

Chapter 1

Newtonian Dynamics

1.1 Introducing Probability

Dynamics of particles is among the most fundamental problems in physics, and for a long time its complete understanding could have been deduced from three axioms, also known as the three Newton's axioms of motion. These axioms in their original form are the following.

Axiom 1 *If a particle is at rest or moving at a constant speed in a straight line, it will remain so unless a force acts it upon.*

This axiom is important because it says that the basic state of particle is to be at rest or to move uniformly in a straight line. The concept of force is introduced, as the cause that changes this state.

Axiom 2 *The time rate of change of the velocity is directly proportional to the force and inversely proportional to the mass of the particle.*

This axiom is important because it defines a quantitative relationship between the time rate at which that basic state of particle changes and the known quantities: the force and the mass.

The third axiom in its original form is: *The action of two particles upon each other is always equal and directly opposite, i.e. reaction is always equal and opposite of action.* Although those axioms were the basis of *classical mechanics* it was discovered during the course of time that they were not sufficient for an adequate description of all the phenomena in Nature. They needed amendments, and one of them was necessary in the third axiom. In its original form it expressed the symmetry of action and reaction among particles and its importance is that without it the energy conservation law would have been violated. Its fundamental defect was that it implicitly assumed that action among particles travels at infinite speed. For the finite

speed of action the axiom is not correct and in the version that takes this into account its formulation is as follows:

Axiom 3 *If two particles are not in relative motion during sufficiently long time then their action upon each other is always equal and directly opposite i.e. reaction is always equal and opposite of action.*

In Axiom 3 the term *sufficiently long time* is used rather than *infinitely long time*. If the action travels at a finite speed then the stationary interaction between particles can be achieved after finite time and therefore it is not necessary to use the more drastic term *infinite*. The axiom also applies to any reference frame (coordinate system), because the term *relative motion* is emphasized.

Those three axioms of classical mechanics were not sufficient to describe all accumulated evidence about the processes in Nature. At least the following two are missing.

Axiom 4 *From the basic laws of Nature it is not possible to suggest experiment that would determine absolute motion of a coordinate system.*

This axiom is often put together with the axiom about the constancy of the speed of light, and the result is the *theory of relativity*, or the unified theory of space and time (there are many textbooks on the special relativity, but a good historic account and the physical insight is given by [Born (1962)]). The axiom about the velocity of light is not necessary because it is derived from the Axiom 4. The light, being the carrier of one of the fundamental forces in Nature, the electromagnetic force, if it were to change its speed from one coordinate system to another would suggest experiment for determining absolute motion.

Axiom 5 *At any time the product of standard deviations for one coordinate and the component of momentum of particle in this coordinate is always larger, or at most equal, to a universal constant.*

In this axiom the term *momentum* is used which will be defined shortly. The axiom is known as the law that is derived from the principles of *quantum mechanics*, and depending on the author the basic ones are either the wave-particle dualism or the observable-operator correspondence. Besides those principles of quantum theory there are a number of others, from which the *law of uncertainty* is derived. Here this law is suggested as one of the axioms that expands traditional three of classical mechanics.

The consequences of those five axioms will be analyzed, and the book is segmented so that the three original ones are gradually expanded. The

shift of emphases in the traditional dynamics of particles is made in this chapter, without going further than the first three axioms allow. This step is essential when at a later stage additional axioms are introduced, in particular Axiom 5.

Axiom 2 is in the form of equation, which for a long time epitomized physics. In a somewhat expanded form this equation (*Newton equation*) is

$$\frac{d\vec{p}}{dt} = \vec{F}, \quad \frac{d\vec{r}}{dt} = \frac{\vec{p}}{m} \quad (1.1)$$

and in essence it relates the rate at which a particle changes its position to the two quantities that are assumed known: the force \vec{F} on and the mass m of the particle. This relationship is somewhat obscured by introducing an intermediate variable \vec{p} , called *momentum*. It is a vector quantity that is proportional to the velocity, and the factor of proportionality is mass m . The significance of momentum is in its very convenient properties, as it will become evident in further developments, but apart from that there is no deeper meaning to it. The space that is spanned by the coordinate vector \vec{r} and the momentum vector \vec{p} is called *phase space*.

The equations (1.1) define initial value problem, meaning that given initial position and velocity of the particle its past and future movement is entirely determined, of course if the force is known and the mass. Based on this character of equations classical mechanics is associated with the deterministic view of Nature, regardless of the fact that initial conditions are always determined from experiment, and therefore they are never given precisely. This is overlooked when applying classical mechanics in the implicit belief that these uncertainties do not affect very much the prediction for the time movement of the particle. In other words, it is believed that if from the experiment the mean value for the position and velocity is chosen for the solution of the equations then their uncertainty stays the same in the course of time.

