Mon. Not. R. astr. Soc. (1989) **238,** 503–521

A phase-space representation of Friedmann–Lemaître universes containing both dust and radiation and the inevitability of a big bang

J. Ehlers and W. Rindler ★ *Max-Planck-Institut für Physik und Astrophysik, Institut für Astrophysik, Karl-Schwarzchild-Strasse 1, D-8046 Garching, FRG*

Accepted 1988 November 24. Received 1988 November 7; in original form 1988 July 12

Summary. We classify all non-static Friedmann–Lemaître universes containing dust and radiation (or, as limiting cases, vacuum), and exhibit their histories as orbits in a three-dimensional (phase-) space $\mathscr{S}$, which faithfully reflects the topology of the set of states. As coordinates in $\mathscr{S}$ we take the dimensionless parameters Ω, ω and λ representing, respectively, the energy density of the dust, the radiation and the vacuum (corresponding to the cosmological constant). Three open subspaces of $\mathscr{S}$ contain, respectively, the orbits representing oscillating, inflectional and bouncing universes; their boundaries and edges contain the orbits of all the other universes. The classification proceeds from initial values (Ω_0, ω_0) in terms of (a, λ_0) diagrams $(a = R/R_0)$. An alternative set of 'invariant' parameters is also introduced, characterizing whole models rather than states of models. Finally we show that present-day redshift and density data exclude all models not having a big bang, without using assumptions about the origin of the microwave background radiation and without *a priori* assumptions about the value of the cosmological constant.

1 Introduction

In presentations of the homogeneous and isotropic general-relativistic cosmologies attention has mostly been restricted to models containing either dust or radiation only. Such models are kinematically specified by their Robertson–Walker (RW) metrics – and these, in turn, by a curvature index k $(= \pm 1$ or $0)$ and an expansion function ('radius of the universe') $R(t)$; additionally, one needs information about the model's contents: a mass constant ρR^3 in the case of dust, or an entropy constant uR^4 in the case of radiation (u = energy density). The metric and the model's contents are related by Einstein's field equations, here taken *with* a cosmological constant Λ. To relate a model to observations one must also know which is to be regarded as its present epoch t_0; this can be specified by the present value R_0 of the expansion function or the present value H_0 of the Hubble parameter.

★ Permanent address: University of Texas at Dallas, Richardson, TX 75083-0688, USA.

Though the traditional four parameters k, Λ, ρR^3 or uR^4, R_0 or H_0, uniquely determine such a model-plus-initial-data, the converse is not true: in the case $k=0$, R_0 and consequently ρR^3 or uR^4 can be rescaled. Therefore, these parameters do not properly coordinatize the three-dimensional space $\mathscr{S}$ of states, i.e. models-plus-initial-data, in neighbourhoods of the two-dimensional subspace corresponding to spatially flat ($k=0$) models: consequently they do not reflect the topology of that space. This drawback is particularly disturbing since all the realistic models are apparently contained in this very region, the curvature sign being not known from observations.

In this paper we consider the set of all homogeneous and isotropic relativistic world models whose substratum consists of an interaction-free mixture of cold, non-relativistic matter idealized as 'dust' and hot, relativistic matter idealized as 'radiation' (separately satisfying the respective conservation equations $\rho R^3 = \text{constant}$ and $uR^4 = \text{constant}$), and which expand (i.e. have positive Hubble parameter) at one instant – e.g. 'now'. It will turn out that models passing through both expanding and contracting states (the 'oscillating' and the 'bouncing' models) possess exactly one expanding and one contracting phase, and that these are mirror images of each other in cosmic time about the moment of arrest. No essential information is therefore lost by considering only the expanding phase of *every* model, which convention in fact we shall adopt in our discussion of $\mathscr{S}$, in order to avoid sign ambiguities. With each expanding state of a model we associate six numbers, Ω, ω, λ, H, κ and q, as follows (G being the Newtonian constant of gravitation):

$$\Omega := \frac{8\pi G}{3H^2}\rho \geq 0 \tag{1}$$

is the dimensionless dust mass density in units of the critical density;

$$\omega := \frac{8\pi G}{3H^2}u \geq 0 \tag{2}$$

is the corresponding quantity for radiation (we are using units such that $c=1$);

$$\lambda := \frac{\Lambda}{3H^2} = \frac{8\pi G}{3H^2}\cdot\frac{\Lambda}{8\pi G} \qquad (-\infty < \lambda < \infty) \tag{3}$$

is a dimensionless representation of the cosmological constant Λ – if $\Lambda > 0$, it can be regarded as the density parameter associated with the 'vacuum energy density' $\Lambda/8\pi G$;

$$H := \frac{dR/dt}{R} > 0 \tag{4}$$

is the Hubble parameter;

$$\kappa := \frac{k}{H^2R^2} \tag{5}$$

is the dimensionless curvature parameter; and

$$q := -\frac{R\,d^2R/dt^2}{(dR/dt)^2} = \frac{d}{dt}(H^{-1}) - 1 \tag{6}$$

is the dimensionless deceleration parameter. Only four of these parameters are independent since the field equations are – in terms of the preceding definitions – equivalent to

$$\kappa = \Omega + \omega + \lambda - 1 \tag{7}$$

and

$$q = \tfrac{1}{2}\,\Omega + \omega - \lambda. \tag{8}$$

[In fact, (8) is the difference of 'the' two field equations while (7) is the extension of the familiar Friedmann equation to include radiation; it is sometimes called the *Lemaître equation*. See, e.g. Rindler (1977, equations 9.72–9.76), or Misner, Thorne & Wheeler (1973, section 27.7). The coefficients $\frac{1}{2}$, 1, -1 in equation (8) reflect the fact that the relative acceleration between fundamental cosmic particles is caused, according to general relativity, by the source variable $\rho + 3p$. For dust, having $p = 0$, this equals ρ; for radiation, with $p = \frac{1}{3}\rho$, it equals 2ρ; and for the vacuum, with $p = -\rho$, it equals -2ρ. Equation (7), on the other hand, comes from an energy equation where only the ρs enter.]

It might seem preferable to use as basic parameters Ω, ω, q and H instead of Ω, ω, λ and H since the present values Ω_0, ω_0, q_0 and H_0 are, in principle, observable in contrast to λ_0. But since this complicates most formulae, we generally use the latter.

The parameters Ω, ω, λ and H play different roles. H fixes the single overall scale: If $\tilde{ds}$ is measured in some standard atomic unit, e.g. s, and ds is measured in units of the Hubble time H_0^{-1} at some instant t_0 of a particular RW model, one has

$$\tilde{ds}^2 = dt^2 - R^2(t)\, d\sigma_k^2 = H_0^{-2}\{dT^2 - a^2(T)\, d\sigma_{\kappa_0}^2\} = H_0^{-2}\, ds^2, \tag{9}$$

where T and $a(T)$ are defined by the first two of the following equations:

$$T := H_0 t,$$

$$a(T) := \frac{R(t)}{R_0} = \frac{1}{1+z},$$

$$\frac{da}{dT} = \frac{HR}{H_0 R_0}. \tag{10}$$

The dimensionless expansion function $a(T)$ is related to the redshift $z = \Delta\lambda/\lambda$, and its derivative to the Hubble parameter, as shown. In (9) $d\sigma_K^2$ denotes the metric of a three-space of constant Gaussian curvature $K (= k$ or $\kappa_0)$. The dimensionless metric ds^2 is determined by $(\Omega_0, \omega_0, \lambda_0)$ via equation (7) and equation (12) below. [Only the topology of a model's three-space is not determined by these parameters. The variety of possible space forms (Wolf 1972) depends on the sign of κ. However, in each case there is only *one* simply connected space form, this being $\mathbb{R}^3$ if $\kappa \leq 0$ and $\mathbb{S}^3$ if $\kappa > 0$.]

For many purposes it is useful to disregard the scale and to work with the dimensionless RW metrics ds^2. Then it must be kept in mind that a dimensionless metric of a model refers to a choice of 'present' time; metrics belonging to different 'present' times differ by constant conformal factors. We shall henceforth use the term (cosmological) *model* for 'dimensionless model' and accordingly shall mean from now on by a (cosmological) state that information contained in an initial data set which is dimensionless (scale-free). States in this sense are in one-to-one correspondence with number triplets $(\Omega, \omega, \lambda)$. The set $\mathcal{S}$ of states is thus a three-dimensional manifold-with-boundary-and-edge,

$$\mathcal{S} = \{(\Omega, \omega, \lambda) \mid \Omega \geq 0,\ \omega \geq 0\}.$$

The boundaries $\omega = 0$ and $\Omega = 0$ represent the dust states and the radiation states, respectively, and the edge $\omega = \Omega = 0$ represents the states of empty models.

