

Chapter 2

Classical Physics

What is classical physics? In fact it has become the name for non-quantum physics. This begs the question: What is quantum physics in contrast to classical physics? One readily finds the statement that in classical physics the world is described by classical notions, like particles moving around in space, while in modern physics, i.e., quantum mechanics, the classical notions are no longer adequate, so there is no longer such a “naive” description of what is going on. In a more sophisticated version, quantum mechanics is physics in which the position and momentum of a particle are operators. But such a statement as it stands is meaningless. One also reads that the difference between quantum physics and classical physics is that the former has a smallest quantum of action, viz., Planck’s constant \hbar , and that classical physics applies whenever the action is large compared to \hbar , and in many circumstances, this is a true statement.

But our own viewpoint is better expressed as follows. Classical physics is the description of the world when the interference effects of the Schrödinger wave, evolving according to Schrödinger’s equation, can be neglected. This is the case for a tremendously wide range of scales from microscopic gases to stellar matter. In particular it includes the scale of direct human perception, and this explains why classical physics was found before quantum mechanics. Still, the viewpoint just expressed should seem puzzling. For how can classical motion of particles emerge from a wave equation like the Schrödinger equation? This is something we shall explain. It is easy to understand, once one writes down the equations of motion of Bohmian mechanics. But first let us discuss the theory which governs the behavior of matter across the enormous range of classical physics, namely, Newtonian mechanics. In a letter to Hooke, Newton wrote: “If I have seen further it is by standing on the shoulders of giants.”

2.1 Newtonian Mechanics

Newtonian mechanics is about point particles. What is a point particle? It is “stuff” or “matter” that occupies a point in space called its position, described mathematically by $\mathbf{q} \in \mathbb{R}^3$. The theory describes the motion of point particles in space. Mathematically, an N -particle system is described by the positions of the N particles:

$$\mathbf{q}_1, \dots, \mathbf{q}_N, \quad \mathbf{q}_i \in \mathbb{R}^3,$$

which change with time, so that one has trajectories $\mathbf{q}_1(t), \dots, \mathbf{q}_N(t)$, where the parameter $t \in \mathbb{R}$ is the time.

Newtonian mechanics is given by equations – the physical law – which govern the trajectories, called the equations of motion. They can be formulated in many different (but more or less equivalent) ways, so that the physical law looks different for each formulation, but the trajectories remain the same. We shall soon look at an example. Which formulation one prefers will be mainly a matter of taste. One may find the arguments leading to a particular formulation more satisfactory or convincing than others.

To formulate the law of Newtonian mechanics one introduces positive parameters, called masses, viz., m_1, \dots, m_N , which represent “matter”, and the law reads

$$m_i \ddot{\mathbf{q}}_i = \mathbf{F}_i(\mathbf{q}_1, \dots, \mathbf{q}_N). \quad (2.1)$$

\mathbf{F}_i is called the force. It is in general a function of all particle positions. Put another way, it is a function of the *configuration*, i.e., the family of all coordinates $(\mathbf{q}_1, \dots, \mathbf{q}_N) \in \mathbb{R}^{3N}$. The set of all such N -tuples is called *configuration space*. The quantity $\dot{\mathbf{q}}_i = d\mathbf{q}_i/dt = \mathbf{v}_i$ is the velocity of the i th particle, and its derivative $\ddot{\mathbf{q}}_i$ is called the acceleration.

Newtonian mechanics is romantic in a way. One way of talking about it is to say that particles accelerate each other, they interact through forces exerted upon each other, i.e., Newtonian mechanics is a theory of interaction. The fundamental interaction is gravitation or mass attraction given by

$$\mathbf{F}_i(\mathbf{q}_1, \dots, \mathbf{q}_N) = \sum_{j \neq i} G m_i m_j \frac{\mathbf{q}_j - \mathbf{q}_i}{\|\mathbf{q}_j - \mathbf{q}_i\|^3}, \quad (2.2)$$

with G the gravitational constant.

All point particles of the Newtonian universe interact according to (2.2). In effective descriptions of subsystems (when we actually use Newtonian mechanics in everyday life), other forces like harmonic forces of springs can appear on the right-hand side of (2.1). Such general forces need not (and in general will not) arise from gravitation alone. Electromagnetic forces will also play a role, i.e., one can sometimes describe electromagnetic interaction between electrically charged particles by the Coulomb force using Newtonian mechanics. The Coulomb force is similar to (2.2), but may have a different sign, and the masses m_i are replaced by the charges e_i which may be positive or negative.

One may wonder why Newtonian mechanics can be successfully applied to subsystems like the solar system, or even smaller systems like systems on earth. That is, why can one ignore all the rest of the universe? One can give various reasons. For example, distant matter which surrounds the earth in a homogeneous way produces a zero net field. The force (2.2) falls off with large distances and the gravitational constant is very small. In various general situations, and depending on the practical task in hand, such arguments allow a good effective description of the subsystem in which one ignores distant matter, or even not so distant matter.

Remark 2.1. Initial Value Problem

The equation (2.1) is a differential equation and thus poses an initial value problem, i.e., the trajectories $\mathbf{q}_i(t)$, $t \in \mathbb{R}$, which obey (2.1) are only determined once initial data of the form $\mathbf{q}_i(t_0)$, $\dot{\mathbf{q}}_i(t_0)$ are given, where t_0 is some time, called the initial time. This means that the future and past evolution of the trajectories is determined by the “present” state $\mathbf{q}_i(t_0)$, $\dot{\mathbf{q}}_i(t_0)$. Note that the position alone is not sufficient to determine the state of a Newtonian system.

It is well known that differential equations need not have unique and global solutions, i.e., solutions which exist for all times for all initial values. What does exist, however, is – at least in the case of gravitation – a local unique solution for a great many initial conditions, i.e., a solution which exists uniquely for some short period of time, if the initial values are reasonable. So (2.1) and (2.2) have no solution if, for example, two particles occupy the same position. Further, for the solution to exist, it must not happen that two or more particles collide, i.e., that they come together and occupy the same position. It is a famous problem in mathematical physics to establish what is called the existence of dynamics for a gravitating many-particle system, where one hopes to show that solutions fail to exist globally only for exceptional initial values. But what does “exceptional” mean? We shall answer this in a short while. ■

We wish to comment briefly on the manner of speaking about interacting particles, which gives a human touch to Newtonian mechanics. We say that the particles *attract each other*. Taking this notion to heart, one might be inclined to associate with the notion of particle more than just an object which has a position. But that might be misleading, since no matter how one justifies or speaks about Newtonian mechanics, when all is said and done, there remains a physical law about the motion of point particles, and that is a mathematical expression about changes of points in space with time. We shall explore one such prosaic description next.

2.2 Hamiltonian Mechanics

One can formulate the Newtonian law differently. Different formulations are based on different fundamental principles, like for example the principle of least action. But never mind such principles for the moment. We shall simply observe that it is

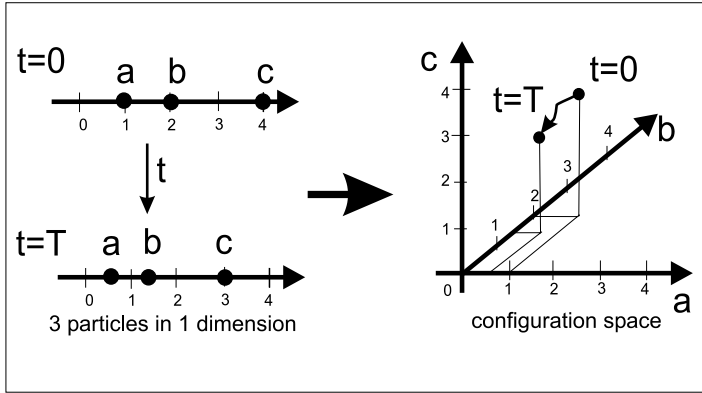


Fig. 2.1 Configuration space for 3 particles in a one-dimensional world

mathematically much nicer to rewrite everything in terms of configuration space variables:

$$\mathbf{q} = \begin{pmatrix} \mathbf{q}_1 \\ \vdots \\ \mathbf{q}_N \end{pmatrix} \in \mathbb{R}^{3N},$$

that is, we write the differential equation for all particles in a compact form as

$$m\ddot{\mathbf{q}} = \mathbf{F}, \quad (2.3)$$

with

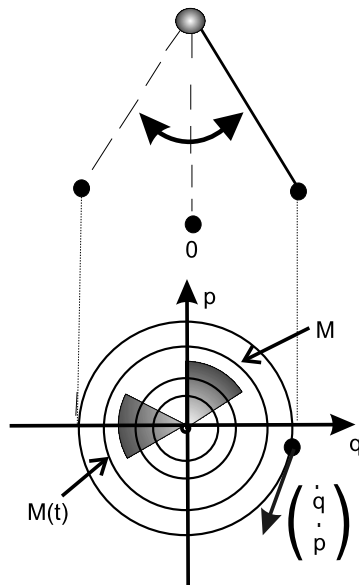
$$\mathbf{F} = \begin{pmatrix} \mathbf{F}_1 \\ \vdots \\ \mathbf{F}_N \end{pmatrix},$$

and the mass matrix $m = (\delta_i^j m_j)_{i,j=1,\dots,N}$.

Configuration space cannot be depicted (but see Fig. 2.1 for a very special situation), at least not for a system of more than one particle, because it is 6-dimensional for 2 particles in physical space. It is thus not so easy to think intuitively about things going on in configuration space. But one had better build up some intuition for configuration space, because it plays a fundamental role in quantum theory.

A differential equation is *by definition* a relation between the *flow* and the *vector field*. The flow is the mapping along the solution curves, which are integral curves along the vector field (the tangents of the solution curves). If a physical law is given by a differential equation, the vector field encodes the physical law. Let us see how this works.

Fig. 2.2 Phase space description of the mathematically idealized harmonically swinging pendulum. The possible trajectories of the mathematically idealized pendulum swinging in a plane with frequency 1 are concentric circles in phase space. The sets M and $M(t)$ will be discussed later



The differential equation (2.3) is of second order and does not express the relation between the integral curves and the vector field in a transparent way. We need to change (2.3) into an equation of first order, so that the vector field becomes transparent. For this reason we consider the *phase space* variables

$$\begin{pmatrix} \mathbf{q} \\ \mathbf{p} \end{pmatrix} = \begin{pmatrix} \mathbf{q}_1 \\ \vdots \\ \mathbf{q}_N \\ \mathbf{p}_1 \\ \vdots \\ \mathbf{p}_N \end{pmatrix} \in \mathbb{R}^{3N} \times \mathbb{R}^{3N} = \Gamma,$$

which were introduced by Boltzmann,¹ where we consider positions and velocities. However, for convenience of notation, the latter are replaced by momenta $\mathbf{p}_i = m_i \mathbf{v}_i$. One point in Γ represents the present state of the entire N -particle system. The phase space has twice the dimension of configuration space and can be depicted for one particle moving in one dimension, e.g., the pendulum (see Fig. 2.2).