Various forces act among particles, but there are only few fundamental ones and these will be of the main interest. The oldest known is gravity¹, but in many respects it has some very peculiar properties about which more will be said later. Electromagnetic force is the one which dominates most of the processes in Nature, and will be of the main interest. Nuclear and weak forces are as important but unfortunately relatively little is known about them, in the context of the classical concepts of the force. Therefore

¹In fact the oldest known is magnetism, but here it is referred to the forces that have been studied scientifically.

the implicit assumption is that the equations (1.1) describe dynamics of charged particles, and hence they should be an integral part of what one would call the unified dynamics of charges and electromagnetic field. In that respect those equations are not complete because the equations for the electromagnetic field are missing. This fact was not realized for a long time, because the theory for the electromagnetic field developed separately from the theory for particles. Furthermore when referring to a particle it was often meant an object that is a conglomerate of a huge number of elementary constituents of matter. The forces that act among such objects can indeed have various forms, but they are nothing but certain averages of the electromagnetic force. Once this is realized then the goal of the theory is more transparent. It is primarily the theory that describes dynamics of the elementary constituents of matter in their mutual interaction that is mediated by the electromagnetic force. In order to achieve this goal two essential steps need to be made. One is to move the emphases from the deterministic character of classical mechanics, and the second is to unify dynamics of particles with dynamics of the electromagnetic field.

Classical mechanics ([Goldstein (1981)]) is analyses of various problems whose solution is obtained from the set of equations (1.1). The problems are distinguished by specifying initial conditions for the position \vec{r} and momentum \vec{p} or the velocity $\vec{v} = d\vec{r}/dt$ of the particle. There is in fact the third arbitrary parameter that needs to be determined, but often it is not mentioned: the initial instant in time. This parameter is not discussed because it is believed not to be very important: dynamics of a closed system is invariant to the choice of the instant from which time is measured, and therefore it can be taken $t = 0$. In conclusion, if the initial values for \vec{r} and \vec{p} are specified at $t = 0$ then the equations (1.1) have unique solution $\vec{r}(t)$, which is called trajectory. This means that no two trajectories that start with different initial conditions at $t = 0$ can have the same coordinate and velocity at the same time later.

Important constants (invariants) of motion are derived from the set of equations (1.1). One of them is obtained by multiplying the first equation with \vec{p} , in which case the following relationship is obtained

$$d_t \left(\frac{1}{2} \vec{p}^2 \right) - m d_t \vec{r} \cdot \vec{F} = 0 \quad (1.2)$$

where the second equation of the set (1.1) was also used (the symbol d_u^n from now on stands for the derivative d^n/dt^n). In general the force \vec{F} can have any functional form, in particular it can also be velocity (momentum)

dependent. The dependence of this kind is characteristic of the electromagnetic force (more specifically of its magnetic component), but at this stage it will not be considered. The forces with only the coordinate dependence, and time dependence when explicitly mentioned, will be taken into account, and for them a theorem can be used. It says that any vector can be decomposed into two components: one component whose divergence is zero (perpendicular component) and the other whose rotor is zero (parallel component). This means that the force can be written as

$$\vec{F} = -\nabla V + \nabla \times \vec{A} \quad (1.3)$$

where the scalar function V is called (scalar) potential, and by analogy \vec{A} is called the vector potential. However in most cases the force is derived from only the scalar potential, but forces that also include the vector potential will also be used, mainly for modelling purpose. The latter are most intimately connected with the electromagnetic force, but in somewhat more complicated form. If the force \vec{F} , without the vector potential, is replaced in (1.2) then the following relationship is obtained

$$d_t \left(\frac{1}{2m} p^2 + V \right) = \partial_t V \quad (1.4)$$

and for potentials V that are not explicitly time dependent it follows that

$$\frac{1}{2m} \vec{p}^2 + V = E \quad (1.5)$$

where E is called the total energy of the system, which is a constant (invariant) of motion. For the force with the vector potential component, of the kind (1.3), the energy conservation law cannot be derived.

Another important invariant is obtained by multiplying the first equation of (1.1) by \vec{r} , as a vector product. In this case

$$\vec{r} \times d_t \vec{p} = d_t (\vec{r} \times \vec{p}) = \vec{r} \times \nabla V$$

which is zero for a spherically symmetric potential. For such potentials the quantity

$$\vec{L} = \vec{r} \times \vec{p}$$

is invariant and it is called the total angular momentum of the particle.

The set of equations (1.1) may have other invariants, but they are specific to a particular form of force. The total energy and the total angular

momentum are invariants that are the best known and the most widely used.

The set of equations (1.1) can be partially solved if the force is velocity and time independent. By scalar-multiplying the second equation with \vec{p} then

$$dt = m \frac{\vec{p} \cdot d\vec{r}}{p^2}$$

where the momentum is a function of \vec{r} because it is calculated from the energy conservation law (1.5). The time is then a line integral

$$t = m \int_{\vec{r}_0}^{\vec{r}} \frac{\vec{p}(\vec{u}) \cdot d\vec{u}}{p^2(\vec{u})} \quad (1.6)$$

where it is assumed that $\vec{r} = \vec{r}_0$ at $t = 0$. This is an implicit equation for trajectory, and in some analysis it is very useful.

