What is especially relevant here, is its genesis in the pre-neutron era, notably in the period preceding Fermi's solution to the problem of beta decay, that definitely banished electrons from the nuclear realm. In this sense Landau's paper is acting like a prism refracting different controversies both in the physical and the astrophysical realms. Landau's interests were extremely wide, moreover, in the mid-1930s, as he himself explained, "theoretical physics, unlike experimental physics, is a small science open to perception in its entirety by any theorist" (Gorelik 2005). Landau's excursion in the astrophysical realm of stellar theory, was triggered by the debate about the Milne-Eddington controversy. Moreover, during his stay in Copenhagen, Landau interacted with Bengt Strömgren, who acknowledged Landau's assistance in the first paper he wrote for the *Zeitschrift für Astrophysik* (Strömgren 1931) where he discussed Milne's ideas of a stellar nucleus of extreme density and temperature along lines differing from those followed by Eddington and examined the question of the existence of stellar configurations with a nucleus of this character. Strömgren, who had studied physics in Bohr's Institute for theoretical physics, was working at the Copenhagen Observatory, but he frequently attended conferences there, thus having the occasion to become familiar with foreign visitors. During his studies for the Master's degree, he became much impressed with the latest developments: "I had the idea that the time was ripe for applications of the new quantum mechanics to astrophysical situations". He also remembered that (Strömgren 1978) "Landau was a frequent visitor, and he was deeply interested in these questions, but had his own views that differed radically from those of other people [. . .] And his idea was that you have a very high density core, in the sun, and that release of gravitational energy therefore plays a role. For this reason he was inclined to disregard all of Eddington's work". Strömgren eventually became one of the leading theoretical astrophysicists in the world. Landau's motivation thus aroused within the hot topic of the stellar equilibrium in gravity, particularly the maximum mass of white dwarfs, where electron degeneracy pressure stands against gravity. Milne's theory aspired to explain the existence of all types of stars including the white dwarfs, explain the energy generation and eliminate the problem of the stellar absorption coefficient, by supposing that the mass, the luminosity, and the absorption coefficient were completely independent. Landau started criticizing Milne's arguments against Eddington's mass-luminosity relation, also annoyed by his excessive reliance on what he considered 'mathematical eccentricities', far from physics: "The astrophysical methods usually applied in attacking the problems of stellar structure are characterised by making physical assumptions chosen only for the sake of mathematical convenience" like, for instance, Mr. Milne's proof of the impossibility of a star consisting throughout of classical ideal gas". This proof, added Landau, rested on the assertion that, "for arbitrary L and M, the fundamental equations of a star consisting of classical ideal gas admit, in general, no regular solution". Landau then stressed that Milne seemed "to have overlooked the fact, that this assertion results only from the assumption of opacity being constant throughout the star, which assumption *is made only for mathematical purposes and has nothing to do with reality* [emphasis added]". Only in the case of this assumption, recalled Landau, the radius R disappears from the relation between L, M and R, "which relation would be quite exempt from the physical criticisms put forward against Eddington's mass-luminosity-relation". Once clarified his position towards astrophysicists' way of tackling such problems, Landau declared: 'It seems reasonable to try to attack the problem of stellar structure by methods of theoretical physics, i.e. to investigate the physical nature of stellar equilibrium" (Landau 1932, p. 285). As Cowling commented much later, "Both observing astronomers and physicists tend to wax critical of the mathematician, and sometimes with reason. Mathematicians try to construct models of stars: I remember Milne saying here in 1930 that he would no longer speak of stars, but only of spherical masses of gas" (Cowling 1966, p. 121).

Landau criticized Milne's introduction of "a condensed inner part of the system" as an *ad hoc* hypothesis, without explaining the reason why such condensations could appear at all, so that the connection between the condensed state and the normal state remained "rather mysterious". Then, independently, followed Chandra's in using the Emden-Lane equation investigating the statistical equilibrium of a given mass without generation of energy, and showing that in the case of classical ideal gas, there is no equilibrium at all: "Every part of the system would tend to a point". However, "the state of affairs becomes quite different when we consider the quantum effects". He then discussed the extreme-relativistic case finding that a star of fixed mass would have to either expand or collapse to a point to attain a minimum of the energy and reach an equilibrium state. In order to find the criterion separating the two cases he solved the $n = 3$ polytropic equation of Emden finding that an equilibrium state is reached only for masses smaller than a critical mass of about 1.5 solar masses, again of the order of values found by Stoner and Chandra. However, for masses greater than the critical mass, Landau remarked that "there exists in the whole quantum theory no cause preventing the system from collapsing to a point".

However, he continued, "As in reality such masses exist quietly as stars and do not show any such ridiculous tendencies we must conclude that *all stars heavier than* $1.5M_\odot$ *certainly possess regions in which the laws of quantum mechanics* (and therefore of quantum statistics) *are violated*". Landau stressed that there was no reason to believe that stars could be divided into two physically different classes according to the condition of having a mass greater or smaller than the critical mass, so that he supposed that all stars should possess those "pathological regions" avoiding the necessity of such division, and even that "just the presence of these regions makes stars stars".

But if this is the case, reasoned Landau, there was no need to suppose that the radiation of stars might be due to "some mysterious process of mutual annihilation of protons and electrons" (he is here referring to Jean's old ideas on annihilation as a possible source of stellar energy, also mentioned by Milne) because protons and electrons in atomic nuclei are very close together and "they do not annihilate themselves", even being *both* constituents of the nucleus, according to current ideas about nuclear matter. In dismissing astrophysicists' vague ideas on the sources of stellar energy, recently tackled from a physical point of view by Atkinson and Houtermans, Landau is mentioning "a beautiful idea of Prof. Niels Bohr's" according to which one could be able to believe that "the stellar radiation is due simply to a violation of the law of energy, which law, as Bohr has first pointed out, is no longer valid in the relativistic quantum theory, when the laws of quantum mechanics break down"[27]. At that time there was the big problem of apparent non conservation of energy in β-decay. Bohr speculated on the idea that perhaps energy conservation is not strictly valid in microscopic processes related to such nuclear transformations and that this might *also* even explain mechanisms related to the production of stellar energy. Such a problem, in turn, posed another completely unsolved mystery, together with nuclear structure, β-spectra and, last but not least, the famous Klein paradox, according to which an electron could not be confined within nuclei, a problem much debated since 1929. All this led Bohr to conclude that, "As soon as we inquire [...] into the constitution of even the simplest nuclei the present formulation of quantum mechanics fails completely". As emphasized by Pais, in anticipating "such drastic revisions of physics",

[27] In this regard, see Gamow's interview recalling Bohr's unpublished theory (Gamow 1968), Bohr's manuscript in his *Collected Works* (Bohr 1986, Vol. 9, p. 88) and (Gorelik 2005, pp. 63–82).

Bohr was looking for a comprehensive point of view that would all at once explain these four puzzles (Pais 1993, p. 367). Pauli, who definitely disagreed with Bohr, was reflecting on the possibility that there would be agreement with experiments if a new neutral particle took part in the beta-disintegration process carrying away the excess of energy and angular momentum. To Bohr's proposal about energy in stars Pauli thus answered: "let the stars radiate in peace!"[28] Rutherford, on his side, decided to wait and see before expressing an opinion, feeling that "there are more things in Heaven and Earth than are dreamt of in our philosophy", as he wrote to Bohr in November 1929[29].

In following Bohr, Landau thought he was killing two birds with one stone: not only was he avoiding catastrophic collapse to a point invoking non-conservation of energy, but was also obtaining a source of stellar energy. In January 1931, Landau, then in Zurich, had written with Peierls an article where they had already based on Bohr's idea in arguing that Pauli principle and thus ordinary quantum theory, did not apply in the nucleus, where special relativistic effects become relevant (Landau and Peierls 1931)[30].

Landau expected that the breakdown of quantum mechanics would occur "when the density of matter becomes so great that atomic nuclei come in close contact, forming *one gigantic nucleus* [emphasis added]". Landau did not specify what particles were involved, even if he must have clearly referred to nuclei as built out of protons and electrons, as they were still generally considered at that time. We have here a definite transition from Fowler's dense matter of a white dwarf, described as "analogous to a giant molecule" to a core of highly condensed matter forming "a single giant nucleus" surrounded by matter in ordinary state within the central region of the star. In the end, Landau supported Milne's idea about the central region of the star consisting of a core of highly condensed matter. However, Milne's theoretical 'collapsed configuration' was transformed in Landau's hands in a full-fledged physical system on which physicists could theorize. The price to be paid was to reject the possibility that stars' evolution might depend on their mass.

To summarize: in this rather short note, Landau is pursuing very ambitious aims: finding conditions for the equilibrium of a star, establishing the existence of a limit mass, finding a source of energy for stellar radiation and trying to develop a theory of stellar structure. Analyzing it in hindsight, many critics and comments could be put forward, that should be discussed within the state of physics at the time.

[28] Pauli to Bohr 17 July 1929, reprinted in (Bohr 1986, Vol. 6, p. 447).

[29] It is well known how Fermi took Pauli's idea so seriously, to incorporate 'Pauli's neutron', in the meantime renamed neutrino, in his ground-breaking theory of beta-decay, in which a new interaction was introduced using the language of quantum field theory. Based on the proton-neutron model of the nucleus, the mechanism of particle creation – the electron-neutrino pair – solved the problem both of the pathological 'nuclear electrons' and of the missing energy in the decay process. But immediate reactions were not exactly enthusiastic and only gradually the theory was generally accepted.

[30] However, according to Gamow, later it was shown by Landau himself that "the rejection of the conservation law for energy will be connected with very serious difficulties in the general gravitational theory, according to which the mass present inside a certain closed surface is entirely defined by the gravitational field on this surface" (Gamow 1934, p. 747). Gorelik has mentioned that at that time Bronstein realized the need for 'a relativistic quantum theory + the theory of gravitation in astrophysics' explaining it in a very simple way: "If the sun were compressed to nuclear density, its radius would be comparable with the gravitational radius" (Gorelik 2005, p. 1042). By that time cosmology was becoming for Bronstein the real great challenge: "a solution to the cosmological problem requires first to create a unified theory of electromagnetism, gravity, and quanta". Such a brilliant mind became one of the many victims of Stalin's Great Purges and was executed in 1938.

It is however to be remarked that, even being aware that the gravitational collapse was a consequence of his calculations, Landau rejected this possibility, heavily contributing with his influence to block acceptance of this catastrophic phenomenon. Without any doubt, this article was appealing to physicists, because it spoke their language, and for this reason it was widely cited during the years and opened the way to fruitful theoretical developments for reasoning on superdense matter in stars. The first clue to a fundamental difference in the evolution and final stages of low and high mass stars had been provided, but at the same time it had become clear that the analysis has to be shifted to the still basically unknown realm of nuclear matter.

8 Interlude: dense matter and the early universe. Georges Lemaître and the primeval super-atom

If a dense plasma of nuclei and electrons could exist within white dwarfs, "like a gigantic molecule in its lowest quantum state", forming a "Star-Atom", according to Eddington's colourful expression (Eddington 1927, p. 127), a super-compact atomic nucleus having a weight equal to the entire mass of the universe could well be at the origin of the whole universe itself, according to a proposal put forward by the Belgian physicists and cosmologist Georges Lemaître (Lemaître 1931a). During his university studies Lemaître had already tackled the general theory of relativity, and for this reason he decided to use a grant he had received in the summer of 1923 to go to Cambridge and study under Eddington, whose influential personality as a scientist and especially as an expert in relativity, inspired him to address his research interests to what appeared to him as a most fascinating field. During his later stay in the United States, in the period 1924-1925, Lemaître prepared for a Ph.D in astronomy at MIT and being attached to Harvard Observatory he was also introduced to the latest developments in astronomy and in particular experienced the impact of Hubble's observations of the early 1920s according to which the spiral nebulae are galaxies outside the Milky Way. Being convinced of the relevance of this new perspective and of the redshift-distance relation for relativistic cosmology, Lemaître visited both Vesto Slipher at the Lowell Observatory in Arizona, who had been the first to discover in 1917 that most spiral galaxies have considerable redshifts, and Hubble himself at Mount Wilson Observatory (Kragh 2013).

Unaware of Alexander Friedmann's work of 1922 (Friedmann 1922), showing that Einstein's equations have dynamical solutions, Lemaître's formulated the same cosmological differential equations. He proposed a dynamical cosmological model in his "Un univers homogène de masse constante et de rayon variable rendant compte de la vitesse radiale des nébuleuses extra-galactiques" (Lemaître 1927). But his approach was quite different, because, contrarily to Friedmann, who did not compare the models with astronomical data, Lemaître addressed the cosmological implications of general relativity combining mathematical results with the physical reality, in particular, with astronomical observations of the recession of the nebulae, that he viewed as a "cosmical effect of the expanding universe" (Lemaître 1931, p. 489). "The expansion of the universe is a matter of astronomical facts interpreted by the theory of relativity" stressed Lemaître in October 1931, during a meeting of the British Association for the Advancement of Science dedicated to the evolution of the universe, to which de Sitter, Eddington, Millikan, and Milne participated. His 1927 theory went rather unnoticed and was 'rediscovered' around 1930 when Eddington and De Sitter contributed to make it widely known (Kragh 1996, ch. 2). At first, both Friedmann and Lemaître were ignored. Lemaître himself became aware of Friedmann's work when he attended the 1927 Solvay Conference, during discussions with Einstein. Einstein recognized the similarity between the two theories, and had no objection in this sense, but his conclusive comment was unfavorable: he considered it definitely "abominable" from the physical point of view (Deprit 1984, p. 370) (Kragh 1987, p. 125) (Eisenstaedt 1993, p. 8). But actually Lemaître, in telling Einstein about the recessional velocities of galaxies, had the impression that the latter was not really informed about astronomical facts.

Lemaître's physical cosmology, in connection with current views of dense matter in bulk subject to quantum laws, that most probably concurred to inspire him, led to a proposal which Lemaître presented in a short note to *Nature*. In 'The beginning of the world from the point of view of quantum theory' (Lemaître 1931a) Lemaître answered to Eddington's contribution 'The end of the world: from the standpoint of mathematical physics' (Eddington 1931) published on the same journal, where the latter had clearly stated that "the notion of a beginning of the present order of Nature is repugnant". Lemaître proposed instead that he was "inclined to think that the present state of quantum theory suggests a beginning of the world very different from the present order of Nature". Thermodynamic principles, he said, require that "(1) energy of constant total amount is distributed in distinct quanta" and that "(2) the number of distinct quanta is ever increasing. If we go back in the course of time we must find fewer and fewer quanta, until we find all the energy of the universe packed in a few *or even in a unique quantum* [emphasis added]".

If an atomic nucleus could be counted as a unique quantum, "the atomic number acting as a kind of quantum number," one could conceive the beginning of the universe in the form of a unique atom, the atomic weight of which is the total mass of the universe". This highly unstable universe-atom "would divide in smaller and smaller atoms by a kind of super-radioactive process". He thus believed that the primeval atom hypothesis provided a physical beginning of the universe and that its subsequent evolution was the result of a disintegration (Lemaître 1931c, pp. 113-114): "In the atomic realm, we know a spontaneous transformation that can give us some idea of the direction of the natural evolution; it is the transformation of radioactive bodies [...] an uranium atom is eventually transforming into a lead atom and seven or eight helium atoms. This is a transformation from a more condensed to a less condensed [...] The natural tendency of matter to break in more and more numerous particles, which shows itself in so striking a way in the radioactive transformations, can be observed also in the grains of light or photons that form the different forms of radiation". The 'super-radioactive' processes he mentioned suggest a kind of matter very similar to nuclear matter, consisting of electrons and especially of alpha particles, which were considered a sort of building blocks of the nucleus, because of their recognized stability as entities deriving from the decay of radioactive elements, in particular very heavy elements such as uranium and thorium, whose half-lives were of the order of billion years. Some remnant of this process, recalled Lemaître following Jeans's idea, might still be fostering the heat of the stars. All this found an experimental base on what can be considered the nuclear physics of the time, which was on the verge of entering its modern era with the detection of the neutron, but whose knowledge still derived mainly from the study of radioactive decays, which on the other hand were connected to the formation of new chemical elements, that had been studied since the early years of radiochemistry.

Lemaître's primeval matter appears to be quite similar to the stuff of which dense white dwarfs were supposed to be made. But actually, he did not specify the nature of the 'primeval atom', a term that is probably to be interpreted as something similar to the basic primordial entity very common in ancient cosmogonies. A hint about the nature of the primordial super-atom is provided by his contribution to the mentioned discussion of October 1931 at the meeting of the British Association for the Advancement of Science, a longer contribution in which he fully outlined his views about the physical universe – "The expansion of the universe is a matter of astronomical facts interpreted by the theory of relativity" – and its origin from the disintegration of the primeval atom: "We want a 'fireworks' theory of evolution. The last two thousand million years are slow evolution: they are ashes and smoke of bright but very rapid fireworks". He suggested that big stars were remnants of the successive splittings of the primeval atom and that, with their fireworks of radiation, they were the source of cosmic rays of high energy (Lemaître 1931b, p. 704-705). The key of the problem, according to Lemaître, was afforded by the discovery of cosmic rays: "the energy of cosmic rays is comparable in amount to the whole energy of matter [...] If the cosmic rays originated chiefly before the actual expansion of space, their original energy was even bigger [...] The only energy we know which is comparable to the energy of the cosmic rays is the matter of the stars. *Therefore it seems that the cosmic rays must have originated from the stars* [emphasis added]". Inspired by Jeans' ideas admitting the possible existence of atoms of considerably higher atomic weight than the known end decay products of radioactive decays of the heavies atomic elements, Lemaître stated that "Cosmogony is atomic physics on a large scale – large scale of space and time – why not large scale of atomic weight? Radioactive disintegration is a physical fact, cosmic rays are like the rays from radium. Have they not escaped from a big scale super-radioactive disintegration, the disintegration of an atomic star, the disintegration of an atom of weight comparable to the weight of a star". Cosmic rays would be "glimpses of the primeval fireworks of the formation of a star from an atom, coming to us after their long journey through free space".

He immediately suggested that "a possible test of the theory is that, if I am right, cosmic rays cannot be formed uniquely of photons, but must contain, like the radioactive rays, fast beta rays and alpha particles, and even new rays of greater masses and charges"[31].

Whether this was "wild imagination or physical hypothesis", it could not be said. In order to solve the problem two things were needed, according to Lemaître: "First, a theory of nuclear structure sufficient to be applied to atoms of extreme weights [...] The second thing we want is a better knowledge of the nature of the cosmic rays".

What is relevant in our context is that Lemaître's ambitious theory was relating the mathematical universe of General Relativity to an evolutionary physical universe whose nature as a physical system was being discovered by astronomers: "A really complete cosmogony should explain both atoms and suns" (Lemaître 1931c, p. 113). But he also showed how the theory of the expansion of the universe could be adapted to the idea of a primeval atom through three different phases: a first period of rapid expansion during which the universe-atom breaks in star-atoms, a period of slowdown, followed by a third phase of accelerated expansion, that we are living now, which is responsible of the separation of stars in extra-galactic nebulae (Lemaître 1931c, p. 119).

[31] According to Millikan's opinion cosmic rays were "the birth cries of the elements", high-energy photons arising from the building-up of elements in the depths of space. His theory had recently been challenged by the Bothe and Kolhörster's experiment published in 1929 (Bothe and Kolhörster 1929), showing that cosmic rays were charged particles and not 'ultra-gamma rays'. But it cannot be excluded that the 'cosmic birth' context summoned by Millikan's theory, played someway a role in Lemaître's reflections leading to the primeval atom theory. In any case, during his stay at MIT, Lemaître collaborated with the Mexican physicist Manuel Vallarta in complicated calculations of the energies and trajectories of charged particles in the Earth's magnetic field, making use of MIT's differential-analyzer computer developed by Vannevar Bush. They concluded that both Arthur Compton's data deriving from his world campaign, that had verified the existence of the latitude effect ("showing that the cosmic rays contain charged particles") and their own computer calculations, were providing "some experimental support to the theory of super-radioactive origin of the cosmic radiation" (Lemaître and Vallarta 1933, p. 91).

In his ambition to explain the Universe at a macroscopic and microscopic level as a physical system in continuous evolution, Lemaître put quantum theory and thermodynamics in connection with a state of superdense concentration of matter, having such universal character to give origin to all the observed distribution of matter in the universe: all the atomic nuclei were produced by disintegrations of the primeval quantum. Moreover, for the second time, after Eddington's observation that general relativity must be connected to the observed spectra of dense stars like white dwarfs, Einstein's theory was connected to a primeval dense concentration of matter giving origin to the whole physical universe.

It is difficult to assess the overall impact of Lemaître's speculations related to his physical cosmogony[32]. It is quite clear that he influenced further bold speculations put forward by physicists like Fritz Zwicky and especially Gamow, who had been Friedmann's student and had a knowledge of general relativity since the beginning of his research activity.

His "wild imagination" was offering such cosmic fireworks to physicists who had the same bold attitude and whose minds resonated on Lemaître's words (Lemaître 1931a, p. 706): "Our world is now understood to be a world where something really happens; the whole story of the world need not have been written down in the first quantum like a song on the disc of a phonograph. The whole matter of the world must have been present at the beginning, but the story it has to tell may be written step by step".

9 Sterne 1933: neutronization of superdense matter in stars

When the neutron officially became a new constituent of the nucleus – even if it was not immediately clear whether it was or not a bound state of proton and electron – it opened a new era in nuclear physics and in particular in its application to the astrophysical stage. The long-standing problem of the origin of elements and of stellar energy could be discussed on a new base. As it had happened in the case of the new statistics in connection with metals and white dwarfs, now dense stellar matter became a testing ground for nuclear reactions.

It appears that the first to propose a systematic discussion on the equation of state of nuclear matter, and to apply it to stars interiors, was Theodor E. Sterne, who received his PhD from Cambridge University in the summer of 1931 with Fowler as his supervisor (like in Stoner's case, Fowler is again acting behind the scene...).

Sterne started his investigations on what was considered "One of the most important problems requiring solution" at that time, that is the production of energy in stars (Sterne 1933a). It was generally agreed that the principal, if not almost the entire, source of this energy must be subatomic. In 1932, disintegrations produced by artificially produced fast protons had been observed at Cavendish Laboratory by Cockcroft and Walton with large production of energy, as well as transmutations produced by bombardment of fast alpha particles, resulting in the emission of neutrons capable in turn of further transmutations in striking other nuclei. The possibility of induced transmutations had thus been established beyond any reasonable doubt by strong experimental evidence, and considerable absorption or liberation of subatomic energy were expected in most cases. These energies, said Sterne, must be intimately related to the abundances of the elements in the stellar matter during the changes between the different states. In March 1933, he announced his program (Sterne 1933a, p. 585): "It is possible to consider by statistical mechanics an assembly containing

[32] Attention has been given by Kragh to responses to Lemaître's theory of the expanding universe (see (Kragh 2013) and references therein).

radiation, atomic nuclei, electrons, and neutrons; when all possible transmutations of the nuclei occur without the 'annihilation' of any ultimate particles. One can calculate the abundances of the nuclei of the various sorts in such an assembly, when it is in equilibrium, in terms of the atomic masses and packing fractions"[33].

In three further papers appearing in cascade in MNRAS (Sterne 1933b) (Sterne 1933c) (Sterne 1933d) Sterne discussed the formation of the chemical elements by nuclear reactions in stars and the liberation of energy by transmutations, apparently being the first systematic investigation in this sense[34]. In (Sterne 1933b, p. 748), he investigated the gradual contraction of a star, with equilibrium composition gradually shifting as the density and temperature increased. He pointed out that, as determined by Chadwick, neutrons had packing fractions which are considerably greater than the packing fractions of other kinds of nuclei. Applying the Darwin-Fowler method to the statistical equilibrium among nuclei, he arrived at the conclusion that "At sufficient enormous densities [greater than approximately 2.3×10^{10} g/cm^3 when $T \ll 6 \times 10^7 \rho^{1/3}$] [...] the assembly at low temperatures should contain a preponderance of neutrons [...] At these high densities, matter at low temperatures would be literally squeezed together into the form of neutrons". (Sterne 1933b, p. 750)[35].

He concluded the article expressing the hope that "the statistical theory here developed may prove to be of assistance to astrophysicists"[36].

In parallel with Sterne's theoretical work in which it was clarified that compression of cold matter to high densities would induce neutronization, the role of neutrons in the structure of stars was widely discussed in a PhD dissertation written under Max Born in Göttingen by Siegfried Flügge (Flügge 1933). While Sterne was more relying on the idea that after all a neutron was a bound state of a proton and an electron, Flügge specified that as during $\beta-$decay processes a neutron is transmuted in a proton + an electron, one could imagine that an evaluation of the number of neutrons in stars could be done through a "thermodynamical equation according to the Synthesis Proton+Electron = Neutron + Energy" (Flügge 1933, p. 278). He also examined, "as a curiosity" what would be the characteristic of a star consisting only of neutrons ("*ein Stern, der nur aus Neutronen bestünde*") and speculated how neutron capture by heavy nuclei could explain the production of stellar energy (Flügge 1933, p. 282).

Neutrons were beginning to become the great protagonists of nuclear processes taking place in stars. It is thus not surprising that speculations on the existence of

[33] At that time physicists still discussed whether the neutron was a real elementary particle and whether the positron, that Sterne included in his discussion, was identical with Dirac's 'holes'. In this regard Sterne, mentioned Carl D. Anderson's observation of the positive electron at Caltech, as well as cloud-chamber experiments performed at Cavendish Laboratory by Patrick Blackett and Giuseppe Occhialini, who had observed the phenomenon of electron-positron pair production producing a strong support to Dirac's theory.

[34] By that time Sterne was at Jefferson Physical Lab Cambridge, Mass. He thanked Cecilia Payne and Ralph Fowler, who communicated the papers to the Royal Astronomical Society.

[35] And indeed, in a short note on *Nature* (Sterne 1933a) he had presented his preliminary investigations on the equilibrium property of an assembly containing radiation, atomic nuclei, electrons, and neutrons based on the "hypothesis that nuclei (*and neutrons*) are made of electrons and protons [emphasis added]". In (Sterne 1933b), instead, he also considered the possibility that the neutron could be "an ultimate particle".

