1. Classical Physics

What is classical physics? It became the name for non quantum physics. This begs the question: What is quantum physics? One easily finds a statement like: It is the physics were the position and momentum of a particle are operators. That is as senseless as it reads. Some time ago it was fashionable to say that classical physics describes the world with classical notions like particles moving around in space, while in modern physics, i.e. quantum mechanics, the classical notions are not anymore adequate. Why? Who knows! Here is a more serious statement: Classical physics is the description of the world when the interference effects of the Schrödinger wave function are so small that they can be neglected. That is the case for a tremendously wide range of scales from microscopic gases to stellar matter, in particular it includes the scale of our direct human perception, which is why classical physics was found before quantum mechanics. A theory which governs in this enormous range the behavior of matter is Newtonian mechanics. No wonder that Newton considered himself a giant of physics.

1.1 Newtonian Mechanics

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

$$\boldsymbol{q}_1, \ldots, \boldsymbol{q}_N, \qquad \boldsymbol{q}_i \in \mathbb{R}^3$$

which change with time, so that one has trajectories $\boldsymbol{q}_1(t), \ldots, \boldsymbol{q}_N(t); t \in \mathbb{R}$, where the parameter $t \in \mathbb{R}$ means usually the time.

Newtonian mechanics is given by equations—the physical law—which govern the trajectories, called the equations of motions. Equations of motions 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 mostly be 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" $m_1, \ldots, m_N$, which represent "the matter" and the law reads

$$m_i \ddot{\boldsymbol{q}}_i = \boldsymbol{F}_i(\boldsymbol{q}_1, \ldots, \boldsymbol{q}_N). \tag{1.1}$$

$\boldsymbol{F}_i$ is called the "force". That is in general a function of all particle positions i.e. it is a function of the *configuration* , that is the family of all coordinates $(\boldsymbol{q}_1, \ldots, \boldsymbol{q}_N) \in \mathbb{R}^{3N}$ (the set of all such tuples is called *configuration space* .) $\dot{\boldsymbol{q}}_i = \mathrm{d}\boldsymbol{q}_i/\mathrm{d}t = \boldsymbol{v}_i$ is the velocity of the $i-th$ particle and it's derivative $\ddot{\boldsymbol{q}}_i$ is called 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 with each other: Newtonian mechanics is a theory of interaction. The fundamental interaction is gravitation or mass-attraction given by

$$\boldsymbol{F}_i(\boldsymbol{q}_1, \ldots, \boldsymbol{q}_N) = \sum_{j \neq i} G m_i m_j \frac{\boldsymbol{q}_j - \boldsymbol{q}_i}{\|\boldsymbol{q}_j - \boldsymbol{q}_i\|^3}, \tag{1.2}$$

with G the gravitational constant.

All point particles of the Newtonian universe interact according to (1.2).

In effective descriptions of subsystem (in the way we actually use Newtonian mechanics in everyday life), other forces like harmonic forces of springs can appear on the left hand side of (1.1). Such general forces need not (and in general will not) arise from gravitation alone, electromagnetic forces will also play a role, i.e. we describe electromagnetic interaction between electrically charged particles by the Coulomb force as approximation within Newtonian mechanics. The Coulomb force is similar to (1.2) 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? There are various reasons one can give. For example far away matter which homogeneously surrounds the earth produces a zero net field. Also the force (1.2) falls off with large distances (especially for astronomically large distances it is negligible). That then allows the effective description of a subsystem ignoring far away matter. But of course the effect of distant matter is never exactly zero.

Anmerkung 1.1.1. Initial value problem
The equation (1.1) is a differential equation and thus poses an initial value problem: The trajectories $\boldsymbol{q}_i(t), t \in \mathbb{R}$, which obey (1.1), are only determined once "initial data" are given $\boldsymbol{q}_i(t_0), \dot{\boldsymbol{q}}_i(t_0)$ where t_0 is some time, called the "initial time". This means that the future and the past evolution of the trajectories are determined by the "present" state $\boldsymbol{q}_i(t_0), \dot{\boldsymbol{q}}_i(t_0)$. 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 exists for all times for all initial values. What does exist however is — at least for the gravitation case — 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: (1.1) und (1.2) have no solution if for example two particles occupy the same position. Also for the solution to exist it must not happen that two or more particles collide. It is a famous mathematical physics problem to establish the so called "existence of dynamics" result for a many particle gravitating systems, where one hopes to show, that only for "exceptional initial values" the solutions fail to exist globally. What does "exceptional" mean? We shall answer this in short while.

We wish to shortly comment on the manner of speaking of "interacting" particles, which gives a human touch to Newtonian mechanics: *The particles attract each other*. Taking this notion to heart one might be inclined to associate with the notion of particle more than simply an object which has a position. 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 the change of points in space with time. We shall explore one such prosaic description next.

1.2 Hamilton 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, we simply observe now that it is mathematically much nicer to rewrite everything in terms of configuration space variables

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

i.e. we write the differential equation for all particles in a compact form as

$$m\ddot{\boldsymbol{q}} = \boldsymbol{F} \tag{1.3}$$

with

$$\boldsymbol{F} = \begin{pmatrix} \boldsymbol{F}_1 \\ \cdot \\ \cdot \\ \boldsymbol{F}_N \end{pmatrix}$$

and the mass matrix

$$m = \begin{pmatrix} m_1 & & & & \\ & m_2 & & 0 & \\ & & m_3 & & \\ & & & m_4 & \\ & 0 & & & m_5 \\ & & & & & \ddots \end{pmatrix}.$$

Configuration space cannot be depicted ((see 1.1) 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 better builds some intuition for configuration space because it plays a fundamental role in quantum theory.

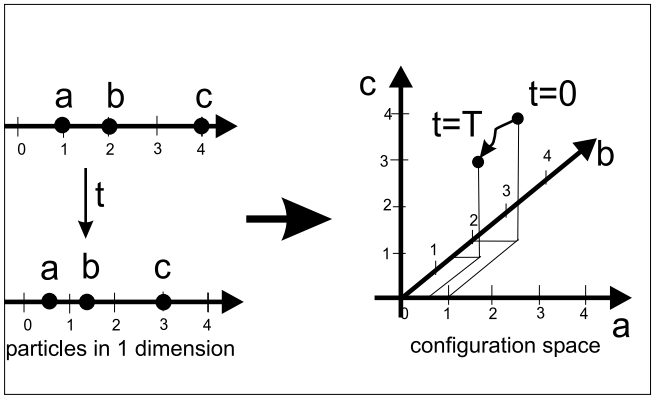

Abb. 1.1. configuration space for 3 particles in a one dimensional world

A differential equation is *by definition* a relation between the *flow* and the *vector field*. The flow is the mapping along the 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.

The differential equation (1.3) is of second order and does not express transparently the relation between integral curves and vector field. We need to change (1.3) into an equation of first order so that the vector field becomes transparent.

For that one considers the *phase space* variables

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

which was introduced by Boltzmann[1], where we consider positions and velocities, the latter however for convenience of notation as momenta– $\boldsymbol{p}_i = m_i \boldsymbol{v}_i$. One point in Γ represents the present state of the entire N-particle system. The dimension of phase space has twice the dimension of configuration space and can be depicted for one particle moving in one dimension, for example the pendulum (See figure 1.2).

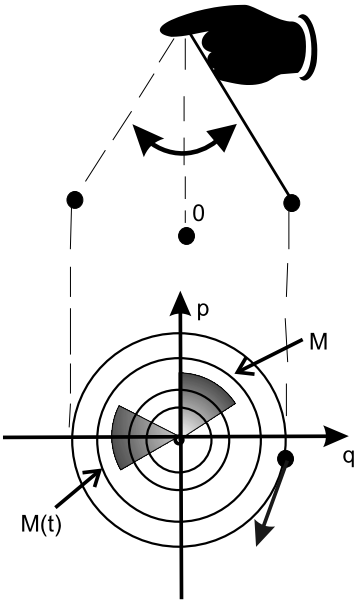

Abb. 1.2. Phase space description of the mathematically idealized harmonically swinging pendulum. The possible trajectories of the mathematically idealized pendulum swinging in a plane with frequency 1 are concentric circles in phase space. The sets M und $M(t)$ will be discussed later.

(1.3) becomes obviously

. (1.4)

[1] The notion of phase space was used by Ludwig Boltzmann (1844–1906) as synonymous to state space, the phase being the collection of variables which uniquely determine the physical state. The physical state is uniquely determined if it's future and past evolution in time is uniquely determined given the physical law.

The state of the N-particle system is completely determined by $\begin{pmatrix} \boldsymbol{q} \\ \boldsymbol{p} \end{pmatrix}$ because (1.4) and the "initial values" $\begin{pmatrix} \boldsymbol{q}(t_0) \\ \boldsymbol{p}(t_0) \end{pmatrix}$ determine uniquely the phase space trajectory (if the initial value problem allows for a solution.)

