# 1   From Particles to Fields

SYNOPSIS To introduce some basic concepts of field theory, we begin by considering two simple model systems – a one-dimensional "caricature" of a solid and a freely propagating electromagnetic wave. As well as exemplifying the transition from discrete to continuous degrees of freedom, these examples introduce the basic formalism of classical and quantum field theory as well as the notions of elementary excitations, collective modes, symmetries and universality – concepts that will pervade the rest of the text.

One of the appealing facets of condensed matter physics is that phenomenology of remarkable complexity can emerge from a Hamiltonian of comparative simplicity. Indeed, microscopic "condensed matter Hamiltonians" of high generality can be constructed straightforwardly. For example, a prototypical metal or insulator may be described by the **many-particle Hamiltonian**, $H = H_\mathrm{e} + H_\mathrm{i} + H_\mathrm{ei}$, where

many-particle Hamiltonian

$$H_\mathrm{e} = \sum_i \frac{\mathbf{p}_i^2}{2m} + \sum_{i<j} V_\mathrm{ee}(\mathbf{r}_i - \mathbf{r}_j),$$

$$H_\mathrm{i} = \sum_I \frac{\mathbf{P}_I^2}{2M} + \sum_{I<J} V_\mathrm{ii}(\mathbf{R}_I - \mathbf{R}_J), \tag{1.1}$$

$$H_\mathrm{ei} = \sum_{iI} V_\mathrm{ei}(\mathbf{R}_I - \mathbf{r}_i).$$

Here, $\mathbf{r}_i$ ($\mathbf{R}_I$) denotes the coordinates of valence electrons (ion cores), while $H_\mathrm{e}, H_\mathrm{i}$, and H_ei describe the dynamics of electrons, ions and the interaction of electrons and ions, respectively (see the figure). Of course, the Hamiltonian (1.1) can be made more realistic by, for example, remembering that electrons and ions carry spin, adding potential disorder, or introducing host lattices with multi-atomic unit cells. However, for developing our present line of thought, the prototype Hamiltonian H will suffice.
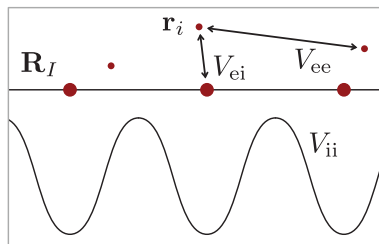
The fact that a seemingly innocuous Hamiltonian like Eq. (1.1) is capable of generating a plethora of phenomenology can be read in reverse: normally, one will not be able to make progress theoretically by approaching the problem in an

3

"**ab initio**" manner, i.e., by an approach that treats all microscopic constituents as equally relevant degrees of freedom. How, then, can successful analytical approaches be developed? The answer lies in several basic principles inherent in generic condensed matter systems.

**reduction
principle**

1. **Structural reducibility**: Not all components of the Hamiltonian (1.1) need to be treated simultaneously. For example, when our interest is in the vibrational motion of the ion lattice, the dynamics of the electron system can often be neglected or, at least, treated in an "effective" manner. Similarly, the dynamics of the electron system can often be considered independent of the ions, etc.

2. In the majority of condensed matter applications, one is interested not so much in the full profile of a given system but, rather, in its energetically low-lying dynamics. This is motivated partly by practical aspects (in daily life, iron is normally encountered at room temperature and not at its melting point), and partly by the tendency of large systems to behave in a "universal" man-
**universality
principle** ner at low temperatures. Here, **universality** implies that systems differing in microscopic detail (i.e., with different types of interaction potentials, ion species, etc.) exhibit common collective behavior at low energy or long length scales. As a physicist, one will normally seek for unifying principles in collective phenomena rather than to describe the peculiarities of individual elements or compounds. However, universality is equally important in the *practice* of condensed matter theory. In particular, it implies that, at low temperatures, system-specific details of microscopic interaction potentials are often of secondary importance, i.e., one may employ *simple* model Hamiltonians.

3. For most systems of interest, the number of degrees of freedom $N = \mathcal{O}(10^{23})$ is formidably large. However, contrary to first impressions, the magnitude of this figure is rather an advantage: in addressing condensed matter problems,
**statistical
principles** the **principles of statistics** imply that statistical errors tend to be negligibly small.[1]

symmetries 4. Finally, condensed matter systems typically possess intrinsic **symmetries**. For example, the Hamiltonian (1.1) is invariant under the simultaneous translation and/or rotation of all coordinates, which expresses the global Galilean invariance of the system (these are continuous symmetries). Invariance under spin rotation (continuous) or time reversal (discrete) are other examples of common symmetries. The general importance of symmetries cannot be overemphasized: symmetries support the conservation laws that simplify any problem. Yet, in

---

[1] The importance of this point is illustrated by the empirical observation that the most challenging systems in physical sciences are of *medium*, and not large, scale, e.g., metallic clusters, medium-sized nuclei or large atoms consisting of some $\mathcal{O}(10^1$–$10^2)$ fundamental constituents. Such systems are beyond the reach of few-body quantum mechanics while not yet accessible to reliable statistical modeling. The only viable path to approaching systems of this type is often through numerical simulation or the use of phenomenology.

condensed matter, symmetries are even more important. A conserved observable is generally tied to an energetically low-lying excitation. In the universal, low-temperature, regime in which we will typically be interested, it is precisely the dynamics of these excitations that govern the gross behavior of the system. Generally, the identification of fundamental symmetries is the first step in the sequence "symmetry $\mapsto$ conservation law $\mapsto$ low-lying excitations" and one that we will encounter time and again.

To understand how these basic principles can be used to formulate and explore effective low-energy field theories of solid state systems, we will begin by focusing on the **harmonic chain**, a collection of atoms bound by a harmonic potential. In doing so, we will observe that the universal characteristics encapsulated by the low-energy dynamics[2] of large systems relate naturally to concepts of **field theory**.

## 1.1 Classical Harmonic Chain: Phonons

[Classical Harmonic Chain: Phonons]

Returning to the prototype Hamiltonian (1.1), let us focus on the dynamical properties of the positively charged *core ions* that constitute the host lattice of a crystal. For the moment, we will neglect the fact that atoms are quantum objects and treat the ions as *classical* entities. To further simplify the problem, let us consider a one-dimensional atomic chain rather than a generic d-dimensional solid. In this case, the positions of the ions can be specified by a sequence of coordinates with average lattice spacing a. Relying on the structural reducibility principle 1, we will first argue that, to understand the behavior of the ions, the dynamics of the conduction electrons are of secondary importance, i.e., we will set $H_\mathrm{e} = H_\mathrm{ei} = 0$.

At zero temperature, the system freezes into a regularly spaced array of ion cores at coordinates $R_I = \bar{R}_I \equiv Ia$. Any deviation from a perfectly regular configuration incurs a potential energy cost. For low enough temperatures (principle 2), this energy will be

approximately quadratic in the small deviation of the ion from its equilibrium position. The "reduced" low-energy Hamiltonian of the system then reads

$$H = \sum_{I=1}^{N} \left[\frac{P_I^2}{2m} + \frac{k_\mathrm{s}}{2}(R_{I+1} - R_I - a)^2 \right], \tag{1.2}$$
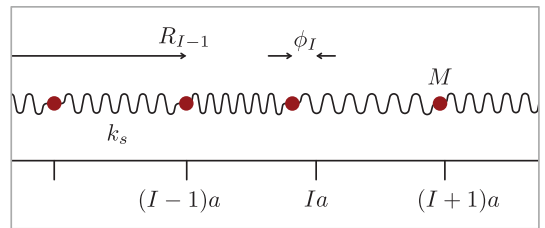
[2] In this text, we will focus on the *dynamical* behavior of large systems, as opposed to their *static* structural properties. In particular, we will not address questions related to the formation of definite crystallographic structures in solid state systems.

where the coefficient k_s determines the steepness of the lattice potential. Note that H can be interpreted as the Hamiltonian of N point-like particles of mass m connected by elastic springs with spring constant k_s (see the figure).

1.1.1 Lagrangian formulation and equations of motion

classical
harmonic
chain

What are the elementary low-energy excitations of the **classical harmonic chain**? To answer this question we might, in principle, attempt to solve Hamilton's equations of motion. Indeed, since H is quadratic in all coordinates, such a program is feasible. However, few of the problems encountered in solid state physics enjoy this property. Further, it seems unlikely that the low-energy dynamics of a macroscopi-

Joseph-Louis Lagrange
1736–1813
was a French mathematician and astronomer (though born in Turin) who excelled in all fields of analysis, number theory and celestial mechanics. In 1788, he published *Mécanique Analytique*, which summarized the field of mechanics since the time of Newton, and is notable for its use of the theory of differential equations. In this text, he transformed mechanics into a branch of mathematical analysis.

cally large chain – which we know from our experience will be governed by *large-scale* wave-like excitations – is adequately described in terms of an "atomistic" language; the relevant degrees of freedom will be of a different type. We should, rather, draw on the basic principles 1–4 set out above. Notably, so far, we have paid attention to neither the intrinsic symmetry of the problem nor the fact that the number of ions, N, is large.

To reduce a microscopic model to an effective low-energy theory, often the Hamiltonian is not a very convenient starting point. Usually, it is more efficient to start from the **classical action**, S. In the present case, $S = \int dt \, L(R, \dot{R})$, where $(R, \dot{R}) \equiv \{R_I, \dot{R}_I\}$ represent the coordinates and their time derivatives. The corresponding **classical Lagrangian** L related to the Hamiltonian (1.2) is given by

classical
action

classical
Lagrangian

$$L = T - U = \sum_{I=1}^{N} \left[\frac{m}{2} \dot{R}_I^2 - \frac{k_s}{2} (R_{I+1} - R_I - a)^2 \right], \qquad (1.3)$$

where T and U denote, respectively, the kinetic and potential energy.

Since we are interested in the properties of the large-N system, we can expect boundary effects to be negligible. In this case, we may impose periodic boundary conditions, making the identification $R_{N+1} = R_1$. Further, anticipating that the effect of lattice vibrations on the solid is weak (i.e., long-range atomic order is maintained), we may assume that the deviation of ions from their equilibrium position is small (i.e., $|R_I(t) - \bar{R}_I| \ll a$). For $R_I(t) = \bar{R}_I + \phi_I(t)$, with $\phi_{N+1} = \phi_1$, the Lagrangian (1.3) assumes the simplified form

$$L = \sum_{I=1}^{N} \left[\frac{m}{2} \dot{\phi}_I^2 - \frac{k_s}{2} (\phi_{I+1} - \phi_I)^2 \right].$$

To make further progress, we will now make use of the fact that we are not concerned with behavior on "atomic" scales. For such purposes, our model would,

in any case, be much too primitive! Rather, we are interested in experimentally observable behavior that becomes manifest on macroscopic length scales (principle 2). For example, one might wish to study the specific heat of the solid in the limit of infinitely many atoms (or at least a macroscopically large number, $\mathcal{O}(10^{23})$). Under these conditions, microscopic models can usually be simplified substantially (principle 3). In particular, it is often permissible to subject a discrete lattice model to a **continuum limit**, i.e., to neglect the discreteness of the microscopic entities and to describe the system in terms of effective continuum degrees of freedom.

continuum limit

In the present case, taking a continuum limit amounts to describing the lattice fluctuations ϕ_I in terms of smooth functions of a continuous variable x. (See the figure, where the [horizontal] displacement of the point particles is plotted along the vertical axis.)
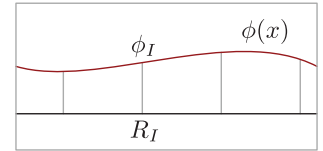
Clearly such a description makes sense only if the relative fluctuations on atomic scales are weak (for otherwise the smoothness condition would be violated). However, if this condition is met – as it will be for sufficiently large values of the stiffness constant k_s – the continuum description is much more powerful than the discrete encoding in terms of the "vector" $\{\phi_I\}$. The steps that we will need to take to go from the Lagrangian to concrete physical predictions will then be much easier to formulate.

Introducing continuum degrees of freedom $\phi(x)$, and applying a first-order Taylor expansion,[3] let us define

$$\phi_I \rightarrow a^{1/2} \phi(x)\Big|_{x=Ia}, \quad (\phi_{I+1} - \phi_I) \rightarrow a^{3/2} \partial_x \phi(x)\Big|_{x=Ia}, \quad \sum_{I=1}^{N} \rightarrow \frac{1}{a} \int_0^L dx,$$

where $L = Na$. Note that, as defined, the functions $\phi(x,t)$ have dimensionality $[\text{length}]^{1/2}$. Expressed in terms of the new degrees of freedom, the continuum limit of the Lagrangian then reads

$$L[\phi] = \int_0^L dx \ \mathcal{L}(\partial_x \phi, \dot{\phi}), \qquad \mathcal{L}(\partial_x \phi, \dot{\phi}) = \frac{m}{2}\dot{\phi}^2 - \frac{k_s a^2}{2}(\partial_x \phi)^2, \qquad (1.4)$$

Lagrangian density

where the **Lagrangian density** $\mathcal{L}$ has dimensionality $[\text{energy}]/[\text{length}]$. Similarly, the classical action assumes the continuum form

$$\boxed{S[\phi] = \int dt \ L[\phi] = \int dt \int_0^L dx \ \mathcal{L}(\partial_x \phi, \dot{\phi})} \qquad (1.5)$$

classical field

We have thus succeeded in abandoning the N point-particle description in favor of one involving *continuous* degrees of freedom, a **(classical) field**. The dynamics of the latter are specified by the **functionals** L and S, which represent the continuum generalizations of the discrete classical Lagrangian and action, respectively.

---

[3] Indeed, for reasons that will become clear, higher-order contributions to the Taylor expansion are immaterial in the long-range continuum limit.

field   **INFO** The continuum variable ϕ is our first encounter with a **field**. Before proceeding with our example, let us pause to make some preliminary remarks on the general definition of these objects. This will help to place the subsequent discussion of the atomic chain in a broader context. Formally, a field is a smooth mapping

$$\phi : M \to T$$
$$z \mapsto \phi(z)$$

field manifold   from a certain manifold M,[4] often called the "**base manifold**," into a "**target**" or "**field manifold**" T (see the figure).[5] In our present example, $M = [0, L] \times [0, t] \subset \mathbb{R}^2$ is the product of intervals in space and time, and $T = \mathbb{R}$. In fact, the factorization $M \subset \mathcal{R} \times \mathcal{T}$ into a space-like manifold $\mathcal{R}$ multiplied by a one-dimensional time-like manifold $\mathcal{T}$ is inherent in most applications of condensed matter physics.[6]
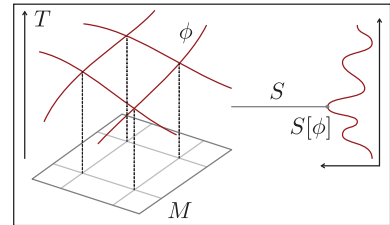
However, the individual factors $\mathcal{R}$ and $\mathcal{T}$ may, of course, be more complex than in our prototypical problem above. As for the target manifold, not much can be said in general; depending on the application, the realizations of T range from real or complex numbers over vector spaces and groups to the "fanciest objects" of mathematical physics.

functionals   In applied field theory, fields appear not as final objects, but rather as input to **functionals**. Mathematically, a functional $S : \phi \mapsto S[\phi] \in \mathbb{R}$ is a mapping that takes a field as its argument and maps it into the real numbers. The functional profile $S[\phi]$ essentially determines the character of a field theory. Notice that the argument of a functional is commonly indicated in square brackets [ ].

