Understanding Polaritonic Chemistry from Ab Initio Quantum Electrodynamics

Published as part of the Chemical Reviews virtual special issue “Polaritonic Chemistry”.

Michael Ruggenthaler,* Dominik Sidler,* and Angel Rubio*

ABSTRACT: In this review, we present the theoretical foundations and first-principles frameworks to describe quantum matter within quantum electrodynamics (QED) in the low-energy regime, with a focus on polaritonic chemistry. By starting from fundamental physical and mathematical principles, we first review in great detail ab initio nonrelativistic QED. The resulting Pauli-Fierz quantum field theory serves as a cornerstone for the development of (in principle exact but in practice) approximate computational methods such as quantum-electrodynamical density functional theory, QED coupled cluster, or cavity Born–Oppenheimer molecular dynamics. These methods treat light and matter on equal footing and, at the same time, have the same level of accuracy and reliability as established methods of computational chemistry and electronic structure theory. After an overview of the key ideas behind those ab initio QED methods, we highlight their benefits for understanding photon-induced changes of chemical properties and reactions. Based on results obtained by ab initio QED methods, we identify open theoretical questions and how a so far missing detailed understanding of polaritonic chemistry can be established. We finally give an outlook on future directions within polaritonic chemistry and first-principles QED.

CONTENTS

1. Introduction
2. A Theory of Light and Matter: Quantum Electrodynamics
   2.1. Relativistic Origins
   2.2. Quantizing the Light Field
   2.3. Coupling Light and Matter
3. The Pauli-Fierz Quantum-Field Theory
   3.1. Nonrelativistic QED
   3.2. Mathematical Properties of the Theory
   3.3. Approximations
      3.3.1. Cavity as Modification of Local Mode Structure: Dipole Approximation
4. First-Principles Approaches to Nonrelativistic QED
   4.1. Quantum-Electrodynamical Density Functional theory
   4.2. Exact Results
   4.3. Quantum Electrodynamics Coupled Cluster Theory for Electronic Strong Coupling
   4.4. Nuclear/Ionic Dynamics in the Generalized Born-Huang Picture
5. Polaritonic Chemistry from First-Principles
   5.1. Restructuring the Electromagnetic Field Modes
      5.1.1. Modified Fluctuations and Fields
5.1.2. Chemical Consequences of Cavity-Restructured Modes
6. Conclusion and Outlook

Received: November 10, 2022
1. INTRODUCTION

"Until the beginning of the 20th century, light and matter have been treated as different entities, with their own specific properties [...] The development of quantum mechanics has enabled the theoretical description of the interaction between light-quantum and matter."

M. Hertzog in ref 1.

Chemistry investigates, very broadly spoken, how matter arranges itself under different conditions (temperature, pressure, chemical environment, etc.) and how these arrangements lead to various functionalities and phenomena. The basic building blocks of chemical systems, as we understand them today, are the various atoms of the periodic table of elements. Combining these basic building blocks then leads to the formation of molecules and solids, and the arrangement of the atoms determines much of the emerging properties of these complex matter systems. Light, or more generally, the electromagnetic field, usually appears in this context in two distinct capacities: First, as an external (classical) agent that drives the matter system out of equilibrium. External driving is then used to either spectroscopically investigate matter properties, such as when recording an absorption or emission spectrum, or to force the matter system into a different (transient) state. Second, as a (quantized) part of the system, such as in the case of the longitudinal electric field between two charged particles, which gives rise to the Coulomb interaction and determines how the atoms are arranged.

Light as an external, classical probe and control field is widely used in chemistry nowadays. However, the potential to employ the quantized light field as part of the system to modify and probe chemical properties has only began to be explored in the last years. In order to achieve control over the internal light field one can use photonic structures, such as optical cavities, and in this way control the local electromagnetic field of a molecular system. The resulting restructuring of the electromagnetic modes has very fundamental consequences, since it changes the building blocks of light: the electromagnetic vacuum modes and with this the notion of photons in quantum electrodynamics. Keeping in mind that the interaction between charged particles is mediated via the exchange of photons, it becomes clear that such modifications can in principle influence the properties of atomic, molecular and solid-state systems. Even more so, if we realize that the basic building blocks of matter (electrons, nuclei/ions, atoms,...) are themselves hybrid light–matter systems that depend on the photonic environment (see also discussion after eq 1).

Although optical cavities have been used in atomic physics and quantum optics routinely since several decades to interrogate and change the behavior of (an ensemble of) atoms, it came as a surprise to many that cavities could also influence complex chemical and solid-state processes. The main reason being that in quantum optics, or more precisely in cavity and circuit quantum electrodynamics (QED), which focus on the properties of the photons and a limited set of matter degrees of freedom, often ultralow temperatures and ultrahigh vacuua are needed in order to observe the influence of the changed electromagnetic vacuum modes. Such very specific external conditions are not often considered in chemistry and materials science, and hence, it was assumed that there would be no observable effect on chemical and material properties upon changing the photonic environment at ambient conditions. Yet there is by now a multitude of seminal experimental results that show that indeed the restructuring of the electromagnetic environment by optical cavities can influence chemical and material properties at ambient conditions, even if there is no external illumination and the effects are driven mainly by vacuum and thermal fluctuations (for an overview, see various reviews, e.g., refs 1, 5, 14, 36–47). We here only highlight, as exemplifications, changes in energy and charge transport, the appearance of exciton-polariton condensates at room temperature, and the modification of the phases of solids. In the following, we will focus on changes in chemical properties of (finite) molecular systems upon modifying the photonic environment and do not go into detail on changes observed and induced in extended solid-state systems.

This new flavor of chemistry, which uses the modification of the photonic environment as an extra control knob, has been named QED or polaritonic chemistry. The latter notion is derived from the quasi particle polariton, which is a mixed light–matter excitation (see also Figure 4), and whose appearance in absorption or emission spectra is often assumed to be a prerequisite for observing changes in chemical properties. Polaritonic chemistry is a highly interdisciplinary field with often conflicting perspectives on the same physical concepts. From a (quantum) optics perspective, for instance, the role of light and matter is reversed compared to chemistry. Matter is used to either interrogate or change the properties of the electromagnetic field. This clash of perspectives, which arises due to the artificial subdivision into different research fields and their unification via QED, see also Figure 1), makes it a scientifically very rewarding field of research since it constantly challenges one’s basic conceptions. A plethora of theoretical methods from (quantum) optics and (quantum) chemistry are employed and combined to capture and understand the emerging novel functionalities when changes in the electromagnetic environment lead to strong coupling between light and matter.

While (quantum) optics methods are geared to capture details of the electromagnetic field and photonic states, the (quantum) chemical methods are naturally focused on a detailed description of the matter system. Many currently employed combinations of such methods are able to capture certain effects, but fail in important situations, such as to describe (even only qualitatively) the observed changes in ground-state chemical reactions under vibrational strong coupling. On a first glance, owing to the complexity of the systems under study (a large number of complex molecules in solvation at ambient conditions strongly coupled to an optical cavity with many photonic modes), this might not come as a surprise, since already the accurate theoretical description of a single complex molecule in vacuum at zero temperature is highly challenging. Even simple working principles of polaritonic chemistry, which single out the most
important ingredients to control chemistry via changed electromagnetic environments, remain elusive so far. On a second, more careful glance, however, there might be a more fundamental reason for why currently employed approaches, which combine (quantum) chemistry and (quantum) optics methods, are not able to describe some of the experimentally observed effects. Our most fundamental description of how light and matter interact, QED, does not allow for a strict distinction between light and matter (see also Figure 1). Considering one aspect without the other can lead to inconsistencies, and for a consistent description always both (quantum light and quantum matter aspects) must be treated at the same time.

Indeed, if we reconsider the basic building blocks of matter from a QED perspective, we realize that already electrons and atoms are hybrid light–matter objects themselves, and their properties depend on various assumptions. Take, for instance, the hydrogen atom as described by the nonrelativistic Schrödinger equation in Born–Oppenheimer approximation in SI units (used throughout this review)

$$\left(-\frac{\hbar^2}{2m_e}\nabla^2 - \frac{e^2}{4\pi\varepsilon_0|r|}\right)\Psi(r) = E\Psi(r)$$

(1)

where $\hbar$ is the reduced Planck’s constant, $m_e$ is the physical mass of the electron, $e$ is the elementary charge, and $\varepsilon_0$ is the permittivity of the free electromagnetic vacuum. However, from a QED perspective, the electron of a hydrogen atom has a mass that depends on the structure of the electromagnetic vacuum surrounding the atom, and also the Coulomb attraction depends on the form of the surrounding electromagnetic vacuum. Indeed, the physical mass of the electron has two contributions

$$m_e = m + m_{ph}$$

(2)

where the bare mass $m$ depends on how the electromagnetic vacuum modes decay when going to higher and higher frequencies (ultraviolet regularization) and the photon-induced mass $m_{ph}$ comes from the energy due to the interaction of a moving electron with the photon field (see discussion in Sections 3.2 and 3.3 for more details). In addition, the form as well as the strength of the Coulomb interaction is determined solely by the structure of the vacuum modes (see the discussion in Section 3.3 for more details). In other words, what we call a hydrogen atom is defined with respect to a specific photonic environment, i.e., in this case the free electromagnetic vacuum. Similarly, the photonic environment dictates how a laser or thermal radiation interacts with matter. Hence, it becomes clear that when we restructure the electromagnetic environment with the help of an optical cavity or other setups, we might need to rethink what are the basic building blocks of matter, which statistics they obey, how they interact among each other and how they couple to external perturbations.

Admittedly, having in mind the many other aspects that might have an influence in QED chemistry (see also Section 5), such fundamental considerations might seem on a first glance like a theoretical nuisance. However, it is important to realize which assumptions are made and which theoretical inconsistencies (at least with respect to ab initio QED, see Appendix A for mathematical details) can arise when combining methods from (quantum) optics and (quantum) chemistry or electronic structure theory. Especially, since we do not yet have simple and reliable rules for how polaritonic chemistry operates, what are the basic factors that determine the observed changes and how to control them. Furthermore, in recent years, theoretical methods have been devised that avoid the common a priori division into light and matter, allowing approximate solutions to QED in the low-energy regime directly. These first-principles QED methods have already provided important insights into certain aspects of polaritonic chemistry and strong light–matter coupling for molecular and solid-state systems.

In this review, we will focus on these first-principles QED methods and on the basic ab initio description of coupled light–matter systems under the umbrella of QED in the low-energy regime. We do not attempt to discuss the many alternative theoretical methods successfully applied within polaritonic chemistry, but refer the interested reader to various available reviews on this topic, e.g., refs 37, 38, 45. The considerations presented here allow us to address several important (and often very subtle) fundamental topics that arise in the context of describing polaritonic chemistry and materials science and that are decisive to find the main physical mechanisms observed in experiment. The first main question to answer is how to devise a (physically and mathematically) consistent theory of interacting light and matter that treats all basic degrees of freedom of the low-energy regime, i.e., photons, electrons and nuclei/ions, on the same quantized and nonperturbative footing. We will give the basic principles and a concise derivation of such a theory in Section 2 and discuss the resulting Hamiltonian formulation for fundamentally polaritonic quantum matter in Section 3. The next important topic that arises is how the gauge choice influences what we call light and what we call matter. This topic has a direct impact on consistently combining methods from quantum optics and quantum chemistry. As we discuss in more detail at the end of Section 3.2, this topic has provoked many debates, and gauge-inconsistencies can even predict wrong and unphysical effects. The next main question is how to find approximations that allow a reduction of complexity and a straightforward combination of different theoretical methodologies without introducing too many uncontrolled assumptions. We will discuss this in Section 3.3 and specifically highlight the long
wavelength approximation and its implicit assumptions. Sometimes the implicit assumptions of this common approximation lead to misunderstandings and can therefore be a barrier for new people in the field of QED chemistry and materials sciences. A further important issue is how changing the photonic environment leads to modified vacuum and thermal fluctuations, specifically when considering changes of chemical properties under ambient conditions. We highlight under which conditions the modified vacuum or thermal fluctuations become important in Section 5.1 and might induce noncanonical equilibrium conditions for the matter subsystem. A final question to address in polaritonic chemistry is then the difference between single-molecule strong coupling, also called local strong coupling, and collective strong coupling. We discuss the topic of local/collective strong coupling in Section 5.2, and we highlight how an effective single-molecule picture suggests itself.

Despite the internal complexity and depth of this review, we try to keep it structured modularly, and the different sections are largely self-contained. This will help the reader, allowing them to, for instance, skip the first few sections, which detail the theoretical foundations of ab initio QED, and jump directly to the later sections which focus more on polaritonic chemistry. Yet a better understanding of many arguments (as highlighted above) necessitate detailed discussions, and hence we have provided many cross-links between various sections. In Section 2, we give a concise introduction into QED with a focus on the description of the electromagnetic field. In Section 3, we introduce the basic Hamiltonian of ab initio QED, discuss its many important properties, and provide its most commonly employed approximations. In Section 4, we discuss various first-principles QED methods. In Section 5, we discuss polaritonic chemistry from an ab initio perspective. Finally, in Section 6, we give a conclusion and outlook on how to influence photonic environment as an extra control knob to influence chemical and material properties. We note that we also provide an extensive appendix that addresses many subtle mathematical details about ab initio QED, which become important when developing computationally highly efficient ab initio methods, such as quantum-electrodynamical density functional theory or similar approaches.

2. A THEORY OF LIGHT AND MATTER: QUANTUM ELECTRODYNAMICS

"In a hydrogen atom an electron and a proton are bound together by photons (the quanta of the electromagnetic field). Every photon will spend some time as a virtual electron plus its antiparticle, the virtual positron [...]"

G. Kane in ref 67.

QED is a cornerstone of modern physics, and Feynman, Tomonaga and Schwinger were awarded the Nobel prize in physics in 1965 for their contributions to this theory.68 It tells us on the most fundamental level how light and charged particles interact and how their coupling leads to the emergence of the observable electrons/positrons and photons.69–71 The beauty of QED is that it can be derived from a few very basic principles. However, it is also plagued by several mathematical issues that restrict the applicability of full QED to perturbative high-energy scattering processes.72 Yet, in certain limits, most notably when the charged particles are treated nonrelativistically, QED allows for a beautiful and mathematically well-defined formulation that is very similar to standard electronic quantum mechanics.72 The resulting nonrelativistic QED theory in Coulomb gauge will form the foundation of ab initio QED chemistry and will be discussed in Section 3. But before, we will briefly summarize how QED can be derived from basic principles.

2.1. Relativistic Origins

There are different formulations of the basic equations of QED as well as various different ways to derive them.73–75,69,70 e.g., in a Lagrangian description a formulation in terms of path integrals and associated scattering amplitudes suggests itself.71 Let us follow here a Hamiltonian route that at the same time highlights that both sectors of the theory, that is, the light and the matter parts, follow from the same reasoning and that the coupling between the sectors enforces a strong consistency between the light and matter sector. As a first step, we want the matter as well as the light sector to individually obey special relativity in the form of the energy-momentum relation:71,112

\[ E^2 = m^2c^4 + p^2c^2 \]  

(3)

This relation can be derived from the assumption of a highest possible velocity \( c \) which we call the speed of light in vacuum. We note that eq 3 implies that we think about the flat (Euclidean) space \( \mathbb{R}^4 \) or its extension including time, the Minkowski space.11,12 Its homogeneity, i.e., that no point is special, and its isotropy, i.e., that no direction is special, are very important since these symmetries determine the basic building blocks of our theories. These symmetries are connected directly to the position-momentum and energy-time uncertainty relations,71,11,12,73 i.e., the translations in space are connected to momentum operators and the translations in time to the energy operator. Thus, the basic building blocks are (self-adjoint realizations of) the momentum \( -i\hbar \nabla \) and position \( r \) operators and the energy \( i\hbar \partial_t \) and time \( t \) operators (see Appendices A.1 and A.2 for more details). And the basic wave functions describing matter and light, respectively, should obey eq 3, but with the substitution \( E \rightarrow i\hbar \partial_t \) and \( p \rightarrow -i\hbar \nabla \). Just using the resulting second-order equation to determine the basic wave functions leads, however, to several problems.71,11,13,74 A possible way out is to recast the second-order equation in terms of a first-order Hamiltonian equation, i.e., an evolution equation for the energy. Following Dirac’s seminal idea, we can use for spin-1/2 particles the four-component Dirac equation

\[ i\hbar \partial_t \psi(r) = [-i\hbar \alpha \cdot \nabla + \alpha_0 m c^2] \psi(r) \]  

(4)

where the vector of matrices \( \alpha \) and \( \alpha_0 \) are the 4 \( \times \) 4 Dirac matrices.11,13,74 Applying the Dirac equation twice, we recover the operator form of eq 3 as intended. Eq 4 is then used to describe the matter sector of QED. If we use a vector of spin-1 matrices \( S \) instead, we find the Riemann-Silberstein equation75–78

\[ i\hbar \partial_t f(r) = -i\hbar S \cdot \nabla f(r) \]  

(5)

for a three-component wave function \( f \) with zero mass and the necessary side condition

\[ \nabla \cdot f(r) = 0 \]  

(6)

This side condition ensures that wave function \( f \) has only two transverse degrees of freedom, as to be expected for free electromagnetic fields, which have two independent polarizations. Eqs 5 and 6 are then used to describe the electromagnetic sector of QED and recover the usual Maxwell equations in the classical limit, as discussed below.
2.2. Quantizing the Light Field

The main issue with these two relativistic equations is that, since they are first order, they necessarily have besides positive-also negative-energy eigenstates. That this is an issue becomes immediately clear from the Riemann-Silberstein wave function \( f \), which should be a quantum version of the electromagnetic energy expression in terms of the electric field \( E(rt) \) and magnetic field \( B(rt) \), i.e.,

\[
E_{ph} = \frac{\epsilon_0}{2} \int \left( E^2(rt) + c^2 B^2(rt) \right) \, dr
\]

with strictly positive eigenenergies. To resolve this issue, we follow a further seminal idea of Dirac. We reinterpret the single-particle equations as actually being equations for two particles. That is, the positive-energy states are the particles and the negative-energy states are the corresponding antiparticles.\(^{11,12,74}\) For the photon, we find that it is its own antiparticle, where positive-energy states are associated with positive helicity and negative-energy states with negative helicity.\(^{12,13,13}\) To translate this idea into a mathematical prescription, we perform a second quantization step. In more detail, we use the distributional eigenstates of the respective equations (plane waves with momentum \( k \) times the corresponding Dirac spinors for matter, or times circular polarization vectors for light), define creation and annihilation field operators for particles and antiparticles, and effectively exchange the meaning of creation and annihilation for the antiparticles such that the energy becomes manifestly positive.\(^{11,12,74}\)

In the case of the electromagnetic field quantization the respective field operators obey, due to being spin-1 particles, the (bosonic) equal-time commutation relations:

\[
[\hat{a}(\mathbf{k}', \lambda'), \hat{a}^\dagger(\mathbf{k}, \lambda)] = \delta_{\lambda, \lambda'} \delta_2(\mathbf{k} - \mathbf{k}')
\]

(8)

Here we interpret \( \lambda = 1 \) as having positive helicity and \( \lambda = 2 \) as having negative helicity.\(^{11-13}\) With this, we find the quantized form of eq 7 to be

\[
\hat{H}_{ph} = \sum_{\lambda=1}^{2} \int \hbar \omega_0 \hat{a}^\dagger(\mathbf{k}, \lambda) \hat{a}(\mathbf{k}, \lambda) \, dk
\]

(9)

where \( \omega_0 = c|k| \) (dispersion of the light cone) and we have discarded the trivial and unobservables, yet infinite vacuum contribution \( \sum_\lambda \frac{1}{\hbar} \hbar \omega_0 \, dk/2 \), i.e., we have assumed normal ordering.\(^{11-15}\)

In this very condensed derivation of the quantized electromagnetic Hamiltonian (we do not give further details of the electronic part of relativistic QED, because we will consider nonrelativistic charged particles only) we have made some important implicit choices that need to be highlighted. First, we used a quantization procedure based on the vector curl, i.e., \(-i \mathbf{S} \cdot \nabla = \nabla \times \), its distributional eigenfunctions are also distributional eigenfunctions for the static vector-potential formulation of the homogeneous Maxwell equation (see also eq 24):

\[
-\nabla A(r) = (\nabla \times \nabla \times - \nabla^2) A(r) = k^2 A(r)
\]

(10)

Here, the left-hand side is just due to a vector identity of the vector Laplacian and we note that the longitudinal part is zero by construction due to the side condition of eq 6, i.e., only the transverse part (first term) is nontrivial. The quantization in terms of the vector potential is an important choice, since in the context of the Riemann-Silberstein formulation one often uses a quantization procedure based on the electric and magnetic fields instead.\(^{20,79,80}\) We will comment on this and further connections to classical electrodynamics a little below. Furthermore, since we have only considered the transverse eigenfunctions of eq 10, we have implicitly chosen the Coulomb gauge, i.e., \( \mathbf{V} \cdot \mathbf{A} = 0 \). Consequently, the electromagnetic vector potential

\[
\mathbf{A}(r) = \frac{\hbar c^2}{\epsilon_0} \sum_{\lambda=1}^{2} \int \frac{1}{\sqrt{2\omega_0}} \left[ \hat{a}(\mathbf{k}, \lambda) \exp(i\mathbf{k}\cdot\mathbf{r}) \epsilon(\mathbf{k}, \lambda) + \hat{a}^\dagger(\mathbf{k}, \lambda) \exp(-i\mathbf{k}\cdot\mathbf{r}) \epsilon^*(\mathbf{k}, \lambda) \right] \, dk
\]

(11)

given here in units of Volts, to agree with relativistic notation,\(^{11,12,81}\) has only the two physical transverse components. If we had chosen a different gauge instead, we would have to take care of unwanted longitudinal and time-like degrees of freedom by employing quite intricate technical methods, such as Gupta-Bleuler or ghost-field methods.\(^{11,20,71}\) The main drawback of the Coulomb gauge is that it is not explicit Lorentz covariant, i.e., if we perform a Lorentz transformation to a new reference frame the Coulomb condition is violated in general.\(^{11}\) However, since we usually have a preferred reference frame for our considerations, i.e., the lab frame, this is a minor restriction in practice. The second point we want to mention is that we have so far chosen, in accordance to the distributional eigenfunctions of eq 5, circularly polarized vectors \( \epsilon(\mathbf{k}, \lambda) \).\(^{11,12,71}\) But for the quantization of the electromagnetic field we can equivalently choose any other polarization vectors that obey

\[
\epsilon(\mathbf{k}, \lambda) \cdot \mathbf{k} = \epsilon(\mathbf{k}, 1) \cdot \mathbf{k} = 2 = 0
\]

(12)

and are normalized, i.e., \( \epsilon^*(\mathbf{k}, \lambda) \epsilon(\mathbf{k}, \lambda) = 1 \). Indeed, in the following, we will assume the standard choice of linearly polarized vectors if nothing else is stated because the linearly and the circularly polarized representation are connected by a canonical transformation that leaves everything invariant. For the following theoretical considerations, it is sufficient to overload the meaning of \( \hat{a}(\mathbf{k}, \lambda) \) and \( \epsilon(\mathbf{k}, \lambda) \) to correspond to the respective linearly polarized objects as well. The only formal difference is that we can take \( \epsilon(\mathbf{k}, \lambda) \) outside the brackets in eqs 11 and 25 since in this case it is a real-valued three-dimensional vector. We note that in certain cases the linear polarization will be important, e.g., for the derivation of the length gauge Hamiltonian of eq 39. We will come across an electromagnetic field given in terms of circularly polarized (also called chiral) modes only at the very end, i.e., in the outlook presented in Section 6.