For(1.2) and many other effective forces there exists a function V on $\mathbb{R}^{3N}$, the so called potential energy function such that $\boldsymbol{F} = -\text{grad}_q V = -\frac{\partial V}{\partial \boldsymbol{q}} = -\nabla V$. Using this we may write (1.4) as

$$\begin{pmatrix} \dot{\boldsymbol{q}} \\ \dot{\boldsymbol{p}} \end{pmatrix} = \begin{pmatrix} \frac{\partial H}{\partial \boldsymbol{p}}(\boldsymbol{q}, \boldsymbol{p}) \\ \\ -\frac{\partial H}{\partial \boldsymbol{q}}(\boldsymbol{q}, \boldsymbol{p}) \end{pmatrix}, \tag{1.5}$$

$$\text{where } H(\boldsymbol{q}, \boldsymbol{p}) = \frac{1}{2}(\boldsymbol{p} \cdot m^{-1}\boldsymbol{p}) + V(\boldsymbol{q}) \tag{1.6}$$

$$= \frac{1}{2}\sum_{i=1}^{N} \frac{\boldsymbol{p}_i^2}{m_i} + V(\boldsymbol{q}_1, \ldots, \boldsymbol{q}_N).$$

Now we have the Newtonian law in the form of a transparent differential equation (1.5), expressing the relation between the integral curves (on the left hand side differentiated to have tangent vectors) and the vector field on the right hand side(which are the tangent vectors specifying the physics). The way we wrote it, the vector field is actually generated by a function H (1.6) on phase space, called Hamilton function after its inventor William Rowan Hamilton (1805-1865), who in fact did introduce the Symbol H in honor of the physicist Christiaan Huygens (1629-1695) and we shall say later what the "wave man" Huygens has to do with all this. The role of the Hamilton function $H(\boldsymbol{q}, \boldsymbol{p})$ is to give the vector field

$$\boldsymbol{v}^H(\boldsymbol{q}, \boldsymbol{p}) = \begin{pmatrix} \frac{\partial H}{\partial \boldsymbol{p}} \\ -\frac{\partial H}{\partial \boldsymbol{q}} \end{pmatrix}. \tag{1.7}$$

and the Hamiltonian dynamics is simply given by

$$\begin{pmatrix} \dot{\boldsymbol{q}} \\ \dot{\boldsymbol{p}} \end{pmatrix} = \boldsymbol{v}^H(\boldsymbol{q}, \boldsymbol{p}). \tag{1.8}$$

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 shortly see examples. The equations (1.5) and (1.6) with Hamilton function $H(\boldsymbol{q}, \boldsymbol{p})$ define a Hamiltonian system.

The integral curves along this vector field (1.7) are the possible system trajectories in phase space, i.e. the are solutions $\begin{pmatrix} \boldsymbol{q}(t, (\boldsymbol{q}, \boldsymbol{p})) \\ \boldsymbol{p}(t, (\boldsymbol{q}, \boldsymbol{p})) \end{pmatrix}$ of (1.8) to

given initial values $\begin{pmatrix} \boldsymbol{q}(0,(\boldsymbol{q},\boldsymbol{p})) \\ \boldsymbol{p}(0,(\boldsymbol{q},\boldsymbol{p})) \end{pmatrix} = \begin{pmatrix} \boldsymbol{q} \\ \boldsymbol{p} \end{pmatrix}$. Note that that requires existence and uniqueness of solutions of the differential equations (1.8). One possible evolution of the entire system is represented by one curve in phase space (see figure 1.3) — 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 within time t under the evolution (if that evolution is defined, see *Remark* 1.1.1) :

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

We say something more to the flow map later on. The flow can be thought of pictorially as flow of a material fluid in Γ with the (system!-)trajectories as flow lines.

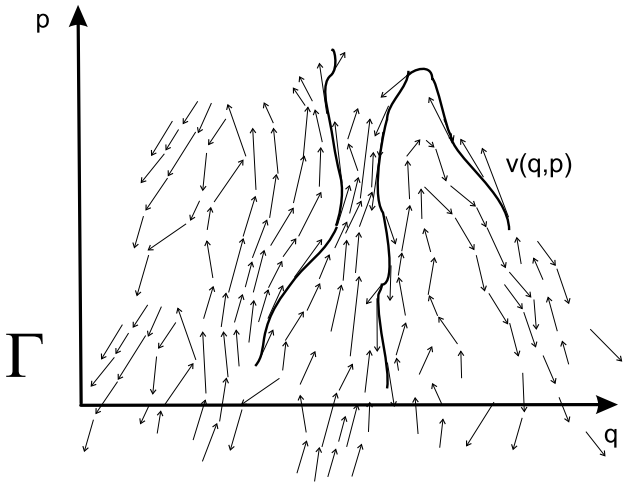

Abb. 1.3. The Hamilton function generates a vector field on the $6N$-dimensional phase space of an N-particle system in physical space. The integral curves are the possible trajectories of the entire system in phase space. Each point in phase space is the collection of all the positions and velocities 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 in every point of phase space there is a unique vector. Trajectories in phase space do not interact with each other! No way! They are not the trajectories of particles!

Hamiltonian mechanics is another way to talk about Newtonian mechanics; it is a prosaic way to talk about the motion of particles, the only romance left is the secret how to write down the physically relevant H, once that is

done, the romance is over and one has before one the laws of mechanics written in mathematical language. That is it. The advantage of the Hamiltonian form is that it directly expresses the law as differential equation in its very meaning (1.8). And it has the advantage that it allows one to talk simultaneously about all possible trajectories of a system. That will be helpful to define a *typical* trajectory of the system, which we must and shall do later.

This however does not mean that we should forget the Newtonian way altogether. By no means. 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 intuition about that. Moreover we should not loose sight of what we are interested in: The behavior of the system in physical space. Although we did not touch 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 one confidence in the correctness of the law. Of course one also achieves confidence by checking whether the theory describes correctly what we see, but since we can usually only see a tiny fraction of what a theory says, confidence is mostly 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 equations (1.8) with (1.7), i.e. $H(\boldsymbol{q},\boldsymbol{p})$ can be a completely general function of $(\boldsymbol{q},\boldsymbol{p})$ and need not be the function (1.6). In such generality one calls $\boldsymbol{p}$ the canonical momentum which is not anymore velocity times mass. Now, conservation of energy means that the value of the Hamilton function does not change along trajectories. That is easy to see: Let $(\boldsymbol{q}(t),\boldsymbol{p}(t)),\ t\in\mathbb{R}$ be a solution of (1.8) then:

$$\frac{\mathrm{d}}{\mathrm{d}t}H(\boldsymbol{q}(t),\boldsymbol{p}(t)) = \dot{\boldsymbol{q}}\frac{\partial H}{\partial \boldsymbol{q}} + \dot{\boldsymbol{p}}\frac{\partial H}{\partial \boldsymbol{p}} = \frac{\partial H}{\partial \boldsymbol{p}}\frac{\partial H}{\partial \boldsymbol{q}} - \frac{\partial H}{\partial \boldsymbol{q}}\frac{\partial H}{\partial \boldsymbol{p}} = 0. \tag{1.9}$$

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

$$\frac{\mathrm{d}}{\mathrm{d}t}f(\boldsymbol{q}(t),\boldsymbol{p}(t)) = \dot{\boldsymbol{q}}\frac{\partial f}{\partial \boldsymbol{q}} + \dot{\boldsymbol{p}}\frac{\partial f}{\partial \boldsymbol{p}} = \frac{\partial H}{\partial \boldsymbol{p}}\frac{\partial f}{\partial \boldsymbol{q}} - \frac{\partial H}{\partial \boldsymbol{q}}\frac{\partial f}{\partial \boldsymbol{p}} =: \{f,H\}, \tag{1.10}$$

abstractly

$$\frac{\mathrm{d}}{\mathrm{d}t}f\circ\Phi_t = \{f\circ\Phi_t, H\}.$$

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 Hamilton function with Φ_t^g as the flow generated by g

$$\{f,g\} = \frac{\mathrm{d}}{\mathrm{d}t}f\circ\Phi_t^g = \frac{\partial g}{\partial \boldsymbol{p}}\frac{\partial f}{\partial \boldsymbol{q}} - \frac{\partial f}{\partial \boldsymbol{q}}\frac{\partial f}{\partial \boldsymbol{p}}. \tag{1.11}$$