Dynamics of particle is obtained as solution of the equations (1.1). Deviation between the measurement and the theoretical prediction is attributed to the failure of measuring precisely the position and the velocity of particle rather than the inadequacy of the equations (1.1). A very simple analysis dispels this attitude.² The question that needs to be answered is how a small uncertainty in the initial conditions affects certainty of the time evolution of a trajectory. In other words, if a small change $\vec{\Delta}_0$ is made in the initial position of the particle and $\dot{\vec{\Delta}}_0$ (by convention the dot over the function will designate time derivative) in its velocity then the trajectory is assumed reliable if this change stays within the bounds of the order Δ_0 . In mathematical terms if $\vec{r}(t)$ is the trajectory that is determined by the initial position \vec{r}_0 and the initial velocity \vec{v}_0 then at any time later the trajectory that starts with slightly different initial conditions is estimated as

$$\vec{r}_{\Delta}(t) = \vec{r}(t) + \vec{\Delta}(t). \quad (1.7)$$

By replacing this in the set of equations (1.1), and expanding the force in the powers of $\vec{\Delta}$, the approximate equation for the vector $\vec{\Delta}(t)$ is

$$m \ddot{\vec{\Delta}}(t) = \vec{\Delta} \cdot \nabla \vec{F}[\vec{r}(t)]. \quad (1.8)$$

This is approximate equation because in the expansion of the force only the first power in $\vec{\Delta}$ was retained. If the derivatives of the force \vec{F} , with respect

²Original analysis was done by Lyapunov in 1892.

to the coordinates, are assumed to be constant in a small time interval then the equation can be transformed into

$$m \ddot{\vec{\delta}} = \vec{\delta} \cdot \vec{\omega}$$

where $\vec{\delta}$ and $\vec{\Delta}$ are related by a linear combination that transforms the right side of the equation (1.8) into a scalar product of two vectors. A general solution of the equation for the component δ_n is

$$\delta_n = a \sinh\left(t\sqrt{\frac{\omega_n}{m}}\right) + b \cosh\left(t\sqrt{\frac{\omega_n}{m}}\right)$$

which is an exponentially increasing function of time, if the component ω_n is positive. In most cases this is true, one exception being harmonic oscillator for which ω_n are negative. The assumption therefore that $\vec{\Delta}(t)$ in (1.7) is small if $\vec{\Delta}_0$ and $\dot{\vec{\Delta}}_0$ are small is not correct, because in general the separation of the two trajectories increases exponentially in time. For the harmonic oscillator the separation stays within the bounds of the initial one.

Rapid increase in separation of two, initially very close, trajectories indicate that the same impact would be on the other parameters of the system, e.g. on the total energy or the angular momentum, the two most important invariants. Although qualitative arguments could be given that this is not the case, it is necessary to show this more precisely. The total energy can be calculated for two initially close trajectories, and the difference between them is

$$\delta E = \frac{\left(\vec{p} + m \dot{\vec{\Delta}}\right)^2}{2m} + V(\vec{r} + \vec{\Delta}) - \frac{p^2}{2m} - V(\vec{r}) \approx \vec{p} \cdot \dot{\vec{\Delta}} + (\vec{\Delta} \cdot \nabla)V(\vec{r})$$

where only the first correction in Δ was retained. The time rate of this difference is

$$\delta \dot{E} = \dot{\vec{p}} \cdot \dot{\vec{\Delta}} + \vec{p} \cdot \ddot{\vec{\Delta}} + \dot{\vec{\Delta}} \cdot \nabla V + \vec{\Delta} \cdot \nabla \dot{\vec{r}} \cdot \nabla V = 0 \quad (1.9)$$

where in the last step the equations (1.8) and (1.1) were used. The result indicates that at least in the first order of Δ the initial small change in the total energy remains small at any later time. Therefore the total energy of the particle is not very affected by the separation of two trajectories.

The same reasoning could be applied to the angular momentum. The time rate of the difference between the angular momenta for the two tra-

jectories is

$$\delta \vec{L} \approx \dot{\vec{r}} \times \dot{\vec{\Delta}} + \vec{r} \times \ddot{\vec{\Delta}} + \dot{\vec{\Delta}} \times \vec{p} + \vec{\Delta} \times \ddot{\vec{p}} = -\vec{\Delta} \cdot \nabla \vec{r} \times \nabla V$$

where again the equations (1.8) and (1.1) were used. The rate is equal to zero for a spherically symmetric potential, in the case when the angular momentum is conserved. Otherwise the two angular momenta will follow the same pattern of separation as the trajectories.