Going back to the Riemann-Silberstein eq 5, we recognize that there is a well-known classical equation associated with it, in contrast to the Dirac equation. Indeed, if we reinterpret the three-component wave function and give it the units of an energy wave function, i.e., \( \sqrt{CV/m^2} \) where \( C \) is Coulomb, \( V \) Volts and \( m \) meters, we can associate

\[
F(rt) = \frac{\epsilon_0}{2} \left( E(rt) + i\mathbf{B}(rt) \right)
\]

(13)

Using this (classical) Riemann-Silberstein vector, eqs 5 and 6 become the four Maxwell equations without sources:
\[
\frac{1}{c} \partial_t E(rt) = \nabla \times B(rt) \quad (14)
\]
\[
\partial_t B(rt) = -\nabla \times E(rt) \quad (15)
\]
\[
\nabla \cdot E(rt) = 0 \quad (16)
\]
\[
\nabla \cdot B(rt) = 0 \quad (17)
\]

In this reinterpretation of eq 5, the operator \(-i\hbar \mathbf{A} \nabla\) no longer refers to an energy but rather to power since we can cancel the \(\hbar\) on both sides of eq 5. Further, the energy of eq 7 is given by the norm of the Riemann-Silberstein vector:

\[
E_{\text{ph}} = \int F^\dagger(\mathbf{r}) \cdot F(\mathbf{r}) \, d\mathbf{r} \quad (18)
\]

To connect the classical Maxwell equations back to the above second quantization procedure, we note that the vector potential representation of eqs 14–17 in an arbitrary gauge is

\[
-\nabla^2 \phi(rt) - \frac{1}{c} \partial_t (\nabla \cdot \mathbf{A}(rt)) = 0 \quad (19)
\]

\[
\left[ \frac{1}{c^2} \partial_t^2 - \nabla^2 \right] \mathbf{A}(rt) + \nabla \left( \nabla \cdot \mathbf{A}(rt) + \frac{1}{c} \partial_t \phi(rt) \right) = 0 \quad (20)
\]

where the four-potential vector is given by \((\phi(rt), \mathbf{A}(rt))\) and we have the association

\[
E(rt) = -\nabla \phi(rt) - \frac{1}{c} \partial_t \mathbf{A}(rt) \quad (21)
\]

\[
\mathbf{B}(rt) = \frac{1}{c} \nabla \times \mathbf{A}(rt) \quad (22)
\]

Choosing now the Coulomb gauge, i.e., \(\nabla \cdot \mathbf{A}_\perp(rt) = 0\), the above equations become

\[
\nabla^2 \phi(rt) = 0 \quad (23)
\]

\[
\left[ \frac{1}{c^2} \partial_t^2 - \nabla^2 \right] \mathbf{A}_\perp(rt) = 0 \quad (24)
\]

The only zero solution of eq 23 is \(\phi(rt) = 0\), and all zero solutions of eq 24, i.e., freely propagating Maxwell fields, can be constructed with the help of the distributional eigenstates of eq 10. The Coulomb gauge is a maximal gauge, since it removes all gauge ambiguities (compare eqs 19 and 20) that would still be allowed in other gauges. We further note that we recover the classical equations from the above vector-potential-based second-quantized formulation by using the Heisenberg equations of motions, where \(\dot{\mathbf{B}}(r) = \frac{1}{c} \nabla \times \mathbf{A}_\perp(r)\) and

\[
\dot{\mathbf{A}}_\perp(r) = \left( \frac{\hbar}{\varepsilon_0} \right)^2 \sum_{x=1}^2 \int \frac{i \omega_k}{\sqrt{2 \omega_k}} \left( \delta(k, \lambda) \exp(i \mathbf{k} \cdot \mathbf{r}) \epsilon(k, \lambda) - \delta(k, \lambda) \exp(-i \mathbf{k} \cdot \mathbf{r}) \epsilon^*(k, \lambda) \right) \, dk \quad (25)
\]

Finally, we mention that one can also do a second quantization based on the interpretation of eq 13 without resorting to the vector potential formulation. This has the advantage that the resulting basic objects of the theory are gauge-independent. On the other hand, as we will see next, the coupling between light and matter is based on the gauge principle, and hence, at that point usually the vector potential formulation appears again.

### 2.3. Coupling Light and Matter

Let us next couple the two sectors of the theory. Not surprisingly, there are again various ways to derive how photons and quantized charged particles couple.\(^{11,12,20,22,23,71}\) We will use a further symmetry argument here to couple light and matter. The Dirac and Riemann-Silberstein equations are intimately connected to symmetries. One specifically important symmetry is connected to the local conservation of charge (or probability if we do not include the elementary charge \(|e|\) in the arguments below). Indeed, from eq 4 we find that the Dirac charge density \(\rho(rt) = -|e| \psi^\dagger(\mathbf{r}) \psi(\mathbf{r})\) and the Dirac charge current \(j(rt) = -|e| \psi^\dagger(\mathbf{r}) \mathbf{v} \psi(\mathbf{r})\), where and \(-|e|\) is the charge of the electron, obey the continuity equation:

\[
\partial_t \rho(rt) + \nabla \cdot j(rt) = 0 \quad (26)
\]

This equation guarantees that locally charge cannot be destroyed or created; it can only flow from one point to another. Since in the above equation the phase of the wave function becomes irrelevant, we realize that this conservation law holds even if we change the phase of the wave function \(\psi(\mathbf{r}) \rightarrow \exp(i \varphi(\mathbf{r})) \psi(\mathbf{r})\). In order to enforce that this phase change does not affect any physical observable, we have to change \(\varphi(\mathbf{r}) \rightarrow \varphi(\mathbf{r}) + \theta\) and \(-i \mathbf{v} \rightarrow -i \mathbf{v} - \nabla \varphi(\mathbf{r})\) in eq 4. One therefore interprets the resulting linearly coupled fields \((\partial_t \varphi(rt), \nabla \varphi(rt))\) as having no physical effect on the charged particle. Indeed, if we determine the Maxwell energy that such fields would correspond to, we find that the four vector potential \((\frac{\hbar}{\varepsilon_0} \partial_t \varphi(rt), -i \mathbf{v} \varphi(rt))\) leads to zero physical fields (compare to eqs 21 and 22) and thus to zero energy (compare to eq 7). The phase of the wave function therefore corresponds to the gauge freedom of the electromagnetic field. This suggests that we should couple a general (nonzero) electromagnetic field in the same linear (minimal) manner, i.e.,

\[
i \hbar \partial_t \mathbf{A}_\perp \rightarrow i \hbar \partial_t \mathbf{A}_\perp + |e| \phi(rt)
\]

\[
-\hbar \nabla \mathbf{v} \rightarrow -i \hbar \nabla + \frac{|e|}{c} \mathbf{A}(rt)
\]

This adapted derivative is then called a gauge-covariant derivative.\(^{11,12,71}\) All of this can be formalized much more elegantly in a Lagrangian representation of the problem, where the gauge-covariant derivative makes the local charge conservation explicit.\(^{11,12,71}\)

Let us next see what that prescription entails for light. For this we look at the (still classical) light–matter interaction energy expression that we recover from the above prescription which is

\[
E_{\text{int}} = -\frac{1}{c} \int j(rt) \cdot \mathbf{A}(rt) \, d\mathbf{r} + \int \rho(rt) \phi(rt) \, d\mathbf{r}
\]

Varying this energy expression with respect to the four vector potential, we can derive the corresponding contributions to the Maxwell equation.\(^{11}\) If we choose the Coulomb gauge we thus find compactly

\[
-\nabla^2 \phi(rt) \equiv \rho(rt) \quad (29)
\]

\[
\left[ \frac{1}{c^2} \partial_t^2 - \nabla^2 \right] \mathbf{A}_\perp(rt) = \mu_0 c I_\perp(rt)
\]
where due to the inner product in eq 29 only the transverse part of the charge current contributes. We have thus derived the Maxwell equations including sources that obey the continuity of eq 26. For completeness and later reference we further give the inhomogeneous Maxwell equations as

\begin{align}
\nabla \times \mathbf{B}(\mathbf{r}) - \frac{1}{c^2} \partial \mathbf{E}(\mathbf{r}) &= \mu_0 \mathbf{J}(\mathbf{r}) \\
\partial \mathbf{B}(\mathbf{r}) + \nabla \times \mathbf{E}(\mathbf{r}) &= 0 \\
\nabla \cdot \mathbf{E}(\mathbf{r}) &= \frac{\rho(\mathbf{r})}{\varepsilon_0} \\
\nabla \cdot \mathbf{B}(\mathbf{r}) &= 0
\end{align}

(32) (33) (34) (35)

If we next assume that the only sources for the electromagnetic fields are the (quantized) charged particles, the longitudinal part of the fields, i.e., those corresponding to \( \phi(\mathbf{r}) \) in eq 30, can be expressed purely in terms of the charge density itself, i.e., the Hartree potential

\[ \phi(\mathbf{r}) = \int \frac{\rho(\mathbf{r}')}{4\pi \varepsilon_0 |\mathbf{r} - \mathbf{r}'|} \, d\mathbf{r}' \]

(36)

If we combine this longitudinal interaction energy with the longitudinal contribution in \( E_{\phi} \) we obtain the well-known Coulomb interaction between the (quantized) charged particles.\(^{11}\) So upon second quantization of the electromagnetic field, the longitudinal contributions in Coulomb gauge are only affected by the quantization of the particles and we are left by just replacing \( A_\perp(\mathbf{r}) \rightarrow \hat{A}_\perp(\mathbf{r}) \) (in the Schrödinger picture\(^{11}\)).

Before we give the basic Hamiltonian of nonrelativistic QED in the next section, we want to highlight the intimate relation between the geometry of (real) space, the light and the matter sector, the gauge choice, and the interaction. Changing any of these ingredients needs to be accompanied by a careful re-evaluation of the basic theory. First, we highlight that if we restrict to only a part of \( \mathbb{R}^3 \), we need to carefully re-evaluate the basic symmetries in the theory. This is relevant for practical implementations of nonrelativistic QED and derivation of corresponding approximate models. For instance, a box with periodic boundary conditions, where all three edges have the same length, keeps all the basic symmetries intact (see also Appendix A.2). One finds that the resulting theory, where the plane wave solutions of the various differential operators become normalizable eigenfunctions, converges to the free-space formulation that we have discussed so far. One therefore often uses these two settings interchangeably. Already just choosing other boundary conditions, for instance, zero boundary conditions, might imply subtle differences (see also Section 3.3). We further note that both basic equations, i.e., eqs 4 and 5, are based on the same differential operators and hence share the same (distributional) eigenfunctions. This consistency is highlighted again in the gauge principle of eqs 27 and 28, where the differential operator and the fields obey the same boundary conditions. Thus, changing the modes of the light field independently from the matter can violate, for instance, the basic gauge principle and the Maxwell equations. We will comment on this also later in Section 3.3 (see also Appendix A.4). Finally, the gauge choice influences what we call matter and what we call light. This can be nicely seen from the fact that in Coulomb gauge the longitudinal and time-like photons are absent and subsumed in the Coulomb interaction between the charged particles. This will be further discussed in Section 3.2.

### 3. THE PAULI-FIERZ QUANTUM-FIELD THEORY

“The claimed range of validity of the Pauli-Fierz Hamiltonian is flabbergasting. To be sure, on the high-energy side, nuclear physics and high-energy physics are omitted. On the long-distance side, we could phenomenologically include gravity on the Newtonian level, but anything beyond that is ignored. As the bold claim goes, any physical phenomenon in between, including life on Earth, is accurately described through the Pauli-Fierz Hamiltonian […].”

H. Spohn in ref 22.

We discussed earlier how the (quantized) electromagnetic field can be deduced and how it can be coupled to a quantized matter description. Yet, if we treat matter on the same relativistic level as light, we encounter various conceptual and mathematical issues. Performing a second quantization of also the Dirac equation and coupling it to a second-quantized Maxwell equation via the above gauge-coupling prescription, leads to several divergences.\(^{12,69,71}\) Full QED treats these divergences by regularizing and then renormalizing scattering theory.\(^{15,13,71}\) The simplest realization of a regularization introduces several energy cutoffs in the theory (largest and smallest energy scales for the different particles and their interactions), and it is then shown that the results of perturbative calculations do not depend on how the cutoffs are removed upon renormalization of the theory. In the following, however, we go beyond perturbation theory and consider, for instance, spatially and temporally resolved how a molecule changes during a chemical reaction. In other words, we solve a Schrödinger-type equation that gives us access to such processes.

#### 3.1. Nonrelativistic QED

Indeed, within the last decades tremendous progress has been made to reformulate QED as a nonperturbative ab initio quantum theory in several limits.\(^{22,84–86}\) The most important situation for our purpose is the nonrelativistic limit for the matter sector (while keeping the photon sector fully relativistic), which allows for a mathematical formulation that is similar to standard electronic quantum mechanics (see also Appendix A for more details on the mathematical setting of ab initio quantum physics).\(^{72,73}\) So instead of the Dirac equation, we are mainly interested in the electronic part of matter and assume that the electrons have small momenta (with respect to relativistic scales). In other words, we discard the positrons and replace the Dirac momentum by the nonrelativistic momentum and hence assume that the electrons are well described by the Schrödinger equation. Because this also implies matter particle conservation (no electron-positron pair creation is possible anymore) we do not need to second-quantize the matter sector. This avoids many of the pitfalls of full QED that arise from working with mathematically problematic field operators (see Appendix A.3).\(^{83,87}\) The resulting Hamiltonian, where light and matter couple via the exact minimal coupling prescription from above, is the generalized Pauli-Fierz Hamiltonian\(^{22,81}\)
\[
\hat{H}_{\text{pq}} = \sum_{\sigma=1}^{N} \left( \frac{1}{2m} \left( -\hbar \nabla_{\gamma} + \frac{\hbar c}{2} \mathbf{A}_{\gamma}(\tau) \right)^2 + \frac{1}{2m} \mathbf{F}(\tau) \right) + \sum_{\nu=1}^{N} \left( \frac{1}{2m} \left( -\hbar \nabla_{\eta} - \frac{Z_{\nu} e}{c} \right)^2 + \frac{1}{2m} \mathbf{B}(\eta) \right)
\]

The explicit interaction with the photons, on the other hand, makes it necessary in general to work with \textit{bare electronic and nuclear/ionic masses} \( m \) and \( M \), respectively. That is, the masses in eq 37 are not the observable masses that one uses in quantum mechanics. The physical masses of the particles in quantum mechanics are recovered from nonrelativistic QED by tracing out the photon part which leads, e.g., for the electronic mass to \( m_e = m + m_{ph} \), as also highlighted in the introduction. Here the photon contribution, \( m_{ph} \), is due to the electromagnetic energy that is created by the charged particle itself. When considering the dispersion of a free particle in nonrelativistic QED, we realize that the bare mass is necessarily smaller than in quantum mechanics, i.e., \( m_{ph} > 0 \). This is because the free charged particle generates extra energy due to coupling to the photons when having nonzero momentum and is thus effectively slowed down, i.e., the electron is dressed by the photon field (see also Figure 1 for an artistic view on dressed particles in QED). We will give an explicit expression for the photonic mass (of single particles in the dipole approximation) and comment on further implications of this mass renormalization in Section 3.3. Irrespective of the specific choice of (the positive and finite) bare mass, however, the Pauli-Fierz Hamiltonian has some very nice properties. It is self-adjoint, which guarantees that we can uniquely solve the corresponding static and time-dependent \textit{Schrödinger-type equations}

\[
i\hbar \partial_t \Psi(t) = \hat{H}_{\text{pq}} \Psi(t)
\]
electromagnetic field, also has a ground state in nonrelativistic QED.99–104 This is exactly the property we need in order to discuss the equilibrium properties of a coupled light–matter system. An important difference, however, is that all excited states turn into resonances in nonrelativistic QED, i.e., excited states are no longer eigenstates but have a finite lifetime.99,101,105 This feature, which is also termed spontaneous emission, is missing in standard ab initio electronic structure theory, where excited states have the unphysical property of being infinitely long-lived. Indeed, if one just looks at the spectrum of the Pauli-Fierz Hamiltonian, one will usually just find one eigenstate, i.e., the ground state and then a continuum above the ground state. Thus, the spectrum alone does not provide much insight into the properties of the coupled light–matter system.22,101,105 On the other hand, due to the inclusion of the continuum of photon modes and all the nuclear/ionic degrees of freedom, we have included all dissipation and decoherence channels that are physically present for the subsystems of the total light–matter system, and no external baths or non-Hermitian terms need to be added to mimic those processes. In other words, despite the theory being self-adjoint, i.e., closed, the infinite amount of degrees of freedom includes also the physical bath degrees of freedom by radiating light from the molecules to the far field and hence being lost to the molecular subsystem. So we can conclude that we have found a fully nonperturbative and mathematically consistent ab initio quantum theory of light and matter (see Appendix A for further details on ab initio quantum physics), which answers the first fundamental question from the introduction.

One final important comment addresses the possibility of working with a different gauge, which relates to the second fundamental question of the introduction. Performing a gauge transformation on the Pauli-Fierz Hamiltonian is far from trivial, since the choice of gauge alters the structure of the underlying Hilbert spaces. This becomes even more problematic because the introduced ultraviolet cutoff does not commute in general with the gauge fixing; i.e., exact gauge equivalence is usually lost once a cutoff has been introduced. We will find one notable exception in the case of the dipole approximation of the Pauli-Fierz Hamiltonian below in Section 3.3. Furthermore, to the best of our knowledge, only the Pauli-Fierz Hamiltonian in the Coulomb gauge has been shown to have all the above desirable mathematical properties within an ab initio QED framework. Using other gauges to quantize the theory needs careful considerations, as novel problematic terms and divergences arise.107,108 In addition, one has to note that for other gauges, e.g., the Lorentz gauge, the Coulomb interaction is mediated directly via the (time-like and longitudinal) photons. Consequently even a “quantum-mechanical calculation” that takes into account only the longitudinal Coulomb interaction needs infinitely many quantized modes that need to fulfill certain consistency conditions, such as enforced by the Gupta-Bleuler method.11,20 Therefore, the Coulomb gauge seems to be the most relevant and practical gauge on a nonperturbative Hamiltonian level, and it connects seamlessly with standard quantum mechanics, which is implicitly always assuming the Coulomb gauge.15,12,22 Consequently it is important to choose the Coulomb gauge if combining theoretical methods for the quantized light field with standard theoretical approaches to quantum matter. This avoids implicit gauge inconsistencies such as double counting the longitudinal interactions between charged particles.

### 3.3. Approximations

Nonrelativistic QED allows to work with (polaritonic) wave functions $\Psi$ of the fully coupled light–matter system,5,22,81 which makes it very similar to standard quantum mechanics. However, the corresponding wave function depends not only on $N_e$ electronic and $N_n$ nuclear/ionic coordinates anymore but also on a full continuum of photon modes as well. Thus, even for a single particle in free space, a wave function solution of eq 38 is practically unfeasible. Note furthermore that we might need to describe the photonic structure as part of the quantum system in minimal coupling; e.g., the mirrors of an optical cavity are described with the Pauli-Fierz Hamiltonian as well. As will be discussed below, just approximating the cavity structure with a different level of theory runs the risk of introducing severe inconsistencies. Thus, on this highest level of theory, for any calculation, we first need to fix a cutoff for the free-space continuum of modes, adjust the bare mass of the particles to agree with their experimentally observed free-space dispersions, and describe the photonic structure as well as the matter system that is coupled to this structure with the same Pauli-Fierz Hamiltonian. To date it remains unknown whether the results of the Pauli-Fierz theory depend on the specific choice of the cutoff and the corresponding bare mass or whether, similar to its scalar counterpart,109 taking the cutoff to infinity corresponds to a mere infinite energy shift, i.e., that the theory is nonperturbatively renormalizable. Irrespective of these details, this level of theory on the wave function level is impractical. So, how can we make the Pauli-Fierz theory applicable? A first slight simplification is found by realizing that we can discretize the photon continuum, and consider then a continuum limit.110,111 A good enough discretization (for our setup, a very large quantization box with periodic boundary conditions) is virtually indistinguishable from a real continuum. However, this does not really resolve the problem of the still humongous amount of coordinates in the wave function. One way is to reformulate the Pauli-Fierz theory as a density functional theory (see discussion in Section 4.1), which allows minimal-coupling simulations in practice (for examples see Section 5.1). Yet, before we discuss potential first-principles approaches in Section 4, we want to focus on the linear formulation of the problem in terms of wave functions. Therefore, one has to cut back drastically on the amount of coordinates if one is interested in a nonperturbative solution of the Pauli-Fierz Hamiltonian. For perturbative approaches many alternative strategies exist such as to subsume the continuum of modes in a mass-renormalization from the start, i.e., one works with the physical masses of the particles, and everything else is taken into account by, e.g., Wigner-Weisskopf theory.89 We will focus here on the nonperturbative ab initio QED approaches that are beneficial for the strong-coupling regime of polaritonic chemistry, which we are interested in in the following.