[36] Sterne's pioneering article was cited by Gamow in 1939 (Gamow 1939a), at a time when nuclear astrophysics had already developed into a research field attracting physicists with a competence in theoretical nuclear physics. Gamow acknowledged that: "It was first indicated by Sterne that, at very high densities and not-too-high temperatures, the formation of a large number of neutrons must take place because the free electrons are, so to speak, squeezed into the nuclei by the high pressures". Gamow is also suggesting to look at Hund's review article of 1936 (Hund 1936), which will be discussed later.

exotic stars consisting only of neutrons, mentioned by Flügge as a curiosity, were quickly incorporated in a theory on the most catastrophic cosmic event known at the time: the explosion of a star.

10 A not so lonely sailor: Fritz Zwicky

As outlined in the previous sections, during the 1920s many physicists addressed astrophysical problems, exploring the properties of very dense stars in order to derive basic properties of matter in conditions that could not be obtained in any terrestrial laboratory. The growing relevance of the problem of stellar energy, and the related difficulties faced by physicists in their attempt to account for the actual production of such energy, went in parallel with the shifting of interest towards the nuclear realm during the 1930s, especially after the strong impact deriving from the confirmed existence of the neutron that opened the way to brand new theoretical and experimental investigations.

Theories about the stellar interiors included the new particle in discussions about the structure, equilibrium and generation of energy in stars. Papers on the phenomenon of neutronization of matter in stars with increasing density certainly did not escape the attention of Fritz Zwicky, a Swiss theoretical physicist working at the California Institute of Technology since the 1920s[37]. He was familiar with quantum theory, as well as with dense matter in metals and crystals, a field in which he was still working during the early 1930s.

At the same time, the Caltech campus is near the Mount Wilson Observatory, which had the world's largest telescope, and where Edwin Hubble was working since the end of the 1910s. In 1929, Zwicky was intrigued by Hubble's results (Hubble 1929) showing a roughly linear correlation between the apparent velocity of recession and the distance of galaxies (Zwicky 1929) and his interest in astrophysics grew with the arrival of the German astronomer Walter Baade from Hamburg in 1931. Baade was studying novae and together they came to the conclusion that the population of novae consists of two types: the ordinary novae and the 'supernovae', which are very rare but much more energetic. In December 1933, during the annual meeting of the American Physical Society at Stanford, they proposed that "In the *supernova* process *mass in bulk is annihilated*. In addition the hypothesis suggests itself that *cosmic rays are produced by supernovae*". Basing on the assumption that "in every nebula one supernova occurs every thousand years" they accordingly evaluated the expected intensity of cosmic rays, comparing it with Millikan and Regener's observed flux. They concluded the abstract with a bold proposal: "With all reserve we advance the view that supernovae represent the transitions from ordinary stars into *neutron stars* which in their final stages consist of extremely closely packed neutrons" (Baade and Zwicky 1933)[38].

Such a star, they explained in a more detailed article, "may possess a very small radius and an extremely high density. As neutrons can be packed much more closely than ordinary nuclei and electrons, the 'gravitational packing' energy in a *cold* neutron star may become very large, and, under certain circumstances, may far exceed the

[37] Born in Bulgaria in 1898, Zwicky grew up in Switzerland, and studied in Zurich. He studied solid-state physics and worked in crystallography research before moving to California on an International Education Board post-doctoral fellowship in 1925.

[38] According to a review article by Zwicky (Zwicky 1940, p. 85), he and Baade introduced the term supernovae in seminars and an astrophysics course at Caltech in 1931 then used it publicly in 1933 during the just mentioned meeting of the American Physical Society held at Mount Wilson Observatory.

ordinary nuclear packing fractions. A neutron star would therefore represent the most stable configuration of matter as such" (Baade and Zwicky 1934a, p. 263). They were fully aware that their suggestion carried with it "grave implications regarding the ordinary views about the constitution of stars" and therefore would require "further careful studies" (Baade and Zwicky 1934b, p. 77).

Speculations on planetary nebulae, as originating in novae, with their gaseous expanding shells as the remains of past outbursts, even suggesting an origin in outbursts of several stars, provided a well defined scenario – on a large space-time scale – of a phenomenon suggesting a process in which matter expanded after an explosion. Already in 1923, for example, J.H. Reynolds concluded an article on gaseous planetary nebulae with the following words: "The old idea that the gaseous nebulae were the primitive forms of matter from which stars were evolved must, it seems, be given up for the exactly contrary hypothesis that they had their origin in stellar outbursts, where matter passed from complex to simpler forms by atomic disintegration under the stress of extreme temperature development" (Reynolds 1923). As already mentioned, the idea of stellar explosions associated with collapse to a superdense configuration had been already suggested in connection with discussions on white dwarfs. In 1926, in comparing the nuclei of planetary nebulae to white dwarfs, Donald H. Menzel said in a section entitled 'The physical state of the nuclear stars (white dwarfs)': "Novae arise from giants and dwarfs, that is they are outbursts from dwarf stages of stars, that are probably experiencing these outbursts many hundred times during their history" (Menzel 1926, p. 307) However, the first very explicit description of the idea of stellar explosions associated with collapse to a dense configuration can be found in Milne's talk at the meeting of the British association of October 1931 (Discussion on the Evolution of the Universe) (Milne 1931b, p. 716). Milne had recalled that during the contraction a star is losing gravitational energy, which is set free as heat and light, this shrinking must thus be "the actual origin of the brightening [...] Since the rate of brightening is very rapid, we infer that the process of shrinkage is very rapid – in fact cataclysmic. The process of shrinkage is a veritable collapse. In a nova outburst the star is seen to be collapsing on itself; and the suddenness of the collapse, and the resulting enormous amount of gravitational energy that must be got rid of in the short time available, conspire to produce the huge brightening of the star as observed. This sudden liberation of energy produces enormously increased radiation, which in turn expels the outer layers of gas. Such is the probable explanation of the origin of novae, or 'new stars"'. Milne also specified that "the mass of the star, after the outbursts, is practically the same as before, yet it occupies a much smaller volume, hence its mean density must be much larger than before [...] The gases expelled from the star during the outburst are chips from the old block; but the star itself does not remain an old block; it becomes very much of a new block – a very dense block". Of course Milne immediately mentioned other dense stars, known as white dwarfs, and the nuclei of the planetary nebulae, both having probably undergone the process of collapse: "It is reasonable to assume [...] that every white dwarf has been at one time a nova".

These speculations provided the astrophysical background, while the novelties derived by the new status of nuclear matter inspired Zwicky's further conjectures which resulted in an attempt to fill the collapse idea discussed by Milne and others with a more physical content. It is rather plausible that this part of their proposal came from Zwicky himself. His experience with dense matter in crystals and metals most probably led him in a most natural way to reason on super dense neutronic matter in stars. The close packing of neutrons within dense stellar cores could explain the energy release in supernovae which he estimated to be equivalent to the annihilation of the order of several tenths of a solar mass. However, he could only guess at the scenario for forming neutron stars; all the physical mechanisms of the implosion, including the behaviour of matter in the core during the process and the actual emission of energy, remained completely unknown. What they estimated was the evaluation of energy involved in supernova explosions as if produced by particles or photons that in turn was compared to the observations of the intensity of cosmic rays made by Regener, and by Millikan and his collaborators. Lemaître's hypothesis of cosmic rays "as remnants of some super radioactive process which took place a long time ago" was mentioned by Zwicky exactly at that time.

What has always been duly termed a 'prescient' idea, was thus not coming out of the blue. It cannot be excluded that many of Zwicky's reflections about neutrons were inspired by the work of his colleague Langer, who was especially interested in the properties of neutrons, and also in the origin of cosmic rays, topics that he discussed at the same Stanford meeting of December 1933 in three different talks. The guiding concept in Baade and Zwicky's proposal appears in fact to be the problem of the origin of cosmic rays, seen as a mysterious radiation whose 'cosmic' nature was still attracting the main attention, notably at Caltech, because of Millikan's presence. Millikan, the director of the Norman Bridge Laboratory of Physics at Caltech, had since the 1920s advocated that cosmic rays were high-energy gamma rays produced during the birth of elements in the universe, and had undertaken a major study of the radiation. Zwicky was thus definitely familiar with the problem. That same 1932, a worldwide measurement campaign investigating a possible dependence of the rate on magnetic latitude was led by Arthur H. Compton and established beyond any doubt that a part of the primary radiation consists of charged particles. Moreover, parallel experiments also proved the existence of the east-west effect, hypothesized in 1930 by the Italian cosmic ray physicist Bruno Rossi. According to his prediction there should be an azimuthal asymmetry in the intensity of cosmic rays that would depend on the sign of the charge of the primary particles. Both the charged nature of cosmic rays (also verified by the latitude effect) *and* the sign of the charge, were determined by such experiments (Bonolis 2014). Research on cosmic rays was already becoming strongly related to the emerging field of elementary particle physics, and the problem of their origin was gradually less investigated, at least up to the 1940s, when it was possible to establish the nature of the primary radiation. At that time the problem of their origin again became a hot subject, also in connection with other astrophysical developments.

Zwicky, Baade, and all other astronomers in Pasadena were following Hubble's work and had witnessed Lemaître's lectures on the expanding universe and the primeval-atom hypothesis during his journey in the U.S. Already in early September 1932, during the Fourth General Assembly of the International Astronomical Union, which took place at Cambridge, Massachusetts. There, Eddington's public lecture on the expanding universe was a climax event and Lemaître's "fireworks theory of the beginning of things" was widely discussed (Deprit 1984, p. 373–375). Lemaître remained for some time working with Vallarta on his hypothesis for the origin of cosmic rays and both participated to the meeting of the American Physical Society that same November, where Arthur Compton presented the preliminary results of his survey of the intensity of cosmic radiation at a large number of stations scattered all over the world, widely confirming previous observations and ruling out the hypothesis that the radiation consisted of photons alone and that it was made up at least partly of charged particles. This question, according to Lemaître, was very likely bound up with general cosmogonical problems, even if the question as to their origin remained unanswered. Moreover, in November Lemaître was invited by Percy H. Robertson to give a seminar on his cosmology in Princeton, obviously attended by Einstein, and in December he moved to Caltech, where he also met Hubble. His seminars in which he discussed his astounding theories on the expanding universe and on the cosmic rays as the remains of the primordial universe, were widely spread by a long article on the *New York Times* Magazine appearing in February1933. By that time Zwicky had already begun his speculations on the origin of cosmic rays, and the red-shift phenomenon of far away galaxies.

In January 1933, Zwicky investigated the problem of the origin in an article entitled 'How far do cosmic rays travel?' in which he tried "to establish a relation between them and the red shift of extragalactic-nebulae" examining two entirely different hypotheses: the one suggesting that cosmic rays must be of local origin (upper atmosphere, planetary system, etc.) and the second one, especially advanced by Robert A. Millikan, that they were produced throughout interstellar or intergalactic spaces (Zwicky 1933a). Zwicky had in fact concluded from the results of observations on the red-shift of extragalactic nebulae, that the amount of dark matter in the Universe must be grater than that of luminous matter, and he thus tried to establish a connection between these two phenomena (Zwicky 1933b)[39]. The connection between the origin of cosmic rays and the redshift phenomenon related to the expanding universe in Zwicky's research, is strongly suggesting that Lemaître's ideas on the expansion of the universe and especially about the primeval atom and its explosive nuclear processes provided a strong conceptual platform as a starting point for reflections on relativistic cosmology and in particular on the problem of cosmic rays, eventually leading to the theory of supernovas. Baade and Zwicky mentioned the possibility that either the cosmic rays "originate in intergalactic space or that they are survivors from a time when physical conditions in the universe were entirely different from what they are now (Lemaître)", but they considered both hypotheses to be very unsatisfactory and for this reason they made "an entirely new proposal" removing some of the major difficulties concerning the origin of cosmic rays (Baade and Zwicky 1934a, p. 260). In 1931 Regener, too, had speculated on cosmic rays as a remain of an original explosion in connection with Einstein's closed universe (Regener 1931).

Lemaître's theory of a dense primeval state whose "explosive" expansion could gave origin also to cosmic rays, in connection with the growing role of neutrons in astrophysical realm, might well explain why a star consisting only of neutrons, that Flügge had considered a mere 'curiosity', became a basic assumption in Baade and Zwicky's theory of neutron stars as remnants of supernova explosions, that in turn became the source for high energy cosmic rays. Milne himself had suggested (Milne 1930a) that novae resulted from the collapse of stellar cores, then becoming white dwarfs, that is very dense stars. In turn, the collapse to a superdense configuration had led to Sterne's and Flügge's suggestions that compressed matter in stars would result in neutronization. All this was part of Zwicky's conscious and unconscious imagination.

Baade and Zwicky did not mention Landau and Chandrasekhar, or any other work about the maximum mass of white dwarfs. Any connection would require a far deeper knowledge of nuclear theory and nuclear reactions. In any case, no relationship was established at the moment between these two compact objects: white dwarfs and

[39] Zwicky measured the velocity dispersion of the galaxies in the Coma cluster and found that there must be about 100 times more dark, or hidden, matter as compared with visible matter in the cluster. In this article Zwicky discussed redshift in connection with cosmological theories and explicitly mentioned: "Another important proposal was made by Friedmann, Tolman, Lemaître and Eddington, whose work shows that according to the theory of relativity a static space is dynamically unstable and therefore tends to contract or expand. This result was interpreted by him to imply that the redshift would correspond to a factual expansion of space". In his editorial note to the English translation of Zwicky's paper (Zwicky 2009), Jürgen Ehlers suggests that Zwicky did not specify which of the four names he meant, but that in reality this proposal was first made by Lemaître (Ehlers 2009). Actually Tolman himself became really involved in cosmology around 1930-1931, in connection with Hubble's results about the red-shifts of the extragalactic nebulae being proportional to their distances (Hubble 1929) and when Lemaître's work became widely known also in the United States.

the hypothetical neutron stars. However, as astronomers, they had recognized the existence of a special class of stars, the supernovae, that during several weeks radiate as much energy as a whole galaxy of stars. This suggested that observation of these unique objects would furnish valuable information on fundamental problems such as the generation of energy in stars, the evolution of stars and stellar systems, the origin and characteristics of cosmic rays. Baade and Zwicky thus felt strongly motivated to start a systematic search of supernovae, that promised to be particular significant.

11 Chandrasekhar and the final fate of a white dwarf

Towards the end of 1931, Chandra began to feel uneasy. His results on model stellar photospheres presented at the January 1932 meeting of the Royal Astronomical Society were much appreciated by both Milne and Eddington, who were following his work with great attention, apparently because they hoped that new results would confirm their own theories. However, he was still a PhD student, and in trying to measure up to such established and incredibly influential astrophysicists such as Eddington and Milne he was in reality an outsider within this small scientific community. Moreover, he felt that: “Physics, was at the center, not astrophysics” (Wali 1990, p. 98). Later Chandra recalled that Dirac told him (Chandrasekhar 1977): “Well, if I were you, I would be interested in relativity, rather than astrophysics”. Chandra then asked him: “One time you did write a paper on astrophysics…” and Dirac answered: “Oh, that was before quantum mechanics”. All this made Chandra feel afraid that astrophysics was considered inferior by most physicists. He felt alone and even thought of entering the field of theoretical physics. He greatly admired Dirac, with whom he had developed a friendly relationship, and told him how unhappy he was in Cambridge, so that Dirac suggested him to spend some time at Niels Bohr’s Institute in Copenhagen, where Chandra went during his final year, before the end of his Government scholarship. He stayed there from August 1932 to May 1933, finding a friendly, informal and international atmosphere. During this period he established a strong relationship in particular with Léon Rosenfeld, who was much interested in Chandra’s work, and at the same time could discuss common research issues with Bengt Strömgren, who very often visited Bohr’s institute and had a strong physical background. Both Chandra and Strömgren represented, even if in different perspectives, a new figure of astrophysicist, strongly familiar with the physicists’ community also because of university education.

At the time Bohr told Chandra that “Well, I’ve always been interested in astrophysics, but the first question I should like to know about the sun is: where does the energy come from? And since I can’t answer that question, I do not think a rational theory of the stellar structure is possible”. In recalling this conversation, Chandra added (Chandrasekhar 1977): “Well, great as Bohr is, that remark of Bohr’s is invalid. Later on, if one found the right nuclear reactions, it was because one had found out earlier the right temperatures and physical conditions by their ingenuity”. Here Chandra is certainly referring to what Bethe himself recognized about his theory on stellar energy and how it was inspired by the insight coming from Strömgren’s work, that will be explained later.

In a report written by Bohr in October 1933, concerning the work of Chandrasekhar during his stay in Copenhagen from August 1932 to May 1933, he declared: “I am glad to take this opportunity for expressing my high appreciation of the scientific work which Mr. Chandrasekhar has performed in the course of his studies in this institute since September 1st 1932. During this time he has been successfully engaged in the theoretical treatment of a number of important astrophysical problems, and as well in the choice of these problems as in the methods used for their solution he has shown great ingenuity and ability. In my opinion he may be regarded as one of the most competent among the younger astrophysicists, as to whose future scientific activity great expectations are justified"[40].

By the end of 1932 Chandra had published four papers on rotating self-gravitating polytropes, which became his Ph.D. thesis. In (Chandrasekhar 1932) he considered stars whose mass exceeds the critical mass and concluded that for these stars "*the perfect gas equation does not break down, however high the density may become, and the matter does not become degenerate. An appeal to Fermi-Dirac statistics to avoid the central singularity cannot be made*". The only way out of the singularity, added Chandra, "is to assume that there exists a maximum density ρ_{max} which matter is capable of". However, at the very end of the article he wrote: "We may conclude that great progress in the analysis of stellar structure is not possible before we can answer the following fundamental question: *Given an enclosure containing electrons and atomic nuclei, (total charge zero) what happens if we go on compressing the material indefinitely?*"[41].

In October 1933 he was elected to a Trinity Fellowship, "one of the most gratifying events that can happen to one", as remarked by Milne in a letter he hastened to send him as soon as the news was announced (Wali 1990, p. 109). The Fellowship put him in contact with the Cambridge scientific society and he also got invitations from abroad. In particular from Boris P. Gerasimovič, who had just become the director of the Pulkovo Observatory, near Saint Petersburg[42].

They had been in contact for some time and Chandra was eager to see Russia. During this four-week trip, he met Landau and Viktor A. Ambartsumyan and gave two lectures at Pulkovo, one of which about his work on white dwarfs and the limiting mass. The brilliant Ambartsumyan, who was organizing the Soviet Union's first department of astrophysics, fully grasped the significance of Chandra's work on dwarf stars and suggested that he investigate the problem in greater detail working out the exact, complete theory of white dwarfs, (i.e., by direct radial integration of the equations, using the complete pressure-density relation), devoid of some simplifying assumptions, and to examine the entire range of densities, within the framework of relativistic quantum statistics and the improved knowledge of stellar interiors. Chandrasekhar felt again encouraged to tackle such immense problem.

Since the beginning, Chandra's work had actually been related to fundamental issues involved in the Milne-Eddington controversy on the nature of the boundary conditions one should use in determining the equilibrium configurations of stars. The existence of a limiting mass contradicted Milne's idea that *all* stars had a degenerate core surrounded by outer layers of stellar material obeying the perfect gas equation of state. During the period 1932-1934, Chandra had been occupied with finishing his degree, moreover there had not been so much impact from his work. But now, Ambartsumyan's suggestion to explore again the problem represented a new challenge that might also settle the controversy. Eddington, who was personally interested in this new work, hoping that his ideas would prevail, even lent him a Brunsviga hand calculator, that was a fundamental tool for solving numerically the differential equations related to the equations of hydrostatic equilibrium for each white-dwarf star of his sample.

[40] Archives for the History of Quantum Physics, Bd. AHQP/BSC 19: Niels Bohr. Scientific correspondence, 1930–1945.

[41] In this article, written in Copenhagen, Chandra cited (Landau 1932) and thanked Strömgren for advice. The latter most probably attracted his attention on Landau's paper.

[42] Later Stalinist purges in 1936-1937 devastated Russian astronomy and destroyed Pulkovo as an active research institute and the effect on Russian astronomy was to be felt for a very long time (Eremeeva 1995).

By the end of 1934 Chandra had completed a detailed analysis on the problem of the limiting mass, distinguishing between dense matter obeying the equation $p \sim \rho^{5/3}$ and ultradense matter which obeys the equation $p \sim \rho^{4/3}$. He reached a conclusion that a limiting mass is obtained only for the ultradense case, which he stated in the following terms (Chandrasekhar 1934a, pp. 373-377): "*The life-history of a star of small mass must be essentially different from the life-history of a star of large mass. For a star of small mass the natural white-dwarf stage is an initial step towards complete extinction. A star of large mass cannot pass into the white-dwarf stage and one is left speculating on other possibilities* [emphasis added]"[43].

On January 1, 1935, Chandra completed the paper "The highly collapsed configurations of a stellar mass (Second paper)" (Chandrasekhar 1935a), a follow up of his (Chandrasekhar 1931c), where he is clearly showing that the existence of a limiting mass (that for a mean molecular weight per electron = 2 was 1.44 solar masses) meant that a white-dwarf state does not exist for stars that are more massive. This paper includes a figure (Fig. 2) exhibiting the mass-radius relation deduced on the basis of the exact equation of state allowing for the effects of special relativity of which equations $M = constant \times R^{-3}$ and $p = k_2(n_e)^{4/3}$ are the appropriate limiting forms, where k_2 is an atomic constant and n_e is the electron concentration. The effect of special relativity is to reduce the power of the pressure dependence on density from 5/3 to 4/3. This limiting form of the equation of state has a dramatic effect on the predicted mass-radius relation: the radius must tend to zero as a certain limiting mass is reached.

He remarked how one could notice clearly from these two curves "how marked the deviations from the limiting curves become even for quite small masses," and how the relativistic effects are quite significant even for small masses. "These completely collapsed configurations, continued Chandra, have a natural limit, and our exact treatment now shows how this limit is reached". He extended the discussion in a second paper dated January 4, and concluded that the developed methods and the results obtained "would have to be extended for more general stellar models before any very definite conclusions could be drawn" (Chandrasekhar 1935c).

Chandra gave an account of this work in the January 1935 meeting of the Royal Astronomical Society, of course showing Figure 2, a clear definitive demonstration of what might happen to a white dwarf exceeding Chandra's maximum mass. Eddington attacked him frontally (Eddington 1935a, p. 38): "Chandrasekhar shows that a star of mass greater than a certain limit remains a perfect gas and can never cool down. The star has to go on radiating and radiating and contracting and contracting until, I suppose, it gets down to a few kilometres radius when gravity becomes strong enough to hold the radiation and the star can at last find peace. Dr. Chandrasekhar had got this result before, but he has rubbed it in his latest paper; and, when discussing it with him, I felt driven to the conclusion that this was almost a *reductio ad absurdum* of the relativistic degeneracy formula. Various accidents may intervene to save the star, but I want more protection than that. I think that there should be a law of nature to prevent the star from behaving in this absurd way". Eddington recognized that Chandra had worked out correctly the astrophysical consequences of relativistic degeneracy, according to the current interpretation (Eddington 1935b, p. 195): "I do not think that any flaw can be found in the usual mathematical derivation of the formula. But its physical foundation does not inspire confidence, since it is a combination of relativistic mechanics with non relativistic quantum theory". In contending that the relativistic formula rested on a misconception ("It must at least rouse suspicion as to the soundness of its foundation"), Eddington examined this "unholy alliance" concluding that the 'relativistic' formula was "erroneous" and

[43] See also 'Stellar configurations with degenerate cores' (Chandrasekhar 1934b).

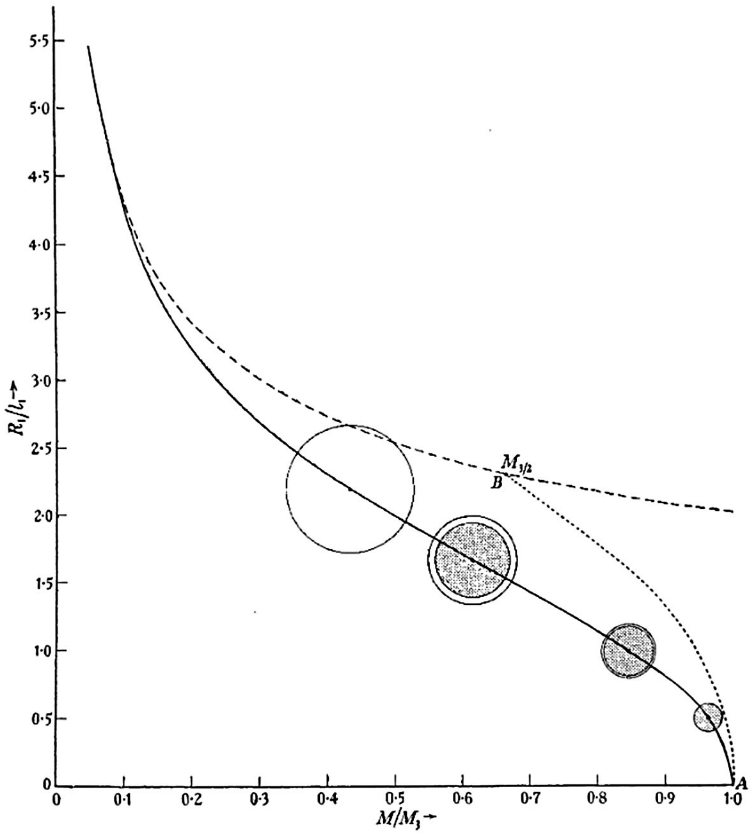


Fig. 2. The mass of the white dwarf along the abscissa, is measured in units of the limiting mass (denoted by M_3) for a stable white dwarf, that is 5.728 divided by the average molecular weight squared, a ratio directly emerging from his theory. The full line curve represents Chandra's theory, showing the exact (mass-radius) relation for completely degenerate configurations, showing stars with highly collapsed configurations at different stages. This curve tends asymptotically to the dotted line curve. As the mass of the white dwarf (M) approaches the maximum mass (M_3), the star shrinks while the radius R becomes zero. The dashed curve represents the relation $M = constant \times R^{-3}$ that follows from the non-relativistic equation of state $p = k_1(n_e)^{5/3}$ (low densities), thus representing Fowler's theory. The curve continues forever, thus showing that Fowler's theory does not predict a maximum mass; at the point B along this curve, the threshold momentum p_0 of the electrons at the centre of the configuration is exactly equal to mc. Along the exact curve, at the point where a full circle (with no inner circle) is drawn, p_0 (at the centre) is again equal to mc; the inner circles of the other circles represent the regions in these configurations where the electrons may be considered to be relativistic ($p_0 \geq mc$). The dotted line shows the transition from the core in Fowler's theory to the one in Chandra's (Chandrasekhar 1935a, p. 219). (This figure is subject to copyright protection and is not covered by a Creative Commons license.)

again correctly described the fate of a white dwarf with mass in excess of the critical value[44].