While these formulations may appear unnecessarily abstract, remembering the mathematical backbone of the theory often helps to avoid confusion. At any rate, it takes some time and practice to get used to the concept of fields and functionals. Conceptual difficulties in handling these objects can be overcome by remembering that any field in condensed matter physics arises as the limit of a *discrete* mapping. In the present example, the field $\phi(x)$ is obtained as a continuum limit of the discrete vector $\{\phi_I\} \in \mathbb{R}^N$; the functional $L[\phi]$ is the continuum limit of the function $L : \mathbb{R}^N \to \mathbb{R}$, etc. While, in practice, fields are usually easier to handle than their discrete analogs, it is sometimes helpful to think about problems of field theory in a discrete language. Within the discrete picture, the mathematical apparatus of field theory reduces to finite-dimensional calculus.

equations of motion   Although the Lagrangian (1.4) contains the full information about the model, we have not yet learned much about its actual behavior. To extract concrete physical information from the Lagrangian, we need to derive **equations of motion**. At first sight, it may not be entirely clear what is meant by the term "equations of motion" in the context of an infinite-dimensional model: the equations of motion relevant for

---

[4] If you are unfamiliar with the notion of manifolds (for a crash course, see appendix section A.1), think of M and T as subsets of some vector space. For the moment, this limitation won't do any harm.

[5] In some (rare) cases it becomes necessary to define fields in a more general sense (e.g., as sections of mathematical objects known as fiber bundles). However, in practically all condensed matter applications, the more restrictive definition above will suffice.

[6] By contrast, the condition of Lorentz invariance implies the absence of such factorizations in relativistic field theory. In classical statistical field theories, i.e., theories probing the thermodynamic behavior of large systems, M is just space-like.

the present problem are obtained as the generalization of the conventional Lagrange equations of N point-particle classical mechanics to a model with infinitely many degrees of freedom. To derive these equations, we need to generalize Hamilton's extremal principle (i.e., the route from an action to the associated equations of motion) to infinite dimensions. As a warm-up, let us briefly recapitulate how the extremal principle works for a system with one degree of freedom.

Suppose the dynamics of a classical *point* particle with coordinate $x(t)$ is described by the classical Lagrangian $L(x, \dot{x})$ and action $S[x] = \int dt\, L(x, \dot{x})$. **Hamilton's extremal principle** states that the configurations $x(t)$ that are *realized* are those that extremize the action, $\delta S[x] = 0$, i.e., for any smooth curve $t \mapsto y(t)$,

Hamilton's extremal principle

$$\lim_{\epsilon \to 0} \frac{1}{\epsilon}(S[x + \epsilon y] - S[x]) = 0. \tag{1.6}$$

(For a more rigorous discussion, see section 1.2 below.) To first order in ϵ, the action has to remain invariant. Applying this condition, one finds that it is fulfilled if and only if x satisfies **Lagrange's equation of motion**

Lagrange's

$$\boxed{\frac{d}{dt}(\partial_{\dot{x}} L) - \partial_x L = 0} \tag{1.7}$$

EXERCISE Recapitulate the derivation of (1.7) from the classical action.

In Eq. (1.5) we are dealing with a system of infinitely many degrees of freedom, $\phi(x, t)$. Yet Hamilton's principle is general and we may see what happens if Eq. (1.5) is subjected to an extremal principle analogous to Eq. (1.6). In this case, we require the action (1.5) to be invariant under variations $\phi(x, t) \to \phi(x, t) + \epsilon \eta(x, t)$, to first order

in ϵ. Note that field variations must respect boundary conditions, if present. For example, if $\phi|_{\text{boundary}} = \text{const.}$, then $\eta|_{\text{boundary}} = 0$ (see the figure). When applied to the specific Lagrangian (1.4), substituting the "varied" field leads to
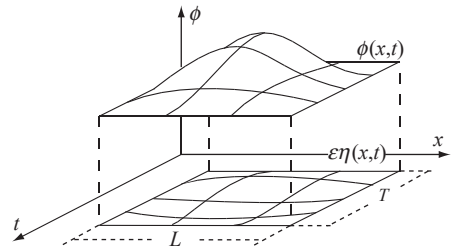
$$S[\phi + \epsilon \eta] = S[\phi] + \epsilon \int dt \int_0^L dx \left(m\dot{\phi}\,\dot{\eta} - k_s a^2 \partial_x \phi\, \partial_x \eta \right) + \mathcal{O}(\epsilon^2).$$

Integrating by parts and requiring the contribution linear in ϵ to vanish, one obtains

$$\lim_{\epsilon \to 0} \frac{1}{\epsilon}\left(S[\phi + \epsilon \eta] - S[\phi]\right) = -\int dt \int_0^L dx \left(m\ddot{\phi} - k_s a^2 \partial_x^2 \phi \right) \eta \stackrel{!}{=} 0.^7$$

(Notice that the boundary terms vanish identically.) Now, since η was defined to be an arbitrary smooth function, the integral above can vanish only if the factor in

---

[7] Here and throughout $a \stackrel{!}{=} b$ means "we require a to be equal to b."

parentheses is globally vanishing. Thus the equation of motion takes the form of a
classical wave equation

$$\left(m\partial_t^2 - k_s a^2 \partial_x^2\right)\phi = 0 \tag{1.8}$$

The solutions of (1.8) have the form
$\phi_+(x-vt)+\phi_-(x+vt)$, where $v = a\sqrt{k_s/m}$,
and $\phi_\pm$ are arbitrary smooth functions of

the argument. From this we can deduce that the low-energy **elementary excitations** of our model are lattice vibrations propagating as **classical sound waves** to the left or right at a constant velocity v (see figure).[8] The trivial behavior of the model is, of course, a direct consequence of its simplistic definition – no dissipation, dispersion, or other nontrivial ingredients. Adding these refinements leads to the general classical theory of lattice vibrations (see, e.g., Ashcroft and Mermin[9]). Finally, notice that the elementary excitations of the chain have little in common with its "microscopic" constituents (the atoms). Rather they are **collective excitations**, i.e., elementary excitations comprising a macroscopically large number of microscopic degrees of freedom.

INFO The "relevant" **excitations of a condensed matter system** can, but need not, be of collective type. For example, the interacting electron gas (a system to be discussed in detail below) supports microscopic excitations – charged quasi-particles standing in 1:1 correspondence with the electrons of the original microscopic system – while the collective excitations are plasmon modes of large wavelength that involve many electrons. Typically, the nature of the fundamental excitations cannot be straightforwardly inferred from the microscopic definition of a model. Indeed, the mere *identification* of the relevant excitations often represents the most important step in the solution of a condensed matter problem.

1.1.2 Hamiltonian formulation

An important characteristic of any excitation is its *energy*. How much energy is stored in the sound waves of the harmonic chain? To address this question, we need to switch back to a Hamiltonian formulation. Once again, this is achieved by generalizing standard manipulations from point-particle

Sir William Rowan Hamilton 1805–1865 was an Irish mathematician credited with the discovery of quaternions, the first non-commutative algebra to be studied. He also made seminal contributions to the study of geometric optics and classical mechanics.

mechanics to the continuum. Remembering that, in the Lagrangian formulation of

---

[8] Strictly speaking, the modeling of our system enforces a periodicity constraint $\phi_\pm(x + L) = \phi_\pm(x)$. However, in the limit of large system sizes, this aspect becomes inessential.
[9] N. W. Ashcroft and N. D. Mermin, *Solid State Physics* (Holt–Saunders International, 1983).

point particle mechanics, $p \equiv \partial_{\dot{x}} L$ is the momentum conjugate to the coordinate x, let us consider the Lagrangian *density* and define the field[10]

$$\pi(x) \equiv \frac{\partial \mathcal{L}}{\partial \dot{\phi}} \tag{1.9}$$

canonical momentum

as the **canonical momentum** associated with ϕ. In common with $\phi(x)$, the momentum $\pi(x)$ is a continuum degree of freedom. At each point in space, it may take an independent value. Notice that $\pi(x)$ is nothing but the continuum generalization of the lattice momentum P_I of Eq. (1.2). (Applied to P_I, a continuum approximation like $\phi_I \to \phi(x)$ would produce $\pi(x)$.) The **Hamiltonian density** is then defined as usual through the Legendre transformation,

Hamiltonian density

$$\mathcal{H}(\partial_x \phi, \pi) = \left(\pi \dot{\phi} - \mathcal{L} \right)\big|_{\dot{\phi} = \dot{\phi}(\phi, \pi)} \tag{1.10}$$

from which the full Hamiltonian is obtained as $H = \int_0^L dx \, \mathcal{H}$.

EXERCISE Verify that the transition $L \to H$ is a straightforward continuum generalization of the Legendre transformation of the N-particle Lagrangian $L(\{\phi_I\}, \{\dot{\phi}_I\})$.

Having introduced a Hamiltonian, we are in a position to determine the energy of the sound waves. Application of Eqs. (1.9) and (1.10) to the Lagrangian of the atomic chain yields $\pi(x) = m\dot{\phi}(x)$ and

$$H[\pi, \phi] = \int dx \left[\frac{\pi^2}{2m} + \frac{k_{\mathrm{s}} a^2}{2} (\partial_x \phi)^2 \right]. \tag{1.11}$$

Considering, say, a right-moving sound-wave excitation, $\phi(x,t) = \phi_+(x - vt)$, we find that $\pi(x,t) = -mv\partial_x \phi_+(x - vt)$ and $H[\pi, \phi] = k_{\mathrm{s}} a^2 \int dx [\partial_x \phi_+(x - vt)]^2 = k_{\mathrm{s}} a^2 \int dx \, [\partial_x \phi_+(x)]^2$, i.e., a positive-definite time-independent expression, as one would expect.

INFO For completeness, we mention that the **Hamiltonian representation of the action** (1.5) is given by $S[\phi, \pi] = \int dt \int_0^L dx (\pi\dot{\phi} - \mathcal{H})$. From here, the Hamiltonian version of the equations of motion can be derived by independent variations in ϕ and π, just as in classical mechanics. As an exercise, carry out this variation for the harmonic chain and verify that you obtain equations equivalent to the wave equation (1.8).[11] Whether one prefers to work in a Hamiltonian or Lagrangian formulation of a field theory depends on the context and is often decided on a case-by-case basis.

Hamiltonian action

Before proceeding further, let us note an interesting feature of the energy functional: in the limit of an infinitely shallow excitation, $\partial_x \phi_+ \to 0$, the energy vanishes. This sets the stage for principles 4, hitherto unconsidered, **symmetry**. The Hamiltonian

---

[10] In field theory literature, it is traditional to denote the momentum by a Greek letter.

[11] Variation of the action in ϕ and π leads to (invert this to check the result) $\phi \frac{x}{c} \varrho_c v^s \eta = \psi$ ' $\frac{uu}{u} = \phi$. Differentiation of the first equation in time followed by substitution into the second equation yields the wave equation.

of an atomic chain is invariant under simultaneous translation of all atom coordinates by a fixed increment: $\phi_I \to \phi_I + \delta$, where δ is constant. This expresses the fact that a global translation of the solid as a whole does not affect the internal energy. Now, the ground state of any specific realization of the solid is defined through a static array of atoms, each located at a fixed coordinate $R_I = Ia \Rightarrow \phi_I = 0$. We say that the above translational symmetry is "spontaneously broken," i.e., the solid has to decide where exactly it wants to rest. However, spontaneous breakdown of a symmetry does not imply that the symmetry has disappeared. On the contrary, infinite-wavelength deviations from the pre-assigned ground state come close to global translations of (macroscopically large portions of) the solid and, therefore, cost a vanishingly small amount of energy. This is the reason for the vanishing of the sound-wave energy in the limit $\partial_x \phi \to 0$. It is also our first encounter with the aforementioned phenomenon that continuous symmetries lead to the formation of soft, i.e., low-energy, excitations. A much more systematic exposition of these connections will be given in chapter 5.

To conclude our discussion of the classical harmonic chain, let us consider the **specific heat**, a quantity directly accessible in experiment. A rough estimate of this quantity can be obtained from the microscopic Hamiltonian (1.2). According to the principles of statistical mechanics, the thermodynamic energy density is given by

$$u = \frac{1}{L} \frac{\int d\Gamma \; e^{-\beta H} H}{\int d\Gamma \; e^{-\beta H}} = -\frac{1}{L} \partial_\beta \ln \mathcal{Z},$$

where $\beta \equiv 1/k_\mathrm{B} T$, $\mathcal{Z} \equiv \int d\Gamma e^{-\beta H}$ is the **Boltzmann partition function**, and the phase space volume element $d\Gamma = \prod_{I=1}^{N} dR_I dP_I$. (Hereafter, for simplicity, we set $k_\mathrm{B} = 1$.) The specific heat is then obtained as $c = \partial_T u$. To determine the temperature dependence of c, we can make use of the fact that, upon rescaling of the integration variables, $R_I \to \beta^{-1/2} X_I$, $P_I \to \beta^{-1/2} Y_I$, the exponent $\beta H(R,P) \to H(X,Y)$ becomes independent of temperature (a property that relies on the quadratic dependence of H on both R and P). The integration measure transforms as $d\Gamma \to \beta^{-N} \prod_{I=1}^{N} dX_I \; dY_I \equiv \beta^{-N} d\Gamma'$. Expressed in terms of the rescaled variables, one obtains the energy density as $u = -L^{-1} \partial_\beta \ln(\beta^{-N} K) = \rho T$, where $\rho = N/L$ is the density of the atoms and we have made use of the fact that the constant $K \equiv \int d\Gamma' \; e^{-H(X,Y)}$ is independent of temperature. We thus find a temperature independent specific heat $c = \rho$. Notice that c is fully universal, i.e., independent of the material constants m and k_s determining H. (In fact, we could have anticipated this result from the equipartition theorem of classical mechanics, i.e., the law that in a system with N degrees of freedom, the energy scales as $U = NT$.)

How do these findings compare with experiment? Figure 1.1 shows the specific heat of the insulating compound $EuCoO_3$.[12] At large temperatures, the specific heat approaches a constant, which is consistent with our analysis. However, at

[12] Note that, in metals, the specific heat due to lattice vibrations exceeds the specific heat of the free conduction electrons for temperatures larger than a few degrees kelvin.
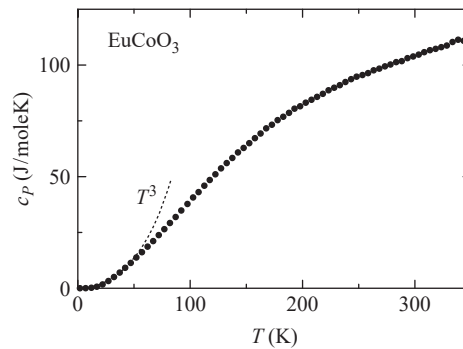
Fig. 1.1   Specific heat c_p of the insulator $EuCoO_3$. At large temperatures, c_p approaches a constant value, as predicted by analysis of the classical harmonic chain. However, for small temperatures, deviations from $c_p = \text{const.}$ are substantial. Such deviations can be ascribed to quantum effects. (Courtesy of M. Kriener, A. Reichl, T. Lorenz and, A. Freimuth.)

lower temperatures, the specific heat shows deviations from $c = \text{const.}$ Yet, this temperature dependence does not reflect a failure of the simplistic microscopic modeling. Rather, the deviation is indicative of a **quantum phenomenon**. Indeed, so far, we have neglected the quantum nature of the atomic variables. In the next chapter we will discuss how an effective low-energy theory of the harmonic chain can be promoted to a quantum field theory. However, before doing so, let us pause to introduce several mathematical concepts that surfaced above, in a way that survives generalization to richer problems.