Clearly, (2.3) becomes

¹ The notion of phase space was taken by Ludwig Boltzmann (1844–1906) as synonymous with the state space, the phase being the collection of variables which uniquely determine the physical state. The physical state is uniquely determined if its future and past evolution in time is uniquely determined by the physical law.

$$\begin{pmatrix} \dot{\mathbf{q}} \\ \dot{\mathbf{p}} \end{pmatrix} = \begin{pmatrix} m^{-1}\mathbf{p} \\ \mathbf{F}(\mathbf{q}) \end{pmatrix}. \quad (2.4)$$

The state of the N -particle system is completely determined by $\begin{pmatrix} \mathbf{q} \\ \mathbf{p} \end{pmatrix}$, because (2.4) and the initial values $\begin{pmatrix} \mathbf{q}(t_0) \\ \mathbf{p}(t_0) \end{pmatrix}$ uniquely determine the phase space trajectory (if the initial value problem allows for a solution).

For (2.2) and many other effective forces, there exists a function V on \mathbb{R}^{3N} , the so called potential energy function, with the property that

$$\mathbf{F} = -\text{grad}_{\mathbf{q}} V = -\frac{\partial V}{\partial \mathbf{q}} = -\nabla V.$$

Using this we may write (2.4) as

$$\begin{pmatrix} \dot{\mathbf{q}} \\ \dot{\mathbf{p}} \end{pmatrix} = \begin{pmatrix} \frac{\partial H}{\partial \mathbf{p}}(\mathbf{q}, \mathbf{p}) \\ -\frac{\partial H}{\partial \mathbf{q}}(\mathbf{q}, \mathbf{p}) \end{pmatrix}, \quad (2.5)$$

where

$$\begin{aligned} H(\mathbf{q}, \mathbf{p}) &= \frac{1}{2}(\mathbf{p} \cdot m^{-1}\mathbf{p}) + V(\mathbf{q}) \\ &= \frac{1}{2} \sum_{i=1}^N \frac{\mathbf{p}_i^2}{m_i} + V(\mathbf{q}_1, \dots, \mathbf{q}_N). \end{aligned} \quad (2.6)$$

Now we have the Newtonian law in the form of a transparent differential equation (2.5), expressing the relation between the integral curves (on the left-hand side, differentiated to yield tangent vectors) and the vector field on the right-hand side (which are the tangent vectors expressing the physics). The way we have written it, the vector field is actually generated by a function H (2.6) on phase space. This is called the Hamilton function, after its inventor William Rowan Hamilton (1805–1865), who in fact introduced the symbol H in honor of the physicist Christiaan Huygens (1629–1695). We shall see later what the “wave man” Huygens has to do with all this. The role of the Hamilton function $H(\mathbf{q}, \mathbf{p})$ is to give the vector field

$$\mathbf{v}^H(\mathbf{q}, \mathbf{p}) = \begin{pmatrix} \frac{\partial H}{\partial \mathbf{p}} \\ -\frac{\partial H}{\partial \mathbf{q}} \end{pmatrix}, \quad (2.7)$$

and the Hamiltonian dynamics is simply given by

$$\begin{pmatrix} \dot{\mathbf{q}} \\ \dot{\mathbf{p}} \end{pmatrix} = \mathbf{v}^H(\mathbf{q}, \mathbf{p}) . \quad (2.8)$$

The function H allows us to focus on a particular structure of Newtonian mechanics, now rewritten in Hamiltonian terms. Almost all of this section depends solely on this structure, and we shall see some examples shortly. Equations (2.5) and (2.6) with the Hamilton function $H(\mathbf{q}, \mathbf{p})$ define a Hamiltonian system.

The integral curves along this vector field (2.7) represent the possible system trajectories in phase space, i.e., they are solutions $\begin{pmatrix} \mathbf{q}(t, (\mathbf{q}, \mathbf{p})) \\ \mathbf{p}(t, (\mathbf{q}, \mathbf{p})) \end{pmatrix}$ of (2.8) for given initial values $\begin{pmatrix} \mathbf{q}(0, (\mathbf{q}, \mathbf{p})) \\ \mathbf{p}(0, (\mathbf{q}, \mathbf{p})) \end{pmatrix} = \begin{pmatrix} \mathbf{q} \\ \mathbf{p} \end{pmatrix}$. Note that this requires existence and uniqueness of solutions of the differential equations (2.8). One possible evolution of the entire system is represented by one curve in phase space (see Fig. 2.3), which is called a flow line, and one defines the Hamiltonian flow by the map $(\Phi_t^H)_{t \in \mathbb{R}}$ from phase space to phase space, given by the prescription that, for any t , a point in phase space is mapped to the point to which it moves in time t under the evolution (as long as that evolution is defined, see Remark 2.1):

$$\Phi_t^H \left(\begin{pmatrix} \mathbf{q} \\ \mathbf{p} \end{pmatrix} \right) = \begin{pmatrix} \mathbf{q}(t, (\mathbf{q}, \mathbf{p})) \\ \mathbf{p}(t, (\mathbf{q}, \mathbf{p})) \end{pmatrix} .$$

We shall say more about the flow map later on. The flow can be thought of pictorially as the flow of a material fluid in Γ , with the system trajectories as flow lines.

Hamiltonian mechanics is another way of talking about Newtonian mechanics. It is a prosaic way of talking about the motion of particles. The only romance left is the secret of how to write down the physically relevant H . Once that is done, the romance is over and what lies before one are the laws of mechanics written in mathematical language. So that is all that remains. The advantage of the Hamiltonian form is that it directly expresses the law as a differential equation (2.8). And it has the further advantage that it allows one to talk simultaneously about all possible trajectories of a system. This will be helpful when we need to define a *typical* trajectory of the system, which we must do later.

However, this does not by any means imply that we should forget the Newtonian approach altogether. To understand which path a system takes, it is good to know how the particles in the system interact with each other, and to have some intuition about that. Moreover, we should not lose sight of what we are interested in, namely, the behavior of the system in physical space. Although we have not elaborated on the issue at all, it is also important to understand the physical reasoning which leads to the mathematical law (for example, how Newton found the gravitational potential), as this may give us confidence in the correctness of the law. Of course we also achieve confidence by checking whether the theory correctly describes what we see, but since we can usually only see a tiny fraction of what a theory says, confidence is mainly grounded on theoretical insight.

The fundamental properties of the Hamiltonian flow are conservation of energy and conservation of volume. These properties depend only on the form of the equa-

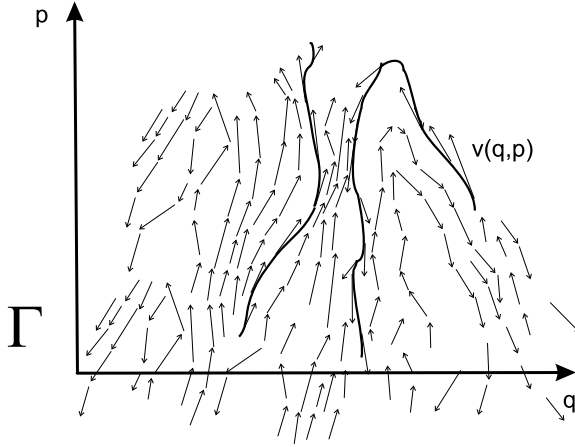


Fig. 2.3 The Hamilton function generates a vector field on the $6N$ -dimensional phase space of an N -particle system in physical space. The integral curves are the possible trajectories of the entire system in phase space. Each point in phase space is the collection of all the positions and velocities of all the particles. One must always keep in mind that the trajectories in phase space are not trajectories in physical space. They can never cross each other because they are integral curves on a vector field, and a unique vector is attached to every point of phase space. Trajectories in phase space do not interact with each other! They are not the trajectories of particles

tions (2.8) with (2.7), i.e., $H(\mathbf{q}, \mathbf{p})$ can be a completely general function of (\mathbf{q}, \mathbf{p}) and need not be the function (2.6). When working with this generality, one calls \mathbf{p} the canonical momentum, which is no longer simply velocity times mass. Now, conservation of energy means that the value of the Hamilton function does not change along trajectories. This is easy to see. Let $(\mathbf{q}(t), \mathbf{p}(t))$, $t \in \mathbb{R}$, be a solution of (2.8). Then

$$\frac{d}{dt}H(\mathbf{q}(t), \mathbf{p}(t)) = \dot{\mathbf{q}} \frac{\partial H}{\partial \mathbf{q}} + \dot{\mathbf{p}} \frac{\partial H}{\partial \mathbf{p}} = \frac{\partial H}{\partial \mathbf{p}} \frac{\partial H}{\partial \mathbf{q}} - \frac{\partial H}{\partial \mathbf{q}} \frac{\partial H}{\partial \mathbf{p}} = 0. \quad (2.9)$$

More generally, the time derivative along the trajectories of any function $f(\mathbf{q}(t), \mathbf{p}(t))$ on phase space is

$$\frac{d}{dt}f(\mathbf{q}(t), \mathbf{p}(t)) = \dot{\mathbf{q}} \frac{\partial f}{\partial \mathbf{q}} + \dot{\mathbf{p}} \frac{\partial f}{\partial \mathbf{p}} = \frac{\partial H}{\partial \mathbf{p}} \frac{\partial f}{\partial \mathbf{q}} - \frac{\partial H}{\partial \mathbf{q}} \frac{\partial f}{\partial \mathbf{p}} =: \{f, H\}. \quad (2.10)$$

The term $\{f, H\}$ is called the Poisson bracket of f and H . It can also be defined in more general terms for any pair of functions f, g , viewing g as the Hamilton function and Φ_t^g the flow generated by g :

$$\{f, g\} = \frac{d}{dt}f \circ \Phi_t^g = \frac{\partial g}{\partial \mathbf{p}} \frac{\partial f}{\partial \mathbf{q}} - \frac{\partial g}{\partial \mathbf{q}} \frac{\partial f}{\partial \mathbf{p}}. \quad (2.11)$$

Note, that $\{f, H\} = 0$ means that f is a *constant of the motion*, i.e., the value of f remains unchanged along a trajectory ($df/dt = 0$), with $f = H$ being the simplest example.