Note, that $\{f, H\} = 0$ means that f is a *constant of motion* i.e. the value of f remains unchanged along a trajectory ($\mathrm{d}f/\mathrm{d}t = 0$), $f = H$ 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 be pictured as a "fluid flow" in Γ with the (system!-) trajectories as flow lines, which are the integral curves along the Hamiltonian vector field $\boldsymbol{v}(\boldsymbol{q}, \boldsymbol{p})$ (1.7). These flow lines have neither sources nor sinks, i.e. the vector field is divergence free:

$$\begin{aligned} \mathrm{div}\boldsymbol{v}(\boldsymbol{q},\boldsymbol{p}) &= \left(\frac{\partial}{\partial \boldsymbol{q}}, \frac{\partial}{\partial \boldsymbol{p}}\right) \begin{pmatrix} \frac{\partial H}{\partial \boldsymbol{p}} \\ -\frac{\partial H}{\partial \boldsymbol{q}} \end{pmatrix} \\ &= \frac{\partial^2 H}{\partial \boldsymbol{q}\partial \boldsymbol{p}} - \frac{\partial^2 H}{\partial \boldsymbol{p}\partial \boldsymbol{q}} = 0 \ . \end{aligned} \tag{1.12}$$

This important (though rather trivial) mathematical fact is known as Liouville's theorem (Liouville (1809–1882)) for the Hamiltonian flow (it has nothing to do with Liouville's theorem of complex analysis.) One calls a fluid which flows divergence free "incompressible", a behavior differently from air in an air 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:

Anmerkung 1.2.1. On the time evolution of measures.
The notion of "volume" deserves a bit of elaboration. Clearly, since phase space is very high dimensional, "volume" is more abstract than volume of a three dimensional object. In fact we shall use later on a notion of "volume" which is not simply the trivial extension of three dimensional volume. Volume means *measure*, the size or weight or whichever intuitive expression one might favor of sets, where in general one may want to consider a biased weight. The most famous such measure and in fact the mother of all measures is the generalization of the volume of a cube to arbitrary subsets, the Lebesgue-measure λ. We shall say more to that later. Not needed now[2]! If one feels intimidated by the name Lebesgue measure, then take $|A| = \int_A \mathrm{d}^n x$—usual Riemann integration—as Lebesgue measure of A (for all practical purposes correct). The measure may in a more general sense be also thought of as some kind of weight distribution where the Lebesgue measure gives equal

[2] We need to deal with the curse of the continuum: Not all subsets of $\mathbb{R}^n$ 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 are mathematically existent but 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 nevertheless they are there and thus must be dealt with properly and we do so in section **??**

(i.e. unbiased) weight to every point, which for a continuum of points is a somewhat demanding notion, but one nevertheless may feel what is meant. For the time being we may require that the measure be an additive positive set function, i.e. a function which gives positive values to sets and which is additive on disjoint sets: $\mu(A \cup B) = \mu(A) + \mu(B)$. The role of the measure will eventually be to tell us the size of a set, which sets are small and which are big. Big sets are important small sets are not—almost humane.

It may be best to think for now of a measure in general as some abstract way of giving "mass" to subsets and the question is: How does a measure (or the mass) change 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 phase space. Next we consider a general (not necessarily Hamiltonian) flow map on phase space (see (1.19)) below for a definition.

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

$$\mu_t = \mu \circ \Phi_{-t} \tag{1.13}$$

which means

$$\mu_t(A) = \mu(\Phi_{-t}A) \tag{1.14}$$

(or $\mu_t(\Phi_t A) = \mu(A)$) 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 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.

Very important is the notion of stationary measure . That is a measure which does not change under the flow Φ_t:

$$\mu_t(A) = \mu(A), \forall t \in \mathbb{R}, \forall A. \tag{1.15}$$

The stationary measure plays a distinguished role for justifying probabilistic reasoning. It's importance was presumably first discovered by Boltzmann and we shall spend some time on Boltzmann's general ideas on statistical physics which are valid for all kind of theories. But not now.

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

$$\lambda(\Phi_{-t}A) = \lambda(A). \tag{1.16}$$

with λ the Lebesgue measure on phase space. That means, that under the Hamiltonian flow sets change their shape but not their volume, which may also be referred to as Liouville's theorem, since it is a direct consequence of (1.12), as we shall show next. For the pendulum in figure (1.2) one sees that immediately since the "slices of pie" just rotate and the general situation is depicted in figure 1.4.

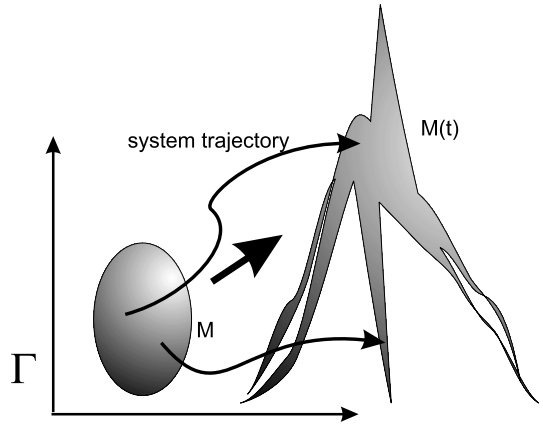

Abb. 1.4. The preservation of volume of the Hamiltonian flow. The set M in phase space changes under the flow but its volume remains the same.
We shall now show that come about.

Anmerkung 1.2.2. Continuity equation

One can derive from the change of measure (1.14) for a general flow generated by a vector field a differential equation, governing the change of the *density* of the measure. That differential equation is called continuity equation. If the measure has a density (you may think of a mass density) $\varrho(\boldsymbol{x})$ then the change of measure with time defines a time dependent density $\varrho(\boldsymbol{x}, t)$ and one has the logical relation[3]

$$\begin{aligned} \mu_t(A) &=: \int \chi_A(\boldsymbol{x})\varrho(\boldsymbol{x},t)\mathrm{d}^n x \\ = \mu(\Phi_{-t}(A)) &=: \int \chi_{\Phi_{-t}(A)}(\boldsymbol{x})\varrho(\boldsymbol{x})\mathrm{d}^n x \\ &= \int \chi_A(\Phi_t\boldsymbol{x})\varrho(\boldsymbol{x})\mathrm{d}^n x\,, \end{aligned} \tag{1.17}$$

where χ_A means the characteristic function of the set A, i.e. the function which is 1 on A and zero otherwise. Furthermore Φ_t is the solution flow map of some vector field $\boldsymbol{v}(\boldsymbol{x})$ on $\mathbb{R}^n$ (or some general phase space), i.e.

$$\frac{\mathrm{d}}{\mathrm{d}t}\Phi_t(\boldsymbol{x}) = \boldsymbol{v}(\Phi_t(\boldsymbol{x})). \tag{1.18}$$

Actually the family $(\Phi_t)_{t\in\mathbb{R}}$ is a one parameter group (the "time" t) of maps, i.e.

[3] Note aside that $\varrho(\boldsymbol{x}, t)$ can be computed from an obvious change of variables in the last integral: $\varrho(\boldsymbol{x}, t) = \varrho(\boldsymbol{x})|\frac{\partial\Phi_{-t}\boldsymbol{x}}{\partial\boldsymbol{x}}|$

$$(\Phi_t(\boldsymbol{x}))_{t\in\mathbb{R}},\ \ \boldsymbol{x}\in\mathbb{R}^n:\ \Phi_t\circ\Phi_s(\boldsymbol{x})=\Phi_{t+s}(\boldsymbol{x}),\ \ \Phi_0(\boldsymbol{x})=\boldsymbol{x}. \tag{1.19}$$

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

$$\frac{\partial}{\partial t}\varrho(\boldsymbol{x},t)+\operatorname{div}(\boldsymbol{v}(\boldsymbol{x})\varrho(\boldsymbol{x},t))=0. \tag{1.20}$$

To see that replace in (1.17) the indicator function by a smooth function f which has compact support

$$\int f(\Phi_t(\boldsymbol{x}))\varrho(\boldsymbol{x})\mathrm{d}^n x=\int f(\boldsymbol{x})\varrho(\boldsymbol{x},t)\mathrm{d}^n x\,, \tag{1.21}$$

Differentiate now (1.21) with respect to t and get

$$\int\frac{\mathrm{d}\Phi_t(\boldsymbol{x})}{\mathrm{d}t}\cdot(\nabla f(\Phi_t(\boldsymbol{x})))\varrho(\boldsymbol{x})\mathrm{d}^n x=\int f(\boldsymbol{x})\frac{\partial}{\partial t}\varrho(\boldsymbol{x},t)\mathrm{d}^n x. \tag{1.22}$$

Replace $\frac{\mathrm{d}\Phi_t(\boldsymbol{x})}{\mathrm{d}t}$ with the right hand side of (1.18) and use (1.21) again with f replaced by $\boldsymbol{v}(\boldsymbol{x},t)\cdot(\nabla f(\boldsymbol{x}))$ (that is a wonderful trick) to get for the left hand side of (1.22)

$$\int\boldsymbol{v}(\boldsymbol{x},t)\cdot(\nabla f(\boldsymbol{x}))\varrho(\boldsymbol{x},t)\mathrm{d}^n x \tag{1.23}$$

and after partial integration

$$-\int f(\boldsymbol{x})\nabla\cdot(\boldsymbol{v}(\boldsymbol{x},t)\varrho(\boldsymbol{x},t))\mathrm{d}^n x. \tag{1.24}$$

This being equal to the right hand side of (1.22) we conclude (since f is arbitrary) that (1.20) holds.