One feasible approximation is to use conditional-wave function approaches112–114,117 to disentangle the different degrees of freedom. In other words, we could apply a Born–Oppenheimer-type of approach, i.e., evolve the nuclei/ions quantum mechanically on a potential-energy surface that is provided by the electrons and/or photons.112,113 One then needs to choose whether to group the photons with the electrons or nuclei/ions (see also Section 4.4) and ensure that there is no double-counting due to the coupling of the photons to both matter degrees of freedom. Further, one needs to take
into account that the photons also mediate new couplings between the electrons, the nuclei/ions and between the electrons and the nuclei/ions. In the dipole approximation (see discussion in Sections 3.3 and 4.4) such extended Born–Oppenheimer approaches have already been investigated and employed in practice. An even further simplification would then be to treat the nuclei/ions classically, which leads to a coupled Ehrenfest-Pauli-Fierz problem.\(^\text{3,115,116}\) This option is discussed in a little more detail in Section 4.4. Another possibility to reduce our problem size is to disentangle different parts of the problem by position, which employs the real-space nature of the Pauli-Fierz Hamiltonian. For instance, we can imaging a common cavity setup, where metallic surfaces constitute the optical cavity and we have the matter system of interest in the middle of this cavity (such as in Figure 2).

![Figure 2. Common Fabry-Pérot cavity setup. If we assume that the molecules of interest are far removed from the cavity mirrors and localized around the center, one can approximate the main cavity frequencies due to the mirror distance \(L\) by \(\omega_0 = c/2L\), where we have subsumed the effect of the continuum of free-space modes (perpendicular to \(L\)) in the effective/observed mass of the particles. The coupling strength \(g_{\text{cd}}\) for the two independent polarization directions \(\lambda\) then increases with \(\sqrt{1/L}\) if we keep the low-energy (continuous free-space) modes fixed and take their effect into account by the physical mass of the particles (see Section 3.3.1).](image)

If the surfaces are far enough from the molecular system of interest, the mirrors of the cavity can be described with an effective theory that accounts for changes in the local mode structure of the electromagnetic field instead of describing the (macroscopic) cavity as part of the (cavity+molecular) system. Such a procedure is commonly done, for instance, in macroscopic QED, where the modes of some photonic structures are quantized based on linear-response theory.\(^\text{21,117}\) Such an approximation procedure can lead, however, to problems. Keeping in mind our discussion about the necessary consistency between light and matter in Section 2, where we saw that the mode structures of both sectors are the same, we can break various exact relations, such as energy and momentum conservation, if we change the (Fourier) mode structure of light and matter independently. An instructive example is found if we take periodic boundary conditions for matter but perfect-conductor boundary conditions for light\(^\text{21}\) to simulate a cavity structure. In this case the gauge principle of eq 28 tells us that just adding \(\exp\left(i\frac{2\pi}{L}x\right)\) to the wave function on \(x \in [0, L]\) corresponds to a pure gauge, and the resulting pure gauge field is proportional to \(\frac{2\pi}{L}\), i.e., a constant field. The Maxwell energy with the perfect-conductor boundary conditions of a constant field is, however, infinite. This can be seen either by a basis expansion or by realizing that a self-adjoint differential operator always knows about the boundary conditions and hence interprets that the constant field drops instantaneously to zero at the boundary, which is not differentiable\(^\text{118,119}\) (for more details see Appendix A.4). Such issues are avoided once we make the dipole-coupling approximation, where the mode consistency between light and matter becomes irrelevant and we can indeed replace the cavity by a local modification of the electromagnetic modes. We discuss this in more detail below in Section 3.3.1.

A different type of simplification follows from a clever basis choice, such as the eigenfunctions of the uncoupled problem, and then to assume that only a few such matter and light states contribute significantly to the solution of the Pauli-Fierz equation. This is a very common way in quantum optics\(^\text{19,24}\) but it clearly needs already a very detailed understanding or intuition of the subsystems and the physics involved in the light–matter coupling. Moreover, one also needs knowledge about the representation of these states in the original basis of the Pauli-Fierz Hamiltonian to model the proper coupling among the new (many-body) states and the potentially complex photonic states. This knowledge is commonly not available. The many-body methods needed for large systems do not provide the states directly. We will also discuss this issue below in the context of first-principle methods of the Pauli-Fierz Hamiltonian (see Section 4). To circumvent the issues of having the many-body states available, again the dipole approximation comes in handy since dipole transition moments are readily available for many different systems from various theoretical ab initio methodologies.

### 3.3.1. Cavity as Modification of Local Mode Structure: Dipole Approximation

For a straightforward simplification of the Pauli-Fierz problem, one usually goes directly to the dipole approximation thanks to its many desirable properties. The basic assumption implies that all relevant modes of the electromagnetic field have a wavelength \(2\pi/|k|\) that is much larger than the extent of the localized matter system. This clearly requires that we need to adjust the cutoff to low enough frequencies. Indeed, for most calculations one usually reduces the number of modes to only a few effective ones.\(^\text{120}\) We will discuss the resulting implications below. Following the above assumption, we replace \(\hat{A}_\lambda(r) \rightarrow \hat{A}_\lambda(0) \equiv \hat{A}_\lambda\) in eq 37, where we have also assumed implicitly that the matter system is localized (center of charge) at the origin of the coordinate system. An alternative way to arrive at the same approximation is to assume \(\exp(ik \cdot r) \approx 1\) in eq 11. Besides becoming problematic when the wavelength of the considered modes becomes comparable with the size of the matter system or when retardation effects become important, we also discard in the dipole approximation any direct influence due to the magnetic part of the quantized photon field on the spin degrees of freedom. We further note that we do not use a multicenter dipole approximation, as often assumed in perturbative or model approaches, where different particles see different fields,\(^\text{23}\) since this would a priori violate the fundamental indistinguishability criterion of quantum particles.
Only upon interacting with an environment can we attain distinguishability and classicality, which is discussed in Section 5.3. The resulting (single-center) Hamiltonian is then often also called to be in velocity gauge, which is just the dipole-approximated Coulomb-gauged Pauli-Fierz Hamiltonian. Its form highlights a few important properties that make the dipole approximation so versatile. While eq 37 is translationally and rotationally invariant only in the full configuration space of light and matter, in the dipole approximation the Hamiltonian is translationally and rotationally invariant also with respect to the matter subsystem. Thus, we find the nice and practical feature that the Pauli-Fierz Hamiltonian can be efficiently expanded in the usual matter-only Bloch states in dipole approximation, in contrast to the full minimal coupling Hamiltonian. Hence, one usually works in the dipole approximation for solid-state systems. How to properly include beyond dipole contributions for extended systems remains an active topic of research.

Specifically in the context of symmetries, it is important to highlight that there is a second, unitarily equivalent form of the dipole-approximated Pauli-Fierz Hamiltonian. In more detail, upon performing a unitary transformation \( \exp(\frac{i}{\hbar}\mathbf{A} \cdot \mathbf{R}) \), where \( \mathbf{R} = -\sum_{\alpha} N^\alpha \mathbf{\hat{r}}_\alpha + \sum_{\alpha} N^\alpha \mathbf{\hat{Z}}_\alpha \mathbf{\hat{R}}_\alpha \) is the total dipole operator, and a swapping of conjugate photon variables, one finds the length gauge Pauli-Fierz Hamiltonian

\[
\hat{H}_{\text{PF}} = -\frac{N}{2m} \sum_{\alpha} \mathbf{\hat{r}}_\alpha^2 \nabla^2_{\mathbf{\hat{r}}_\alpha} + \frac{1}{2} \sum_{\alpha,\beta} \frac{e^2}{4\pi\epsilon_0 |\mathbf{\hat{r}}_\alpha - \mathbf{\hat{r}}_\beta|} \\
- \sum_{i=1}^{N^\alpha} \sum_{m=1}^{N_{\alpha,m}} \frac{Z_{\alpha,m}^2}{4\pi\epsilon_0 |\mathbf{\hat{r}}_\alpha - \mathbf{\hat{R}}_{m}|} - \sum_{i=1}^{N^\alpha} \frac{\mathbf{\hat{r}}_i^2}{2M_i} \nabla^2_{\mathbf{\hat{r}}_i} \\
+ \frac{1}{2} \sum_{\alpha,\beta} \frac{Z_{\alpha,m}^2}{4\pi\epsilon_0 |\mathbf{\hat{r}}_\alpha - \mathbf{\hat{R}}_{m}|} \\
+ \sum_{\alpha=1}^{M^\alpha} \left[ -\frac{\mathbf{\hat{r}}_\alpha^2}{2\partial^2_{\mathbf{\hat{r}}_\alpha^2}} + \frac{\omega_a^2}{2} \left( q_a - \frac{e_a^2}{\omega_a^2} \mathbf{\hat{R}} \right)^2 \right]^{\frac{3}{2}}
\]

Here we have already assumed a discretized continuum of \( M^\alpha \) modes (given in terms of displacement coordinates \( q_a \) in units of \( \sqrt{\hbar/\alpha} \)) labeled by \( \alpha \), where each \( \alpha \) is associated with a specific frequency \( \omega_{\alpha} \), coupling strength \( \gamma_{\alpha} \), and polarization \( e_{\alpha} \). In the free space case with a quantization volume \( L^3 \), these quantities would be associated with \( k_\alpha = 2\pi n/L, \alpha \equiv (k_\alpha, \lambda) \), \( \omega_{\alpha} = c|k_\alpha| \), and \( \gamma_{\alpha} = 1/c_0 L^3 \), where \( n \in \mathbb{Z}^3_+ \). However, now we can adapt the frequencies, coupling strengths, and polarizations to match a given cavity structure without breaking fundamental symmetries since the actual spatial mode structure and the momentum matching (no momentum is transferred in the dipole approximation) are no longer important. For a simple example see Figure 2.

Upon first glance, the form of eq 39 seems to break the above-discussed symmetries and has an unusual self-interaction term proportional to \((e_{\alpha} \cdot \mathbf{R})^2\). This seeming conundrum can be resolved by carefully analyzing the unitary transformation and realizing that one has changed explicitly the conjugate variables of the photonic theory and mixed light and matter. Indeed, \( q_a \) does not correspond to a pure photonic quantity anymore but is connected to the auxiliary displacement field of the macroscopic Maxwell equations. The macroscopic Maxwell equations are equivalent to the microscopic Maxwell equations discussed in Section 2, yet they use the auxiliary displacement and magnetization fields that stem from a division of the charge currents and densities into bound and free ones. For completeness and for later reference let us briefly consider how these auxiliary quantities arise. We thus first define

\[
\rho(r_t) = \rho_{\text{bound}(r_t)} + \rho_{\text{free}(r_t)}
\]

\[
J(r_t) = J_{\text{bound}(r_t)} + J_{\text{free}(r_t)}
\]

and then introduce the polarization \( P(r_t) \) and magnetization \( M(r_t) \) due to the bound matter by

\[
J_{\text{bound}}(r_t) = \nabla \times M(r_t) + \frac{\partial P(r_t)}{\partial t}
\]

\[
\rho_{\text{bound}}(r_t) = -\nabla \cdot P(r_t)
\]

We note that these equations are equivalent to those in eqs 32 and 34. If we then make a corresponding division in the electromagnetic fields

\[
c_0 E(r_t) = D(r_t) - P(r_t)
\]

\[
B(r_t) = H(r_t) + M(r_t)
\]

and apply these definitions to eqs 32 and 34, we find

\[
\nabla \times H(r_t) - \frac{\partial D(r_t)}{\partial t} = J_{\text{free}(r_t)}
\]

\[
\nabla \cdot D(r_t) = \rho_{\text{free}(r_t)}
\]

Consequently, displacement \( D(r_t) \) and magnetization fields \( H(r_t) \) describe only the free part of the charges. We note that homogeneous eqs 33 and (35) are usually obeyed by the bound and free auxiliary fields individually. This formal reshuffling is useful in connecting Maxwell theory to a theory that describes a bound system and its reaction to electromagnetic fields. Thus, this formulation is often used in conjunction with approximate (matter-only) linear response theory in terms of constitutive relations. In our case, where light and matter are treated self-consistently and we have captured the reaction due to (bound) longitudinal fields exactly by using the Coulomb gauge, we are only left with transverse displacement and polarization fields. In the dipole coupling limit, where the magnetization is disregarded, we therefore find that \( \sum_{\alpha} c_0 \epsilon_a g_{a2}^2 (e_{a,R}) e_a = \hat{P}_L \) and \( \sum_{a} c_0 \epsilon_a g_{a2}^2 (e_{a,R}) e_a = \hat{D}_L \), such that

\[
c_0 \epsilon_a \hat{E}_a = \hat{D}_L - \hat{P}_L
\]

is the transverse electric field operator. Thus, the last line in eq 39 corresponds to the mode-resolved \( \hat{E}_a^\perp + \epsilon^2 \hat{B}_a^\perp \) and quadratic self-interaction terms naturally arise when coupling to light in terms of displacement and magnetization fields. Notice that also the matter coordinates have now a different meaning, since we have mixed light and matter (as we originally defined with respect to the Coulomb gauge). For instance, the translational symmetry is now found along a combined coordinate, i.e., one shifts not only \( r_t \) and \( R_t \) but at the same time also all \( q_a \). In addition, other observables, e.g., the number of photons, have now a different representation too. This issue has spawned a lot of misunderstandings, mainly in connection with what is called a superradiant phase.
transition. In more detail, the transverse electric field is by construction zero for any eigenstate, which follows from eq 21 in the Coulomb gauge. Yet the displacement field expectation value can be nonzero for an eigenstate. This merely means that one has a nonzero polarization, i.e., a nonzero total dipole of the system. However, the nonzero displacement field has been often misinterpreted as being the electric field, which led to the wrong conclusion that one can find radiating ground states, i.e., a photonic instability. Due to the symmetries of the Pauli-Fierz Hamiltonian we know that any ground state of atoms, molecules or solids has, by construction, in total zero transverse electric field expectation value. Nevertheless, one could still have a macroscopic amount of virtual photons in the ground state. A macroscopic amount of virtual photons in the ground state, e.g., in form of a constant macroscopic magnetic field, could alternatively be interpreted as a superradiant phase.

Let us note for completeness that the length gauge form of the Pauli-Fierz Hamiltonian of eq 39 can also be derived from the Power-Zienau-Woolley gauge in dipole approximation, assuming that this gauge had the same longitudinal Coulomb interaction. Yet beyond the dipole situation, both gauges are, as discussed above, formally different theories. A further reason for this discrepancy can be found in the fact that no multipole expansion exists for unbounded operators. That is, the common argument that a Coulomb-gauged field can be multipole expanded and in this way connected to the Power-Zienau-Woolley gauge only holds perturbatively and not on the level of operators in ab initio QED (see also Appendix A.3 for further details). In the context of working with operators instead of with perturbation theory we note that we have implicitly assumed that we are on $\mathbb{R}^3$ and instead of boundary conditions on the matter wave functions we have imposed normalizability to have self-adjoint operators. This is the standard setting of ab initio quantum physics (see also Appendix A.2). If we would restrict the matter domain, e.g., choose genuine periodic boundary conditions in the velocity gauge, the length gauge transformation changes these boundary conditions as well in a nontrivial manner again highlighting subtle differences when working with different gauges.

After these important technical details, let us return to the main advantage of the dipole approximation. That is, we can treat the photonic environment implicitly by changing the mode structure of the electromagnetic field at the position of the matter subsystem. In our case, we chose the origin as the center of charge. Therefore, one can take now the mode structure of a photonic environment, e.g., from a Maxwell calculation or from experiment, and adapt the $\omega_{\alpha}, \epsilon_{\alpha}$ and $g_{\alpha}$ in eq 39 accordingly. We note that one needs to use the corresponding displacement modes instead of the electric modes in the length gauge, i.e., in eq 39. A further important detail is that, in principle, when changing the mode structure, also the induced longitudinal interaction would change. For a better understanding of this aspect, let us first highlight how the usual Coulomb interaction arises based on the free-space mode structure. The Coulomb kernel in eq 36 is connected to the inverse of the longitudinal modes of the electromagnetic field, i.e., the (distributional) eigenfunctions of $-V^2$ from eq 10. Due to the high consistency between light an matter (see also Section 3.3 and Appendix A.4), we can express the longitudinal interaction simply in terms of the scalar (distributional) eigenfunctions and hence find for eq 36 the usual Coulomb kernel

$$\frac{1}{4\pi |\mathbf{r} - \mathbf{r}'|} = \int \frac{\mathbf{r} \langle \mathbf{k} | \mathbf{k}' \rangle \langle \mathbf{k}' | \mathbf{r}' \rangle}{\omega_k^2 + \omega_{\alpha}^2} \frac{e^{-\omega_{\alpha} (\mathbf{k} - \mathbf{k}' - \mathbf{r} + \mathbf{r}')}}{(2\pi)^3} d\mathbf{k}$$

(49)

Now changing the mode structure will also affect the longitudinal eigenfunctions and with this lead to a modified Coulomb interaction. Thus, in eq 39, we might need to replace the Coulomb kernels by a modified kernel that takes into account this change of interaction. In certain cases it is argued that this modification would be the main difference to free space. Alternatively, especially for nanoplasmonic cavities, one might instead take into account just one or a few quantized longitudinal modes of the photonic structure explicitly. We will comment on this a little later below.

Changing the mode structure in the dipole approximation, however, has a few further subtle consequences. First, if we have a (discretized) continuum of modes we will have to work with bare masses as already discussed in Section 3.2. In dipole approximation, the connection between the (single-particle) bare mass $m$ and physical mass $m_p$ is known nonperturbatively as

$$m_p = m + \frac{c^2}{4\pi\epsilon_0\hbar} \Lambda$$

(50)

where the term in the parentheses is the fine structure constant and $\Lambda$ is the ultraviolet cutoff wavenumber. This already implies that the cutoff should not be chosen too small else in our case we would need an unphysical negative bare mass, i.e., in dipole approximation nonrelativistic QED is not fully renormalizable (for a single electron the energy where this happens is, however, gigantic). If we change the mode structure, the connection between bare and physical mass will change, in general. In most cases of polaritonic chemistry, it is, however, tacitly assumed that the changes in the mode structures are not so severe as to modify this completely. Hence one usually subsumes the continuum of modes in the physical mass and only keeps a few “enhanced” modes explicitly in the calculations. Indeed, usually just one mode is kept. On the other hand, if we use a discretized continuum, we have included radiative dissipation and decoherence. In other words, since we have very many photonic degrees of freedom, the quantum revival time tends to infinity and hence we have effectively irreversible processes. This is broken once we use the physical mass of the particles and merely keep a few effective modes. To reintroduce the irreversibility, often artificial baths are included in a few mode calculation. But in principle such open-system approaches are not needed in nonrelativistic QED as it would contain all dissipation channels explicitly.

One last subtle but very important point concerns the self-polarization term $(\epsilon_\alpha \mathbf{R})^2$. While often one might hope that this term, which causes the difference between the electric and the displacement field, is not very important, it turns out that without this term the ab initio theory becomes unstable and leads to unphysical results. Indeed, no basis-set limit exists without self-polarization, i.e., the theory has no eigenstates that could be approximated by a finite basis expansion, and an unphysical coordinate- and gauge-dependence is introduced. Thus, the results can become highly unphysical for a finite number of basis states, such as having
alleged ground states with nonzero transverse (propagating) electric fields. Physically that is easy to understand, since one can only discard this term if one had a perfectly localized system of the form $\delta^2(r)$, which is impossible in quantum mechanics$^{2,14,22}$ (see Appendix A.4 for further details). Therefore, this assumption is equivalent to that of a classical particle at the origin of the coordinate system with some internal structure. Consequently, the self-interaction term must be included for a physical ab initio quantum theory in the length gauge. This statement holds true, of course, also if the mode structure is changed, as discussed above. We note that the effect of the dipole self-energy term is often not to change the result of a purely dipolar (perturbative or few-level) calculation but to stabilize it and guarantee a unique basis-set limit. Yet it depends on the specific setup and the quantities under investigation whether a decisive difference between a perturbative/few-level and a full ab initio calculation can be observed.\textsuperscript{108} Importantly, for longitudinal modes that are potentially due to, e.g., a nanoplasmonic cavity, self-polarization terms also need to be taken into account in an ab initio description. This becomes clear from the fact that in principle also longitudinal interactions can be treated in terms of the auxiliary displacement and magnetization fields (see eqs 42 and 43). However, this leads to several mathematical issues for a full continuum of modes and one therefore usually assumes that such terms can be replaced by the usual Coulomb interaction in free space.\textsuperscript{107} Yet for individual longitudinal modes, which are changed, e.g., due to a nanocavity, such a procedure is straightforward. Because of the manifest positive energy of the photon field, we must include a self-polarization term, otherwise one could lower the energy indefinitely and no basis set limit is possible.\textsuperscript{108,121} (see also Appendix B for a simple proof of this fact). In practice, this issue can often be circumvented by restricting space to a finite simulation box with certain boundary conditions, which then serves the same purpose as a self-interaction term. The size of the box, however, then becomes a parameter of the ab initio theory and must be chosen with care. Which way we ever turn it, a stable ab initio quantum theory dictates to include quadratic (beyond linear) terms and the only difference with respect to the transverse case of eq 39 is that the quadratic contribution might be different (since nonzero longitudinal fields are physically possible even for static eigenstates). The same condition appears in any other coupled ab initio quantum systems, such as electron–phonon systems.\textsuperscript{143,144}

To conclude this extensive part of the discussion: For a practical first-principles calculation on the level of the dipole-approximated Pauli-Fierz theory one needs to choose the mode structure of the photonic environment at the position of the matter system (obtained from experiment, a separate simulation, or other theories) and choose a cutoff for these modes and the corresponding bare mass. We note again that, in contrast to the full minimal-coupling theory, dipole-approximated Pauli-Fierz theory is known to be not fully renormalizable.\textsuperscript{22} This is, however, simple to understand on physical grounds: Infinitely high photon momenta directly contradict the basic assumption that the wavelength of these modes is large compared to that of the matter subsystem.