A model consists of a one-parameter sequence of states. The expanding states of almost all models are mutually distinct. However, for exactly four models the state does not change in time: For the parabolic Einstein–de Sitter dust model (with $\Omega = 1$, $\omega = \lambda = 0$), the corresponding radiation model (with $\omega = 1$, $\Omega = \lambda = 0$), for Milne's empty model (with $\Omega = \omega = \lambda = 0$) and, of course, for the steady state (de Sitter) model (with $\Omega = \omega = 0$, $\lambda = 1$). (These models are characterized by the property that in each of them any two hypersurfaces of constant cosmic time can be mapped onto each other by a homothetic map of the underlying spacetime; therefore their states – in the new sense – are constant. They are also characterized as the only ones having constant deceleration parameters $\frac{1}{2}$, 1, 0, -1, respectively. The first three of these models are further characterized as the only ones whose $a(T)$ is given by a power law, or whose dimensionless ages $T_0 = H_0 t_0$ are constant during their evolution.)

The (expanding) states of a model form a curve in $\mathscr{S}$ which degenerates, for the four models just described, to a point. The space $\mathscr{M}$ of models thus arises from $\mathscr{S}$ as a two-dimensional quotient (or orbit) space. The corresponding spaces which include scale information are obtained simply as $\mathbb{R} \times \mathscr{S}$ and $\mathbb{R} \times \mathscr{M}$, respectively, an appropriate additional coordinate being H in the first case and $\lambda H^2 = \frac{1}{3}\Lambda$ in the second case.

The arrangement of the body of this paper is as follows: In Section 2 we classify models for fixed initial values (Ω_0, ω_0) in terms of (a, λ_0)-diagrams. Section 3 deals with the space of all cosmological states and its phase flow. Section 4 introduces an alternative set of 'invariant' model parameters. Finally in Section 5 (which is independent of Sections 3 and 4) we show that present-day redshift and density data exclude all but big-bang models, without using assumptions about the origin of the microwave background radiation and without *a priori* assumptions about the value of Λ.

2 Classification of models with given initial data $(\Omega_0, \omega_0, \lambda_0)$ using (a, λ_0) diagrams

Let us choose an initial state characterized by $(\Omega_0, \omega_0, \lambda_0)$, e.g. the present state. From the definitions (1)–(5) and the conservation laws $\rho R^3 = \text{constant}$, $uR^4 = \text{constant}$ it follows that

$$\Omega H^2 R^3 = \Omega_0 H_0^2 R_0^3,$$
$$\omega H^2 R^4 = \omega_0 H_0^2 R_0^4,$$
$$\lambda H^2 = \lambda_0 H_0^2,$$
$$\kappa H^2 R^2 = \kappa_0 H_0^2 R_0^2. \tag{11}$$

Using (11), we easily reobtain from (7) the original Lemaître differential equation in terms of the dimensionless expansion factor $a(T)$ of (10) (denoting differentiation with respect to T by a dot):

$$\dot{a}^2 = \omega_0(a^{-2} - 1) + \Omega_0(a^{-1} - 1) + \lambda_0(a^2 - 1) + 1. \tag{12}$$

At the initial instant T_0, $a_0 = \dot{a}_0 = 1$. In order to survey the set of all solutions of equation (12) which have arbitrary parameter values, we pick a pair (Ω_0, ω_0) with $\Omega_0 \geq 0$, $\omega_0 \geq 0$ and represent the range of each possible solution $a(T)$ corresponding to these data and an initial value λ_0, as a horizontal line segment in an (a, λ_0)-plane. For this purpose it is necessary to mark the boundary of the 'permitted region' $a > 0$, $\dot{a}^2 \geq 0$ in each diagram. The boundary corresponding to the first condition is of course the λ_0-axis $a = 0$, and we must stay to the right of that. The boundary corresponding to the second condition is found from (12) to have equation

$$\lambda_0 = \lambda_0(a) := \frac{(a-1)[a\Omega_0 + (a+1)\omega_0] - a^2}{a^2(a-1)(a+1)} = \frac{a^2(\Omega_0 + \omega_0 - 1) - a\Omega_0 - \omega_0}{a^2(a^2-1)}. \tag{13}$$

Fig. 1 shows qualitatively this boundary curve and the regions which it excludes. Each model corresponds to a maximally extended line segment $\lambda_0 =$ constant which includes the (present)

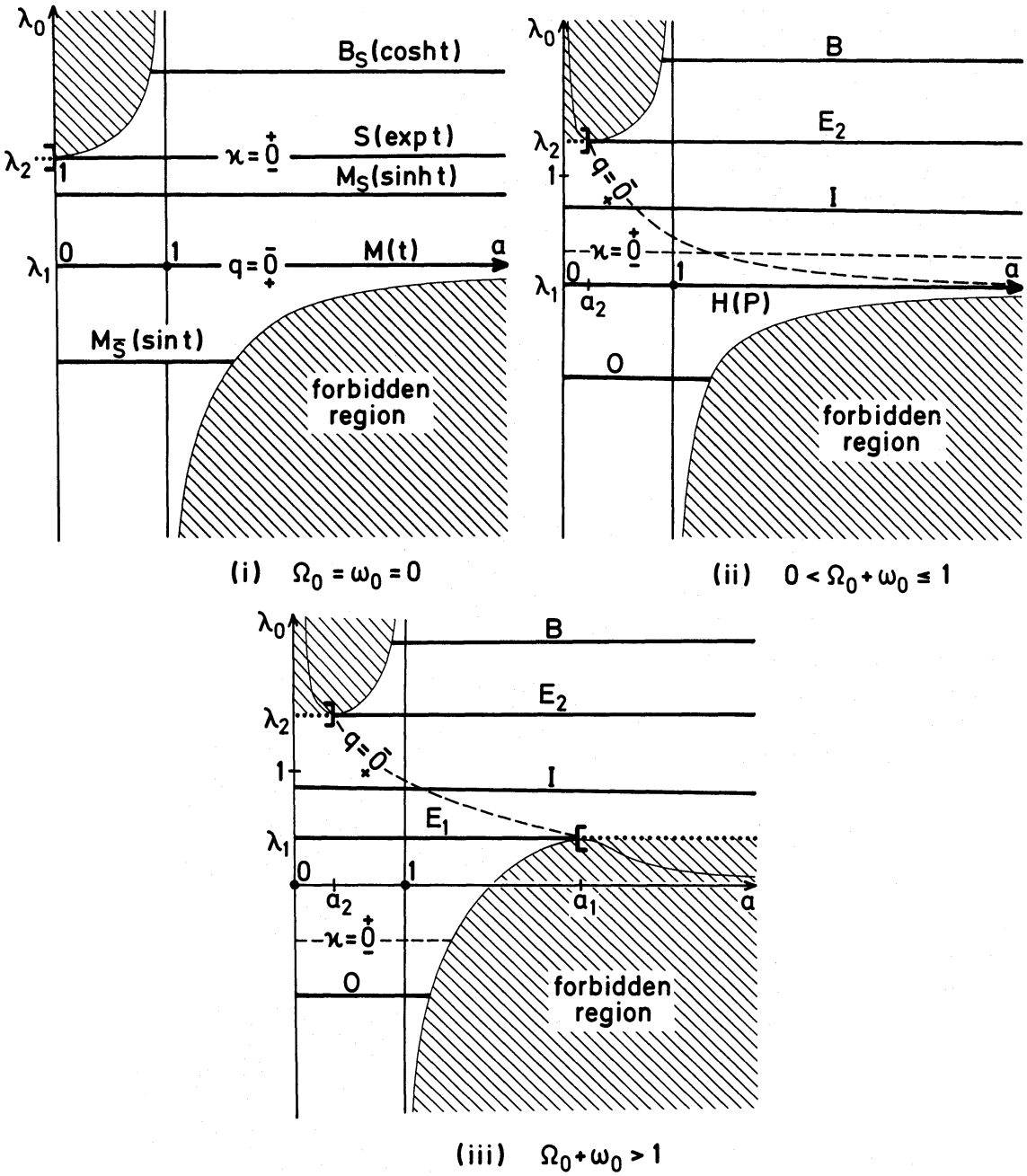

Figure 1. The (a, λ_0) diagrams for the cases (i) $\Omega_0 = \omega_0 = 0$, (ii) $0 < \Omega_0 + \omega_0 \leq 1$, (iii) $\Omega_0 + \omega_0 > 1$. Each horizontal line segment, maximally extended from $a = 1$, represents an FL model: B = bouncing model, I = inflectional model, O = oscillating model; $B_S = B$ in de Sitter space [a representative $R(t)$ is indicated in parenthesis in (i)], S = de Sitter model, M = Milne model, M_S = Milne-type model in de Sitter space, $M_{\bar{S}}$ = Milne-type model in anti-de Sitter space, E_1 and E_2 = models expanding to and from a stationary state, respectively, H = hyperbolic model (reducing to P = parabolic model when $\Omega_0 + \omega_0 = 1$). Note the (stippled) division lines $\kappa = 0$ separating positive- from negative-curvature states, and $q = 0$ separating accelerating ($q < 0$) from decelerating states; the $q = 0$ locus coincides with the λ_0 and a axes in (i).

value $a=1$. If that segment is bounded on the left by the locus $a=0$, the model has a big bang (except for the de Sitter model), and if it touches the locus $\lambda_0=\lambda_0(a)$ $(\dot{a}=0)$ on the right or on the left, the model has a turn-around from expansion to contraction or vice versa, respectively; if the segment approaches $\lambda_0=\lambda_0(a)$ at an extremal, say at $a=a_j$, the model approaches a static state, $a(T)\to a_j$, in the infinite future or in the infinite past.