Example 1 and Example 2 show that predictions based on slightly inaccurate initial conditions are in most cases meaningless. There are two aspects to this conclusion. One is practical and the other theoretical. It can be argued that while the practical determination of the initial conditions is a matter of technological advances, the theoretical part has been solved, only one needs to have powerful enough computing devices to solve the equations. Neither argument is correct in the light of what was shown in the previous examples. Even the most imaginative experimental devices would never achieve the absolute accuracy. However, given the possibility that such do exist the prediction requires solving the dynamics equations numerically, because only in a very few cases it is possible to solve them in a closed form. Numerical solution requires accurate numerical code, and high precision computing devices. Both requirements can be fulfilled to a finite precision, and hence the errors inevitable come in the form of the numerical inaccuracies. As it was shown, these inaccuracies propagate with the exponential increase in time. Therefore it can be safely argued that the following question is academic: what is position and velocity of particle at certain time if its initial position and velocity are known? There is no way this question can be answered, only approximately for a short time interval. There is no universal criterion what this short time interval might be. It must be defined for a particular example, and it must be based on the analysis of the solution of equation (1.8).

In the light of what has been shown the conceptual approach to dynamics of particles must be changed in its basis. Instead of the question of certainty another one must be asked which takes into account inaccuracies in the knowledge of the time evolution of trajectories. And uncertain knowledge of the time evolution of trajectories inevitably introduces the concept of probability into dynamics. If this is the case then the first thing that must be changed is the attitude towards the initial conditions. Instead of requiring that they are known with certainty one asks what is probability with which they are determined, i.e. one defines a function of

these parameters which gives probability to measure them. The function is written as $\rho_0(\vec{r}, \vec{p})$ and will be called the phase space probability density, because probability to measure position \vec{r} and momentum \vec{p} of the particle in a small volume element $\delta^3 r \delta^3 p$ around these values is $\rho_0(\vec{r}, \vec{p}) \delta^3 r \delta^3 p$. How is the phase space probability density obtained? Any experiment that seeks to determine initial conditions is repeated many times, and as the rule the more it is repeated the better results are obtained. Out of these experiments the data are produced that are randomly distributed around some mean value in the phase space, and after N such measurements there will be n outcomes in some volume element in the phase space. The ratio n/N then gives the probability of finding in the next measurement the initial position and velocity that will be in the same volume element. It is convenient to normalize this probability by the volume of this phase space element, and this is the required phase space probability density $\rho_0(\vec{r}, \vec{p})$. For each measurement the particle is followed until time t , when again its position and velocity is measured and by the same procedure another phase space probability density is obtained, which is written as $\rho(\vec{r}, \vec{p}, t)$ to emphasize its time dependence. The former can be referred to as the initial phase space probability density, and the latter is the phase space probability density at time t . The two are related, and the principal question to be answered is: how is this done theoretically? Or to put it differently: given the initial phase space density (the epithet “probability” will be omitted in further reference to this function for the reason that will be explained later) $\rho_0(\vec{r}, \vec{p})$ how this function evolves in time if each trajectory satisfies the set of equations (1.1)? This question is of much greater relevance for the analysis of dynamics of a particle than the one that is asked for a single trajectory.

By shifting the emphases from trajectory to phase space density the meaning of a single experiment should be clarified. Single measurement of initial conditions produces certain outcome, but this certainty is after the measurement was done. The theory predicts, and therefore it cannot say a priori what the outcome of measurement will be. It can only give the probability that certain initial conditions will be measured and from that the probability of finding particle at certain location, and with certain velocity, after time t . However, with this shift of emphases the meaning of experiment is also changed because if its aim is to reproduce theoretical prediction then it cannot be based on a single event. The experiment must be repeated, in which case, if the theory is correct the phase space density would be reproduced. The parallel can be drawn with another,

much simpler experiment: tossing a coin. The theory predicts that given no preconditions the outcome of getting heads is one half, however, once the coin is thrown into the air the outcome is certainty. If the outcome is tails the theory is not invalidated, because its prediction is probability not certainty of a single event. The experiment must be repeated many times in order to ensure good statistical sample, and then the result should be compared with what the theory predicts. If the outcome is different from what the theory predicts then the answer may be in the initial phase space density that is biased towards certain initial conditions, and the theory should answer the question why is this the case.

It was demonstrated that there is a need to shift the emphases from the concept of trajectory to the concept of probability, but the question is; what are the relevant dynamics equations and how to solve them? To answer the first part of the question a very important property of trajectories will be utilized: no two trajectories in the phase space can cross at the same instant in time. This property results from the fact that any trajectory is uniquely specified by its initial conditions. If two trajectories would cross (in the phase space) at the same instant in time then with these initial conditions two trajectories are defined, and this is not possible. The result is that a boundary of any volume element in the phase space stay connected for all time, because no trajectory from inside can cross outside and vice versa. This also means that the number of trajectories within a phase space volume element stays constant, and also the volume of this element. In short, the phase space density, along a trajectory, stays constant. This is the property from which the dynamics equation for the phase space density will be derived.