Finally, after having assumed the dipole approximation, subsuming the continuum of modes in the physical masses of the particles and keeping only one effective mode (this means integrating over the part of the continuum that has been enhanced and thus deducing an effective single-mode coupling), we arrive at the starting point of most currently employed theoretical models in polaritonic chemistry. Upon reducing the matter state to just two states, i.e., a ground and excited state irrespective of whether one considers electronic, rotational or vibrational excitations, one reaches the Rabi model.\textsuperscript{13} With these approximations, the dipole self-energy term becomes a constant offset and is therefore often discarded. Making then the rotating-wave approximation one finds the famous Jaynes-Cummings model \textsuperscript{47} that is virtually always invoked when discussing QED chemistry.\textsuperscript{1,5,14,39,46,47,48} If one wants to consider an ensemble of two-level systems, one then often employs the further approximated Dicke or Tavis-Cummings models. The latter becomes equivalent to an effectively scaled Jaynes-Cummings model.\textsuperscript{1,5,38,39,48} The Dicke or Tavis-Cummings models assume that the individual physical systems, e.g., molecules, are so far apart that they do not interact with each other directly but only couple via the cavity mode. Yet in the model, the dipole self-energy term, which necessarily arises in an ab initio theory in the length gauge beyond only two levels, is discarded (perfect localization of the whole ensemble is assumed), and no spatial information on the individual systems is kept. We note that also on this level of approximation, the choice and knowledge of the gauge is crucial. If the Dicke or Tavis-Cummings model is interpreted in terms of the length gauge without the dipole self-energy, it is possible to find the unphysical case of nonzero transverse electric field in the ground state. If the Dicke or Tavis-Cummings model is interpreted in terms of the Coulomb gauge, then such unphysical results are avoided.

Of course, there are many more advanced models and alternative theoretical approaches, and they are discussed in many of the available reviews highlighted throughout this work. Yet to keep the amount of theoretical approaches to be covered in detail tractable, we in the following focus on ab initio approaches (based on the mathematical nomenclature defined in Appendix A.5) and only highlight more advanced models and other theoretical approaches where necessary.

4. FIRST-PRINCIPLES APPROACHES TO NONRELATIVISTIC QED

“To better understand the properties of the hybrid states, further development of QED chemistry calculation methods, akin to those in quantum chemistry, would be extremely valuable.”

T.W. Ebbesen in ref 14.

If we do not want to rely on the many (potentially) restrictive assumptions underlying many of the currently employed models, as introduced at the end of the previous section, then we need to find alternative approaches to handle the extreme complexity of the Pauli-Fierz Hamiltonian. For this purpose, we will rewrite the problem of nonrelativistic QED in convenient ways that allow (in practice approximate) solutions of the general Pauli-Fierz Hamiltonian numerically. This means that we want to solve eq 38 either for the Hamiltonian of eq 37 or of eq 39 without using too much apriori knowledge or assumptions, e.g., which matter or light states are assumed to be the most important ones. However, before we continue, we generalize the Pauli-Fierz Hamiltonians even further. This is helpful for several reasons: First, for density functional methods (see Section 4.1) we need to include external fields to establish the necessary mappings.\textsuperscript{92,145} Second, external fields are natural to calculate, e.g., absorption spectra or to
investigate how a laser would induce nonequilibrium dynamics. Third, in various approximations, e.g., the cavity Born–Oppenheimer approach (see Section 4.4), internal degrees of freedom become effective external fields and hence it is helpful to see how (and which) external fields are included in the Pauli-Fierz Hamiltonian. Therefore, in the full minimal-coupling eq 37 we replace
\[ \hat{A}_\perp(r) \rightarrow \hat{A}_\perp(r) + A_{\text{ext}}(rt) \]
and add the terms
\[ \sum_{l=1}^{N} -|e|\psi_{\text{ext}}(r,t) + \sum_{l=1}^{N} Z_{l}|\psi_{\text{ext}}(r,t) \]
and
\[ \frac{-1}{c} \int_{\text{ext}}(rt) \cdot \hat{A}_\perp(r) \]
This means we now include external classical electromagnetic fields \((\psi_{\text{ext}}(r), A_{\text{ext}}(rt))\) to act directly on the matter subsystem, and an external classical current \(J_{\text{ext}}(rt)\) to act directly on the photons. In the Pauli-Fierz Hamiltonian we can even define the (fully quantized) laser pulse by the chosen initial state of the photon subsystem. This ambiguity raises interesting questions about how to best describe, for instance, a laser pulse and what are the differences in the descriptions.\(^{146}\) We further note that we have here subsumed the zero component of the external charge current, i.e., \(\rho_{\text{ext}}(rt)\), in \(\psi_{\text{ext}}(rt)\) since in Coulomb gauge we can just use eq 36 to connect both. Further, due to the Coulomb gauge we could even restrict to only the transverse part of \(J_{\text{ext}}(rt)\) in accordance to the quantized field being only transverse\(^{81,94}\) (compare also to eq 31). We note in passing that the moment we consider also external fields, we effectively gain a second gauge freedom. The physical results will not depend on the choice of the gauge of the external field, and we do not necessarily need to choose the internal and the external fields to have the same gauge. In contrast to the gauge choice of the internal fields (see Section 3.2 for further details), it is straightforward to change the gauge of the classical external fields. Having included such general external time-dependent fields leads to an explicitly time-dependent Hamiltonian \(H_{\text{PP}}(t)\).

For the dipole-approximated theory of eq 39 in length gauge we add merely
\[ \sum_{l=1}^{N} -|e|\psi_{\text{ext}}(r,t) + \sum_{l=1}^{N} Z_{l}|\psi_{\text{ext}}(r,t) \]
and
\[ \sum_{\alpha=1}^{M_{l}} q_{\alpha}j_{\alpha}(t) \]
where the last term corresponds to eq 53. There are, however, several transformations in between\(^{94,124}\) and so \(j_{\alpha}(t)\) is proportional to the mode-resolved time-derivative of \(J_{\text{ext}}(rt)\). And accordingly we find in this case an explicit time-dependent dipole-approximated Pauli-Fierz Hamiltonian \(H_{\text{PP}}(t)\).

In the following we want to present different first-principles methods for nonrelativistic QED. Similarly to ab initio methods in quantum mechanics, every approach has certain advantages and drawbacks. Which one to use will depend not only on the system under study or the investigated effects but also on the level of detail, e.g., whether the full wave function should be accessible (at least approximately) or whether reduced physical quantities suffice. The good thing is that many of these methods have overlapping fields of application and can hence be used to validate results obtained with a different ab initio QED approach.\(^{147,148}\) All of these approaches are extensions of quantum-mechanical methods that have been applied successfully in theoretical chemistry and electronic structure theory for many decades. These approaches therefore aim at describing molecular systems coupled to photons on the same level of detail as their quantum-mechanical (matter-only) counterparts. We note that there are many advanced models and alternative theoretical methods for molecular polaritons (see, e.g., refs \(^{149–163}\)) that have a more quantum-optical background and hence are geared more toward photonic observables. They are discussed in detail in various reviews on QED chemistry, e.g., refs.\(^{37,58,45}\)

Notice that although the focus of this review is on ab initio QED, the authors do not imply that first-principles QED approaches are scientifically superior to other theoretical methods applied in polaritonic chemistry. On the contrary, as also discussed in the introduction, any of those methods serves its purpose with different intrinsic advantages and disadvantages. Indeed, an important goal of theoretical polaritonic chemistry is to develop a comprehensive picture of the underlying mechanisms that subsumes the different theoretical viewpoints and encompasses (quantum) optics, (quantum) chemistry, and electronic structure theory (see also discussion in Appendix A.5).

4.1. Quantum-Electrodynamical Density Functional theory
Quantum-electrodynamical density functional theory (QEDFT) follows the seminal ideas originally developed by Kohn, Hohenberg, and Sham for the electronic ground state\(^{164,165}\) and later by Runge and Gross for the time-dependent situation of electronic quantum mechanics.\(^{160,167}\) While the fundamental theorems for the static and the time-dependent situation use different quantities we want to follow here the more general time-dependent perspective which encompasses the static case as well.\(^{168–170}\)

The basic idea is to replace the high dimensional wave function as a descriptor of the system with a reduced/collective physical variable. This is a ubiquitous idea in physics. For instance, in classical mechanics, the description of a fluid is not based on the humongous phase space of all the individual particles but on density and velocity fields such as in the Navier–Stokes equations. A different example is the use of reduced Green’s functions in many-body quantum physics.\(^{3,171}\) The main advantage of a density-functional reformulation is that we can do this reformulation in an exact manner. That is, we want to guarantee that we can recover the exact results of the wave function formulation, at least in principle. In more technical terms, we want to have a bijective mapping between the set of all physical wave functions and the set of collective variables.\(^{118,164,166}\) In this way, once we know the values of the collective variables, we can uniquely identify the corresponding wave function and determine all observables from it (see Appendix C for details on the basic QEDFT mappings). The existence of such a mapping can be recast into the question of whether one can find a closed set of equations that are deduced from the Hamiltonian description in terms of wave functions and that only include the collective variables. In the
case of eq 37 we find these two equations that form a closed set to be 

\[ \partial_t J^\gamma (r, t) = \frac{i}{\hbar} (\langle \Psi (t) | [\hat{\mathcal{H}}_{\text{ex}}, \hat{\mathcal{F}}^\gamma (t)] | \Psi (t) \rangle + \langle \Psi (t) | \partial_t \hat{\mathcal{F}}(r, t) \rangle | \Psi (t) \rangle ) \]

(56)

\[ \left( \frac{1}{c^2} \nabla^2 - \mathbf{V} \right) A^\nu (r, t) = \mu_0 c \mathbf{J}^\nu (r, t) \]

(57)

where \( \mathbf{J}^\nu (r, t) \) is the total charge current density operator that is explicitly time-dependent even in the Schrödinger picture if we have a time-dependent external vector potential. Eq 56 is a local force equation, and eq 57 is the Maxwell equation in the Coulomb gauge of the internal fields induced by the (transverse part of the) charge current density.

Of course, the problem is that we do not know all the terms on the right-hand side of eq 56 explicitly in terms of \( (\mathcal{F}(r), A^\nu (r, t)) \). So in practice we have to resort to approximations, similar to the case of standard electronic density functional theories. Note, however, that for eqs 56 and 57 gauge and relativistic invariance become much easier to enforce then for the wave function formulation, and indeed on a QEDFT level it might be beneficial to employ these facts for more accurate approximation strategies in the future. Yet here we stay in Coulomb gauge and follow the seminal ideas of Kohn and Sham, who proposed that in order to approximate such complicated momentum-stress and interaction-stress terms we should use an auxiliary system, which is as close as possible to the original problem, yet is still numerically tractable.

So in practice a system of noninteracting electrons, nuclei/ions, and photons is usually solved that generate the same equations and, on the other hand, that the subsystems (electrons, nuclei/ions, and photons) have vastly different energy/time and length/momentum scales. This makes the development of special multisystem and multiscale methods necessary. An important technical aspect, that connects back to the introduction of the Riemann-Silberstein formulation of classical electrodynamics (see Section 2), is to recast everything as first-order equations in time such as to (re)use the same numerical propagation routines.

The first-order equations of the different particle species then need to be solved self-consistently, i.e., the full feedback between the different subsystems (electrons, nuclei/ions, and photons) is included. Another technical aspect, specifically with respect to the Maxwell’s equation, is to simulate free space by working in a finite simulation box and to use perfectly matched layers.

This gives rise to radiative dissipation and decoherence from first-principles; i.e., if the system of interest and its photonic environment are enclosed within the Maxwell simulation box, photons that reach the boundary of this simulation box are emitted (lost) to the far field. In this way the transient nature of of photonic excitations in realistic nanophotonic environments can be captured (see Section 5.1 for further details). Finally, owing to the difference in mass between the nuclei/ions and electrons, one often makes a further approximation and simulates the nuclei/ions by classical statistical methods, e.g., multitrajectory Ehrenfest methods. It is within this approximation for the nuclei/ions that QEDFT for eq 37 has been successfully applied.

Of course, in many practical situations, especially in the case of molecular systems, a full minimal-coupling description is not always needed (although it is still desirable to have such high-level solutions even in such cases in order to justify approximations). So one often uses QEDFT in the long-wavelength (dipole) approximation, where eqs 56 and 57 reduce to the corresponding equations for the Hamiltonian of eq 39. QEDFT can indeed seamlessly connect to this and various other limiting cases. Before we discuss...
specifically QEDFT in the dipole-coupling limit, we want to highlight a related methodology that can be applied in an intermediate regime. For two-dimensional materials, one can approximate the in-plane and out-of-plane coupling differently. Such an ansatz was considered by the authors of ref. 182 which investigated two-dimensional materials weakly coupled to a cavity and the arising Purcell effect, i.e., the cavity-induced faster spontaneous emission of photons. They employed macroscopic QED to quantify the field of the cavity and then coupled it with the help of Wigner-Weisskopf theory (only zero or a single photon in each mode and effective particle masses) to (electron-only) density functional Kohn–Sham wave functions. Since in this approach light and matter are treated separately, e.g., matter is described in Coulomb-gauged density-functional theory while the Maxwell field is quantized in Weyl gauge, extra care has to be taken to not generate unphysical effects. We note, however, that with the help of the fundamental mapping theorems of QEDFT (see also Appendix C), one could unambiguously connect macroscopic QED with a Pauli-Kohn–Sham wave function and provide advanced light–matter interaction functionalities.

The separate quantization of light and matter becomes less error prone if we consider the interaction with the transverse electromagnetic modes in the dipole approximation (see Section 3.3 for details). Within dipole-approximated QEDFT124,145,183–186 dissipation and decoherence is still included184,186,187 if the discretized continuum of photon modes is kept, and one can thus investigate, e.g., the superradiance of a collection of molecules, mass-renormalization effects and changes in the spontaneous emission.184 To reduce the numerical costs even further, one can either reduce the mode number to a few (or merely one) effective modes or one can, for example, approximate the photon modes by an instantaneous radiation-reaction potential.188,189 Most of the results in polaritonic chemistry obtained with QEDFT-related methods employ one of these limits (see Section 5). The radiation-reaction approach is specifically efficient in including simple Markovian dissipation and allows, in combination with linear-response theory, to reach the macroscopic collective-coupling limit and explore its implications for real molecules in the dilute gas limit.190 For plasmonic situations one can either include the plasmonic structure itself or (more approximately) some quantized effective (potentially longitudinal) modes (see also Appendix B) or even just modify the Coulomb interaction (see also Section 3.3). We finally note that once we take the coupling to the (now only few) transverse modes of the photonic structure to zero, QEDFT recovers standard (time-dependent) density-functional theory.181 Time-dependent density-functional theory is then often sufficient to capture strong-coupling effects to longitudinal modes of plasmonic cavities if the plasmonic nanostructure is treated explicitly.191–195

All in all, QEDFT is highly versatile and allows access to electronic, photonic, and nuclear/ionic quantities and their self-consistent interplay on various levels of approximation. The main disadvantages involve that it is not easy to assess the error of an approximate density-functional for a given level of theory, and it is not straightforward to access observables that are not trivially given by the auxiliary Pauli-Kohn–Sham wave functions.

4.2. Exact Results

While QEDFT is able to treat the different forms of the Pauli-Fierz Hamiltonian efficiently, in one way or another, the results are usually approximate. For validation purposes and elementary insights, it would be good to have exact results. However, for coupled light–matter systems, not many exact results (analytic or numerical) are available. To the best of our knowledge, only for dipole coupling some exact reference results are known, whose main insights are summarized in the following.

First, we assume the dipole approximated light–matter coupling of eq 39 and consider a single particle trapped in a harmonic potential. It can be shown192 that the time-dependent dipole moment of this particle can be computed by just solving the classical equation of motion of the harmonically trapped particle coupled to the Maxwell’s equation, instead of solving the full quantum field problem. This example is also a good rationalization of QEDFT, where the coupled eqs 56 and 57 are directly reduced to these classical equations for this case. The computed time evolution of the dipole moment allows to access, e.g., the lifetimes of the excited states and absorption/emission spectra.

Staying with a harmonic potential, recently analytically exact results of the influence of the photon field with many (identical) interacting particles have been presented and implications discussed, e.g., that even for a ground state resonant behavior can be observed.196 Furthermore, exact analytical results are available for free particles (electrons)199 which have been used to devise approximations within QEDFT.170 Besides others, interesting effects on the linear response of the system have been highlighted (e.g., the appearance of plasmon-polariton resonances and a decrease of the Drude peak), and mass renormalization effects due to the thermodynamic limit of the photon field have been shown. In both cases, the authors have used that in velocity gauge the photon field couples only to the center of charge of the total system directly and that this then leads to only an indirect modification of the relative degrees of freedom.

A different example concerns the one-mode approximation. In this case, numerically exact results are available for a quantum three-body system coupled to this effective mode. For He, HD+ and H3+ one can reformulate the 10 dimensional problem in a problem-specific coordinate system and solve for the lowest lying eigenstates.197,198 This seemingly simple problem already provides many new insights and effects that we will highlight in Section 5. Suffice it to say that already for the simple one-mode case, the eigenvalues of the problem, without any further knowledge, lose the simple interpretation they have in standard quantum mechanics (see also the discussion in Section 3.2 concerning the loss of excited states in QED). Because one has access to the lowest-lying eigenstates in this numerically exact approach, one can also calculate the exact thermal (canonical) ensemble and deduce cavity-modified thermal properties, which will be discussed in Section 5.

The main drawback of either analytically or numerically exact ab initio results is that they are only available for very specific situations and cannot be applied to different, chemically more relevant cases.
### 4.3. Quantum Electrodynamics Coupled Cluster Theory for Electronic Strong Coupling

A compromise between generality and accuracy can be found if we restrict to eq 39 in the static case from the start and additionally treat the nuclei as external (clamped) quantities. Afterward quantum electrodynamics coupled cluster (QED-CC) theory can be employed for electronic strong coupling conditions, which has become another important first-principle QED method. In contrast to many-body methods, such as QEDPT, QED-CC theory tries to approximate the many-body wave function of electrons and photons directly. We note that alternative wave function-based methods are available (see e.g. refs 201, 202), but we will not elaborate further on those in this review. The exact electron–photon wave function in QED-CC is re-expressed by applying a cluster (excitation) operator \( \hat{T} \) on a reference wave function

\[
|\Psi\rangle = \exp(\hat{T})|R\rangle
\]

(62)

where \( |R\rangle \) is usually the tensor product of the electronic Hartree–Fock wave function with the vacuum states of the modes \( \alpha \). In addition to standard coupled cluster theory the cluster operator now also contains photonic contributions and reads for a single cavity mode as

\[
\hat{T} = \sum_{\mu,n} t_{\mu,n} \hat{b}_{\mu}^\dagger \hat{a}^\dagger \hat{a}^n
\]

(63)

Here \( \hat{b}_{\mu}\in\{\hat{b}_{\alpha}, \hat{b}_{\gamma}\} \) are the electronic excitation operators of rank \( \mu \), \( n \) is the number of photons in the mode, and the unknown parameters (amplitudes) \( t_{\mu,n} \) are to be determined. Also, when comparing to standard coupled cluster theory, one might wonder whether the bosonic nature of the photons implies some sort of symmetrization of the mode wave functions. Yet in eq 39 the bosonic nature of the photons is made explicit by the quantum harmonic oscillators \( \alpha \). This happens because the excitations of a quantum harmonic oscillator \( \alpha \) are connected to the number of photons (note that, as discussed in Section 3.3, the "length gauge photons" are not the physically observed photons) in this mode \( \alpha \); i.e., we can have infinitely many photons (bosons) in one mode. The expression of eq 62 for the wave function \( |\Psi\rangle \) leads to an expansion in the number of electronic and photonic excitations that, even if truncated early, gives very accurate results provided that the exact ground state \( |\Psi\rangle \) is dominated by the single reference wave function \( |R\rangle \).

The choice of truncation is important and in practice the number of electronic excitations is chosen as two (with potentially perturbative triples) and the mixed electronic-photonic and pure photon excitations in each mode is either one or two.\(^{148,203,204}\) This truncation allows us to perform practical calculations for relatively large systems. Embedding approaches allow to reach larger systems,\(^{205}\) where only part of the problem is treated on the QED-CC level and other parts with, e.g., a Hartree–Fock-type approximation. Ultrastrong and deep-strong coupling,\(^{39}\) where many more than just one excitation per mode arise, need a truncation at higher excitations. Here reformulations of the problem in an adapted basis\(^{170,206}\) might prove helpful. Further, as can be inferred from eq 63, treating many different photon modes can become numerically costly, and no extension to full minimal-coupling (see eq 38 with the nuclei/ions treated classically) has been devised as of yet. However, the reformulation of our QED eigenvalue problem in terms of (unitary) coupled cluster theory allows for a relatively straightforward implementation\(^{207}\) on noisy intermediate scale quantum devices\(^{208,209}\) employing variational quantum eigensolvers.\(^{209,210}\) The representation on a quantum computer has the appealing feature that in principle, many (entangled) photon modes could efficiently be represented in contrast to classical devices.