Now we ask whether there is a stationary measure (1.15). In terms of the densities the question is: Does there exist a stationary density that is a density independent of time, satisfying (1.20)? Since the time derivative part in (1.20) vanishes: $\frac{\partial}{\partial t}\varrho(\boldsymbol{x})=0$, the density must satisfy the partial differential equation

$$\operatorname{div}(\boldsymbol{v}(\boldsymbol{x})\varrho(\boldsymbol{x}))=0. \tag{1.25}$$

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. There the answer is trivial! Because of (1.12)! Setting $\boldsymbol{x}=\begin{pmatrix}\boldsymbol{q}\\ \boldsymbol{p}\end{pmatrix}$ (1.12) reads $\operatorname{div}\boldsymbol{v}(\boldsymbol{x})=0$ for all $\boldsymbol{x}$. In this case (1.20) becomes (after using product rule)

$$\frac{\partial}{\partial t}\varrho(\boldsymbol{x},t)+\boldsymbol{v}(\boldsymbol{x})\cdot\nabla\varrho(\boldsymbol{x},t)=0. \tag{1.26}$$

$\varrho(\boldsymbol{x},t)=$ const. is obviously stationary.

Put now $f = \chi_A$, $A \subset \Gamma$. Then $\varrho = 1$ yields for (1.21)

$$\int \chi_A(\Phi_t(\boldsymbol{x}))\mathrm{d}^n x = \int \chi_{\Phi_{-t}A}(\boldsymbol{x})\mathrm{d}^n x = \int \chi_A(\boldsymbol{x})\mathrm{d}^n x = \lambda(A), \tag{1.27}$$

where $\Phi_{-t}A = \{(\boldsymbol{q},\boldsymbol{p}) \in \Gamma \mid \Phi_t(\boldsymbol{q},\boldsymbol{p}) \in A\}$.

In short, (1.27) says

$$\lambda(\Phi_{-t}A) = \lambda(A). \tag{1.28}$$

We may as well put into (1.28) instead of A the future set $\Phi_t A$,and use $\Phi_{-t}\Phi_t = \mathrm{id}$ (identity) so that

$$\lambda(A) = \lambda(\Phi_t A) \tag{1.29}$$

Concluding, *Liouville's theorem* asserts that the Lebesgue measure ($\equiv$ volume) is stationary for Hamiltonian flows.

Anmerkung 1.2.3. Time dependent vector fields
The continuity equation holds also for time dependent vector fields $\boldsymbol{v}(\boldsymbol{x},t)$, 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 in (1.20) the vector field by the time dependent expression $\boldsymbol{v}(\boldsymbol{x},t)$, the proof goes through verbatim. Now the notion of stationary measure seems unfitting, 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) does not change under the flow.

Anmerkung 1.2.4. On the initial value problem (1.1.1)
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: One is happy, if one can show 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, is small in the sense of measure but not necessarily small in the sense of cardinality (number of points in the set). The famous Cantor set as subset of the interval $[0,1]$ has as much numbers as the reals but has zero Lebesgue measure.

We close this section with a for modern physicists heretical thought, namely we imagine a Newtonian universe. That is not physical (we know there is Quantum mechanics and much more), but we nevertheless can think of it and it is good enough to ask a question which in an appropriate sense can be asked in any other theory. *Which initial values give rise to OUR Newtonian universe? According to which criteria were the initial values of OUR universe chosen?* We do not ask, who chose the initial conditions. We ask: Which physical law determines them? One possibility to answer this could be that our universe is nothing special, that it could be *typical* in the sense that almost all (in the sense that the Lebesgue measure of the set which does not

give rise to a universe like ours is very small) initial conditions would give rise to a universe very much like ours. That is however not the case, but that we shall address later.

1.2.1 Hamilton-Jacobi Formulation

The Hamiltonian structure and phase space are intimately connected with symplectic geometry. We shall say a bit about that at the end of the chapter.

But let us move on to another question. The Hamiltonian formulation of Newtonian mechanics is prosaic and brings out the particular structure shared by Newtonian mechanics with 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 tried for an analogy of mechanics with wave optics (Huyghens' principle , that's Hamilton's H) and Fermat's extremal principle of geometric optics, according to which light rays takes the path of the shortest time and moreover they 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 $\delta \int L\mathrm{d}t = 0$, $L(\boldsymbol{q}, \dot{\boldsymbol{q}})$ being the so called Lagrange function. The mechanical trajectories (the Newtonian trajectories) between $t_0, \boldsymbol{q}_0$ and $t, \boldsymbol{q}$ (note that instead of initial position and initial velocity we consider initial position and end position) are characterized as the extremals of

$$\int_{t_0}^{t} L(\boldsymbol{q}(t'), \dot{\boldsymbol{q}}(t'))\mathrm{d}t' \,.$$

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

$$L(\boldsymbol{q}, \dot{\boldsymbol{q}}) = \frac{1}{2}\dot{\boldsymbol{q}} \cdot m\dot{\boldsymbol{q}} - V(\boldsymbol{q}) \,, \tag{1.30}$$

which yield the Newtonian equations as Euler-Lagrange equations. Lagrange function and Hamilton function are Legendre transforms of another[4](what is remarkable here is that almost all great mathematicians of the 19th century have left some trace in theoretical mechanics). Starting with H you get L by changing from the variable $\boldsymbol{p}$ to $\dot{\boldsymbol{q}}$, meaning that $(\boldsymbol{q}, \boldsymbol{p})$ gets replaced by $(\boldsymbol{q}, \dot{\boldsymbol{q}})$, using the implicitly given function $\dot{\boldsymbol{q}} = \frac{\partial H(\boldsymbol{q},\boldsymbol{p})}{\partial \boldsymbol{p}}$, where the equation is solved by $\boldsymbol{p}$ as function of $\dot{\boldsymbol{q}}$.

[4] Here is the definition: Let $f(x)$ be convex and let z be a given slope, search the point $x(z)$, in which the tangent on the graph of f has slope z. You find $x(z)$ by minimizing $F(x, z) = f(x) - xz$ in x, which by convexity of f is uniquely determined. The Legendre transform of f is $g(z) = F(x(z), z)$.

For "normal" Hamilton functions (quadratic in the momentum) that solution is immediate and looking at (1.30) the Legendre transform pops right up:

$$L(\boldsymbol{q}, \dot{\boldsymbol{q}}) = \boldsymbol{p} \cdot \dot{\boldsymbol{q}} - H(\boldsymbol{q}, \boldsymbol{p}). \tag{1.31}$$

Note aside that if one starts with the least action principle as being fundamental one can guess the form of the Lagrange function from so called first principles like symmetry, homogeneity, simplicity. But that is not what we wish to recount here.

We wish to come to Huyghens principle and the definition of waves $S_{\boldsymbol{q}_0}(\boldsymbol{q}, t)$ which guide mechanical trajectories starting at $\boldsymbol{q}_0$ moving along the vector field $\boldsymbol{p}(\boldsymbol{q}, t) = \nabla S_{\boldsymbol{q}_0}(\boldsymbol{q}, t)$.

Hamilton suggested

$$S_{\boldsymbol{q}_0, t_0}(\boldsymbol{q}, t) := \int_{t_0}^{t} L(\gamma, \dot{\gamma}) \mathrm{d}t \tag{1.32}$$

where $\gamma : \boldsymbol{q}_0, t_0 \longrightarrow \boldsymbol{q}, t$ is the extremum of the action principle, i.e. the Newtonian path.

This definition leads unfortunately in general to a multi-valued function: Take for example the harmonic oscillator with period T. There are very many extremal trajectories for a harmonic oscillator having period T which go from $(0, T)$ to $(0, 2T)$, that is, S is not uniquely defined. Or think of a ball which may bounce off a wall. The position $\boldsymbol{q}$ before the wall can always be reached within a given time on two ways. Once with and once without reflection on the wall.

However the definition is good for short enough times but never mind such nuisance, let us finish the thought: Ignoring the dependence on $(\boldsymbol{q}_0, t_0)$ and considering [5]

$$\mathrm{d}S = \frac{\partial S}{\partial \boldsymbol{q}} \mathrm{d}\boldsymbol{q} + \frac{\partial S}{\partial t} \mathrm{d}t,$$

we can identify in view of (1.31)

$$\mathrm{d}S = \boldsymbol{p} \mathrm{d}\boldsymbol{q} - H(\boldsymbol{q}, \boldsymbol{p}) \mathrm{d}t$$

and thus by comparison

$$\boldsymbol{p}(\boldsymbol{q}, t) = \frac{\partial S}{\partial \boldsymbol{q}}(\boldsymbol{q}, t) \tag{1.33}$$

hence

$$\frac{\partial S}{\partial t}(\boldsymbol{q}, t) + H\left(\boldsymbol{q}, \frac{\partial S}{\partial \boldsymbol{q}}\right) = 0. \tag{1.34}$$

[5] We may ignore that dependence because we assume uniqueness of the trajectory which yields that $S_{\boldsymbol{q}_0, t_0}(\boldsymbol{q}, t) = S_{\boldsymbol{q}_1, t_1}(\boldsymbol{q}, t) + S_{\boldsymbol{q}_0, t_0}(\boldsymbol{q}_1, t_1)$ i.e. the vector field (1.33) does not depend on the choice of $(\boldsymbol{q}_0, t_0)$.