This representation is somewhat different from the more usual one which relies on an (R, Λ)-diagram for fixed mass parameter ρR^3 (for dust models) – and which dates back at least to Robertson (1933, see also Stabell & Refsdal 1966, or Rindler 1977) – though in essence it is of course equivalent to it. (We note, in particular, that models with the same Λ (and Ω_0, ω_0) but different R-ranges, e.g. the two expanding models approaching the Einstein static universe in the past or in the future, now correspond to segments at different λ_0 levels, according to their different H_0's.) Two advantages of the present representation are that it is dimensionless and that the boundary of the (a, λ_0) domain occupied by solution-segments is now uniquely determined by (Ω_0, ω_0), whereas in the traditional scheme it depends also on the value of the curvature index k, another unknown.

The salient features of the boundary curve $\lambda_0(a)$ in the region of interest, $a>0$, are: a 'well' without zeros between $a=0$ and $a=1$ (whose left branch becomes vertical in the limit $\Omega_0=\omega_0=0$), and a 'cliff' rising from $\lambda_0=-\infty$ at $a=1$ to $\lambda_0=0$ at $a=\infty$ with or without a single prior hump according as

$$\Omega_0+\omega_0>1 \quad \text{or} \quad \Omega_0+\omega_0\le 1. \tag{14}$$

In the special case $\Omega_0=\omega_0=0$ the curve (13) is evidently as shown in Fig. 1(i). For the rest of the discussion we assume $\Omega_0+\omega_0>0$. From (13)(i), when $a\to 0$, $\lambda_0(a)\sim\omega_0/a^2$ if $\omega_0\neq 0$ and $\lambda_0(a)\sim\Omega_0/a$ otherwise; when $a\to 1$, $\lambda_0(a)\sim 1/2(1-a)$; and when $a\to\infty$, $\lambda_0(a)\to 0$. This establishes the asymptotic behaviour of $\lambda_0(a)$ near $a=0$, 1 and ∞. Consider next the zeros of $\lambda_0(a)$. If $0<a<1$ then $\lambda_0>0$, hence the well has no zeros. If $a>1$ and $\Omega_0+\omega_0\le 1$, $\lambda_0<0$ and the cliff has no zeros either; if $\Omega_0+\omega_0>1$, the numerator in (13)(ii) is -1 at $a=1$ and tends to infinity as $a\to\infty$, hence the 'cliff' has exactly one zero. To discuss the extremals of $\lambda_0(a)$ we first show that 'positive' horizontals $\lambda_0=\text{constant}>0$ intersect its graph twice (counting with proper multiplicity) or not at all. The intersection points are given by the roots a_i of the quartic (or cubic, after division by a, if $\omega_0=0$)

$$a^4\lambda_0-a^2(\lambda_0+\Omega_0+\omega_0-1)+a\Omega_0+\omega_0=0.$$

We read off that $\Sigma a_i=0$, so not all roots are positive, and $\Pi a_i=\omega_0/\lambda_0>0$ $(\Omega_0/\lambda_0<0$ if $\omega_0=0)$, so there are two positive roots or none, as asserted. This shows (i) that the minimum of the 'well' lies above the maximum of the 'cliff', and (ii) that neither of these branches can have more than one extremal for $\lambda_0>0$. That there are no extremals for $\lambda_0\le 0$ follows most easily from the fact that all extremals lie on the curve

$$\lambda_0=\frac{a\Omega_0+2\omega_0}{2a^4}, \tag{15}$$

which is positive for $a>0$. [To justify (15) as the equation for this locus, denote the rhs of equation (12) by $f(a, \lambda_0)$; then $f=0$ on $\lambda_0=\lambda_0(a)$ and $d\lambda_0/da=-(\partial f/\partial a)/(\partial f/\partial\lambda_0)$. Setting $\partial f/\partial a=0$ results in equation (15).] This completes our discussion of the general features of the 'permitted region' $a>0$, $\dot{a}^2\ge 0$ in the (a, λ_0) diagram.

The curve (15), shown as a stippled line in the (a, λ_0) diagrams, also plays another role: it corresponds to the locus $\ddot{a}=0$, since $\partial f/\partial a=\partial\text{lhs}(12)/\partial a=2\dot{a}\ddot{a}/\dot{a}=2\ddot{a}$. As such it divides decelerating epochs $(\ddot{a}<0)$ from accelerating epochs $(\ddot{a}>0)$, the former lying below it, the latter above. (In the case $\Omega_0=\omega_0=0$, $\partial f/\partial a=2a\lambda_0$, and so the locus of $\ddot{a}=0$ coincides with the

axes of a and λ_0.) Another useful division line in the (a, λ_0) diagram for a given choice (Ω_0, ω_0) is the horizontal

$$\lambda_0 = 1 - \Omega_0 - \omega_0, \tag{16}$$

corresponding to $\kappa_0 = 0$ [*cf.* equation (7)]; it separates models with positive curvature (lying above this line) from those with negative curvature (below it). This line lies below the 'well' and above the 'cliff' when $\Omega_0 + \omega_0 \leq 1$ [i.e. when the cliff has no hump – *cf.* (14)], and it intersects the cliff below $\lambda_0 = 0$ otherwise. For, any intersection points of the curves (13) and (16) must satisfy

$$a^4\lambda_0 = -a\Omega_0 - \omega_0,$$

i.e. $\lambda_0 < 0$ in the region of interest $a > 0$; but when the rhs of (16) is negative, the cliff has a hump [*cf.* (14)] and our assertion is established.

In order to allow a graphical determination of the expansion function $a(T)$ corresponding to a given set of parameters $(\Omega_0, \omega_0, \lambda_0)$, the (a, λ_0) diagram should be enriched by the contour lines $\dot{a}^2 = \text{constant}$ obtainable from equation (12), so that for each value of a along the horizontal segment $\lambda_0 = \text{constant}$ characterizing the model, $\dot{a}$ can be read off. But for a mere discussion of model-types that is not necessary. [Numerically computed graphs of $a(T)$ for various dust models are displayed in Felten & Isaacman 1986.] We have already indicated [after (13), before (16)] how $a(T)$ behaves when the segment approaches the boundaries $a = 0$ or $\lambda_0 = \lambda_0(a)$, and also when the segment crosses the locus $\ddot{a} = 0$. If, then, in an (a, λ_0) diagram,

$$a_1(\Omega_0, \omega_0), \lambda_1(\Omega_0, \omega_0) \quad \text{and} \quad a_2(\Omega_0, \omega_0), \lambda_2(\Omega_0, \omega_0) \tag{17}$$

denote the maximum of the 'cliff' (which we define to be zero when there is no hump), and the minimum of the 'well', respectively, we observe: segments having $\lambda_0 < \lambda_1$ represent 'oscillating' ('*O*') universes (decelerating all the way from a big bang to a big crunch); those with $\lambda_1 < \lambda_0 < \lambda_2$ represent 'inflectional' ('*I*') universes (decelerating after a big bang until some critical time, after which the expansion accelerates); segments with $\lambda_0 > \lambda_2$ represent 'bouncing' ('*B*') universes (contracting from infinite extension to a finite minimum and then expanding again to infinite extension); when the 'cliff' has a proper maximum $\lambda_1 > 0$, the segment $\lambda_0 = \lambda_1$, $0 < a < a_1$ represents a big-bang universe 'E_1' whose $a(T)$ monotonically increases to a_1 as $T \to \infty$; when $\lambda_1 = 0$ the segment $\lambda_0 = \lambda_1 = 0$, $0 < a < \infty$ represents a decelerating but nevertheless indefinitely expanding (hyperbolic, '*H*', or parabolic, '*P*') big-bang universe; and the segment $\lambda_0 = \lambda_2$, $a_2 < a < \infty$, represents a universe 'E_2' which expands indefinitely, approaching a static state $a = a_2$ as $T \to -\infty$.

We observe that *the classification of models* – i.e. the number and qualitative properties of the respective types and the regions they occupy in the (a, λ_0) diagram – *is the same for all pairs* $(\Omega_0, \omega_0) \neq (0, 0)$; in particular, it is the same for dust models, radiation models and 'mixed' models, with the following exception: the division line $\lambda_0 = \lambda_1$ itself represents E_1 models if $\Omega_0 + \omega_0 > 1$, 'parabolic' models if $\Omega_0 + \omega_0 = 1$, and 'hyperbolic' ones if $0 < \Omega_0 + \omega_0 < 1$ (see Figs 1 and 2). (For parabolic models $\dot{a} \to 0$ as $T \to \infty$, for hyperbolic models $\dot{a} \to \text{constant} > 0$.) We finally note that the types *B*, *I* and *O* are 'generic' in the sense that they fill two-dimensional regions of the (a, λ_0)-plane whereas the others are 'special', representing boundaries separating models of generic types.