Now we come to the conservation of volume. Recall that the Hamiltonian flow $(\Phi_t^H)_{t \in \mathbb{R}}$ is best pictured as a fluid flow in Γ , with the system trajectories as flow lines. These are the integral curves along the Hamiltonian vector field $\mathbf{v}^H(\mathbf{q}, \mathbf{p})$ (2.7). These flow lines have neither sources nor sinks, i.e., the vector field is divergence-free:

$$\operatorname{div} \mathbf{v}^H(\mathbf{q}, \mathbf{p}) = \left(\frac{\partial}{\partial \mathbf{q}}, \frac{\partial}{\partial \mathbf{p}} \right) \begin{pmatrix} \frac{\partial H}{\partial \mathbf{p}} \\ -\frac{\partial H}{\partial \mathbf{q}} \end{pmatrix} = \frac{\partial^2 H}{\partial \mathbf{q} \partial \mathbf{p}} - \frac{\partial^2 H}{\partial \mathbf{p} \partial \mathbf{q}} = 0. \quad (2.12)$$

This important (though rather trivial) mathematical fact is known as Liouville's theorem for the Hamiltonian flow, after Joseph Liouville (1809–1882). (It has nothing to do with Liouville's theorem in complex analysis.) A fluid with a flow that is divergence-free is said to be incompressible, a behavior different from air in a pump, which gets very much compressed. Consequently, and as we shall show below, the “volume” of any subset in phase space which gets transported via the Hamiltonian flow remains unchanged. Before we express this in mathematical terms and give the proof, we shall consider the issue in more general terms.

Remark 2.2. On the Time Evolution of Measures.

The notion of volume deserves some elaboration. Clearly, since phase space is very high-dimensional, the notion of volume here is more abstract than the volume of a three-dimensional object. In fact, we shall later use a notion of volume which is not simply the trivial extension of three-dimensional volume. Volume here refers to a *measure*, the size or weight of sets, where one may in general want to consider a biased weight. The most famous measure, and in fact the mother of all measures, is the generalization of the volume of a cube to arbitrary subsets, known as the Lebesgue measure λ . We shall say more about this later.² If one feels intimidated by the name Lebesgue measure, then take $|A| = \int_A d^n x$, the usual Riemann integral, as the (fapp-correct) Lebesgue measure of A . The measure may in a more general sense be thought of as some kind of weight distribution, where the Lebesgue measure gives equal (i.e., unbiased) weight to every point. For a continuum of points, this is a somewhat demanding notion, but one may nevertheless get a feeling for what is meant. For the time being we require that the measure be an additive nonnegative set function, i.e., a function which attributes positive or zero values to sets, and which is additive on disjoint sets: $\mu(A \cup B) = \mu(A) + \mu(B)$. The role of the measure will

² We need to deal with the curse of the continuum, which is that not all subsets of \mathbb{R}^n actually have a volume, or as we now say, a measure. There are non-measurable sets within the enormous multitude of subsets. These non-measurable sets exist mathematically, but are not constructible in any practical way out of unions and intersections of simple sets, like cubes or balls. They are nothing we need to worry about in practical terms, but they are nevertheless there, and so must be dealt with properly. This we shall do in Sect. 4.3.1.

eventually be to tell us the size of a set, i.e., which sets are small and which are big. Big sets are important, while small ones are not.

It may be best to think for now of a measure in general as some abstract way of attributing “mass” to subsets. One is then led to ask how the measure (or the mass) changes with a flow. That question was first asked for the Hamiltonian flow, but it can be asked, and in fact has been asked, for flows of a general character. We shall do the same now. Let μ be a measure on the phase space Γ , which we take for simplicity as being \mathbb{R}^n . We consider now a general (not necessarily Hamiltonian) flow map on phase space, that is, a one-parameter family of maps $(\Phi_t)_{t \in \mathbb{R}}$, with parameter “time” t :

$$(\Phi_t(\mathbf{x}))_{t \in \mathbb{R}}, \quad \mathbf{x} \in \mathbb{R}^n: \quad \Phi_t \circ \Phi_s(\mathbf{x}) = \Phi_{t+s}(\mathbf{x}), \quad \Phi_0(\mathbf{x}) = \mathbf{x}. \quad (2.13)$$

In general, any flow Φ_t on \mathbb{R}^n (or some general phase space) naturally defines the time evolution of the measure μ_t on \mathbb{R}^n :

$$\mu_t = \mu \circ \Phi_{-t}, \quad (2.14)$$

which means

$$\mu_t(A) = \mu(\Phi_{-t}A), \quad \text{or} \quad \mu_t(\Phi_t A) = \mu(A), \quad (2.15)$$

for all (measurable) sets A and all $t \in \mathbb{R}$. Behind this definition is the simple logic that, if a measure μ is given at time $t = 0$, then the measure μ_t of a set A is the measure μ of the set from which A originated by virtue of the flow. In other words the measure changes only because the set changes.

The notion of stationary measure is very important. This is a measure that does not change under the flow Φ_t , i.e.,

$$\mu_t(A) = \mu(A), \quad \forall t \in \mathbb{R}, \quad \forall A. \quad (2.16)$$

The stationary measure plays a distinguished role in justifying probabilistic reasoning. Its importance was presumably first discovered by Boltzmann, and later on we shall spend some time considering Boltzmann’s general ideas about statistical physics, which are valid for a whole range of theories.

The above-mentioned preservation of volume for Hamiltonian flows as a consequence of Liouville’s theorem refers to phase space volume and is the assertion that

$$\lambda(\Phi_{-t}A) = \lambda(A), \quad (2.17)$$

with λ the Lebesgue measure on phase space. This means that, under the Hamiltonian flow, sets change their shape but not their volume. This may also be referred to as Liouville’s theorem, since it is a direct consequence of (2.12), as we shall show next. For the pendulum in Fig. 2.2, one sees this immediately, since the slices of the pie just rotate. The general situation is depicted in Fig. 2.4. ■

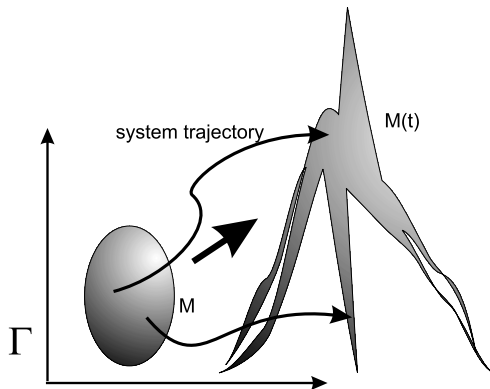


Fig. 2.4 The preservation of volume of the Hamiltonian flow. The set M in phase space changes under the flow but its volume remains the same

Remark 2.3. Continuity Equation

From the change in the measure (2.15) for a general flow generated by a vector field, one can derive a differential equation which governs the change in the *density* of the measure. That differential equation is called the continuity equation. If the measure has a density $\rho(\mathbf{x})$ (you may think of a mass density), then the change of measure with time defines a time-dependent density $\rho(\mathbf{x}, t)$, and one has the logical relation³

$$\begin{aligned} \mu_t(A) &=: \int \chi_A(\mathbf{x}) \rho(\mathbf{x}, t) d^n x \\ &\parallel \\ \mu(\Phi_{-t}A) &=: \int \chi_{\Phi_{-t}A}(\mathbf{x}) \rho(\mathbf{x}) d^n x \\ &= \int \chi_A(\Phi_t(\mathbf{x})) \rho(\mathbf{x}) d^n x, \end{aligned} \quad (2.18)$$

where χ_A is the characteristic function of the set $A \subset \Gamma$, also called the indicator function of the set A , i.e., the function which is 1 on A and zero otherwise, and $\Phi_{-t}A = \{\mathbf{x} \in \Gamma \mid \Phi_t(\mathbf{x}) \in A\}$. Furthermore Φ_t is the solution flow map of some vector field $\mathbf{v}(\mathbf{x})$ on \mathbb{R}^n (or some general phase space), i.e.,

$$\frac{d}{dt} \Phi_t(\mathbf{x}) = \mathbf{v}(\Phi_t(\mathbf{x})). \quad (2.19)$$

We shall now show that the density $\rho(\mathbf{x}, t)$ satisfies the continuity equation:

$$\frac{\partial}{\partial t} \rho(\mathbf{x}, t) + \operatorname{div} [\mathbf{v}(\mathbf{x}) \rho(\mathbf{x}, t)] = 0. \quad (2.20)$$

³ Note in passing that $\rho(\mathbf{x}, t)$ can be computed from an obvious change of variables in the last integral, namely, $\rho(\mathbf{x}, t) = \rho(\Phi_{-t}(\mathbf{x})) |\partial \Phi_{-t}(\mathbf{x}) / \partial \mathbf{x}|$.

To see this, replace the indicator function in (2.18) by a smooth function f with compact support:

$$\int f(\Phi_t(\mathbf{x}))\rho(\mathbf{x})d^n x = \int f(\mathbf{x})\rho(\mathbf{x},t)d^n x. \quad (2.21)$$

Now differentiate (2.21) with respect to t to get

$$\int \frac{d\Phi_t(\mathbf{x})}{dt} \cdot [\nabla f(\Phi_t(\mathbf{x}))] \rho(\mathbf{x})d^n x = \int f(\mathbf{x}) \frac{\partial}{\partial t} \rho(\mathbf{x},t)d^n x. \quad (2.22)$$

Replacing $d\Phi_t(\mathbf{x})/dt$ by the right-hand side of (2.19) and using (2.21) again with f replaced by $\mathbf{v}(\mathbf{x}) \cdot [\nabla f(\mathbf{x})]$ (a wonderful trick), the left-hand side of (2.22) becomes

$$\int \mathbf{v}(\mathbf{x}) \cdot [\nabla f(\mathbf{x})] \rho(\mathbf{x},t)d^n x, \quad (2.23)$$

and after partial integration,

$$- \int f(\mathbf{x}) \operatorname{div} [\mathbf{v}(\mathbf{x})\rho(\mathbf{x},t)]d^n x. \quad (2.24)$$

Since this is equal to the right-hand side of (2.22) and since f is arbitrary, we conclude that (2.20) holds. ■

Now we ask whether there is a stationary measure (2.16). In terms of densities, the question is: Does there exist a stationary density, that is, a density independent of time, satisfying (2.20)? Since the time derivative part of (2.20) vanishes, i.e.,

$$\frac{\partial}{\partial t} \rho(\mathbf{x}) = 0,$$

the density must satisfy the partial differential equation

$$\operatorname{div} [\mathbf{v}(\mathbf{x})\rho(\mathbf{x})] = 0. \quad (2.25)$$

This is in general a very difficult question. In particular, it is almost impossible to find the solution for a general vector field. However, not so in the Hamiltonian case, where the answer turns out to be trivial! This is a consequence of (2.12). Setting

$$\mathbf{x} = \begin{pmatrix} \mathbf{q} \\ \mathbf{p} \end{pmatrix},$$

equation (2.12) reads $\operatorname{div} \mathbf{v}(\mathbf{x}) = 0$ for all \mathbf{x} . In this case, (2.20) becomes (after using the product rule)

$$\frac{\partial}{\partial t} \rho(\mathbf{x},t) + \mathbf{v}(\mathbf{x}) \cdot \nabla \rho(\mathbf{x},t) = 0. \quad (2.26)$$

Now $\rho(\mathbf{x}, t) = \text{constant}$, which we may choose to be unity, is obviously stationary. Putting $f = \chi_A$, $A \subset \Gamma$, and taking into account $\rho = 1$, (2.21) yields

$$\int \chi_A(\Phi_t(\mathbf{x})) d^n x = \int \chi_{\Phi_{-t}A}(\mathbf{x}) d^n x = \int \chi_A(\mathbf{x}) d^n x = \lambda(A). \quad (2.27)$$

In short, (2.27) says

$$\lambda(\Phi_{-t}A) = \lambda(A). \quad (2.28)$$

We may as well put the future set $\Phi_t A$ into (2.28), instead of A , and use $\Phi_{-t}\Phi_t = \text{id}$ (identity), whence

$$\lambda(A) = \lambda(\Phi_t A). \quad (2.29)$$

In conclusion, *Liouville's theorem* implies that the Lebesgue measure (\equiv volume) is stationary for Hamiltonian flows.