### 4.4. Nuclear/Ionic Dynamics in the Generalized Born-Huang Picture

If we want to investigate properties of the nuclear/ionic degrees of freedom when strongly coupled to a cavity, we usually take the dipole-coupling approximation and hence start from eq 39. In this case we first perform a generalization of the Born-Huang expansion.\(^{58,112,113}\) That is, we re-express the fully correlated wave function of electrons, nuclei/ions and photons

\[
\Psi(r_1, \sigma_1, \ldots, r_N, \sigma_N, R_1, S_1, \ldots, R_{N_c} S_{N_c}, q_1, \ldots, q_{M_r})
\]

\[
= \sum_{\epsilon} \sum_{\mu} \frac{\chi_\mu(R)}{|R\rangle}
\]

\[
\times \sum_{\epsilon} \sum_{\mu} \frac{\psi_\mu(r; \{ R \})}{|\psi_\mu(r; \{ R \})\rangle}
\]

in terms of conditional wave functions. There are now several ways of how to perform this expansion, i.e., which subsystem depends conditionally on the others. While the generalized Born-Huang expansion is exact irrespective of partitioning, the

---

**Figure 3.** Two main forms of the generalized Born-Huang expansion for coupled light–matter systems discussed in the main text. While the cavity Born–Oppenheimer partitioning is geared toward ground-state chemical reactions, the polaritonic energy surface partitioning is more geared toward photochemical processes.
The choice of partitioning is important when performing approximations. We will here focus on the two most relevant choices for investigating the nuclear and ionic degrees of freedom (see Figure 3). Which choice is more appropriate for approximating the problem at hand depends on the physical setup and is thus determined by the chemical system inside the photonic environment and the properties of the photonic environment itself, i.e., how do the matter excitations compare to the modes of the cavity structure. We will discuss this in more detail below. We note that there are also alternative schemes, such as exact factorization approaches, that we will not go into further detail here.

The first choice, which we call the cavity Born–Oppenheimer approach, is to group the photons with the nuclei/ions and to make the electrons depend parametrically on $\mathbf{R}$ and $\mathbf{q}$. In this case, in order to find the exact solution for

$$
E\Psi(\mathbf{r}, \mathbf{R}, \mathbf{q}) = \hat{H}_{\mathbf{pp}}\Psi(\mathbf{r}, \mathbf{R}, \mathbf{q})
$$

via the generalized Born-Huang expansion, we have to solve the equations

$$
E_{\mathbf{q}}(\mathbf{R}, \mathbf{q})\psi(\mathbf{r}; (\mathbf{R}, \mathbf{q})) = \hat{H}_{\mathbf{pp}}(\mathbf{R}, \mathbf{q})\psi(\mathbf{r}; (\mathbf{R}, \mathbf{q}))
$$

where $\hat{H}_{\mathbf{pp}}(\mathbf{R}, \mathbf{q})$ is the Hamiltonian of eq 39 parametrically dependent on $\mathbf{R}$ and $\mathbf{q}$ and the kinetic nuclear/ionic and photonic parts are set to zero (treated classically), together with

$$
E_{\mathbf{q}}(\mathbf{R}, \mathbf{q}) = E_{\mathbf{q}}(\mathbf{R}, \mathbf{q})\chi_{\mathbf{q}}(\mathbf{R}, \mathbf{q}) + \sum_{\mathbf{q}=1}^{\mathbf{q}} \int d\mathbf{r} \psi^{*}(\mathbf{r}; (\mathbf{R}, \mathbf{q})) \left( \sum_{\alpha=1}^{M_{R}} \frac{\hbar^{2}}{2m_{\alpha}} \mathbf{R}_{\alpha}^{2} + \sum_{\alpha=1}^{M_{q}} \frac{\hbar^{2}}{2m_{\alpha}} \mathbf{q}_{\alpha}^{2} \right) \psi(\mathbf{r}; (\mathbf{R}, \mathbf{q})) \chi_{\mathbf{q}}(\mathbf{R}, \mathbf{q})
$$

We note that the last term in eq 67 contains all derivatives and thus also all non-adiabatic couplings between the polaritonic nuclear/ionic wave functions $\chi_{\mathbf{q}}(\mathbf{R}, \mathbf{q})$. We note that in the original work of ref 58, the off-diagonal nonadiabatic couplings were discarded. Furthermore, with respect to the usual case without photonic degrees of freedom, also the electronic potential-energy surfaces $E_{\mathbf{q}}(\mathbf{R}, \mathbf{q})$ are now changed, since they depend explicitly on $\mathbf{q}$. Thus, to distinguish them, we call them cavity (Born–Oppenheimer) potential energy surfaces. The cavity Born–Oppenheimer expansion is specifically efficient if we are interested in ground-state chemical reactions under vibrational strong coupling. If the ground-state cavity potential energy surface is well-separated from the first excited cavity potential energy surface (usually if the relevant bare cavity frequencies are much lower than the bare first electronic excitation), we can make the cavity Born–Oppenheimer approximation $\Psi(\mathbf{r}, \mathbf{R}, \mathbf{q}) \approx \chi_{\mathbf{q}}(\mathbf{R}, \mathbf{q})\psi(\mathbf{r}; (\mathbf{R}, \mathbf{q}))$. However, even the resulting simplified equations are far from trivial and we discuss various first-principles approaches to approximately solve them below.

A second important partitioning is to choose the electrons grouped with the photons (see Figure 3), such that the resulting potential energy surfaces $E_{\mathbf{R}}(\mathbf{R})$ are polaritonic energy surfaces. If we partition also the electron-photon conditional wave function, we can solve the photonic part analytically. We therefore still have only two coupled equations, one for nuclei/ions on polaritonic energy surfaces and one for electrons, yet the analytic solution of the photons leads to novel (analytically known) nonadiabatic coupling elements among electronic states as well as among nuclear/ionic states. These new analytically known nonadiabatic coupling elements are akin to the couplings in Floquet theory, i.e., they connect states with different number of excited photons. The partitioning chosen here, which leads to the polaritonic potential energy surfaces, is now specifically efficient if one is interested in photochemistry, where the cavity modes are in resonance with electronic excitations and we consider the influence of electronic strong coupling on chemistry. In this case, if we assume that the novel nonadiabatic couplings in the nuclear/ionic sector are negligible and the photons only couple efficiently to the electronic sector, we find the explicit polariton approximation. In this case, the nuclear/ionic degrees of freedom are only indirectly modified by the photon degrees of freedom due to changes in the potential energy surfaces. If we further assume that in the electronic sector the coupling to the cavity modes acts only perturbatively, we recover polaritonic potential energy surfaces as originally introduced in ref 217.

Either way, in order to determine the influence of the cavity modes on the nuclear/ionic subsystem, we, in principle, need to solve high-dimensional coupled ab initio quantum equations. A similar problem appears also for the usual electron–nucleus/ion dynamics, and various approaches have been developed to approximately solve such situations. However, when compared to the traditional electron–nucleus/ion-only problem, the inclusion of the photonic modes implies novel nonadiabatic coupling terms which might become important to faithfully describe certain effects.

If we assume only a very few nuclear/ionic degrees of freedom to be relevant, one can cut back on the dimensionality of the problem and perform numerically exact simulations. We note, however, that a priori it is not clear whether the same nuclear/ionic degrees of freedom are relevant as outside a photonic structure, e.g., that the cavity can correlate nuclear/ionic degrees of freedom that are largely uncorrelated outside the cavity. Hints toward this issue are highlighted in Section 5.3. For this often an adiabatic to quasi-diabatic basis transformation is performed, which makes the treatment of nonadiabatic couplings and (potentially cavity-induced) conical intersections simpler. Such simulations show that the influence of the cavity also on the electronic (nonadiabatic couplings) degrees and a consistent treatment of the dipole self-energy terms (see Section 3.3.1) can be decisive.

If a strong a priori reduction to merely a few nuclear/ionic degrees of freedom is not possible then one can, for instance, extend the multiconfigurational Hartree approach to polaritonic problems. Alternatively, the use of path-integral methods and ring-polymer quantization of light and the nuclear/ionic degrees of freedom allows to investigate higher-dimensional (in terms of photonic and nuclear-ionic degrees of freedom) cases. Simplifying even further, especially in the case of thermally driven chemical reactions, extensions of a semiclassical methods or surface-hopping approaches to coupled nuclei/ion-phonon systems are possible. Here the use of cavity Born–Oppenheimer potential energy surfaces as introduced in eqs 66 and 67 seems the best choice to formulate a generalization of molecular-dynamics simulations for coupled cavity-nuclei/ions systems. One should, however, be careful regarding the
treatment of the nuclear/ionic and photonic degrees of freedom. The displacement field dynamics in eq 67 can be orders of magnitude faster than the nuclear/ionic dynamics and hence might necessitate the use of adapted Langevin/open-system approaches.66 We will comment in more detail on the physically relevant implications later in Section 5.

Notice that commonly the free-space electronic surfaces or force fields are employed instead of cavity potential energy surfaces. This implies a further approximation since in principle the displacement coordinates also influence the reduced energy eigenvalues. Aside from this, it is important to note that (not only for nuclear/ion-phonon dynamics) a basis truncation has to be performed in practice, which can introduce an artificial gauge dependence in such calculations. That is, if we performed a simulation in velocity gauge and one in length gauge (see Section 3.3 for details) then at the same level of truncation we might find different results.242,243 Only for converged results we should compare different gauges (see also Appendix A for the importance of basis-set considerations in ab initio quantum physics). While one can mitigate such effects between the two relevant (length and velocity) dipole-coupled gauges,244,245 we recall (see Secs. 3.2 and 3.3) that for the original minimal coupling Hamiltonian mainly the Coulomb gauge seems practically relevant for ab initio QED. If we finally make further assumptions, e.g., that only zero- or one-photon states can be occupied and that we are in a perturbative limit such that we can discard the dipole self-energy terms (see also Section 3.3), then we recover common Dicke-type interaction models.246,247

Overall we can conclude that to accurately describe the influence of a strongly coupled photon mode on the nuclear/ionic degrees of freedom, we need access to cavity Born–Oppenheimer or polaritonic potential energy surfaces and potentially their nonadiabatic couplings. Which one of those to use is dictated by the details of the cavity and the matter system to which the cavity modes couple to. The usage of potential energy surfaces from a bare matter problem (accessible with standard quantum chemistry software) is a widely applied approximation, which neglects the modifications of the electrons by the photon field entirely, and important effects might be missing. Let us finally note that in quantum chemistry (outside of cavities) the potential energy surfaces are always with respect to a single system undergoing a chemical reaction, and the full ensemble of reacting molecules is treated statistically (as is also assumed in transition-state and Marcus theory). However, this approach is no longer straightforward to apply, considering that many molecules are collectively coupled via cavity modes. In polaritonic chemistry sometimes the concept of a “supermolecule” is invoked, with a potential energy surface that now encompasses the full ensemble. We will comment on this controversial concept that commonly assumes (quantum) coherence among a macroscopic amount of molecules later in Section 5.3.

5. POLARITONIC CHEMISTRY FROM FIRST-PRINCIPLES

“It has been argued that the Rabi splitting experienced by each molecule involved in the collective coupling is not $\hbar \Omega_A$ but $\hbar \Omega_A / \sqrt{N}$. If this were the case, the splitting would be tiny, and it is unlikely that any molecular or material property would be modified as observed experimentally.”

T.W. Ebbesen in ref 14.

Let us now turn to the main focus of this review: the modification of chemical and material properties by strong light–matter coupling. As already highlighted in the introduction, we will present here a perspective on QED chemistry, which is based on first-principles results. For more traditional perspectives on polaritonic chemistry based on various model considerations or alternative theoretical methods, we refer the reader to the many reviews available, e.g., refs 37, 38, 45 and references therein. In the following, we assume that we can capture the observed effects by employing either the Hamiltonian of eq 37, where we describe also the cavity as part of the system, or we can use the approximate Hamiltonian of eq 39, where the cavity is taken into account by modifying the mode structure of the electromagnetic field. The presented results are then obtained by solving the Schrödinger-type equations with one of the above-described first-principles methods (see Section 4 and Appendix A for the mathematical framework of ab initio QED). We want to relate the various results with each other but at the same time also highlight explicitly the underlying assumptions. Such questions of consistent assumptions turn out to be very important for various reasons, as will become clear in the next sections. First of all, QED chemistry is a novel research discipline, and many assumptions are still under debate and not yet generally accepted. Moreover, the strong coupling between light and matter can potentially invalidate accepted assumptions of theoretical chemistry, which have been successfully applied for decades outside of photonic structures. In addition, the increased theoretical complexity of polaritonic chemistry includes many additional ingredients, which makes the choice of reasonable assumptions even more delicate. For example, in most applications we have to account for

1. The chemical complexity of the (individual) molecular system under study.
2. The effect of nonzero temperature.
3. Potential chemical effects from the solvent in which the molecular system under study is contained.
4. The self-consistent interaction with the restructured (quantized) electromagnetic field.
5. The collective/cooperative effects due to an ensemble/solvent or by the photonic structure itself.

Already without a photonic structure, when only points 1–3 are relevant, the complexity is staggering. Combining the first three points encompasses most issues describing reactivity in theoretical chemistry.249–251 Adding the last two points is the origin of the observed changes in chemical properties, but also the origin of even more theoretical complexity. In more detail, they can potentially change the basic ingredients of chemistry, as has been highlighted already in the introduction, which in turn also affects how points 1–3 combine. Let us try to unravel these aspects and their connections a little more from an ab initio perspective in the following.

5.1. Restructuring the Electromagnetic Field Modes

The first fundamentally new ingredient is that a photonic structure, e.g., an optical cavity or some plasmonic structure,15–17,26 will modify locally the modes of the electromagnetic field from simple plane waves (see also Sections 2 and 3.3) to more complex forms. Of course, this restructuring is automatically contained in nonrelativistic QED if we explicitly include the photonic structure as part of the physical system.
A nice demonstration of this fact is found in, e.g., ref 81 where the time-resolved field structure between plasmonic nanospheres is considered (each nanosphere contains 297 sodium atoms and their dynamics self-consistently coupled to the Maxwell field upon irradiation with an external short laser pulse is simulated up to 40 fs). It is also shown how longitudinal and transverse electromagnetic modes are modified at the same time for such very small cavities that are explicitly treated as part of the system (see also Section 2.2 for the usual free-space distinction). It is no surprise that such near-field effects can have a strong influence on the properties and dynamics of molecules. Physically it is quite simple to understand that the (large) charge densities and currents of the nanospheres lead to a modified electromagnetic mode structure and that the fluctuations of these charge densities and currents are connected to the fluctuations of the electromagnetic field inside the cavity. Abstracting further, the photon-field fluctuations can be understood as current—current correlators between the charged particles of the cavity and the molecule inside the cavity, in analogy to the arguments that can be made for the Casimir forces.21,248 This idea also underlies the theory of macroscopic QED, where the photon field fluctuations are expressed in terms of currents obtained from linear-response functions of the cavity material.21,249 Furthermore, it is nice to observe that the local photon modes lead to strong radiative dissipation, since exciting them transfers energy from the near to the far field and this energy is effectively lost from the localized (cavity-molecule) system.81,184 One should, however, be aware that strictly speaking the photonic structure does not really generate new photon modes, but the transient nature of the excitation in the cavity material rather leads to quasi modes.21,249−251 So it is a theoretical abstraction/simplification to denote the cavity-induced local changes in the electromagnetic field as new modes.

Keeping this cautionary note in mind, we will still use the (approximate) picture of changed electromagnetic modes due to a photonic structure in the following. This becomes specifically handy when we want to unite various different physical situations where strong light–matter coupling appears. For instance, often strong coupling is not considered in a nanocavity but rather on a surface and the molecules coupled to an evanescent wave, a surface plasmon-polariton, which itself is actually a light–matter hybrid state.25 Overall the strong-coupling effects in these different physical situations are quite similar,27,81,191 at least in a coarse-grained view (see also Section 3.3.1). Now, putting one or a few molecules in contact with these modified local electromagnetic modes can have strong effects on the molecules. Such situations are commonly called single-molecule or local strong coupling. It is simple to accept that, for instance, plasmonic near-field modes, which (if excited) can generate very strong local fields, can transiently affect molecular properties or change chemical reactions.252−255 An important point is that one does not need to excite these modes externally but also at equilibrium they can have a strong influence. Indeed, the main interest in the following, as already highlighted as one of the main questions in polaritonic chemistry in the introduction, lies in the equilibrium fluctuations of these modes and their impact on the molecular properties. These fluctuations can be either of quantum nature or of thermal nature, as we will discuss in the next section.

5.1.1. Modified Fluctuations and Fields. Assuming that our coupled cavity-molecule system is completely isolated and we consider the coupled ground state (see also Section 3.2 about the existence of ground states in ab initio nonrelativistic QED), the fluctuations of these quasi modes inside the cavity are purely quantum in nature. If we then focus on the equilibrium molecule inside our photonic structure, any changes with respect to free-space equilibrium can then be attributed to the changed mode structure and its changed vacuum fluctuations. Instead, if we start from an excited state, which then can decay due to (radiative and potentially also vibronic/phononic) dissipation, we expect to observe different dynamics due to the changed mode structure. Nevertheless, in this case the main driving force will be the induced nonzero electromagnetic (near) fields and not so much the coupling to the fluctuations.146,178,184,195 Certainly, the dominant mechanism will depend on the amount of energy transferred from the molecular system to the cavity modes. If we now bring our cavity-molecule system in contact with a thermal reservoir, the mode fluctuations will additionally get a thermal component. Depending on the temperature and the energy range of the cavity coupling, e.g., ro-vibrational, vibrational, or electronic, the thermal fluctuations can dominate over the vacuum contributions. There is now, however, a simple but important point to be highlighted. While the thermal state of the total system is canonical, this is not necessarily the case anymore for the (nuclear/ionic) dynamics of the strongly coupled molecular system inside the cavity. Similarly, the thermal cavity mode fluctuations can also be very different to the empty-cavity thermal fluctuations. Only in the limit of weak coupling between light and matter can we expect to reach a canonical state for the molecular subsystem. Such effects have been observed for simple molecular systems coupled to a single cavity mode.198

A different way to quantify cavity induced modifications is to measure the impact on the basic molecular building blocks, i.e., on the electrons or on the nuclei/ions. For example, by measuring the dispersion relation and determining its curvature at zero momentum, one can determine the mass of the free particles.22,97,139 If the same measurement is performed inside a cavity, then the photonic environment will alter the dispersion. In addition, the loss of isotropy in the cavity (think about mirrors that restrict the x direction, as displayed in Figure 2) will imply that one has (slightly) different masses in different directions.22 An example of such a mass renormalization can be found, e.g., in ref 139. Notice that the mode restructuring can not only affect particle masses but also imply that the longitudinal (Coulomb) interaction between the charged particles of the molecules gets modified (see also Section 3.3).

5.1.2. Chemical Consequences of Cavity-Restructured Modes. What are the chemical consequences of the restructured photonic modes? Considering the impact on the electronic sector first, where (room) temperature effects are usually assumed negligible, one finds that the electronic ground state can get modified appreciably only for quite strong coupling, i.e., when the relevant modes correspond to large local fields if excited189,202,256 (This does, however, not mean that perturbative/few-level calculations are correct within ab initio QED for small changes, since these changes can strongly vary locally113,174). In the common dipole approximation of eq 39 this happens very roughly when for some modes α (or the sum of all the enhanced modes) we
have $g_\Omega \Delta \Omega |\psi| \approx 1$ comparable to the free-space Coulomb interaction, where $\Delta \Omega$ is the relevant (microscopic) length scale of the localized quantum system.\textsuperscript{108} It has to be highlighted (see also Section 3.3) that for the combined ground state of the cavity-molecule system no real (propagating) fields are generated but the mode occupation is virtual; i.e., it is the vacuum fluctuations of these modes that lead to changes. The hybridized nature of the ground state in a cavity can not only modify the energy or the ionization potential\textsuperscript{204,237} but also the electronic density of the ground state.\textsuperscript{174,176,200,256} For a fixed coupling strength, the magnitude of these effects also depends on the position of the (clamped) nuclei/ions, i.e., since the relevant length scale $\Delta l$ from above is also modified. For instance, if dissociating molecules are considered, a cavity mode can lead to strong effects due to novel long-range correlations\textsuperscript{258} and it can modify van der Waals interactions substantially.\textsuperscript{148}

For **time-dependent and excited state properties** modifications can be observed already for much smaller couplings compared to ground state effects. Notice that (time-dependent) excitations typically also imply further delocalization with respect to the ground state. For example, electronic (usually vacuum) Rabi splitting, the hallmark of strong coupling (see Figure 4), can usually already be observed for coupling regimes where the electronic ground state still remains unaffected.\textsuperscript{43,58,184} In most cases, the calculated Rabi splitting shows an asymmetric behavior\textsuperscript{184–186} and one also recovers the super/sub-radiant features (radiative lifetime is shorter/longer than free-space counterpart) of these polaritonic states.\textsuperscript{184,186,187} This is a nice consistency check with respect to experimental evidence. To include the radiative losses these time-dependent simulations either need to take into account the continuum of modes for a specific environment (see also Section 3.3) or explicitly include dissipation phenomenologically.\textsuperscript{188,239,259,260} Specifically interesting for chemistry is the appearance of new nonadiabatic couplings between (excited) electronic surfaces and novel conical intersections\textsuperscript{199,223,226,261} (see also Section 4.4).