1.2 Functional Analysis and Variational Principles

Let us revisit the derivation of the equations of motion associated with the harmonic chain, Eq. (1.8). Although straightforward, the calculation was neither efficient, nor did it reveal general structures. In fact, what we did – expanding explicitly to first order in the variational parameter ϵ – has the same status as evaluating derivatives by explicitly taking limits: $f'(x) = \lim_{\epsilon \to 0} \frac{1}{\epsilon}(f(x+\epsilon) - f(x))$. Moreover, the derivation made explicit use of the particular form of the Lagrangian, thereby being of limited use with regard to a general understanding of the construction scheme. Given the importance attached to extremal principles in the whole of field theory, it is worthwhile investing some effort in constructing a more efficient scheme for the general variational analysis of continuum theories. To carry out this program, we first need to introduce the mathematical tool of functional analysis – the concept of **functional differentiation**.

functional differentiation

In working with functionals, one is often concerned with how a given functional behaves under (small) variations of its argument function. In particular, given a certain function f that we suspect would make a functional $F[f]$ stationary, one

would like to find out whether the functional remains invariant under variations $f \to f + h$, where h is an infinitely small increment function. In ordinary analysis, questions of this type are commonly addressed by exploring *derivatives*, i.e., we need to generalize the concept of a derivative to functionals. This is achieved by the following definition: a functional F is called differentiable if

$$F[f + \epsilon g] - F[f] = \epsilon \, DF_f[g] + \mathcal{O}(\epsilon^2)$$

differential   where the **differential** DF_f is a linear functional (i.e., one with $DF_f[g_1 + g_2] = DF_f[g_1] + DF_f[g_2]$), ϵ is a small parameter, and g is an arbitrary function. The subscript indicates that the differential depends generally on the "base argument," f. A functional F is said to be **stationary** on f if, and only if, $DF_f = 0$.

In principle, the definition above answers our question concerning a stationarity condition. However, to make use of the definition, we still need to know how to compute the differential DF, and how to relate the differentiability criterion to the concepts of ordinary calculus. To understand how these questions can be systematically addressed, it is helpful to return temporarily to a discrete way of thinking, i.e., to interpret the argument f of a functional $F[f]$ as the limit $N \to \infty$ of a discrete vector $\mathbf{f} = \{f_n \equiv f(x_n), n = 1, \ldots, N\}$, where $\{x_n\}$ denotes a discretization of the support of f (cf. the harmonic chain, $\phi \leftrightarrow f$). Prior to taking the continuum limit, $N \to \infty$, $\mathbf{f}$ has the status of an N-dimensional vector and $F(\mathbf{f})$ is a function defined over N-dimensional space. After taking the continuum limit, $\mathbf{f} \to f$ becomes a function itself and $F(\mathbf{f}) \to F[f]$ becomes a functional.

Now, within the discrete picture, it is clear how the variational behavior of functions is to be analyzed. For example, the condition that, for all ϵ and all vectors $\mathbf{g}$, the linear expansion of $F(\mathbf{f} + \epsilon \mathbf{g})$ ought to vanish is simply to say that the ordinary differential, $dF_{\mathbf{f}}$, defined through $F(\mathbf{f} + \epsilon \mathbf{g}) - F(\mathbf{f}) = \epsilon \cdot dF_{\mathbf{f}}(\mathbf{g}) + \mathcal{O}(\epsilon^2)$, must be zero. In practice, one often expresses conditions of this type in terms of a certain basis. In a Cartesian basis of N unit vectors, $\mathbf{e}_n, n = 1, \ldots, N$, $dF_{\mathbf{f}}(\mathbf{g}) \equiv \langle \nabla F_{\mathbf{f}}, \mathbf{g} \rangle$, where $\langle \mathbf{f}, \mathbf{g} \rangle \equiv \sum_{n=1}^{N} f_n g_n$ denotes the standard scalar product, and $\nabla F_{\mathbf{f}} = \{\partial_{f_n} F\}$ represents the gradient, with the partial derivative defined as

$$\partial_{f_n} F(\mathbf{f}) \equiv \lim_{\epsilon \to 0} \frac{1}{\epsilon} \left[F(\mathbf{f} + \epsilon \mathbf{e}_n) - F(\mathbf{f}) \right]. \tag{1.12}$$

From these identities, the differential is identified as

$$dF_{\mathbf{f}}(\mathbf{g}) = \sum_n \partial_{f_n} F(\mathbf{f}) g_n. \tag{1.13}$$

The vanishing of the differential amounts to the vanishing of all partial derivatives, $\partial_{f_n} F = 0$.

Equations (1.12) and (1.13) can now be straightforwardly generalized to the continuum limit, whereupon the summation defining the finite-dimensional scalar product translates to an integral,

$$\langle \mathbf{f}, \mathbf{g} \rangle = \sum_{n=1}^{N} f_n g_n \to \langle f, g \rangle = \int dx \; f(x) g(x).$$

The analog of the nth unit vector is a δ-distribution, $\mathbf{e}_n \to \delta_x$, where $\delta_x(x') \equiv \delta(x - x')$, as can be seen from the following correspondence:

$$f_n \overset{!}{=} \langle \mathbf{f}, \mathbf{e}_n \rangle = \sum_{m=1}^{N} f_m (e_n)_m \to f(x) \overset{!}{=} \langle f, \delta_x \rangle = \int dx' \, f(x') \delta_x(x').$$

Here $(e_n)_m = \delta_{nm}$ denotes the mth component of the nth unit vector. The correspondence (unit vector $\leftrightarrow$ δ-distribution) is easy to memorize: while components of $\mathbf{e}_n$ vanish, save for the nth component, which equals unity, δ_x is a function that vanishes everywhere, save for x where it is *infinite*. That a unit component is replaced by infinity reflects the fact that the support of the δ-distribution is infinitely narrow; to obtain a unit-normalized integral $\int \delta_x$, the function must be singular.

As a result of these identities, (1.13) translates to the continuum differential,

$$DF_f[g] = \int dx \, \frac{\delta F[f]}{\delta f(x)} g(x), \tag{1.14}$$

where the generalization of the partial derivative,

$$\frac{\delta F[f]}{\delta f(x)} \equiv \lim_{\epsilon \to 0} \frac{1}{\epsilon} \left(F[f + \epsilon \delta_x] - F[f] \right) \tag{1.15}$$

is commonly denoted by δ instead of ∂. Equations (1.14) and (1.15) establish the connection between ordinary and functional differentiation. Notice that we have not yet learned how to calculate the differential practically, i.e., to evaluate expressions like Eq. (1.15) for concrete functionals. Nevertheless, the identities above are very useful, enabling us to generalize more complex derivative operations of ordinary calculus by straightforward extrapolation. For example, the generalization of the **chain rule** standard **chain rule**, $\partial_{f_n} F(\mathbf{g}(\mathbf{f})) = \sum_m \partial_{g_m} F(\mathbf{g}) \big|_{\mathbf{g} = \mathbf{g}(\mathbf{f})} \partial_{f_n} g_m(\mathbf{f})$ reads

$$\frac{\delta F[g[f]]}{\delta f(x)} = \int dy \, \frac{\delta F[g]}{\delta g(y)} \bigg|_{g=g[f]} \frac{\delta g(y)[f]}{\delta f(x)}.$$

Here $g[f]$ is the continuum generalization of an $\mathbb{R}^m$-valued function, $\mathbf{g} : \mathbb{R}^n \to \mathbb{R}^m$, a function whose components $g(y)[f]$ are functionals by themselves. Furthermore, given some functional $F[f]$, we can construct its **Taylor expansion** as

$$F[f] = F[0] + \int dx_1 \, \frac{\delta F[f]}{\delta f(x_1)} \bigg|_{f=0} f(x_1) + \int \frac{dx_1 \, dx_2}{2} \, \frac{\delta^2 F[f]}{\delta f(x_2)\delta f(x_1)} \bigg|_{f=0} f(x_1)f(x_2) + \cdots,$$

where (exercise)

$$\frac{\delta^2 F[f]}{\delta f(x_2)\delta f(x_1)}$$
$$= \lim_{\epsilon_{1,2} \to 0} \frac{1}{\epsilon_1 \epsilon_2} \left(F[f + \epsilon_1 \delta_{x_1} + \epsilon_2 \delta_{x_2}] - F[f + \epsilon_1 \delta_{x_1}] - F[f + \epsilon_2 \delta_{x_2}] + F[f] \right)$$

generalizes a two-fold partial derivative. The validity of these identities can be made plausible by applying the prescription given in table 1.1 to the corresponding relations of standard calculus. To actually verify the formulae, one has to find the continuum limit of each step taken in the discrete variant of the corresponding

Table 1.1 Summary of basic definitions of discrete and continuum calculus.

Entity	Discrete	Continuum
Argument	vector $\mathbf{f}$	function f
Function(al)	multidimensional function $F(\mathbf{f})$	functional $F[f]$
Differential	$dF_{\mathbf{f}}(\mathbf{g})$	$DF_f[g]$
Cartesian basis	$\mathbf{e}_n$	δ_x
Scalar product $\langle\,,\,\rangle$	$\sum_n f_n g_n$	$\int dx\ f(x)g(x)$
"Partial derivative"	$\partial_{f_n} F(\mathbf{f})$	$\delta F[f]/\delta f(x)$

proofs. Experience shows that it takes some time to get used to the concept of functional differentiation. However, after some practice, it will become clear that this operation is not only useful but is as easy to handle as conventional partial differentiation.

We finally address the question how to compute functional derivatives in practice. In doing so, we will make use of the fact that, in all but a few cases, the functionals encountered in field theory are of the structure

$$S[\phi] = \int_M dx\, \mathcal{L}(\phi^i, \partial_\mu \phi^i) \tag{1.16}$$

Here, we assume the base manifold M to be parameterized by an m-dimensional coordinate vector $x = \{x_\mu\}$. (In most practical applications $m = d + 1$ and $x = (x_0, x_1, \ldots, x_d)$ contains one time-like component $x_0 = t$ and d space-like components $x_k, k = 1, \ldots, d.$[13]) We further assume that the field manifold has dimensionality n and that $\phi^i, i = 1, \ldots, n$, are the coordinates of the field. Functionals of this type are called **local functionals**.

local functional

What makes the functional $S[\phi]$ easy to handle is that all of its information is stored in the *function* $\mathcal{L}$. Owing to this simplification, the functional derivative can be related to an ordinary derivative of $\mathcal{L}$. To see this, all we have to do is to evaluate the general definition (1.14) on the functional S:

$$S[\phi + \epsilon\theta] - S[\phi] = \int_M dx \left[\mathcal{L}(\phi + \epsilon\theta, \partial_\mu\phi + \epsilon\partial_\mu\theta) - \mathcal{L}(\phi, \partial_\mu\phi) \right]$$

$$= \int_M dx \left[\frac{\partial \mathcal{L}}{\partial \phi^i}\theta^i + \frac{\partial \mathcal{L}}{\partial(\partial_\mu\phi^i)}\partial_\mu\theta^i \right] \epsilon + O(\epsilon^2)$$

$$= \int_M dx \left[\frac{\partial \mathcal{L}}{\partial \phi^i} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu\phi^i)} \right) \right] \theta^i \epsilon + O(\epsilon^2),$$

where in the last line we have assumed that the field variation vanishes on the boundary of the base manifold, $\theta\,|_{\partial M} = 0$. Comparison with Eq. (1.14) identifies the functional derivative as

$$\frac{\delta S[\phi]}{\delta \phi^i(x)} = \frac{\partial \mathcal{L}}{\partial \phi^i(x)} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu\phi^i(x))} \right).$$

[13] Following standard convention, we denote space-like components by small Latin indices $k = 1, \ldots, d$. By contrast, space–time indices are denoted by Greek indices $\mu = 0, \ldots, d$.

We conclude that the **stationarity of the functional** (1.16) is equivalent to the condition

$$\forall x, i : \quad \frac{\partial \mathcal{L}}{\partial \phi^i(x)} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi^i(x))} \right) = 0 \tag{1.17}$$

Euler–Lagrange equation

Equation (1.17) is known as the **Euler–Lagrange equation** of field theory. In fact, for $d = 0$ and $x_0 = t$, Eq. (1.17) reduces to the familiar Euler–Lagrange equation for a point particle in n-dimensional space. For $d = 1$ and $(x_0, x_1) = (t, x)$, we get back to the stationarity equations discussed in the previous section. In the next section we will apply the formalism to a higher-dimensional problem.

1.3 Maxwell's Equations as a Variational Principle

REMARK This section requires familiarity with the basic notions of special relativity such as the concepts of 4-vectors, Lorentz transformations, and covariant notation.[14]

classical electrodynamics

As a second example, let us consider *the* archetype of classical field theory, **classical electrodynamics**. As well as exemplifying the application of continuum variational principles for a familiar problem, this example illustrates the unifying potential of the approach: That problems as different as the low-lying vibrational modes of a crystalline solid and electrodynamics can be described by almost identical language indicates that we are dealing with a

> **James Clerk Maxwell 1831–1879**
> was a Scottish theoretical physicist and mathematician who made seminal contributions to the study of electricity, magnetism, optics, and the kinetic theory of gases. In particular, he is credited with the formulation of the theory of electromagnetism, synthesizing seemingly unrelated experiments and equations of electricity, magnetism and optics into a consistent theory. He is also known for creating the first true color photograph in 1861!

Maxwell's equations

useful formalism. Specifically, our aim will be to explore how the equations of motion of electrodynamics, the inhomogeneous **Maxwell's equations**,

$$\nabla \cdot \mathbf{E} = \rho, \qquad \nabla \times \mathbf{B} - \partial_t \mathbf{E} = \mathbf{j}, \tag{1.18}$$

can be obtained from variational principles. For simplicity, we restrict ourselves to a vacuum theory, i.e., $\mathbf{E} = \mathbf{D}$ and $\mathbf{B} = \mathbf{H}$. Further, we have set the velocity of light to unity, $c = 1$. Within the framework of the variational principle, the homogeneous equations,

$$\nabla \times \mathbf{E} + \partial_t \mathbf{B} = 0, \qquad \nabla \cdot \mathbf{B} = 0, \tag{1.19}$$

are regarded as *ab initio* constraints imposed on the degrees of freedom $\mathbf{E}$ and $\mathbf{B}$.

Lorentz invariance

INFO As preparation for the following discussion, let us briefly recapitulate the notion of **Lorentz invariance**. In this text, we will work mostly in non-relativistic contexts

---

[14] For a summary of the covariant notation used in this text, see the Info block on 524.

where the time coordinate t and the d space coordinates x_i are bundled into a $(d+1)$-dimensional vector $x = x^\mu = x_\mu = (t, x_i)$ and $\mu = 0, \ldots, d$. In this case, t and x_i may be considered as coordinates of a Euclidean space. Field theories defined in such spaces are called **Euclidean field theories**. By contrast, in relativistic theories we are working in a space–time continuum with a **Minkowski metric**

Euclidean field theory

$$
\eta = \{\eta^{\mu\nu}\} = \begin{pmatrix} -1 & & & \\ & +1 & & \\ & & +1 & \\ & & & +1 \end{pmatrix}. \tag{1.20}
$$

Here, too, we denote space–time coordinate vectors by $x = x^\mu = (t, x_i)$. However, now the – contravariant or "upstairs" – positioning of the index becomes an essential part of the notation; see Info block on 524 for a summary of the notation conventions of relativity.