Remark 2.4. Time-Dependent Vector Fields

The continuity equation also holds for time-dependent vector fields $\mathbf{v}(\mathbf{x}, t)$, in which case the flow map is a two parameter group $\Phi_{t,s}$ advancing points from time s to time t . All one needs to do is to replace the vector field in (2.20) by the time-dependent expression $\mathbf{v}(\mathbf{x}, t)$, and the proof goes through verbatim. But now the notion of a stationary measure seems unsuitable, since the velocity field (representing the physical law) changes with time. But remarkably the notion still applies for Hamiltonian flows, i.e., even in the case where the Hamiltonian is time dependent (energy is not conserved), the volume (Lebesgue measure of a set) remains unchanged under the flow. ■

Remark 2.5. On the Initial Value Problem

The Lebesgue measure in phase space plays a distinguished role for the Hamiltonian flow. It is thus natural to weaken the problem of initial values in the sense of the measure, so that one is happy if it can be shown that the bad set of initial conditions for which no global solutions exist has Lebesgue measure zero. Be warned however, that a set which has measure zero may be small in the sense of the measure, but it is not necessarily small in the sense of cardinality (number of points in the set). The famous Cantor set, a subset of the interval $[0, 1]$, has as many members as the reals, but has Lebesgue measure zero.

We close this section with what is a heretical thought for modern physicists, namely, a Newtonian universe. This is not physical (we know it ignores quantum mechanics and much more), but we can nevertheless conceive of it, and it is a good enough framework in which to ask a question which in an appropriate sense could be raised in any other theory: *Which initial values give rise to OUR Newtonian universe?* Put another way: *According to which criteria were the initial values of OUR universe chosen?* We do not ask *who* chose the initial conditions, but rather: *Which physical law determines them?* One possible answer to this could be that our universe is nothing special, that it could be *typical* in the sense that almost all initial

conditions would give rise to a universe very much like ours (where “almost all” means that the Lebesgue measure of the set which does not give rise to a universe like ours is very small). It turns out that this is not the case, but we shall address this issue later. ■

2.3 Hamilton–Jacobi Formulation

The Hamiltonian structure and phase space are intimately connected with symplectic geometry. We shall say more about that at the end of the chapter. We wish to move on to another question. The Hamiltonian formulation of Newtonian mechanics is prosaic and brings out the particular structure shared by Newtonian mechanics and all Hamiltonian flows: conservation of energy (if H is time independent) and phase space volume. But that was not Hamilton’s aim. He had a much deeper vision for mechanics. He was looking for an analogy between mechanics and wave optics, namely Huygens’ principle and Fermat’s extremal principle of geometric optics, according to which light rays take the path requiring the shortest time, and moreover follow the normals of wave fronts. Could mechanics be formulated by a similar guidance principle, where the mechanical paths are determined by the normal vectors of wave fronts? The extremal principle replacing Fermat’s is the least action principle $\delta \int L dt = 0$, where $L(\mathbf{q}, \dot{\mathbf{q}})$ is called the Lagrange function. The mechanical (Newtonian) trajectories between t_0, \mathbf{q}_0 and t, \mathbf{q} (note that instead of initial position and initial velocity, we consider initial position and end position) are characterized as the extremals of

$$\int_{t_0}^t L(\mathbf{q}(t'), \dot{\mathbf{q}}(t')) dt'.$$

We omit the derivation of the Euler–Lagrange equation, which is standard, but we recall that for Newtonian mechanics

$$L(\mathbf{q}, \dot{\mathbf{q}}) = \frac{1}{2} \dot{\mathbf{q}} \cdot m \dot{\mathbf{q}} - V(\mathbf{q}). \quad (2.30)$$

For this Lagrange function, the Euler–Lagrange equations are the usual Newtonian equations. The Lagrange function and Hamilton function are Legendre transforms of one another.⁴ (It is remarkable that almost all the great mathematicians of the 19th century have left some trace in theoretical mechanics.) Starting with H , we get L by changing from the variable \mathbf{p} to $\dot{\mathbf{q}}$, so that (\mathbf{q}, \mathbf{p}) gets replaced by $(\mathbf{q}, \dot{\mathbf{q}})$, using the implicitly given function

$$\dot{\mathbf{q}} = \frac{\partial H(\mathbf{q}, \mathbf{p})}{\partial \mathbf{p}},$$

⁴ Here is the definition. Let $f(x)$ be convex and let z be a given slope. Then look for the point $x(z)$ at which the tangent to the graph of f has slope z . You find $x(z)$ by minimizing $F(x, z) = f(x) - xz$ in x . By convexity of f , this is uniquely determined. The Legendre transform of f is $g(z) = F(x(z), z)$.

where the equation is solved by \mathbf{p} as a function of $\dot{\mathbf{q}}$.

For “normal” Hamilton functions (quadratic in the momentum), that solution is immediate, and looking at (2.30), the Legendre transform pops right up:

$$L(\mathbf{q}, \dot{\mathbf{q}}) = \mathbf{p} \cdot \dot{\mathbf{q}} - H(\mathbf{q}, \mathbf{p}) . \quad (2.31)$$

Note in passing that if one starts with the least action principle as being fundamental, one can guess the form of the Lagrange function from basic principles such as symmetry, homogeneity, and simplicity. But that is not what we wish to discuss here.

We come now to Huygens’ principle and the definition of waves $S_{\mathbf{q}_0, t_0}(\mathbf{q}, t)$ which guide mechanical trajectories starting at \mathbf{q}_0 and moving along the vector field $\mathbf{p}(\mathbf{q}, t) = \nabla S_{\mathbf{q}_0, t_0}(\mathbf{q}, t)$. Hamilton suggested

$$S_{\mathbf{q}_0, t_0}(\mathbf{q}, t) := \int_{t_0}^t L(\gamma, \dot{\gamma}) dt , \quad (2.32)$$

where $\gamma : \mathbf{q}_0, t_0 \longrightarrow \mathbf{q}, t$ is the extremum of the action principle, i.e., the Newtonian path. This function is often called the Hamilton–Jacobi function.

Unfortunately, this definition generally leads to a multivalued function. Take for example the harmonic oscillator with period T . There are many extremal trajectories for a harmonic oscillator with period T which go from $(0, T)$ to $(0, 2T)$, so S is not uniquely defined. Or again, think of a ball which bounces off a wall. The position \mathbf{q} in front of the wall can always be reached within a given time by two routes, one with and one without reflection from the wall.

However, the definition is good for short enough times. So never mind this difficulty, let us pursue the thought. Ignoring the dependence on (\mathbf{q}_0, t_0) and considering⁵

$$dS = \frac{\partial S}{\partial \mathbf{q}} d\mathbf{q} + \frac{\partial S}{\partial t} dt ,$$

and in view of (2.32) and (2.31), we can identify

$$dS = \mathbf{p} d\mathbf{q} - H(\mathbf{q}, \mathbf{p}) dt .$$

Then by comparison

$$\mathbf{p}(\mathbf{q}, t) = \frac{\partial S}{\partial \mathbf{q}}(\mathbf{q}, t) , \quad (2.33)$$

whence

$$\frac{\partial S}{\partial t}(\mathbf{q}, t) + H\left(\mathbf{q}, \frac{\partial S}{\partial \mathbf{q}}\right) = 0 . \quad (2.34)$$

⁵ We may ignore that dependence because we assume uniqueness of the trajectory, and this implies that $S_{\mathbf{q}_0, t_0}(\mathbf{q}, t) = S_{\mathbf{q}_1, t_1}(\mathbf{q}, t) + S_{\mathbf{q}_0, t_0}(\mathbf{q}_1, t_1)$.

This is known as the Hamilton–Jacobi differential equation. For Newtonian mechanics, where $\dot{\mathbf{q}}_i = \mathbf{p}_i/m_i$, we then obtain the following picture. On the configuration space \mathbb{R}^n for N particles ($n = 3N$), we have a function $S(\mathbf{q}, t)$ (“unfortunately” multivalued), whose role it is to generate a vector field

$$\mathbf{v}(\mathbf{q}, t) = m^{-1} \nabla S(\mathbf{q}, t) \quad (2.35)$$

on configuration space. Integral curves $\mathbf{Q}(t)$ along the vector field are the possible trajectories of the N -particle system, i.e., they solve

$$\frac{d\mathbf{Q}}{dt} = \mathbf{v}(\mathbf{Q}(t), t) .$$

The function $S(\mathbf{q}, t)$ is itself dynamical and solves the nonlinear partial differential equation

$$\frac{\partial S}{\partial t}(\mathbf{q}, t) + \frac{1}{2} \sum_{i=1}^N \frac{1}{m_i} \left(\frac{\partial S}{\partial \mathbf{q}_i} \right)^2 + V(\mathbf{q}) = 0 . \quad (2.36)$$

This picture is, as we said, not quite right, because S is in general not well defined for mechanics. On the other hand, it is almost quantum mechanics. We shall soon understand which little quantum is missing to get the picture right.

2.4 Fields and Particles: Electromagnetism

Many hold the view that the particle is not a good concept for physics. They see it as a classical Newtonian concept which has been made obsolete by quantum mechanics. Fields on the other hand are generally well accepted, because relativity teaches us that the right physical theory will be a field theory, eventually quantized of course. To understand whether fields do work as well as one hopes, we shall have a quick look at electromagnetism, where dynamical fields come into play as fundamental objects needed to describe interactions between particles which carry a “charge”. Electromagnetic fields act on a particle at position $\mathbf{q} \in \mathbb{R}^3$ via the Lorentz force:

$$m\ddot{\mathbf{q}} = e \left[\mathbf{E}(\mathbf{q}, t) + \frac{\dot{\mathbf{q}}}{c} \times \mathbf{B}(\mathbf{q}, t) \right] , \quad (2.37)$$

where $\mathbf{E}(\mathbf{q}, t)$ und $\mathbf{B}(\mathbf{q}, t)$ are the electric and magnetic fields, and c is the velocity of light. While the fields act on particles as described, in electromagnetism the fields are not independent agents living in a kingdom of their own, for they are themselves generated by particles. They are generated by particles and they act on particles, which is why one may say that they are there to represent the interaction between particles. But when the particles are point particles, which is the most natural rela-

tivistic possibility, this does not go down well with the fields. We shall explain this now. We shall also use the opportunity to introduce relativistic physics.