For the rotational and vibrational degrees of freedom, the effects of (room) temperature can become decisive to describe chemistry. In this case the (energetically) relevant modes of the photonic structure might have a non-negligible thermal occupation. For photochemical reactions these modified thermal fluctuations might typically be less important than the new cavity-induced nonadiabatic couplings and conical intersections, but in general the interplay of these cavity-induced effects will alter chemical properties.\textsuperscript{197,198,240,241,262–264} Notice, however, that one can observe already very interesting changes in simple photochemical reactions, even when disregarding these thermal contributions.\textsuperscript{190,203,218,220,228,239,265,266} On the other hand, the modification of the thermal fluctuations are expected to be specifically important for ground-state chemical reactions\textsuperscript{4,66} and many other phenomena of materials in cavities (e.g., quantum phase transitions\textsuperscript{267,268}). We will discuss this issue in Section 5.3 in more detail for a specific case. For the generic situation, we want to highlight that the common simplification to describe classically the thermal fluctuations of the (relatively heavy) nuclei/ions is not necessarily appropriate for the fluctuations of the modes even at ambient conditions (depending on the chosen cavity frequency). The mode fluctuations can still have strong nonclassical contributions of vacuum and quantum thermal nature\textsuperscript{188} (see also Figure 5).

### 5.2. Collectivity and Cooperativity

The second fundamentally novel aspect (as also highlighted in the introduction) that becomes decisive inside a photonic structure is that the cavity can facilitate strong collective or cooperative effects. "Collective" here means that similar physical entities, e.g., the same type of molecules, start to interact with each other via the cavity and potentially synchronize, while “cooperative” means that such cross talking happens between different physical entities, e.g., solute and solvent. Strictly speaking, any effect that we observe is cooperative, due to the cavity being a different physical entity than the material inside, but this distinction inside the photonic structure is common.\textsuperscript{44,47}

In order to construct cavities that have a particular strong coupling to molecules, it is often helpful to further fill the cavity with a highly polarizable medium.\textsuperscript{14,36,41,45,48} Indeed, in many cases of QED chemistry one simply employs the molecules of interest themselves to increase the coupling effect.\textsuperscript{14,36,45,48} Clearly this cannot be done ad infinitum, since even in the gas phase the molecules get densely packed at one point and lose their individual character and hence will respond very differently. Ab initio QED simulations can nicely reproduce this behavior and recover the well-known $\sqrt{N_{mol}}$ increase of the vacuum Rabi-splitting by the number of coupled molecules $N_{mol}$\textsuperscript{196,260,269} (see also Figure 4). As we will also highlight later, this does not necessitate quantum coherence between the different molecules though. Such

---

**Figure 4.** A free-space molecule (a) has specific electronic transitions of frequency $\omega$ from its ground state $|g\rangle$ to some excited state $|e\rangle$. These transitions show up in an absorption (or emission) spectrum where some external probe pulse interacts. If the molecule is placed inside a Fabry-Pérot cavity (b) with the same resonance frequency $\omega$, one observes that the degenerate (matter and photon) excitations turn into an avoided crossing. This is due to the coupling between light and matter, and instead of one peak one finds now two peaks, i.e., the upper $|up\rangle$ and lower $|lp\rangle$ polaritons, which are split by the Rabi frequency $\Omega_R$. From the simple Jaynes-Cummings (for a single molecule) or the Tavis-Cummings (many identical molecules) model (see the end of Section 3.3) one infers that the vacuum Rabi splitting depends inversely on the volume of the Fabry-Pérot cavity, is proportional to the dipole matrix element of the individual molecules, and scales with the square root of the number of molecules as well as photons. Reproduced with permission from ref 5. Copyright 2018 Springer Nature.
collective effects are not only observed for the excited states but also for the collective ground state of molecules. One of the interesting aspects of the collective coupling situation is the appearance of dark states. That is, the ensemble of molecules can attain a collective state which does not couple to external (dipole) radiation and hence is “dark” for absorption spectra. These states will only be thermally populated and can modify the relaxation dynamics or they can also act as a thermal reservoir for the coupled “bright” collective states. An alternative approach to modify chemistry collectively is by resonantly tuning on the solvent (or highly polarizable plasmonic structures) which yields a density-dependent Rabi splitting with respect to the solvent concentration. The difference is that one hopes that the strongly coupled solvent either induces strong single-molecule coupling to the solute or that the cooperative behavior of the solvent leads in some other way to observable changes in the solute. To describe theoretically the mesoscopic amount of molecules that is present in experimental ensembles, one usually needs to make some further approximations, e.g., that the molecules (assumed in gas phase) only couple with each other via the cavity in a semiclassical way. In this way, first-principle simulations are able to also consider the macroscopic limit.

There are now two important observations to be made regarding collective and cooperative effects. First, once a molecule out of the ensemble is slightly modified, e.g., due to the onset of a chemical reaction, the distinction between collectivity and cooperativity even inside the cavity becomes fuzzy again. Indeed, ab initio simulations have shown that a strongly coupled solvent either induces strong single-molecule effects on a modified molecule similar to cooperative strong coupling. Second, for phenomenological models it is often argued that the collective effects are quantum in nature and that a robust and collectively delocalized (over a mesoscopic amount of molecules) polaritonic quantum state is generated (see Section 5.3 for more details). From an ab initio perspective, a mesoscopic quantum collective mechanism does not seem to be necessary, and in certain cases it even becomes problematic. For instance, the response of the collective system, together with the dark state configurations, can be captured purely semiclassically. Therefore, the term “state of the ensemble” does not need to imply a quantum state, since also the response of classical dipoles will show such configurations. Furthermore, it has been shown that the quantum entanglement between light and matter vanishes rapidly above zero degrees Kelvin even for simple molecular systems. On the other hand, at least for the common long wavelength approximations, the light–matter Hamiltonian is not size-extensive (where in the length gauge the main contribution comes from the dipole self-energy term). That is, the more molecules are fully quantum coherently coupled, the stronger the effect of the modes becomes (even if these molecules are arbitrarily far apart). As a simple consequence of this, the cavity modes would be strongly blue-shifted from the alleged mesoscopic amount of quantum-coherently coupled molecules, which is, however, not observed in experiments (see also Section 5.3 for an example).

To conclude, ab initio approaches provide access to collective and cooperative coupling regimes, and they reproduce the well-known effects from commonly applied models and alternative theoretical methods. However, at the same time, ab initio results suggest rather a semiclassical mechanism than a fully quantum collective/cooperative origin of the experimentally observed effects at ambient conditions.

5.2.1. Chemical Consequences of Collective Coupling. With these caveats in mind, we can ask what chemical consequences can be expected that originate from collectivity or cooperativity? First of all, essentially all previously mentioned effects in Section 5.1 can in principle arise (and even be collectively enhanced), since the coupled ensembles can mediate single-molecule strong coupling. However, we can now find additional, nontrivial modifications that emerge specifically due to having ensembles with a large number of molecules. Such effects include, for instance, ensemble-induced changes in lifetimes, modified intermolecular interactions and enhanced transport properties. In addition, how an ensemble changes local molecular properties can have a nontrivial dependence on the number of molecules in the ensemble. Of course, the probably most relevant effect for chemical applications will be the site/bond selective modifications and control of chemical reactions in an ensemble of molecules without external driving, i.e., in thermal equilibra.
We note that chemistry is local, i.e. the electronic and nuclear structure is modified on a single-molecule or nearest-neighbor level. However, in the case of collective/cooperative strong coupling, this prevalent paradigm is challenged, since chemical reactions seemingly become dependent on the total ensemble. For example, a priori it is unclear if the reaction mechanism in a cavity is altered due to a quantum-collective state, cavity-mediated intermolecular interactions, cavity-modified thermal fluctuations or single-molecule strong-coupling effects. To quantify the extent and origin of these modifications is currently one of the main goals of QED chemistry. This understanding will allow reaching a qualitative and quantitative theoretical understanding, and accurate predictions become feasible that can significantly advance experiments and applications of polaritonic chemistry.

5.3. Cavity-Modified Chemical Reactions

As pointed out before, the cavity induced contributions to the chemical complexity offer many tantalizing opportunities yet make a detailed understanding even more challenging. Their additional interplay with (single-molecule) symmetries and external probes is just getting explored and might lead to further very interesting effects. Let us next focus on a specific experiment to reduce the immense amount of possibilities and thus complexity. This paradigmatic example highlights how ab initio theory can help to unravel the main mechanisms of cavity-modified chemistry.

The seminal experiment that we consider in the following is the ground-state deprotection reaction of 1-phenyl-2-trimethylsilylacetylene (PTA) under vibrational strong coupling. The PTA molecules are mixed with tetra-n-butylammonium fluoride (TBAF) in methanol. In the ensuing deprotection reaction, fluoride ions released from TBAF interact with PTA, forming an intermediate complex, which makes the breaking of the Si–C bond in the PTA molecule more likely (see Figure 6). The Fabry-Perot cavity is then set on resonance with Si–C stretching modes at roughly 856 cm$^{-1}$. It is important to note that the cavity is not pumped except of the thermal effects due to ambient conditions. To verify the vibrational strong coupling condition, the transmission spectrum is observed and shows a large Rabi splitting. Eventually, one finds that the deprotection reaction rate is strongly suppressed for the nonpumped resonantly coupled system, when compared to free-space or off-resonant coupling. The measured suppression is also strongly dependent on the temperature of the total system, such that from fitting simple equilibrium rate models even a qualitative change in the transitions state would be predicted. This observation was interpreted as potential evidence for cavity-induced nonequilibrium effects. For the further interpretation of this experiment one important remark has to be made: The vacuum Rabi splitting of the mixture of molecules, which is the usual way to identify strong coupling situations, depends on the density of PTA molecules and their products alike, since both contain the same Si–C stretching modes. Therefore, the Rabi splitting stays constant throughout the minutes-long reaction and is a self-adapting mixture of collectivity and cooperativity.

Most of the interpretations of this experimental result follow a two-step procedure. First, a perfect ensemble of aligned PTA molecules in the gas phase with a small single-molecule coupling constant (obtained from the coupling constant of the empty Fabry–Perot cavity) is assumed. Then, by assuming zero temperature, the Dicke or Tavis-Cummings model (see Section 3.3 for all the other assumptions that go into this model) is used to determine the number of (two-level)
molecules that are quantum-collectively coupled to the vacuum of the cavity mode. Based on these assumptions, a fit suggests a mesoscopic number of quantum-collectively coupled two-level molecules on the order of $10^7$ molecules. In a second step, concepts of quantum chemistry are applied on this quantum collective state, i.e., assuming that a single collective “supermolecule” is formed with many dark states. In contrast to usual quantum chemistry, where only a single molecule and its potential energy surface is considered, the “supermolecule” has now a potential energy surface that is formed by the $10^7$ molecules (for each molecule reduced to the main free-space reaction-coordinate) plus the single-excitation subspace of the cavity mode.

This new humongous potential energy surface is then assumed to change the chemistry, since now all molecules move in a concerted motion and no longer statistically independently. However, this common combination of quantum optics and quantum chemistry concepts cannot explain (even qualitatively) the experimentally observed findings. Moreover, from a rigorous theoretical perspective even a single quantum-mechanical molecule would never attain, e.g., a permanent dipole moment or specific internal structures without coupling to the environment. It is the interaction with the environment that leads to a specific realization of the molecular structure, e.g., a certain orientation of the pyramid of the NH$_3$ molecule. In a “supermolecule”, all these (exactly similar) realization of the individual molecular structures are assumed to happen simultaneously and fully quantum-coherently due to coupling to the cavity even at ambient conditions.

Can ab initio polaritonic chemistry now help to understand this stark discrepancy between theory (based on a simplified model calculation) and experiment and maybe hint at a potential mechanism? Let us first fix the basic level of theory that we deem sufficient and computationally feasible to investigate the PTA experiment theoretically. We assume that the dipole-coupled Hamiltonian of eq 39 with one effective mode and the physical masses of the particles (see Section 3.3 for more details) is a sufficient framework for describing polaritonic chemistry in a Fabry-Pérot cavity. From the chosen Hamiltonian, the standard Hamiltonian of quantum chemistry can be directly recovered for zero coupling strength. In addition, the quantum-optical Dicke and Tavis-Cummings model can also be deduced from it. After having made this theory choice, we immediately realize that a fundamental inconsistency arises with the alleged number of quantum-collectively coupled molecules, which are suggested by the Dicke or Tavis-Cummings model (see also ref 192 for related problems with phenomenological models in plasmonic cavities). Not surprisingly, the matter inside the cavity modifies the frequency of the cavity mode, which will be accounted for in Pauli-Fierz theory. This means the enhanced refractive index of the filled cavity will shift the bare (empty cavity) frequency toward smaller wave numbers. However, the assumption of $10^7$ quantum-collectively coupled molecules would lead to a diamagnetic shift of the cavity frequencies, which is an order of magnitude larger than the experimentally observed frequency of 856 cm$^{-1}$. This discrepancy suggests that at the Pauli-Fierz level of theory, we need to restrict quantum coherence to a much smaller length scale (closer to the common understanding of chemistry as being local) and potential collectivity/cooperativity effects on a macroscopic scale will rather be semiclassical in nature. This seems reasonable since the amount of degrees of freedom (translational, rotational, vibrational, and electronic) which can lead to decoherence in a real chemical system in solvation is so breathtaking that a quantum coherence at ambient conditions over large distances seems implausible in practice.

An alternative interpretation arises if one keeps in mind that the observed Rabi splitting is not an absolute, but rather a statistical quantity, i.e., not all molecules contribute with the same amount. Therefore, there is no reason to assume that all molecules experience coupling to the cavity mode in the same way. Indeed, as discussed above, single molecules can experience strong local coupling effects in a collectively/cooperatively coupled environment. Note again that in the experiment, the (constant) Rabi splitting is by construction a mixture of collectivity and cooperativity. Consequently, it seems plausible that a fraction of the PTA molecules in the cavity could feel strong single-molecule effects, specifically in the case that they undergo a chemical reaction. Taking into account that chemical reactions are rare events and that the likelihood of these events is determined by the temperature, this fraction can become decisive for the observed rate change. This setting suggests that the cooperative/collective coupling can effectively be interpreted in terms of a highly polarizable and strongly frequency-dependent medium in the vicinity of a reacting PTA molecule. Based on this (simplified) ab initio picture, recent QEDFT simulations were able to reproduce the experimental PTA results qualitatively and they could also reproduce other predictions in connection to solvent effects.

Overall, these simulations suggest that the cavity can correlate various intramolecular vibrational modes and hence can transfer energy from the bond-breaking stretching modes to other internal motions, thus effectively strengthening the Si–C bond in the PTA experiment. This indicates that restricting to the main cavity-free degree of freedom of a potential energy surface in vibrational strong coupling simulations could miss important contributions (see the discussion in Section 4.4).

Of course, this simple local model, which infers a frequency-dependent polarizable environment from the collective/cooperative ensemble, is not the end of the story. The ab initio simulations also suggest—again in agreement with the original interpretation of the experiment—that the cavity might induce nonequilibrium effects. In the context of chemical reactions, this means that the nuclear/ionic system might follow a noncanonical (classical) thermal distribution. In contrast, for the uncoupled, bare matter system, the thermal state is usually well-described by a classical canonical distribution. Nonequilibrium dynamics for the coupled matter system is not surprising, since it is a strongly coupled subsystem; i.e., tracing out the cavity degrees of freedom will usually induce a noncanonical/nonstationary distribution for the subsystem. However, what might be more exceptional is that even for ambient conditions it is not correct to treat the cavity degrees of freedom (particularly the fluctuations) purely classically and assume that the thermal fluctuations are uncorrelated (see also Figure 5). Furthermore, it has been argued that such noncanonical dynamics of classical particles (nuclei/ions) can lead to stochastic resonances, which could explain on the ensemble level, why the experiment sees a strong frequency dependence (resonance effect) in the polaritonic reaction rates, without any external periodic driving. At the same time, stochastic resonances are quite delicate and they seem to arise only under very special conditions. This could also rationalize why in many experimental situations of strong coupling no changes in
6. CONCLUSION AND OUTLOOK

"As more researchers enter the field, influx of new viewpoints will ensure rapid development of polaritonic chemistry concepts and further pioneering cross-disciplinary breakthroughs."

K. Hirai in ref 44.

If you followed this review chronologically, then it has been a real tour-de-force. It encompasses very basic considerations of relativistic quantum physics (how relativity, symmetries and spin lead to the Maxwell equations and their coupling to matter) in Section 2, the basic Hamiltonian of nonrelativistic QED (properties and potential approximations) in Section 3, ab initio QED methods in Section 4, and their applications on relevant research questions of polaritonic chemistry in Section 5. Clearly, many of the details that were highlighted might not be relevant for a specific experiment in QED chemistry, where the restructuring of the local electromagnetic modes can modify chemical properties. However, as highlighted in the introduction, in the absence of established simple mechanistic rules for polaritonic chemistry, which challenges the locality assumption prevalent in common chemistry, a re-evaluation of all the intrinsic assumptions in our theoretical modeling is needed. This hopefully helps to select among the existing models and combinations (quantum) optics and (quantum) chemistry approaches the most reliable ones and allows the development of more accurate models in the future, in order to get an intuitive understanding of the relevant mechanisms in polaritonic chemistry.

Let us repeat in this context the main aspects of the different sections and their answers to the main questions raised in the introduction, i.e., the basic Hamiltonian, the choice of gauge, the implications of the dipole approximation, cavity-induced changes in vacuum and thermal fluctuations, and the interplay of local and collective strong coupling. In Section 2 we have shown how the light and matter sectors follow from the same basic principles and need to be treated consistently, especially when they interact. On the most basic level, one cannot even distinguish between light and matter degrees. Changing one sector can have a strong influence on the other and might even break basic physical principles. This should serve as a guidance on how to carefully recombining theoretical methods describing matter, e.g., quantum chemistry methods, and theoretical tools for photons, e.g., quantum optics methods. In Section 3 we have presented the basic Hamiltonians of nonrelativistic QED, which form the basis of a consistent and nonperturbative ab initio theory of light and matter. We highlighted that the Pauli-Fierz Hamiltonian guarantees the stability of matter, that excited states turn into resonances with finite lifetimes, and that we have to work with the bare masses of the charged particles. Furthermore, we have discussed that the Coulomb gauge is the natural gauge to work in (at least on the wave function level), because it guarantees internal consistency between quantum mechanics and quantum optics and that only in the dipole-coupling limit can we easily replace a photonic structure by a local modification of modes. We have also spelled out the various assumptions (extension of localized matter small when compared to the cavity wavelengths, single charge center to have indistinguishability, small enough frequency cutoff to avoid nonrenormalizability, linear and quadratic coupling terms to have stable theory) that go into the dipole approximation. In Section 4 we have highlighted the necessity of first principle methods to be able to cope in an unambiguous way with the humongous amount of degrees of freedom (photonic, electronic and nuclear/ionic) of a realistic coupled light–matter system. Depending on the specific question and/or coupled systems, different theoretical methods become more appropriate than others (e.g., QEDFT as a general-purpose approach, QED-CC methods for electronic strong coupling, or the cavity Born–Oppenheimer partitioning for cavity-modified ground state chemical reactions). In Section 5 we have discussed polaritonic chemistry from an ab initio QED perspective. We have highlighted the two main differences to chemistry outside of cavities, i.e., the self-consistent interaction with the restructured (quantized) electromagnetic field and collective/cooperative effects due to an ensemble/solvent, and presented several results obtained with various ab initio QED methods. Based on these results, we have argued for the importance of modified quantum/thermal fluctuations that can induce noncanonical equilibrium conditions for the matter subsystem. Furthermore, simulations suggest that collective ensemble/solvent effects are mainly semiclassical and can be approximated as a frequency-dependent modification of the local polarizable medium. We have then scrutinized a paradigmatic experiment in polaritonic chemistry and found that an interplay of single-molecule coupling with semiclassical nonequilibrium effects is indeed a natural explanation for the observed changes in chemical reactions.

Clearly, we cannot yet provide a general and universally accepted answer for all of the microscopic mechanisms at work when chemical properties are changed by a photonic structure. Many other possible effects, which have not been taken into account in the different ab initio simulations, might be important as well. The most obvious shortcoming is that a macroscopic ensemble of molecules inside a cavity cannot be simulated at a full ab initio level. However, several recent developments have made it possible to also get approximate results for the macroscopic case. Yet we believe that statistical

https://doi.org/10.1021/acs.chemrev.2c00788
Chem. Rev. XXX, XXX, XXX--XXX
and thermal effects dominate on a macroscopic scale at ambient conditions, in analogy to chemistry outside of cavities. Therefore, a semiclassical description should be appropriate to recover the observed effects. This suggests that it will be paramount to develop adapted statistical methods in the future that can faithfully include the contributions of the (quantized) cavity mode. A further issue that is often disregarded for simplicity is the effect of the solvent on chemical properties. Indeed, there are recent experiments\textsuperscript{46,47} which show that strongly coupled solvents can have different effects on chemical processes than their uncoupled counterparts. Although this fits into the simplified picture of collective/cooperative coupling as a frequency dependent polarizable surrounding for molecules, actual solvent effects can be much more intricate. Another obvious shortcoming of most considerations so far is the simplified treatment of the cavity as an effective single- or few-mode structure. Specifically for nanocavities, where a few molecules couple to plasmonic excitations, a detailed treatment of the cavity as an active physical entity, which can efficiently dissipate energy, might become crucial. A further aspect that might become important for the specific design of chemical properties is to go beyond the dipole-coupling approximation in our theoretical description. Dipole coupling implies that no momentum is transferred between light and matter and also that we lose locality (at least approximate for nonrelativistic particles) and no retardation effects are included. Beyond-dipole contributions can become specifically important once we take into account the exact structure of the modes of a cavity, e.g., when we couple strongly to chiral (circularly polarized) light modes. To disentangle which details are important, we call for a combined theoretical and experimental effort. Besides theoretical developments on the ab initio and model side, new experimental setups and observables need to be identified to unravel the influence of the above highlighted issues. It is clear that merely considering the Rabi splitting is not enough to understand the mechanisms at work, and spatially as well as temporally resolved experimental investigations are key for the future development of QED chemistry.