Lorentzian field theory

Field theories in space–times with a Minkowski metric are called **Lorentzian field theories**. Recall that a linear coordinate transformation $x^\mu \to x'^\mu \equiv \Lambda^\mu{}_\nu x^\nu$ is a Lorentz transformation if it leaves the Minkowski metric invariant: $x^\mu \eta_{\mu\nu} x^\nu = x'^\mu \eta_{\mu\nu} x'^\nu$. In the covariant notation of relativity, covariant components, x_μ, are obtained from contravariant components, x^μ, by index lowering via the Minkowski metric, $x_\mu = \eta_{\mu\nu} x^\nu$ (this is why the positioning is relevant) and the invariance condition assumes the form $x_\mu x^\mu = x'_\mu x'^\mu$. Expressed as a condition for the Lorentz transformations, this reads $\eta_{\mu\nu} \Lambda^\mu{}_{\mu'} \Lambda^\nu{}_{\nu'} = \eta_{\mu'\nu'}$.

Hendrik Antoon Lorentz
1853–1928
was a Dutch physicist who shared the 1902 Nobel Prize in Physics with Pieter Zeeman "in recognition of the extraordinary service they rendered by their researches into the influence of magnetism upon radiation phenomena." Lorentz derived the transformation equations subsequently used by Albert Einstein to describe space and time.

In cases where we are discussing material which does not depend on the realization of the metric, covariant notation will be used. The Euclidean field theory is then represented by the unit metric $\eta_{\mu\nu} = \delta_{\mu\nu}$.

The representation of Maxwell's theory as a variational principle is best formulated in the language of **relativistically invariant electrodynamics**. As a starting point, we require (1) a field formulated in a set of suitably generalized coordinates and (2) its action. Regarding coordinates, the natural choice will be the coefficients of the electromagnetic (EM) 4-potential, $A^\mu = (\phi, \mathbf{A})$, where ϕ is the scalar potential and $\mathbf{A}$ is the vector potential. The 4-potential A is unconstrained and uniquely determines the fields $\mathbf{E}$ and $\mathbf{B}$ through the standard equations $\mathbf{E} = -\nabla\phi - \partial_t \mathbf{A}$ and $\mathbf{B} = \nabla \times \mathbf{A}$. (In fact, the set of coordinates A_μ is "overly free" in the sense that gauge transformations $A_\mu \to A_\mu + \partial_\mu \Gamma$, where Γ is an arbitrary function, leave the physical fields invariant. Later we will comment explicitly on this point.) The connection between A and the physical fields can be expressed in a more symmetric way by introducing the EM field tensor,[15]

---

[15] Notice that the field tensor (1.21) differs from that in many textbooks on electromagnetism by a sign change, $E_i \leftrightarrow -E_i$. The reason is that in this text we work with a different sign convention for the Minkowski metric, $\eta \leftrightarrow -\eta$; see p.524 for a discussion of this point.

$$F = \{F_{\mu\nu}\} = \begin{pmatrix} 0 & -E_1 & -E_2 & -E_3 \\ E_1 & 0 & -B_3 & B_2 \\ E_2 & B_3 & 0 & -B_1 \\ E_3 & -B_2 & B_1 & 0 \end{pmatrix}. \tag{1.21}$$

The relation between fields and potential now reads $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$, where $x_\mu = (-t, \mathbf{x})$ and $\partial_\mu = (\partial_t, \nabla)$.[14]

EXERCISE Confirm that this relation follows from the definition of the vector potential. To verify that the constraint (1.19) is automatically included in the definition (1.21), compute the construct $\partial_\lambda F_{\mu\nu} + \partial_\mu F_{\nu\lambda} + \partial_\nu F_{\lambda\mu}$, where $(\lambda\nu\mu)$ represent arbitrary but *different* indices. This produces four different terms, identified as the left-hand side of Eq. (1.19). Evaluation of the same construct on $F_{\mu\nu} \equiv \partial_\mu A_\nu - \partial_\nu A_\mu$ produces zero, by the symmetry of the right-hand side.

To obtain the structure of the action $S[A]$, we can proceed in different ways. One option would be to regard Maxwell's equations as fundamental, i.e., to construct an action that produces these equations upon variation (by analogy with the situation in classical mechanics where the action functional is designed to reproduce Newton's equations). However, we can also be a little bit more ambitious and ask whether the structure of the action can be motivated independently of Maxwell's equations. In fact, there is just one principle in electrodynamics that is as fundamental as Maxwell's equations: **symmetry**. A theory of electromagnetism must be Lorentz invariant, i.e., invariant under relativistic coordinate transformations.

Aided by the symmetry criterion, we can try to conjecture the structure of the action from three basic assumptions, all independent of Maxwell's equations. The action should be invariant under (i) Lorentz transformations, (ii) gauge transformations, and (iii) it should be simple! The most elementary choice compatible with these conditions is

$$S[A] = \int d^4x \, (c_1 \, F_{\mu\nu} F^{\mu\nu} + c_2 \, A_\mu j^\mu), \tag{1.22}$$

where $d^4x = dt \, dx_1 \, dx_2 \, dx_3$ denotes the measure, $j^\mu = (\rho, \mathbf{j})$ the 4-current, and $c_{1,2}$ are undetermined constants. Indeed, up to quadratic order in A, (1.22) defines the only possible structure consistent with gauge and Lorentz invariance.

EXERCISE Using the continuity equation $\partial_\mu j^\mu = 0$, verify that the Aj-coupling is gauge invariant. (Hint: Integrate by parts.) Verify that a contribution like $\int A_\mu A^\mu$ would *not* be gauge invariant.

Having defined a trial action, we can apply the variational principle (1.17) to compute equations of motion. In the present context, the role of the field ϕ is taken by the four components of A. Variation of the action with respect to A_μ gives four equations of motion,

$$\frac{\partial \mathcal{L}}{\partial A_\mu} - \partial_\nu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\nu A_\mu)} \right) = 0, \quad \mu = 0, \dots, 3, \tag{1.23}$$

where the Lagrangian density is defined by $S = \int d^4 x \, \mathcal{L}$. With the specific form of $\mathcal{L}$, it is straightforward to verify that $\partial_{A_\mu} \mathcal{L} = c_2 j^\mu$ and $\partial_{(\partial_\nu A_\mu)} \mathcal{L} = -4 c_1 F^{\mu\nu}$. We substitute these building blocks into the equations of motion to obtain $4 c_1 \partial_\nu F^{\nu\mu} = c_2 j^\mu$. Comparing this with the definition of the field tensor (1.21), and setting $c_1 / c_2 = -1/4$, we arrive at Maxwell's equations (1.18). The overall multiplicative constant c_1 $(= c_2/4)$ can be fixed by requiring that the Hamiltonian density associated with the Lagrangian density $\mathcal{L}$ reproduce the known energy density of the EM field (see problem 1.8.2). This leads to $c_1 = -1/4$, so that we have identified

$$\boxed{\mathcal{L}(A_\mu, \partial_\nu A_\mu) = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + A_\mu j^\mu} \qquad (1.24)$$

electro-
magnetic
Lagrangian

as the **Lagrangian density of the electromagnetic field**. The corresponding action is given by $S[A] = \int d^4 x \, \mathcal{L}(A_\mu, \partial_\nu A_\mu)$.

At first sight, this result does not look surprising. After all, Maxwell's equations can be found on the first page of most textbooks on electrodynamics. However, our achievement is actually quite remarkable. By invoking only symmetry, the algebraic structure of Maxwell's equations has been established unambiguously. We have thus proven that Maxwell's equations are relativistically invariant, a fact not obvious from the equations themselves. Further, we have shown that Eqs. (1.18) are the *only* equations of motion linear in the current-density distribution and consistent with the invariance principle. One might object that, in addition to symmetry, we have also imposed an *ad hoc* "simplicity" criterion on the action $S[A]$. However, later we will see that this was motivated by more than mere aesthetics.

Finally, we note that the symmetry-oriented modeling that led to Eq. (1.22) is illustrative of a popular construction scheme in modern field theory. The symmetry-oriented approach stands as complementary to the "microscopic" formulation exemplified in section 1.1. Broadly speaking, these are the two principal approaches to constructing effective low-energy field theories.

▷ **Microscopic analysis:** Starting from a microscopically defined system, one projects onto those degrees of freedom that one believes are relevant for the low-energy dynamics. Ideally, this "belief" is backed up by a small expansion parameter stabilizing the mathematical parts of the analysis. *Advantages:* The method is rigorous and fixes the resulting field theory completely. *Disadvantages:* The method is time consuming and, for complex systems, not even viable.

▷ **Symmetry considerations:** One infers an effective low-energy theory on the basis of only fundamental symmetries of the physical system. *Advantages:* The method is fast and elegant. *Disadvantages:* It is less explicit than the microscopic approach. Most importantly, it does not fix the coefficients of the different contributions to the action.

Thus far, we have introduced some basic concepts of field-theoretical modeling in condensed matter physics. Starting from a microscopic model Hamiltonian, we have illustrated how principles of universality and symmetry can be applied to distill

effective continuum field theories, capturing the low-energy content of the system. We have formulated such theories in the language of Lagrangian and Hamiltonian continuum mechanics, and shown how variational principles can be applied to extract concrete physical information. Finally, we have seen that field theory provides a unifying framework whereby analogies between seemingly different physical systems can be uncovered. In the next section we discuss how the formalism of classical field theory can be elevated to the quantum level.

1.4 Quantum Chain

Previously, from measurements of the specific heat, we have seen that at low temperatures the excitation profile of the classical atomic chain differs drastically from that observed experimentally. Generally, in condensed matter physics, low-energy phenomena with pronounced temperature sensitivity are indicative of a quantum mechanism at work. To introduce and exemplify a general procedure whereby quantum mechanics can be incorporated into continuum models, we next consider the low-energy physics of the quantum atomic chain.

The first question to ask is conceptual: how can a model like (1.4) be quantized in general? Indeed, there exists a standard procedure for quantizing continuum theories, which closely resembles the quantization of Hamiltonian point mechanics. Consider the defining equations (1.9) and (1.10) for the canonical momentum and the Hamiltonian, respectively. Classically, the momentum $\pi(x)$ and the coordinate $\phi(x)$ are canonically conjugate variables: $\{\pi(x), \phi(x')\} = -\delta(x - x')$, where $\{\,,\,\}$ is **Poisson bracket** the **Poisson bracket** and the δ-function arises through continuum generalization of the discrete identity $\{P_I, R_{I'}\} = -\delta_{II'}$, $I, I' = 1, \ldots, N$.[16] The theory is quantized by generalization of the canonical quantization procedure for the discrete pair of conjugate coordinates (R_I, P_I) to the continuum: (i) promote $\phi(x)$ and $\pi(x)$ to operators, $\phi \mapsto \hat{\phi}$, $\pi \mapsto \hat{\pi}$, and (ii) generalize the canonical commutation relation $[P_I, R_{I'}] = -i\hbar\delta_{II'}$ to[17]

$$\boxed{[\hat{\pi}(x), \hat{\phi}(x')] = -i\hbar\delta(x - x')} \qquad (1.25)$$

---

[16] Recall that for conjugate coordinates (R_I, P_I) the Poisson bracket is defined by

$$\{f, g\} = \sum_{I=1}^{N} \left(\frac{\partial f}{\partial R_I} \frac{\partial g}{\partial P_I} - \frac{\partial f}{\partial P_I} \frac{\partial g}{\partial R_I} \right).$$

[17] Note that the dimensionalities of both the quantum and the classical continuum field are compatible with the dimensionality of the Dirac δ-function, $[\delta(x - x')] = [\text{length}]^{-1}$, i.e., $[\phi(x)] = [\phi_I] \times [\text{length}]^{-1/2}$ and similarly for π.

Table 1.2 *Relations between discrete and continuum canonically conjugate variables or operators.*

	Classical	Quantum
Discrete	$\{P_I, R_{I'}\} = -\delta_{II'}$	$[\hat{P}_I, \hat{R}_{I'}] = -i\hbar\delta_{II'}$
Continuum	$\{\pi(x), \phi(x')\} = -\delta(x - x')$	$[\hat{\pi}(x), \hat{\phi}(x')] = -i\hbar\delta(x - x')$

quantum field Operator-valued functions like $\hat{\phi}$ and $\hat{\pi}$ are generally referred to as **quantum fields**. For clarity, the relevant relations between canonically conjugate classical and quantum fields are summarized in Table 1.2.

INFO By introducing quantum fields, we have departed from the conceptual framework laid out on page 8: being operator-valued, the quantized field no longer represents a mapping into an ordinary differentiable manifold.[18] It is thus legitimate to ask why we bothered to give a lengthy exposition of fields as "ordinary" functions. The reason is that, in the not too distant future, after the framework of functional field integration has been introduced, we will return to the comfortable ground of the definition on page 8.

Employing these definitions, the classical Hamiltonian density (1.10) becomes the quantum operator

$$\hat{\mathcal{H}}(\hat{\phi}, \hat{\pi}) = \frac{1}{2m}\hat{\pi}^2 + \frac{k_s a^2}{2}(\partial_x\hat{\phi})^2. \tag{1.26}$$

The Hamiltonian above represents a quantum field-theoretical *formulation* of the problem, but not yet a solution. In fact, the development of a spectrum of methods for the analysis of quantum field-theoretical models will represent a major part of this text. At this point our objective is merely to exemplify the way in which physical information can be extracted from models like (1.26). As a word of caution, let us mention that the following manipulations, while mathematically straightforward, are conceptually deep. To disentangle different aspects of the problem, we will first concentrate on the plain operational aspects. Later in this section, we will reflect on "what has really happened."

As with any function, operator-valued functions can be represented in a variety of ways. In particular, they can be subjected to Fourier transformation,

$$\begin{Bmatrix} \hat{\phi}_k \\ \hat{\pi}_k \end{Bmatrix} \equiv \frac{1}{L^{1/2}}\int_0^L dx\ e^{\{\mp ikx\}}\begin{Bmatrix} \hat{\phi}(x) \\ \hat{\pi}(x) \end{Bmatrix}, \quad \begin{Bmatrix} \hat{\phi}(x) \\ \hat{\pi}(x) \end{Bmatrix} = \frac{1}{L^{1/2}}\sum_k e^{\{\pm ikx\}}\begin{Bmatrix} \hat{\phi}_k \\ \hat{\pi}_k \end{Bmatrix}, \quad (1.27)$$

where $\sum_k$ represents the sum over all Fourier coefficients indexed by the quantized momenta $k = 2\pi m/L$, $m \in \mathbb{Z}$ (not to be confused with the operator momentum $\hat{\pi}$). Note that the *real* classical field $\phi(x)$ quantizes to a *hermitian* quantum field $\hat{\phi}(x)$, implying that $\hat{\phi}_k = \hat{\phi}^\dagger_{-k}$ (and similarly for $\hat{\pi}_k$). The corresponding Fourier representation of the canonical commutation relations reads (exercise)

$$[\hat{\pi}_k, \hat{\phi}_{k'}] = -i\hbar\delta_{kk'}. \tag{1.28}$$

---

[18] At least if we ignore the mathematical subtlety that a linear operator can also be interpreted as an element of a certain manifold.