[44] For a detailed discussion on Eddington and the controversy over relativistic degeneracy see (Mestel 2004).

Having realized that relativistic degeneracy was incompatible with his theory, and yet having understood the alarming implications of Chandra's conclusions, Eddington paradoxically did not follow his own physical insight, accepting the physical reality deriving from relativistic degeneracy: in his eyes Chandra had actually revived the very same apparent difficulty solved by Fowler. Actually, already ten years before, Eddington had exactly described what would be the relativistic effects of a very powerful gravitational field exerted by a very big star with a mass between 10 and 100 times greater than the sun: "It is rather interesting to notice that Einstein's theory of gravitation has something to say on this point. According to it a star of 250 million km. radius could not possibly have so high a density as the sun. Firstly, the force of gravitation would be so great that light would be unable to escape from it, the rays falling back to the star like a stone to the earth. Secondly, the red-shift of the spectral lines would be so great that the spectrum would be shifted out of existence. Thirdly, the mass would produce so much curvature of the space-time metric that space would close up round the star, leaving us outside (i.e. nowhere)". Eddington then added that the same argument could be found in the writing of Laplace (*Système du Monde*, Book 5, Cp. VI): "A luminous star, of the same density as the earth, and whose diameter should be two hundred and fifty times larger than that of the sun, would not, in consequence of its attraction, allow any of its rays to arrive at us; it is therefore possible that the largest luminous bodies in the universe may, through this cause, be invisible" (Eddington 1926, p. 6). Eddington of course perfectly knew the Schwarzschild solution, but the above arguments again show that he did not believe in its physical reality.

In his paper 'Stellar configurations with degenerate cores' (Chandrasekhar 1935c), Chandra thanks McCrea, von Neumann, Rosenfeld and Strömgren "for the encouraging interest they have taken in these studies and for many stimulating discussions". All of them were his personal friends. However physicists did not want to enter openly the arena of such controversy, in part because Eddington was a most influential scientist, but also because they did not take Eddington seriously any more and thought that it was not worthwhile losing time in sterile discussions of what they considered completely wrong ideas. Moreover, astrophysics was still a field far away from the exciting new issues coming from theoretical and experimental physics of the early 1930s. On the other hand, Eddington was still admired as an authority by astronomers. So that on both sides, people chose not to be involved, or thought it was not worthwhile being involved, even if we know that physicists completely agreed with Chandra's work. As Chandra later recalled (Chandrasekhar 1977): " [...] all these people who supported me never came out publicly. It was all private". Actually, it was not completely like that. There was a solidarity from his young colleagues under the form of collaboration in articles. The more explicit one was one with Christian Møller (Chandrasekhar and Møller 1935). As Eddington had questioned the validity of the relativistic equation of state for degenerate matter, which by that time was generally accepted, they used Dirac's relativistic wave equation presenting arguments providing grounds "for not abandoning the accepted form of the equation of state". Eddington reacted to their article defending the relativistic degeneracy formula with a Note on 'relativistic degeneracy' (Eddington 1935c, p. 20): "In recent papers I have contended that the 'relativistic' degeneracy formula is erroneous. This has led Møller and Chandrasekhar to publish a note defending it. They give a derivation of the formula which is doubtless more up to date than those which I criticized. It therefore seems desirable that I should amplify my attack on the formula by showing why I am unable to accept Møller and Chandrasekhar's proof".

Chandra's relationships with young physicists is also testified by an investigation he carried on with Léon Rosenfeld on the deviation from perfect laws arising from causes other than degeneracy like the production of electron pairs and that resulted in a work published at that same time (Chandrasekhar and Rosenfeld 1935)[45]. Later Rudolf Peierls recalled: "I did not know any physicist to whom it was not obvious that Chandrasekhar was right in using relativistic Fermi-Dirac statistics, and who was not shocked by Eddington's denial of the obvious, particularly coming from the author of a well-known text on relativity. It was therefore not a question of studying the problem, but of countering Eddington. It was for this purpose that I wrote my paper in the *Monthly Notices* (Peierls 1936) [...] I do not believe Eddington ever took any notice of my paper" (Wali 1990, p. 135)[46].

Many years later Chandra told his biographer (Wali 1990, p. 143): "Kamesh, suppose, just for a moment, Eddington had accepted my result. Suppose he had said, 'Yes, clearly the limiting mass does occur in the Newtonian theory in which it is a point mass. However, general relativity does not permit a point mass. How does general relativity take care of that? If he had asked this question and worked on it, he would have realized that the first problem to solve in that connection is to study radial oscillations of the star in the framework of general relativity. It's a problem I did in 1964, but Eddington could have done it then in the mid-1930s! Not only because he was capable of doing it – he certainly had mastered general relativity – but also because his whole interest in astrophysics originated from studying pulsations of stars. And if he had done it, he would have found that the white dwarf configuration constructed on the Newtonian model became unstable before the limiting mass was reached. He would have found that there was no *reductio ad absurdum*, no stellar buffonery! *He would simply have found that stars became unstable before they reached the limit and that a black hole would ensue. Eddington could have done it.* When I say he could have done it, I am not just speculating. It was entirely within his ability, entirely within the philosophy which underlies his work on internal constitution of stars. And if Eddington had done that, he would stand today as *the greatest theoretical astronomer of this century*, because he would have predicted and talked about collapsed stars in a completely and totally relativistic fashion. It had to wait thirty years' ".

Such an exploration, commented Wali, "was not outside Chandra's ability either[47]. He reported some of his work on rotating white dwarfs at the 1939 Paris meeting,

[45] In Chandra's biography (Wali 1990, pp. 129–131) a correspondence with Rosenfeld, who was working with Bohr in Copenhagen, is mentioned in relationship to that period, January 1935. Bohr, too, expressed the opinion that there was nothing wrong in Chandra's formulation. On January 29, 1935, Rosenfeld wrote Chandra, also on Bohr's behalf: "Would you agree for us to forward confidentially Eddington's manuscript to Pauli, together with a statement of the circumstances and asking for an 'authoritative reply'?" About Eddington's manuscript, Rosenfeld remarked: "After having courageously read Eddington's paper twice, I have nothing to change in my previous statements; it is the wildest nonsense". Pauli declared that "Eddington did not understand physics", but, as Chandra wrote to Rosenfeld, "astronomers continued to believe in Eddington".

[46] Peierls was referring to a controversy arisen as to whether the pressure-density relation of a degenerate relativistic gas enclosed in a certain volume would be independent of the shape of the volume (Eddington 1935d, p. 258). According to Peierls, "This might seem sufficiently obvious to make a proof unnecessary" but in view of the controversy it was worthwhile to give a proof... So that he assumed that "the present form of quantum mechanics applies to the problem", and only proves that from this theory one obtains the usual equation of state.

[47] See (Wali 1990, pp. 135–146) for a description of Chandra's relationship with Eddington and the circumstances that led him to change his field of interest and go into something else: "It was a personal decision I made at the time". He definitely felt "totally discredited by the astronomical community".

and the paper in the Synge volume published in 1972 contains an almost verbatim account of the work he had done in 1935"[48].

In this regard, the most interesting of Chandra's friend of that time was John von Neumann. According to Chandra's later recollections, they became quite friendly during the period 1934-1935, when von Neumann was in Cambridge, on leave of absence from Princeton. This happened exactly at the time when Chandra had his controversy with Eddington. Chandra acknowledged that (Chandrasekhar 1977) "Neumann was one of the people who privately supported me against Eddington [...] I got to know Neumann rather well. I was a fellow at Trinity at that time, Neumann used to visit me in my rooms in Trinity quite frequently. I think he was rather lonely in those days, so he would quite often come up to my rooms in the college and sit down and work in my rooms, and so I got to know him rather well [...] We used to go out for walks". In the spring of 1935, they discussed Eddington's objections (Wali 1990, p. 143): "John said, 'If Eddington does not like stars to recede inside the Schwarzschild radius, one probably should try to see what happens if one uses the absolute, relativistic equation of state'. We started working on that together, but to go on we had to study equilibrium conditions within the framework of general relativity". In 1934 von Neumann had discussed with Abraham H. Taub and Oswald Veblen the extension of the Dirac equation to general relativity (Taub et al. 1934), and was thus in the right position to recognize that Chandra's problem of the limiting mass almost naturally led to apply general relativity, as on the other hand Eddington's acrimonious comments were implying. As Chandra further recalled: "at that time we started to work on some problems in relativistic gas spheres; it didn't go very far. I do remember our discussions of that year, and I did some work and published a paper in the late early seventies, on precisely the problem which Neumann and I discussed in 1934 – the problem of isothermal gas spheres in general relativity. In a way, it shows Neumann's great insight. He said, 'If objects are going to collapse, then they must collapse to smaller dimensions. We ought to look at it in the framework of general relativity...'. We were in the right direction. And in this instance I must say that it was Neumann who took the initiative"[49].

However, soon von Neumann left Cambridge and probably involved in different researches abandoned his work on the problem. Chandra on the other hand "got

[48] Wali is referring to 'A limiting case of relativistic equilibrium' (Chandrasekhar 1972). And actually, in 1962 Chandra decided to turn to general relativity – a subject he was first introduced to during his first year as a graduate student in Cambridge. In 1964 he worked out the theory of pulsation of spherical stars in the framework of general relativity, proving their relativistic instability against gravitational collapse. This most cited work marked Chandra's entry into the 'seventh period' of his scientific life, which started around 1960, when he began to study general relativity thus being ready to work in relativistic astrophysics in coincidence with the discovery of quasars (Friedman 1996).

[49] In 'Stellar configurations with degenerate cores (second paper)' (Chandrasekhar 1935a), Chandra cited an unpublished result of von Neumann, who "has shown that the *very* ultimate EOS (Equation of State) for matter should *always* be $P = \frac{1}{3}c^2\rho$". And actually, in von Neumann manuscripts, there are notes written in 1935, which were published by Abraham H. Taub in the 6th volume of his *Collected Works* (von Neumann 1961). In the first note (p. 172), where he studied the nature of the 'Static solutions of Einstein field equations for perfect fluid with $T^\rho_\rho = 0$', the space-time was assumed to be a static spherically symmetric one. The discussion of such solutions was reduced to the discussion of a differential equation in which pressure and density satisfied $\rho = 3p$ and the result was compared with that obtained in the classical theory. In the second note (p. 173), 'On relativistic gas-degeneracy and the collapsed configurations of stars', von Neumann is approximating the equation of state of degenerate matter presumably occurring in white dwarf stars by different equations for various ranges of the density.

sufficiently discouraged with the situation to leave the problem alone". So, all this turned into a lost occasion.

In spite of his relationships, Chandra was still very young and moreover all the questions appeared to most physicists as a side problem respect the growing field of nuclear physics, the fundamental issue of the sources of stellar energy and other relevant theoretical developments like quantum electrodynamics and the emerging topic of particle physics. As Chandra told Wali (Wali 1990, p. 145): "I felt that astronomers without exception thought that I was wrong. They considered me as a sort of Don Quixote trying to kill Eddington". Wali, Chandra's biographer, immediately commented (Wali 1990, p. 144): "The moral is that a certain modesty of approach toward science always pays in the end. These people [Eddington, Jeans, Milne], terribly clever, of great intellectual ability, terribly perceptive in many ways, lost out because they did not have the modesty to say, 'I am going to learn from what physics teaches me.' They wanted to dictate how physics should be". As a matter of fact, Chandra's work had been "completely and totally discredited by the astronomical community", so that he decided "to change the field of interest and go into something else". In fall 1935 he received an offer from Harvard to lecture in 'Cosmic Physics' and on November 30 he sailed from Liverpool bound to the New World, leaving behind his frustrating involvement in this clash of giants.

Despite Chandra's feelings, theoretical astrophysics emerged during this period as a specialty dedicated to the physical interpretation of celestial phenomena. The strong connection established between the new generation of astrophysicists like himself and Strömgren with the physicists' community, was instrumental in their capacity of bringing new results from physics to bear on stellar problems. In turn, this interaction between the two communities, stimulated some theoretical physicists to tackle astrophysical problems from the point of view of nuclear physics, an exploding frontier field materialized by the new perspectives opened by the neutron. However, the extreme consequence of the limiting mass was still to be explored and this further fundamental step would be triggered by more systematic investigations on the pressing issue of generation of energy in stars, which during the 1930s evolved into a hot research topic within the physicists' community. An important premise in this sense were laid down by studies systematically analyzing the properties of neutronic matter in stars, a study inaugurated by Flügge in his dissertation (Flügge 1933).

12 Hund and Kothari: neutronic matter in stars

As early as 1936, an extensive review on the status of the theory of matter under high pressure and temperature was prepared by Friedrich Hund (Hund 1936). Hund's relevant work in the quantum theory of solids and in the electrons in crystal lattices, as well as his interest in the field of nuclear physics, led him to analyze such physical aspects, using stars as cosmic laboratories providing information about the actual existence of such extreme states. At the same time, regularities in the observed properties of stars could provide support for the relevant laws of matter. Basing on fundamental physical considerations, Hund systematically investigated the properties of a gas of electrons, nuclei, protons, and neutrons, when the temperature and density are extremely high: "From what is known about the β decay of nuclei, one can conclude that protons can transform into neutrons by absorbing electrons or emitting positrons. Based on (Sterne 1933b,c,d), Hund remarked that "at high pressures it can prove to be favorable for the electrons and the nuclei together to transform into neutrons". (Hund 1936, p. 230).

He then considered a gas consisting only of neutrons (see in particular the section 'Das Neutronengas', p. 227) and the transformation processes occurring in regions

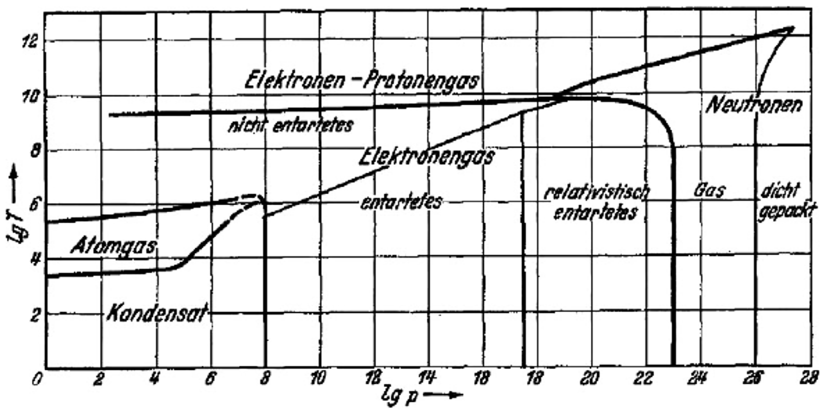


Fig. 3. Phase diagram of matter from (Hund 1936, p. 232), showing the results of calculations for the following systems: (a) nonrelativistic nondegenerate electrons and heavy particles; (b) nondegenerate electrons and heavy particles, relativistic electrons, non-relativistic heavy particles; c) relativistic degenerate electrons, nonrelativistic nondegenerate heavy particles; d) relativistic degenerate electrons, nonrelativistic degenerate heavy particles. (This figure is subject to copyright protection and is not covered by a Creative Commons license.)

of different equations of state of the particles. He was thus able to plot the boundaries between the different areas of electrons and nuclei, electrons and protons, and neutrons (Fig. 3). He found that beyond a certain value of the pressure the transformation of matter into neutrons occurred quite suddenly so that "the nuclei and electrons rapidly disappear" and "matter behaves as a neutron gas".

After having systematically explored the properties of matter, Hund applied his investigations to make some order-of-magnitude predictions about the pressures and temperatures which occur within celestial bodies, such as planets, ordinary stars and, finally, 'dense stars', as Hund named white dwarfs, probably in one of the first explicit uses of this expression (see section 'Die dichten Sterne', p. 253). He took for granted the existence of the limiting mass (Hund 1936, p. 254): "Chandrasekhar has calculated the structure of a star, for which the temperature is no longer important, with a more exact equation of state valid for the nonrelativistic and the relativistic electron gas. As the mass increases, the radius decreases; for the mass of the sun, the radius is approximately equal to that of the earth, and at even higher masses the radius tends rapidly to zero. The zero radius is reached for a finite mass only slightly larger than the mass of the sun. This last result should not be taken too literally, because, for calculating the equation of state, it was assumed to have unlimited validity for high pressures. This collapse to a zero radius (*or to the corresponding value in the general relativity theory* [emphasis added]) stems from the high compressibility of matter in the state of the relativistic degenerate electron gas. If a star above the limiting mass were to have finite radius, the pressure would of course increase, but not fast enough to meet the corresponding increase in the weight of the above-lying layers". Hund correctly remarked that the transformation of matter into neutrons would result in a greater limiting mass, but concluded that stars with sufficient mass could reach radii of the order of 10 km. How to avoid the 'small radii problem'? A possible solution for stars with high masses would be to radiate large amounts of the gravitational energy set free in the process of contracting, reducing its mass significantly. "As a *possible final state in the evolution of stars*, concluded Hund, we are thus led to expect stars of moderate mass with very high densities".

As a matter of fact, Hund was providing a physical base for the concept of neutron matter in stars. But, it was clear that at those pressures and densities the equation of state of nuclear matter was still far from being understood. And the generally spread hope was that something would intervene and save the star from a catastrophic collapse that was still a 'black box', both in term of the properties of matter in such extreme conditions and from the point of view of the collapse process itself, that nobody had still tackled. Hund, had someway briefly touched on the subject when he had mentioned the Schwarzschild radius, but only in brackets, as a side comment.

A step forward in this path was taken the following year by Daulat Singh Kothari, a student of the renown astrophysicist Megh Nad Saha at Allahabad University and later of Ralph Fowler at Cambridge University (Vardya 1994). He had written several papers on degeneracy and dense matter in celestial bodies and independently by Hund introduced the neutronization of matter in the interior of white dwarf stars by inverse beta decay process (Kothari 1937) calculating, for example, the value of the mass for which the electron concentration would reach the maximum possible value beyond which all free electrons would combine with protons to form neutrons. But he was much more explicit in investigating implications deriving from the transition to the neutron phase within superdense matter, thus setting the stage for a major role of neutrons within stellar cores.

13 The superdense core and the problem of stellar energy

By the 1920s it had become clear that gravitational energy was insufficient as a source for powering stars. The radiation of the sun could not be maintained through a period of more than a billion years (the age of the earth at the time was estimated to be 3 billion years) solely through the release of gravitational energy. The release of nuclear energy through the transformation of hydrogen into helium was regarded as a likely mechanism.

During the 1930s considerable progress was made in the field of nuclear physics, both through laboratory experiments and through further development of theory. The theory of stellar interiors had reached a point where the temperature, density and chemical composition of the central regions of main-sequence stars could be specified fairly accurately. Now the task was to compute, or estimate, which nuclear processes would be effective under such circumstances, what the reaction rates were, and how much nuclear energy would be produced per gram per second. The physics of the nuclear processes in the sun naturally stood at the center of interest. Discussions on nuclear synthesis and stellar radiation were now based on neutrons, as units from which nuclei are built together with protons, and from which elements are formed in stellar interiors. A brief mention of the view "that the stars contain central cores consisting largely of free neutrons" since the early life of stars, where such large amount of free neutron would produce light and heavier elements by nuclear reactions, was made by Harold J. Walke in an article on nuclear synthesis and stellar radiation, (Walke 1935, p. 365). He proposed "a complete theory of nuclear synthesis by neutron capture and $\beta-$radioactivity", regarding the neutron "as a fundamental nuclear component, just as the electron is the fundamental extranuclear component". On this theory therefore protons and α-particles would be formed mainly within nuclei as a result of the $\beta-$radioactivity. He also suggested that "the initial condition of the universe" consisted of a uniform distribution of neutrons and gamma-radiation. This primaeval gas, as previously suggested by Jeans, would be gravitationally unstable, and according to Walke it would condense "to form huge non luminous nebulae". As a result, hydrogen would be produced from the more frequent collisions between neutrons. Walke is also mentioning Baade and Zwicky (Baade and Zwicky 1934b): neutrons would accumulate at the centre of a star and thus, he concluded, element formation must take place in stellar interiors, where also cosmic rays could originate (Walke 1935, p. 362).

Later, von Weizsäcker, in remarking that at that time there wasn't very close contact between astronomy and physics, also added that, of course, in astronomy there was one great problem : "every physicist who was working in fields like ours, like, for instance nuclear physics, knew that the problem of the interior of the stars was probably solved by Eddington, with the exception of the problem of the energy and that this was a problem of physics was clear, too. It was not clear how it was to be solved [...] we liked discussing this, of course [...] I would say that people like say Nordheim, who at that time was also in Göttingen, or – Placzek, Weisskopf, Bethe, the whole group, Bloch – they all would have taken some academic interest. I mean, not an active interest, but some general interest in astronomical questions. But none of them, I think, had the idea that he would be working in astronomy". While visiting Bohr's Institute in Copenhagen, von Weizsäcker himself had discussed astrophysical issues with Strömgren, and had suggested in his monograph on atomic nuclei completed in the summer of 1936 that the quickly growing knowledge of nuclear reactions would suffice to resolve the stellar-energy problem (von Weizsäcker 1937a). From these reflections arose his interest in seeking to explain how thermonuclear reactions could build elements up to their present abundances, thus opening the race to find a solution of the stellar-energy problem (von Weizsäcker 1937b) (von Weizsäcker 1938)[50].

By 1937-1938 it was a spread knowledge that energy-generation in stars is the conversion of hydrogen into helium. What was not established were the thermonuclear reactions involved in such process. A turning point in these developments, was Gamow's growing interest for astrophysical issues, a new era in his scientific life. Already in 1933 he had written with Landau a paper investigating the process of thermal transformation of light elements in stars (Gamow and Landau 1933) and was thus invited to give a talk in Paris on the evolution of stars. After participating to a meeting in London in 1934, he then emigrated to the United States. The issue of nuclear reactions powering stars, and the connected fundamental problem of the origin of chemical elements, was discussed by Gamow in a lecture at Ohio University, and later published in the *Ohio Journal of Science*, a rather obscure journal (Gamow 1935). After discussing nuclear transformations especially investigated by Fermi's group in Rome, Gamow shifted his attention from the experimental evidence obtained in the laboratory to the processes happening in the interior of stars. Apart from trying to outline the mechanisms for the building of elements, he also came "to one of the most interesting questions concerning the physical state of the matter deep inside of stars [...] a mixture of two ideal gases: nuclear gas and electronic gas". Basing on Landau's theory of 1932, according to which most stars included a core of superdense 'neutronic' matter of nuclear density, i.e. about $10^{12} g/cm^3$, Gamow gave a short account of Landau's calculations related to the equilibrium problem between the pressure of the electronic gas in the star's interior and the gravitational pressure of the outside layer and showed a diagrammatic representation for three different masses of the star (Fig. 4).

Gamow observed that as far as the momentum is small compared with mc the pressure P of the ideal electronic gas is directly proportional to $\rho^{5/3}$. For larger densities velocities become relativistic and the pressure varies as $\rho^{4/3}$. The outside pressure P' due to gravitation is proportional to $\rho^{4/3}$ and the coefficient of proportionality depends on the total mass M of the star. If M is small (curve II') the inside pressure will be always larger; for somewhat larger mass (curve II'') there is a state of stable

[50] See (Shaviv 2009) for a detailed discussion of (von Weizsäcker 1937b) and (von Weizsäcker 1938).

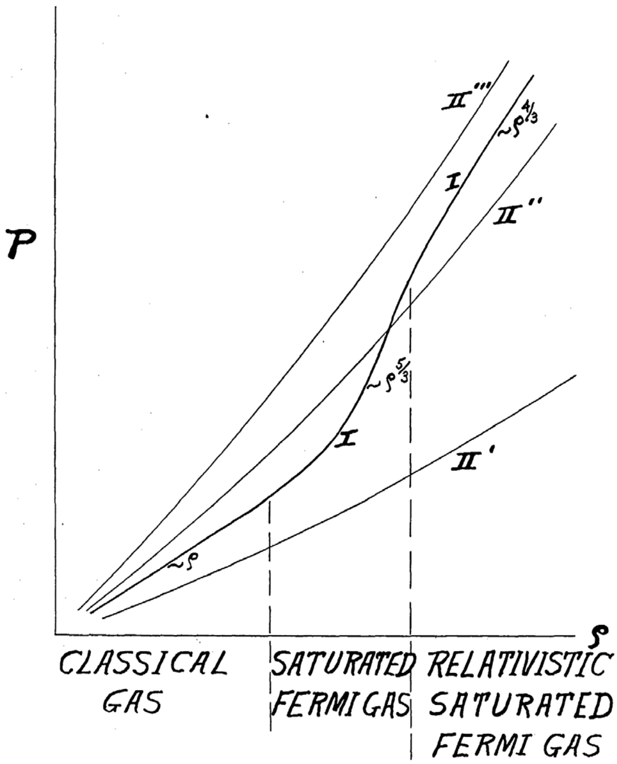


Fig. 4. Diagrammatic representation of the outer pression-density curves corresponding to three different masses of the star (curves II', II'' and II''') compared with the pressure-density relation for an ideal electronic gas (curve I) (Gamow 1935, p. 412). (This figure is subject to copyright protection and is not covered by a Creative Commons license.)

equilibrium between P and P' and the star will have inside a region filled up with non-relativistic saturated Fermi-gas. For still larger masses (curve II''') the inside pressure would "never be able to oppose the weight of stellar substance and the star would collapse into a mathematical point (!) unless, the further compression would be stopped by intranuclear repulsive forces between the particles of nuclear gas". Here, the evolving knowledge on nuclear matter suggested Gamow a 'nuclear argument' to avoid the collapse that Landau had prevented by using Bohr's views about non conservation of energy in nuclear processes. Gamow then mentioned Landau's calculations on the limiting mass, and proposed that *all stars possess such nuclei which evidently represent the sources of the stellar* energy radiated in such large amount into interstellar space". Of course, he added that the question of the mechanism of energy-liberation was not yet quite clear. Moreover, proposed Gamow (Gamow 1935, p. 413), one could "easily imagine that the stellar nucleus may not be considered as an inactive globe. The eruptive processes from the surface of the stellar nucleus will throw out the small pieces of nuclear substance which coming into the outside layer of the star will immediately disintegrate giving rise to the nuclei of different stable and radioactive elements". Gawow expressed the hope that further investigations might clarify "the relative importance of various processes and lead to a complete explanation of the relative abundance of different elements in the universe".