This is the known as Hamilton-Jacobi-differential equation. For Newtonian mechanics, where $\dot{\boldsymbol{q}}_i = \boldsymbol{p}_i/m_i$ we obtain then the following picture: We have defined on configuration space $\mathbb{R}^n$ of N particles ($n = 3N$) a function $S(\boldsymbol{q}, t)$ ("unfortunately" multi-valued) whose role it is to generate a vector field

$$\boldsymbol{v}(\boldsymbol{q}, t) = m^{-1}\nabla S(\boldsymbol{q}, t) \tag{1.35}$$

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

$$\mathrm{d}\boldsymbol{Q}/\mathrm{d}t = \boldsymbol{v}(\boldsymbol{Q}(t), t)\,.$$

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

$$\frac{\partial S}{\partial t}(\boldsymbol{q}, t) + \frac{1}{2}\sum_{i=1}^{N}\frac{1}{m_i}\left(\frac{\partial S}{\partial \boldsymbol{q}_i}\right)^2 + V(\boldsymbol{q}) = 0. \tag{1.36}$$

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

1.3 Fields and Particles: Electromagnetism

Many hold the view that a 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 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 give a short look at electromagnetism where dynamical fields come into play as fundamental objects which are needed to describe interaction between particles, which carry a "charge". Electro-magnetic fields act on a particle at position $\boldsymbol{q} \in \mathbb{R}^3$ via the Lorentz force:

$$m\ddot{\boldsymbol{q}} = e\left(\boldsymbol{E}(\boldsymbol{q}, t) + \frac{\dot{\boldsymbol{q}}}{c} \times \boldsymbol{B}(\boldsymbol{q}, t)\right), \tag{1.37}$$

where $\boldsymbol{E}(\boldsymbol{q}, t)$ und $\boldsymbol{B}(\boldsymbol{q}, t)$ are electric and magnetic field and c is the velocity of light. While the fields act on particles as described, in electro magnetism the fields are not independent agents which have a kingdom of their own but they are generated by particles. They are generated by particles and they act on particles, that is why one may say that they are there to handle the interaction between particles. But the particles are point particles (what else could they be) and that does not go well with fields. We shall explain this now. We shall use the opportunity to also introduce relativistic physics.

Albert Einstein (1879–1955) deduced from Maxwell's equations of electromagnetism that space and most importantly time change differently from

Galilean physics when one changes to moving frames. The nature of their change is governed by the fact, that the velocity of light does not change when one changes to a moving frame of reference. This led to the understanding (and soon to the four dimensional description by Minkowski(1864–1909)) that space and time are "of the same kind": A particle needs for its specification not only a position in space but also one in time, meaning that the coordinates of a particle in relativistic physics are space and time coordinates. This is a revolution in Newtonian mechanics, where we of course are used to particles having different positions but not different times. So one must get used in relativistic physics to the space-time description of particles, each particle having its own space-time coordinates. Like there is no natural simultaneity of particle positions (the particles would all sit on top of each other) there is now no natural simultaneity in time either (which there is Newtonian mechanics, called the absolute time). In other words, the configuration space of classical mechanics where we collect all positions of the particles of a system at the same time, plays no natural role anymore. Instead Einstein believed that the overturn of absolute time brought physics closer to the true description of nature and he believed that physical theories must be local in the sense, that no physical effects can move faster than light. John Bell showed that that is wrong. Until today that seems to be the only scientific mistake Einstein did. We shall devote to this a whole chapter later but now we must move on.

Minkowski introduced the so called four dimensional space-time with a particular scalar product[6]. In space time particles do not move anymore in a Newtonian way, but according to new dynamics.

A particle's position is now x^μ, $\mu = (0, 1, 2, 3)$ where x^0 is selected as time coordinate since it is distinguished by the "signature" of the so called Minkowski-length[7] ($\mathrm{d}s^2 = \mathrm{d}{x^0}^2 - d\boldsymbol{x}^2 = \mathrm{d}{x^0}^2 - \sum_{i=1}^3 \mathrm{d}{x^i}^2$). In Newtonian mechanics we are used to parameterize paths by time, which is not anymore natural. A natural parameter is now length—Minkowski length—normalized

[6] Minkowski suggested to use as space-time coordinates for the time coordinate imaginary numbers $(x_0 = ict, \boldsymbol{x})$ because then the formal Euclidean scalar product yields the Minkowski-Metrik $s^2 = {x^0}^2 + \boldsymbol{x}^2 = -ct^2 + \sum_{i=1}^3 {x^i}^2$). The advantage is that now all congruences, i.e. the transformations which leave the scalar product invariant (the so called Lorentz group) appear in Euclidean disguise and hence are to be seen as four dimensional rotations and reflections. This is different from the Galilean case where the change to relatively moving frames (Galilean boost) must be dealt with separately. In the Minkowski case the corresponding Lorentz boost is simply a "rotation" but with an imaginary angle, which one understands when one considers the changes of x_0 and, say, x_1 then the 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 in the (')-frame (moving with relativ velocity v the point $x_1 = 0$, then $\frac{{x_1}'}{ct'} = \frac{v}{c} = \tanh\psi$, which yields the well known formula for the Lorentz boost.

[7] The sign of $\mathrm{d}s^2$ is convention. $\mathrm{d}s^2 < 0$ means "space like " distance, $\mathrm{d}s^2 > 0$ "time like " distance.

with $1/c$, i.e. on each trajectory we select somewhere a zero mark and go from there with the length element of the i-th particle

$$\mathrm{d}\tau_i = \frac{\mathrm{d}s_i}{c} = \frac{1}{c}\sqrt{\left(\frac{\mathrm{d}x_i^0}{\mathrm{d}\tau_i}\right)^2 - \left(\frac{\mathrm{d}\boldsymbol{x}_i}{\mathrm{d}\tau_i}\right)^2}\,\mathrm{d}\tau_i\,.$$

Thus

$$\dot{x}_i^2 = g_{\mu\nu}\dot{x}_i^\mu \dot{x}_i^\nu = c^2, \quad g_{\mu\nu} = \begin{pmatrix} 1 & 0 \\ 0 & -E_3 \end{pmatrix}, \tag{1.38}$$

where we use the Einstein convention of summation and sum automatically over those indices which appear more than once. The "$\dot{}$" means derivative with respect to Minkowski length τ_i, the "proper time" of the i-th particle. (In the frame where the particle is at rest $x_i^0 = c\tau_i$). The metric tensor $g_{\mu\nu}$ which defines here the Minkowski scalar product can be used to lower (and raise with the inverse of the metric tensor) indices:

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

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 $\boldsymbol{v}$. So we have a new function

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

for which we get by chain rule

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

or taking the Minkowski square

$$1 - \frac{\boldsymbol{v}^2}{c^2} = c^2\left(\frac{\mathrm{d}\tau}{\mathrm{d}x^0}\right)^2, \tag{1.39}$$

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 space time path from x to y as the path with the shortest Minkowski length. That is the variation

$$\delta \int_x^y \mathrm{d}s = \delta \int_{\tau(x)}^{\tau(y)} (\dot{x}^\mu \dot{x}_\mu)^{\frac{1}{2}}\,\mathrm{d}\tau = 0.$$

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 come to Newtonian mechanics as 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 $S = mc\int ds$ the canonical momentum is

$$p^\mu = mc\frac{\dot{x}^\mu}{\sqrt{\dot{x}^2}} = m\dot{x}^\mu.$$

Its Minkowski length is by (1.38)

$$g_{\mu\nu}p^\mu p^\nu = p^\mu p_\mu = m^2\dot{x}^\mu\dot{x}_\mu = m^2c^2. \tag{1.40}$$

Parameterizing with $x^0 = ct$ we have

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

Taking $|\boldsymbol{v}| \ll$ c (i.e. $x^0 \approx c\tau$) we get the "Newtonian limit", from which we recognize m as "rest mass", because

$$\boldsymbol{p} = \frac{m}{\sqrt{1-\frac{v^2}{c^2}}}\boldsymbol{v} = \tilde{m}\boldsymbol{v}$$

becomes $\boldsymbol{p} \approx m\boldsymbol{v}$. Furthermore

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

so that we put $E = p^0c$. With(1.40) we obtain thus the so called energy momentum relation:

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

Now we have for N particles the space-time-trajectories $q_i = (q_i^\mu(\tau_i))$, $\mu = (0, 1, 2, 3)$, $i = 1, \ldots, N$. Let us introduce a "force" K^μ which accelerates the particles. By virtue of (1.38) we have

$$\ddot{x^\mu}\dot{x_\mu} = 0$$

and that suggests

$$K^\mu\dot{x_\mu} = 0.$$

Thus force must be orthogonal (in the sense of Minkowski metric) to velocity. The simplest way to realize that is to put

$$K^\mu = F^{\mu\nu}\dot{x_\nu}$$

with $F^{\mu\nu} = -F^{\nu\mu}$, an anti symmetric tensor of rank 2 (an anti symmetric 4×4-matrix). A way to generate such a tensor is with the help of a "four potential" A^μ :

$$F^{\mu\nu} = \frac{\partial}{\partial x_\mu} A^\nu - \frac{\partial}{\partial x_\nu} A^\mu. \tag{1.41}$$

The Maxwell-Lorentz-theory of electro magnetic interaction has the force act on the particles where the law involves not only masses as parameters, but also "charges" e_i:

$$m_i \ddot{q}_i^\mu = \frac{e_i}{c} F^\mu_{\ \nu}(q_i) \dot{q}_i^\nu \tag{1.42}$$

where in view of (1.37) one names the matrix elements as follows

$$F^\mu_{\ \nu}(x) = \begin{pmatrix} 0 & E_1 & E_2 & E_3 \\ E_1 & 0 & B_3 & -B_2 \\ E_2 & -B_3 & 0 & B_1 \\ E_3 & B_2 & -B_1 & 0 \end{pmatrix}.$$

(Recall that indices are lowered or raised by action of $g^{\mu\nu} = g_{\mu\nu}$, i.e. $F^\mu_{\ \nu} = F^{\mu\lambda} g_{\lambda\nu}$.)

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

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

where " $\dot{}$ " still refers to the derivative with respect to τ_i, but for small velocities (compared to velocity of light) this is close to (1.37).

But the fields – the $F^\mu_{\ \nu}$ – are themselves generated by the particles, which is supposed to give the interaction between the charges.

The equation which describes the generation of fields is not difficult to guess in analogy to the gravitational potential being 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 sense of Minkowski suggests the four dimensional potential equation (invariant under "Minkowskian congruences")

$$\left(\left(\frac{\partial}{\partial x^0} \right)^2 - \left(\frac{\partial}{\partial \boldsymbol{x}} \right)^2 \right) A^\mu = \Box A^\mu = \frac{4\pi}{c} j^\mu. \tag{1.43}$$

j^μ is the current originating from the moving particles with charges e_i.We discuss that current below.

Note aside, that (1.41) does not determine the vector potential uniquely, a term of the form $\frac{\partial}{\partial x^\mu} f$ with a nice f can always be added to A^μ without changing the forces. This is called gauge invariance. (1.43) determines the potential in the so called Lorentz gauge

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

The current is a distribution, because the charges are concentrated on 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 (**??**)): A ball would not remain a ball when moving[8].

The current of a point charge is by itself unproblematic. It has the following (frame independent, i.e. relativistic) form (since we used boldface notation for three dimensional vectors we use x for the four vector $x = (x^0, \boldsymbol{x})$) :

$$j^\mu(x) = \sum_i e_i c \int_{-\infty}^{\infty} \delta^4(x - q_i)\dot{q}_i^\mu \mathrm{d}\tau_i \tag{1.44}$$

with

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

Better known is the form in a frame, which we can get by integration. With $x^0 = ct$ we obtain (in the second equality below we write the integral as "line-integral" along the trajectory curve)

$$\begin{aligned} j^\mu(x) &= \sum_i e_i c \int_{-\infty}^{\infty} \delta^4(x - q_i(\tau_i))\dot{q}_i^\mu(\tau_i)\mathrm{d}\tau_i \\ &= \sum_i e_i c \int\limits_{\substack{\gamma_i = q_i(\tau_i) \\ \tau_i \in (-\infty,\infty)}} \delta^4(x - q_i)\mathrm{d}q_i^\mu \\ &= \sum_i e_i c \int_{-\infty}^{\infty} \delta(\boldsymbol{x} - \boldsymbol{q}_i(t_i))\delta(ct - ct_i)\frac{\mathrm{d}q_i^\mu}{\mathrm{d}t_i}(t_i)\mathrm{d}t_i \\ &= \sum_i e_i \delta(\boldsymbol{x} - \boldsymbol{q}_i(t))\frac{\mathrm{d}q_i^\mu}{\mathrm{d}t}(t). \end{aligned}$$

From (1.44) we get easily the continuity equation

$$\frac{\partial}{\partial x^\mu} j^\mu = 0,$$

[8] Suppose one would not care about relativistic invariance, and one takes 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), which however 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, hence it is effectively part of the electron mass. [6].

using that the trajectories are time like and with the same step as above

$$\begin{aligned}\frac{\partial}{\partial x^\mu} j^\mu &= \sum_i e_i c \int_{-\infty}^{\infty} \frac{\partial}{\partial x^\mu} \delta^4(x - q_i) \frac{\mathrm{d}q_i^\mu}{\mathrm{d}\tau_i} \mathrm{d}\tau_i \\ &= \sum_i e_i c \int\limits_{\substack{\gamma_i = q_i(\tau_i) \\ \tau_i \in (-\infty, \infty)}} \left(-\frac{\partial}{\partial q^\mu}\right) \delta^4(x - q_i) \mathrm{d}q_i^\mu \\ &= \sum_i e_i c \left(\lim_{\tau_i \to \infty} \delta^4(x - q_i(\tau_i)) - \lim_{\tau_i \to -\infty} \delta^4(x - q_i(\tau_i)) \right) \\ &= 0 \quad \forall x \in \mathbb{R}^4.\end{aligned}$$

The system (1.41)-(1.43) of differential equations defines the theory of charges interacting via fields.

But: (1.42) is unproblematic if the field $F^\mu_{\ \nu}$ is *given* as a nice function. The linear partial differential equation (1.43) is unproblematic if j^μ is *given*, even as distribution, like in (1.44). Then one has a "Cauchy-problem" to solve, i.e. to solve (1.43) we need 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 for finding a solution.

But now we have to solve (1.42)-(1.43) together, not separately, and that does now work. The system of differential equations are only formal expression, i.e. there are no functions $q^\mu(\tau),\ A^\mu(x)$, whose derivatives would fulfill the equations. It even does not matter whether we have more than one particle. Let us take one particle to see what goes wrong. First solve

$$A^\mu(x) = \left(\Box^{-1} \frac{4\pi}{c} j^\mu \right)(x) = \int \Box^{-1}_{x,x'} \frac{4\pi}{c} j^\mu(x') \mathrm{d}^4 x' \tag{1.45}$$

with a Greens function $\Box^{-1}_{x,x'}$ given by

$$\Box \Box^{-1}_{x,x'} = \delta^4(x - x') \tag{1.46}$$

The Greens function is not unique, they all differ by solutions of the homogenous equation ($j = 0$), i.e. by functions in the kernel of $\Box$. A symmetric choice is[9]

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

Why is that symmetric? Using

$$\delta(f(x)) = \sum_k \frac{1}{|f'(x_k)|} \delta(x - x_k) \tag{1.47}$$

[9] It is quite natural that the Greens function is like this. It is the most natural relativistic function one can write down. The points which have Minkowski distance zero from each other form the light cones (backward and forward) and they are kind of special. So the function is not eccentric in any way.

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

$$\begin{aligned}\Box^{-1}_{x,y} = \delta\left((x-y)^2\right) &= \delta\left((x^0-y^0)^2 - (\boldsymbol{x}-\boldsymbol{y})^2\right) \\ &= \frac{1}{2}\frac{\delta((x^0-y^0)-|\boldsymbol{x}-\boldsymbol{y}|)}{|\boldsymbol{x}-\boldsymbol{y}|} + \frac{1}{2}\frac{\delta((x^0-y^0)+|\boldsymbol{x}-\boldsymbol{y}|)}{|\boldsymbol{x}-\boldsymbol{y}|},\end{aligned}$$

which is the sum of retarded and advanced Greens function, a notion which becomes clear in a minute. Any linear combination of these parts is a possible $\Box^{-1}_{x,y}$, and one commonly uses only the retarded part to have only to deal with radiation "emitted in the past". We say something more to that later. Convince yourself (by formal manipulations) that[10] (1.46) holds for this or any other linear combination.

Now let us come to the end of the story. With (1.45) and (1.44) we get

$$A^\mu(x) = e\int \delta\left((x-q(\tau))^2\right)\dot{q}^\mu(\tau)\mathrm{d}\tau \tag{1.48}$$

and with(1.47) 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

$$(x-q(\tau))^2 = 0 \Leftrightarrow ct - c\tau = \mp|\boldsymbol{x}-\boldsymbol{q}(\tau)|.$$

We see here the names: The retarded/advanced time is given by the intersection point of forward/backward light cone based at $\boldsymbol{x}$ with the trajectory of the particle (see figure 1.5, where you should think of $\boldsymbol{x}$ as being an arbitrary point).