When expressed in the Fourier representation, making use of the identity,

$$\int_0^L dx\,(\partial_x\hat{\phi})^2 = \sum_{k,k'}(-ik\hat{\phi}_k)(-ik'\hat{\phi}_{k'})\overbrace{\frac{1}{L}\int_0^L dx\,e^{-i(k+k')x}}^{\delta_{k+k',0}} = \sum_k k^2\hat{\phi}_k\hat{\phi}_{-k},$$

together with a similar relation for $\int_0^L dx\,\hat{\pi}^2$, the Hamiltonian $\hat{H} = \int_0^L dx\,\mathcal{H}(\hat{\phi},\hat{\pi})$ assumes the near diagonal form,

$$\hat{H} = \sum_k\left[\frac{1}{2m}\hat{\pi}_k\hat{\pi}_{-k} + \frac{m\omega_k^2}{2}\hat{\phi}_k\hat{\phi}_{-k}\right], \qquad (1.29)$$

where $\omega_k = v|k|$ and $v = a\sqrt{k_\mathrm{s}/m}$ denotes the classical sound wave velocity. In this form, the Hamiltonian can be identified as nothing but a superposition of independent **quantum harmonic oscillators**.[19] This result is easy to understand (see the figure). Classically, the system supports a discrete set of wave excitations, each indexed by a wave number $k = 2\pi m/L$. (In fact, we could have performed a Fourier transformation of the classical fields $\phi(x)$ and $\pi(x)$ to represent the Hamiltonian function as a superposition of classical harmonic oscillators.) Within the quantum picture, each of these excitations is described by an oscillator Hamiltonian with a k-dependent frequency. However, it is important not to confuse the atomic constituents, also oscillators (albeit coupled), with the independent *collective* oscillator modes described by $\hat{H}$.
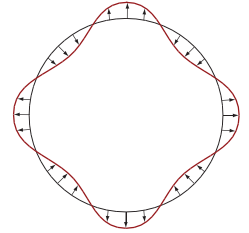
The description above, albeit perfectly valid, still suffers from a deficiency: the analysis amounts to explicitly describing the effective low-energy excitations of the system (the waves) in terms of their microscopic constituents (the atoms). Indeed the different contributions to $\hat{H}$ correspond to details of the microscopic oscillator dynamics of individual k-modes. However, it would be much more desirable to develop a picture where the relevant excitations of the system, the waves, appear as fundamental units without an explicit account of the underlying microscopic details. (As with hydrodynamics, information is encoded in terms of collective density variables rather than through individual atoms.) As preparation for the construction of this improved formulation, let us temporarily focus on a single oscillator mode.

1.4.1 Revision of the quantum harmonic oscillator

harmonic oscillator

Consider a standard **harmonic oscillator** (HO) Hamiltonian

$$\hat{H} = \frac{\hat{p}^2}{2m} + \frac{m\omega^2}{2}\hat{x}^2.$$

[19] The only difference between Eq. (1.29) and the canonical form of an oscillator Hamiltonian $\hat{H} = \hat{p}^2/2m + m\omega^2\hat{x}^2/2$ is the presence of the subindices k and $-k$ (a consequence of the relation, $\hat{\phi}_k^\dagger = \hat{\phi}_{-k}$). As we will show shortly, this difference is inessential.

The first few energy levels $\epsilon_n = \omega(n + 1/2)$ and the associated Hermite polynomial eigenfunctions are displayed schematically in the figure. (To simplify the notation we henceforth set $\hbar = 1$.) The HO has the status of a single-particle problem. However, the equidistance of its energy levels suggests an alternative interpretation: a given state ϵ_n may be thought of as an accumulation of n elementary entities, or **quasi-particles**, each having energy ω. What can be said about the features of these new objects? First, they are structureless, i.e., the only "quantum number" identifying the quasi-particles is their energy ω (since otherwise n-particle states formed of the quasi-particles would not be equidistant in energy). This implies that the quasi-particles must be *bosons*. (The same state ω can be occupied by more than one particle.)

quasi-particles
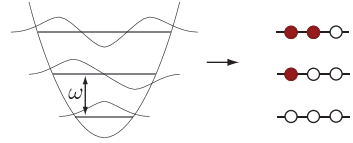
This idea can be formulated in quantitative terms by employing the formalism of ladder operators, in which the operators $\hat{p}$ and $\hat{x}$ are traded for the pair of hermitian adjoint operators $\hat{a} \equiv \sqrt{m\omega/2}(\hat{x} + (i/m\omega)\hat{p})$ and $\hat{a}^\dagger \equiv \sqrt{m\omega/2}(\hat{x} - (i/m\omega)\hat{p})$. Up to a factor of i, the transformation $(\hat{x}, \hat{p}) \to (\hat{a}, \hat{a}^\dagger)$ is canonical, i.e., the new operators obey the canonical commutation relation

$$[\hat{a}, \hat{a}^\dagger] = 1. \tag{1.30}$$

More importantly, the a-representation of the Hamiltonian is very simple, viz.

$$\hat{H} = \omega\big(\hat{a}^\dagger \hat{a} + 1/2\big), \tag{1.31}$$

as can be checked by direct substitution. Suppose, now, that we had been given a zero-eigenvalue state $|0\rangle$ of the operator $\hat{a}$: $\hat{a}|0\rangle = 0$. As a consequence, $\hat{H}|0\rangle = (\omega/2)|0\rangle$, i.e., $|0\rangle$ is identified as the ground state of the oscillator.[20] The hierarchy of higher-energy states can then be generated by setting $|n\rangle \equiv (1/\sqrt{n!})\, (\hat{a}^\dagger)^n |0\rangle$.

EXERCISE Using the canonical commutation relation (1.30), verify that $\hat{H}|n\rangle = \omega(n + 1/2)|n\rangle$ and $\langle n|n\rangle = 1$.

Formally, the construction above represents yet another way of constructing eigenstates of the quantum HO. However, its real advantage is that it naturally affords a many-particle interpretation. To this end, let us declare that $|0\rangle$ represents a "vacuum" state, i.e., a state with zero particles. Next, imagine that $\hat{a}^\dagger |0\rangle$ is a state with a single featureless particle (the operator $\hat{a}^\dagger$ does not carry any quantum number labels) of energy ω. Similarly, $(\hat{a}^\dagger)^n |0\rangle$ is considered as a many-body state with n particles; i.e., within the new picture, $\hat{a}^\dagger$ is an operator that "creates" particles. The total energy of these states is given by $\omega \times$ (occupation number). Indeed, it is

[20] Switching to a real space representation of the ground state equation, verify that its solution is the familiar ground state wave function $\langle x|0\rangle = \sqrt{m\omega/2\pi}e^{-m\omega x^2/2}$. can be verified by explicit construction. As an exercise, switching to a real space representation of the ground state equation, $0 = \langle 0|x\rangle[(m\omega)/\partial_x \varrho + x]$ and verify that its solution is the familiar ground state wave function $\langle x|0\rangle = \sqrt{m\omega/2\pi}e^{-m\omega x^2/2}$.

straightforward to verify (see the exercise above) that $\hat{a}^\dagger \hat{a}|n\rangle = n|n\rangle$, i.e., the Hamiltonian effectively counts the number of particles in the state. While at first sight, this may look unfamiliar, the new interpretation is internally consistent. Moreover, it achieves what we asked for above: it allows an interpretation of the HO states as a superposition of independent structureless entities.

INFO The representation above shows that we can think about individual quantum problems in **complementary pictures**. This principle finds innumerable applications in modern condensed matter physics. The existence of different interpretations of a given system is by no means heterodox but, rather, reflects a principle of quantum mechanics: there is no "absolute" system that underpins the phenomenology. The only thing that matters is observable phenomena. For example, we will see later that the "fictitious" quasi-particle states of oscillator systems *behave* as "real" particles, i.e., they have dynamics, can interact, can be detected experimentally, etc. From a quantum point of view, these objects can be considered as "real" particles.

1.4.2 Quasiparticle interpretation of the quantum chain

Returning to the oscillator chain, one can transform the Hamiltonian (1.29) to a form analogous to (1.31) by defining the ladder operators[21]
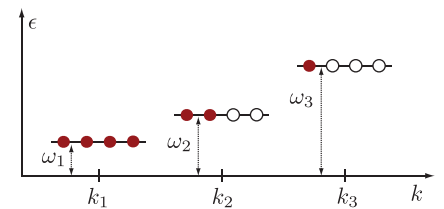
$$\hat{a}_k \equiv \sqrt{\frac{m\omega_k}{2}} \left(\hat{\phi}_k + \frac{i}{m\omega_k} \hat{\pi}_{-k} \right), \quad \hat{a}_k^\dagger \equiv \sqrt{\frac{m\omega_k}{2}} \left(\hat{\phi}_{-k} - \frac{i}{m\omega_k} \hat{\pi}_k \right). \qquad (1.32)$$

With this definition, applying the commutation relations (1.28), one finds that the ladder operators obey commutation relations generalizing Eq. (1.30):

$$[\hat{a}_k, \hat{a}_{k'}^\dagger] = \delta_{kk'}, \qquad [\hat{a}_k, \hat{a}_{k'}] = [\hat{a}_k^\dagger, \hat{a}_{k'}^\dagger] = 0. \qquad (1.33)$$

Expressing the operators $(\hat{\phi}_k, \hat{\pi}_k)$ in terms of $(\hat{a}_k, \hat{a}_k^\dagger)$, it is now straightforward to bring the Hamiltonian into the quasi-particle oscillator form (exercise)

$$\hat{H} = \sum_k \omega_k (\hat{a}_k^\dagger \hat{a}_k + 1/2). \qquad (1.34)$$

Equations (1.34) and (1.33) represent the final result of our analysis. The Hamiltonian $\hat{H}$ takes the form of a sum of harmonic oscillators with characteristic frequencies ω_k. In the limit $k \to 0$ (i.e., long wavelengths), we have $\omega_k \to 0$; excitations with this property are said to be **massless**.

massless excitation

An excited state of the system is indexed by a set $\{n_k\} = (n_1, n_2, \dots)$ of quasi-particles with energy $\{\omega_k\}$ (see the figure). Physically, the quasi-particles of the

---

[21] As to the consistency of these definitions, recall that $\hat{\phi}_k^\dagger = \hat{\phi}_{-k}$ and $\hat{\pi}_k^\dagger = \hat{\pi}_{-k}$. Under these conditions, the second of the definitions in Eq. (1.32) follows from the first upon taking the hermitian adjoint.
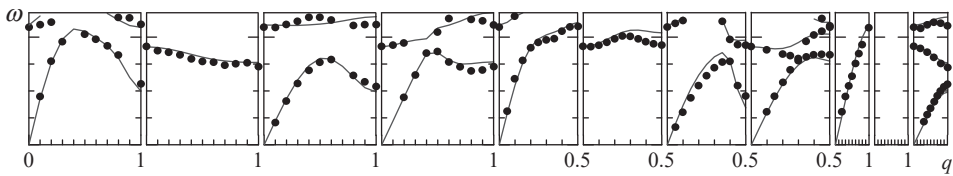
Fig. 1.2   Phonon spectra of the transition metal oxide Sr_2RuO_4 along different axes in momentum space. Notice the approximate linearity of the low-energy branches (**acoustic phonons**) at small momenta q. Superimposed at high frequencies are various branches of **optical phonons**. (Source: Courtesy of M. Braden, University of Cologne.)

phonon

harmonic chain are identified with the **phonon modes** of the solid. A comparison with measured phonon spectra (fig. 1.2) reveals that, at low momenta, we have $\omega_k \sim |k|$ in agreement with our simplistic model (in spite of the fact that the spectrum was recorded for a three-dimensional solid with a nontrivial unit cell – universality!). While the linear dispersion was already a feature of the classical sound wave spectrum, the low-temperature specific heat reflected non-classical behavior. It is left as an exercise (problem 1.8.3) to verify that the quantum nature of the phonons resolves the problem with the low-temperature specific heat discussed in section 1.1.2. (For further discussion of phonon modes in atomic lattices we refer to chapter 2 of the text by Kittel.[22])

EXERCISE Classically, the ground state of the atomic chain comprises a regular array of ions. In the quantum chain, the distance between neighboring ions fluctuates even in the ground state, $|0\rangle$. Using the results above, show that

$$\langle 0|[\phi(x) - \phi(0)]^2|0\rangle = \frac{1}{mL} \sum_k \frac{1 - \cos(kx)}{\omega_k}.$$

In the limit $|x| \gg a$, show that $\langle 0|[\phi(x) - \phi(0)]^2|0\rangle \sim (1/a\sqrt{k_s m}) \ln|x/a|$. What does this imply for the stability of crystalline order in the one-dimensional chain?

1.5 Quantum Electrodynamics

The generality of the procedure outlined above suggests that the quantization of the EM field (1.24) proceeds in a manner analogous to the phonon system. However, there are a number of practical differences that make this task harder (but also more interesting!). First, the vectorial character of the potential, in combination with the condition of relativistic covariance, gives the problem a nontrivial internal geometry. Closely related, the gauge freedom of the vector potential introduces redundant degrees of freedom whose removal on the quantum level is not easily achieved. For example, quantization in a setting where only physical degrees of freedom are kept – i.e., the two polarization directions of the transverse photon field – is technically cumbersome, the reason being that the relevant gauge condition is not relativistically covariant. In contrast, a manifestly covariant

[22] C. Kittel, *Quantum Theory of Solids*, 2nd edition (Wiley, 1987).

scheme, while technically more convenient, introduces spurious "ghost degrees of freedom" that are difficult to remove. To circumvent a discussion of these issues, we will not discuss the problem of EM field quantization in detail.[23] On the other hand, the photon field plays a much too important role in condensed matter physics for us to drop the problem altogether. We will therefore aim at an intermediate exposition, largely insensitive to the problems outlined above, but sufficiently general to illustrate the main principles.

1.5.1 Field quantization

Coulomb gauge

Consider the Lagrangian of the matter-free EM field, $L = -\frac{1}{4} \int d^3x \; F_{\mu\nu}F^{\mu\nu}$. As a first step towards quantization of this system, a gauge choice must be made. In the absence of charge, a particularly convenient choice is the **Coulomb gauge** $\nabla \cdot \mathbf{A} = 0$, with scalar component $\phi = 0$. (Keep in mind that, once a gauge has been set, we cannot expect further results to display "gauge invariance.") Using the gauge conditions, one may verify that the Lagrangian assumes the form

$$L = \frac{1}{2} \int d^3x \left[(\partial_t \mathbf{A})^2 - (\nabla \times \mathbf{A})^2 \right].\qquad(1.35)$$

By analogy with the atomic chain, we would now proceed to "decouple" the theory by expanding the action in terms of eigenfunctions of the Laplace operator. The difference to our previous discussion is that we are dealing (i) with the full three-dimensional Laplacian (instead of a simple second derivative) acting on (ii) the vector quantity $\mathbf{A}$ that is (iii) subject to the constraint $\nabla \cdot \mathbf{A} = 0$. It is these aspects that lead to the complications outlined above.

We can circumvent these difficulties by considering cases where the geometry of the system reduces the complexity of the eigenvalue problem. This restriction is less artificial than it might appear. For example, in anisotropic electromagnetic waveguides, the solutions of the eigenvalue equation can be formulated as[24]

$$-\nabla^2 \mathbf{R}_k(\mathbf{x}) = \lambda_k \mathbf{R}_k(\mathbf{x}),\qquad(1.36)$$

where $k \in \mathbb{R}$ is a *one-dimensional* index and the vector-valued functions $\mathbf{R}_k$ are real and orthonormalized: $\int \mathbf{R}_k \cdot \mathbf{R}_{k'} = \delta_{kk'}$. The dependence of the eigenvalues λ_k on k is governed by details of the geometry (see Eq. (1.38) below) and need not be specified for the moment.

waveguide

INFO An **electromagnetic waveguide** is a quasi-one-dimensional cavity with metallic boundaries (see fig. 1.3). The practical use of waveguides is that they are good at confining EM waves. At large frequencies, where the wavelengths are of order meters or less, radiation loss in conventional conductors is high. In this frequency domain, hollow conductors provide the only practical way of transmitting radiation. Field propagation inside a waveguide is constrained by boundary conditions. Assuming the walls of the system to be perfectly conducting,

$$\mathbf{E}_{\parallel}(\mathbf{x}_b) = 0, \qquad \mathbf{B}_{\perp}(\mathbf{x}_b) = 0,\qquad(1.37)$$

where $\mathbf{x}_b$ is a point at the system boundary and $\mathbf{E}_{\parallel}$ ($\mathbf{B}_{\perp}$) is the parallel (perpendicular) component of the electric (magnetic) field.