Albert Einstein (1879–1955) deduced from Maxwell’s equations of electromagnetism that space and most importantly time change in a different way from Galilean physics when one changes between frames moving with respect to each other. The nature of this change is governed by the fact that the velocity of light does not change, when one moves from one frame to another. This led to the understanding, soon to be formalised in the four-dimensional description by Hermann Minkowski (1864–1909), that space and time are “of the same kind”. That is, a particle needs for its specification not only a position in space, but also a location in time, implying that the coordinates of a particle in relativistic physics should be space and time coordinates. This is a revolution in Newtonian mechanics, where we are of course used to particles having different positions, but not different times. So in relativistic physics one must get used to the spacetime description of particles, with each particle having its own spacetime coordinates. In other words, the configuration space of classical mechanics, where we collect all positions of the particles of a system at the same time, no longer plays a fundamental role. Instead, Einstein showed that the overthrow of absolute time brings physics closer to a true description of nature. On this basis, he believed that physical theories must be local, in the sense that no physical effect can move faster than light. John Bell showed that this is wrong. We shall devote a whole chapter to this later, but now we must move on.

Minkowski introduced the so-called four-dimensional spacetime with a particular scalar product.⁶ In spacetime, particles no longer move in a Newtonian way, but according to new dynamics. The particle position is now x^μ , $\mu = (0, 1, 2, 3)$, where x^0 is selected as the time coordinate, since it is distinguished by the “signature” of the so-called Minkowski length⁷

$$ds^2 = (dx^0)^2 - d\mathbf{x}^2 = (dx^0)^2 - \sum_{i=1}^3 (dx^i)^2.$$

In Newtonian mechanics, we are used to parameterizing paths by time, which is no longer natural. A natural parameter now is length – Minkowski length – normalized by $1/c$, i.e., on each trajectory we select a zero mark somewhere and go from there

⁶ Minkowski suggested using imaginary numbers for the time coordinate in the spacetime coordinates ($x_0 = ict, \mathbf{x}$), because then the formal Euclidean scalar product yields the Minkowski metric $s^2 = (x^0)^2 + \mathbf{x}^2 = -ct^2 + \sum_{i=1}^3 (x^i)^2$. The advantage is that all congruences, i.e., transformations leaving the scalar product invariant (which form the so-called Lorentz group) appear in Euclidean guise, and so can be viewed as four-dimensional rotations and reflections. This differs from the Galilean case, where the change to relatively moving frames (Galilean boosts) must be dealt with separately. In the Minkowski case, the corresponding Lorentz boost is simply a “rotation”, but with an imaginary angle. When one considers the changes of x_0 and, say, x_1 , such a rotation yields $x_0' = x_0 \cos \phi - x_1 \sin \phi$, $x_1' = x_0 \sin \phi + x_1 \cos \phi$, which requires in the Minkowski case $\phi = i\psi$, an imaginary angle. Following the point $x_1 = 0$ in the primed frame (moving with relative velocity v), one then has $x_1'/ct' = v/c = \tanh \psi$, which yields the well-known formula for the Lorentz boost.

⁷ The sign of ds^2 is conventional: $ds^2 < 0$ implies a spacelike distance, while $ds^2 > 0$ implies a timelike distance.

with the length element of the i th particle

$$d\tau_i = \frac{ds_i}{c} = \frac{1}{c} \sqrt{\left(\frac{dx_i^0}{d\tau_i}\right)^2 - \left(\frac{d\mathbf{x}_i}{d\tau_i}\right)^2} d\tau_i.$$

Thus

$$\dot{x}_i^2 = g_{\mu\nu} \dot{x}_i^\mu \dot{x}_i^\nu = c^2, \quad g_{\mu\nu} = \begin{pmatrix} 1 & 0 \\ 0 & -E_3 \end{pmatrix}, \quad (2.38)$$

where E_3 is the 3×3 identity matrix, and we use the Einstein summation convention according to which we sum automatically over those indices appearing more than once. The dot over x_i indicates the derivative with respect to Minkowski length τ_i , also called the proper time, of the i th particle (in the frame where the particle is at rest $\dot{x}_i^0 = c\tau_i$). The metric tensor $g_{\mu\nu}$ which defines here the Minkowski scalar product can be used to lower indices:

$$x_\mu := g_{\mu\nu} x^\nu,$$

while the inverse of the metric tensor denoted by $g^{\mu\nu}$ is used to raise indices.

For us it is natural to parameterize the trajectory by the coordinate time $x^0 = ct$ of our rest frame where we see the particle move with velocity \mathbf{v} . We thus have a new function

$$\bar{x}^\mu(x^0) = (x^0, \mathbf{x}(x^0)) = x^\mu(\tau(x^0)),$$

for which we get by the chain rule

$$\frac{d\bar{x}^\mu}{dx^0} = \left(1, \frac{\mathbf{v}}{c}\right) = \dot{x}^\mu \frac{d\tau}{dx^0},$$

or, taking the Minkowski square,

$$1 - \frac{\mathbf{v}^2}{c^2} = c^2 \left(\frac{d\tau}{dx^0}\right)^2, \quad (2.39)$$

which allows us to switch between proper time and coordinate time.

The relativistic dynamics of a “free” particle may be defined by an extremal principle which determines the physical spacetime path from x to y as the path with the shortest Minkowski length. This means that the variation

$$\delta \int_x^y ds = \delta \int_{\tau(x)}^{\tau(y)} (\dot{x}^\mu \dot{x}_\mu)^{1/2} d\tau = 0.$$

If we wish to talk about relativistic mechanics in Newtonian terms, i.e., if we wish to use Newtonian concepts like energy, mass, and force and the like – and we might wish to do that because it may then be easier to arrive at Newtonian mechanics

as a limiting description of relativistic mechanics – we can multiply the integral by dimensional constants to get an action integral, so that the terms in the Euler–Lagrange equation read analogously to the Newtonian terms. That is, writing $S = -mc \int ds$ and using (2.39) in the $x^0 = ct$ parametrization, the Lagrange function becomes

$$L(\dot{\mathbf{q}}) = -mc^2 \sqrt{1 - \frac{v^2}{c^2}} .$$

The Euler–Lagrange equations lead to the canonical momentum

$$\mathbf{p} = \frac{m}{\sqrt{1 - v^2/c^2}} \mathbf{v} =: \tilde{m} \mathbf{v} ,$$

from which we recognize m as the rest mass, because $\mathbf{p} \approx m \mathbf{v}$ when $v \ll c$.

The canonical momentum can be taken as the vector part of a canonical four-momentum

$$p^\mu = m \dot{x}^\mu ,$$

the Minkowski length of which is [see (2.38)]

$$g_{\mu\nu} p^\mu p^\nu = p^\mu p_\mu = m^2 \dot{x}^\mu \dot{x}_\mu = m^2 c^2 . \quad (2.40)$$

Hence, parameterizing by $x^0 = ct$, we find

$$p^\mu = m \dot{x}^\mu = \frac{m}{\sqrt{1 - v^2/c^2}} (c, \mathbf{v}) .$$

Observing further that

$$\frac{mc}{\sqrt{1 - v^2/c^2}} \approx mc \left(1 + \frac{1}{2} \frac{v^2}{c^2} + \dots \right) = mc + \frac{m}{2} \frac{v^2}{c} + \dots ,$$

we are led to set $E = p^0 c$. With (2.40), we thus obtain the energy–momentum relation:

$$E = \sqrt{\mathbf{p}^2 c^2 + m^2 c^4} = \tilde{m} c^2 .$$

Now we have for N particles the spacetime trajectories $q_i = (q_i^\mu(\tau_i))$, $\mu = 0, 1, 2, 3$, $i = 1, \dots, N$. Let us introduce a force K^μ , which accelerates the particles. By virtue of (2.38), we have

$$\ddot{x}^\mu \dot{x}_\mu = 0 ,$$

and this suggests

$$K^\mu \dot{x}_\mu = 0 ,$$

that is, the force should be orthogonal to the velocity in the sense of the Minkowski metric. The simplest way to achieve this is to put

$$K^\mu = F^{\mu\nu} \dot{x}_\nu ,$$

with $F^{\mu\nu} = -F^{\nu\mu}$, an antisymmetric tensor of rank 2, i.e., an antisymmetric 4×4 matrix. One way to generate such a tensor is to use a four-potential A^μ :

$$F^{\mu\nu} = \frac{\partial}{\partial x_\mu} A^\nu - \frac{\partial}{\partial x_\nu} A^\mu . \quad (2.41)$$

The Maxwell–Lorentz theory of electromagnetic interaction has the force act on the particles through a law that involves not only masses as parameters, but also charges e_i :

$$m_i \ddot{q}_i^\mu = \frac{e_i}{c} F^\mu{}_\nu(q_i) \dot{q}_i^\nu . \quad (2.42)$$

In view of (2.37), one names the matrix elements as follows:

$$F^\mu{}_\nu(x) = \begin{pmatrix} 0 & E_1 & E_2 & E_3 \\ E_1 & 0 & B_3 & -B_2 \\ E_2 & -B_3 & 0 & B_1 \\ E_3 & B_2 & -B_1 & 0 \end{pmatrix} , \quad (2.43)$$

recalling that indices are lowered or raised by action of $g^{\mu\nu} = g_{\mu\nu}$, whence $F^\mu{}_\nu = F^{\mu\lambda} g_{\lambda\nu}$.

For the three-vector $\mathbf{q}_i(\tau_i)$, we then obtain

$$m_i \ddot{\mathbf{q}}_i = e_i \left(\mathbf{E} + \frac{\dot{\mathbf{q}}_i}{c} \times \mathbf{B} \right) ,$$

where dots over symbols still refer to derivatives with respect to τ_i . For small velocities (compared to the velocity of light), this is close to (2.37).