3 The space $\mathscr{S}$ of cosmological states and its phase flow

In order to construct the curve of states representing – according to our convention – the expanding phase of the model having at the instant T_0 the state $(\Omega_0, \omega_0, \lambda_0)$, we begin by

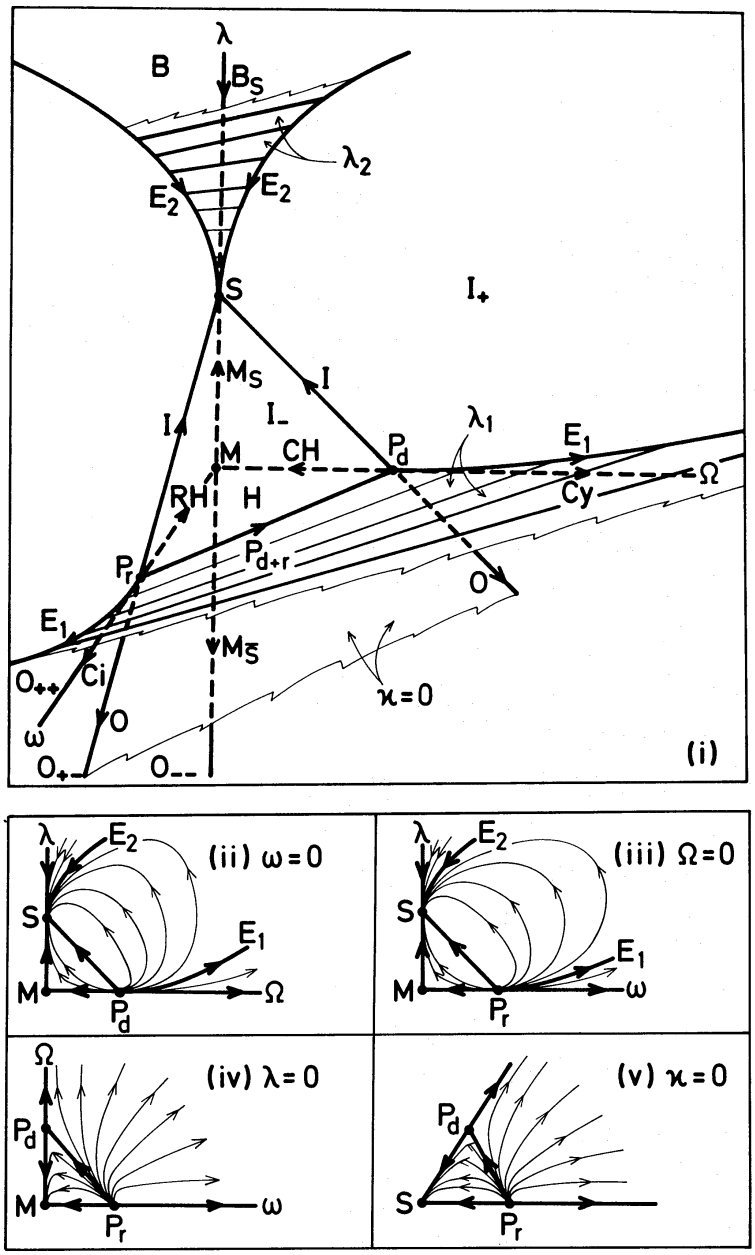

Figure 2. States of the non-static FL models in $(\Omega, \omega, \lambda)$ coordinates. Each model corresponds to a phase curve, arrows along phase curves indicate increasing $a(T)$. (i) The full three-dimensional phase space with its two main dividing (and also invariant) surfaces: $\lambda = \lambda_2$, separating B- from I-models, and $\lambda = \lambda_1$ (including the triangle $P_r P_d M$) separating I- from O-models. The other special invariant surfaces are $\Omega = 0$, $\omega = 0$, $\lambda = 0$ and $\kappa = 0$: models having one state in one of them have all their states in it. Intersections of invariant surfaces are invariant lines, *their* intersections, invariant points: both represent single models. (ii)–(v) Qualitative illustrations of flow line patterns in the four invariant planes. The subscripts on I indicate the sign of κ, those on O the signs of κ, λ respectively; the subscripts on P stand for dust (d), radiation (r) or both (d + r). For definitions of *Cy*, *Ci*, *CH* and *RH* see text, or Figs 3 and 4.

substituting (10)(iii) into the lhs of the Lemaître equation (12), thus obtaining an expression for H^2 at all a-values:

$$H^2 = \frac{H_0^2}{a^4}[a^4\lambda_0 - a^2(\Omega_0 + \omega_0 + \lambda_0 - 1) + a\Omega_0 + \omega_0]. \tag{18}$$

The relations (11)(i–iii) then yield the following parametric equation for the curve of states, with a as parameter:

$$(\Omega, \omega, \lambda) = \frac{(a\Omega_0, \omega_0, a^4\lambda_0)}{[a^4\lambda_0 - \alpha^2(\Omega_0 + \omega_0 + \lambda_0 - 1) + a\Omega_0 + \omega_0]}. \tag{19}$$

Any other point on this curve can theoretically serve as the 'present' point and be re-labelled (Ω_0, ω_0, λ_0). This leads to scale changes in a and T but it does not affect $R(t)$, t and k [*cf.* (19)] nor, thus, Ω, ω, λ and H. The tangent vector to the curve (19) at the (arbitrary) point (Ω_0, ω_0, λ_0) is

$$\left(\frac{d\Omega}{da}\right)_{a=1}\frac{\partial}{\partial\Omega} + \left(\frac{d\omega}{da}\right)_{a=1}\frac{\partial}{\partial\omega} + \left(\frac{d\lambda}{da}\right)_{a=1}\frac{\partial}{\partial\lambda}.$$

Performing the indicated differentiations in (19), and then dropping the suffix zero, we find the vector field

$$(-1+\Omega+2\omega-2\lambda)\Omega\frac{\partial}{\partial\Omega} + (-2+\Omega+2\omega-2\lambda)\omega\frac{\partial}{\partial\omega} + (2+\Omega+2\omega-2\lambda)\lambda\frac{\partial}{\partial\lambda} \tag{20}$$

as the generator of the phase flow in $\mathscr{S}$. It vanishes at exactly four points, the *fixed points* of the flow:

$$\begin{aligned} M &= (0, 0, 0), \\ P_{\mathrm{d}} &= (1, 0, 0), \\ P_{\mathrm{r}} &= (0, 1, 0), \\ S &= (0, 0, 1). \end{aligned} \tag{21}$$

They represent the exceptional models mentioned already in the Introduction: The Milne model M, the parabolic dust (Einstein–de Sitter) and radiation models, P_{d} and P_{r}, and the de Sitter model S. The expanding phases of all other models are represented in $\mathscr{S}$ as (one-dimensional) oriented curves, the orientation corresponding to the arrow of time. This proves that on all non-exceptional models, a triplet (Ω, ω, λ) determines uniquely an instant of time.

There are six noteworthy surfaces in $\mathscr{S}$ which are invariant under the phase flow: The three coordinate planes $\Omega = 0$, $\omega = 0$, $\lambda = 0$, respectively, the zero-curvature plane $\kappa = \Omega + \omega + \lambda - 1 = 0$, as well as the surfaces given by $\lambda = \lambda_1(\Omega, \omega)$, $\lambda = \lambda_2(\Omega, \omega)$, respectively, defined as in (17). For, each of the six conditions $\rho = 0$, $u = 0$, $\Lambda = 0$, $k = 0$, $\lambda = \lambda_1$, $\lambda = \lambda_2$ characterizes a model permanently. The last two surfaces can alternatively be represented by choosing a and ω as parameters and exhibiting Ω and λ as functions of them, using equations (13) and (15). This yields

$$\begin{aligned} \Omega &= \frac{a^4 - \omega(1-a^2)^2}{a(1-a)^2(a+\frac{1}{2})}, \\ \lambda &= \frac{a^2 - \omega(1-a)^2}{2a^2(1-a)^2(a+\frac{1}{2})}, \end{aligned} \tag{22}$$

The 'λ_1-surface' consists of those points which obey equation (22) with $1 < a < \infty$ and the triangular region of the plane $\lambda = 0$ for which $0 \leq \Omega + \omega \leq 1$ (see our discussion in Section 2).

(One easily verifies that the surface is smooth even at the join $\Omega+\omega=1$, $\lambda=0$.) The 'λ_2-surface' is obtained from (22) for $0<a<1$. Since Ω and λ depend linearly on ω for fixed a, both surfaces are ruled by those straight lines which connect the points

$$\Omega=\frac{a^3}{(1-a)^2(a+\frac{1}{2})},$$

$$\omega=0,$$

$$\lambda=\frac{1}{2(1-a)^2(a+\frac{1}{2})} \tag{23}$$

of the plane $\omega=0$ with the points

$$\Omega=0,$$

$$\omega=\frac{a^4}{(1-a^2)^2},$$

$$\lambda=\frac{1}{(1-a^2)^2} \tag{24}$$

of the plane $\Omega=0$, for the respective a-ranges given above. These surfaces are indicated in Fig. 2.

The intersections of the invariant surfaces, also shown in Fig. 2, represent models of the types indicated; the arrows correspond to increasing time and thus (since we only consider the expanding phases of each model) to increasing volume.