The phase space density is a function of the form $\rho(\vec{r}, \vec{p}, t)$ and from the previous property it follows that its change in time, along a trajectory, is zero. In mathematical terms this property is

$$d\rho(\vec{r}, \vec{p}, t) = (d\vec{r} \cdot \nabla_r)\rho + (d\vec{p} \cdot \nabla_p)\rho + dt\partial_t\rho = 0 \quad (1.10)$$

where ∂_s means partial derivative with respect to the variable s and the index of ∇ designates the variable with respect to which the derivatives are taken. This change must be zero along a trajectory, which means that the increments $d\vec{r}$ and $d\vec{p}$ per unit time must be given by the equations of motion (1.1), hence

$$\partial_t\rho + \frac{1}{m}\vec{p} \cdot \nabla_r\rho + \vec{F} \cdot \nabla_p\rho = 0. \quad (1.11)$$

The equation is also known as the Liouville equation, although its original derivation was motivated by the study of large number of non-interacting particles (there are relatively few studies based on Liouville equation, and one is in galactic dynamics as described in [Binney (1987)]). In here, it should be stressed, it represents an equation for the time evolution of the probability density in the phase space, for a single particle. There are a number of useful properties of this equation. If the equation is integrated with respect to the momentum variables, and \vec{F} is not a velocity dependent function then

$$\partial_t P(\vec{r}, t) + \nabla_r \cdot \vec{J}(\vec{r}, t) = 0 \quad (1.12)$$

which is called the continuity equation. In this equation two very important quantities are defined: the probability density $P(\vec{r}, t)$ for coordinates and the probability current $\vec{J}(\vec{r}, t)$, which are defined as

$$P(\vec{r}, t) = \int d^3 p \rho(\vec{r}, \vec{p}, t), \quad \vec{J}(\vec{r}, t) = \int d^3 p \frac{\vec{p}}{m} \rho(\vec{r}, \vec{p}, t). \quad (1.13)$$

The continuity equation is important because it ensures that the probability density $P(\vec{r}, t)$ is conserved, i.e. the probability of finding particle anywhere in the space is always one, or mathematically the volume integral over $P(\vec{r}, t)$ is always one. To ensure that $P(\vec{r}, t)$ is conserved means that trajectories cannot disappear or be created without any cause.

Another type of continuity equation is obtained by integrating (1.11) in the coordinate variables, when one obtains

$$\partial_t Q(\vec{p}, t) + \nabla_p \cdot \vec{K}(\vec{p}, t) = 0 \quad (1.14)$$

where now

$$Q(\vec{p}, t) = \int d^3 r \rho(\vec{r}, \vec{p}, t), \quad \vec{K}(\vec{p}, t) = \int d^3 r \vec{F}(\vec{r}, t) \rho(\vec{r}, \vec{p}, t)$$

and expresses the conservation of the probability density $Q(\vec{p}, t)$.

By shifting the emphasis from the concept of trajectory to the concept of probability the quantities that one works with are averages. For example, the momentum of particle no longer has the meaning as for a single trajectory; instead its average is defined as

$$\vec{p}_{\text{ave}}(t) = \int d^3 r \vec{p}(\vec{r}, t) = \int d^3 r d^3 p \vec{p} \rho(\vec{r}, \vec{p}, t).$$

The time derivative of this average momentum is

$$d_t \vec{p}_{\text{ave}}(t) = \int d^3 r \partial_t \vec{p}(\vec{r}, t) = \int d^3 r \vec{F}(\vec{r}, t) P(\vec{r}, t) \quad (1.15)$$

which is generalized equation of motion, where the force is replaced by its average.

One of the fundamental quantities in dynamics of a particle is its total energy. By introducing the concept of probability it is defined as the average

$$E_{\text{ave}} = \int d^3 r d^3 p \left[\frac{p^2}{2m} + V(r) \right] \rho(\vec{r}, \vec{p}, t) \quad (1.16)$$

which is also a conserved quantity. To show this its time derivative is calculated

$$d_t E_{\text{ave}} = \int d^3 p \frac{p^2}{2m} \partial_t Q(\vec{p}, t) + \int d^3 r V(r) \partial_t P(\vec{r}, t)$$

and if the two continuity equations are used, and after partial integration, then

$$d_t E_{\text{ave}} = \int d^3 p \frac{\vec{p}}{m} \cdot \vec{K} + \int d^3 r \nabla V \cdot \vec{J} = 0 \quad (1.17)$$

which proves that indeed the total energy (1.16) is conserved. This is a very important result because it shows that the shift of the emphases is consistent with one of the basic laws of nature: the total energy conservation law. However, the law is not defined for a trajectory but for the average quantity in the phase space.