But the fields $F^\mu{}_\nu$ are themselves generated by the particles, and this is supposed to give the interaction between the charges. The equation which describes the generation of fields is not difficult to guess, by analogy with the gravitational potential which is given by the potential equation $\Delta V = \nabla \cdot \nabla V = \delta(x)$ for a point mass at the origin. (Note that the scalar product construction of the law is a good trick for making the law invariant under Euclidean congruences.) Taking the four-dimensional ∇ in the Minkowski sense suggests the four-dimensional potential equation (invariant under Minkowskian congruences)

$$\left[\left(\frac{\partial}{\partial x^0} \right)^2 - \left(\frac{\partial}{\partial \mathbf{x}} \right)^2 \right] A^\mu = \square A^\mu = \frac{4\pi}{c} j^\mu , \quad (2.44)$$

where j^μ is the current originating from moving particles carrying charge e_i . We discuss the current below.

Note in passing that (2.41) does not determine the vector potential uniquely, because a term of the form $\partial f / \partial x^\mu$ for some well-behaved f can always be added to A^μ without changing the forces. This is called gauge invariance. Equation (2.44) determines the potential in what is known as the Lorentz gauge, where

$$\frac{\partial}{\partial x^\mu} A^\mu = 0 .$$

The current is a distribution, because the charges are concentrated at the positions of the particles, which are points. One could think of smeared out charges, but a charge with a certain extension (a ball, for example) would not be a relativistic object because of Lorentz contraction, i.e., a ball would not remain a ball when moving.⁸ Lorentz contraction is an immediate consequence of the loss of simultaneity, since the length of a rod is defined by the spatial distance of the end points of the rod when taken at the same time.

The current of a point charge is by itself unproblematic. It has the following frame-independent, i.e., relativistic, form:

$$j^\mu(x) = \sum_i e_i c \int_{-\infty}^{\infty} \delta^4(x - q_i(\tau_i)) \dot{q}_i^\mu(\tau_i) d\tau_i , \quad (2.45)$$

with

$$\delta^4(x) = \prod_{\mu=0}^3 \delta(x^\mu) ,$$

and we use x for the four vector $x = (x^0, \mathbf{x})$, since we have used the boldface notation for three-dimensional vectors. Better known is the form in a coordinate frame. With $x^0 = ct$, we obtain (writing the integral as a line integral along the trajectory in the second equality below)

⁸ Suppose one did not care about relativistic invariance, and took a small ball in the rest frame of the electron (for instance, with radius 10^{-16} cm, which seems to be an upper bound from experimental data). Unfortunately, this yields an effective mass of the electron larger than the observed electron mass. The rough argument is that the Coulomb energy of a concentrated charge (infinite for a point charge) yields by the energy–mass relation a field mass which, since it moves with the electron must be accelerated as well, and is effectively part of the electron mass. Extended Lorentz invariant charge distributions entail dynamical problems as well, when strong accelerations occur [1–4].

$$\begin{aligned}
j^\mu(x) &= \sum_i e_i c \int_{-\infty}^{\infty} \delta^4(x - q_i(\tau_i)) \dot{q}_i^\mu(\tau_i) d\tau_i \\
&= \sum_i e_i c \int_{\{q_i(\tau_i) | \tau_i \in \mathbb{R}\}} \delta^4(x - q_i) dq_i^\mu \\
&= \sum_i e_i c \int_{-\infty}^{\infty} \delta(\mathbf{x} - \mathbf{q}_i(t_i)) \delta(ct - ct_i) \frac{dq_i^\mu}{dt_i}(t_i) dt_i \\
&= \sum_i e_i \delta(\mathbf{x} - \mathbf{q}_i(t)) \frac{dq_i^\mu}{dt}(t) .
\end{aligned}$$

From (2.45) we easily obtain the continuity equation

$$\frac{\partial}{\partial x^\mu} j^\mu(x) = 0 .$$

Using the fact that the trajectories are timelike, we have

$$\begin{aligned}
\frac{\partial}{\partial x^\mu} j^\mu(x) &= \sum_i e_i c \int_{-\infty}^{\infty} \frac{\partial}{\partial x^\mu} \delta^4(x - q_i) \frac{dq_i^\mu}{d\tau_i} d\tau_i \\
&= \sum_i e_i c \int_{\{q_i(\tau_i) | \tau_i \in \mathbb{R}\}} - \frac{\partial}{\partial q_i^\mu} \delta^4(x - q_i) dq_i^\mu \\
&= \sum_i e_i c \left[\lim_{\tau_i \rightarrow \infty} \delta^4(x - q_i(\tau_i)) - \lim_{\tau_i \rightarrow -\infty} \delta^4(x - q_i(\tau_i)) \right] \\
&= 0, \quad \forall x \in \mathbb{R}^4 .
\end{aligned}$$

The system of differential equations (2.41)–(2.44) defines the theory of charges interacting via fields. Now (2.42) is unproblematic if the field F^μ_ν is *given* as a well-behaved function. The linear partial differential equation (2.44) is likewise unproblematic if j^μ is *given*, even as a distribution, as in (2.45). Then one has a Cauchy problem, i.e., in order to solve (2.44), one needs initial data $A^\mu(0, x^1, x^2, x^3)$ and $(\partial A^\mu / \partial t)(0, x^1, x^2, x^3)$, and there is no obstacle to finding a solution.

However, we now have to solve (2.42)–(2.44) together, rather than separately, and this does not work. The system of differential equations is only formal, i.e., there are no functions $q^\mu(\tau)$ and $A^\mu(x)$, whose derivatives would satisfy the equations. It does not even matter whether we have more than one particle. Let us take one particle to see what goes wrong. First solve

$$A^\mu(x) = \left(\square^{-1} \frac{4\pi}{c} j^\mu \right) (x) = \int \square_{x,x'}^{-1} \frac{4\pi}{c} j^\mu(x') d^4 x' , \quad (2.46)$$

with a Green's function $\square_{x,x'}^{-1}$ given by

$$\square \square_{x,x'}^{-1} = \delta^4(x - x') . \quad (2.47)$$

The Green's function is not unique. The different possible Green's functions differ by solutions of the homogeneous equation ($j = 0$), i.e., by functions in the kernel of \square . A symmetric choice is⁹

$$4\pi\square_{x,x'}^{-1} = \delta((x-x')^2) = \delta((x^\mu - x'^\mu)(x_\mu - x'_\mu)) .$$

Why is that symmetric? Using

$$\delta(f(x)) = \sum_k \frac{1}{|f'(x_k)|} \delta(x - x_k) , \quad (2.48)$$

where x_k are the single zeros of f we get

$$\delta(x^2 - y^2) = \frac{1}{2} \frac{\delta(x-y)}{y} + \frac{1}{2} \frac{\delta(x+y)}{y} ,$$

and thus

$$\begin{aligned} 4\pi\square_{x,y}^{-1} &= \delta((x-y)^2) = \delta((x^0 - y^0)^2 - (\mathbf{x} - \mathbf{y})^2) \\ &= \frac{1}{2} \frac{\delta((x^0 - y^0) - |\mathbf{x} - \mathbf{y}|)}{|\mathbf{x} - \mathbf{y}|} + \frac{1}{2} \frac{\delta((x^0 - y^0) + |\mathbf{x} - \mathbf{y}|)}{|\mathbf{x} - \mathbf{y}|} , \end{aligned}$$

which is the sum of retarded and advanced Green's function, a notion which will become clear shortly. Any linear combination of these parts is a possible $\square_{x,y}^{-1}$, and one commonly uses only the retarded part, using the argument that one experiences only radiation emitted in the past. We shall say more about that later. One may convince oneself by formal manipulations¹⁰ that (2.47) holds for this or any other linear combination.

Now let us come to the end of the story. With (2.46) and (2.45), we get

$$A^\mu(x) = e \int \delta((x - q(\tau))^2) \dot{q}^\mu(\tau) d\tau , \quad (2.49)$$

⁹ It is quite natural for the Green's function to be like this. It is the most natural relativistic function one can write down. The points which have Minkowski distance zero from one another form the (backward and forward) light cones and they are special in a certain sense. So the function is not eccentric in any way.

¹⁰ For example,

$$\begin{aligned} &\left[\left(\frac{\partial}{\partial x^0} \right)^2 - \Delta \right] \frac{\delta(x^0 - |\mathbf{x}|)}{|\mathbf{x}|} \\ &= \frac{1}{|\mathbf{x}|} \delta'' - \Delta \left(\frac{1}{|\mathbf{x}|} \right) \delta(x^0 - |\mathbf{x}|) - \frac{1}{|\mathbf{x}|} \Delta \delta(x^0 - |\mathbf{x}|) - 2\nabla \left(\frac{1}{|\mathbf{x}|} \right) \cdot \nabla \delta(x^0 - |\mathbf{x}|) \\ &= \frac{1}{4\pi} \delta(\mathbf{x}) \delta(x^0 - |\mathbf{x}|) , \end{aligned}$$

using the chain rule on $\nabla \delta$.

and with (2.48), we find

$$A^\mu = e \frac{\dot{q}^\mu(\tau_a)}{2(x^\mu - q^\mu(\tau_a))\dot{q}_\mu(\tau_a)} + e \frac{\dot{q}^\mu(\tau_r)}{2(x^\mu - q^\mu(\tau_r))\dot{q}_\mu(\tau_r)},$$

where τ_a and τ_r are the solutions of

$$(x - q(\tau))^2 = 0 \quad \Longleftrightarrow \quad ct - c\tau = \mp |\mathbf{x} - \mathbf{q}(\tau)|.$$

We can now understand the terminology used. The retarded/advanced time is given by the intersection of the forward/backward light cone based at \mathbf{x} with the trajectory of the particle (see Fig. 2.5, where \mathbf{x} should be thought of as an arbitrary point).

We have thus derived a field (see for instance [5], if some manipulations seem unclear) A^μ which is well-behaved everywhere *except* at points x lying on the world-line q of the charge generating the field. For $x = q(\tau)$, we have $\tau_a = \tau_r = \tau$, and we see that the denominator is zero. But this is now the end of the story, since these are the x values which are needed in (2.42).

This problem is well known by the name of the electron self-interaction. The field which the electron generates acts back on the electron, and this back-reaction is mathematically ill-defined, since the electron is a point. Hence, the field idea for managing interactions between point charges does not work, unless one introduces formal manipulations like renormalization [6], or changes electromagnetism on small scales [7].

The Maxwell–Lorentz theory of electromagnetism works well (in the sense of describing physical phenomena correctly) when the fields are generated by smeared out charges (charge clouds), so one can describe the radiation from an antenna. It also works when the fields are given as “external” fields, which act on charges by the Lorentz force equation (see [3, 8, 9] for mathematical proofs concerning existence and uniqueness). In short, electromagnetism is fine for most non-academic life. One may ask why this is so. The reason may be that the Maxwell–Lorentz theory is the macroscopic description of the fundamental theory of electromagnetism we shall describe next. But that theory does not contain fields on the fundamental level.