To study the general flow pattern in $\mathscr{S}$, consider first those models for which $a\to\infty$. In (19), if $\lambda_0\neq 0$, $[\,]\sim a^4\lambda_0$ whence $(\Omega,\omega,\lambda)\to(0,0,1)=S$. So the orbits of all indefinitely expanding models with $\Lambda\neq 0$ end at S. In fact, S is a 'sink' of the flow, since near S $\lambda_0\neq 0$ and all states belong to models of the type here discussed. If $\lambda_0=0$ but $\Omega_0+\omega_0\neq 1$, $[\,]\sim a^2(1-\Omega_0-\omega_0)$ and $(\Omega,\omega,\lambda)\to(0,0,0)=M$, which is thus the end point of all H models. Similarly, if $\lambda_0=\Omega_0+\omega_0-1=0$ but $\Omega_0\neq 0$ (i.e. for P_{d+r}), $(\Omega,\omega,\lambda)\to P_d$; and, of course, $\lambda_0=\Omega_0+\omega_0-1=\Omega_0=0$ implies $(\Omega,\omega,\lambda)\equiv P_r$.

Consider next those models for which $a\to 0$ in the past. In (19), if $\omega_0\neq 0$ (i.e. for all big-bang models with radiation), $[\,]\to\omega_0$, whence $(\Omega,\omega,\lambda)\to(0,1,0)=P_r$. So all corresponding orbits 'come from' P_r. In fact, P_r is a 'source' of the flow, since near it $\omega_0\neq 0$ and all states belong to models of the type here discussed. Similarly, if $\omega_0=0$ but $\Omega_0\neq 0$ (pure dust big-bang models), $(\Omega,\omega,\lambda)\to P_d$; if $\omega_0=\Omega_0=0$ but $\lambda_0\neq 1$, $(\Omega,\omega,\lambda)\to M$; and of course $\omega_0=\Omega_0=0$, $\lambda_0=1$ implies $(\Omega,\omega,\lambda)\equiv S$. We observe that M and P_d are 'neutral' points of the flow, since orbits both end and start there.

Finally consider those models (O, B, E_1, E_2) for which $\dot{a}\to 0$, $a\to$ constant $\neq 0$ in the future or in the past. Since, by (18), $[\,]=a^4H_0^{-2}H^2$ and $H\to 0$ with $\dot{a}\to 0$, $[\,]\to 0$ and so $(\Omega,\omega,\lambda)\to(\infty,\infty,\infty)$, unless one or more of the parameters vanishes identically. Thus E_1 and E_2 models go to, respectively come from, infinity, while the orbits of O and B models – if for once we consider their entire history – start at P_r, P_d or S, go to infinity, and then double back on themselves.

Fig. 2 illustrates these facts. Note, especially from Figs 2(iv) and (v), that 'most' big-bang models with dust *and* radiation never come close to P_d, which is often held to approximate to the present state of the universe. This is the basis of the 'fine-tuning objection' raised against

standard cosmology. Essentially the same objection can be raised, however, against inflationary cosmologies (see Ellis 1988; Madsen & Ellis 1988; and our remarks at the end of this section). However, whether any particular history is considered as 'improbable' depends on a – metaphysical, since in principle untestable – choice of probability measure on $\mathscr{S}$.

Fig. 2 also shows the six invariant open domains which correspond, rather obviously in light of what has gone before, to certain generic types of models: The region O_{--} below the triangle P_r, P_d, M and the plane $\kappa=0$ (corresponding, as the suffixes indicate, to O models with $\kappa<0$, $\lambda<0$); the region O_{+-} ($\kappa>0$, $\lambda<0$) above $\kappa=0$ and below $\lambda=0$; the region O_{++} between $\lambda=0$ and $\lambda=\lambda_1$; the interior I_- ($\kappa<0$) of the tetrahedron P_r, M, P_d, S; the region I_+ ($\kappa>0$) between the λ_1- and λ_2-surfaces and above the triangle P_r, P_d, S; and the region B above the λ_2-surface.

As shown by equations (7) and (8), the level surfaces of the curvature parameter κ and the deceleration parameter q are planes in $\mathscr{S}$, not shown in Fig. 2 but indicated in Figs 3 and 4. It is perhaps worth observing that in this diagram the level surfaces of all five parameters Ω, ω, λ, κ and q are plane and linearly spaced relative to each other. Only the level surfaces for the ages T_0 of the states of the big-bang models (in units of the Hubble time H_0^{-1}) are curved. They can be computed from the formula

$$T_0=\int_0^1\frac{da}{\dot{a}}=\int_0^1\frac{ada}{[\omega_0(1-a^2)+\Omega_0 a(1-a)-\lambda_0 a^2(1-a^2)+a^2]^{1/2}}, \tag{25}$$

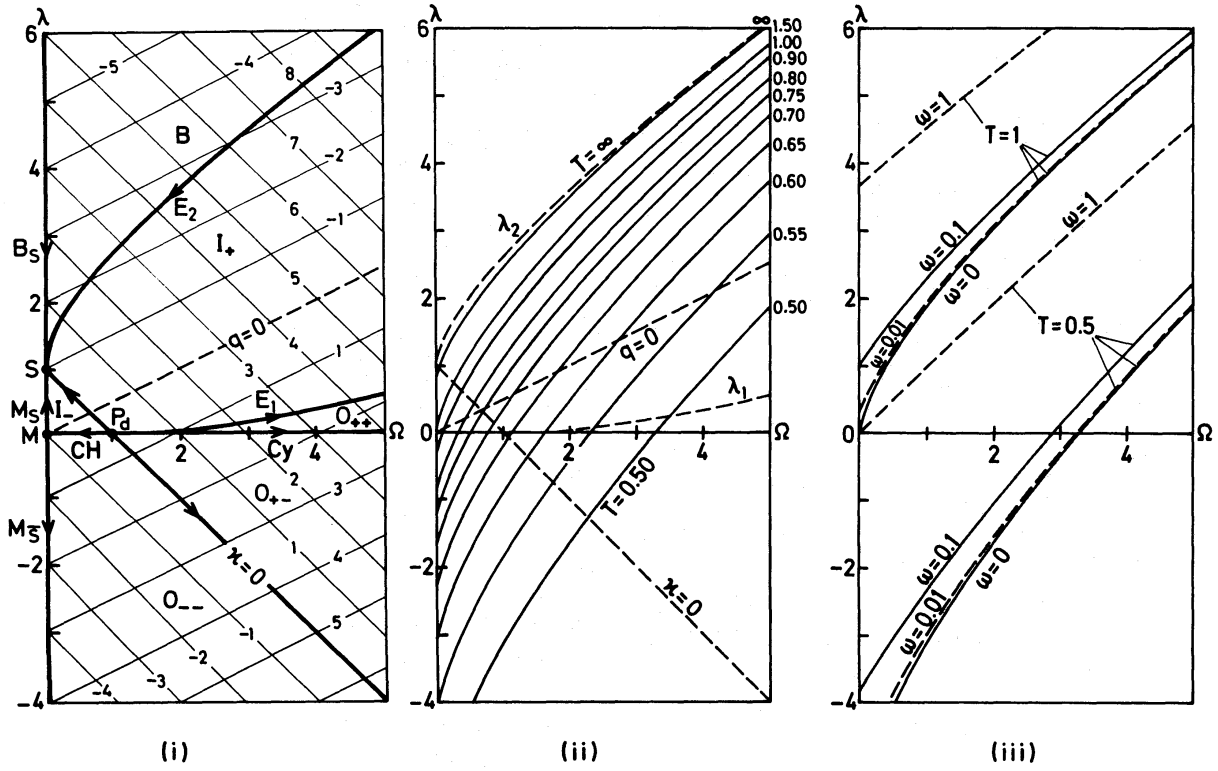

Figure 3. More detailed views of the $\omega=0$ coordinate plane of Fig. 2. (i) Level curves of κ and q at one-unit intervals. E_2 (here, where $\omega=0$) = Eddington–Lemaître model, Cy = cycloidal model, CH = cycloidal hyperbolic model ('cycloid' with circular functions replaced by hyperbolic functions). (ii) Some age contour lines from $T=0.5$ to $T=1.5$ and $T=\infty$ (coincident with E_2). (iii) What happens to two representative age contours of (ii), $T=0.5$ and $T=1.0$, as we move in the ω direction to the planes $\omega=0.01$, 0.1, 1. (The contours at $\omega=0.001$, not shown here, are already indistinguishable, on this scale, from those at $\omega=0$, except within about 0.3 units from the λ axis, where a small visible difference develops.)