We have thus derived a field (see for instance [2] if some manipulations are unclear) A^μ which is everywhere nice *except* on points x which lie on the world line q of the charge which is generating the field. For $x = q(\tau)$, then $\tau_a = \tau_r = \tau$) and the denominator is zero. But this is now the end of the story since these are the x–values which are needed in (1.42).

This is well known under the name of "self interaction" of the electron. 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 to manage interaction between point charges does not work.

[10] e.g.:$(\frac{\partial}{\partial x^0})^2 - \Delta)(\frac{1}{|\boldsymbol{x}|}\delta(x^0-|\boldsymbol{x}|) = \frac{1}{|\boldsymbol{x}|}\delta'' - \Delta(\frac{1}{|\boldsymbol{x}|})\delta(x^0-|\boldsymbol{x}|) - \frac{1}{|\boldsymbol{x}|}\Delta\delta(x^0-|\boldsymbol{x}|) - 2\nabla(\frac{1}{|\boldsymbol{x}|})\cdot\nabla\delta(x^0-|\boldsymbol{x}|) = \delta(\boldsymbol{x})\delta(x^0-|\boldsymbol{x}|)$, where on needs to use the chain rule on $\nabla\delta(\cdot)$.

For reasons outside of Maxwell's theory of electromagnetism (and may be for the reason that the theory which we describe next is the true electromagnetic theory) the theory works well (in the sense of describing physical phenomena correctly) when the fields are generated done by smeared out charges (charge clouds) to describe the radiation of an antenna or when the fields are "external" given fields which act on charges by the Lorentz force equation. In short, electromagnetism is fine for most of non academic life.

1.3.1 No fields, only particles: Electromagnetism

What is bad about fields to describe interaction? That the field is naturally acting everywhere and thus also on the particle itself which generates the field. But taking the idea of relativistic interaction between particles seriously then why don't the particles interact directly relativistically without introducing fields? In a sense one does this in solving (1.43) for A^μ and puts that into (1.42). Fokker thought that way [5] and discovered a relativistic particle theory, which was later rediscovered by Wheeler and Feynman [8], to explain retarded radiation. How can the particles interact directly in a relativistic way?

There is no natural simultaneity for a force to act at the same time between particles but we know already: The simplest choice is to take Minkowski-distance and to say that particles interact when they have distance zero to each other. Hence the particle at space-time point q interacts with all other particles at space-time points which are intersection points of light cones based at q with the other trajectories, that is when

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

i.e. 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 and one directed towards the past (figure 1.5) (Of course the physical law does not care about such notions like past and future.)

It is rather clear that dynamics which are defined by future times and past times cannot anymore be given by differential equations of the ordinary kind. Nevertheless some differential equations can 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:

$$S = \sum_i \left[-m_i c \int ds_i - \sum_{j>i} \frac{e_i e_j}{c} \int\int \delta\left((q_i - q_j)^2\right) \mathrm{d}q_i^\mu \mathrm{d}q_{j\mu} \right] . \qquad (1.49)$$

Writing the trajectories $q_i^\mu(\lambda_i)$ with arbitrary parameters λ_i and denoting $\dot{q}_i^\mu = \mathrm{d}q_i^\mu/\mathrm{d}\lambda_i$ we obtain for S

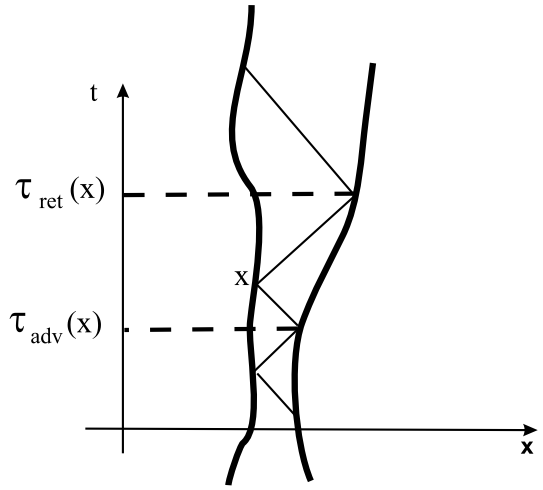

Abb. 1.5. In Feynman-Wheeler-electromagnetism particles interact along backward and forward light cones (here we set $c = 1$).

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

Most noteworthy is that there are no diagonal terms in the double sum. That is the major difference to the Maxwell-Lorentz-theory where morally the diagonal terms are present. The contribution of the i-th particle in the interaction reads

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

with the "field" [11]

$$A_{i\mu}(x) = \sum_{j \neq i} e_j \int \delta\left((x - q_j)^2\right) \dot{q}_{j\mu} \mathrm{d}\lambda_j \, . \tag{1.50}$$

In the Wheeler-Feynman formulation fields (like (1.50)) would only be introduced as a suitable macroscopic description, good for certain situations like handling capacitors.

[11] We computed this in (1.48), but it is important to understand that this is simply a mathematical expression, which plays no role unless x is a point on the word lines of the other particles. There is no field in this theory.

(1.48) shows that both advanced and retarded Greens function appear in A_μ. But it is the "emission of radiation" which we typically see and which is solely described by the retarded Greens function. Wheeler-Feynman as well as Maxwell-Lorentz electromagnetism are time reversible, i.e. the theory does not favor the emission before absorption. The typicality of emission has been coined 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 consider therefore the thermodynamic description of the particle system, i.e. they consider a distribution of charges distributed throughout the universe which "absorb all radiation", which in terms of the theory means that the sum of the differences of 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 like) the time directed radiation phenomena and in particular the observed radiation damping of an accelerated charge is reduced to Boltzmann's explanation of irreversibility, see section **??**.

We emphasize again that the Wheeler-Feynman electro-magnetic theory is a mathematically consistent relativistic theory with interaction, actually the only such mathematically well defined theory existing so far and it is about particles—no fields. The theory is however cumbersome, since it is not of the usual Cauchy-data form, i.e. given by differential equations (ordinary or partial) which determine solutions for given initial values. Why? Because the Euler-Lagrange equations of (1.49) are not ordinary differential equations, since there appear advanced and retarded times in them In comparison Maxwell-Lorentz-theory is formally of the ordinary type, but with the serious drawback that the fields make the equations mathematically undefined.

Anmerkung 1.3.1. On the nature of reality
Reality is a curious notion. Physics takes the view that something "out there" exists, that the world is "made out of something". That's not curious at all. But what it is that the world is made of, that is not easy to say, since the only access we have to the world is by 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, were 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. Often the connection between the entities of the theory and our experience is complicated, often not even spelled out in any detailed way. Nevertheless one has some kind of good 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, which the physical theory is about.

In Maxwell-Lorentz electro-magnetism fields are ontological. Switch on your radio. What better explanation is there than that, that the fields are out there which get absorbed as radio waves by the the radio antenna and the radio transform those into air waves? Music to your ears. But in Wheeler-Feynman electro-magnetism there are no fields and only particles. It explains the music as well. But the explanation is different.

If Maxwell-Lorentz-theory (with point charges) would be mathematically consistent, we could chose 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. Maybe we find in the future a simpler and nicer theory then what we have now which is solely about fields. Then only fields will be "real".

"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 electro-magnetism the electro-magnetic field cannot be measured. Why? 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 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 which one may think about from time to time. Although all variables which are 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, the make up the so called primitive ontology in [1]. They must be there: a particle theory without particle positions is not thinkable. 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 call the particle positions a priori and the other variables a posteriori. Like the electric field. In fact, secondary variables can possibly be replaced by other variables or can even be dispensed with. Like in Wheeler-Feynman electro-magnetism. Another example which we did not touch at all is general relativity which makes the Newtonian force obsolete.

1.4 Secondary Remarks

Anmerkung 1.4.1. On the symplectic structure of the phase space
With the understanding of tensors and forms, in particular differential forms, not only mathematics advanced but also our insight in physics advanced. It was better appreciated what the objects "really" are. We have an example in our relativistic description of electro-magnetism (1.3): The electric and

magnetic field strengths are not vector fields as one learns in school but rather coordinates of an antisymmetric second rank tensor. Mathematical abstraction helps one to get down to the basis of things and that is nice. One such mathematical abstraction is symplectic geometry. We say a few things here mainly to make sure, that no important mysteries are uncovered by further mathematical abstraction which we should take note of.