In arriving to United States, Gamow was employed at George Washington University, where he always gave two regular courses, advanced courses

(Gamow 1968): “relativity, quantum theory, nuclear physics”. But by this time he was more interested in applications of nuclear physics: “Nuclear physics as such became boring for me, became too complicated, with all these complicated experiments and complicated theory, and I was doing nuclear astrophysics, so to speak – the evolution of stars – so I was mostly connected with astronomers, with people like Baade and Hubble [...] And whenever I went to California I was always going to talk to astronomers. I was in much closer intercourse with astronomers than with physicists at this time”. At the same time, added Gamow, “there was always this hostile feeling that astronomers, especially theoretical astronomers, didn’t like me to invade their ground, because actually all these thermonuclear reactions in the stars were done by physicists – me and Bethe and Houtermans and Weizsäcker – because astronomers didn’t know about nuclear physics. They were sitting on their astronomical things [...] I started nuclear physics because in 1928 everybody was doing atomic and molecular structure, and van der Waals forces and doublets and triplets and spin and so on – it was too much. I didn’t want to get mixed up with all this, so I decided to choose myself a corner where nobody was doing anything, so I chose nuclear physics. And in time nuclear physics blew up into a big thing, so I moved to nuclear astronomy, to nuclear astrophysics, cosmology”.

In the mid 1930s, Gamow thus fully focused on stars as a playground for his skills in the fast growing field of nuclear physics of which he had been a pioneer with his 1928 theory of quantum tunnelling. He had a great physical sense and great imagination. Focusing on nuclear processes, he explored different stellar models in which problems of inner structure, energy sources and formation of elements in stars were all intermingled[51]. But apart from the fate of his models, what is relevant for this narrative is that he continued to cultivate the idea that all stars might have a superdense core in their interiors. In his volume *Atomic nuclei and nuclear transformations* (Gamow 1937) – an upgrading of his *Constitution of Atomic Nuclei And Radioactivity*, the very first textbook on nuclear physics (Gamow 1931) – Gamow discussed the nuclear state of matter in the interior of a star in the preface, dated May 1, 1936: “For still higher densities [$> 10^8$ g/cm^3] electrons will probably be absorbed by the nuclei (an inverse β-decay process) and the mixture will tend to a state which can be described very roughly as a gas of neutrons”. For densities of the order of magnitude $\rho \sim 10^{12}$ g/cm^3, average density of atomic nuclei, the conditions in the gas will become analogous to the conditions inside an atomic nucleus, pointed out Gamow, then citing Chandra and Landau in connection with the problem of pressure of degenerate matter in stars. He, too, mentioned the problem of ‘unlimited contraction’ beyond a mass of about 1.5 solar masses, without any further comment.

In showing that a gas of neutrons could be compressed to a much higher density than a gas of nuclei and electrons, Gamow was calling such an extreme state of matter ‘the nuclear state’ and and the region of the star occupied by such nuclear matter the ‘stellar nucleus’.

Although Gamow did not refer to it, because he completed the book in spring 1936, the microscopic descriptions of the equation of state of nuclear matter in beta equilibrium had also been independently given by Hund (Hund 1936). In the chapter ‘The new star model’ of his *Habilitationsschrift* published in 1936, Wilhelm Anderson, too, had talked of the formation in a few millions of years of a neutron core inside a star (Anderson 1936, p. 72)[52].

---

[51] See (Nadyozhin 1995) and (Cenadelli 2010) for an analysis of Gamow’s theorizing on stellar structure and evolution.

[52] The heavier neutrons would sink towards the center leaving behind a gas of electrons and protons. In this neutron-gas sphere, about half of the whole mass of the star would concentrate reaching enormous density and temperature unthinkable in the same condition

In the final pages of the book, Gamow then considers the conditions under which such stellar nuclei can really be formed. He then mentions (Landau 1932) and (Chandrasekhar 1931a) arriving at discussing "*the final state of the star*" up to the relativistic case, finding of course that for masses $M > M_0 \sim 1.5$ solar masses, "equilibrium will never be possible for larger densities and the compression will proceed without limit" (Gamow 1937, p. 237). He does not speculate further about the possibility "for such unlimited contraction", but he immediately remarks that "unlimited contraction may start already for smaller masses than M_0, if we take into account the exchange attractive forces between particles [...] Thus we see that most of the stars, and possibly all stars, if the limiting mass M_0 is lowered by intranuclear forces, are subject to the formation of matter in the nuclear state in their interior at some period of their existence". Milne's dense collapsed cores, similar to white dwarfs tucked within stars, had now been transformed by Gamow into superdense neutron cores, possibly playing a role in fundamental nuclear processes within stars. "The question whether most stars actually possess such nuclei cannot, however, be answered definitely until the relevant astronomical evidence has been thoroughly examined, but there seems to be no reason why they should not" concluded Gamow.

In the very last lines of the volume (Gamow 1937, pp. 234–238) he proposed the theory already exposed in the Ohio lecture, according to which "eruptive processes of different types may go on continuously over the boundary between a large stellar nucleus and the surrounding matter in the ordinary gaseous state" thus forming the nuclei of different elements. Moreover, one could easily see "that pure gravitational energy liberated in the contraction to such immense densities will already be quite enough to secure the life of the star for a very long period of time". This statement concluded the volume that during the following years certainly contributed to the diffusion of views about the possible role of neutron cores in nuclear stellar processes and especially put a seal on the possibility of their existence. It represented a further important step towards the construction of a well founded physical model for Zwicky's speculative neutron stars.

In the meantime, the problem of stellar interiors and all the connected issues, especially the source of stellar energy, were being widely discussed within the physicists' community. On November 5, 1937, Landau sent to Bohr the English version of an article in which he proposed an upgraded version of his 1932 super dense core now transformed in a neutron core and asked him to send it to *Nature*, if he would find that "it contains some physical thoughts". And added that he would be very glad to hear his opinion on the article.

On December 6 Bohr wrote to Landau, enclosing the proof of his letter to *Nature* (Landau 1938): "As I think you know from my letter to Kapitsa, we were all in the Institute much impressed by the beauty of the idea and its promise. In the meantime we have, however, had a number of discussions on astrophysical problems, in which our attention has been directed to two reports in the Ergebnisse der Exakten Naturwissenschaften for 1936 and 1937, written by F. Hund (Hund 1936) and B. Strömgren respectively (Strömgren 1937)[53].

---

for other kind of gases. In this way, so much contraction energy would be set free that it would be superfluous to look for other sources of energy for the sun. He then went on calculating on this model the contraction energy in the new model of star.

[53] Strömgren throws some light on the connection between these two reviews recalling that "one of those who came frequently to the Bohr Institute was Hund, and we discussed questions of stellar matters with him, and in the end it was agreed that he would write an article for the Ergebnisse on the physics of stellar-interior matter, and I would write the corresponding astrophysics review article [...] I found that, in the thirties, this is where they [physicists] got acquainted with stellar interiors, rather than through Eddington's book. For

Landau answered on December 17 that he had added a citation to Hund's article, but on January 14, having received from Bohr the article of Strömgren, he stressed that after reading it he had not been able to find anything connected with his own work: "Nur astrophysikalische Pathologie und etwas bekannte Kernumwandlungsphysik!" [Only astrophysical pathology and some well known nuclear transformation physics]. On January 14 Landau was again writing to Bohr, after receiving a letter from Møller again discussing such topics on which he had reflected once more: "The Strömgren's claims are unfortunately based on the wild Eddington's pathology which, as well known, is false not on one point but on all points. To unmask such pathology in a *Nature* note is completely impossible, such unmasking would be longer and more complicated than the whole article". As recalled by Peierls (Peierls 1997, p. 163): "He was very critical, as was most of our generation of theoreticians, and the comment 'falsch oder trivial' about suspect papers, used often by Landau, was in common use. He was also fond of the term 'pathologists' for people who wrote pathological papers, i.e. nonsense". In spite of Landau's harsh comments, Strömgren's review was instrumental in introducing physicists to the problem of stellar interiors.

In the starting lines of his article, Landau immediately stated that "in bodies of very large mass" the degenerate electron gas does not lead to extremely great densities, because of the 'quantum pressure'. On the other hand, continued Landau citing (Hund 1936), "it is easy to see that matter can go into another state which is much more compressible – the state where all the nuclei and electrons have combined to form neutrons [...] It is easy to compute the critical mass of the body for which the 'neutronic' state begins to be more stable than the 'electronic' state [...] When the mass of the body is greater than the critical mass, then in the formation of the 'neutronic' phase an enormous amount of energy is liberated, and we see that the conception of a 'neutronic' state of matter gives an immediate answer to the question of the sources of stellar energy. The sun during its probable time of radiation (about 2×10^9 years according to general relativity theory) must have emitted something of the order of magnitude of 3×10^{50} ergs. The liberation of this amount of energy requires the transition of only about 2 per cent of the mass of the sun (with the assumption of constant density) or even only $3 \times 10^{-3} M_\odot$ (with the Fermi gas model) to the 'neutronic' phase [...] Thus we can regards a star as a body which has a neutronic core the steady growth of which liberates the energy which maintains the star at its high temperature" Landau then expressed the hope that "The detailed investigation of such a model should make possible the construction of a consistent theory of stars".

As regards the question of how the initial core could be formed, Landau had already shown in 1932 that "the formation of a core must certainly take place in a body with a mass greater than $1.5 M_\odot$". However, he now concluded the article with a challenging question regarding the stars with smaller mass, for which "the conditions which make the formation of the initial core possible have yet to be made clear".

The last letter from Landau to Bohr is dated February 1, 1938 and he never replied to a letter by Bohr of July 5: "As you know all here have been very interested in your most suggestive idea about stellar-constitution, and we have lately followed very closely the discussions about it, which have taken place among astrophysicists. We are all very eager to learn what progress you have made with it yourself". He then continues with the exciting new perspectives "about the origin of the nuclear forces opened by the discovery of the heavy electron [...] It would surely be most pleasant

instance, a footnote by Tolman, shows how physicists got to know about the problem. There was also a limitation – it was in German. But in those years, even in America, obviously German was studied [...] it was so necessary... German, in those years, when quantum mechanics was developing". (Strömgren 1978).

and instructive to all of us to discuss these various prospects with you and we hope very much indeed that you this year will be able once again to take part in our annual conference for the old and present collaborators of the Institute, which is planned to take place in the first week of October"[54].

In the meantime, Gamow, together with Merle Tuve and Edward Teller, was organizing the Fourth Annual Conference on Theoretical Physics sponsored by the George Washington University and the Carnegie Institution of Washington to discuss the burning problem of nuclear energy in stars[55]. The fourth conference, devoted to 'The problem of stellar energy and nuclear processes', was held in Washington, D.C., on March 21–23. It represented Gamow's official entering into the astrophysical realm, a circumstance well reflected by the mixed character of the invited scientists: astrophysicists studying the internal constitution of the stars (S. Chandrasekhar, B. Strömgren, T.E. Sterne, D. Menzel and others) and physicists working on different branches of nuclear physics (H. Bethe, G. Breit, G. Gamow, J.V. Neumann, E. Teller, M. Tuve, L. Hafstad, N. Heydenburg and others). Chandra was at the time completing his book *An introduction to the study of stellar structure*, and Strömgren had just written his review article on the theory of stellar interior and stellar evolution (Strömgren 1937), both had recently moved to the University of Chicago. The stage was set for discussing astronomical observations, astrophysical theories and theoretical physics within a common perspective and establish a collaboration between astronomers, astrophysicists, and experimental nuclear physicists which led to the emergence of nuclear astrophysics as an established research field.

It is not by chance that the first meeting was opened by Strömgren, who outlined in some detail the mathematical treatment and current status of the problem of temperature and density distribution and chemical composition in the interior of stars, with special reference to the critical features of the various particular stellar models used for these calculations. The bearing of current knowledge of nuclear reactions on the evaluation of the behavior of stars with nuclear sources of energy was reported by Gamow, while Bethe reported on the study of particular nuclear reactions which would lead to liberation of energy and to the building up of heavy elements. By that time Bethe's wide knowledge had just been displayed in his 'trilogy' that later became known as the 'Bethe Bible', presenting a complete coverage of nuclear physics published in the *Reviews of Modern Physics* written between 1936 and 1937 in collaboration with his colleagues Bacher and Livingston.

Another question which brought about much discussion during the conference concerned the degree of central condensation of stars, together with the possible existence of a super-dense stellar nucleus, at least in some stars, as recently proposed by Landau. Chandra reported his investigations concerning the possibility of high central condensation in various known stars. His results lead to the conclusion that, whereas for giant stars the degree of central condensation is necessarily slight, there are stars for which as much as 90 per cent of the total mass is concentrated within less than half the radius from the center. Another aspect of the problem of central

[54] Archives for the History of Quantum Physics, Bd. AHQP/BSC 19: Niels Bohr. Scientific correspondence, 1930–1945.

[55] When George Gamow had been employed at George Washington University, the joint meetings organized by the Carnegie Institution and the University had come about as a condition for his employment in order to avoid the isolation from other theorists. Their style was obviously inspired by the conferences organized by Bohr in Copenhagen, having the same informal character and being very limited in size with no published proceedings. The first one, held in 1935, was devoted to a discussion of the latest problems of nuclear physics. The 1936 conference focused on molecular physics, and the third one on problems of the properties and interactions of elementary particles and the related questions of nuclear structure.

condensation was given by Sterne, who indicated the possibility of direct estimates of the density-distribution of double-star components from the observed characteristics of their orbits. The study of the stellar model having a highly condensed neutron-core of the Landau type was reported by Teller. By direct integration of the equations of stellar equilibrium, one arrived for such models at extremely high temperatures ($\sim 10^9$ C) and densities ($\sim 10^9$ g/cm^3) near the surface of the core. Since under such conditions the already-known nuclear reactions would proceed with extremely high velocities, it was concluded that such a star model is inherently unstable. Thus, as far as astrophysical evidence was concerned, the model of a star with a heavy stellar nucleus at the centre was *not* confirmed, except possibly for supergiants, according to the report on the conference published on *Nature* by Gamow, Chandra, and Tuve (Chandrasekhar et al. 1938).

They made an interesting final remark: "Valuable contributions to the discussion of such superdense state of matter in a stellar interior from the point of view of general theory of gravitation was given by Neumann" thus providing hints of a follow up of von Neumann's reflections on these issues going back to his discussions with Chandra at Cambridge in 1935[56].

At this conference Hans Bethe was inspired, especially by Strömgren's new estimates of the solar interior temperature, to investigate those processes that produce energy in massive stars[57]. His paper published in March 1939 (Bethe 1939), in which he showed that the most important source of energy in ordinary stars are the reactions of carbon and nitrogen with protons forming a cycle in which the original nucleus is reproduced, was a landmark paper that formed the basis of much work in astrophysics for decades.

These results demonstrated how farsighted had been the organizers of the conference in gathering together nuclear physicists and astrophysicists for the first time. The researches of the previous two decades into the constitution of the stars had resulted in considerable advance in the understanding of the physical processes in stellar interiors. The chief success of the investigations was the establishing of a mass-luminosity relation. This relation had been obtained without reference to the actual nuclear reactions that are the source of stellar energy, merely from consideration of the mechanical and thermodynamical equilibrium of the star. The problem of stellar energy had to be tackled by nuclear physicists who had devoted all their time to the field to sort it

[56] At that time, Chandra included further unpublished results by von Neumann about the point-source model (in which it is assumed that the entire source of energy is liberated at the center of the star) in a dedicated section of his book *An Introduction to the Study of Stellar Structure*, published in 1939. At page 332 he emphasized that "Von Neumann's treatment of this problem is very powerful". Two manuscript notes related to this issue, entitled 'The point-source model' and 'The point-source-solution, assuming a degeneracy of the semi-relativistic type, $p = K\rho^{4/3}$ over the entire star', were published in von Neumann's *Collected Works* (von Neumann 1961, pp. 175–176).

[57] See what Bethe says at p. 2 of his autobiography 'My life in astrophysics' (Bethe 2003): "Strömgren, a well-known Scandinavian astrophysicist, reported that the central temperature of the sun was now estimated as 15 million degrees, not Eddington's 40. This is still the estimate. This change came as a result of assuming that the sun was predominantly hydrogen with approximately 25% helium, rather than assuming it had about the same chemical composition as the earth". However, Bethe is opening this contribution stating that his first involvement with astrophysics "came as a result of Carl Friedrich von Weizsäcker's suggestion to investigate the fusion of two protons to form a deuteron, namely $H + H \to D + e^+ + \nu$". Actually Gamow had suggested to one of his graduate students, Charles Critchfield, to calculate the proton-proton reaction and in early 1938 the work was submitted to Bethe. The latter found the calculations to be correct and they wrote a joint paper (Bethe and Critchfield 1938) paving the way to Bethe's celebrated paper of March 1939 (Bethe 1939).

out. At the same time, they needed convincing results communicated by astrophysicists: the temperature, the density, the composition... without using at all the input from energy production. The chief success of the investigations is the establishing of a mass-luminosity relation. As Strömgren recalled (Strömgren 1978): "It was simply due to this situation that, whatever the mechanism, it must be one that gives a high degree on concentration of energy production in the central region. Then there's no doubt about the model and that fixes the temperature. Once this was understood by the physicists who were ready to accept this, in spite of what, shall we say, Landau said, then that communication was easier than the other. There were so many things that were very difficult for one who wasn't a nuclear physicist to appreciate". This turned the study of stellar structure from one containing a substantial degree of arbitrariness to one in which definitive models could be derived for any given star in any given state of evolution.

A conclusion had been thus reached during the conference that stellar models with a concentrated nuclear core could not represent standard stars[58]. However, shortly after the conference, Gamow relaunched the subject (Gamow 1938a) remarking that, as the stars of the giant class are distributed in the Hertzsprung-Russell diagram "in a very peculiar way, very different from the main sequence" the energy source in giants must be entirely different, probably due for example "to the beginning of the formation of a dense neutron core in the centers of these stars" since all giants "have the masses larger than the critical mass of Chandrasekhar and Landau". In this case, suggested Gamow in his tentative theory of novae (Gamow 1938b), the formation and growth of a neutron core, "representing a practically unlimited source of energy, should be expected. The growth of such a core [...] may bring the star into the Giant branch of the H-R diagram"[59]. At this point, an explosion of these massive stars would occur, leading to extremely bright novae, that Gamow identified "with the so-called super-novae of Baade and Zwicky".

Gamow was in fact reacting to a paper by Zwicky concluding a wide search for super-novae that the latter had carried out during the last two years using an improved telescope (Zwicky 1938a). The most important conclusion which he drew from these new observational results was that "the existence of *two classes of temporary stars*, super-novae and common novae, has been established beyond doubt". The idea that a certain stage of contraction one might expect that the formation of a large amount of free neutrons would lead to a rapid collapse of the whole star and to the liberation of tremendous amounts of gravitational energy was again related by Gamow to the Baade and Zwicky supernova mechanism the following year (Gamow 1939b).

Gamow's pioneering role in connecting nuclear physics and astrophysics, is testified by a long report about the nuclear transformations as energy sources in stars submitted on May 25 to the *Zeitschrift für Astrophysik*, where he also discussed very dense stars and the accretion of neutrons into an extraordinarily dense and well delimited core ("*Landauschen Kern*"), which could not have been formed from "usual contraction processes" of a less dense material (Gamow 1938c, p. 155). One must thus suppose that they are produced by some external forces when the star was born. In particular he put forward the idea that if these cores existed in all star, one might speculate that, "according to the theory of the *expanding universe*, the whole space in the past must be of a quite small dimension and filled with matter of exceedingly high density. During the expansion process, this 'seed of the world' ['Weltkern']

---

[58] Within one month, the question was also discussed by Gamow and Teller during the APS Meeting of April 28-30 (Gamow and Teller 1938).

[59] On the other hand, a star deprived of any source of nuclear energy would progressively contract and eventually become a white dwarf, for masses smaller than the critical mass.

would disintegrate in smaller pieces, that now, embedded in less dense atmospheres are observed as shining stars".

This was an early hint of Gamow's commitment in the cosmological problem connected with the building of elements that would become a major topic immediately after the war, and shows the influence of Lemaître's physical cosmology, which had been again recently discussed during a conference on the physics of the universe and the nature of primordial particles [cosmic rays] organized at the University of Notre Dame, Indiana, on 2–3 May by Arthur Haas, most probably in collaboration with Lemaître himself, who was at the time a visiting professor there (Department of Physics Notre Dame University 1938). It gathered about a hundred scientists and was one of the first in which cosmology was a main focus. As a student Gamow had been especially fascinated by special and general theory of relativity, and for this reason he had followed Friedmann's course entitled 'Mathematical foundations of the theory of relativity', and "at first hand, directly from him", he had learned the theory of the expanding universe. However, Friedmann prematurely died: "This ruined my plans to continue my work on relativistic cosmology" recalled Gamow in his autobiography (Gamow 1970, p. 45). Although he was only 21 years old when Friedmann died, he continued to consider himself a pupil of Friedmann. All this was now resurfacing and shows how he could well be inspired by Lemaître's theory of a 'colossal explosion' of the primeval atom. Gamow's concluding lines of his report explicitly expressed the hope that "the close collaboration between astronomers and physicists" would soon lead to an answer to the question of the evolution of stars (Gamow 1938c, p. 160).

During that same 1938, von Weizsäcker had published a new article on the problem of energy production in stars in which he also proposed as origin of the universe the result of a cosmic explosion from a superdense compressed nuclear state. His physical cosmology is very similar to Lemaître's primeval atom, however, as emphasized by Kragh (Kragh 1996, p. 99), like Gamow would do later, von Weizsäcker "did not refer to general-relativistic models and did not try to combine his nuclear-historical sketch with the geometrical history of the universe as given by the Friedmann-Lemaître equations. In this sense, it was only half a big-bang hypothesis". However, its strong nuclear physics content would later provide an inspiring key for Gamow's later big-bang cosmology.

14 Oppenheimer and Serber: the stability of a neutron core

During that 'hot' summer of 1938, the stage was set for Oppenheimer's entrance into the still open problem of stellar energy, at which Bethe had begun to work after the Washington conference, and for which he would soon provide a solution.

The road to Oppenheimer's work on the problem of massive stars has been beautifully reconstructed by Hufbauer (Hufbauer 2006) with plenty of interesting details.

According to this reconstruction, before 1938 Oppenheimer came in contact in several occasions with problems belonging to theoretical astrophysics, starting of course from the already mentioned circumstance of his sojourn as a postgraduate student in 1925–1926 at Cambridge University, where he had Fowler as chief mentor. Apart from a series of interesting occasions described by Hufbauer which are forming a convincing background motivating Oppenheimer's interest in stellar theorizing, it is to be further emphasized that during the early 1930s, nuclear physics, cosmic ray physics, and the emerging field of particle physics were still very much part of the same scenario, in which Oppenheimer was actively working.

With the increasing knowledge on the nuclear realm Oppenheimer and many others continued to keep an open eye on the problem of reactions in stars, because of the possibility of understanding how a whole series of nuclear processes that could not be reproduced in terrestrial laboratories, took place in stellar interiors. In a similar manner, the first accelerators used to bombard the nucleus helped in having experimental beams of particles, but could not compete with the high energies typical of the cosmic rays. These were thought to provide information of nuclear processes, in particular after the detection of the mesotron in cosmic rays, which for several years was identified with Yukawa's meson, the carrier of nuclear forces, only during the late 1940s becoming the weak interacting muon. Since 1933 Oppenheimer had been deeply involved in work on the positron, on collisions processes generated by the interaction of high-energy particles or radiation with matter, as well as on the mechanisms underlying the formation of showers and 'nuclear stars' following interactions generated by cosmic ray particles of very high energy and in general in problems related to the interaction between neutrons and the nucleus. In 1937–1938 several of his contributions written with his collaborator Robert Serber, focused on mesons and cosmic rays. A couple of articles appearing between August 1937 and April 1938 on nuclear reactions involving transmutations of light nuclei tackled problems which were not very far from the processes which were at the time being discussed as possible sources of stellar energy. As emphasized in (Thorne 1994, p. 187) Robert Oppenheimer was in the the habit of reading with care every scientific article published by Landau. Thus, Landau's article on neutron cores in the 19 February 1938 issue of *Nature* caught his immediate attention.

In the meantime, the 4th Washington Conference focusing on the problem of stellar energy had taken place and things were becoming ripe for its solution. During the summer, following the Conference, Oppenheimer invited Bethe to lecture to his students and collaborators, and so there was plenty of time for talking about the interior of stars. Landau's model of the neutron core was widely discussed, especially in connection with its possible role in giant stars. Apart from this, neutron cores were in itself very appealing for physicists: It was nuclear matter, after all, resembling a giant nucleus made up mainly by neutrons, so that Oppenheimer thought that it was worthwhile exploring the physics involved. He was of course well aware of Gamow's arguments about a superdense core in stars discussed in his *Structure of atomic nuclei and nuclear transformations* (pp. 232–238) that had been published the previous year, as well as of his latest articles published between spring and early summer. Landau's article must have been widely discussed with Serber and with Oppenheimer's brilliant student George Michael Volkoff, who appears to have been involved because of his longstanding interest in astronomy (Hufbauer 2006, p. 38)[60].

In that same June 1938, Oppenheimer had moved from Berkley's Physics department to Caltech, as he used to do every year, and by that time his interest in the problem of stellar energy had ripened up to the point that together with William A. Fowler, working at Caltech, and Rudolf Minkowski (Carnegie Institution of Washington, Mount Wilson Observatory), he had organized a symposium dedicated to nuclear transformations and their astrophysical significance within the annual summer meeting of the Astronomical Society of the Pacific that was held in San Diego, California, on June 22–23, as a joint session with the American Physical Society (Anonymous 1938).

Oppenheimer was scheduled to give a talk on 'The physical problem of stellar energy', while Minkowski discussed 'The Composition of Stellar Atmospheres'[61].

---

[60] Volkoff became a graduate student of Oppenheimer in 1936 and between 1938 and 1939 he was completing his thesis on 'The equilibrium of massive neutron cores'. Because of his Russian origin, and his proficiency in his native language, he became an important bridge between the scientific communities of East and West during the cold war years (Volkoff 1990).

[61] Minkowski, whose uncle Hermann had been the famous mathematician at Zurich and Göttingen, had done his doctoral studies in Breslau on spectroscopic problems. After a year

William Fowler, who was working with Charles C. Lauritsen at Caltech's Kellogg Radiation Laboratory, endowed with a large high voltage X-ray tube accelerating charged particles, gave a talk on 'Nuclear Reactions as a Source of Energy'[62].

Working in a laboratory that was at the cutting edge of nuclear physics, Fowler thus found himself as one of the founders of the emerging science of nuclear astrophysics[63].