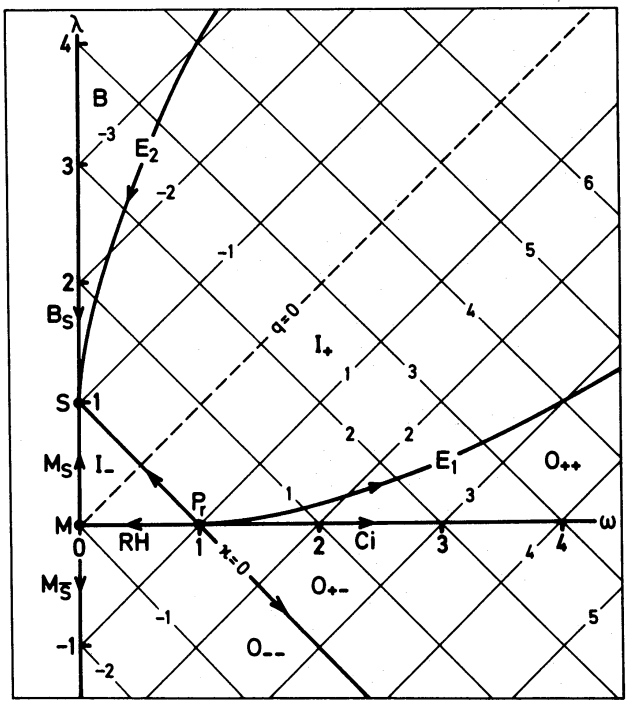

Figure 4. A more detailed view of the $\Omega = 0$ coordinate plane of Fig. 2, with level curves of κ and q at one-unit intervals. *Ci* = circular (Tolman-) radiation model ($t^2 + R^2 = t_0^2$); *RH* = rectangular hyperbolic radiation model ($t^2 - R^2 = t_0^2$); P_r = parabolic radiation model ($R = t^{1/2}$).

which follows from equation (12). T_0 increases monotonically with λ_0 and decreases if ω_0 or Ω_0 increase (see Fig. 3).

The models contained in the coordinate planes $\omega = 0$, $\Omega = 0$ of $\mathscr{S}$ deserve special attention. The former contains all the familiar non-static dust models (see Fig. 3). The intersections of this plane with the λ_1- and λ_2-surfaces are given parametrically by equations (23) for the respective parameter ranges $1 < a < \infty$ and $0 < a < 1$. Elimination of a between those equations yields the cubic

$$(\Omega + \lambda - 1)^3 = \tfrac{27}{4}\Omega^2\lambda, \tag{26}$$

which can, if desired, be solved explicitly for λ by Cardan's method [*cf.* Felten & Isaacman 1986, whose equation (19) is our (26)]. The asymptotic behaviour for $\Omega \to \infty$ of both branches of this cubic curve, in the quadrant $\Omega \geq 0$, $\lambda \geq 0$, is evidently given by $\lambda \sim \frac{1}{2}\Omega$. The (straight) level lines of κ and q, having respective slopes -1 and $\frac{1}{2}$ [*cf.* equations (7) and (8)] are also indicated in Fig. 3, as are some age curves.

By specializing our previous discussion of Fig. 2 to the plane $\omega = 0$, we recognize the latter's main features at once: to begin with, we have the region *I* of inflectional states between the lines $\lambda = \lambda_2$ and $\lambda = \lambda_1$ (augmented by the Ω axis between 0 and 1), the oscillating states *O* below that, and the bouncing states *B* above. The three invariant points (models) *M*, P_d and *S* of Fig. 2 naturally reappear in Fig. 3. The line $\lambda = \lambda_2$, as intersection of two invariant surfaces, represents a single model: the pure-dust E_2 (Eddington–Lemaître) model. Likewise the line $\lambda = \lambda_1$, issuing from P_d: it represents the pure-dust E_1 model. Again, the two segments $0 < \Omega < 1$ and $\Omega > 1$ of the Ω axis are invariant and represent, respectively, the 'cycloidal' hyperbolic ($k = -1$) and the cycloidal ($k = +1$) zero-Λ model (marked *CH* and *Cy*, respectively). *CH*, P_d and *Cy* are often referred to as 'the' three standard models. (For derivations of

closed forms of the expansion functions and further discussion of these and the other special models of Fig. 3 see, e.g. Rindler 1977, pp. 231–234.) The λ axis corresponds to the empty models and exhibits the richest variety, namely three invariant segments separated by the two invariant points M and S. These segments and points are marked in the diagram by representative forms of their expansion functions $R(t)$, these being, from the top down: cosh t, exp t, sinh t, t, sin t. The first three of these models all share (all or part of) the de Sitter spacetime background, while t (i.e. M) occupies part of Minkowski spacetime and sin t part of anti-de Sitter spacetime. Both sinh t and sin t can be regarded as analogues of the Milne model (an expanding ball of test-dust bounded by a spherical light-front) in their respective spacetimes, and are accordingly denoted by M_S, $M_{\bar{S}}$.

The pattern of *states* in the $\omega = 0$ plane as shown in Fig. 3 is qualitatively typical for that in any plane containing the λ axis (though only the two coordinate planes are invariant, i.e. are planes of *models*.) Let us rotate through 90° to the $\Omega = 0$ plane of pure-radiation models (Fig. 4). We then obtain from equations (24), as the analogue of (26), the simple explicit equations

$$\lambda_{1,2} = \omega + 1 \mp 2\sqrt{\omega}. \tag{27}$$

As before, λ_1 (as the lower boundary of inflectional states) is to be taken as zero in the interval $0 \leq \omega \leq 1$ and (27) gives λ_1 for $\omega \geq 1$. The level lines of κ and q now have slopes -1 and 1, respectively [*cf.* equations (7) and (8)], and, as in the case $\omega = 0$, the λ_1- and λ_2-lines ultimately have the same slope as the level lines of q. The invariant segments $0 < \omega < 1$ and $\omega > 1$ of the ω axis corresponds to models whose expansion factor $R(t)$ is part of a rectangular hyperbola (*RH*), respectively of a circle (*Ci*) ('Tolman universe'), while at P_r, $R(t) = t^{1/2}$.

The above discussion should serve to illustrate the 'topologically faithful' nature of the representation of Friedmann–Lemaître states and models in $\mathscr{S}$, in the sense that 'related' states and models are proximate, and that generic models go over smoothly into their various limiting cases.

Realistic models presumably begin their 'standard phase', i.e. the one with $\Lambda = \Lambda_0 = \text{constant}$, somewhere close to P_r (see Section 5 below), cross the plane $\Omega = \omega$ which separates radiation dominated ($\omega > \Omega$) from matter dominated ($\omega < \Omega$) states, and reach a present state somewhere near the region of the (Ω, λ) plane defined by the inequalities $-1.6 < \lambda_0 - \Omega_0 < -0.4$, $-1.5 < \lambda_0 + \Omega_0 < 7.1$ (Loh 1986) and $0.05 < \Omega_0$ (Peebles 1986). Measurements of the Hubble constant and the temperature of the microwave background radiation show that, at least for *it*, the radiation density parameter is now very small, $3 \times 10^{-5} \lesssim \omega_0 \lesssim 3 \times 10^{-4}$. But these limits on ω_0 could be drastically unrealistic if the 'missing' cosmic mass contained an as yet unidentified kind of radiation.

Inflationary models assume a change of Λ associated with a phase transition from a false to a true vacuum state (see, e.g. the reviews by Guth and Linde in Hawking & Israel 1987). The histories of such models can be represented in $\mathscr{S}$ by two orbits with $\Lambda = \text{constant}$, an early and a late one, connected by a curve representing the phase transition. Let the latter be idealized, for simplicity, as a discontinuous change of state (Ω, ω, λ). Then the general-relativistic junction conditions require $\kappa = \Omega + \omega + \lambda - 1$ to be continuous even at this transition (Ellis 1987). The early piece of the phase orbit including the inflation era ($\ddot{a} > 0$) must have $\Lambda = \Lambda_i = \text{constant} > 0$; it can either be part of an inflectional model, starting close to P_r (see Fig. 2) and moving up towards S, or be part of a bouncing model coming down towards S. Having arrived near S it jumps down to a state near P_r, keeping its κ value, and then follows part of an orbit $\Lambda = \Lambda_0$ (often assumed to be zero) of the type described in the last paragraph. The phase diagrams of Fig. 2, especially (iv), show that if $\Lambda_0 = 0$, any pair Ω_0, ω_0 of 'present' values can be reconciled with either kind of pre-transition orbit. This confirms a conclusion recently reached by Ellis (1988). It may also be remarked that the phase-plane diagrams displayed in Madsen &

Ellis (1988), showing for various inflationary and non-inflationary models the graphs of the function $a \to \Omega_{\text{total}} := \Omega + \omega + \lambda$ [*cf.* (18)], complement the diagrams given in the present paper.

4 An alternative representation of the phase flow in the interior of $\mathscr{S}$ based on invariants

Consider now the three alternative dimensionless parameters K, L, x for $\mathring{\mathscr{S}}$, the interior of $\mathscr{S}$, defined for

$$\Omega > 0, \qquad \omega > 0 \tag{28}$$

as follows:

$$K := \frac{\omega^3 \lambda}{\Omega^4},$$

$$L := \frac{\omega(\Omega + \omega + \lambda - 1)}{\Omega^2} = \frac{\omega\kappa}{\Omega^2},$$

$$x := \frac{\Omega}{\omega}. \tag{29}$$

(These definitions cannot be extended to the boundaries $\Omega = 0$ or $\omega = 0$ of $\mathscr{S}$.) K and L are constant along any model orbit, and indeed they are constructed from the lhs of equations (11) for just this purpose. Also from (11), x is proportional to a for each model and thus serves as parameter *along* each orbit, increasing with a. K and L could be expected to serve as convenient coordinates for $\mathring{\mathscr{M}}$, the interior of the space $\mathscr{M}$ of models (orbits). But we shall see that they do not do this uniquely: they coordinate either the big-bang models or the non-big-bang models.