2.5 No fields, Only Particles: Electromagnetism

What is bad about fields when it comes to describing interactions? The problem is that the field naturally acts everywhere, and thus also on the very particle which generates the field. But taking the meaning of relativistic interaction between particles seriously, why does one need fields at all? Why not get rid of the field and have the particles interact directly? In a sense one does this when solving (2.44) for A^μ and putting that into (2.42). Fokker thought that way [10], like many others before him, including Gauss [11], Schwarzschild [12], and Tetrode [13]. He wrote down the variational principle for a relativistic particle theory, which was later rediscov-

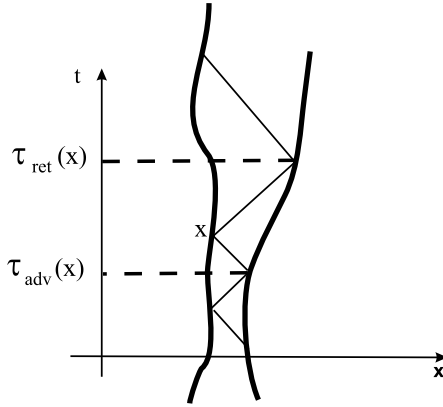


Fig. 2.5 In Feynman–Wheeler electromagnetism particles interact along backward and forward light cones. Here we set $c = 1$

ered by Wheeler and Feynman [14, 15], to explain retarded radiation. How can the particles interact directly in a relativistic way?

There is no natural notion of simultaneity whereby a force can act at the same time between particles, but we already know that the simplest choice is to take the Minkowski distance and to say that particles interact when there is distance zero from one to the other. Hence the particle at spacetime point q interacts with all other particles at spacetime points which are intersection points of light cones based at q with the other trajectories, that is, when

$$(q_i^\mu - q_j^\mu)(q_{i\mu} - q_{j\mu}) = (q_i - q_j)^2 = 0 ,$$

or put another way, when $\delta((q_i - q_j)^2)$ is not zero.

Note, that there are always two light cones based at one point, one directed towards the future and one directed towards the past (see Fig. 2.5), although of course the physical law is not concerned about such notions as past and future.

It is rather clear that dynamics which is defined by future times and past times can no longer be given by differential equations of the ordinary kind. But some differential equations can nevertheless be written down from a variational principle. The Fokker–Wheeler–Feynman action S is the simplest relativistic action one could think of describing interacting particles:

$$S = \sum_i \left[-m_i c \int ds_i - \sum_{j>i} \frac{e_i e_j}{c} \iint \delta((q_i - q_j)^2) dq_i^\mu dq_{j\mu} \right] . \quad (2.50)$$

Writing the trajectories $q_i^\mu(\lambda_i)$ with arbitrary parameters λ_i and using the notation $\dot{q}_i^\mu = dq_i^\mu/d\lambda_i$, we obtain

$$S = \sum_i \left[-m_i c \int (\dot{q}_i^\mu \dot{q}_{i\mu})^{1/2} d\lambda_i - \sum_{j>i} \frac{e_i e_j}{c} \int \int \delta((q_i - q_j)^2) \dot{q}_i^\mu \dot{q}_{j\mu} d\lambda_i d\lambda_j \right].$$

The most noteworthy feature is that there are no diagonal terms in the double sum. This is the major difference with the Maxwell–Lorentz theory, where the diagonal terms should normally be present. The contribution of the i th particle to the interaction reads

$$-\frac{e_i}{c} \int d\lambda_i \dot{q}_i^\mu \int \sum_{j \neq i} e_j \delta((q_i - q_j)^2) \dot{q}_{j\mu} d\lambda_j = -\frac{1}{c^2} \int j_i^\mu(x) A_{i\mu}(x) dx,$$

with the “field”¹¹

$$A_{i\mu}(x) = \sum_{j \neq i} e_j \int \delta((x - q_j)^2) \dot{q}_{j\mu} d\lambda_j. \quad (2.51)$$

In the Wheeler–Feynman formulation, fields [like (2.51)] would only be introduced as a suitable macroscopic description, good for everyday applications of electromagnetism, like handling capacitors. On the fundamental level there is no radiation field and there is no radiation. Therefore, opposite charges may orbit each other without “losing energy” due to radiation. Famous solutions of that kind are known as Schild solutions [16].

Equation (2.49) shows that both the advanced and the retarded Green’s functions appear in A_μ . But it is the “emission of radiation” which we typically see and which is solely described by the retarded Green’s function. Wheeler–Feynman and also Maxwell–Lorentz electromagnetism are time reversible, i.e., the theory does not favor emission before absorption. The typicality of emission has become known as the problem of the electromagnetic arrow of time. The original motivation of Wheeler and Feynman was to reduce this arrow of time to the thermodynamic one, which had been so successfully explained by Boltzmann, by supposing a special initial distribution of the particles in the phase space of the universe. (We shall address this further in the chapter on probability.)

Wheeler and Feynman therefore considered the thermodynamic description of the particle system, i.e., they considered a distribution of charges throughout the universe which “absorb all radiation”. In terms of the theory, this means that the sum of the differences of the retarded and advanced forces over all particles vanish. This is called the absorber condition. This macroscopic theory is still time-symmetric. But supposing further that at some early time the initial distribution of the particles was special (non-equilibrium in some sense), then time-directed radiation phenomena and in particular the observed radiation damping of an accelerated charge are reduced to Boltzmann’s explanation of irreversibility (see Sect. 4.2).

¹¹ We computed this in (2.49), but it is important to understand that this is simply a mathematical expression, which plays no role unless x is a point on the worldlines of the other particles. There is no field in this theory.

We emphasize that the Wheeler–Feynman electromagnetic theory is a mathematically consistent relativistic theory with interaction. In fact it is the only such mathematically well defined and physically relevant theory existing so far, and it is about particles, not fields. The statistical mechanics of the theory leads to the well known description by electromagnetic fields and smeared out charges and is thus experimentally indistinguishable from Maxwell–Lorentz electromagnetism, whenever the latter is well defined, i.e., whenever no point charges are considered to generate fields.

The theory is, however, unfamiliar since its dynamical law is not of the familiar form. The intuition one has from solving differential equations with Cauchy data fails. So why is this? In fact, the Euler–Lagrange equations of (2.50) are not ordinary differential equations, since advanced and retarded times appear in them.¹² In contrast, the Maxwell–Lorentz theory is formally of the ordinary type, but with the serious drawback that the fields render the equations mathematically undefined when they are interpreted as fundamental, i.e., when point charges are considered.

Remark 2.6. On the Nature of Reality

Reality is a curious notion. Physics takes the view that something “out there” exists, and that the world is “made out of something”. This is not curious at all. But it is not so easy to say what it is that the world is made of, since the only access we have to the world is through our senses and our thinking, and communication about the experience we have. What the world is made of is specified by our physical theory about the world, and it is only there where we can see what the world is made of. When our physical theory is about point particles and how they move, then there are point particles out there – if what the theory says about their motion is consistent with our experience, of course. The connection between the entities of the theory and our experience is often complicated, often not even spelt out in any detailed way. Nevertheless, one has some kind of feel for how it works, and a bit of pragmatism in this is alright.

When we wish to explain a physical phenomenon, we reduce it (in the ideal case) to the behavior of the ontological quantities the physical theory is about. In Maxwell–Lorentz electromagnetism, fields are ontological. Switch on your radio. What better explanation is there than to say that the fields are out there, and they get absorbed as radio waves by the radio antenna, and that the radio transforms them back into air waves? Music to your ears. But in Wheeler–Feynman electromagnetism, there are no fields and only particles. It explains the music as well. But the explanation is different [13, 14].

If the Maxwell–Lorentz theory (with point charges) were mathematically consistent, we could choose between fields and particles as being “real”, or only particles as being “real”. Since both would describe the macroscopic world as we see it, our choice would then have to be made on the grounds of simplicity and beauty of the theories. Perhaps in the future we shall find a simpler and nicer theory than the ones we have now, one which is solely about fields. Then only fields will be “real”.

¹² It is an intriguing problem of mathematical physics to establish existence and uniqueness of solutions.

“Reality” thus changes with the nature of our physical theory, and with it the elements which can be measured. In the Wheeler–Feynman theory of electromagnetism, the electromagnetic field cannot be measured. This is because it is not there. It is not part of the theory. That is trivial. Less trivial may seem the understanding that the theory also says what elements are there and how those elements can be measured. In Maxwell–Lorentz theory, the electric field is measured, according to the theory, by its action on charges.

Here is another point one may think about from time to time. Although all variables needed to specify the physical theory are “real”, there is nevertheless a difference. In a particle theory, the particle positions are primitive or primary variables, representing what may be called the *primitive ontology*.¹³ They must be there: a particle theory without particle positions is inconceivable. Particle positions are what the theory is about. The role of all other variables is to say how the positions change. They are secondary variables, needed to spell out the law. We could also say that the particle positions are a priori and the other variables a posteriori. An example of the latter might be the electric field. In fact, secondary variables can be replaced by other variables or can even be dispensed with, as in Wheeler–Feynman electromagnetism. Another example which we did not touch upon at all is general relativity, which makes the Newtonian force obsolete. ■

2.6 On the Symplectic Structure of the Phase Space

With the understanding of tensors and forms, in particular differential forms, not only did mathematics move forward, but so did our insight into physics. It was better appreciated what the objects “really” are. We have an example in our relativistic description of electromagnetism (2.43). The electric and magnetic field strengths are not vector fields, as one learns in school, but rather components of an antisymmetric second rank tensor. Mathematical abstraction helps one to get down to the basics of things, and that is satisfying. One such mathematical abstraction is symplectic geometry. We shall say a few things here mainly to make sure that further mathematical abstractions do not lead to a deeper understanding of the physics which we have presented so far.