Fowler had actually followed Oppenheimer's course on theoretical nuclear physics at Caltech: "that was really one of the highlights, because Robert was an excellent teacher, and he knew what was going on in nuclear physics". Oppenheimer was also deeply involved in the research activities of the Kellogg's laboratory since 1933, when Lauritsen had been able to produce artificial neutrons with accelerators and played an enormous role in teaching them the theoretical implications of their results: "he understood so much more completely than either Charlie or I, or even Tolman, the meaning of what we were doing [...] he understood all the quantum mechanics and special and general relativity in a very deep way [...] He was able to translate what we were finding in the laboratory into useful contributions to physics [...] If it hadn't been for Oppenheimer, I think we would have missed [laughter] practically all of the significance of what we were doing [...] Robert almost certainly was the first one to tell us that Bethe had pointed out the importance of these reactions in the sun and other stars" (Fowler 1983a,b)[64].

This last sentence is suggesting that the idea of the symposium was most probably triggered by Oppenheimer himself. What is to be emphasized is that the title of the symposium, 'Nuclear transformations and their astrophysical significance' – for the first time explicitly connecting nuclear physics and astrophysics – represented a further step along the road of an integration of the two scientific communities, someway inaugurated by the fourth Washington meeting.

In the meantime, Bethe, together with Gamow and Teller's student Charles L. Critchfield, published in August the already mentioned paper addressing the proton-proton reactions into deuterons and developing a quantitative scheme of a theory for stellar energy production (Bethe and Critchfield 1938). This paper was clearly discussed with Oppenheimer, who was mentioned in a footnote. The same "interesting discussion of these questions" were also acknowledged in a short note submitted by Oppenheimer and Serber on September 1 (Oppenheimer and Serber 1938). They started acknowledging Bethe and Critchfield's recent work which "could be made to account successfully for the main sequence stars", but not for the enormous energy output of very massive stars such as the red giant Capella, that had a much lower density and temperature than the Sun. In his talk at the San Diego Meeting, Oppenheimer had already presented "the details of the theory of the possible nuclear changes in the lighter elements and the possibility of their application to the interior of stars" suggesting "a new model with a high central concentration of neutrons"

in Göttingen with James Frank and Max Born, he moved to the University of Hamburg, as an associate professor of physics and where he became Baade's friend and collaborator. He remained there until 1935 when he was dismissed by the nazi racial laws. In 1936 Baade invited him to work at Mount Wilson Observatory, where his competence as spectroscopist was especially appreciated leading to a close and fruitful collaboration.

[62] William Fowler had got his Ph.D. in 1936 studying nuclear reactions of protons with the isotopes of carbon and nitrogen in the laboratory, the very reactions in the CN-cycle, that with Weizsäcker and Bethe's proposals were already revealing their key role within processes governing energy production in stars.

[63] For his theoretical and experimental studies of the nuclear reactions of importance in the formation of the chemical elements in the universe, Fowler would be awarded the 1983 Nobel Prize in physics, jointly with Chandrasekhar.

[64] See (Goodstein and Greenberg 1983) on the beginning of nuclear astrophysics at Caltech.

(Anonymous 1938, p. 210). But now Oppenheimer and Serber left aside the problem of energy generation for standard stars, tackled by Bethe and Critchfield: it was the neutron core, investigated by Landau and Gamow, that represented an intriguing and challenging physical system. A condensed core appeared to be still an interesting hypothesis in connection with the giant stars, where it would form after all the thermonuclear sources of energy, at least for the central material of the star, had been exhausted. And in fact, they immediately clarified that, in this regard, it still seemed of some interest Landau's "suggestion of a condensed neutron core, which would make essential deviations from the Eddington model possible even for stars so light that without a core of highly degenerate central zone could not be stable". The last provocative lines of Landau's paper had hit the bull's eye... Thus Oppenheimer decided to play the game, and started it with Serber, tackling a fundamental problem, which was essential for a discussion of the role of such a core, that is "the estimate of the minimum mass for which it will be stable". At that time Landau was like many others a victim of the Great Terror and languished in Stalin's prisons since April of that year[65].

According to Oppenheimer, Landau's evaluation appeared "to be wrong". Landau had not properly taken account of the attractive forces of gravity and had not considered the role of nuclear forces between neutrons, a force that actually was not fully understood, but on which some guess could be made, based on the phenomenology of experimental work, of which Oppenheimer was deeply aware. Both his expertise as a theoretical physicist and his constant involvement in the interpretation of experimental results were crucial in what will be now discussed. In investigating the stability of such a core, Oppenheimer and Serber were reasoning on a large assembly on nuclear particles, confined by gravitational forces, comparing this system with an actual nucleus: "The question of the actual stability of core models thus involves a consideration of the contribution of nuclear forces to the core-binding. The forces which must be known are those acting between a pair of neutrons; and no existing nuclear experiment or theory gives a complete answer to this question". Based on different assumptions on the nuclear forces also derived from investigations by Critchfield and Teller (Critchfield and Teller 1938)[66], they concluded that "even in the heaviest stars no core will be formed until practically all sources of nuclear energy have been, at least for the central material of the star, exhausted". The arguments given did show that the nuclear forces considered precluded the existence of a core for stars with masses comparable to that of the sun. It was thus clear that Landau's idea – originally inspired by the old Milne's proposal – that a large neutron core could be tucked away in stars like the sun keeping it hot – was definitely wrong. But this did not rule out the possibility of neutron cores in larger stars. Oppenheimer's investigations definitely shifted the attention to the possibility that such a core could form only when "practically all sources of nuclear energy have been, at least in the central material of the star, exhausted".

Oppenheimer and Serber had thus showed that one cannot build a viable model of the Sun with its energy coming from a neutron core, and later Bethe definitely showed that reactions of carbon and nitrogen with protons are "the most important source of energy in ordinary stars", so that interest in such stellar model declined.

[65] Landau's dear friend Matvey Bronstein had already been killed. Only the following year, after a courageous intervention of Pyotr Kapitsa, who wrote a letter directly to Molotov and Stalin threatening to quit the Institute for Physical Problems, Landau was released. He would have barely survived even a short period of further imprisonment. As promised by Kapitza, he was the right person to solve the mystery of superfluidity, an achievement for which Landau was awarded the Nobel Prize for Physics in 1962.

[66] The authors thanked Bethe, Fermi, Gamow and Oppenheimer "for helpful discussions".

However, Bethe himself, in his 1939 paper (Bethe 1939) left the problem of energy production in giant stars open. In that case it seemed rather difficult to account for the large energy production by nuclear reactions (Bethe 1939, p. 450-451): "The only other source of energy known is gravitation, which would require a core model for giants". In this regard he cited (Landau 1938), but according to Bethe, he got this suggestion also by Gamow in a letter (Bethe 1939, footnote 41)[67].

15 General relativity officially enters the stage of compact stars: Tolman and Zwicky

In the meantime, during that same summer 1938, Zwicky had again entered the scene with a follow up of his proposal about the existence of neutron stars (Baade and Zwicky 1934a) (Baade and Zwicky 1934b). At that time very little was known definitely about supernovae and it seemed certainly premature to discuss in any detail the formation of neutron stars as a possible cause for supernovae. However, since 1934 Zwicky had initiated with Baade the first systematic sky survey, and confirmed that a number of historical novae were indeed supernovae. Through the discovery with the 18-inch Schmidt telescope on Palomar mountain of eight supernovae, the existence of supernovae as a new special class of temporary stars might "be regarded as established beyond reasonable doubt " (Zwicky 1938a, p. 727)[68]. These new observations, that unquestionably confirmed the existence of violent events in the universe, urged Zwicky to pick up again the topic. Now that neutron cores were being discussed quantitatively by influential scientist like Gamow, Landau and also by Oppenheimer, based on nuclear physics, he felt that such theoretical investigations might well apply to his old idea of a collapsed neutron stars as remnants of supernovae explosions. It is to be remarked that Oppenheimer himself did not cite Baade and Zwicky's 1934 article. Probably, he wanted to take a distance from such issues that on one side had not affected astronomers' interests being too speculative and had also been completely ignored by physicists because they definitely lacked a physical base. On the other hand, as we have seen, Oppenheimer's interest into the matter was aroused within a completely different context, much more related to the development of the neutron core idea as a support to the problem of stellar energy, with no apparent relationship to Zwicky's remnants of catastrophic explosions. Only vague hints connecting the two scenarios were actually existing up to that moment.

In any case, Zwicky made a brand new very bold attempt along this path employing general relativity. Already during the June San Diego meeting, he presented a talk entitled "On neutron stars" (Zwicky 1938b), where he focused on collapsed neutron stars as representing "*states* of *lowest energy* that matter may assume without being completely transformed into radiation". The very rapid "transformation of stellar matter into the neutron state" might provide an explanation to the "stupendous rate at which energy was liberated in some of the recently observed super-novae". The abstract also mentioned the old Eddington's argument related to gravitational red-shift

[67] It appears that Fermi, too, who always had a great interest in astrophysical issues, speculated on the problem suggesting that normal stars with neutron cores would have the luminosity and spectral characteristics of red supergiants. The circumstance was mentioned by Kip Thorne (Thorne 1989) who added that "nobody seems to have built detailed models of such stars and verified Fermi's suggestion until the work of Thorne and Żytkow". Starting from 1976, Kip Thorne and Anna Żytkow wrote in fact a series of several papers devoted to the question: What are the possible equilibrium states for a star consisting of a massive nondegenerate envelope surrounding a degenerate neutron core?

[68] See list of published articles on these discoveries in footnote 3 of this article.

originating on the surface of dense stars as white dwarfs that he now presumed might be observable also in super-novae and in their remains, neutron stars, where the phenomenon would be much more pronounced. Last but not least, Zwicky made a second stronger connection to general relativity, according to which "the mass of a star of given density cannot surpass a certain critical value (Schwarzschild limit)". He then mentioned the energy that might be liberated at this limit "because of gravitational packing" as being $0.58Mc^2$, where M is the "*proper mass* of the star". The derivation of this result, explained Zwicky, had been obtained in discussion with Richard C. Tolman, and would be communicated elsewhere. From this remark, it appears that Tolman and Zwicky had been discussing the necessity of taking into account general relativity in the case of Zwicky's 'neutron stars' at least in late spring – early summer 1938.

Zwicky discussed more in detail the big issues mentioned in the previous abstract in an article dated August 8, 1938, entitled 'On Collapsed Neutron Stars' (Zwicky 1938c). Here the divide between Zwicky's and Oppenheimer's approach – and Landau's and Gamow's as well – can be measured by the distance Zwicky took in this article from the idea of neutron stars "regarded as a giant nucleus composed of separate neutrons of precisely the same character as free neutrons". Instead Zwicky specified that he used the term neutron star "simply to designate a highly collapsed star, the average density of which is of the order of the density of matter existing inside of ordinary atomic nuclei". Therefore the neutron composition of such a star should be rather taken "as a short designation for an extended state of matter of nuclear density in which every region whose linear dimensions d are larger than about $\delta = e^2/m_ec^2 = 2.8 \times 10^{-13}$ cm is essentially electrically neutral". The paper was divided in two parts. In the first part he discussed the Schwarzschild solution and the second part dealt with the possibility of actually observing the formation of collapsed neutron stars. But Zwicky did not know enough nuclear physics and general relativity to tackle the problem in detail and rigorously, basing on such theoretical tools. Even if he had a longstanding familiarity with general relativity topic since when he studied in Zurich, also with Hermann Weyl.

In his views there were some properties of neutron stars, that appeared to support his hypothesis: First of all, "a) Cold neutron stars, according to present knowledge, represent the states of lowest energy that matter may assume without being completely transformed into radiation. b) According to the general theory of relativity, a *limiting mass* of stars exists for every given average density (Schwarszschild limit)..." (here he cited (Tolman 1934)). Zwicky provided without proof the "energy liberated because of gravitational packing": $E = (1 - 4/3\pi)Mc^2 = 0.58Mc^2$ (where M is the proper mass of the star), and mentioned as well the existence of a limiting mass ('$M_L = 6.4 \times 10^{34}$ g') for an average density $\rho = 10^{14}$ g/cm^3, announcing that "The derivation of these results which was obtained in discussion with Professor R.C. Tolman will be communicated in a joint paper with Professor Tolman". However, there is no trace in the published literature of such joint paper, but as a matter of fact Zwicky had discussed the problem with Tolman, his colleague at Caltech, at least before June, according to the announcement made in the abstract of his talk at San Diego (Zwicky 1938b).

These circumstances, on the other hand, are clearly confirmed in a later paper by Tolman ('Static Solutions of Einstein's Field Equations for Spheres of Fluid') where the latter, in an unusually long footnote (Tolman 1939a, footnote 2, p. 365) acknowledged that: "My own present interest in solutions of Einstein's field equations for static spheres of fluid is specially due to conversations with Professor Zwicky of this Institute, and with Professor Oppenheimer and Mr. Volkoff of the University of California, who have been more directly concerned with the possibility of applying such solutions to problems of stellar structure". We will soon come back to the crucial connection between Tolman and Oppenheimer. Tolman then continued with his description about

Zwicky's recent attempt to introduce general relativity in the treatment of his neutron stars: "Professor Zwicky [...] has suggested the use of Schwarzschild's interior solution for a sphere of fluid of constant density as providing a model for a 'collapsed neutron star"'. Tolman then expressed the hope that "the considerations given in this article may be of assistance in throwing light on the questions that concern him".

Applying to his case an argument which had already been used within Newtonian gravity by John Michell and by Pierre-Simon Laplace during the 1780s (Israel 1987) (and which was later recalled by Eddington both in 1926 (Eddington 1926, p. 6) and during his controversy with Chandrasekhar), Zwicky concluded his paper remarking that (Zwicky 1938c, p. 523–525): "A star which has reached the Schwarzschild limiting configuration must be regarded as an object between which and the rest of the universe practically no physical communication is possible [...] It is, therefore, impossible to observe physical conditions in stellar bodies which have reached the Schwarzschild limit. It should, however, be possible to observe stellar bodies in stages intermediate between the ordinary configurations and the collapsed configurations of limiting mass just described, provided that such are accessible". As a consequence, pointed out Zwicky: "If supernovae are transitions from ordinary stars into neutron stars, the observation of light-curves and spectra of supernovae should furnish us with direct evidence of the neutron-star hypothesis [...] the surface of the central star of a supernova should be exceedingly hot, the acceleration of gravity very high, the light coming from this surface should be subject to enormous gravitational redshifts"[69].

The redshift effect had been mentioned by Ernest J. Öpik in his theory of giant stars (Öpik 1938, p. 3), as a phenomenon which "may asymptotically tend to reduce the luminosity of a superdense contracting star to zero" because in such cases stars "should possess a superdense core containing the major fraction of the mass" and the red-shift effect might be considerable. As white dwarfs had provided in 1915 a new test of Einstein's theory of general relativity well outside the solar system, the neutron-star hypothesis, in conjunction with observations on supernovae might now lead to a further and far-reaching test of the general theory of relativity, in two different astrophysical situations: Zwicky's neutron stars and neutron cores in giant stars. A connection between these still theoretical astrophysical entities was thus established, and Zwicky's old speculations were beginning to transform into a more tangible reality.

16 The Tolman-Oppenheimer-Volkoff equation

A specially relevant question for this narrative is related to the strong personal and scientific friendship between Oppenheimer and Tolman, going back to the period 1929–1930, when Tolman was embarking on his project on the connection between general relativity and thermodynamics following the cosmological issues arising from Hubble's discovery of the distance-redshift relation for galaxies and especially connected to Lemaître's proposal of an expanding universe. Being a physical chemist, but with a strong interest in astronomy and relativity, as early as 1922 he had also investigated the possibility of explaining the relative abundances of hydrogen and helium through chemical equilibrium reactions in what Helge Kragh (Kragh 2013, p. 43) has duly termed "a pioneering contribution to nuclear astrophysics".

That same 1922 he had moved to Caltech. His early interest in the theory of relativity, later led him to tackle the cosmological implications of general relativity that culminated in the publication of his seminal book *Relativity, Thermodynamics,*

[69] In this regard, Zwicky is mentioning Minkowski's forthcoming huge contribution on the spectra of supernovae (Minkowski 1939).

and Cosmology a most cited text up to post-war years[70]. Tolman had always kept contacts with astronomers working at Mount Wilson Observatory near Pasadena, in particular with Hubble, with whom he published a paper on the nature of 'nebular redshifts' (Hubble and Tolman 1935). Tolman was thus a very peculiar figure, whose wide interests integrated experimental, observational and theoretical aspects. Working at the boundary of different fields, he became crucial during the summer-fall 1938 in reorienting Oppenheimer's interests towards a new perspective for investigations on the neutron core, from which most of the nuclear stellar energy mechanisms had been stripped out, apart from some special situations. In this regard, the open questions that could be tackled were related to the evolution and in particular to the final fate of stars. In discussions with Tolman, the nature of the neutron core as a compact physical object definitely emerged: it was clear that it had to be tackled using general relativity. In the case of white dwarfs it was not so compelling, but now to really explore the behaviour and eventually ultimate fate of a superdense assembly of neutrons, Einstein's theory could not be avoided.

In the above mentioned footnote of his paper 'Static Solutions of Einstein's Field Equations for Spheres of Fluid' (Tolman 1939a) in which he acknowledged discussions with Zwicky, Tolman outlined the respective research paths: "Professor Oppenheimer and Mr. Volkoff have undertaken the specific problem of obtaining numerical quadratures for Einstein's field equations applied to spheres of fluid obeying the equation of state for a degenerate Fermi gas, with special reference to the particular case of neutron gas. Their results appear elsewhere in this same issue. My own solutions of the field equations, as given in the immediately following, can make only an indirect contribution to the physically important case of a Fermi gas, since it will be seen that they correspond to equations of state which cannot be chosen arbitrarily. My thinking on these matters has, however, been largely influenced by discussions with Professor Oppenheimer and Mr. Volkoff, and it is hoped that the explicit solutions obtained will at least assist in the general problem of developing a sound intuition for the kind of results that are to be expected from the application of Einstein's field equations to static spheres of fluid"[71].

Tolman's paper, as well as the new contribution written by Oppenheimer in collaboration with his student Volkoff, 'On Massive Neutron Cores' (Oppenheimer and Volkoff 1939), had been actually received in the *Physical Review* the same day, January 3, 1939. Both papers appeared in the same February 15 issue and contained the derivation of the equation of hydrostatic equilibrium for a spherically symmetric star in the framework of general relativity, since then called the Tolman-Oppenheimer-Volkoff equation.

Oppenheimer and Volkoff took up the problem where Chandrasekhar and von Neumann had left it: they studied "the gravitational equilibrium of masses of neutrons,

---

[70] During the mid 1930s Robert Marshak and other students of Columbia College at Columbia University, New York, including Julian Schwinger, Herbert Anderson, Norman Ramsey and Henry Primakoff, all unsatisfied by the too formal approach they were taught, "wanted to learn the physics of relativity" and thus formed the Undergraduate Physics Club lecturing to one another (Marshak 1970). They discovered Tolman's book "which was very physical" and Marshak remembered that "they went through that book very thoroughly".

[71] See correspondence between Oppenheimer and Tolman, courtesy of Caltech archives: Tolman to Oppenheimer, October 19, 1938 (discussing the "paradoxical character of the Schwarzschild interior solution"); Tolman to Oppenheimer, November 9, 1938 (about having found "two more possible solutions for the gravitational field of a static sphere of perfect fluid"); Oppenheimer to Tolman, from Berkeley, no date. See also (Tolman 1939b) (Tolman 1939c), presenting a series of analytical solutions of Einstein's equations allowing to better understand the origin of the limiting mass, that Tolman discussed in two subsequent papers on the *Astrophysical Journal*.

using the equation of state for a cold Fermi gas, and general relativity". In the view, which seemed plausible by that time, that the principal sources of stellar energy are thermonuclear reactions, at least in main sequence stars, then the limiting case considered by Landau in 1932 again became of interest in the discussion of what would eventually happen to a normal main sequence star *after all the elements available for thermonuclear reactions are used up*. Landau had showed that for a model consisting of a cold degenerate Fermi gas, a mixture of electrons and nuclei, there exist no stable equilibrium configurations for masses greater than 1.5 solar masses, all larger masses tending to collapse (Landau 1932). Chandra had clearly highlighted the importance of such issue, when he remarked that the life history of a star of small mass must be essentially different from that of a star of large mass (Chandrasekhar 1934b, p. 377): the latter cannot pass into a white dwarf stage and "one is left speculating on other possibilities".

When gravity becomes the sole and key governing force, a sufficiently massive star collapses under its own gravity, but at that time this was not felt as a fundamental key problem in astronomy and astrophysics, even if the collapsing process had been an ingredient of astrophysical theorizing on the structure of stars. But the growing role of nuclear processes had in the meantime completely transformed the issue of the limiting mass into a problem related to nuclear matter at densities beyond those found inside a nucleus.

Both Landau and Gamow had recently suggested that in sufficiently massive stars after all the thermonuclear sources of energy, at least for the central material of the star, have been exhausted a condensed neutron core would be formed (Gamow 1937, p. 234) (Landau 1938). Oppenheimer and Serber, taking into account some effects of nuclear forces had made a reasonable estimation of the minimum mass for which such a core would be stable (approximately 0.1 solar masses) (Oppenheimer and Serber 1938). A neutron core with a mass less than about 0.1 solar masses would disintegrate into nuclei and electrons. The gradual growth of such a core, with the accompanying liberation of gravitational energy, had been suggested by Landau, and in this connection it seemed now interesting "to investigate whether there is an upper limit to the possible size of such a neutron core".

Landau had found un upper limit of about 6 solar masses, beyond which the core would not be stable but would tend to collapse. Two objections might be raised against this result: "*One is that it was obtained on the basis of Newtonian gravitational theory while for such high masses and densities general relativistic effects must be considered* (emphasis added)". The second one was related to the assumptions used by Landau for the Fermi gas, now that the theory had to be applied to the case of a neutron gas. They thus wanted to establish "*what differences are introduced into the result if general relativistic gravitational theory is used instead of Newtonian and if a more exact equation of state is used* (emphasis added)"[72]. Chandra had used a Braunschweiger calculator to compute the white-dwarf structure, now Volkoff used a Marchant for the numerical integrations of some equations that could not be carried out analytically (Oppenheimer and Volkoff 1939, p. 377–378), as Tolman was trying to do for some specific cases. It was the beginning of what became known as computational relativity.

They found that for a cold neutron core (Oppenheimer and Volkoff 1939, p. 380) "there are no static solutions, and thus no equilibrium, for core masses greater than

[72] Tolman's *Relativity, Thermodynamics and Cosmology*, of which especially cited were pp. 239–247, provided the theoretical basis for the discussion of the general relativistic treatment of the equilibrium of spherically symmetric distributions of matter, and the subsequent treatment of the special ideal case of a cold neutron gas. (Chandrasekhar 1935a) is cited, too.

$m \sim 0.7 M_{\odot}$ [...] Since neutron cores can hardly be stable (with respect to formation of electrons and nuclei) for masses less than $\sim .1 M_{\odot}$, and since, even after thermonuclear sources of energy are exhausted, they will not tend to form by collapse of ordinary matter for masses under $1.5 M_{\odot}$ (Landau's limit), it seems unlikely that static neutron cores can play any great part in stellar evolution". As this limit was lower than the Chandrasekhar mass limit of white dwarfs, $1.44 M_{\odot}$, their limit appeared to create difficulties with the formation of neutron cores in ordinary stars.

Moreover, they added that "the question of what happens, after energy sources are exhausted, to stars of mass greater than $1.5 M_{\odot}$ still remains unanswered"[73].

The conclusion was that there would then seem to be only two answers possible to the question of the final behavior of very massive stars: either the equation of state they had used failed to describe the behavior of matter at density higher than nuclear density (so that their extrapolation of the Fermi equation of state could "hardly rest on a very sure basis"), or the star would "continue to contract indefinitely, never reaching equilibrium". Both alternatives, concluded the authors required "serious consideration".

They were beginning to lay a theoretical framework for investigating the fate of collapsing stars, even if many doubts persisted about the equation of state of highly compressed nuclear matter, that is, the extent to which matter at supranuclear density might successfully resist further compression.

A series of theoretical consideration about nuclear forces, even in the case of $\rho > 10^{15}$ g/cm^3 having the extreme effect of making $p = \frac{1}{3}\rho$ in such a 'critical' core, led them to conclude that it seemed likely that their limit of $\sim 0.7 M_{\odot}$ was "near the truth". This limit would be modified by future developments after the war, but conceptually it confirmed the existence of the mass limit within their theoretical frame[74].

However, it appeared that "for an understanding of the long time behavior of actual heavy stars a consideration of non-static solutions must be essential". Among all spherical non-static solutions one would hope "to find some for which the rate of contraction, and in general the time variation, become slower and slower, so that these solutions might be regarded, not as equilibrium solutions, but as quasi-static". But as a final conclusion to the discussion they stated that "For high enough central densities it is no longer justified to neglect even a very slow time variation; and the singular solutions which presumably represent very massive neutron cores cannot be obtained unless this is taken into account. *These solutions are now being investigated* [emphasis added]".

This new chapter of stellar structure differed from the preceding ones because, contrarily to what had happened with white dwarfs, all these models were derived

[73] These startling results were already announced by Volkoff at the annual meeting of the APS, held at UCLA on December 19, 1938 ('On the equilibrium of massive neutron cores') (Volkoff 1939a): "No physically plausible modifications of the equation of state seem essentially to alter this conclusion, or to change radically the order of magnitude of M_2 [$M_2 \sim 0.75 M_{\odot}$]".

[74] The equation of state $p = \frac{1}{3}\rho$ had been used by von Neumann's in his 1935 notes "Static solutions of Einstein field equations for perfect fluid with $T^{\rho}_{\rho} = 0$" and 'On relativistic gas-degeneracy and the collapsed configurations of stars' (von Neumann 1961, pp. 172–173). It appears that information about von Neumann's results continued to circulate within the scientific community even after the war. At the beginning of chapter 3 of Wheeler and collaborators' contribution to the first Texas meeting (Harrison et al. 1965) ('Hydrostatic equilibrium and extremal mass-energy'), in citing (Tolman 1934) as well as Oppenheimer and Volkoff's article, they wrote: "We have been told that in unpublished work at Cambridge in 1935 John von Neumann integrated the general relativity equation of hydrostatic equilibrium for the special case $p^* = \rho^*/3$".

as purely theoretical constructs, without any observed astronomical objects known at that time to which they might actually apply. These 'stellar neutron cores' were developed within a completely different research strand and far from any connection with Zwicky's 'neutron stars'.