[23] Readers interested in learning more about EM field quantization are referred to, e.g., L. H. Ryder, *Quantum Field Theory* (Cambridge University Press, 1996).

[24] More precisely, one should say that Eq. (1.36) defines the set of eigenfunctions relevant for the low-energy dynamics of the waveguide. More-complex eigenfunctions of the Laplace operator exist, but they carry much higher energy.
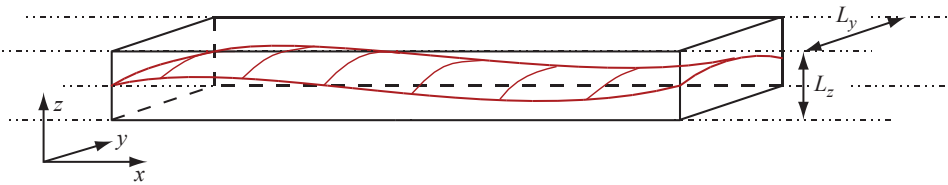
EM waveguide with rectangular cross-section. The structure of the eigenmodes of the field is determined by the boundary conditions at the walls of the cavity.

Regarding the problem of field quantization, let us for concreteness consider a cavity with uniform rectangular cross-section $L_y \times L_z$. To conveniently represent the Lagrangian of the system, we need to express the vector potential in terms of eigenfunctions $\mathbf{R_k}$ that are consistent with the boundary conditions (1.37). A complete set of functions fulfilling this condition is given by

$$\mathbf{R_k} = \mathcal{N}_k \begin{pmatrix} c_1 \cos(k_x x) \sin(k_y y) \sin(k_z z) \\ c_2 \sin(k_x x) \cos(k_y y) \sin(k_z z) \\ c_3 \sin(k_x x) \sin(k_y y) \cos(k_z z) \end{pmatrix}.$$

Here, $k_i = n_i \pi / L_i$ with positive integer n_i, $\mathcal{N}_k$ is a factor normalizing $\mathbf{R_k}$ to unit modulus, and the coefficients c_i are subject to the condition $c_1 k_x + c_2 k_y + c_3 k_z = 0$. Indeed, it is straightforward to verify that a general superposition of the type $\mathbf{A}(\mathbf{x}, t) \equiv \sum_\mathbf{k} \alpha_\mathbf{k}(t) \mathbf{R_k}(\mathbf{x})$, with $\alpha_\mathbf{k}(t) \in \mathbb{R}$, is divergenceless and generates an EM field compatible with (1.37). Substitution of $\mathbf{R_k}$ into Eq. (1.36) identifies the eigenvalues as $\lambda_\mathbf{k} = k_x^2 + k_y^2 + k_z^2$. In the physics and electronic engineering literature, eigenfunctions of the

field modes   Laplace operator in a quasi-one-dimensional geometry are commonly described as **modes**. As we will see shortly, the energy of a mode (i.e., the Hamiltonian evaluated on a specific mode configuration) grows with $|\lambda_\mathbf{k}|$. In cases where one is interested in the low-energy dynamics of the EM field, only configurations with small $|\lambda_\mathbf{k}|$ are relevant. If we consider a massively anisotropic waveguide with $L_z < L_y \ll L_x$, the modes with smallest $|\lambda_k|$ are those with $k_z = 0$, $k_y = \pi / L_y$, and $k_x \equiv k \ll L_{z,y}^{-1}$. (Consider why it is not possible to set both k_y and k_z to zero.) With this choice,

$$\lambda_k = k^2 + (\pi / L_y)^2 \tag{1.38}$$

and a scalar index k suffices to label both eigenvalues and eigenfunctions $\mathbf{R}_k$. A schematic of the spatial structure of the functions $\mathbf{R}_k$ is shown in fig. 1.3. The dynamical properties of these configurations will be discussed in the text.

Returning to the problem posed by Eq. (1.35) and (1.36), one can expand the vector potential in terms of eigenfunctions $\mathbf{R}_k$ as $\mathbf{A}(\mathbf{x}, t) = \sum_k \alpha_k(t) \mathbf{R}_k(\mathbf{x})$, where the sum runs over all allowed values of the index parameter k. (In a waveguide, $k = \pi n / L$ where $n \in \mathbb{N}$ and L is the length of the guide.) Substituting this expansion into Eq. (1.35) and using the normalization properties of $\mathbf{R}_k$, we obtain $L = \frac{1}{2} \sum_k \left(\dot{\alpha}_k^2 - \lambda_k \alpha_k^2 \right)$, i.e., a decoupled representation where the system is described in terms of independent dynamical systems with coordinates α_k. From this point on, quantization proceeds along the lines of the standard algorithm, as follows.

First, define momenta through the relation $\pi_k = \partial_{\dot{\alpha}_k} L = \dot{\alpha}_k$. This yields the Hamiltonian $H = \frac{1}{2} \sum_k (\pi_k \pi_k + \lambda_k \alpha_k \alpha_k)$. Next, quantize the theory by promoting fields to operators, $\alpha_k \to \hat{\alpha}_k$ and $\pi_k \to \hat{\pi}_k$, and declare that $[\hat{\pi}_k, \hat{\alpha}_{k'}] = -i\delta_{kk'}$. The quantum Hamiltonian operator, again of harmonic oscillator type, then reads

$$\hat{H} = \frac{1}{2} \sum_k \left(\hat{\pi}_k \hat{\pi}_k + \omega_k^2 \hat{\alpha}_k \hat{\alpha}_k \right),$$

where $\omega_k^2 = \lambda_k$. Following the same logic as that marshaled in section 1.4.2, we then define ladder operators

$$a_k \equiv \sqrt{\frac{\omega_k}{2}} \left(\hat{\alpha}_k + \frac{i}{\omega_k} \hat{\pi}_k \right), \qquad a_k^\dagger \equiv \sqrt{\frac{\omega_k}{2}} \left(\hat{\alpha}_k - \frac{i}{\omega_k} \hat{\pi}_k \right),$$

whereupon the Hamiltonian assumes the now familiar form

$$\hat{H} = \sum_k \omega_k (a_k^\dagger a_k + 1/2). \tag{1.39}$$

For the specific problem of the first excited mode in a waveguide of width L_y, $\omega_k = [k^2 + (\pi/L_y)^2]^{1/2}$. Equation (1.39) represents our final result for the quantum Hamiltonian of the EM waveguide. Before concluding this section let us make a few comments on the structure of the result.

▷ The construction above parallels almost perfectly our previous discussion of the harmonic chain.[25] The structural similarity between the two systems finds its origin in the fact that the free field Lagrangian (1.35) is quadratic and, therefore, bound to map onto an oscillator-type Hamiltonian. That we obtained a simple *one-dimensional* superposition of oscillators is due to the boundary conditions specific to a narrow waveguide. For less restrictive geometries, e.g., free space, a more complex superposition of vectorial degrees of freedom in three-dimensional space would have been obtained. However, the principle that the free EM field is mapped onto a superposition of oscillators is independent of geometry.

photon ▷ Physically, the quantum excitations described by Eq. (1.39) are, of course, the **photons** of the EM field. The unfamiliar appearance of the dispersion relation ω_k is, again, a peculiarity of the waveguide geometry. However, in the limit of large longitudinal wave numbers, $k \gg L_y^{-1}$, the dispersion approaches the form $\omega_k \sim |k|$, i.e., the relativistic dispersion of the photon field. Also, notice that, owing to the equality of the Hamiltonians (1.34) and (1.39), all that has been said about the behavior of the phonon modes of the atomic chain carries over to the photon modes of the waveguide.

▷ As with their phononic analog, the oscillators described by Eq. (1.39) exhibit zero-point fluctuations. It is a fascinating aspect of quantum electrodynamics that these oscillations, caused by quantization of the ultra-relativistic photon field, have various manifestations in non-relativistic physics:

INFO Without going into detail, let us mention some manifestations of **vacuum fluctuations in the phenomenology of condensed matter systems**. One of the most important phenomena induced by vacuum fluctuations is the **Casimir effect**.[26] Two parallel conducting plates embedded into the vacuum exert an attractive force on each other. This phenomenon is not only of conceptual importance – it demonstrates that the vacuum is "alive" – but also of practical relevance. For example, the force balance of hydrophobic

Casimir effect

[25] Technically, the only difference is that, instead of index pairs $(k, -k)$, all indices (k, k) are equal and positive. This can be traced to the fact that we have expanded in terms of the real eigenfunctions of the closed waveguide instead of the complex eigenfunctions of the circular oscillator chain.

[26] H. B. G. Casimir and D. Polder, *The influence of retardation on the London–van der Waals forces*, Phys. Rev. **73**, 360 (1948); H. B. G. Casimir, *On the attraction between two perfectly conducting plates*, Proc. Kon. Nederland. Akad. Wetensch. **51**, 793 (1948).

suspensions of particles of size $0.1 - 1\,\mu\mathrm{m}$ in electrolytes is believed to be strongly influenced by Casimir forces. Qualitatively, the origin of the Casimir force is readily understood. In common with their classical analog, quantum photons exert a certain radiation pressure on macroscopic media. The difference to the classical case is that, due to zero-point oscillations, even the quantum vacuum is capable of creating radiation pressure. For a single conducting body embedded into the infinite vacuum, the net pressure vanishes by symmetry. However, for two parallel plates, the situation is different. Mode quantization arguments similar to those used in the previous section show that the density of quantum modes between the plates is lower than in the semi-infinite outer spaces. Hence, the force (density) created by outer space exceeds the counter-pressure from the inside; the plates "attract" each other.

Van der Waals forces

A second context where vacuum fluctuations play a role is the physics of **van der Waals forces**. Atoms or molecules attract each other by a potential that, at small separation r, scales as r^{-6}. While a detailed discussion of the unusually high power at which this force decays would lead us too far astray, the essence of the argument is as follows. The zero-point fluctuations of the EM field may induce a dipole moment in atoms, which in turn generate a dipole–dipole interaction between close-by atoms, whose detailed evaluation[27] leads to the r^{-6} power–law dependence. Seen in this way, geckos and spiders owe their ability to climb walls to a deeply microscopic principle of quantum field theory.[28]

1.6 Noether's Theorem

It is a basic paradigm of physics that every continuous symmetry entails a conservation law.[29] Conservation laws, in turn, simplify greatly the solution of any problem, which is why one gets acquainted with the correspondence (symmetry $\leftrightarrow$ conservation law) at a very early stage of the physics curriculum, e.g., the connection between rotational symmetry and the conservation of angular momentum. However, it is not trivial to see (at least within the framework of Newtonian mechanics) that the former entails the latter. One needs to know what to look for (viz. angular momentum) to identify the corresponding conserved quantity (rotational invariance). A major advantage of Lagrangian over Newtonian mechanics is that it provides a tool – Noether's theorem – to automatically identify the conservation laws generated by the symmetries of classical mechanics.

What happens when one advances from point to continuum mechanics? Clearly, multidimensional continuum theories leave more room for the emergence of complex symmetries but, even more so than in classical mechanics, we are in need of a tool to identify the corresponding conservation laws.

---

[27] P. W. Milonni, *The Quantum Vacuum* (Academic Press, 1994).

[28] The feet of geckos and spiders are covered with bushels of ultra-fine hair (about three orders of magnitude thinner than human hair). The tips of these hairs come close enough to the atoms of the substrate material to make the van der Waals force sizable. Impressively, this mechanism provides a force of about two orders of magnitude larger than that required to support a spider's full body weight. Both spiders and geckos have to "roll" their feet off the surface to prevent getting stuck by the enormous power of the forces acting on their many body hairs!

[29] Before exploring the ramifications of symmetries and conservation laws for fields, it may be instructive to recapitulate Noether's theorem in the context of classical point-particle mechanics – see, e.g., L. D. Landau and E. M. Lifshitz, *Classical Mechanics* (Pergamon, 1960).

Fortunately, it turns out that Noether's theorem of point mechanics affords a more or less straightforward generalization to higher dimensions. Starting from the general form of the action of a continuum system, Eq. (1.16), the continuum version of Noether's theorem will be derived below. In that we do not refer to a specific physical problem, our discussion will be somewhat dry. This lack of physical context is, however, more than outweighed by the general applicability of the result. The generalized form of Noether's theorem can be – without much further thought – applied to generate the conservation laws of practically any physical symmetry. In this section, we will illustrate the application of the formalism on the simple (yet important) example of space–time translational invariance. A much more intriguing case study will be presented in section 3.6 after some further background of quantum field theory has been introduced.

Amalie E. Noether **1882–1935** was a German mathematician known for her ground-breaking contributions to abstract algebra and theoretical physics. Alive at a time when women were not supposed to attend college preparatory schools, she was often forbidden from lecturing under her own name. Despite these obstacles, Noether became one of the greatest algebraists of the century. Described by Albert Einstein as the most significant creative mathematical genius thus far produced since the higher education of women began, she revolutionized the theories of rings, fields and algebras. In physics, Noether's theorem explains the fundamental connection between symmetry and conservation laws. In 1933, she lost her teaching position owing to her being a Jew and a woman, and was forced out of Germany by the Nazis.

1.6.1 Symmetry transformations

The symmetries of a physical system are manifest in the invariance of its action under certain transformations. Mathematically, symmetry transformations are described by two pieces of input data: first, a mapping $f : M \to M$, $x \mapsto f(x) \equiv x'(x)$ that assigns to any point of the base manifold some "transformed" point; second, the field configurations themselves may undergo some change, i.e., there may be a mapping $(\phi : M \to T) \mapsto (\phi' : M \to T)$ that defines a transformed "new field" ϕ' in terms of the "old" ϕ. In principle, there is unlimited freedom in defining such transformations. However, for most applications it is sufficient to consider

$$\phi'(x') = F(\phi(x)), \qquad (1.40)$$

where F is a function: the new field in the transformed space–time coordinates is obtained as a function of the old field at the original coordinates. With $x' = f(x)$, this correspondence may be equivalently represented as $\phi'(x) = F(\phi(f^{-1}(x)))$. However, irrespective of the representation, it is important to understand that the two operations, $x \mapsto x'$ and $\phi \mapsto \phi'$ may, in general, be independent of each other. The working of such transformations is best illustrated on a few examples:

The invariance properties of a theory under translations in space–time are probed by the mapping $x' = x + a$, $a = $ const., $\phi'(x') = \phi(x)$. This describes the translation of a field by a fixed offset a in space–time (see the figure). The system is translationally invariant if $S[\phi] = S[\phi']$ for all fields ϕ. As a second example, let us probe the rotational
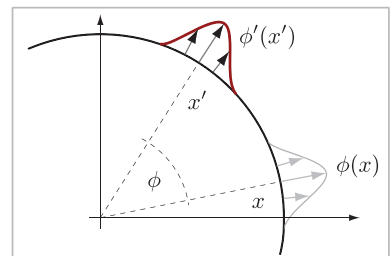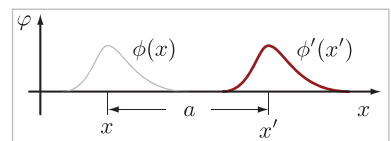
symmetry, $x' = Rx$, where $R \in \mathrm{O}(m)$ is a rotation of Euclidean space–time. In this case it would, in general, be unphysical to define $\phi'(x') = \phi(x)$. To illustrate this point, consider the example of a vector field in two dimensions $n = m = 2$ (see the figure). A properly rotated field configuration is defined by $\phi'(x') = R\phi(x)$, i.e., the field amplitude actively participates in the operation. In fact, one does often consider symmetry operations where only the fields are transformed while the base manifold is left untouched.[30] For example, the intrinsic[31] rotational invariance of a magnet is revealed by setting $x' = x$, $m'(x) = R \cdot m(x)$, where the vector field m describes the local magnetization. Conversely, a scalar field $\phi \in \mathbb{R}$ will transform as $\phi(x') = \phi(x)$. These examples show how the extrinsic effects of rotation, $x \mapsto Rx$, and the intrinsic effects, $\phi \mapsto R\phi$, may appear in all sorts of combinations.