Another fundamental quantity is the angular momentum, which is defined as

$$\vec{L}_{\text{ave}} = \int d^3 r d^3 p \vec{r} \times \vec{p} \rho(\vec{r}, \vec{p}, t) \quad (1.18)$$

and it is also invariant of motion. By taking time derivative

$$d_t \vec{L}_{\text{ave}} = \int d^3 r d^3 p \vec{r} \times \vec{p} \partial_t \rho(\vec{r}, \vec{p}, t)$$

and after using the Liouville equation

$$d_t \vec{L}_{\text{ave}} = \int d^3 r d^3 p \vec{r} \times \vec{F}(\vec{r}, t) \rho(\vec{r}, \vec{p}, t) \quad (1.19)$$

which is zero only for the central force (the force which is obtained from the spherically symmetric potential). This result also expresses, but for the

average quantity, another fundamental law: conservation of the angular momentum.

In this way it was shown that the fundamental laws are not violated by shifting the emphases in the dynamics of particle from a trajectory to a probability concept. The only difference is that the laws do not apply to the single trajectory but to the averages in the phase space. Therefore, from now the subscript that indicates the average quantities, such as in (1.15) and (1.16), will be omitted and they will be called without the adjective “average”.

Working with the newly defined quantities does not have meaning if another important concept is not introduced, the mean deviation (error) from the average or the *standard deviation*. For a quantity A it is defined as

$$\Delta A = \sqrt{(A^2)_{\text{ave}} - (A_{\text{ave}})^2} \quad (1.20)$$

where

$$A_{\text{ave}} = \int d^3r d^3p A(\vec{r}, \vec{p}) \rho(\vec{r}, \vec{p}, t).$$

The concept of error will be essential in the later developments.

Those were foundations on the basis of which dynamics of a particle can be built upon. This should be done consistently, and expect that sometime the results may not have an adequate counterpart in the concept of trajectory. The whole way of thinking must be adapted to this approach. Instead of emphasizing trajectory as the basis for understanding Nature, it is the concept of probability that takes the central role. This also means that the deep rooted determinism, which was built around the notion that “given initial positions and velocities in the Universe its future can be predicted”, must be replaced by uncertainty, even on the much more modest scale.

1.2 Solving Dynamics Problems

Solving dynamics problems, based on equation (1.11) consists of three separate steps. In the first the initial phase space density is chosen. In the second the phase space density is propagated in time, and finally the phase space is analyzed. In the traditional analysis for single trajectory that separation was unnecessary because the initial conditions could be singled out immediately, analysis of the final outcome very simple, only the propagation in time required more effort.

1.2.1 *Initial conditions*

Perhaps the most difficult task, and the most important one, is determining initial conditions for the phase space density. The difficulty will become even greater when electromagnetic field is introduced, but this problem will be encountered later. Why is determining the initial phase space density so difficult and important? The answer to the latter is simple: it determines the physics of the problem. Depending on the choice of the initial conditions the answer will be either meaningful or meaningless. Why is so difficult? The initial phase space density reflects all effects that contribute to the uncertainty in determining the initial conditions for a particle. They can be numerous and very difficult to collect all, but the simplest way to avoid thinking about them is to do repeated measurements. Each measurement is a point in the phase space, and after N such measurements one could define the density of such points around a coordinate in the phase space, which will be the phase space density. Although quite straightforward repeated measurements are often impossible to perform. For example, position and velocity of Earth can be determined once, but doing it next time can only be made after certain short time interval, after which it had moved to another place. Therefore one can make an educated guess about the initial phase space density from a single measurement by thinking about all possible sources of errors and how they affect it. Another example is the electron moving around a nucleus. Again repeated measurements are impossible (setting aside much deeper reason why is that so) because the electron moves very fast (of the order 10^{16} times around the nucleus in one second) and in a such a tiny space that no experiment would produce a meaningful result. For the electron, therefore, to define the initial phase space density it is not a straightforward task, and its choice should be based on certain criterion. The important one is that if the electron is left unperturbed for a long time (compared to its typical revolution time around the nucleus) the corresponding and meaningful phase space density should be independent of time (time independent or stationary). Whether such a phase space density can be achieved, and what is its form, will be discussed at length in later analysis.

Those two examples described some difficulties in the choice of the initial phase space density, however, when the electromagnetic field is included the problem becomes very serious. At this stage it is not clear why is that so, but in the later development it will be necessary to assume that the phase space density interacts with itself through the delayed electromag-

netic interaction. This means that a force on the phase space density at one point has a component that arises from its other part but at some earlier time. By definition the initial phase space density is defined at a certain time instant, which is a dividing point: the past of the phase space density before that instant is irrelevant for its future dynamics. However, with the delayed interaction such an initial instant is impossible to define, unless the phase space density is of particular kind in its past. In other words, once the delayed self interaction is introduced the choice of the initial conditions becomes a truly difficult task.