To invert equations (29), we first note that they imply

$$\frac{\lambda}{\omega} = Kx^4,$$

$$\frac{\kappa}{\omega} = Lx^2, \tag{30}$$

and when these expressions are substituted into (7) we find

$$\frac{1}{\omega} = Kx^4 - Lx^2 + x + 1 =: P(x). \tag{31}$$

In terms of $P(x)$ as here defined, equations (29)(iii), (31) and (30)(i) then yield, respectively,

$$\Omega = \frac{x}{P(x)},$$

$$\omega = \frac{1}{P(x)},$$

$$\lambda = \frac{Kx^4}{P(x)}. \tag{32}$$

While both K and L evidently range from minus to plus infinity, x is restricted via (28) and (32) to values satisfying

$$x > 0 \quad \textit{and} \quad P(x) > 0. \tag{33}$$

Given $(K, L) \in \mathbb{R}^2$, the set of x values obeying the inequalities (33) consists of either one or two open intervals: the first is $0 < x < x_1(K, L)$, where $x_1(K, L)$ denotes the smallest positive zero of $P(x)$ or ∞. A second range $x_2(K, L) < x < \infty$ $(x_1 \leq x_2)$ may or may not exist.

To justify these assertions, note first:

$$P(0) = P'(0) = 1,$$

$$P''(0) = -2L,$$

$$P''(x) = 0 \longleftrightarrow x = \pm\sqrt{L/6K} \qquad (K \neq 0).$$

Thus at $x = 0$, $P(x)$ is positive with positive slope and, if $L > 0$, it is concave down. On the (fourth) closed quadrant $K \geq 0$, $L \leq 0$ of the K, L plane, $P(x) > 1$ for all $x > 0$, so $P(x)$ has no positive zeros and $x_1 = \infty$. On the open half-plane $K < 0$, $P(x) \sim -|K|x^4$ for large x, so there must be one or three positive zeros (possibly with coincidences); but there cannot be three, since $P''(x)$ has at most one positive zero, so there is exactly one, $x_1 < \infty$. On the open line segment $K = 0$, $L > 0$, $P(x) \sim -Lx^2$, so again there is exactly one positive zero $x_1 < \infty$. Lastly, consider the open (first) quadrant $K > 0$, $L > 0$. $P''(x)$ now has one positive zero, hence $P(x)$ has two positive zeros (possibly coincident) or none. For the critical case of coincident zeros, we need

$$P(x) = P'(x) = 0. \tag{34}$$

Let us combine these equations to form the *quadratic* equation

$$xP'(x) - 4P(x) = 2Lx^2 - 3x - 4 = 0, \tag{35}$$

whose positive root is

$$x_0 = \frac{3 + \sqrt{9 + 32L}}{4L} =: x_0(L). \tag{36}$$

This x_0 satisfies (34) if and only if $P'(x_0) = 0$, i.e. if

$$K = \frac{8L^3(1 + \sqrt{9 + 32L})}{(3 + \sqrt{9 + 32L})^3} =: K_c(L). \tag{37}$$

Hence $K = K_c(L)$ is the condition for $P(x)$ to have the *critical* shape of dipping down from 1 at $x = 0$ to a double zero at $x = x_0$ and then rising to ∞. If, keeping L fixed, we increase K, it is clear from (32) that for every value of x, $P(x)$ increases, so we get a curve lying entirely above the critical one, having no positive zeros. Likewise $K < K_c(L)$ gives rise to a curve below the critical one, having two distinct zeros x_1 and x_2, outside of which $P(x) > 0$.

The classification of models given in Section 3 can now be restated – at least for models in $\mathring{\mathscr{S}}$ – in terms of (K, L, x), where all orbits appear straightened out, being now line segments in the x direction (see Fig. 5 and Table 1). This representation exhibits more explicitly the orbit structure of $\mathring{\mathscr{S}}$, and the relation between $\mathring{\mathscr{S}}$ and the two-dimensional manifold $\mathring{\mathscr{M}}$ of models.

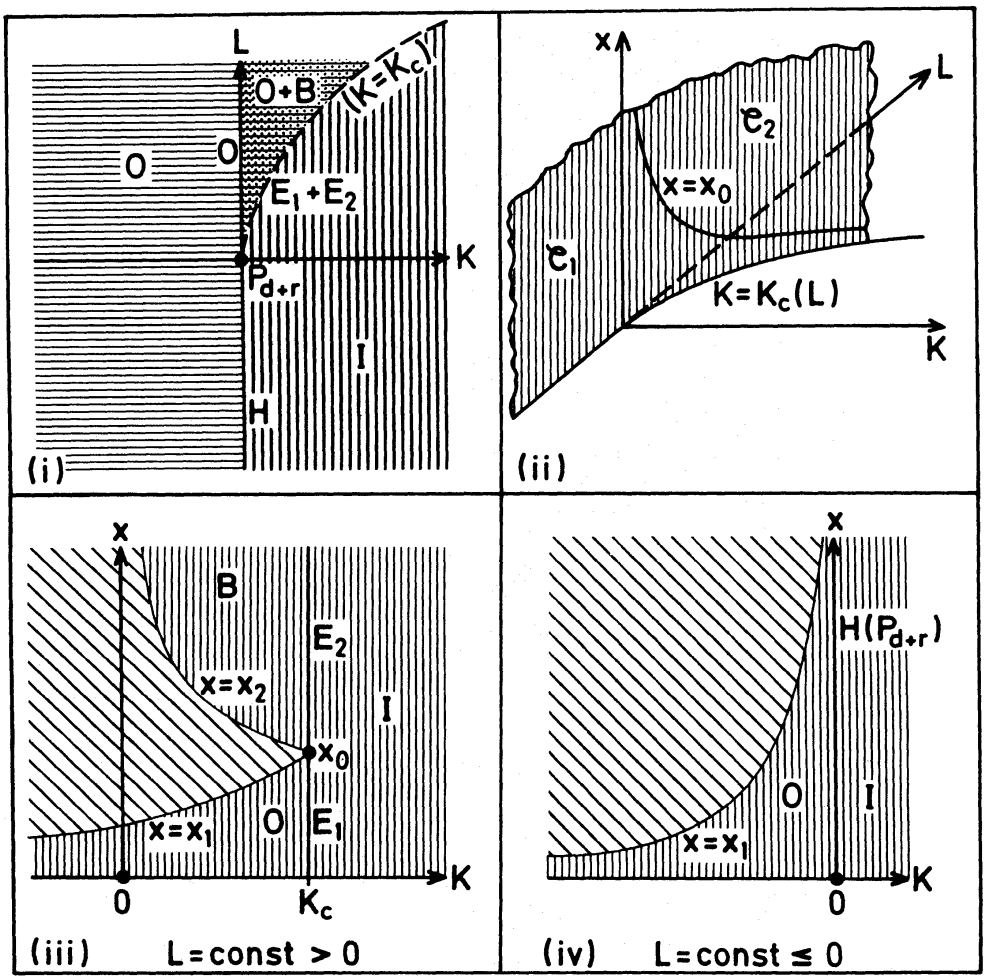

Figure 5. States of non-static dust-and-radiation ($\Omega>0$, $\omega>0$) FL models in (K, L, x) coordinates. Each model corresponds to a line segment in the x direction above the K, L plane shown in (i), x being proportional to $a(T)$. To the left of the surface $\mathscr{C}$ shown in (ii) there is a 'roof' over the K, L plane, $x=x_1$ [shown in cross-section in (iii) and (iv)], which bounds segments corresponding to O models above. Between $\mathscr{C}$ and the $K=0$ coordinate plane there is a second roof, $x=x_2$, shown in (iii), which bounds segments corresponding to B models below. The segments to the right of $\mathscr{C}$ have no roof and correspond to I models. As L decreases to zero and negative values, the point (K_c, x_0) in (iii) moves left towards $K=0$ and up towards $x=\infty$, 'taking' the x_1-surface with it, which results in its typical shape for $L\leq 0$ shown in (iv).

Orbits rising from the K, L plane to infinity generally represent inflectional models [whose $a(T)$ ranges from zero to infinity]; similarly, those rising to a finite height x_1 generally represent oscillating models, while those 'starting' at a height x_2 generally represent bouncing models. Since the signs of K and L coincide with those of Λ and k, respectively, Fig. 5 also provides the subclassification according to these signs. The origin and the negative L axis represent special models: the origin $K=L=0$ ($\lambda=0$, $\kappa=0$), the parabolic dust-plus-radiation model P_{d+r}, and the negative L axis ($\lambda=0$, $\kappa<0$), the hyperbolic models H. The invariant surfaces $\lambda=\lambda_j(\Omega, \omega)$ separating O- from I-, respectively B-models, now reappear as the lower ($j=1$) and upper ($j=2$) parts of the cylindrical surface $\mathscr{C}$ whose generators are orbits and whose cross section in the K, L plane is the curve $K=K_c(L)$ (stippled in Fig. 5) augmented by the negative L axis; $\mathscr{C}$ is cut by the line $x=x_0(L)$ into a lower part $\mathscr{C}_1$ generated by the H-, P_{d+r}- and E_1-models, and an upper part $\mathscr{C}_2$ generated by the E_2-models. [Note that the line $x=x_0(L)$ exists only above the stippled curve, and that $x_0(L)\to\infty$ as $L\to 0$ while $x_0(L)\to 0$ as $L\to\infty$.) The curve separating $\mathscr{C}_1$ from $\mathscr{C}_2$ is also the common edge of the two surfaces $x=x_j(K, L)$ which lie above the region

Table 1. Classification of models in $\mathring{\mathscr{M}}$ by use of the coordinates (K, L, x). This table should be used in conjunction with Fig. 5.