To understand the consequences of a symmetry transformation, it is sufficient to consider its infinitesimal version. (Note that any finite transformation can be generated by successive application of infinitesimal transformations.) Consider the two mappings

$$
\begin{aligned}
x^\mu \to x^{\mu\prime} &= x^\mu + \partial_{\omega_a} x^\mu|_{\omega=0} \omega_a(x), \\
\phi^i(x) \to \phi'^i(x') &= \phi^i(x) + \omega_a(x) F^i_a(\phi(x)),
\end{aligned}
\tag{1.41}
$$

expressing the change of fields and coordinates to first order in a set of parameter functions ω_a characterizing the transformation. (For a three-dimensional rotation, $(\omega_1, \omega_2, \omega_3) = (\phi, \theta, \psi)$ would be the rotation angles, etc.) The functions F^i_a – which need not depend linearly on the field ϕ, and may explicitly depend on the coordinate x – *define* the incremental change $\phi'(x') - \phi(x)$.

We now ask how the action (1.16) changes under the transformation (1.41), i.e., we wish to compute the difference

$$
\Delta S = \int dx' \; \mathcal{L}(\phi'^i(x'), \partial_{x^{\mu\prime}} \phi'^i(x')) - \int dx \; \mathcal{L}(\phi^i(x), \partial_{x^\mu} \phi^i(x)),
$$

where dx is a shorthand for the integration measure over m coordinates x. Inserting Eq. (1.41), using the identity $\partial_{x^\nu} x'^\mu = \delta^\mu_\nu + \partial_{x^\nu}(\omega_a \partial_{\omega_a} x^\mu)$, together with[32] the Jacobian matrix $\det(\partial x'/\partial x) = 1 + \partial_{x^\mu}(\omega_a \partial_{\omega_a} x^\mu) + \mathcal{O}(\omega^2)$, one obtains

$$
\begin{aligned}
\Delta S \simeq & \int dx \,(1 + \partial_{x^\mu}(\omega_a \partial_{\omega_a} x^\mu)) \mathcal{L}((\phi^i + \omega)_a F^i_a, (\delta_\mu{}^\nu - \partial_{x^\mu}(\omega_a \partial_{\omega_a} x^\nu)) \partial_{x_\nu}(\phi^i + \omega_a F^i_a)) \\
& - \int dx \, \mathcal{L}(\phi^i(x), \partial_{x_\mu} \phi^i(x)).
\end{aligned}
$$

So far, we did not use the fact that the transformation was actually meant to be a symmetry transformation. By definition, we are dealing with a symmetry if, for constant parameters ω_a – a uniform rotation or global translation, etc. – the action difference ΔS vanishes. In other words, we may ignore terms in the expansion of ΔS which do not contain derivatives acting on ω_a, as they will not be present in the case where ω_a parameterizes a symmetry. The straightforward expansion of ΔS to leading order in $\partial_\mu \omega_a$ then leads to

$$
\Delta S \overset{\text{sym.}}{=} - \int dx \; j^{a\mu}(x) \partial_\mu \omega_a(x),
\tag{1.42}
$$

---

[30] For example, the standard symmetry transformations of classical mechanics, $q(t) \to q'(t)$, belong to this class: the coordinate vector of a point particle, q (a "field" in $0 + 1$ space–time dimensions) changes while the "base" (time t) does not.

[31] "Intrinsic" means that we rotate just the spins but not the entire system (as we did in our second example, rotational symmetry).

[32] Note that $\det(\partial x'/\partial x) = \exp \mathrm{tr} \ln(\partial x'/\partial x) \simeq \exp[\partial_{x^\mu}(\omega_a \partial_{\omega_a} x^\mu)] \simeq 1 + \partial_{x^\mu}(\omega_a \partial_{\omega_a} x^\mu)$. (Exercise: Show that $\det A = \exp \mathrm{tr} \ln A$, where A is a linear operator.)

Noether current

where the components of the so-called **Noether currents** j^a are given by

$$j^{a\mu} = \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi^i)}\partial_\nu \phi^i - \mathcal{L}\,\delta^\mu_{\ \nu}\right)\frac{\partial x^\nu}{\partial \omega_a}\bigg|_{\omega=0} - \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi^i)}F_a^i(\phi) \tag{1.43}$$

For a general field configuration, not much can be said about the Noether current (no matter whether or not the theory possesses a symmetry). However, if the field ϕ obeys the classical equations of motion *and* the theory is symmetric, the Noether current is locally conserved,

$$\partial_\mu j^{a\mu} = 0 \tag{1.44}$$

This follows from the fact that, for a solution ϕ of the Euler–Lagrange equations, the linear variation of the action in any parameter must vanish. Specifically, integration by parts in Eq. (1.42) leads to $\Delta S = \Delta S[\phi] = \int (\partial_\mu j^{a\mu})\omega_a$. The vanishing of this expression for arbitrary solutions ϕ and arbitrary ω requires Eq. (1.44). (As an exercise in partial differentiation, try to derive this identity directly from Eq. (1.43). You will need to use the Euler–Lagrange equations Eq. (1.17).) It is very important to keep in mind that the conservation law holds only for solutions of the equations of motion. Therefore, in summary, we have Noether's theorem:

> A continuous symmetry entails a classically conserved current.

We call the current "classically conserved" because, as we will discuss later, in section 9.2, quantum fluctuations around classical solutions may spoil the conservation of currents via

anomaly

the so-called **quantum anomaly**.

The local conservation of a current entails the existence of a globally conserved "charge." For a theory with $d + 1$ space–time coordinates $x = (x^0, x^i) = (t, x^i)$, integration over the space-like directions, and application of Stokes' theorem (exercise), gives $d_t Q^a = 0$, where[33]

$$Q^a(t) \equiv \int d^d x \; j^{a0}(t, x^i) \tag{1.45}$$

conserved charge

is the **conserved charge** and we have assumed that the current density vanishes at spatial infinity.

Notice that nowhere in the discussion above have we made any assumption about the internal structure of the Lagrangian. In particular, all results apply equally to the Minkowskian and the Euclidean formulations of the theory.

1.6.2 Examples of symmetries

translational invariance

Condensed matter systems are often translationally invariant, in space and/or in time. **Translational invariance** may hold down to the microscopic level, where it assumes the form of a discrete symmetry under translation by multiples of the lattice spacing, or it may be emergent only at larger length scales. For example, the fluctuating spin configurations of a paramagnet look locally random, however the system becomes translationally invariant on average over mesoscopic volumes containing many spins. In either case, translational invariance appears as a continuous symmetry of the effective theories relevant to the low-energy physics.

The corresponding symmetry transformation is defined by $x'^\mu = x^\mu + a^\mu$, $\phi'(x') = \phi(x)$. The infinitesimal version of this transformation reads $x'^\mu = x^\mu + \omega^\mu$, where we have

---

[33] Notice that the integral involved in the definition of Q runs only over spatial coordinates.

identified the parameter index a with the space–time index μ. Noether's current, which in the case of translational invariance is called the **energy–momentum tensor** or **stress–energy tensor**, is given by $T^\mu_{\ \nu}$:

$$T^\mu_{\ \nu} = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi^i)} \partial_\nu \phi^i - \delta^\mu_{\ \nu} \, \mathcal{L} \tag{1.46}$$

The conserved "charges" corresponding to this quantity are

$$P_\nu \equiv \int d^d x \left(\frac{\partial \mathcal{L}}{\partial(\partial_0 \phi^i)} \partial_\nu \phi^i - \delta^0_{\ \nu} \, \mathcal{L} \right),$$

where P_0 is the energy and P_i, $i = 1, \ldots, d$, the total momentum carried by the system.
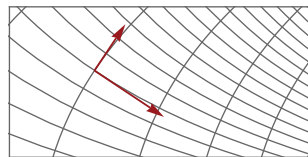
EXAMPLE Evaluation of the zeroth component $T^0_{\ 0}$ for the Lagrangian (1.4) of the harmonic chain yields

$$T^0_{\ 0} = \frac{m}{2} \dot\phi^2 + \frac{k_s a^2}{2} (\partial_x \phi)^2,$$

which is identical to the Hamiltonian density of Eq. (1.11), with $\pi = m\dot\phi$. For a discussion of the momentum density of the chain and of the energy–momentum tensor of the electromagnetic field we refer to problem 1.8.4.

Systems positioned at the critical point of a second-order phase transition are **scale invariant**. Here, the system looks the same at all length scales, a feature formally expressed as symmetry under *dilatation*, $x \to \lambda x$. The ramifications of this symmetry in field theories will be the central theme of chapter 6. However, at the critical point, systems generally show an even larger set of sym-

metries, known as **conformal symmetries**. By definition, conformal transformations of space–time are angle-preserving in that they map curves intersecting at a certain angle onto image curves intersecting at the same angle. For example, the figure shows the conformal image of a rectangular grid. Besides dilatations, *translations* and *rotations* have this feature. The final, and less obviously angle-preserving, representatives of conformal transformations in general dimensions are the *special conformal transformations* $x^\mu \to (x^\mu - b^\mu x^2)/(1 - 2x_\mu b^\mu + b^2 x^2)$. Geometrically, these are a composition of inversion $x^\mu \to x^\mu/x^2$ followed by translation by b^μ and then by another inversion. The set of

all these transformations defines the **conformal group**, a finite-dimensional symmetry group. (Exercise: How many parameters define the group?[34])

Where they exist, conformal symmetries have far-reaching consequences for the physical properties of a theory. This principle is driven to an extreme in the special and important case of two-dimensional conformal invariance (i.e., the physics of critical two-dimensional systems). The reason is that the two-dimensional conformal group is actually infinite dimensional. Referring to appendix section A.3 for a discussion of two-dimensional conformal invariance, here we note only that the existence of infinitely many symmetries, and as many conserved currents, suffices to almost fully characterize two-dimensional conformal theories. The mechanisms by which this happens are discussed in the appendix, which is perhaps best read at a later stage after more concepts of field theory have been introduced.

Translational and conformal symmetry are examples of space–time symmetries. Later,

[34] $1 + p + p(p-1)/2 + 1 = p + p + (p+1)(p+2)/2$ for dilation, translation, rotations, and special transformation, respectively.

when we have introduced field manifolds of richer geometric structure, we will encounter numerous realizations of **internal symmetries**.

1.7 Summary and Outlook

In this chapter we have introduced the general procedure whereby classical continuum theories are quantized. Employing the elementary harmonic oscillator as a example, we have seen that the Hilbert spaces of these theories afford different interpretations. Of particular use is a quasi-particle picture in which the collective excitations of the continuum theories acquired the status of elementary particles. Both examples discussed in this text, the quantum harmonic chain and free quantum electrodynamics, lead to exactly solvable **free field theories**. However, it takes little imagination to foresee that few continuum theories will be as simple. Indeed, the exact solvability of the atomic chain would have been lost had we included higher-order contributions in the expansion in powers of the lattice displacement. Such terms would hinder the free wave-like propagation of the phonon modes. Put differently, phonons would begin to scatter, i.e., interact. Similarly, the free status of electrodynamics is lost once the EM field interacts with a matter field. Needless to say, **interacting field theories** are much more complex, but also more interesting, than the systems considered so far.

Technically, we have seen that the phonon or photon interpretation of the field theories discussed in this chapter could be conveniently formulated in terms of ladder operators. However, the applications discussed so far provide only a glimpse of the advantages of this language. In fact, the formalism of ladder operators, commonly described as "second quantization," represents a central, and historically the oldest, element of quantum field theory. The next chapter is devoted to a more comprehensive discussion of both the formal aspects and applications of this formulation.

1.8 Problems

1.8.1 Electrodynamics from a variational principle

Choosing the Lorentz-gauged components of the vector potential as generalized coordinates, the aim of this problem is to show how the wave equations of electrodynamics can be obtained as a variational principle.

Lorentz gauge

Electrodynamics can be described by Maxwell's equations or, equivalently, by wave-like equations for the vector potential. Working in the **Lorentz gauge**, $\partial_t \phi = -\nabla \cdot \mathbf{A}$, these equations read $(\partial_t^2 - \nabla^2)\phi = \rho$, $(\partial_t^2 - \nabla^2)\mathbf{A} = \mathbf{j}$. Using relativistically covariant notation, the form of the equations can be compressed further to $\partial_\mu \partial^\mu A^\nu = j^\nu$. Starting from the action, $S[A] = -\int d^4x(\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + A_\mu j^\mu)$, obtain these equations by applying the variational principle. Compare the Lorentz gauge representation of the action with that of the elastic chain. What are the differences and parallels?

Answer:

Substituting the EM field tensor $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ and integrating by parts, the action assumes the form

$$S[A] = -\int d^4x \left(-\frac{1}{2} A_\nu \left[\partial_\mu \partial^\mu A^\nu - \partial_\mu \partial^\nu A^\mu \right] + j_\mu A^\mu \right).$$

Owing to the Lorentz gauge condition, the second contribution within the square brackets vanishes, and we obtain $S[A] = -\int d^4x(\frac{1}{2}\partial_\mu A_\nu \partial^\mu A^\nu + j_\mu A^\mu)$, where we have again integrated by parts. Applying the general variational equation (1.17), one obtains the wave equation.

1.8.2 Hamiltonian of electromagnetic field

Here, it is shown that the Hamiltonian canonically conjugate to the Lagrangian of the EM field does indeed coincide with the energy density familiar from elementary electrodynamics.

Consider the EM field in the absence of matter, $j = 0$. Verify that the total energy stored in the field is given by $H \equiv \int d^3x\, \mathcal{H}(\mathbf{x})$ where $\mathcal{H}(\mathbf{x}) = \mathbf{E}^2(\mathbf{x}) + \mathbf{B}^2(\mathbf{x})$ is the familiar expression for the EM energy density. (Hint: Use the vacuum form of Maxwell's equations and the fact that, for an infinite system, the energy is defined only up to surface terms.)

Answer:

Following the canonical prescription, let us first consider the Lagrangian density

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} = \frac{1}{2}\sum_{i=1}^{3}(\partial_0 A_i - \partial_i A_0)^2 - \frac{1}{4}\sum_{i,j=1}^{3}(\partial_i A_j - \partial_j A_i)^2.$$

We next determine the components of the canonical momentum through the relation $\pi_\mu = \partial_{\partial_0 A^\mu}\mathcal{L} : \pi_0 = 0,\ \pi_i = \partial_0 A_i - \partial_i A_0 = -E_i$. Using the fact that $\partial_i A_j - \partial_j A_i$ is a component of the magnetic field, the Hamiltonian density can now be written as

$$\mathcal{H} = \pi_\mu \partial_0 A^\mu - \mathcal{L} = \frac{1}{2}(-2\mathbf{E}\cdot\partial_0\mathbf{A} - \mathbf{E}^2 + \mathbf{B}^2) \overset{(1)}{=} \frac{1}{2}(2\mathbf{E}\cdot\nabla\phi + \mathbf{E}^2 + \mathbf{B}^2)$$
$$\overset{(2)}{=} \frac{1}{2}(2\nabla\cdot(\mathbf{E}\phi) + \mathbf{E}^2 + \mathbf{B}^2),$$

where equality (1) is based on addition and subtraction of a term $2\mathbf{E}\cdot\nabla\phi$ and equality (2) on the relation $\nabla\cdot\mathbf{E} = 0$ combined with the identity $\nabla\cdot(\mathbf{a}f) = \nabla\cdot\mathbf{a}f + \mathbf{a}\cdot\nabla f$ (valid for general vector [scalar] functions $\mathbf{a}$ [f]). Substitution of this expression into the definition of the Hamiltonian yields

$$H = \frac{1}{2}\int d^3x\left(2\nabla\cdot(\mathbf{E}\phi) + \mathbf{E}^2 + \mathbf{B}^2\right) = \frac{1}{2}\int d^3x\left(\mathbf{E}^2 + \mathbf{B}^2\right),$$

where we have used the fact that the contribution $\nabla\cdot(\mathbf{E}\phi)$ is a surface term that vanishes upon integration by parts.