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 [7]. Symplectic geometry needs space of even dimensions. Classical physics provides that: Consider the phase space $\mathbb{R}^{2n}$ with coordinates $(q_1, \ldots, q_n, p_1, \ldots, p_n)$. Given $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^{2n}$ let $q_i(\boldsymbol{x})$ be the projection of $\boldsymbol{x}$ on the q_i-th coordinate axis. Then

$$\omega_i^2(\boldsymbol{x}, \boldsymbol{y}) = q_i(\boldsymbol{y})p_i(\boldsymbol{x}) - q_i(\boldsymbol{x})p_i(\boldsymbol{y}) \tag{1.51}$$

is the area of the parallelogram projected into the (q_i, p_i)-plane, generated by $\boldsymbol{x}, \boldsymbol{y}$. Put

$$\omega^2(\boldsymbol{x}, \boldsymbol{y}) = \sum_{i=1}^{n} w_i^2(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{x} \cdot (-I)\boldsymbol{y} = ((-I)\boldsymbol{y})^t \boldsymbol{x} \tag{1.52}$$

with

$$I = \begin{pmatrix} 0_n & +\mathrm{E}_n \\ -\mathrm{E}_n & 0_n \end{pmatrix}, \qquad \begin{array}{r} \mathrm{E}_n = n - \text{dimensional unit matrix} \\ 0_n = n - \text{dimensional zero matrix,} \end{array}$$

and $\boldsymbol{z}^t = \boldsymbol{z}$ transposed = row vector = element in the dual space of $\mathbb{R}^{2n}$. The 2-Form ω^2 respectively 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 $\frac{\partial f}{\partial \boldsymbol{x}}$ of a function $f(\boldsymbol{x})$, $\boldsymbol{x} \in \mathbb{R}^{\mathrm{d}}$ is in fact a "dual element", i.e. it is the row vector which acts as a linear map on $\boldsymbol{h}$ as $\frac{\partial f}{\partial \boldsymbol{x}}\boldsymbol{h}$ = row times column (matrix multiplication with a vector). Now use ω^2 to identify $\frac{\partial f}{\partial \boldsymbol{x}}$ with a vector $\nabla_\omega f$ (using the Euclidean scalar product this is the normal ∇f): given $\boldsymbol{z} \in \mathbb{R}^{2n}$ $\omega^2(\cdot, \boldsymbol{z})$ is a linear form, i.e. a linear map

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

which we wish to equal $\frac{\partial f}{\partial \boldsymbol{x}}$, hence search for a $\boldsymbol{z}_f$, so that

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

Thus (1.6)can be written as

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

The Hamiltonian flow respects the symplectic geometry, in particular it is area preserving, which means the following: Let C be a closed curve in $\mathbb{R}^{2n}$, define the "enclosed area" as sum of the n areas, which arise from projections of C on the coordinate planes (q_i, p_i) (cf. (1.51) and (1.52)). In the (q_i, p_i)-plane we have a curve C_i with the area

$$\Phi(q_i, p_i) = \begin{pmatrix} q_i \\ p_i \end{pmatrix}, \qquad q_i, p_i \in \text{area}(C_i) = A(C_i).$$

The area can be transformed by Stokes' theorem in two dimensions into a line-integral

$$\begin{aligned} \int\limits_{A(C_i)} \mathrm{d}q_i \mathrm{d}p_i &= \int\limits_{A(C_i)} \text{rot} \begin{pmatrix} 0 \\ q_i \end{pmatrix} \mathrm{d}q_i \mathrm{d}p_i \\ &= \oint\limits_{C_i} \begin{pmatrix} 0 \\ q_i \end{pmatrix} \cdot \begin{pmatrix} \mathrm{d}q_i \\ \mathrm{d}p_i \end{pmatrix} \\ &= \oint\limits_{C_i} q_i \mathrm{d}p_i \,. \end{aligned} \tag{1.53}$$

In general:

$$A = \text{area of } C = \oint\limits_{C} \boldsymbol{q} \cdot \mathrm{d}\boldsymbol{p} = \sum_i \oint\limits_{C_i} q_i \mathrm{d}p_i.$$

Using differential forms

$$\omega_i^2 = \mathrm{d}p_i \wedge \mathrm{d}q_i \text{ and } \omega^2 = \sum_i \mathrm{d}p_i \wedge \mathrm{d}q_i \text{ and } \omega^2 = \mathrm{d}\omega^1 \text{ with } \omega^1 = \sum_i p_i \mathrm{d}q_i.$$

(1.53) is nothing but

$$\int\limits_{A(C_i)} \mathrm{d}\omega^1 = \int\limits_{C_i} \omega^1.$$

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

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}A(t) &= \frac{\mathrm{d}}{\mathrm{d}t}\oint\limits_C \boldsymbol{q}(t)\,\mathrm{d}\boldsymbol{p}(t) = \oint\limits_C \dot{\boldsymbol{q}}\,\mathrm{d}\boldsymbol{p} + \oint\limits_C \boldsymbol{q}\,\mathrm{d}\dot{\boldsymbol{p}} \\ &= \oint\limits_C \dot{\boldsymbol{q}}\,\mathrm{d}\boldsymbol{p} - \oint\limits_C \dot{\boldsymbol{p}}\,\mathrm{d}\boldsymbol{q} \quad \left(+\oint\limits_C d(\boldsymbol{q}\dot{\boldsymbol{p}}) = 0\right) \\ &= \oint\limits_C \frac{\partial H}{\partial \boldsymbol{q}}\,\mathrm{d}\boldsymbol{q} + \oint\limits_C \frac{\partial H}{\partial \boldsymbol{p}}\,\mathrm{d}\boldsymbol{p} = \oint\limits_C dH \\ &= 0.\end{aligned}$$

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 (1.51) yield the volume form: $\omega = \mathrm{d}p_1 \wedge \mathrm{d}q_1 \wedge \ldots \wedge \mathrm{d}p_n \wedge \mathrm{d}q_n$, the Lebesgue measure on phase space (in form language) and thus Liouville's theorem arises from preservation of area.

Transformations of coordinates $(\boldsymbol{q}, \boldsymbol{p}) \xrightarrow{\psi} (\boldsymbol{Q}, \boldsymbol{P})$ are called canonical or symplectic, if the Jacobi matrix $\nabla\psi$ is symplectic, which is a variation of "orthogonal matrix"), i.e.

$$(\nabla\psi)^t I \nabla\psi = I. \tag{1.54}$$

They preserve areas and they fulfill the canonical equations

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

The Poisson bracket (1.11) is invariant under canonical transformations because (1.11) can be written as

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

and if $f(\boldsymbol{Q}, \boldsymbol{P})$, $g(\boldsymbol{Q}, \boldsymbol{P})$ and $(\boldsymbol{Q}, \boldsymbol{P}) = \psi(\boldsymbol{q}, \boldsymbol{p})$ are given and observing

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

we have that

$$\begin{aligned}\{f \circ \psi, g \circ \psi\} &= \nabla(f \circ \psi) \cdot I \nabla(g \circ \psi) \\ &= (\nabla f \circ \psi) \cdot (\nabla\psi)^t I \nabla\psi (\nabla g \circ \psi) \\ &\overset{(1.54)}{=} (\nabla f \circ \psi) \cdot I (\nabla g \circ \psi) \\ &= \{g, f\} \circ \psi.\end{aligned}$$

Clearly

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

and variables which fulfill this are called canonical. Of particular interest are variables $(Q_1, \ldots, Q_n, P_1, \ldots, P_n)$, where $P_1, \ldots, P_n$ do not change with time- and where the Hamilton function achieves the form $\tilde{H}(\boldsymbol{Q}, \boldsymbol{P}) = \sum_{i=1}^{n} \omega_i P_i$. Then $\dot{Q}_i = \omega_i$, i.e. $Q_i = \omega_i t + Q_{i,0}$, and Q_i is like a phase of a harmonic oscillator. Such (P_i, Q_i) are called action-angle-variables. Systems, which allow for such variables are called integrable, since their behaviour in time is in principle completely under control, their motion (in the new coordinates) being that of "uncoupled" harmonic oscillators and thus the solution is found by algebraic manipulation and integration. The Hamiltonian movements 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 $P = H$. Integrability is however *atypical* for Hamiltonian systems.

References

[1] Allori V., T. R. Z. N., Goldstein S.: *On the common structure of bohmian mechanics and the ghirardi-rimini-weber theory*, quant-ph/, 2007.

[2] Barut, A. O.: *Electrodynamics and classical theory of fields & particles*, Dover Publications Inc., New York, 1980, Corrected reprint of the 1964 original.

[3] Born, M.: *Quantenmechanik der stoßvorgänge*, Z. Physik **37** (1926), 863.

[4] Born, M.: *Quantenmechanik der stoßvorgänge*, Z. Physik **38** (1926), 233.

[5] Fokker, A. D., Z. Physik **58** (1929), 386.

[6] Rohrlich, F.: *Classical charged particles*, third ed., World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2007.

[7] Scheck, F.: *Mechanics*, third ed., Springer-Verlag, Berlin, 1999, From Newton's laws to deterministic chaos, Translated from the German.

[8] Wheeler, J. A. and Feynman, R. P.: *Classical electrodynamics in terms of direct inter-particle action*, Rev. Modern Physics **21** (1949), 425–433.