An attempt had been made by Zwicky to provide a theoretical basis to the idea that neutron stars should be born in supernovae explosions, and to relate it to observational astronomy through the redshift of supernova spectra investigated by Minkowski. Interest in these issues was beginning to connect the two aspects, but a large divide still existed between the different cultures. Zwicky's new contribution 'On the theory and observation of highly collapsed stars' (Zwicky 1939), of April 1939, was pursuing a new strategy investigating "the general relativistic solution given by Schwarzschild of the problem of a homogeneous sphere of constant density" and "the possibility of actually observing the formation of collapsed neutron stars". In stating that the hypothesis of the formation of a neutron star "would run into serious difficulties if one should attempt to retain the classical theory of gravitation", Zwicky specified that "in the theory of neutron stars it is necessary to introduce general relativistic effects", according to which he interpreted the redshift in the spectrum of the supernova IC 4182 as a general relativistic gravitational redshift (Zwicky 1939, p. 727), and estimated some of the physical characteristics of the central star of a supernova one year after maximum brightness: radius (100 km), average density (10^{12} g/cm^3) and temperature (greater than 5×10^6 degrees). In making a distinction between the rest mass and the gravitational mass he was able to estimate the binding energy of a neutron star of mass M, and thus evaluate how much energy could be released during the core-collapse of massive stars.

Zwicky, was basing his theoretical investigations on new spectral studies of two bright supernovae (IC 4182 and NGC 1003) performed by his colleague Rudolf Minkowski at Mount Wilson Observatory (Minkowski 1939), that in his eyes fully justified a more detailed examination of the neutron star proposal[75].

But it was again Tolman who inspired investigations towards the application of the general theory of gravitation. In the concluding lines Zwicky thanks him for discussions during which "many of the results given in the first part of this paper were derived", but no mention of Oppenheimer's papers with his collaborators can be found, apart from an article by Volkoff[76].

Among the special reasons for which the study of supernovae might "eventually prove to be of considerable interest", stressed Zwicky in the concluding lines of his new paper, the following was to be singled out: "If the neutron star hypothesis of the origin of supernovae can be proved, it will be possible to subject the general theory of relativity to tests which according to the considerations presented in this paper deal with effects which in order of magnitude are large compared with the tests so far available". Apart from mentioning the possibility that cosmic rays originate in supernovae as an added incentive for pursuing such investigations (Zwicky 1939, p. 743), Zwicky made a further startling statement, that clarifies how deeply aware he

[75] Minkowski made a very detailed discussion of his observation basing on Zwicky's assumption that the observed red shift might be caused "by the increase of gravitational potential at the surface of a collapsing star" (Minkowski 1939, p. 208) and concluded that two different explanations of the red shift, as either a gravitational effect or as Doppler effect, appeared possible. If a more detailed study of the radiative equilibrium did not lead to a rejection of one of these conceptions, a decision might be brought about by a theory of supernovae which could explain the similarity of the red shift in different supernovae.

[76] Volkoff is investigating the difference in behavior between solutions of Einstein's field equations with infinite central pressure (that Schwarzschild had dismissed as physically inadmissible because of this singularity) and the Oppenheimer-Volkoff cold neutron gas model leading to an upper limit on the size of a static sphere (Volkoff 1939b).

had become of the possible implications of his original fascinating idea: "*The general theory of relativity, although profound and exceedingly satisfactory in its epistemological aspects, has so far practically not lent itself to any very obvious and generally impressive applications. This unfortunate discrepancy between the formal beauty of the general theory of relativity and the meagerness of its practical applications makes it particularly desirable to search for phenomena which cannot be understood without the help of the general theory of relativity* (emphasis added)".

This statement is the best possible coeval comment to the Oppenheimer-Volkoff paper, in which for the first time general relativity was deliberately applied to tackle the problem of a compact astrophysical object, and that in Zwicky's words probably acquired a meaning going well beyond Oppenheimer's own intentions and at the same time represents the best introduction for the final phase of Oppenheimer's efforts in this direction in which he explored with his student Snyder the final fate of a collapsing stellar neutron core.

Tolman later had an important role within the Manhattan Project, and, as revealed by Serber himself, it is remarkable that he was the first to put forward the idea of implosion as a way of compressing matter and triggering the explosion process of nuclear weapons (Serber 1992, p. xxxii). The similarity between stellar implosion-explosion problems and the building of nuclear and thermonuclear weapons would in turn attract the attention of a new generation of physicists deeply involved in these activities during World War II – notably John A. Wheeler and Ya B. Zeldovich – towards the connections between general relativity and the interior of a compact star. Such similarity also suggested the adaptation of bomb design codes to simulate stellar implosions (Colgate and White 1966).

17 From the neutron-core to the neutron star. The last chapter of Oppenheimer's trilogy

Officially Zwicky was completely ignored, not being cited in the Oppenheimer-Volkoff paper, that deliberately took a distance from 'neutron stars', considered a fruit of Zwicky's speculations, without any clear physical content[77]. On the other hand, Chandrasekhar's classic white-dwarf work, too, was scarcely credited by Oppenheimer – as well as by Landau – thus favoring an interpretation in the direction of a divide between fields and scientific styles. From Oppenheimer's point of view, all that was restricted to the theory of relativistic electron degeneracy as needed for a full investigation of the white-dwarf problem, a very specific astrophysical problem also having an interest for astronomical observations. In his articles he explicitly mentions (Landau 1932) whose investigations on the physical nature of the equilibrium of a given mass of material had been performed using "a model consisting of a cold degenerate Fermi gas". Following Landau, but with improved knowledge on nuclear matter, Oppenheimer and collaborators used astrophysics, as a realm providing a *physical system* at extremely high density for their investigations about its stability and the existence of an upper limit to its possible size (Oppenheimer and Serber 1938; Oppenheimer and Volkoff 1939).

In any case, because of his commitment to subnuclear processes generated by cosmic-rays and their relationship with the emerging modern particle physics, Oppenheimer could not ignore Zwicky's plausible speculations on supernovae explosions

[77] It is to be remarked that, still in 1964, three years before the discovery of pulsars, when it was "accepted with a reasonable assurance" that the supernova explosion of a star is first triggered by the collapse of its core, Hong-Yee Chiu stated in a review on 'Supernovae, neutrinos, and neutron stars' that the possibility that "neutron stars may be the remnants of supernovae has so far been accepted only with skepticism" (Chiu 1964, p. 368).

as sources of high energy particles, even if such investigations were in turn connected with further speculations on a collapsed compact astrophysical object made up of neutrons.

From the outlined arguments presented up to now, it is quite evident that Oppenheimer's work with his collaborators was being carried on within a wider related context, and especially within the ongoing and spreading interest towards superdense neutron cores in stars. However, Zwicky's idea was gaining momentum as one can see from Gamow's article 'Physical possibilities of stellar evolution' (Gamow 1939a), submitted in November of 1938, where he outlined a picture of stellar evolution on the basis of the Bethe-Weizsäcker theory. Apart from attempts to explain the energy production in red giants, he discussed the "contractive stage where the energy liberation is purely gravitational" and "the possibility of neutron-core formation in heavier stars, in application to the explosion phenomena observed in supernovae"[78]. Gamow recognized that for stars with large masses, no stable finite state does exist and so they "must undergo continuous unlimited contraction". However, he saved the situation proposing that "such a process will never continue indefinitely because, since all stars possess an angular momentum, the centrifugal forces will soon become large and will, most probably, cause the breaking of such a massive star into several smaller pieces with the masses below the critical value. These pieces will then continue to exist indefinitely in the form of white dwarfs". Thus, according to Gamow, the existing white dwarfs did not represent a finite stage of evolution of a single star, but must be considered the fragments resulting from the explosion of heavy stars.

In extending his astrophysical investigations, Gamow, in collaboration with Teller (Gamow and Teller 1939), also addressed the problem of the origin of great nebulae within the framework of an expanding universe: "The type of expansion necessary for the formation of nebulae indicates that space is infinite and unlimitedly expanding". In the last section, they considered the cosmological consequences which they discussed from the form of the fundamental (Friedmann–Lemaître) equation for the expanding universe as given by Richard Tolman in his textbook of 1934 (Tolman 1934). They deduced that the nebulae are the largest assemblies of matter which can be kept together by gravitation against the dispersing effect of the random velocities of the stars. Moreover, they pointed out that the mutual velocities of neighbouring nebulae are of the same order as the random velocities of stars in a nebula. This supported the theory that all nebulae originated from the same very limited region of space: "It seems much more likely that such an odd occurrence as our planetary system might be formed in the original highly condensed state of the universe than in the present dilute one". Since 1937 Gamow had actually shifted his interests from nuclear physics proper and had decided to give a graduate course at George Washington University on general relativity and its connections with cosmology (Hufbauer 2009, p. 21). This work is an early hint of Gamow's developing research interest in cosmology, which, would be officially inaugurated by the Eighth Washington Conference on theoretical physics devoted to 'Stellar Evolution and Cosmology' held in 1942 (Gamow and Fleming 1942), where he spoke about his new ideas on cosmological nucleosynthesis. Immediately after the war, Gamow would fully merge cosmology with his wide competence as a nuclear astrophysicist, formulating what became successfully known as the big-bang theory of the universe further developed with his collaborators Ralph A. Alpher and Robert Herman who predicted in a later paper that the cooled remnant of the hot early phases should be present in the Universe today and estimated that the temperature of this thermal background should be about 5 K (Alpher 2012).

[78] Apart from Baade and Zwicky and Chandrasekhar, Gamow cited Sterne and Hund, so that all the implications contained in these pioneering works were beginning to be fully appreciated by this time.

During that spring-summer 1939, Tolman himself did not resist the temptation of examining the connection between the stability of stellar models and the origin of novae (Tolman 1939c), in an article in which he acknowledged discussions with Oppenheimer. As he remarked in the introduction: "With the help of such studies one might ultimately hope to understand not only the existence of the great majority of stars in steady states and of a limited classes of stars in pulsating states producing variations in luminosity, but also "the existence of some – perhaps nearly all – stars in states that can lead to the occasional formation of novae, or to the related case of supernovae".

In his March 1939 paper, Bethe, too, had tackled the problem of the last stages of stellar evolution: "It is very interesting to ask what will happen to a star when its hydrogen is almost exhausted. Then, obviously, the energy production can no longer keep pace with the requirements of equilibrium so that the star will begin to contract [...] In the white dwarf state, the necessary energy production is extremely small so that such a star will have an almost unlimited life [...] For heavy stars, it seems that the contraction can only stop when a neutron core is formed [...] However, these questions obviously require much further investigation" (Bethe 1939, p. 456)[79].

Novae (and thus supernovae) and white dwarfs, were definitely an issue at stake during 1939, and in fact a specific conference was organized in July in Paris, at the Collège de France, in order to study these two categories of stars, which were in the foreground of the current research. On that occasion, Chandra (Chandrasekhar 1941) stated again his conclusions regarding the limiting mass based on relativistic degeneracy and connected his theory to the supernova phenomenon suggesting that a star which has exhausted its nuclear fuel and whose mass was exceeding such an upper limit would collapse with a huge release of gravitational energy. Such energy would in turn fuel the explosion, leaving a very compact neutron core.

That same July, Félix Cernuschi, working at MIT with Sandoval Vallarta, discussed in three articles the problem of supernovae, 'neutron-core stars' and the origin of cosmic rays, in the new perspective of the discovery of fission (Cernuschi 1939a,b,c). The titles of the three articles ('Super-Novae and the Neutron-Core Stars', 'A Tentative Theory of the Origin of Cosmic Rays' and 'On the Behavior of Matter at extremely high temperatures and pressures') are a clear indication of the constant interest for the dense neutron cores from the point of view of the quick development of nuclear physics, but also in connection with a growing interest for Zwicky's theory of supernovae as sources of cosmic rays and as stellar objects representing the transition of an ordinary star into a neutron star. But his first objection to Zwicky's theory was that "an ordinary star is a gaseous star without neutron core" and thus it appeared difficult to imagine how a supernova could result from such a transformation

[79] In coming back from the 1938 Washington conference organized by Gamow and Teller, Bethe was very excited, and thus triggered his PhD student Robert Marshak's interest in astrophysics and especially in white dwarfs and in their energy source. In his PhD thesis ('Contributions to the Theory of the Internal Constitution of Stars') Marshak investigated in detail the state of matter in the interior of a white dwarf star and concluded that no hydrogen could be present. Under these circumstances the radius of the star is uniquely determined by its mass, according to the theory of degenerate configurations. However, in his calculations Marshak found a serious discrepancy between the theoretical radius of Sirius B ("only 5.7×10^8 cm, as compared with the observed radius of 13.6×10^8 cm): "The present investigation has at least established almost beyond question that the claim of astrophysics is in direct conflict with the claim of nuclear physics and that there really seems to be no simple explanation of the radius discrepancy for Sirius B" (Marshak 1940). Marshak's work later served as a fundamental reference for the understanding of this type of stars (Marshak 1970). Marshak's attention was then diverted from astrophysics and he started working on other problems which were more directly connected with particle physics broadly interpreted.

(Cernuschi 1939a, p. 120). He assumed instead that white dwarfs are stars with an unstable neutron core and that such instability derived from fission processes of very heavy nuclei such as uranium and thorium whose existence in the neutron core he had postulated. In this way it was possible to imagine that the same process of fission of single a giant nucleus of atomic number 10 000 would produce energies of the order of 10^{12} eV, so that , "a super-nova would not be a transition of an ordinary star into a neutron star, but would result from the explosion of the neutron core of a white dwarf" in a cascade of successive fissions (Cernuschi 1939b, p. 121). Under these assumptions it seemed also possible "to imagine a concrete physical mechanism which might underlie the production of cosmic rays". One can see here once more the pervasive influence of Lemaître's primaeval atom.

Neutrinos, too, are beginning to populate the interior of superdense cores: "if the neutrino does exist, it will be of great importance in the internal constitution of the stars, due to the fact that this particle should have an extremely high penetrability and, therefore, under certain conditions the transport of heat resulting from the neutrinos might not be negligible beside the flow of radiation". In this sense neutrinos are beginning to play a role in supernovae explosions. Cernuschi is also trying to reconcile the divide between astronomers and physicists investigating whether Landau's theory might support Zwicky's proposal.

Cernuschi's article on the *Physical Review* is followed back to back by the third paper of the Oppenheimer and collaborators' trilogy, submitted in early July 1939: 'On continued gravitational contraction.' With his student Hartland Snyder, Oppenheimer took general relativity far beyond Zwicky's possibilities and focused on how the neutron core would evolve once it became unstable (Oppenheimer and Snyder 1939).

Volkoff and Oppenheimer had already made clear that assemblies of neutrons are so compact that general relativity is no longer a small correction and can no longer be neglected because it is central to the stability of such astrophysical objects. They had been able to show that the general relativistic field equations do not possess any static solution for a spherical distribution of cold neutrons, if the total mass of the neutrons is greater than $\sim 0.7 M_\odot$, and had established that a star under these circumstances would collapse under the influence of its gravitational field. In the meantime, with Snyder, Oppenheimer had explored the process of gravitational collapse itself, where the full consequences of Einstein's theory of gravitation could be seen at work[80]. Oppenheimer and Snyder were now definitely stripping 'neutron cores' (the Oppenheimer-Volkoff "spherical distribution of cold neutrons") of any outer envelope, openly studying what were actually 'neutron stars'. Even if they did not call them as such, and only referred to 'heavy stars', made of course mainly of neutrons. These investigations were now waiting to officially enter the field of theoretical astrophysics, but had already begun to re-write the chapter of 'compact stars', up to that time only containing theorizing on white dwarfs.

The very first sentence of the abstract itself, once excluding other possible situations, left no hope for a star with a critical value of the mass: "When all thermonuclear sources of energy are exhausted a sufficiently heavy star will collapse. *Unless* fission due to rotation, the radiation of mass, or the blowing off of mass by radiation, reduce the star's mass to the order of that to the sun, *this contraction will continue indefinitely*" [emphasis added]. The concluding lines further emphasized their expectations: "this behaviour will be realized by all collapsing stars which cannot end in a stable stationary state".

[80] Already during his presentation of December 1938 at UCLA meeting, Volkoff had mentioned that nonstatic solutions for the cases of masses beyond the critical mass were being investigated (Volkoff 1939a).

Oppenheimer and Snyder found that, as seen by a distant observer, general relativity predicted that the star would asymptotically shrink to its Schwarzschild radius, light from the surface of the star would be progressively reddened, being able to escape over a progressively narrower range of angles. According to the scenario already outlined by Eddington more than a decade before (Eddington 1926, p. 6), the star would close itself off from the rest of the universe except for its intense gravitational field: "The mass would produce so much curvature [...] that space would close up round the star, leaving us outside (i.e. nowhere)". These inescapable general arguments were confirmed by the study of analytic solution of the field equations for the case that the pressure within the star could be neglected. It showed that, although the collapse would formally take an infinite time when viewed from large distance, the time measured by an observer comoving with the star would be finite as would also be the time until a distant observer would find the star to be undetectably faint as a consequence of the general relativistic effects.

That Oppenheimer and Snyder were venturing into unknown territory is someway testified also from the fact that, apart from the obvious reference to the Oppenheimer-Volkoff paper, the only citation is to an article by Tolman of 1934 (Tolman 1934a). Tolman is also thanked for "making a portion of development available". The simple scenario they had used to describe the collapse process was rather idealized and far from an actual physical model of a collapsing star. But in establishing the physical reality of a phenomenon deeply rooted in the theory of general relativity, the Oppenheimer-Snyder paper gave rise to a startling and unexpected consequence in the real world of astronomy. For the first time, Schwarzschild's purely mathematical solution to the general theory of relativity was systematically discussed within a framework related to a specific physical object.

As stressed in (Eisenstaedt 1993), these results were actually derived using the so called "dust solution", a general solution of the field equations for the case of spherical symmetry and no pressure (Lemaître 1932), and which were well known to Tolman, who had worked with him during Lemaître's stay for two months at Caltech in the early 1930s. In this remarkable contribution, whose 1933 version is cited in (Tolman 1934a), Lemaître also demonstrated that the Schwarzschild singularity is only an apparent singularity. Eisenstaedt emphasizes in his detailed discussion about Lemaître's pioneering results (Eisenstaedt 1993, p. 11), that he tackled this problem once he realized that (Lemaître 1932, p. 200): "The equations of the Friedman universe admit [...] solutions in which the radius of the universe goes to zero. This contradicts the generally accepted result that a given mass cannot have a radius smaller than [...] $2m$" (in natural units: $G = c = 1$).

It is to be emphasized that, in the last section, where Lemaître is discussing the physical interpretation of the "zero value of the radius", he is remarking that matter should find a way to avoid the vanishing of its volume and as matter is formed by stars, this would be 'manifestly impossible". Lemaître is comparing this situation to "the interior of the companion of Sirius": it appears that even for a degenerate gas nothing might oppose to such a condensed form of matter. At distances between atomic nuclei and electrons of the order of 10^{-12}, subatomic forces opposing to penetration between particles would dominate and certainly be able to stop contraction: "The universe would thus be comparable to a giant atomic nucleus". He immediately added that once the contraction is blocked, the process should restart in the opposite verse: "These solutions in which the universe is expanding and then contracting, periodically reducing to an atomic mass having the dimensions of the solar system, definitely have a poetic charm and make us think to the phenix of the legend," concluded Lemaître in the last lines of the paper.

However, even appreciating that the Schwarzschild singularity could be locally eliminated by a coordinate change, Lemaître did not provide an overall picture of collapse to a black hole. Interestingly, in 1934 Synge wrote a paper entitled 'On the Expansion or Contraction of a Symmetrical Cloud under the Influence of Gravity', in which he studied the evolution of a small cloud of particles finding that a collapse beyond the Schwarzschild singularity is possible, at least in the pressure-free case. Eisenstaedt duly remarks (Eisenstaedt 1993, p. 14) that this paper, which might be of great interest for the Oppenheimer-Snyder problem, was almost never cited, even by Synge himself, who actually later found the complete extension of the Schwarzschild solution (Synge 1950).

When the expansion of the universe was becoming an accepted phenomenon, general relativity had potentially revealed something new and quite unexpected on the universe. As Wheeler much later stressed, (Wheeler 1968, p. 6): "No test of Einstein's theory is more dramatic than the expansion of the universe itself, and none has a closer bearing on the phenomenon of collapse" well expressing that between the predicted and observed expansion and the gravitational collapse of a star "there is not one significant difference of principle". However, this flurry of interests connecting general relativity and astrophysics was interrupted by the outbreak of World War II, and the two aspects remained separated up to the end of the 1950s – early 1960s, when a novel closer alliance would be established between the general theory of relativity and the physical universe. By that time, the period of stagnation of the theory, the "low-water-mark" of general relativity (Eisenstaedt 1987a,b) had given way first to the "renaissance" (Blum et al. 2016) and then to the "golden age" of general relativity (Thorne 2003, pp. 74–80), during which relativistic astrophysics was established as a novel research field and consensus about the existence of extreme physical implications of the theory such as gravitational waves and black holes had formed.

Both the theoretical demonstration of an inescapable process such as the gravitational collapse within Einstein's theory, and the established existence of supernovae as a new class of astrophysical objects – a striking evidence for the existence of violent events in the universe – as well as the continuously evolving stage of an expanding universe, combined all together in marking the end of the 'Aristotelian vision' that had dominated astronomy for about 2000 years: the heavens as the domain of an eternal perfect harmony, contrasted with Earth as the realm of conflict and change.

Most astronomers, however, paid little attention to such reality, generally believing that in the final stage of collapse sufficient material would always be ejected to bring the mass of the resulting body down to below the Chandrasekhar limit – or to below the Oppenheimer-Volkoff limit which is the corresponding maximum mass of a neutron stellar core. The awkward character of the questions aroused from the problem is also testified by the lack of any mention of the collapse for masses beyond the limiting mass in Chandra's comprehensive textbook *An introduction to the study of stellar structure* (Chandrasekhar 1939).

However, within a year, Gamow and the young Brazilian physicist Mario Schönberg, who had studied with Fermi and Pauli, investigated for the first time the physical process of 'catastrophic collapse' from a point of view of nuclear physics (Gamow and Schönberg 1940), arguing that rapid cooling due to extensive neutrino losses by what they called, for brevity, "urca-processes" in inverse beta-decay, would result in a catastrophic failure of pressure support near the core, unable to support the weight of the overlying collapsing layers (Gamow and Schönberg 1941, p. 540)[81]. They did not mention Cernuschi's articles, and at the same time rejected Zwicky's hypothesis of the collapse as being due to the formation of a large number of neutrons

[81] They named it the the 'urca-process' because it results in a rapid disappearance of thermal energy from the interior of a star, similar to the rapid disappearance of money from the pockets of the gamblers in the Casino da Urca, in Rio de Janeiro, where they discussed the problem when they first met (Gamow 1970, p. 137).

with subsequent closer "packing" in the central regions. What they wanted was the instantaneous removal of large amounts of gravitational energy produced by contraction, in order to have a collapse "with a velocity comparable to that of 'free fall' independent of the kind of particles existing in its interiors". Processes of absorption and reemission of free electrons could lead to tremendous energy losses through neutrino emission and cause the collapse of the entire stellar body. During the last ten years, neutrinos had gained a considerably important position in nuclear physics, in spite of the fact that all the attempts at their direct observation had failed. However, exactly their capability of passing through many thousands of kilometers of matter without suffering absorption (Bethe and Peierls 1934), made them the right agent to remove the surplus energy from the interior of a contracting star, whose body was completely transparent for neutrinos. They also emphasized that (Gamow and Schönberg 1941, p. 541), "while the neutrinos are still considered as highly hypothetical particles because of the failure of all efforts made to detect them, the phenomena of which we are making use in our considerations are supported by the direct experimental evidence of nuclear physics". In speculating that neutrinos might play an important role in stellar evolution, particularly in the collapse of evolved stars, they ushered in the advent of particle astrophysics. Such a hypothesis was quite bold for the time because neutrinos, which had been proposed by Pauli in 1930, were not directly detected until the mid 1950s. The intense neutrino flux emitted during the process was dramatically confirmed by Supernova 1987A whose observation in 1987 coincided with a burst of 11 neutrinos, detected by Super-Kamiokande in Japan, by 8 further neutrinos registered independently in Ohio, and by 5 events at the Baksan Neutrino Observatory on the Caucasus mountains. The fast removal of energy due to the emission of neutrinos would induce "the collapse of the entire stellar body with an almost free-fall velocity" while rapid contraction would increase the central temperature. Stars possessing a mass larger than the critical mass would undergo a much more extensive collapse and their ever-increasing radiation would drive away more and more material from their surface: "The process will probably not stop until the expelled material brings the mass of the remaining star below the critical value". This process might be compared with the supernovae explosions, in which case the expelled gases would form extensive nebulosities such as the Crab-Nebula (Gamow and Schönberg 1941, p. 546), leaving behind a faint star, that according to its observed properties, was classified at the time as a very dense white dwarf (Minkowski 1942). This view, already clearly expressed by Chandra (Chandrasekhar 1935a,c), was destined to endure for a long time. The following year, Schönberg became Chandra's post-doc student and with him he wrote a paper in which they discussed the problem of what is the maximum mass of a star's hydrogen-exhausted core that can support the overlying layers against gravitational collapse (Chandrasekhar and Schönberg 1942). The result of their investigation was that the helium core reached the maximum mass it could attain without collapsing when just about 10% of the hydrogen had been consumed. This is known now as the Chandrasekhar-Schönberg limit. In the concluding lines they again stressed that "the supernova phenomenon may result from the inability of a star of mass greater than M_3 [upper limit to the mass of degenerate configurations] to settle down to the final state of complete degeneracy without getting rid of the excess mass" thus assuming that the final state would be that of a white dwarf, and without considering the possibility that the star might collapse to nothing.

By 1939, the problem of what happens to a compact star core made entirely of degenerate fermions (electrons and neutrons) had been studied by a handful of researchers. Oppenheimer, however, – and Tolman as well – did not recognize that what they had tackled was conceptually quite similar and had been already anticipated by B. Datt's simple but more general model, published in May 1938, in the *Zeitschrift für*

Physik (Datt 1938)[82]. Working at Presidency College, where Chandrasekar had studied and where later Abdus Salam would move his first steps as a physicist, Datt used general relativity to examine the final fate of an idealized homogeneous pressure-less spherically symmetric, massive cloud with no rotation and internal stresses, collapsing under its own gravitational attraction[83]. This classic scenario became later known as the Oppenheimer-Snyder-Datt model (OSD).