1.8.3 Phonon specific heat

Previously, we stated that the mode quantization of elastic media manifests itself in low-temperature anomalies of the specific heat. In this problem, concepts of elementary quantum statistical mechanics are applied to determine the temperature profile of the specific heat.

Compute the energy density $u = -L^{-1}\partial_\beta \ln \mathcal{Z}$ of one-dimensional longitudinal phonons with dispersion $\omega_k = v|k|$, where $\mathcal{Z} = \mathrm{tr}\, e^{-\beta\hat{H}}$ denotes the quantum partition function. First show that the thermal expectation value of the energy density can be represented as

$$u = \frac{1}{L}\sum_k\left[\frac{\omega_k}{2} + \omega_k n_\mathrm{B}(\omega_k)\right], \tag{1.47}$$

where $n_{\mathrm{B}}(\epsilon) = (e^{\beta\epsilon} - 1)^{-1}$ is the Bose–Einstein distribution. Approximate the sum over k by an integral and show that the specific heat $c_v \equiv \partial_T u \sim T$. At what temperature T_{cl} does the specific heat cross over to the classical result, $c_v = \mathrm{const}$? (Remember that the linear dispersion $\omega_k = v|k|$ is based on a quadratic approximation to the Hamiltonian and, therefore, holds only for $|k| < \Lambda$, where Λ is some cutoff momentum.) Recalling the discussion in section 1.4, for a d-dimensional isotropic solid of volume L^d (with the atomic exchange constants remaining the same in all directions), show that the dispersion generalizes to $\omega_{\mathbf{k}} = v|\mathbf{k}|$, where $\mathbf{k} = 2\pi(n_1, \ldots, n_d)/L$ and $n_i \in \mathbb{Z}$. Show that the specific heat shows the temperature dependence $c_v \sim T^d$.

Answer:

As discussed in the text, the eigenstates of the system are given by $|n_1, n_2, \ldots\rangle$, where n_m is the number of phonons of wavenumber $k_m = 2\pi m/L$, $E_{|n_1, n_2, \ldots\rangle} = \sum_m \omega_{k_m}(n_m + 1/2) \equiv \sum_m \epsilon_m^{n_m}$ the eigenenergy, and $\omega_m = v|k_m|$. In the energy representation, the quantum partition function then takes the form

$$\mathcal{Z} = \mathrm{tr}\, e^{-\beta\hat{H}} = \sum_{\text{states}} e^{-\beta E_{\text{state}}} = \prod_{m=1,2,\ldots} \sum_{n_m=0}^{\infty} e^{-\beta\omega_m(n_m+1/2)} = \prod_{m=1,2,\ldots} \frac{e^{-\beta\omega_m/2}}{1 - e^{-\beta\omega_m}},$$

where n_m is the occupation number of the state with wavenumber k_m. Hence, $\ln \mathcal{Z} = -\sum_m [\beta\omega_m/2 + \ln(1 - e^{-\beta\omega_m})]$. Differentiation with respect to β yields Eq. (1.47) and, making the replacement $\sum_m \to \frac{L}{2\pi} \int dk$, we arrive at $u = C_1 + \frac{1}{2\pi} \int_{|k|<\Lambda} dk\, \frac{v|k|}{e^{\beta v|k|} - 1} = C_1 + \beta^{-2} C_2$, where C_1 is the temperature-independent constant accounting for the "zero-point energies" $\omega_m/2$. In the second equality, we have scaled $k \to \beta k$. This produces a prefactor β^{-2} multiplied by a temperature-independent (up to the temperature dependence of the boundaries $\Lambda \to \beta\Lambda$) integral that we denoted by C_2. Differentiation with respect to T then leads to the relation $c_v = \partial_T u \sim T$. However, for temperatures $T > v\Lambda$ higher than the highest frequencies stored in the phonon modes, the procedure above no longer makes sense (formally, owing to the now non-negligible temperature dependence of the boundaries). Yet, in this regime, we may expand $e^{\beta v|k|} - 1 \simeq \beta v|k|$, which brings us back to the classical result $c_v = \mathrm{const}$.

Consider now a d-dimensional solid with isotropic coupling, $\frac{k_{\mathrm{s}}}{2} \sum_{i=1}^{d} (\phi_{\mathbf{R}+\mathbf{e}_i} - \phi_{\mathbf{R}})^2$ with $\mathbf{e}_i$ a unit vector in the direction i. Taking the continuum limit leads to a contribution $\frac{k_{\mathrm{s}} a^2}{2} (\nabla\phi(\mathbf{x}))^2$. Proceeding as in the one-dimensional system, the relevant excitations are now waves with wavevector $\mathbf{k} = 2\pi(n_1, \ldots, n_d)/L$ and energy $\omega_{\mathbf{k}} = v|\mathbf{k}|$. Setting $\sum_{\mathbf{k}} \sim \int d^d k$ and scaling $k_i \to \beta k_i$ then generates a prefactor $\beta^{-(d+1)}$, and we arrive at the relation $c_v \sim T^d$.

1.8.4 Energy–momentum tensor of the harmonic chain

In this problem we analyze the energy–momentum (EM) tensor of the harmonic chain. We discuss its computation and how to make sense of its components.

(a) Show that the two independent components $T^0{}_0$ and $T^1{}_0$ of the EM tensor of the harmonic chain defined via the Lagrangian (1.4) are given by

$$T^0{}_0 = \frac{m}{2}\dot{\phi}^2 + \frac{k_s a^2}{2}(\partial_x\phi)^2, \qquad T^0{}_1 = m\,\partial_t\phi\,\partial_x\phi.$$

(b) In section 1.6.2 we identified $T^0{}_0$ as the energy density of the system. But what is the meaning of the second component? Turning back to the discrete representation of the

chain, compute the total momentum carried by weak dynamical fluctuations $\phi_I(t)$ of the mass center coordinates and show that it turns into an integral over $T^0_{\ 1}$ in the continuum limit. This construction identifies $T^0_{\ 1}$ as the momentum density of the chain.

Answer:

(a) This part involves a straightforward application of Eq. (1.46). **(b)** We can consider the total momentum of the chain as $P = \sum_I a \times \delta(\text{mass density}) \times \dot{\phi}_I$, where $\delta(\text{mass density})$ are the fluctuations in mass density associated with a deviation profile ϕ_I. The local particle density at site I is given by (one particle) / (distorted particle distance), i.e., $1/(a - \phi_{I+1} + \phi_I) \simeq a^{-1} + a^{-2}(\phi_{I+1} - \phi_I) \simeq a^{-1} + a^{-1/2}\partial_x\phi$, where we used the definition of the continuum variable $\phi(x) = a^{-1/2}\phi_I$. This leads to $\delta(\text{mass density}) \simeq m a^{-1/2}\partial_x\phi$. With the particle velocity $\dot{\phi}_I = a^{1/2}\partial_t\phi(x,t)$, we obtain $P = \sum_I a\, m\, \partial_x\phi\, \partial_t\phi \simeq \int dx\, m\, \partial_x\phi\, \partial_t\phi$.

1.8.5 Stress–energy tensor from variation in metric

This problem is for advanced readers. It requires familiarity with integration over manifolds of nontrivial geometry, as reviewed in section A.1, and fluency in variational calculus. Other readers should not tackle this problem just yet. We offer an interpretation of the stress–energy tensor generalizing that given in section 1.6.2: the stress tensor describing how a field theory responds to variations in the underlying geometry.

In section 1.6.2 we derived the stress–energy tensor by investigating how a theory changes under variations $x^\mu \to x^\mu + \omega^\mu(x)$, where the infinitesimal shift may be coordinate dependent. Such deviations describe a local distortion in the geometry of the base manifold. To substantiate this view, consider a situation where the base manifold has a nontrivial geometry, described by a metric tensor $g = \{g_{\mu\nu}\}$. For example, in the field theories of gravity, the base manifold is the universe, and $g_{\mu\nu}$ is its space–time metric. A more mundane example would be a field theory formulated in curvilinear coordinates, where $g_{\mu\nu}$ is the (square of the) Jacobian describing the transformation from Cartesian coordinates.

The generalization of the Lagrangian Eq. (1.16) to this case is given by

$$S[\phi] = \int dx \sqrt{g}\, \mathcal{L}(\phi, \partial_\mu, \partial^\mu\phi),$$

where $g = |\det(g_{\{\mu\nu\}})|$, and the notation emphasizes that derivatives in the Lagrangian appear in invariant combinations such as $\partial^\mu\phi\partial_\mu\phi$. Their dependence on the metric is hidden in $\partial^\mu\phi = g^{\mu\nu}\partial_\nu\phi$, where $g^{\mu\nu}$ are the coefficients of the inverse of the metric tensor, $g^{\mu\nu}g_{\nu\lambda} = \delta^\mu_{\ \lambda}$. (We have omitted the internal field index ϕ^i to lighten the notation.)
(a) Prove the auxiliary relations $\partial g^{\rho\sigma}/\partial g_{\mu\nu} = -\delta^{\rho\mu}\delta^{\nu\sigma}$, $\partial_{g_{\mu\nu}}\sqrt{g} = \frac{1}{2}\sqrt{g}g^{\mu\nu}$, and $\partial F/\partial(\partial_\mu\phi) = (\partial F/\partial(\partial^\nu\phi))g^{\nu\mu}$.
(b) Show that the stress tensor is obtained by variation of the action in the metric:

$$T^{\mu\nu}(x) = -\frac{2}{\sqrt{g}}\frac{\delta S}{\delta g_{\mu\nu}(x)}. \tag{1.48}$$

(c) As an example, consider the theory of a free scalar field, $\mathcal{L} = -\frac{1}{2}\partial_\mu\phi\partial^\mu\phi$. Compute the stress tensor via Eq. (1.48) and convince yourself that the result is compatible with that of the example below Eq. (1.46) for the harmonic chain Eq. (1.4) in the case where the differentiation is carried out on the two-dimensional Minkowski metric, $g = \text{diag}(-1, 1)$, and the constants are scaled as $m = k_s a^2 = 1$.

Conceptually, Eq. (1.48) demonstrates that the stress tensor answers the question how a theory responds to variations in the geometry of its base manifold. (The terminology

stress tensor underpins this interpretation.) Methodologically, it is often convenient to compute the stress tensor via Eq. (1.48), including in cases where the theory is varied at a trivial metric $g_{\mu\nu} = \delta_{\mu\nu}$.

Answer:

(a) Using that $\partial g_{\rho\sigma}/\partial g_{\mu\nu} = \delta^{\mu}{}_{\rho}\delta^{\sigma}{}_{\nu}$,[35] the first identity is obtained from the matrix relation $0 = \partial_{g_{\mu\nu}}(gg^{-1}) = (\partial_{g_{\mu\nu}}g)g^{-1} + g\partial_{g_{\mu\nu}}g^{-1}$. Written in components, it yields the desired relation. With $g = \pm\det(g)$,[36] the second follows from $\partial_{g_{\mu\nu}}\sqrt{g} = 1/(2\sqrt{g})\partial_{g_{\mu\nu}}(\pm\det(g))$. Using $\det g = \exp\operatorname{tr}\ln g$, and $\partial_{g_{\mu\nu}}\operatorname{tr}\ln g = (g^{-1})^{\nu\mu} = g^{\mu\nu}$, we obtain the relation. The final relation follows from the chain rule applied to $\partial_\mu\phi = g_{\mu\nu}\partial^\nu\phi$.

(b) The metric enters the action in two places, the first being the factor $\sqrt{g}$, the second the derivatives $\phi^\mu = g^{\mu\nu}\phi^\nu$. We thus have

$$
\begin{aligned}
T^{\mu\nu} &= -\frac{2}{\sqrt{g}}\frac{\delta S}{\delta g_{\mu\nu}} = -\frac{2}{\sqrt{g}}\left(\mathcal{L}\frac{\partial\sqrt{g}}{\partial g_{\mu\nu}} + \sqrt{g}\frac{\partial\mathcal{L}}{\partial(\partial^\rho\phi)}\frac{\partial(\partial^\rho\phi)}{\partial g_{\mu\nu}}\right) \\
&= -\frac{2}{\sqrt{g}}\left(\mathcal{L}\frac{\partial\sqrt{g}}{\partial g_{\mu\nu}} + \sqrt{g}\frac{\partial\mathcal{L}}{\partial(\partial^\rho\phi)}\left(\frac{\partial g^{\rho\sigma}}{\partial g_{\mu\nu}}\right)\partial_\sigma\phi\right) = -\mathcal{L}g^{\mu\nu} + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\partial^\nu\phi,
\end{aligned}
$$

where in the final step we used the three relations in (a). Lowering the right index, $T^\mu{}_\nu = T^{\mu\rho}g_{\rho\sigma}$, we get back to Eq. (1.46).

Now considering the relation $\det g = \exp\operatorname{tr}\ln g$, it is varied as $\partial_{g_{\mu\nu}}\sqrt{g} = \frac{1}{2}\sqrt{g}\partial_{g_{\mu\nu}}\operatorname{tr}\ln g = \sqrt{g}(g^{-1})_{\nu\mu} = \sqrt{g}g^{\mu\nu}$, where we have used the symmetry $g_{\mu\nu} = g_{\nu\mu}$ of the metric tensor, and the notation $g^{\mu\nu} = (g^{-1})_{\mu\nu}$ for its inverse.

The differentiation in the second occurrence of the metric, $\partial_\mu\phi = g_{\mu\nu}\partial^\nu\phi$, is done as follows: $\partial_{g_{\mu\nu}}\mathcal{L} = \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\partial_{g_{\mu\nu}}\partial_\mu\phi = \frac{\mathcal{L}}{\partial(\partial_\mu\phi)}\partial^\nu\phi$. Adding the two terms we get

$$
T^{\mu\nu} = \sqrt{g}\left(g^{\mu\nu}\mathcal{L} + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\partial^\nu\phi\right).
$$

(c) For the free field theory in Minkowski space, we have $\sqrt{g} = 1$ and $\partial_{\partial_\mu\phi}\mathcal{L} = -\partial^\mu\phi$. This gives $T^0{}_0 = \frac{1}{2}(\partial_\mu\phi\partial^\mu\phi) - \partial^0\phi\partial_0\phi = \frac{1}{2}(-\partial^0\phi\partial_0\phi + \partial^1\phi\partial_1\phi) = \frac{1}{2}((\partial_0\phi)^2 + (\partial_1\phi)^2))$. Identifying the zero-coordinate with time, and the one-coordinate with space, this equals the Hamiltonian density (kinetic energy+potential energy density) of the harmonic chain.

---

[35] All derivatives are carried out for a general matrix, and then evaluated at the symmetric configuration $g_{\rho\sigma} = g_{\sigma\rho}$. We are not differentiating within the class of symmetric matrices. Think about this difference.

[36] It is common practice to denote the modulus determinant $g = \pm\det(\{g_{\mu\nu}\})$, and the matrix $g = \{g_{\mu\nu}\}$ by the same symbol g. Which is which should always be clear from the context.