While this list of extra complications might seem like spelling doom for a comprehensive understanding of strongly coupled light—matter systems, it at the same time opens the door for many still to be discovered chemical effects. We hope that this review does highlight where seemingly small changes in the photonic environment might lead to novel effects. Take, for example, the quantization of the electromagnetic degrees of freedom in Section 2. By following the Riemann-Silberstein approach, we saw that it is quite natural to quantize dipole contributions to the cavity, e.g., when we couple strongly to chiral (circularly polarized) light modes. To disentangle which details are important, we call for a combined theoretical and experimental effort. Besides theoretical developments on the ab initio and model side, new experimental setups and observables need to be identified to unravel the influence of the above highlighted issues. It is clear that merely considering the Rabi splitting is not enough to understand the mechanisms at work, and spatially as well as temporally resolved experimental investigations are key for the future development of QED chemistry.

APPENDIX A: AB INITIO QUANTUM PHYSICS

In this appendix, we want to define the notion of ab initio quantum physics as we employ it throughout this review. We will highlight important mathematical subtleties that arise in ab initio quantum theories involving light and matter, which are often absent in models and alternative theoretical methods. However, those subtleties become important for the formulation of computationally efficient, nonperturbative first-principles theories such as quantum-electrodynamical density functional theory (QEDFT).

Nevertheless, we will also clarify that neither ab initio nor alternative approaches are scientifically superior, instead they serve complementary purposes and hence necessarily rely on each other.

A.1. Mathematical Setting

Let us start by setting the mathematical stage of ab initio quantum physics. Quantum theories are commonly formulated on infinite-dimensional complex Hilbert spaces.\textsuperscript{22,23} The necessity for infinite dimensions in quantum physics is relatively easy to understand. If we assumed that the Heisenberg uncertainty principle (up to $\hbar$),

\begin{equation}
\delta \hat{p} - \delta \hat{x} = i \hbar \tag{68}
\end{equation}

should hold on a finite dimensional complex Hilbert space, which is equivalent to $\mathbb{C}^N$, then we would find by the cyclic property of the trace the contradiction

\begin{equation}
0 = \text{tr}(\delta \hat{p}) - \text{tr}(\delta \hat{x}) = i \text{tr}(\mathbf{I}) = iN \tag{69}
\end{equation}

This contradiction leads to the conclusion that we cannot represent the basic commutation relations in finite dimensions.\textsuperscript{73} Commonly one further assumes in ab initio quantum physics and constructive quantum field theory separable infinite-dimensional Hilbert spaces,\textsuperscript{69,72} i.e., they allow for a countably infinite basis, such that all of these Hilbert spaces are isometrically isomorphic to the fundamental sequence space $l^2(\mathbb{N})$,\textsuperscript{73} which contains all infinite sequences $\{\psi_1, \psi_2, \ldots\}$ with $\sum_{n=1}^{\infty} |\psi_n|^2 < \infty$. We can thus consider this relatively simple space as a universal representative of the infinite-dimensional case of quantum physics. An element/state in that vector space is denoted by $|\psi\rangle = \sum_{n=1}^{\infty} \psi_n |n\rangle$. We next note that the above argument of the commutation relations is part of the Lemma of Wielandt, which states that the commutation relation can only be fulfilled if at least one of the linear operators is unbounded.\textsuperscript{73}

Unbounded operator means that we can always find a
normalized state in the Hilbert space which will attain an arbitrarily large norm after applying the operator on it. Take, as an example, the unbounded linear operator

$$\hat{N} = \sum_{n=1}^{\infty} \eta(n) |n\rangle \langle n|$$

(70)

If we consider the normalized state $|\Phi\rangle = \frac{1}{\sqrt{\pi}} \sum_{n=1}^{\infty} \frac{1}{n} |n\rangle$ then we find $\langle \hat{N} |\Phi\rangle |\Phi\rangle \rightarrow \infty$. Thus, the action of the operator $\hat{N}$ on $|\Phi\rangle$ is ill-defined and we need to exclude such elements when we work with $\hat{N}$. The set of states for which the operator is defined is called its domain, and the domain is decisive for the properties of the operator.\textsuperscript{72,73,87} Thus, an operator is actually not just the rule, i.e., how it maps states to other states, such as in eq 70, but it also contains its domain in its definition.\textsuperscript{72,73} For certain operators we can instead, of explicitly stating the domain, encode the domain in an alternative representation: its diagonal representation. Indeed, eq 70 is the diagonal representation of the number operator and hence it implicitly defines its unique maximal domain. The most important of such operators in quantum physics, which allow a diagonal representation, are self-adjoint operators. In contrast to finite-dimensional problems, self-adjointness is not the same as symmetric (we will give some relevant examples later), but it also is a statement about the properties of its domain.\textsuperscript{72,73} Yet in the diagonal representation, self-adjointness becomes conveniently equivalent to the fact that the (generalized) eigenvalues of an operator are real.\textsuperscript{72,73}

A.2. Mathematical Realization

After the general mathematical setting, we next need to find an explicit realization of an ab initio quantum theory. This is done by first choosing the physical space $\Omega$ in which the basic entities, i.e., particles or fields, should be described. The physical space should be compatible with the basic symmetries of the theory that we want to turn into a quantum theory (compare with Section 2.1). The basic symmetries of the physical space can then be turned with the help of the Stone-Von Neumann theorem\textsuperscript{73} into uniquely defined self-adjoint operators (generators of these symmetries). For instance, in the case of the physical space $\Omega$ being a flat torus, i.e., a finite volume $L^3 \subset \mathbb{R}^3$ with periodic boundary conditions, translations are a basic symmetry and the (componentwise) self-adjoint operator is then (up to $\hbar$) the momentum operator

$$-iV_{\text{per}} = \sum_{n \in \mathbb{Z}^3} \frac{2\pi}{L} n |n, \text{per}\rangle \langle n, \text{per}|$$

(71)

with \((r |n, \text{per}) = \frac{1}{\sqrt{V}} \exp \left( \frac{2\pi}{L} n \cdot r \right). It is important to note here that while in quantum physics we commonly just write \(-iV\), when we consider this derivative as a self-adjoint operator, we implicitly mean its diagonal representation. Hence without stating either domain or the diagonal representation, the symbol \(-iV\) remains ambiguous. This can have severe consequences. As an example we change our physical space, and assume $L^3 \subset \mathbb{R}^3$ but use zero boundary conditions, i.e., we consider a hard-wall box. While a derivative \(-iV\) is of course defined on that space, there exists no self-adjoint realization of the momentum operator, and thus the concept of momentum is ill-defined with zero boundary conditions.

This is quite easy to understand since a box with hard walls does not have translations as a basic symmetry operation. A further problem appears once we choose a physical space that itself does not have a positive definite metric, i.e., that the only vector of length zero is the zero vector. This is, for instance, the case for the Minkowski space. Indeed, the space of square-integrable functions on the Minkowski space is only a pre-Hilbert space and hence concepts like self-adjointness are no longer easy to define. The standard way to avoid these mathematical issues is to restrict to appropriate subspaces. Hence in Section 2.1 the choice of subspace restriction dictates which quantities can be chosen as self-adjoint operators. While different choices are possible, e.g., the light-cone,\textsuperscript{72,74} in this work we single out a unique time frame and restrict the quantization to the spacial subspace, which has a positive definite metric. This choice goes hand in hand with a Hamiltonian description (energy and time are physically conjugate objects), the Coulomb gauge (since it relies also on a global time) and hence time emerges as a parameter, similar to non-relativistic quantum mechanics. From the perspective of ab initio QED this is the most convenient choice, since we want to have compatibility with ab initio non-relativistic quantum mechanics. We further note that commonly, as also followed in this review, one considers $\Omega = \mathbb{R}^3$. The reason is that for this case the basic symmetry operators have one and only one self-adjoint realization (which also makes the representation of the commutation relations unique) and there are no mathematical ambiguities that can arise by not stating domains.\textsuperscript{72,73}

A.3. Mathematical Subtleties

Self-adjoint operators, which correspond to basic symmetries, are then the building blocks of a specific realization of a quantum theory. As is easy to see from the diagonal representation, the momentum operator of eq 71 is unbounded from above and below, i.e., it does not have a smallest or highest eigenvalue. This complicates things, since for unbounded operators addition and multiplication are not universally defined, i.e., there is no algebra for unbounded operators. A simple example, which becomes relevant for consistency conditions between light and matter, is the case of a box with hard-wall boundary conditions. In contrast to the momentum operator, the Laplacian has a self-adjoint realization of the form

$$-V_{\text{zero}}^2 = \sum_{n \in \mathbb{Z}^3} \left( \frac{2\pi}{L} \right)^2 n^2 |n, \text{zero}\rangle \langle n, \text{zero}|$$

(72)

where $\langle n |n, \text{zero}\rangle = \frac{1}{\sqrt{L^3}} \prod_{k=1}^{3} \sin \left( \frac{2\pi}{L} n_k \right)$. For periodic boundary conditions we find, however, the consistency that \(-V^2\) is just applying eq 71 twice in diagonal representation and thus we have

$$-V_{\text{per}}^2 = \sum_{n \in \mathbb{Z}^3} \left( \frac{2\pi}{L} \right)^2 n^2 |n, \text{per}\rangle \langle n, \text{per}|$$

(73)

The issue of a missing algebra is also the reason why no Taylor-expansion exists for unbounded operators. That is, if we consider a multiplications operator on $\mathbb{R}^3$ of the form $v(r)$, even if the function $v(r)$ might be well approximated by a Taylor expansion, for the operator this is wrong. Implications of this fact for ab initio QED are discussed towards the end of Section 3.3.1 and in Appendix B. Note further, that the self-adjointness of a specific Hamiltonian is important to have
unique solvability of the corresponding time-dependent Schrödinger equation. Using the Stone-Von Neumann theorem, the Hamiltonian then becomes the generator of an evolution (semi-) group\textsuperscript{3,118,295–297} and hence we have a unique solution for any initial state (where the type of solution, however, depends on the properties of the initial state\textsuperscript{118,295–297}). For explicitly time-dependent Hamiltonians the more general concept of evolution systems is employed\textsuperscript{116,295–297}.

Let us also highlight that the mathematical subtleties become even more once we employ field operators to setup a quantum theory. Field operators, despite their name, are not operators in the sense discussed above. Indeed, \(\hat{a}(k, \lambda)\) and \(\hat{a}^\dagger (k, \lambda)\) are not each others adjoint, and \(\hat{a}^\dagger (k, \lambda)\) is rather an operator-valued distribution.\textsuperscript{27} As the name suggests, one can turn the operator-valued distributions into unbounded operators in the above sense, by integrating over a square-integrable form function \(\phi_p(k)\). The simplest version is a step function that becomes zero after a certain wave number cutoff. This is also discussed in Section 3.2. This mathematical necessity for any well-defined ab initio quantum field theory is the origin of regularization and renormalization procedures. And also for field operators, or their regularized unbounded operator versions, addition and multiplication is not universally defined.

A.4. Operator and Boundary-Value Consistency

As we have seen, the operators \(\mathbf{−i∇}\) and \(−\mathbf{∇}^2\) depend on the chosen boundary conditions, which are encoded in the diagonal representation by the behaviour of the corresponding eigenstates at the boundaries. It is important to note that quantum physics is mathematically designed in such a way that this behaviour, i.e., the domain of momentum or kinetic-energy operators, are kept when we add external fields and interactions.\textsuperscript{72,73} Indeed, it can be shown that any eigenstate in quantum mechanics is non-zero almost everywhere on the physical space \(Ω^{297,298}\) which also explains that disregarding the dipole self-energy operator is equivalent to assuming classical, perfectly localized distributional eigenstates of this unbounded operator. That is, any eigenstate knows about its boundary conditions, which is consistent with the fact that ab initio quantum physics is equivalent to solving high-dimensional linear partial-differential equations with chosen boundary conditions.\textsuperscript{299} This is, of course, also true for the Maxwell equation and its subsequent quantization. As is often done in practice, one can quantize the light field and the matter system independently. Yet, due to being fundamentally defined on the same physical space \(Ω\), based on the same symmetries (see Section 2.1 for details of QED), and also due to the coupling scheme based on the gauge principle (see Section 2.3), there is a high consistency between the boundary conditions of light and matter. If we enforce the local gauge principle, only phases of the matter wave functions are allowed that are consistent with the realizations of \(−\mathbf{i∇}\) or \(−\mathbf{∇}^2\). This in turn determines the gauge freedom of the vector potentials of the Maxwell field, and hence the boundary conditions for the light field. On the other hand, the eigenmodes of the Maxwell field are due to the diagonal representation of the vector Laplacian, which in contrast to the scalar Laplacian from, e.g., eq 72, maps from vector-valued square-integrable wave functions to vector-valued square-integrable wave functions (see eq 10). For the case of periodic boundary conditions on matter, the corresponding vector Laplacian of the Maxwell field is then given as

\[
−\nabla^2_{\text{vec,per}} = \sum_{\lambda=1}^{3} \sum_{n \epsilon \mathbb{Z}} \left( \frac{2\pi}{L} \right)^2 n^2 |\lambda, n\rangle \langle n, \lambda|,
\]

(74)

where \(\langle r|\lambda, n\rangle = e(n, \lambda)(r|n\rangle, \text{per})\) and \(e(n, \lambda)\) are the three orthonormal polarization vectors of \(n\).\textsuperscript{11} Thus, the diagonal representations of the vector Laplacian of the Maxwell field (and also the divergence and curl or, equivalently, the Helmholtz decomposition for periodic vector fields, see discussion around eq 10), the scalar Laplacian and the momentum operator of matter are all built upon the same scalar (matter) eigenfunctions \(\langle r|n\rangle, \text{per}\). If we now choose different boundary conditions for the matter system and the light system, these consistencies can be broken. This is alluded to in the example of Section 3.3, where the basic zero-energy mode of the Maxwell field due to the gauge freedom of the periodic matter system is a constant function \(2\pi/L\), yet for perfect-conductor boundary conditions such a field is ill-defined. In more detail, the perfect-conductor boundary conditions for the Maxwell field implies that the self-adjoint vector Laplacian becomes

\[
−\nabla^2_{\text{vec,per}} = \sum_{n \epsilon \mathbb{Z}} \left( \frac{2\pi}{L} \right)^2 n^2 |\lambda, n\rangle \langle n, \lambda|,
\]

(75)

where the scalar zero-boundary eigenfunctions are \(\langle x|\lambda, n\rangle = \sqrt{\frac{L}{\pi}} \sin \left( \frac{\pi n x}{L} \right)\), the zero-derivative boundary eigenfunctions are \(\langle x|\lambda, 0\rangle = \sqrt{\frac{L}{\pi}} \cos \left( \frac{\pi n x}{L} \right)\) and accordingly for \(y \leftrightarrow n_x\) and \(z \leftrightarrow n_y\). To show that a constant \(2\pi/L\) mode is not in the domain of the vector Laplacian of eq 75 and hence ill-defined for the corresponding Maxwell equation, we first consider the expansion of a constant function \(1/\sqrt{L}\) in zero-boundary eigenfunctions, i.e.,

\[
\langle 1/\sqrt{L}|\lambda, n\rangle = \frac{\sqrt{\pi}}{\sqrt{n}} \left( 1 - (-1)^n \right).\textsuperscript{118}
\]

Applying the Laplacian, e.g., the second derivative with respect to \(y\), leads to an ill-defined expression, i.e.,

\[
\left( \frac{1}{\sqrt{L}} \right) - \frac{\partial^2}{\partial y^2} \left( \frac{1}{\sqrt{L}} \right) = \frac{2}{L^2} \sum_{n \epsilon \mathbb{Z}} (1 - (-1)^n)^2 \rightarrow \infty.
\]

(76)

Consequently, the constant gauge field that arises from the periodic matter subsystem (see also Section 2.3), e.g., \(A(r) \propto \epsilon_0\), corresponds to infinite energies for the Maxwell equation with perfect-conductor boundary conditions. In more detail, using the expansion of the constant function in zero-derivative
boundary conditions for $x$ (only $n_2 = 0$ contributes), and the zero boundary conditions for $y$ and $z$ we find
\[
\epsilon_i = \sum_{n_2,n_3} \frac{2\sqrt{L}}{\pi n_2 n_3} (1 - (-1)^{n_2})(1 - (-1)^{n_3})
\]
\[
\begin{pmatrix}
\langle y|p_{x,2}, \text{zero} \rangle \langle z|p_{x,3}, \text{zero} \rangle \\
0 \\
0
\end{pmatrix}
\]
(77)

The corresponding magnetic field is then proportional to
\[
B(r) \propto \sum_{n_2,n_3} \frac{2\sqrt{L}}{\pi n_2 n_3} (1 - (-1)^{n_2})(1 - (-1)^{n_3})
\]
\[
\begin{pmatrix}
0 \\
\langle y|p_{x,2}, \text{zero} \rangle \frac{\pi n_3}{L} \langle z|p_{x,3}, \text{zero} \rangle \\
-\frac{\pi n_2}{L} \langle y|p_{x,2}, \text{zero} \rangle \langle z|p_{x,3}, \text{zero} \rangle
\end{pmatrix}
\]
(78)

Thus, we then have due to eq 7 that
\[
E_{ph} \propto \int B^2(r) \, dr
\]
(79)
\[
\alpha \sum_{n_2,n_3} \frac{(n_2^2 + n_3^2)}{n_2 n_3} (1 - (-1)^{n_2})(1 - (-1)^{n_3}) \to \infty
\]
(80)

A further problem can be the violation of energy and momentum conservation in such a case. Take, for instance, an allowed single-particle matter wave function of the form
\[
\langle r|\Psi_r \rangle = \frac{1}{\sqrt{4\pi}} \exp\left(-i\frac{2\pi}{L} x\right),
\]
which has a momentum proportional to $k_x = \frac{2\pi}{L} \epsilon_i$. If we have as initial state $|k_x, 0\rangle = |k_x, 0\rangle \otimes |0\rangle$ $\otimes |0\rangle \otimes \ldots$, i.e., we have zero photons in all modes, then the linear coupling $A_{\text{pol}}(r) \cdot (-i\mathbf{V}_{\text{pol}})$ generates photons of arbitrary high energy and momentum without any loss of energy or momentum from the matter subsystem. Let us focus on the transverse modes with $n_1 = 0$, i.e., $n_2$ and $n_3$ are non-zero such that only the $x$-component of the vector-potential is non-zero, as also discussed above. If we then consider transition-matrix elements of the form
\[
\langle k_x, 1_{n_2,n_3} | A_{\text{pol}}(r) \cdot (-i\mathbf{V}_{\text{pol}}) | k_x, 0 \rangle
\]
(81)
where $1_{n_2,n_3}$ indicates that we have one photon in the mode $(0, n_2, n_3)$ and zero else, then for any such mode we have
\[
\langle k_x, 1_{n_2,n_3} | A_{\text{pol}}(r) \cdot (-i\mathbf{V}_{\text{pol}}) | k_x, 0 \rangle
\]
\[
\propto \langle k_x, 1_{n_2,n_3} | \hat{A}_{\text{pol}}(r) \cdot (-i\mathbf{V}_{\text{pol}}) | k_x, 0 \rangle
\]
\[
\propto \frac{2\pi}{L} \frac{2}{n_2 n_3} (1 - (-1)^{n_2})(1 - (-1)^{n_3})
\]
(82)

Contrast this with $\hat{A}_{\text{per}}(r) \cdot (-i\mathbf{V}_{\text{per}})$, where only those transition-matrix elements are non-zero for which energy and momentum is conserved for the total physical system. On the other hand, we have also consequences for the matter subsystem. Since, for instance, fields that obey perfect-conductor boundary conditions are not necessarily compatible with periodic boundary conditions, e.g., due to $\langle x|p_{x,3}, \text{zero} \rangle$, these fields lie outside the domain of the momentum operator of eq 71. One could ignore this issue and merely consider the consistency with the domain of the scalar Laplacian, as commonly done to establish self-adjointness of the Hamiltonian.22,73,118,297 In this case, however, the regularity of eigenstates will become too low to be able to fulfill local charge conservation and similar local conservation laws.118,297 This is quite intuitive for the diamagnetic current density, which is proportional to $\rho(r)A(r)$,119,204 such that this current contribution can inherit the violation of the corresponding boundary conditions of the matter system.

### A.5. Ab Initio and Alternative Approaches

Let us next give our working definitions for what we call ab initio methods. We will do so first in an abstract manner based on the proceeding discussions and then provide details on why such a formal definition is reasonable from a natural-science point of view.

We abstractly identify a theoretical approach as ab initio, if the approach is formulated on infinite-dimensional Hilbert spaces for all subsystems and that tests its (in practice) finite-dimensional approximate results with respect to the infinite basis set limit. That is, the stability of the obtained results are tested with respect to including more and more states and with respect to changing the basis set. Alternatively, we can state that an ab initio approach solves partial differential equations for given boundary conditions using some form of a Galerkin method.299 The later definition is most easily connected with the common statement that ab initio methods only rely on fundamental constants of nature, such as in the case of the Schrödinger equation or of the Pauli-Fierz Hamiltonian of eq 37. Using the connection between the basic symmetries of space-time and self-adjoint operators (see Section A.2) we only need a few fundamental constants to make predictions. In practice, though, such as in the case of the dipole-approximated Pauli-Fierz Hamiltonian of eq 39, we often use further information from experiment or other theories, e.g., the mode structure of a given cavity.