But at the moment all this sounded like an exotic problem, and nobody realized, not even Oppenheimer himself, how innovative their contribution was: on one side nuclear matter – and particle physics – were becoming essential for the description of matter at such extreme densities. On the other hand, it had become clear that such superdense objects could be described *only* within Einstein's theory of gravity: the door had been opened on the world of relativistic astrophysics. There was at least one physicist who was deeply aware of the relevance of these results. Landau, who was again free after a year of imprisonment and was working at his celebrated theory of superfluidity, added the Oppenheimer-Snyder paper to his 'Golden List', according to what Evgeny Mikhailovich Lifschitz told Kip Thorne many years later (Thorne 1994, p. 219). It even appears that "So great was Landau's influence that his view took hold among leading Soviet theoretical physicists from that day forward".

Most probably it is not by chance that in May 1939, not long after the appearance of Datt's contribution on the *Zeitschrift für Physik* – but nearly in parallel to the Oppenheimer-Snyder paper – Einstein himself submitted a contribution in which he worked out how a swarm of particles would behave as they collapsed through gravity (Einstein 1939). As stated by Einstein, this investigation arose out of discussions conducted with Howard P. Robertson, Peter G. Bergmann and Valentine Bargmann on the *mathematical* and *physical* significance of the Schwarzschild singularity, which had played a role in Zwicky's paper on the collapse of neutron stars, but especially in Tolman's article written in parallel with Oppenheimer and Volkoff's contribution, that should not have escaped the attention of Einstein and Robertson, both having a longstanding personal relationship with Tolman. In particular, during the 1930s, Robertson had worked on the problem of the Schwarzschild space-time, but he did not publish it (Eisenstaedt 1987a, p. 328). However, according to Bergmann, Einstein was not aware of Oppenheimer's papers. In his *Introduction to the Theory of Relativity*, whose first edition was printed in May 1942, with a foreword by Einstein himself mentioning the many hours spent in discussing the text, Bergmann summarized as follows Robertson's view (Bergmann 1942, p. 203–204): "Robertson has shown that, if a Schwarzschild field could be realized, a test body which falls freely toward the center would take only a finite proper time to cross the 'Schwarzschild' singularity, even though the coordinate time is infinite; and he has concluded that at least part of the singular character of the surface $r = 2m$ must be attributed to the choice of the coordinate system". At the end of the section dedicated to the Schwarzschild singularity, Bergmann introduced a short description of Einstein's article with the following clear-cut sentence: "*In nature, mass is never sufficiently concentrated to permit a Schwarzschild singularity to occur in empty space* [emphasis added]" (Bergmann 1942, p. 204). In any case, no reference to Oppenheimer's works with his collaborators can be found in Bergmann's book.

Starting from the Schwarzschild's solution of the static gravitational field of spherical symmetry, and from the vanishing of the g_{44} term of the equation, Einstein tackled the question whether it was possible "to build up a field containing such singularities with the help of actual gravitating masses or whether such regions with vanishing g_{44}

[82] It was dated Kalkutta, Presidency College, 10 September 1937.

[83] See comment to the English translation of Datt's article by Andrzej Krasiński (Krasiński 1999).

do not exist in cases which have physical reality". As a field-producing mass he chose a system formed by a great number of small gravitating particles moving freely under the influence of the field produced by them all together. The particles in Einstein's "spherical star cluster" all moved in circular orbits around a common center, and he calculated that material particle orbits could not have radii less than one and a half Schwarzschild radii, in Schwarzschild coordinates, so that the essential result of this investigation, concluded Einstein, was "a clear understanding as to why the 'Schwarzschild singularities' do not exist in physical reality. Although the theory given here treats only clusters whose particles move along circular paths it does not seem to be subject to reasonable doubt that more general cases will have analogous results. The 'Schwarzschild singularity' does not appear for the reason that matter cannot be concentrated arbitrarily. And this is due to the fact that otherwise the constituting particles would reach the velocity of light". In the concluding lines of the paper, Einstein drastically stated that it is not possible to attain the Schwarzschild radius in nature and thus the problem of the mathematical and physical significance of the Schwarzschild singularity, "quite naturally leads to the question, answered by this paper in the negative, as to whether physical models are capable of exhibiting such a singularity". As underlined by Jean Eisenstaedt within a discussion on Robertson's work on the Schwarzschild singularity, Einstein's starting hypothesis of circular orbits for his system of gravitating particles, as a matter of fact excluded the possibility of reaching the Schwarzschild singularity, because the radius of a gas cloud described by strictly circular orbits must necessarily be greater than 3/2 the Schwarzschild radius. In this sense, Einstein's article is based on a 'circular logic' (Eisenstaedt 1987a, p. 337).

Oppenheimer and Snyder had emphasized that although the Schwarzschild singularity occurring at radius $r = 2m$ is not actually a singularity, there is still a space-time singularity at the centre, where the density of the dust becomes infinite. According to Roger Penrose (Penrose 1996), "Since there is still a singularity in the Oppenheimer-Snyder collapse model (at $r = 0$), the Chandrasekhar dilemma (on the existence of a maximum mass for white dwarf stars) is not removed by their collapse picture. However many people remained unconvinced that this description would necessarily be the inevitable result of the collapse of a star too massive to be sustainable as either a white dwarf or neutron star. There were a number of good reasons for some scepticism. In the first place, the equations of state inside the matter were assumed to be those appropriate for pressureless dust, which is certainly far from realistic for the late stages of stellar collapse. Moreover, the density was assumed to be constant throughout the body. With realistic material, there are many alternative evolutions to that described by Oppenheimer and Snyder. For example, nuclear reactions set off at the centre could lead to an explosion – a supernova – which might perhaps drive off sufficient mass from the star that a stable equilibrium configuration becomes possible".

In January 1939, the Fifth Washington Conference on Theoretical physics had been held, having low temperatures as a focus for the discussion. However, as it is well known, Bohr, who had just arrived from Europe, brought with him news about the Frisch-Meitner explanation of fission as a physical process, immediately arousing an incredible excitement and putting in motion a series of events which would deeply affect the whole scientific community in connection with the dramatic developments on the world stage. On September 1, when Oppenheimer and Snyder's paper appeared in the *Physical Review*, Nazi troops marched across the Polish border. In the same issue Bohr and Wheeler, working together at Princeton, outlined an account of the mechanism of nuclear fission on the basis of the liquid drop model of nuclei (Bohr and Wheeler 1939). By the end of 1939, actual neutrons in heavy nuclei had already become the protagonists in a completely different realm, eventually leading to the building of the first nuclear reactors and the first nuclear weapons. Oppenheimer himself would be heavily involved in these efforts, heading the Manhattan Project's secret research laboratory at Los Alamos. The curtain apparently closed on the march leading to the first applications of general relativity to an astrophysical compact object, but behind the scenes new premises for a great renewal of interest in superdense matter and compact objects were laid during the war period which would eventually flourish within the new conditions provided by post-war science.

Acknowledgements. I am deeply indebted to Gordon Baym, Wolf Beiglböck, Werner Israel, Roberto Lalli, and an unknown referee, for a critical and constructive evaluation of the manuscript and for many most useful suggestions. Special thanks go to Jürgen Renn for arousing my interest in the fascinating history of astrophysics. I also gratefully acknowledge that this research has made a systematic use of NASA's Astrophysics Data System Bibliographic Services. Open access funding provided by Max Planck Society.

References

Adams, W.S. 1915. The spectrum of the companion of Sirius. *Publications of the Astronomical Society of the Pacific* **27**: 236–237.

Adams, W.S. 1925a. A Study of the Gravitational Displacement of the Spectral Lines in the Companion of Sirius. *Publications of the Astronomical Society of the Pacific* **37**: 158–158.

Adams, W.S. 1925b. The relativity displacement of the spectral lines in the companion of Sirius. *Proceedings of the National Academy of Sciences* **11**: 382–387.

Alpher, V.S. 2012. Ralph A. Alpher, Robert C. Herman, and the Cosmic Microwave Background Radiation. *Physics in Perspective* **14** (3): 300–334.

Ambartsumyan, V. and G.S. Saakyan. 1960. The degenerate superdense gas of elementary particles. *Soviet Astronomy* **4** (2): 187–201. Russian original in *Astronomicheskii Zhurnal* **37** (2): 193–206.

Ambartsumyan, V. and G.S. Saakyan. 1962a. On Equilibrium Configurations of Superdense Degenerate Gas Masses. *Soviet Astronomy* **5** (5): 601–610.

Ambartsumyan, V. and G.S. Saakyan. 1962b. Internal Structure of Hyperon Configurations of Stellar Masses. *Soviet Astronomy* **5** (6): 779–784.

Anderson, W. 1928a. Die 'Entartung' des Elektronengases im Innern einiger Sterne. *Zeitschrift für Physik* **50** (11): 874–877.

Anderson, W. 1928b. Zur Theorie von G.I. Pokrowski über die obere Grenze für die Masse eines Sternes. *Zeitschrift für Physik* **53** (7-8): 597–600.

Anderson, W. 1929a. Die Theorie von G.I. Pokrowski und die Kontraktionsenergie der Sterne. *Zeitschrift für Physik* **55** (5-6): 386–394.

Anderson, W. 1929b. Über die Grenzdichte der Materie und der Energie. *Zeitschrift für Physik* **56** (11-12): 851–856.

Anderson, W. 1936. Existiert eine obere Grenze für die Dichte der Materie und Energie? I. Teil. *Publications of the Tartu Astrophysical Observatory* **29** (1): 1–142.

Anonymous. 1938. The San Diego Meeting of the Astronomical Society of the Pacific June 22–23, 1938. *Publications of the Astronomical Society of the Pacific* **50** (296): 210–231.

Atkinson, R. D'E. and F. Houtermans. 1929a. Transmutation of the Lighter Elements in Stars. *Nature* **123** (3102): 567–568.

Atkinson, R. D'E. and F. Houtermans. 1929b. Zur Frage der Aufbaumöglichkeit der Elemente in Sternen. *Zeitschrift für Physik* **54** (9-10): 656–665.

Baade, W. and F. Zwicky. 1933. Supernovae and cosmic rays. Minutes of the Stanford Meeting, December 15-16, 1933. *Physical Review* **45** (2): 138–138.

Baade, W. and F. Zwicky. 1934a. Cosmic rays from Super-novae. *Proceedings of the National Academy of Sciences* **20** (5): 259–263.

Baade, W. and F. Zwicky. 1934b. Remarks on Super-Novae and Cosmic Rays. *Physical Review* **46** (1): 76–77.

Bates, L.F. 1969. Edmund Clifton Stoner. 1899–1968. *Biographical Memoirs of Fellows of the Royal Society* **15**: 201–237.

Baym, G. 1983. Neutron stars: the first fifty years, in P. Schofield (ed.). *The Neutron and its Applications, 1982*. IOP Conference Series 64. The Institute of Physics, Bristol and London, pp. 45–50.

Beckedorff, D.L. 1962. *Terminal configurations of stellar evolution*. Dissertation (under C. Misner). Princeton University, Department of Mathematics http://www2.physics.umd.edu/~misner/Beckedorff_metadata.htm

Bergmann, P.G. 1942. *Introduction to the Theory of Relativity*. Prentice-Hall, New York.

Bethe, H. 1939. Energy Production in stars. *Physical Review* **55** (5): 434–456.

Bethe, H. 2003. *Annual Review of Astronomy and Astrophysics* **41**: 1–44.

Bethe, H. and R. Peierls. 1934. The "Neutrino". *Nature* **133** (3362): 532–532.

Bethe, H. and C.L. Critchfield. 1938. The Formation of Deuterons by Proton Combination. *Physical Review* **54** (4): 248–254.

Bloch, F. 1928. Über die Quantenmechanik der Elektronen in Kristallgittern. *Zeitschrift für Physik* **52** (7): 555–600.

Blum, A., R. Lalli and J. Renn. 2015. The Reinvention of General Relativity: A Historiographical Framework for Assessing One Hundred Years of Curved Space-time. *Isis* **106** (3): 598–620.

Blum, A., R. Lalli and J. Renn. 2016. The renaissance of General Relativity: How and why it happened. *Annals of Physics* (Berlin) **528** (5): 344–349.

Bohr, N. 1986. *Collected Works*. Volume 9. *Nuclear Physics (1929-1952)*, edited by R. Peierls (Elsevier).

Bohr, N. and J.A. Wheeler. 1939. *Physical Review* **56** (5): 426–450.

Bonolis, L. 2014. From cosmic ray physics to cosmic ray astronomy: Bruno Rossi and the opening of new windows on the universe. *Astroparticle Physics* **53**: 67–85.

Bothe, W. and W. Kolhörster. 1929. Das Wesen der Höhenstrahlung. *Zeitschrift für Physik* **56** (11-12): 751–777.

Bronstein, M. 1929. Zum Strahlungsgleichgewichtsproblem von Milne. *Zeitschrift für Physik* **58** (9): 696–699.

Bronstein, M. 1930. Note on the temperature distribution in the deep layers of stellar atmospheres. *Monthly Notices of the Royal Astronomical Society* **91**: 133–138.

Bronstein, M. 1931. Sovremennoe sostoyanie relyativistskoy kosmologii [The Modern State of Relativistic Cosmology]. *Uspekhi Fizicheskikh Nauk* **11**: 124–184.

Burbidge, G. 1959. Estimates of the Total Energy in Particles and Magnetic Field in the Non-Thermal Radio Sources. *Astrophysical Journal* **129**: 849–852.

Burbidge, G. 1962. Nuclear Astrophysics. *Annual Review of Nuclear and Particle Science* **12**: 507–576.

Burbidge, G., M. Burbidge, W. Fowler and F. Hoyle. 1957. Synthesis of the Elements in Stars and cosmological theories. *Reviews of Modern Physics* **29** (4): 547–650.

Cameron, A.G.W. 1958. Nuclear Astrophysics. *Annual Review of Nuclear and Particle Science* **8**: 299–326.

Cameron, A.G.W. 1959. Neutron Star Models. *Astrophysical Journal* **130**: 884–894.

Cenadelli, D. 2010. Solving the Giant Stars Problem: Theories of Stellar Evolution from The 1930s to The 1950s. *Archive for History of Exact Sciences* **64** (2): 203–267.

Cernuschi, F. 1939a. Super-Novae and the Neutron-Core Stars. *Physical Review* **56** (1): 120–120.

Cernuschi, F. 1939b. A Tentative Theory of the Origin of Cosmic Rays. *Physical Review* **56** (1): 120–121.

Cernuschi, F. 1939c. On the Behavior of Matter at extremely high temperatures and pressures. *Physical Review* **56** (5): 450–455.

Chadwick, J. 1932. Possible Existence of a Neutron. *Nature* **129** (3252): 312–312.

Chandrasekhar, S. 1929. The Compton Scattering and the New Statistics. *Proceedings of the Royal Society* A **125** (796): 231–237.

Chandrasekhar, S. 1931a. The Maximum Mass of Ideal White Dwarfs. *Astrophysical Journal* **74**: 81–82.

Chandrasekhar, S. 1931b. The density of white dwarf stars. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* **11** (70): 592–596, Special Issue.

Chandrasekhar, S. 1931c. The highly collapsed configurations of a stellar mass. *Monthly Notices of the Royal Astronomical Society* **91**: 456–466.

Chandrasekhar, S. 1932. Some Remarks on the State of Matter in the Interior of Stars. *Zeitschrift für Astrophysik* **5**: 321–327.

Chandrasekhar, S. 1934a. The physical state of matter in the interior of stars. *The Observatory* **57**: 93–99.

Chandrasekhar, S. 1934b. Stellar configurations with degenerate cores. *The Observatory* **57**: 373–377.

Chandrasekhar, S. 1935a. The highly collapsed configurations of a stellar mass (Second paper). *Monthly Notices of the Royal Astronomical Society* **95**: 207–225.

Chandrasekhar, S. 1935b. Stellar configurations with degenerate cores. *Monthly Notices of the Royal Astronomical Society* **95**: 226–260.

Chandrasekhar, S. 1935c. Stellar configurations with degenerate cores (Second paper). *Monthly Notices of the Royal Astronomical Society* **95**: 676–693.

Chandrasekhar, S. 1939. *An introduction to the study of stellar structure*. University of Chicago Press.

Chandrasekhar, S. 1941. The white dwarfs and their importance for the theories of stellar evolution in Lundmark et al. (eds.) *Colloque international d'Astrophysique. Tenu au Collège de France, Paris. 17-23 juillet, 1939 sous la Presidence du Professeur Henry Norris Russell. Les novae et les naines blanches. 1. Observations des novae*. Hermann, Paris, pp. 41–48.

Chandrasekhar, S. 1964. The Dynamical Instability of Gaseous Masses Approaching the Schwarzschild Limit in General Relativity. *Astrophysical Journal* **140**: 417–433.

Chandrasekhar, S. 1972. A limiting case of relativistic equilibrium, in *General Relativity* edited by L. O'Raifertaigh, in honor of J.L. Synge. Clarendon Press, Oxford, pp. 185–199.

Chandrasekhar, S. 1977. Interview by Spencer Weart on 1977 May 17, Niels Bohr Library & Archives, American Institute of Physics, College Park, MD USA.

Chandrasekhar, S. and C. Møller. 1935. Relativistic degeneracy. *Monthly Notices of the Royal Astronomical Society* **95**: 673–676.

Chandrasekhar, S. and L. Rosenfeld. 1935. Production of Electron Pairs and the Theory of Stellar Structure. *Nature* **135** (3424): 999–999.

Chandrasekhar, S., G. Gamow and M. Tuve. 1938. The Problem of Stellar Energy. *Nature* **141** (3578): 982–982.

Chandrasekhar, S. and M. Schönberg. 1942. On the Evolution of the Main-Sequence Stars. *Astrophysical Journal* **96**: 161–172.

Chiu, H.-Y. 1964. Gravitational collapse. *Physics Today* **17** (5): 21–34.

Colgate, S.A. and R.H. White. 1966. The Hydrodynamic Behavior of Supernovae Explosions. *Astrophysical Journal* **143**: 626–681.

Cowling, T.G. 1966. The Development of the Theory of Stellar Structure. *Quarterly Journal of the Royal Astronomical Society* **7**: 121–137.

Critchfield, C. and E. Teller. 1938. On the Saturation of Nuclear Forces. *Physical Review* **53** (10): 812–818.

Datt, B. 1938. Über eine Klasse von Lösungen der Gravitationsgleichungen der Relativität. *Zeitschrift für Physik* **108** (5-6): 314–321.

Department of Physics, Notre Dame University. 1938. The Notre Dame Symposium on the Physics of the Universe and the Nature of Primordial Particles. *Science* **87** (2265): 487–490.

Deprit, A. 1984. Monsignor Georges Lemaître. In André Berger (ed.). *The Big Bang and Georges Lemaître. Proceedings of a Symposium in Honour of G. Lemaître Fifty Years after his Initiation of Big-Bang Cosmology, Louvain-la-Neuve, Belgium,10-13 October 1983*. D. Reidel, Dordrecht, pp. 357–392.

Dirac, P.A.M. 1926. On the Theory of Quantum Mechanics. *Proceedings of the Royal Society* A **112** (762): 661–677.

Eddington, A.S. 1922. A Century of Astronomy. *Nature* **109** (2747): 815–817.

Eddington, A.S. 1924a. On the relation between the masses and luminosities of the stars. *Monthly Notices of the Royal Astronomical Society* **84**: 308–332.

Eddington, A.S. 1924b. *Nature*. The Relation between the Masses and Luminosities of the Stars. *Nature* **113** (2848): 786–788.

Eddington, A.S. 1926. *The internal constitution of stars*. Cambridge University Press.

Eddington, A.S. 1927. *Stars and Atoms*. Clarendon Press, Oxford.

Eddington, A.S. 1931. The End of the World: from the Standpoint of Mathematical Physics. *Nature* **127** (3203): 447–453.

Eddington, A.S. 1935a. Relativistic degeneracy. *The Observatory* **58** (729): 37–39.

Eddington, A.S. 1935b. On 'relativistic degeneracy'. *Monthly Notices of the Royal Astronomical Society* **95**: 194–206.

Eddington, A.S. 1935c. Note on 'relativistic degeneracy'. *Monthly Notices of the Royal Astronomical Society* **96**: 20–21.

Eddington, A.S. 1935. The Pressure of a Degenerate Electron Gas and Related Problems. *Proceedings of the Royal Society* **152** (876): 253–272.

Ehlers, J. 2009. Editorial note to (Zwicky 1933b). F. Zwicky The redshift of extragalactic nebulae. *General Relativity and Gravitation* **41** (1): 203–206.

Einstein, A. 1939. On a stationary system with spherical symmetry consisting of many gravitating masses. *Annals of Mathematics* **40** (4): 922–936.

Eisenstaedt, J. 1986. La relativité générale à l'étiage: 1925–1955. *Archive for the History of Exact Sciences* **35** (2): 115–185.

Eisenstaedt, J. 1987a. Trajectoires et impasses de la solution de Schwarzschild. *Archive for History of Exact Sciences* **37** (4): 275–357.

Eisenstaedt, J. 1987b. The low water mark of general relativity, 1925–1955. In Howard D. and Stachel J. (eds.). *Einstein and the History of General Relativity*. Birkhäuser, Boston, pp. 277–292.

Eisenstaedt, J. 1993. Lemaître and the Schwarzschild Solution. In J. Earman, M. Janssen, and J.D. Norton (eds.). *The Attraction of Gravitation: New Studies in the History of General Relativity. Proceedings of the Third International Conference on the History and Philosophy of General Relativity*. Einstein Studies, Vol. 5. Birkhäuser, Boston, pp. 353–389.

Eisenstaedt, J. 2006. *The Curious History of Relativity: How Einstein's Theory of Gravity was Lost and Found Again*. Princeton University Press.

Eremeeva, A.L. 1995. Political Repression and Personality: The History of Political Repression Against Soviet Astronomers. *Journal for the History of Astronomy* **26**: 297–324.

Fermi, E. 1926a. Sulla quantizzazione del gas perfetto monoatomico. *Rendiconti Lincei* **3**: 145–149. [in Italian] (Presented at the Accademia dei Lincei meeting of February 7, 1926); English translation by Alberto Zannoni. 1999. E. Fermi. On the quantization of the monoatomic ideal gas arXiv:cond-mat/9912229

Fermi, E. 1926b. Zur Quantelung des Idealen Einatomigen Gases. *Zeitschrift für Physik* **36** (11): 902–912.

Fermi, E. 1949. On the Origin of the Cosmic Radiation. *Physical Review* **75** (8): 1169–1174.

Flügge, S. 1933. Der Einfluss der Neutronen auf den inneren Aufbau der Sterne. (Veröffentlichungen der Universitäts–Sternwarte zu Göttingen 31). Dissertation. *Zeitschrift für Astrophysik* **6**: 272–292.

Fowler, R.H. 1926a. On Dense Matter. *Monthly Notices of the Royal Astronomical Society* **87** (2): 114–122.

Fowler, R.H. 1926b. The statistical mechanics of assemblies of ionized atoms and electrons. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* **1** (4) Special Issue: 845–875.

Fowler, R.H. 1926c. General Forms of Statistical Mechanics with Special Reference to the Requirements of the New Quantum Mechanics. *Proceedings of the Royal Society* A **113** (764): 432–449.

Fowler, R.H. 1929. *Statistical Mechanics. The theory of the properties of matter in equilibrium*. Cambridge University Press New York, The Macmillan Company.

Fowler, R.H. and E.A. Milne. 1923. The intensities of absorption lines in stellar spectra, and the temperatures and pressures in the reversing layers of stars. *Monthly Notices of the Royal Astronomical Society* **83**: 403–424.

Fowler, R.H. and E.A. Guggenheim. 1925a. Applications of statistical mechanics to determine the properties of matter in stellar interiors. Part I. The mean molecular weight. *Monthly Notices of the Royal Astronomical Society* **85**: 939–960.

Fowler, R.H. and E.A. Guggenheim. 1925b. Applications of statistical mechanics to determine the properties of matter in stellar interiors. Part II. The Adiabatics. *Monthly Notices of the Royal Astronomical Society* **85**: 961–970.

Fowler, W.A. 1983a. Autobiographical note, *Nobel Lectures, Physics 1981-1990*. Editor-in-Charge Tore Frängsmyr, Editor Gösta . World Scientific Publishing Co., Singapore, http://www.nobelprize.org/nobel_prizes/physics/laureates/1983/fowler-bio.html

Fowler, W.A. 1983-1986. Interviews by John Greenberg and Carol Bugé. Pasadena, California, 1983 May 3, 1984 May 31, 1986 October 3. Oral History Project, California Institute of Technology Archives.

Hoyle, F. and W.A. Fowler. 1963a. On the Nature of Strong Radio Sources. *Monthly Notices of the Royal Astronomical Society* **125** (2): 169–176.

Hoyle, F. and W.A. Fowler. 1963b. Nature of strong radio sources. *Nature* **197** (4867): 533–535.

Frenkel, J. 1928. Nouveaux développements de la théorie Électronique des métaux. *Atti del Congresso internazionale dei Fisici (11-20 September 1927, Como-Pavia-Roma)* edited by the Committee for the celebration of the centennial of the death of Alessandro Volta, N. Zanichelli, Bologna, pp. 65–103.

Frenkel, J. 1928a. Zur wellenmechanischen Theorie der metallischen Leitfähigkeit. *Zeitschrift für Physik* **47** (11): 819–834.

Frenkel, J. 1928b. Anwendung der Pauli-Fermischen Elektronengastheorie auf das Problem der Kohäsionskräfte. *Zeitschrift für Physik* **50** (3-4): 234–248.

Frenkel, J. 1928c. Elementare Theorie magnetischer und elektrischer Eigenschaften der Metalle beim absoluten Nullpunkt der Temperatur. *Zeitschrift für Physik* **49** (1-2): 31–45.

Frenkel, V.Y. 1966. *Yakow Ilich Frenkel. His work, life and letters*. Springer, Basel.

Friedmann, A. 1922. Über die Krümmung des Raumes. *Zeitschrift für Physik* **10** (1): 377–386.

Friedman, J.L. 1996. Stability Theory of Relativistic Stars. *Journal of Astrophysics and Astronomy* **17** (3-4): 199–211.

Gamow, G. et al. 1929. Discussion on the Structure of Atomic Nuclei. *Proceedings of the Royal Society* A **123** (792): 373–390.

Gamow, G. 1930. Mass Defect Curve and Nuclear Constitution. *Proceedings of the Royal Society* A **126** (803): 632–644.