$0 < K < K_c(L)$,	$L > 0$,	$x > x_2(K,L)$:	B
$0 < K < K_c(L)$,	$L > 0$,	$x < x_1(K,L)$:	O
$K \leq 0$,	$L > 0$:		O
$K < 0$,	$L \leq 0$:		O
$K > K_c(L)$,	$L > 0$:		I
$K > 0$,	$L \leq 0$:		I
$K = K_c(L)$,		$x < x_0(L)$:	E_1
$K = K_c(L)$,		$x > x_0(L)$:	E_2
$K = 0$,	$L < 0$:		H
$K = 0$,	$L = 0$:		P_{d+r}

marked $O+B$ in Fig. 5. The x_1-surface approaches finite values above the positive L axis (which, however, become unbounded as $L \to 0$) while the x_2-surface approaches infinity all along the positive L axis. The x_1- and x_2-surfaces, their edge, and the region between them, do not belong to the phase space.

The invariants K and L go a long way to coordinatize the open model space $\mathring{\mathscr{M}}$, i.e. those models which contain *both* dust and radiation ($\Omega > 0$, $\omega > 0$). They do form a global coordinate system on the open submanifold of big-bang models and also on the submanifold-with-boundary of non-big-bang models; but not on $\mathring{\mathscr{M}}$ itself: some O- and B-models as well as E_1- and E_2-models have the same values of K and L.

(It may be noted that $\mathring{\mathscr{M}}$, viewed as a quotient space constructed from $\mathring{\mathscr{S}}$, is not a Hausdorff-, but only a T_1-space: two points representing E_1 and E_2 models with the same K, L values do not have disjoint neighbourhoods, only neighbourhoods not containing the other point.)

Finally we observe that it is possible to carry out an analogous classification of models in terms of invariants in the planes $\omega = 0$ and $\Omega = 0$, respectively. We shall not carry this out here, but content ourselves with indicating that the essential tools for that are the invariants

$$M := \frac{(\Omega + \lambda - 1)^3}{\Omega^2 \lambda} = \frac{\kappa^3}{\Omega^2 \lambda} \tag{38}$$

for $\omega = 0$, $\Omega > 0$, $\lambda \neq 0$, and

$$N := \frac{(\omega + \lambda - 1)^2}{\omega \lambda} = \frac{\kappa^2}{\omega \lambda} \tag{39}$$

for $\Omega = 0$, $\omega > 0$, $\lambda \neq 0$.

5 Observational restrictions; in particular, the inevitability of a big bang*

The classification of the dust-plus-radiation models of Section 2 (see especially Fig. 1) shows that the expansion function $a(T)$ of all non-empty models without a big bang have a positive

*A brief note concerning the content of this section, without complete proof, has been published by Börner & Ehlers (1988).

lower bound $a_* < 1$ which is determined by $\dot{a} \to 0$ as $a \to a_*$. Lemaître's equation (12) thus implies

$$\lambda_0 = \frac{1 + \omega_0(a_*^{-2} - 1) + \Omega_0(a_*^{-1} - 1)}{1 - a_*^2}. \tag{40}$$

Also for those models, as $a \to a_*$, $\ddot{a}$ tends to a non-negative limit which vanishes only for the E_2-models. Therefore [see equation (15) and the subsequent paragraph],

$$\lambda_0 \geq \frac{2\omega_0 + a_* \Omega_0}{2a_*^4}. \tag{41}$$

Combining (40) and (41) we obtain

$$\omega_0(1 - a_*^2)^2 + \Omega_0 a_*(\tfrac{1}{2} - \tfrac{3}{2} a_*^2 + a_*^3) \leq a_*^4, \tag{42}$$

with equality holding for the E_2- models only. As a simple consideration shows, not only the first but also the second term on the lhs of (42) is non-negative, so that we can obtain inequalities for Ω_0 and ω_0 separately. If instead of a_* we use the corresponding *maximal redshift* $z_* = a_*^{-1} - 1$ *observable today* [*cf.* (10)(ii)], these inequalities are

$$\Omega_0 \leq \frac{2}{z_*^2(z_* + 3)}, \tag{43}$$

$$\omega_0 \leq \frac{1}{z_*^2(z_* + 2)^2}. \tag{44}$$

The equality signs in (43) and (44) hold precisely for E_2-dust-models and E-radiation-models, respectively. (The equation for the E_2-dust-model was obtained by Crilly 1968.)

According to all recent determinations of which we are aware (e.g. Metzger & Schmid-Burgk 1983; Peebles 1986), $\Omega \geq 0.05$; also six quasar redshifts z larger than 4 have been observed (Warren *et al.* 1987). Together , these values (for Ω_0, even $\Omega_0 > 0.02$ would suffice) violate the inequality (43), which shows that the (generalized Eddington–Lemaître) models E_2 as well as the bouncing models can be excluded as models of the (late phase of the) real universe. Consequently, under our assumptions, a big bang, i.e. an early 'state' or phase with densities and temperatures so large that classical theory does not apply, is unavoidable, whether or not there was an inflationary era prior to that. In contrast to previous arguments (see, e.g. Hawking & Ellis 1968), the present one makes no reference to the cosmic background radiation nor does it require any prior assumptions about Λ ; it hinges instead on the assumption that quasar redshifts are cosmological, and on the empirical datum that $\Omega_0 > 0.02$.

Since B and E_2 models have been eliminated, realistic models must obey

$$\lambda < \lambda_2[\Omega, \omega], \tag{45}$$

and consequently, by (7) and (8),

$$q > \tfrac{1}{2}\Omega + \omega - \lambda_2[\Omega, \omega], \tag{46}$$

$$\kappa < \Omega + \omega - 1 + \lambda_2[\Omega, \omega]. \tag{47}$$

More stringent restrictions are provided by the level surfaces of T_0, but it is not our purpose here to enter into such an analysis.

An argument very similar to the one eliminating B and E_2 models can be applied to restrict – though not eliminate – inflectional models with inflection in the past. For such a model, let $a_* < 1$ denote the expansion function $a(T)$ at the moment of inflection. Then, since in this case $\dot{a}_*^2 \geq 0$, whereas $\ddot{a}_* = 0$, we obtain analogues of (40) and (41), but with '$\lambda \leq$' and '$\lambda =$' instead of '$\lambda =$' and '$\lambda \geq$', respectively. Consequently the three inequalities (42)–(44) apply as before, except that z_* now denotes the maximal redshift of the *accelerating* phase. If we assume $\Omega_0 \geq 0.05$ we find $z_* \leq 2.66$. This shows that if the actual universe is inflectional, its inflection must have occurred *after* its expansion function reached $(3.66)^{-1} = 0.27$ of its present value. With $\Omega_0 = 1$ that ratio increases to 0.58. At any rate, inflection would have occurred, if at all, well after galaxy formation.

Acknowledgments

We thank Dr Martin Lottermoser for his expertise and generous help in computer-generating Fig. 3. WR is grateful to the Division of Astrophysics of the Max-Planck-Institute for Physics and Astrophysics in Garching for their hospitality during the course of this work.

References

Börner, J. & Ehlers, G., 1988. *Astr. Astrophys.*, **204,** 1.
Crilly, A. J., 1968. *Mon. Not. R. astr. Soc.*, **141,** 435.
Ellis, G. F. R., 1987. *Astrophys. J.*, **314,** 1.
Ellis, G. F. R., 1988. *Classical Quantum Grav.*, **5,** 891.
Felten, J. E. & Isaacman, R., 1986. *Rev. mod. Phys.*, **58,** 689.
Hawking, S. & Ellis, G. F. R., 1968. *Astrophys. J.*, **152,** 25.
Hawking, SD. & Israel, W., 1988. *300 Years of Gravitation,* Cambridge University Press.
Loh, E. D., 1986. *Phys. Rev. Lett.*, **57,** 2865.
Madsen, M. S. & Ellis, G. F. R., 1988. *Mon. Not. R. astr. Soc.*, **234,** 67.
Metzger, P. G. & Schmid-Burgk, J., 1983. *Mitt. Astron. Ges.*, **58,** 31.
Misner, C. W., Thorne, K. S. & Wheeler, J. A., 1973. *Gravitation,* Freeman and Co., San Francisco.
Peebles, P. J. E., 1986. *Nature,* **321,** 27.
Rindler, W., 1977. *Essential Relativity,* 2nd edn, Springer-Verlag, New York.
Robertson, H. P., 1933. *Rev. mod. Phys.*, **5,** 62.
Stabell, R. & Refsdal, S., 1966. *Mon. Not. R. astr. Soc.*, **132,** 379.
Warren, S. J., Hewett, P. C., Owen, P. S. & Irvin, M. J., 1987. *Nature,* **330,** 453.
Wolf, J. A., 1972. *Spaces of Constant Curvature,* 2nd edn, Berkeley University Press.