Mathematically deeper than conservation of energy and volume is the symplectic structure of phase space, which goes hand in hand with Hamilton’s formulation of mechanics [18]. Symplectic geometry needs a space of even dimensions, and classical physics provides that. Consider the phase space \mathbb{R}^{2n} with coordinates $(q_1, \dots, q_n, p_1, \dots, p_n)$. Given $\mathbf{x}, \mathbf{y} \in \mathbb{R}^{2n}$, let $q_i(\mathbf{x})$ be the projection of \mathbf{x} on the q_i th coordinate axis. Then

$$\omega_i^2(\mathbf{x}, \mathbf{y}) = q_i(\mathbf{y})p_i(\mathbf{x}) - q_i(\mathbf{x})p_i(\mathbf{y}) \quad (2.52)$$

¹³ The notion and role of primitive ontology has long been ignored, but it has recently been revitalized and emphasized in [17].

is the area of the parallelogram generated by \mathbf{x}, \mathbf{y} when projected into the (q_i, p_i) plane. Set

$$\omega^2(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^n w_i^2(\mathbf{x}, \mathbf{y}) = \mathbf{x} \cdot (-I)\mathbf{y} = ((-I)\mathbf{y})^t \mathbf{x}, \quad (2.53)$$

with

$$I = \begin{pmatrix} 0_n & +E_n \\ -E_n & 0_n \end{pmatrix},$$

where E_n is the n -dimensional unit matrix, 0_n is the n -dimensional zero matrix, and \mathbf{z}^t is the transpose of \mathbf{z} , i.e., the row vector, an element in the dual space of \mathbb{R}^{2n} . The 2-form ω^2 , or equivalently the symplectic matrix I , defines the symplectic structure of phase space \mathbb{R}^{2n} ($n = 3N$ for N particles) and gives (like a scalar product) an isomorphism between the vector space and its dual. From courses in analysis, we know that the gradient $\partial f / \partial \mathbf{x}$ of a function $f(\mathbf{x})$, $\mathbf{x} \in \mathbb{R}^d$, is in fact a dual element, i.e., the row vector which acts as a linear map on \mathbf{h} according to $(\partial f / \partial \mathbf{x})\mathbf{h}$ = row times column (matrix multiplication by a vector). Now use ω^2 to identify $\partial f / \partial \mathbf{x}$ with a vector $\nabla_{\omega} f$ (using the Euclidean scalar product, this is the normal ∇f). Given $\mathbf{z} \in \mathbb{R}^{2n}$, the object $\omega^2(\cdot, \mathbf{z})$ is a linear form, i.e., a linear map

$$\begin{aligned} \mathbb{R}^{2n} &\longrightarrow \mathbb{R} \\ \mathbf{x} &\longmapsto \omega^2(\mathbf{x}, \mathbf{z}) = (-I\mathbf{z})^t \mathbf{x}. \end{aligned}$$

which we wish to be equal to $\partial f / \partial \mathbf{x}$. We thus search for a \mathbf{z}_f such that

$$\begin{aligned} \frac{\partial f}{\partial \mathbf{x}} \mathbf{x} &= (-I\mathbf{z}_f)^t \mathbf{x}, \quad \forall \mathbf{x}, \\ \iff \left(\frac{\partial f}{\partial \mathbf{x}} \right)^t &= (\nabla f) = (-I\mathbf{z}_f) \\ \iff \mathbf{z}_f &= \nabla_{\omega} f = I(\nabla f). \end{aligned}$$

It follows that (2.5) can be written in the form

$$\begin{pmatrix} \dot{\mathbf{q}} \\ \dot{\mathbf{p}} \end{pmatrix} = I(\nabla H) = I \begin{pmatrix} \frac{\partial}{\partial \mathbf{q}} H \\ \frac{\partial}{\partial \mathbf{p}} H \end{pmatrix}.$$

The Hamiltonian flow respects the symplectic structure. In particular, it is area preserving, which means the following. Let C be a closed curve in \mathbb{R}^{2n} and define the enclosed area as the sum of the n areas which arise from projections of C onto the coordinate planes (q_i, p_i) [compare with (2.52) and (2.53)]. In the (q_i, p_i) plane, we have a curve C_i with area $A(C_i)$. The area can be transformed into a line integral by

Stokes' theorem in two dimensions:

$$\begin{aligned}
 \int_{A(C_i)} dq_i dp_i &= \int_{A(C_i)} \text{curl} \begin{pmatrix} p_i \\ 0 \end{pmatrix} dq_i dp_i \\
 &= \oint_{C_i} \begin{pmatrix} p_i \\ 0 \end{pmatrix} \cdot \begin{pmatrix} dq_i \\ dp_i \end{pmatrix} \\
 &= \oint_{C_i} p_i dq_i .
 \end{aligned} \tag{2.54}$$

In general,

$$A = \text{area of } C = \oint_C \mathbf{p} \cdot d\mathbf{q} = \sum_i \oint_{C_i} p_i dq_i .$$

Using differential forms,

$$\omega_i^2 = dp_i \wedge dq_i , \quad \omega^2 = \sum_i dp_i \wedge dq_i ,$$

and

$$\omega^2 = d\omega^1 , \quad \omega^1 = \sum_i p_i dq_i .$$

Equation (2.54) is nothing other than

$$\int_{A(C_i)} d\omega^1 = \int_{C_i} \omega^1 .$$

Transporting C with the Hamiltonian flow yields the area $A(t)$, and preservation of area means that $A(t) = A$. By change of variables, the integration over $A(t)$ can be expressed by $\mathbf{q}(t)$ and $\mathbf{p}(t)$ in the integral, so that

$$\begin{aligned}
 \frac{d}{dt} A(t) &= \frac{d}{dt} \oint_C \mathbf{p}(t) \cdot d\mathbf{q}(t) = \oint_C \dot{\mathbf{p}} \cdot d\mathbf{q} + \oint_C \mathbf{p} \cdot d\dot{\mathbf{q}} \\
 &= \oint_C \dot{\mathbf{p}} \cdot d\mathbf{q} - \oint_C \dot{\mathbf{q}} \cdot d\mathbf{p} \quad \left[+ \oint_C d(\mathbf{q}\dot{\mathbf{p}}) = 0 \right] \\
 &= - \oint_C \frac{\partial H}{\partial \mathbf{q}} d\mathbf{q} - \oint_C \frac{\partial H}{\partial \mathbf{p}} d\mathbf{p} \\
 &= - \oint_C dH = 0 .
 \end{aligned}$$

Furthermore, the volume in even-dimensional vector spaces can be thought of as arising from a product of areas (generalizing the area in \mathbb{R}^2 = width times length), i.e., products of two-forms (2.52) yield the volume form:

$$\omega = dp_1 \wedge dq_1 \wedge \dots \wedge dp_n \wedge dq_n ,$$

the Lebesgue measure on phase space (in form language). Then Liouville's theorem arises from preservation of area.

Transformations of coordinates $(\mathbf{q}, \mathbf{p}) \xrightarrow{\psi} (\mathbf{Q}, \mathbf{P})$ are called canonical or symplectic if the Jacobi matrix $\nabla\psi$ is symplectic, which is a variation on the theme of the orthogonal matrix:

$$(\nabla\psi)^t I \nabla\psi = I . \quad (2.55)$$

They preserve areas and satisfy the canonical equations

$$\begin{pmatrix} \dot{\mathbf{Q}} \\ \dot{\mathbf{P}} \end{pmatrix} = I \begin{pmatrix} \frac{\partial}{\partial \mathbf{Q}} \tilde{H} \\ \frac{\partial}{\partial \mathbf{P}} \tilde{H} \end{pmatrix}, \quad \tilde{H} \circ \psi = H .$$

The Poisson bracket (2.11) is invariant under canonical transformations, because (2.11) can be written as

$$\{f, g\} = \nabla f \cdot I \nabla g ,$$

and if $f(\mathbf{Q}, \mathbf{P})$, $g(\mathbf{Q}, \mathbf{P})$, and $(\mathbf{Q}, \mathbf{P}) = \psi(\mathbf{q}, \mathbf{p})$ are given, since we have

$$\nabla(f \circ \psi) = \nabla\psi(\nabla f \circ \psi) ,$$

it follows that

$$\begin{aligned} \{f \circ \psi, g \circ \psi\} &= \nabla(f \circ \psi) \cdot I \nabla(g \circ \psi) \\ &= (\nabla f \circ \psi) \cdot (\nabla\psi)^t I \nabla\psi (\nabla g \circ \psi) \\ &= (\nabla f \circ \psi) \cdot I (\nabla g \circ \psi) \quad [\text{by (2.55)}] \\ &= \{f, g\} \circ \psi . \end{aligned}$$

Clearly,

$$\{q_i, p_i\} = \delta_{ij}, \quad \{q_i, q_j\} = 0, \quad \{p_i, p_j\} = 0 ,$$

and variables which satisfy this are said to be canonical. Of particular interest are variables $(Q_1, \dots, Q_n, P_1, \dots, P_n)$, where P_1, \dots, P_n do not change with time, and where the Hamilton function takes the form

$$\tilde{H}(\mathbf{Q}, \mathbf{P}) = \sum_{i=1}^n \omega_i P_i .$$

Then $\dot{Q}_i = \omega_i$, i.e., $Q_i = \omega_i t + Q_{i,0}$, and Q_i is like the phase of a harmonic oscillator. Such (P_i, Q_i) are called action-angle variables. Systems which allow for such variables are said to be integrable, since their behavior in time is in principle completely under control, with their motion (in the new coordinates) being that of “uncoupled”

harmonic oscillators. The solution can then be found by algebraic manipulation and integration. The Hamiltonian motions in \mathbb{R}^2 (H not time dependent, one particle in one space dimension) are integrable, since H itself does not change with time, and hence one may choose $P = H$. However, integrability is *atypical* for Hamiltonian systems.

References

1. J. Novik: Ann. Phys. **28**, 225 (1964)
2. F. Rohrlich: *Classical Charged Particles*, 3rd edn. (World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2007)
3. H. Spohn: *Dynamics of Charged Particles and the Radiation Field* (Cambridge University Press, 2004)
4. M.K.H. Kiessling: In: *Nonlinear Dynamics and Renormalization Group* (Montreal, QC, 1999), CRM Proc. Lecture Notes, Vol. 27 (Amer. Math. Soc., Providence, RI, 2001) pp. 87–96
5. A.O. Barut: *Electrodynamics and Classical Theory of Fields and Particles* (Dover Publications Inc., New York, 1980). Corrected reprint of the 1964 original
6. P. Dirac: Proc. Roy. Soc. A **178**, 148 (1938)
7. M. Born, L. Infeld: Commun. Math. Phys. A **144**, 425 (1934)
8. W. Appel, M.K.H. Kiessling: Ann. Physics **289** (1), 24 (2001)
9. G. Bauer, D. Dürr: Ann. Henri Poincaré **2** (1), 179 (2001)
10. A.D. Fokker: Z. Physik **58**, 386 (1929)
11. C. Gauss: A letter to Weber on 19 March 1845
12. K. Schwarzschild: Nachr. Ges. Wis. Gottingen **128**, 132 (1903)
13. H. Tetrode: Z. Physik **10**, 317 (1922)
14. J. Wheeler, R. Feynman: Rev. Mod. Phys. **17**, 157 (1945)
15. J.A. Wheeler, R.P. Feynman: Rev. Mod. Phys. **21**, 425 (1949)
16. A. Schild: Phys. Rev. **131**, 2762 (1962)
17. V. Allori, S. Goldstein, R. Tumulka, N. Zanghi: British Journal for the Philosophy of Science (2007)
18. F. Scheck: *Mechanics*, 3rd edn. (Springer-Verlag, Berlin, 1999). From Newton's laws to deterministic chaos. Translated from the German

Bohmian Mechanics

The Physics and Mathematics of Quantum Theory

Dürr, D.; Teufel, S.

2009, XII, 393 p. 41 illus., Hardcover

ISBN: 978-3-540-89343-1