Obtaining initial $\rho_0(\vec{r}, \vec{p})$ involves simultaneous measurement of the position and velocity (momentum) of the particle. The experiments of this kind are difficult, and it is easier to measure position and velocity separately. In other words, the probability density $P(\vec{r})$ for the position is measured regardless of what the velocity of the particle is. The same applies for the measurement of the probability density for the momentum $Q(\vec{p})$; it is measured regardless of what the position of the particle is. From these two functions one needs to obtain the initial $\rho_0(\vec{r}, \vec{p})$, which are interrelated by

$$P(\vec{r}) = \int d^3p \rho_0(\vec{r}, \vec{p}), \quad Q(\vec{p}) = \int d^3r \rho_0(\vec{r}, \vec{p}). \quad (1.21)$$

The inversion problem of this kind does not produce the unique result. A function $f(\vec{r}, \vec{p})$ with the property

$$0 = \int d^3p f(\vec{r}, \vec{p}), \quad 0 = \int d^3r f(\vec{r}, \vec{p}) \quad (1.22)$$

can always be added to $\rho_0(\vec{r}, \vec{p})$ and the condition (1.21) would always be satisfied. There is infinite number of functions that satisfy the condition (1.22), and in general the information $P(\vec{r})$ and $Q(\vec{p})$ is not sufficient for the precise study of the time dependence $\rho(\vec{r}, \vec{p}, t)$. At best this information can be used as an approximation for the initial phase space density, a model in which it is assumed that $\rho_0(\vec{r}, \vec{p})$ is product separable, i.e. it is given in the form $\rho_0(\vec{r}, \vec{p}) = P(\vec{r})Q(\vec{p})$.

There is, however, one particular form of the function f , with the properties (1.22), which also has significant physical importance. If the function is chosen as

$$f(\vec{r}, \vec{p}) = \hat{n}_0 \cdot (\nabla_r \times \nabla_p) g(\vec{r}, \vec{p}) \quad (1.23)$$

where \hat{n}_0 is an arbitrary (constant) unit vector and g is an arbitrary function (that goes to zero sufficiently fast for large r and p) then

$$\int d^3p f(\vec{r}, \vec{p}) = (\hat{n}_0 \times \nabla_r) \cdot \int d^3p \nabla_p g(\vec{r}, \vec{p}) = 0.$$

Similarly it can be shown that the integral in the coordinates over the function $f(\vec{r}, \vec{p})$ is also zero. It is therefore shown that for the function (1.23) the conditions (1.22) are satisfied, which means that its addition to the phase space density does not have an effect on the probability densities $P(\vec{r})$ and $Q(\vec{p})$, and it appears to have no physical repercussions. This, however, is not true because the addition has important effect on the probability current, which is not zero for the function (1.23). From its definition (1.13) the current is given by

$$\vec{J} = \int d^3p \frac{\vec{p}}{m} [\hat{n}_0 \cdot (\nabla_r \times \nabla_p)] g(\vec{r}, \vec{p}) = -\frac{1}{m} (\hat{n}_0 \times \nabla_r) \int d^3p g(\vec{r}, \vec{p}) \quad (1.24)$$

where partial integration was performed. This form of the probability current indicates rotational flow of the probability density, and therefore angular momentum can be associated with it. The average angular momentum is defined as (1.18), and for the function (1.23) it is given by

$$\vec{L}_{ave} = m \int d^3r \vec{r} \times \vec{J}(\vec{r}) = 2 \hat{n}_0 \int d^3r d^3p g(\vec{r}, \vec{p}) \quad (1.25)$$

and indeed it is not zero (for arbitrary function g). The rotational flow of the probability density manifests itself as a contribution to the angular momentum of the particle, and because it does not have an effect on the probability densities $P(\vec{r})$ and $Q(\vec{p})$ it will be called the *internal rotation*, or the *spin*.

The choice (1.23) does not have effect on the initial probability densities $P(\vec{r})$ and $Q(\vec{p})$, but the question is: does it have on the time evolution of them? To answer it one should check whether the solution of the Liouville equation preserves the function form (1.23) for all times. By assuming that (1.23) is valid at any instant then the Liouville equation (for this particular component of the entire phase space density) is

$$\partial_t f + \frac{\vec{p}}{m} \cdot \nabla_r [\hat{n}_0 \cdot (\nabla_r \times \nabla_p)] g + \vec{F} \cdot \nabla_p [\hat{n}_0 \cdot (\nabla_r \times \nabla_p)] g = 0.$$