Gamow, G. 1931. *Constitution of atomic nuclei and radioactivity*. Clarendon Press, Oxford.

Gamow, G. and L.D. Landau. 1933. Internal Temperature of Stars. *Nature* **132** (3336): 567–567.

Gamow, G. 1934. Modern Ideas on Nuclear Constitution. *Nature* **133** (3368): 744–747.

Gamow, G. 1935. Nuclear transformation and the origin of chemical elements. *Ohio Journal of Science* **35** (5): 406–413.

Gamow, G. 1937. *Atomic nuclei and nuclear transformations*. Clarendon Press.

Gamow, G. 1938a. Tracks of Stellar Evolution *Physical Review* **53** (11): 907–908.

Gamow, G. 1938b. Tentative Theory of Novae. *Physical Review* **54** (6): 480–480.

Gamow, G. 1938c. Zusammenfassender Bericht. Kernumwandlungen als Energiequelle der Sterne. *Zeitschrift für Astrophysik* **16**: 113–160.

Gamow, G. 1939a. Physical Possibilities of Stellar Evolution. *Physical Review* **55** (8): 718–725.

Gamow, G. 1939b. Nuclear Reactions in Stellar Evolution. *Nature* **144** (3648): 575–577.

Gamow, G. 1949. On Relativistic Cosmology. *Reviews of Modern Physics* **21** (3): 367–373.

Gamow, G. 1968. Interview by Charles Weiner on 1968 April 25, Niels Bohr Library & Archives, American Institute of Physics, College Park, MD USA.

Gamow, G. 1970. *My World Line. An informal autobiography.* The Viking Press, New York.

Gamow, G. and M. Schönberg. 1940. The Possible Role of Neutrinos in Stellar Evolution. *Physical Review* **58** (12): 1117–1117.

Gamow, G. and M. Schönberg. 1941. Neutrino Theory of Stellar Collapse. *Physical Review* **59** (7): 539–547.

Gamow, G. and E. Teller. 1938. On the neutron core of stars. Proceedings of the APS, Minutes of the Washington DC Meeting, April 28-30, June 1 1938, *Physical Review* **53** (11): 929–930.

Gamow, G. and E. Teller. 1939. On the Origin of Great Nebulae. *Physical Review* **55** (7): 654–657.

Gamow, G. and J.A. Fleming. 1942. The Eighth Annual Washington Conference of Theoretical Physics. *Science* **95** (2475): 579–581.

Gorelik, G.E. 2005. Matvei Bronstein and quantum gravity: 70th anniversary of the unsolved problem. *Physics-Uspekhi* **48** (10): 1039–1053.

Gorelik, G. and V.Y. Frenkel. 1994. *Matvei Petrovich Bronstein and Soviet theoretical physics in the thirties*. Birkäuser, Basel.

Grant, K. 1926. The constitution of stars. *Nature* **118** (2967): 373–374.

Greenberg, J.L. and J.R. Goodstein. The origins of nuclear astrophysics at Caltech. Humanities Working Paper 97. California Institute of Technology, Pasadena, CA.

Greenstein, J.L., and T.A. Matthews. 1963. Red-shift of the unusual radio source: 3C48. *Nature* **197** (4872): 1041–1042.

Greenstein, J.L., J.B. Oke and H.L. Shipman. 1971. Effective Temperature, Radius, and Gravitational Redshift of Sirius B. *Astrophysical Journal* **169**: 563–566.

Hagihara, Y. 1931. Theory of the Relativistic Trajectories in a Gravitational Field of Schwarzschild. *Japanese Journal of Astronomy and Geophysics* **8**: 67–176.

Hamada, T. and E.E. Salpeter. 1961. Models for Zero-Temperature Stars. *Astrophysical Journal* **134**: 683–698.

Harrison, B.K., M. Wakano and J.A. Wheeler. 1958. Matter-energy at high density; end point of thermonuclear evolution, in R. Stoops (ed.). *Proceedings, 11ème Conseil de Physique de l'Institut International de Physique Solvay: La structure et l'évolution de l'univers: rapports et discussions*.

Harrison, B.K., K.S. Thorne, M. Wakano and J.A. Wheeler. 1965. *Gravitation theory and gravitational collapse*. University of Chicago Press.

Hazard, C., M.B. Mackey and A.J. Shimmins. 1963. Investigation of the Radio Source 3C 273 by the method of lunar occultations. *Nature* **197** (4872): 1037–1039.

Hoddeson, L., G. Baym and M. Eckert. 1987. The development of the quantum-mechanical electron theory of metals: 1928–1933. *Reviews of Modern Physics* **59** (1): 287–327.

Hoddeson, L. et al. 1992. *Out of the Crystal Maze: Chapters from The History of Solid State Physics*. Oxford University Press.

Holberg, J.B. 2009. The Discovery of the Existence of White Dwarf Stars: 1862 to 1930. *Journal for the History of Astronomy* **40**: 137–154.

Holberg, J.B. 2010. Sirius B and the Measurement of the Gravitational Redshift. *Journal for the History of Astronomy* **41** (1): 41–64.

Hubble, E.P. 1929. A Relation between Distance and Radial Velocity among Extragalactic Nebulae. *Proceedings of the National Academy of Sciences* **15** (3): 168–173.

Hubble, E.P. and R.C. Tolman, *Astrophysical Journal* **82** (2): 302–337.

Hufbauer, K. 2006. J. Robert Oppenheimer's path to black holes, C. Carson and D.A. Hollinger (eds.). *Reappraising Oppenheimer*. University of California Press, Berkeley, pp. 36–48.

Hufbauer, K. 2007. Landau's youthful sallies into stellar theory: Their origins, claims, and receptions. *Historical Studies in the Physical and Biological Sciences* **37** (2): 337–354.

Hufbauer, K. 2009. George Gamow. 1904–1968. *Biographical Memoirs of the National Academy of Sciences*.

Hund, F. 1936. Materie unter sehr hohen Drucken und Temperaturen. *Ergebnisse der Exakten Naturwissenschaften* **15**: 189–228. English translation in H. Riffert et al. (eds.). 1996. Matter at high densities in *Astrophysics. Compact Stars and the Equation of State*. Springer, pp. 217–257.

Israel, W. 1987. Dark stars: the evolution of an idea, in S.W. Hawking and W. Israel (eds.). *Three Hundred Years of Gravitation*. Cambridge University Press, pp. 199–276.

Jeans, J.H. 1927. On liquid stars and the liberation of stellar energy. *Monthly Notices of the Royal Astronomical Society* **87**: 400–414.

Kapitza, P.L. and E.M. Lifshitz. 1969. Lev Davydovitch Landau. 1908–1968. *Biographical Memoirs of Fellows of the Royal Society* **15**: 140–158.

Kothari, D.S. 1937. Neutrons, degeneracy and white dwarfs. *Proceedings of the Royal Society* A **162** (911): 521–528.

Kragh, H. 1987. The Beginning of the World: Georges Lemaître and the Expanding Universe. *Centaurus* **30** (2): 114–139.

Kragh, H. 1996. *Cosmology and Controversy*. Princeton University Press.

Kragh, H. 2005. George Gamow and the 'Factual Approach' to Relativistic Cosmology, in A.J. Kox and J. Eisenstaedt (eds.). *The Universe of General Relativity*, Einstein Studies vol. 11. Birkhäuser, Basel, pp. 175–188.

Kragh, H. 2013. 'The Wildest Speculation of All': Lemaître and the Primeval-Atom Universe. in R.D. Holder and S. Mitton (eds.). *Georges Lemaître: Life, Science and Legacy*. Springer-Verlag, Berlin, Heidelberg, pp. 23–38.

Krasinski, A. 1999. Editor's Note: On a Class of Solutions of the Gravitation Equations of Relativity (Datt 1938). *General Relativity and Gravitation* **31** (10): 1615–1618.

Landau, L. 1930. Diamagnetismus der Metalle. *Zeitschrift für Physik* **64** (9-10): 629–637.

Landau, L.D. 1932. On the theory of stars. *Physikalische Zeitschrift der Sowjetunion* **1** (2): 285–288.

Landau, L.D. 1938. Origin of stellar energy. *Nature* **141** (3564): 333–334. English version of the paper published in 1937 on *Doklady Akademii Nauk* SSSR **17**: 301–302.

Landau, L.D. and R. Peierls. 1931. Erweiterung des Unbestimmtheitsprinzips für die relativistische Quantentheorie. *Zeitschrift für Physik* **69** (1-2): 56–69.

Langer, R.M. and N. Rosen. 1931. The neutron. *Physical Review* **37** (12): 1579–1582.

Lee, T.D. 1950. Hydrogen content and energy-production of white dwarfs. *Astrophysical Journal* **111**: 625–640.

Lemaître, G. 1927. Un univers homogène de masse constante et de rayon variable rendant compte de la vitesse radiale des nébuleuses extra-galactiques. *Annales de la Societé Scientifique de Bruxelles* A **47**: 49–59.

Lemaître, G. 1931. Expansion of the universe. A homogeneous universe of constant mass and increasing radius accounting for the radial velocity of extra-galactic nebulae. *Monthly Notices of the Royal Astronomical Society* **91**: 483–490. English translation of Lemaître (1927). Un univers homogène de masse constante et de rayon croissant, rendant compte de la vitesse radiale des nébuleuses extragalactiques. *Annales de la Société Scientifique de Bruxelles* A **47**: 49–59.

Lemaître, G. 1931a. The beginning of the world from the point of view of quantum theory. *Nature* **127** (3210): 706–706.

Lemaître, G. 1931b. Contributions to a British Association Discussion on the Evolution of the Universe. Supplement to *Nature* **128** (3234): 704–706.

Lemaître, G. 1931c. L'Expansion de l'Espace. *Publications du Laboratoire d'Astronomie et de Geodesie de l'Universite de Louvain* **8**: 101–120.

Lemaître, G. 1932. l'Univers en Expansion. *Publications du Laboratoire d'Astronomie et de Geodesie de l'Universite de Louvain* **9**: 171–205, and in G. Lemaître. 1933. *Annales de la Société Scientifique de Bruxelles* **53**, Série A, Sciences mathématiques: 51–85.

Lemaître, G. and M.S. Vallarta. 1933. On Compton's latitude effect of cosmic radiation. *Physical Review* **43** (2): 87–91.

Luyten, W.J. 1960. White dwarfs and stellar evolution. *American Scientist* **48** (1): 30–39.

Marshak, R.E. 1940. The Internal Temperature of White Dwarf Stars. *Astrophysical Journal* **92**: 321–353.

Marshak, R.E. 1970. Interview by C. Weiner on 1970 June 15, Niels Bohr Library & Archives, American Institute of Physics, College Park, MD USA.
Menzel, D.H. 1926. The Planetary Nebulae. *Publications of the Astronomical Society of the Pacific* **38** (225): 295–312.
Mestel, L. 2004. Arthur Stanley Eddington: pioneer of stellar structure theory. *Journal of Astronomical History and Heritage* **7** (2): 65–73.
Migdal, A.B. 1977. Interview by Lillian Hoddeson and Gordon Baym on May 25, 1977, Niels Bohr Library & Archives, American Institute of Physics, College Park, MD USA.
Migdal, A.B. 1959. Superfluidity and the moments of inertia of nuclei. *Nuclear Physics* **13** (5): 655–674.
Miller, J.C. 1993. Relativistic Gravitational Collapse. G. Ellis et al. *The Renaissance of General Relativity and Cosmology. A Survey to Celebrate the 65th Birthday of Dennis Sciama*. Cambridge University Press, pp. 73–99.
Miller, A.I. 2005. *Empire of the Stars. Friendship, Obsession and Betrayal in the Quest for Black Holes*. Little Brown, London.
Milne, E.A. 1929. The masses, luminosities, and effective temperatures of the stars. *Monthly Notices of the Royal Astronomical Society* **90**: 17–54.
Milne, E.A. 1930a. The analysis of stellar structure. *Monthly Notices of the Royal Astronomical Society* **91**: 4–55.
Milne, E.A. 1930b. The analysis of stellar structure. *The Observatory* **53**: 305–308.
Milne, E.A. 1930c. Stellar Structure and the Origin of Stellar Energy. *Nature* **126** (3172): 238–238.
Milne, E.A. 1930d. The masses, luminosities, and effective temperatures of the stars (Second paper). *Monthly Notices of the Royal Astronomical Society* **90**: 678–687.
Milne, E.A. 1931a. Stellar Structure and the Origin of Stellar Energy. *Nature* **127** (3192): 16–18.
Milne, E.A. 1931b. Discussion on the Evolution of the Universe. *Nature* **128** (3234): 715–717.
Milne, E.A. 1945. Ralph Howard Fowler. 1889-1944. *Obituary Notices of Fellows of the Royal Society* **5** (14): 60–78.
Minkowski, R. 1939. The Spectra of the Supernovae in IC 4182 and in NGC 1003. *Astrophysical Journal* **89**: 156–216.
Minkowski, R. 1942. The Crab Nebula. *Astrophysical Journal* **96**: 199–213.
Nadyozhin, D.K. 1995. Gamow and the Physics and Evolution of Stars. *Space Science Reviews* **74** (3-4): 455-461
Nauenberg, M. 2008. Edmund C. Stoner and the discovery of the maximum mass of white dwarfs. *Journal for the History of Astronomy* **39**: 297–312.
von Neumann, J. 1961. *Collected Works*, edited by A.H. Taub, Vol. VI. Pergamon Press, London.
Oke, J.B. 1963. Absolute energy distribution in the optical spectrum of 3C 273. *Nature* **197** (4872): 1040–1041.
Öpik, E. 1938. Composite stellar models, *Publication of the Tartu University Observatory* **30** (4): 1–48.
Oppenheimer, J.R. 1926. On the Quantum Theory of the Problem of the Two Bodies. *Mathematical Proceedings of the Cambridge Philosophical Society* **23**: 422–431.
Oppenheimer, J.R. and R. Serber. On the Stability of Stellar Neutron Cores. *Physical Review* **54** (7): 540–540.
Oppenheimer, J.R. and G.M. Volkoff. 1939. On Massive Neutron Cores. *Physical Review* **55** (4): 374–381.
Oppenheimer, J.R. and H. Snyder. 1939. On Continued Gravitational Contraction. *Physical Review* **56** (5): 455–459.
Pais, A. 1993. *Niels Bohr's times: in Physics, Philosophy, and Polity*. Clarendon Press, Oxford.
Parker, E.N. 1997. Subrahmanyan Chandrasekhar. 1910-1995. *Biographical Memoirs of the National Academy of Sciences* **72**: 28–49.
Pauli, W. 1925. Über den Zusammenhang des Abschlusses der Elektronengruppen in Atom mit der Komplex Struktur der Spektren. *Zeitschrift für Physik* **31** (1): 765–783.

Pauli, W. 1927. Über Gasentartung und Paramagnetismus. *Zeitschrift für Physik* **41** (2): 81–102.

Peebles, P.J.E. 2014. Discovery of the hot Big Bang: What happened in 1948. *The European Physical Journal* H **39** (2): 205–223.

Peierls, R. 1936. Note on the derivation of the equation of state for a degenerate relativistic gas. *Monthly Notices of the Royal Astronomical Society* **96**: 780–784.

Peierls, R. 1977. Interview by Lillian Hoddeson and Gordon Baym on 1977 May 20, Niels Bohr Library & Archives, American Institute of Physics, College Park, MD USA.

Peierls, R. 1997. *Atomic Histories.* American Institute of Physics Press.

Penrose, R. 1996. Chandrasekhar, Black Holes, and Singularities. *Journal of Astrophysics and Astronomy* **17** (3-4): 213–232.

Pokrowski, G.I. 1928. Zur Frage nach der oberen Grenze für die Masse eines Sterns. *Zeitschrift für Physik* **49** (7-8): 587–589.

Regener, V. 1931. Über die Herkunft der Ultrastrahlung (Hesschen Strahlung). *Die Naturwissenschaften* **19** (22): 460–461.

Reynolds, J.H. 1923 Gaseous Nebulae. *Nature* **112** (2810): 375–376.

Robinson, I., A. Schild and E.L. Schucking (eds.). 1965. *Quasi-stellar sources and gravitational collapse. Including the proceedings of the First Texas Symposium on Relativistic Astrophysics.* University of Chicago Press, Chicago.

Rosenfeld, L. 1963. Interview by Thomas S. Kuhn and John L. Heilbron on 1963 July 19, Niels Bohr Library & Archives, American Institute of Physics, College Park, MD USA.

Russell, H.N. 1925a. Relativity Displacement of Spectral Lines and Stellar Constitution. *Nature* **116** (2912): 285–285.

Russell, H.N. 1925b. Remarkable New Tests Favor the Einstein Theory. *Scientific American* **133** (2): 88–88.

Rutherford, E. 1920. Bakerian Lecture: Nuclear Constitution of Atoms. *Proceedings of the Royal Society* A **97** (686): 347–400.

Salpeter, E.E. 1960. Matter at high densities. *Annals of Physics* **11** (4): 393–413.

Salpeter, E.E. 1961. Energy and Pressure of a Zero-Temperature Plasma. *Astrophysical Journal* **134** (3): 669–682.

Sandage, A.R. and M. Schwarzschild. 1952. Inhomogeneous Stellar Models. II. Models with Exhausted Cores in Gravitational Contraction. *Astrophysical Journal* **116**: 463–476.

Sardanashvily, G. 2014. Dmitri Ivanenko (in honor of the 110th year anniversary). *Science Newsletter* **1**: 16–17.

Schmidt, M. 1963. 3C 273: A star-like object with large red-shift. *Nature* **197** (4872): 1040–1040.

Schucking, E. 1989. The first Texas Symposium of Relativistic Astrophysics. *Physics Today* **42** (8): 46–52.

Schweber, S. 2008. *Einstein and Oppenheimer: The Meaning of Genius.* Harvard University Press.

Serber, R. 1992. *The Los Alamos Primer.* University of California Press, Berkeley CA.

Shaviv, G. 2009. *The Life of Stars. The Controversial Inception and Emergence of the Theory of Stellar Structure.* The Hebrew University Magnes Press and Springer-Verlag.

Smith, H.J. and D. Hoffleit. 1963. Light Variations in the Superluminous Radio Galaxy 3C273. *Nature* **198** (4881): 650–651.

Sommerfeld, A. 1928a. Zur Elektronentheorie der Metalle und des Volta-Effektes nach der Fermi'schen Statistik. *Atti del Congresso internazionale dei Fisici (11-20 September 1927, Como-Pavia-Roma)* edited by the Committee for the celebration of the centennial of the death of Alessandro Volta. N. Zanichelli, Bologna, pp. 449–473.

Sommerfeld, A. 1928b. Zur Elektronentheorie der Metalle auf Grund der Fermischen Statistik. *Zeitschrift für Physik* **47** (1-2): 1–32.

Sommerfeld, A. and H. Bethe. 1933. Elektronentheorie der Metalle, in *Aufbau der zusamennhängenden Materie. Handbuch der Physik* **24/2**: 333–622.

Sterne, T.E. 1932. Statistical Mechanics with Particular Reference to the Vapor Pressures and Entropies of Crystals. *Reviews of Modern Physics* **4** (4): 635–722.

Sterne, T.E. 1933a. The equilibrium theory of the abundance of the elements. *Physical Review* **43** (7): 585–586.

Sterne, T.E. 1933b. The equilibrium theory of the abundance of the elements: a statistical investigation of assemblies in equilibrium in which transmutations occur. *Monthly Notices of the Royal Astronomical Society* **93**: 736–766.

Sterne, T.E. 1933c. A note on the liberation of energy by transmutations of nuclei in the stars. *Monthly Notices of the Royal Astronomical Society* **93**: 767–769.

Sterne, T.E. 1933d. The equilibrium of transmutations in stars in which transmutations are an important source of energy. *Monthly Notices of the Royal Astronomical Society* **93**: 770–776.

Stoner, E.C. 1924. The Distribution of Electrons among Atomic Levels. The London, Edinburgh and Dublin *Philosophical Magazine and Journal of Science* **48** (286): 719–736.

Stoner, E.C. 1929. The limiting density in white dwarf stars. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* **7** (41): 63–70.

Stoner, E.C. 1930. The equilibrium of dense stars. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* **9** (60): 944–963.

Strömgren, B. 1931. The Point-Source Model with Coefficient of Opacity $k = k_1 \rho T^{-3.5}$. *Zeitschrift für Astrophysik* **2**: 345–369.

Strömgren, B. 1937. Die Theorie des Sterninnern und die Entwicklung der Sterne. *Ergebnisse der Exakten Naturwissenschaften* **16**: 465–534.

Strömgren, B. 1978. Interview by Karl Hufbauer on 1978 April 24, Niels Bohr Library & Archives, American Institute of Physics, College Park, MD USA.

Suzuki, S. 1931. Constitution of the white dwarf stars. *Nature* **128** (3237): 838–838.

Synge, J.L. 1950. The gravitational field of a particle. *Proceedings of the Royal Irish Academy* A **53**: 83–114.

Taub, A.H., O. Veblen and J. von Neumann. 1934. The Dirac equation in projective relativity. *Proceedings of the National Academy of Sciences of the United States of America* **20** (6): 383–388.

Tauber, G.E. and J.W. Weinberg. 1961. Internal State of a Gravitating Gas. *Physical Review* **122** (4): 1342–1346.

Thirring, H. 1926. Neuere experimentelle Ergebnisse zur Relativitätstheorie. *Die Naturwissenschaften* **14** (7): 111–116.

Thomas, E. 2011. On Stoner and white dwarf stars. *Philosophical Magazine* **91** (26): 3416–3422.

Thorne, K.S. 1989. Giant and supergiant stars with degenerate neutron cores. *Astrophysical Journal* **346**: 277–283.

Thorne, K.S. 1994. *Black Holes and Time Warps. Einstein's Outrageous Legacy.* W.W. Norton & Company, 1st ed.

Thorne, K.S. 2003. Warping spacetime, in G.W. Gibbons et al. (eds.) *The future of theoretical physics and cosmology: celebrating Stephen Hawking's 60th birthday.* Cambridge University Press, pp. 74–104.

Tolman, R.C. 1934. *Relativity, Thermodynamics, and Cosmology.* Clarendon Press, Oxford.

Tolman, R.C. 1934a. Effect of Inhomogeneity on Cosmological Models. *Proceedings of the National Academy of Sciences* **20** (3): 169–176.

Tolman, R.C. 1939a. Static Solutions of Einstein's Field Equations for Spheres of Fluid. *Physical Review* **55** (4): 364–373.

Tolman, R.C. 1939b. On the Stability of Spheres of Simple Mechanical Fluid Held Together by Newtonian Gravitation. *Astrophysical Journal* **90**: 541–567.

Tolman, R.C. 1939c. On the Stability of Stellar Models, with Remarks on the Origin of Novae. *Astrophysical Journal* **90**: 568–600.

Vardya, M.S. 1994. Astrophysics Contributions of Indian Scientists. *Defence Science Journal* **44** (3): 207–213.

Volkoff, G.M. 1939a. On the equilibrium of massive neutron cores. *Physical Review* **55** (4): 421–422.

Volkoff, G.M. 1939b. On the equilibrium of massive spheres. *Physical Review* **55** (4): 413–413.

Volkoff, M. 1990. Interview by Ann Carroll on 1990 May 23, University of British Columbia Archives Audio Recording Collection.

Wali, K.C. 1990. *Chandra. A biography of S. Chandrasekhar*. University of Chicago Press.

Walke, H.J. 1935. Nuclear synthesis and stellar radiation. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* **19** (126): 341–367.

von Weizsäcker, C.F. 1937a. *Die Atomkerne*. Akademische Verlagsgesellschaft, Leipzig.

von Weizsäcker, C.F. 1937b. Über Elementumwandlungen im Innern der Sterne. I (On transformations of elements in the interiors of stars. I). *Physikalische Zeitschrift* **38**: 176–191.

von Weizsäcker, C.F. 1938. Über Elementumwandlungen im Innern der Sterne. II (On transformations of elements in the interiors of stars. II). *Physikalische Zeitschrift* **39**: 633–646.

Wheeler, J.A. 1968. Our Universe: The Known and the Unknown, address before the American Association for the Advancement of Science, New York, 29 Dec. 1967. *American Scientist* **56** (1): 1–20.

Wheeler, J.A. 1998. *Geons, Black Holes, and Quantum Foam: A Life in Physics*. W.W. Norton & Co Inc.

Yakovlev, D.G. 1994. The article by Ya I Frenkel' on 'binding forces' and the theory of white dwarfs. *Physics Uspekhi* **37** (6): 609–612.

Yakovlev, D.G. et al. 2013. Lev Landau and the conception of neutron stars. *Physics Uspekhi* **56** (3): 289–295.

Zeldovich, Y.B. 1962a. The equation of state at ultrahigh densities and its relativistic limitations. *Soviet physics JETP* **14** (5): 1142–1147.

Zeldovich, Y.B. 1962b. The Collapse of a Small Mass in the General Theory of Relativity. *Soviet physics JETP* **15** (2): 446–447.

Zeldovich, Y.B. and I. Novikov. 1966. Relativistic Astrophysics. II. 1966. *Soviet Physics Uspekhi* **8** (4): 522–575.

Zwicky, F. 1929. On the Red Shift of Spectral Lines through Interstellar Space. *Proceedings of the National Academy of Sciences* **15** (10): 773–779.

Zwicky, F. 1933a. How far do cosmic rays travel? *Physical Review* **43** (2): 147–148.

Zwicky, F. 1933b. F. Zwicky, Die Rotverschiebung von extragalaktischen Nebeln. *Helvetica Physica Acta* **6**: 110–127. English translation: F. Zwicky. 2009. Republication of: The redshift of extragalactic nebulae. *General Relativity and Gravitation* **41** (1): 207–224.

Zwicky, F. 1938a. Some Results of the Search for Super-Novae. *Physical Review* **53** (12): 1019–1020.

Zwicky, F. 1938b. On neutron stars, Minutes of the San Diego Meeting, June 22-24, 1938. *Physical Review* **54** (3): 242–242.

Zwicky, F. 1938c. On Collapsed Neutron Stars. *Astrophysical Journal* **88**: 522–525.

Zwicky, F. 1939. On the Theory and Observation of Highly Collapsed Stars. *Physical Review* **55** (8): 726–743.

Zwicky, F. 1940. Types of Novae. *Reviews of Modern Physics* **12** (1): 66–85.

Zwicky, F. 2009. The redshift of extragalactic nebulae. *General Relativity and Gravitation* **41** (1): 203–206.