Clearly, the above definition is not precise in practice. There are many models or alternative theoretical methods that employ specific infinite-dimensional spaces, such as coupled quantum harmonic oscillators. Yet for most models and alternative theoretical methods the previously discussed issues (see Appendices A.1—A.4) related to infinite dimensions and unbounded operators are not considered nor encountered. In contrast, in ab initio theories these issues are omnipresent. As an example we highlight that in a finite-dimensional problem all symmetric linear operators are matrices and thus always have eigenstates with real eigenvalue. Yet for infinite-dimensional problems (as also discussed in Appendix A.1) this is not the case, and in an ab initio approach one has to test how the eigenstates change upon changing the basis set. In this way one finds that in the case of non-self-adjointness real eigenvalues might cease to exist in the basis set limit, like the ground state in the atomic Dirac equation for nuclear charge $Z > 136$, or that complex eigenvalues arise such as in the case of the atomic Dirac equation for nuclear charge $Z > 118$.300 In ab initio QED, for instance, such investigations allow to identify the onset of the dissociation continuum.295 A further point of difference is the accuracy/applicability of perturbation theory. Since for finite dimensions we always have a largest eigenvalue, we can safely assume that perturbation theory captures the resulting changes accurately if we add a further linear operator.
(matrix) with a small pre-factor $c \ll 1$. Yet for infinite dimensions, where we can add unbounded operators that have arbitrarily large (generalized) eigenvalues, this is not necessarily the case. Examples within ab initio QED are discussed in Secs. 3.3 and Appendix B below. Perturbation theory with unbounded operators does usually not converge and hence often only gives reasonable results for low orders. For this reason we focus in this work on nonperturbative solutions to infinite-dimensional problems.

Up until now we have discussed a definition based on mathematical considerations, which are only of concern to theoretical chemists and physicists. But are there more general scientific reasons why we differentiate between ab initio methods and alternative theoretical approaches? Indeed, from the mathematical perspective a first obvious difference is that ab initio methods are designed to test the predictability/accuracy of a realization of quantum physics, as discussed in Appendix A.2. This is also alluded to in the main text, e.g., at the end of Section 3.1. Therefore, one might say that ab initio methods are more general in their applicability than alternative approaches that model specific setups or observables. Yet, even if we disregard the problem of accuracy of ab initio methods in practice, the main drawback of ab initio methods is that they do not directly provide any intuitive insight. They are, for all intents and purposes, numerical experiments. To provide insights and an understanding of a chemical/physical effect it is in practice necessary to develop specific models and alternative theoretical approaches. This distinction is also nicely reflected in the demand of quantitative results for ab initio methods, e.g., “chemical accuracy” in quantum chemistry, while for models qualitative results are the main goal.

For the authors of this review it goes without saying that different theoretical approaches are needed since they supplement each other. Because their purposes are complementary, it does not make sense to consider one or the other to be scientifically superior. If we attain in polaritonic chemistry a similar level of agreement between ab initio methods and alternative approaches, as is the case in quantum chemistry or solid-state physics, we will at the same time have attained a detailed understanding of the basic mechanisms that make this field so exciting.

Let us finally also clarify that, of course, also a realization of quantum physics as defined in Section A.2 is just a model of reality. It is merely an infinite-dimensional model and hence more flexible. On the long run it should be our goal as scientists to refine this types of models to, for instance, get a better connection with general relativity. Thus, irrespective of whether one works with ab initio or alternative approaches, fundamentally we should all keep in mind that

“All models are wrong but some are useful”

G. Box in ref 301.

APPENDIX B: NECESSITY OF LONGITUDINAL DIPOLE SELF-ENERGY TERM

Let us for simplicity and without restriction of generality treat the nuclei/ions clamped and consider (as commonly done) only the electronic interaction between a plasmonic and a molecular system of interest. In this case the longitudinal modes due to a plasmon-molecule interaction are all mediated in Coulomb gauge via

$$\hat{W}_{ee} = \frac{1}{2} \sum_{l \neq m} \frac{e^2}{4\pi \epsilon_0 |r_l - r_m|}$$

as can immediately be inferred from the Pauli-Fierz Hamiltonian of eq 37. It is obvious that for all possible square-integrable wave functions $\Psi$ in the domain of the operator (see Appendix A.1 for further details), irrespective of statistics and indistinguishability, we have

$$\langle \Psi | \hat{W}_{ee} | \Psi \rangle > 0$$

That is, $\hat{W}_{ee}$ is a positive operator. Assume now that we choose instead of $\hat{W}_{ee}$ an operator of the form

$$\hat{W}_{\text{dip}} = \frac{1}{2} \sum_{l \neq m} \sum_{\mu = 1}^{3} g_{\mu}^l g_{\mu}^m$$

where $r_{\mu}^l$ are the different Euclidean coordinates of particle $m$ and $g_{\mu}^m$ are some arbitrary real constants such that we even allow to break the symmetry of the original Coulomb operator. Obviously this operator is not positive and it can be made arbitrarily negative, i.e., $\hat{W}_{\text{dip}}$ is not bounded from below and we can find $\Psi$ such that

$$\langle \Psi | \hat{W}_{\text{dip}} | \Psi \rangle \rightarrow -\infty$$

Not surprisingly, with similar arguments as for the transverse case, we find that an Hamiltonian with purely dipolar interaction has no eigenstates and hence does not have an equilibrium solution.

The reason for this unphysical behavior within ab initio quantum physics is quite obvious since we have broken the very basic condition that the longitudinal electromagnetic energy is positive. As discussed in Section 3.3, we have two options that amount to the same physics: We either keep also quadratic contributions of the form $\sum_{\mu = 1}^{3} (g_{\mu} r_{\mu}^m)^2$ that counter the purely linear coupling, or we restrict to a finite area (finite simulation box with chosen boundary conditions). In both cases the approximate interaction becomes bounded from below and can be made manifestly positive by a finite energy shift. We have thus shown, in yet a different way, that quadratic/counter terms are necessary to have a stable ab initio quantum theory and claims to the contrary in the literature arise from a basic misunderstanding of the difference between perturbative/few-level calculations and solving a (necessarily infinite-dimensional) Schrödinger-type quantum equation. Let us finally note that in the long-wavelength or dipole limit we can, strictly speaking, no longer distinguish between transverse and longitudinal modes. Only the physical context of the dipolar approximation provides us with the information which type of field we consider.

APPENDIX C: QEDFT MAPPINGS AND THE EFFECTIVE FIELDS

Let us first note that QEDFT and its basic mapping theorems are distinct from using (time-dependent) density functional theory and approximately taking into account the interaction with a quantized light field. Indeed, QEDFT is the exact reformulation of the Pauli-Fierz quantum field theory in terms of current densities and vector potentials, in analogy to electronic density functional theory being the exact reformulation of the static Schrödinger theory with scalar external potentials. Further we note that the term "QEDFT" is
used synonymously for the ground-state, the real-time or linear-response time-dependent, minimal-coupling or dipole-coupling density-functional reformulation of the (generalized) Pauli-Fierz field theory.\textsuperscript{81,94,145,183,184,186} The context unambiguously identifies which specific realization/implementation of QEDFT is meant/used. We use the same convention when referring to “density functional theory”\textsuperscript{118,169}

To highlight these details let us compare the basic mappings of density functional theory and QEDFT. For simplicity we assume clamped nuclei to connect to the standard case of density functional theories and avoid issues due to localization finite systems in free space.\textsuperscript{303,304} In this case, we have either due to the Hohenberg–Kohn,\textsuperscript{164} the Runge-Gross\textsuperscript{166} (where we for simplicity assume the ground state as initial state\textsuperscript{305}), or van Leeuwen Laplace-transform\textsuperscript{306} theorems (note the notation convention from eq 64):

\[
\nu(r,t) \leftrightarrow \Psi(r,t) \leftrightarrow n(r,t)
\]  

(87)

For the static case the parameter \( t \) (time) is redundant. This mapping shows that instead of the electronic wave function \( \Psi(r,t) \) we can express everything in terms of \( n(r,t) \), the particle-number density, since we can perform a functional-variable transformation within electronic quantum mechanics of the form\textsuperscript{118,164,166}

\[
\langle \Psi(t)|\hat{O}|\Psi(t)\rangle = \langle \Psi([n],t)|\hat{O}|\Psi([n],t)\rangle = O([n],t)
\]  

(88)

for any observable \( \hat{O} \). Here the notation \( O[n] \) means that the object \( O \) is uniquely determined by \( n \). We have thus replaced the usual quadratic-form structure of quantum physics in terms of wave functions by \textit{exact} nonlinear functionals in terms of the density. But this reformulation has so far no practical relevance, since we do not know how to determine from a given \( v(r,t) \) the corresponding \( n(r,t) \) without going through the wave function \( \Psi(r,t) \) first. In the ground-state case the obvious way would be to reformulate the minimization over all wave functions in terms of densities.\textsuperscript{164,165} Yet it is very hard to express the various contributions in terms of the density only. Alternatively one can consider the local ground-state force equilibrium.\textsuperscript{118,168} In practice (also for the time-dependent case) one tries to approximate the mapping \( \nu(r,t) \leftrightarrow n(r,t) \) for interacting electrons with the help of an auxiliary mapping that is physically close yet numerically tractable. The standard choice is to consider the mapping of noninteracting electrons

\[
\nu_0(r) \leftrightarrow \Phi(r) \leftrightarrow n_0(r)
\]  

(89)

where \( \Phi(r) \) is then (usually) a Slater determinant of single-particle orbitals \( \psi_\sigma(r,\sigma) \), \( n_0(r) = \sum_\sigma \sum_\epsilon |\psi_\sigma(r,\sigma)|^2 \) and the subindex “s” refers to “single particle” indicating that the noninteracting many-body Schrödinger equation can be recast as single-particle Schrödinger equations.\textsuperscript{118} Assuming now that interacting and noninteracting systems generate the same set of densities, we can thus combine both maps and find\textsuperscript{118}

\[
\nu(r) \leftrightarrow \nu_0(r)
\]  

(90)

Thus, we have mapped the interacting problem to a noninteracting problem that generates the same density. This new effective potential is then called the Kohn–Sham potential and is denoted as \( v_{KS}(r) = v_0(U) \). Still we did not gain anything, because the mapping from interacting to non-interacting potentials is even harder to approximate. The final step is to once again use the bijectivity between densities and potentials to re-express the Kohn–Sham potential as\textsuperscript{118}

\[
v_{KS}(\{\nu\}, r) = v(r) + v_{KS}(\{\nu\}, r) - v(r, t)
\]  

\[
= v(r) + v[[n], r,t] - v([n], r,t)
\]  

(91)

This turns the linear single-particle Schrödinger equations in terms of \( v_{KS}(\{\nu\}, r) \) into the well-known nonlinear Kohn–Sham single-particle equations in terms of the Hartree-exchange-correlation potential \( v_{xc}(\{n\}, r,t) \). To approximate the difference between interacting and noninteracting maps in terms of densities, many different highly successful strategies exist.

In the case of QEDFT we have instead of the Schrödinger Hamiltonian the Pauli-Fierz Hamiltonian that does not only depend on the electronic degrees but also on the (continuum of) photonic degrees of freedom. Since there are some subtle differences in the minimal-coupling QEDFT case between the ground-state and the time-dependent case with respect to the suitable functional variables (although these differences allow to cure old issues of ground-state current-density functional theory),\textsuperscript{145} we in the following restrict for simplicity to the dipole-coupled Pauli-Fierz Hamiltonian of eq 39 with the external fields of the form of eqs 54 and 53. In this case, we have with the identification of \( v(r,t) = -|j|\phi_u(r,t) \), the interacting mapping in analogy to the density-functional case\textsuperscript{93,124,145,184}

\[
v(r,t), j(t) \leftrightarrow \Phi(r, q, t) \leftrightarrow (n(r,t), q(t))
\]  

(92)

where \( j(t) \) and \( q(t) \) indicates that we have \( M_p \)-long vectors of mode-resolved external currents and displacement coordinates, respectively. The reason why we have now a pair of functional variables \( (n(r,t), q(t)) \) is that we can change not only \( v(r,t) \) to consider different physical situations but also adapt \( j(t) \) to influence the full system of light and matter. Obviously, even if the electronic Schrödinger equation and the dipole-approximated Pauli-Fierz Hamiltonian have the same external fields, the expectation values of the same operators give different answers in general and we have access in QEDFT to all photonic observables since we have re-expressed everything in terms of

\[
\langle \Psi(t)|\hat{O}|\Psi(t)\rangle = \langle \Psi([n,q],t)|\hat{O}|\Psi([n,q],t)\rangle = O([n,q],t)
\]  

(93)

In analogy, we then find adapted effective fields

\[
v(r,t), j(t) \leftrightarrow (\nu(r,t), j_s(t))
\]  

(94)

if we choose as auxiliary system noninteracting electrons and photons to generate the same density and displacement field.\textsuperscript{94,124,145} Accordingly, we find with the definition of the Maxwell-Kohn–Sham fields \( v_{MKS}(\{\nu_j\}, r,t) = v_j(\{\nu_j\}, r,t) \) and \( j_{MKS}(\{\nu_j\}, t) = j_j(\{\nu_j\}, t) \) that we have

\[
v_{MKS}(\{\nu_j\}, r,t) = v(\{\nu_j\}, r,t) + v_{MKS}(\{\nu_q\}, r,t)
\]  

(95)

\[
j_{MKS}(\{\nu_j\}, t) = j(\{\nu_j\}, t) + j_{MKS}(\{\nu_q\}, t)
\]  

(96)

Here we have denoted the nonlinear terms as mean-field-exchange-correlation (Mxc) fields in order to highlight that besides the Hartree term there are now also mean-field
contributions that can be made explicit in the effective fields. Similarly to the electronic density-functional case there are various ways of approximating the difference of these mappings\(^{170,172,256,307}\) (note that in the dipole case the Mxc current is exactly the mean-field current, i.e., \(j_{\text{Mxc}}(n, q, r) = \frac{\partial H_{\text{xc}}}{\partial n}(n, q, r)\)), and we can distinguish the contribution due to the Coulomb interaction and due to the direct electron-photon interaction as

\[
v_{\text{Mxc}}(n, q, r) = v_{\text{Hxc}}(n, q, r) + v_{\text{pc}}(n, q, r) \quad (97)
\]

While in principle \(v_{\text{Hxc}}(n, q, r) \neq v_{\text{Hxc}}(n, r)\), in practice one usually employs approximations to \(v_{\text{Hxc}}(n, r)\) from electronic density functional theories also in QEDFT. This raises the question of consistency between approximation to the longitudinal (Coulombic) and transverse (photonic) interactions. That is, optimally both functionals are approximated in the same way. Examples of this are the use of optimized effective potentials in the exchange approximation\(^{172,258}\) or the use of local exact exchange approximations\(^{170}\) for Coulomb and photon interactions. Similar consistency considerations arise also in the context of linear-response time-dependent density functional functional theory, where one can potentially use different functionals for the Coulomb interaction in the ground state calculation and for the Coulomb interaction in the Hartree-exchange-correlation kernel.\(^{166}\)

We finally note that one is not restricted to using noninteracting electrons and photons as auxiliary system. Similar to electronic density functional theory, where nonlocal (generalized) Kohn–Sham systems\(^{308–310}\) or strictly correlated electrons\(^{311–313}\) are sometimes used, in QEDFT one can, for instance, use polaritonic orbitals as reference.\(^{147,176}\) The main drawback of using polaritonic orbitals is that we have mixed statistics that need to be handled with extra care to not generate unphysical results.\(^{147,174,176}\)

### AUTHOR INFORMATION

**Corresponding Authors**

- **Michael Ruggenthaler** – Max-Planck-Institut für Struktur und Dynamik der Materie, 22761 Hamburg, Germany; The Hamburg Center for Ultrafast Imaging, 22761 Hamburg, Germany; https://orcid.org/0000-0002-0728-8372; Email: Michael.Ruggenthaler@mpsd.mpg.de
- **Dominik Sidler** – Max-Planck-Institut für Struktur und Dynamik der Materie, 22761 Hamburg, Germany, The Hamburg Center for Ultrafast Imaging, 22761 Hamburg, Germany; https://orcid.org/0000-0002-4732-5713; Email: Dominik.Sidler@mpsd.mpg.de
- **Angel Rubio** – Max-Planck-Institut für Struktur und Dynamik der Materie, 22761 Hamburg, Germany; The Hamburg Center for Ultrafast Imaging, 22761 Hamburg, Germany; Center for Computational Quantum Physics, Flatiron Institute, New York, New York 10010, United States; https://orcid.org/0000-0003-2060-3151; Email: Angel.Rubio@mpsd.mpg.de

Complete contact information is available at [https://pubs.acs.org/10.1021/acs.chemrev.2c00788](https://pubs.acs.org/10.1021/acs.chemrev.2c00788)

**Funding**

Open access funded by Max Planck Society.

### Notes

The authors declare no competing financial interest.

### Biographies

Michael Ruggenthaler is research group leader at the Max-Planck Institute for the Structure and Dynamics of Matter in Hamburg, Germany. He obtained his Ph.D. in theoretical physics jointly from the Max-Planck Institute for Nuclear Physics and the University of Heidelberg, Germany, in 2009. As Schrödinger Fellow of the Austrian Science Fund, he worked at the University of Jyväskylä, Finland, and then established an independent research group at the University of Innsbruck, Austria, before he moved to Hamburg in 2016. His research group works on the theoretical and mathematical foundations of quantum many-body theories, in and out of equilibrium. His current research focus is on first-principles approaches to ab initio quantum electrodynamics, polaritonic chemistry, and cavity-modified material properties.

Dominik Sidler is a postdoctoral researcher at the Max-Planck Institute for the Structure and Dynamics of Matter and the Cluster of Excellence CUI: Advanced Imaging of Matter in Hamburg, Germany. He obtained his Ph.D. in computational chemistry from ETH Zürich, Switzerland, in 2018. He works on the theoretical and computational description of (strong) light–matter interactions in polaritonic chemistry. His current research focus is on first-principles approaches to describe emergent local and collective strong coupling phenomena in optical cavities, as well as the impact of vacuum field fluctuations on molecular nonequilibrium conditions at finite temperatures.

Angel Rubio is the Director of the Theory Department of the Max-Planck Institute for the Structure and Dynamics of Matter in Hamburg, Germany, where his research focuses on the modeling and theory of electronic and structural properties of condensed matter. His department works on developing novel theoretical tools, such as quantum-electrodynamical density functional theory and computational codes for the ab initio description and control of quantum many-body systems, and on characterizing new nonequilibrium states of matter. His work has been recognized by several awards including the 2018 Max Born medal and prize, 2016 Medal of the Spanish Royal Physical Society, the 2014 Premio Rey Jaime I for basic research, the 2006 DuPont Prize in nanotechnology, the 2005 Friedrich Wilhelm Bessel Research Award of the Humboldt Foundation, and two European Research Council advanced grants. He is a fellow of the APS, EPS, and AAAS, and member of the Leopoldina Academy, BBAW, the European Academy of Sciences, the Academia Europaea, and a foreign associate member of the National Academy of Sciences (USA).

### ACKNOWLEDGMENTS

We acknowledge enlightening discussions with Tal Schwartz, Abraham Nitzan, Sharly Fleischer, Thomas W. Ebbesen, Christian Schäfer, Simone Latini, Johannes Flick, Enrico Ronca, Henrik Koch, Fabijan Pavosevic, and Markus Kowalewski. This work was made possible through the support of the 2018 Max Born medal and prize, 2016 Medal of the Spanish Royal Physical Society, the 2014 Premio Rey Jaime I for basic research, the 2006 DuPont Prize in nanotechnology, the 2005 Friedrich Wilhelm Bessel Research Award of the Humboldt Foundation, and two European Research Council advanced grants. He is a fellow of the APS, EPS, and AAAS, and member of the Leopoldina Academy, BBAW, the European Academy of Sciences, the Academia Europaea, and a foreign associate member of the National Academy of Sciences (USA).
REFERENCES


(11) Greiner, W.; Reinhardt, J. Field quantization; Springer Berlin Heidelberg, 1996.


(13) Greiner, W.; Reinhardt, J. Quantum electrodynamics; Taylor & Francis, 2003.


(22) Spohn, H. Dynamics of charged particles and their radiation field; Cambridge University Press, 2004.


(70) Scharf, G. Finite quantum electrodynamics: the causal approach; Courier Corporation, 2014.
(73) Blanchard, P.; Brüning, E. Mathematical methods in Physics: Distributions, Hilbert space operators, variational methods, and applications in quantum physics; Birkhäuser, 2015; Vol. 69.


(108) Schärer, C.; Ruggenthaler, M.; Rokaj, V.; Rubio, A. Relevance of the quadratic diamagnetic and self-polarization terms in cavity quantum electrodynamics. ACS photonics 2020, 7, 975−990.


Dipolar coupling of nanoparticle-molecule assemblies: an efficient quantitative sampling of atomic-scale electromagnetic waveforms.


hot electron transport drives nano-localized chemistry.

Busch, K.; Knorr, A.; Richter, M. Quantization of quasinormal modes, 51, 021301.


D modes on polariton relaxation.

Simulations of molecular polaritons: the effect of multiple cavity chemistry.


(256) Flick, J.; Schafer, C.; Ruggenthaler, M.; Appel, H.; Rubio, A. Ab inito optimized effective potentials for real molecules in optical cavities: photon contributions to the molecular ground state. ACS Photonics 2018, 5, 992–1005.


(295) Pazy, A. Semigroups of linear operators and applications to partial differential equations; Springer Science & Business Media, 2012; Vol. 44.

(296) Reed, M.; Simon, B. II: Fourier analysis, self-adjointness; Elsevier, 1975; Vol. 2.