The second term in the equation can be transformed into

$$\frac{\vec{p}}{m} \cdot \nabla_r [\hat{n}_0 \cdot (\nabla_r \times \nabla_p)] g = [\hat{n}_0 \cdot (\nabla_r \times \nabla_p)] \frac{\vec{p}}{m} \cdot \nabla_r g$$

while the third term, in general, cannot be transformed in the same way. If the force, however, is of a general linear form

$$\vec{F} = \vec{F}_0 + F_1 \vec{r} + \hat{n} F_2 (\hat{n}_0 \cdot \vec{r}) \quad (1.26)$$

where \vec{F}_0 is a constant vector, F_1 a constant scalar and \hat{n} is arbitrary unit vector, the same transformation can be performed. In this case, indeed, the functional form (1.23) is preserved for all times, and its addition into the initial phase space density does not have any effect on the probability densities $P(\vec{r}, t)$ and $Q(\vec{p}, t)$. The impact of the time dependence $f(\vec{r}, \vec{p}, t)$ on the probability current (1.24) is not zero, however, the continuity equation is not violated because

$$\nabla \cdot \vec{J} = 0.$$

In all other circumstances the functional form (1.23) is not preserved in time, which means that the probability density $P(\vec{r}, t)$ will be affected by the addition of (1.23) into the initial phase space density. However, the important property of the function $f(\vec{r}, \vec{p}, t)$ is that its contribution to the probability densities $P(\vec{r}, t)$ and $Q(\vec{p}, t)$ does not affect the normalization because

$$\int d^3r d^3p f(\vec{r}, \vec{p}, t) = 0.$$

This means that the function f must have positive and negative values, and therefore the entire phase space density is not necessarily positive everywhere. By using the argument that the phase space density must be positive in the entire phase space (because it represents the probability density in the phase space) the function f can be dismissed as non-physical. This view is correct but it can be relaxed somewhat by making a compromise. The advantage of the non-physical phase space density is that it has great flexibility for describing various phenomena, and this feature should be accepted. In particular in the later developments it will be imperative to accept this possibility. Therefore a compromise should be encouraged, and it consists in accepting the possibility that the phase space density could be negative, but it is imperative that the functions $P(\vec{r}, t)$ and $Q(\vec{p}, t)$, which

are derived from it, are strictly positive, i.e. they represent probability density. In other words it is assumed that the probability densities $P(\vec{r}, t)$ and $Q(\vec{p}, t)$ are measured, and the phase space density $\rho(\vec{r}, \vec{p}, t)$ is only the mean of calculating their time evolution. One feature that is favorable for such a compromise is that the Liouville equation does not distinguish between “physical” and “non physical” phase space densities, both are solutions of the equation.

If the phase space densities with negative values are accepted as legitimate, there is still a problem to ensure that the probability densities $P(\vec{r}, t)$ and $Q(\vec{p}, t)$ must be positive. For example, one could choose at $t = 0$ arbitrary $g(\vec{r}, \vec{p})$ in (1.23) but it cannot be guaranteed that at any later time those probability densities are positive. One way of avoiding the problem would be to choose only special functions $g(\vec{r}, \vec{p})$ that ensure the criterion for all times, but that could be a very difficult task. On the other hand one could use this as the argument to dismiss phase space densities that contain the terms (1.23), however, those will be of considerable importance later when an additional degree of freedom will be discussed, the spin or the internal rotation of the phase space density. Therefore it is necessary to complete the task of finding the way to ensure that the probability densities $P(\vec{r}, t)$ and $Q(\vec{p}, t)$ are always positive for arbitrary function $g(\vec{r}, \vec{p})$ in (1.23). The most general way to accomplish this, the one that applies to more general force than (1.26), is to enforce the parametrization (1.23) for all times. This can be achieved by assuming that the function $f(\vec{r}, \vec{p}, t)$ is parametrized as

$$f(\vec{r}, \vec{p}, t) = (\nabla_r \times \nabla_p) \cdot \hat{n}(\vec{r}, t) g(\vec{r}, \vec{p}, t)$$

where now the vector \hat{n} is function of time and coordinates. In order to ensure that $f(\vec{r}, \vec{p}, t)$ satisfies the Liouville equation then the vector $\hat{n}(\vec{r}, t)$ should satisfy certain equation. The equation for $\hat{n}(\vec{r}, t)$ will be discussed later, but even at this stage it is obvious that it will be coupled to the equations (1.1) and together they would form generalized equations of motion for a particle. It is simple to show that by accepting parametrization of this kind for the function $f(\vec{r}, \vec{p}, t)$ its contribution to the probability densities $P(\vec{r}, t)$ and $Q(\vec{p}, t)$ is zero. This also means that their time rate is zero, which implies that the divergence of the relevant currents must also be zero. For the probability density $P(\vec{r}, t)$ the contribution to the current

from the function $f(\vec{r}, \vec{p}, t)$ is

$$\vec{J} = \int d^3p \frac{\vec{p}}{m} f(\vec{r}, \vec{p}, t) = \frac{1}{m} \nabla_r \times \int d^3p \hat{n}(\vec{r}, t) g(\vec{r}, \vec{p}, t)$$

and the divergence of \vec{J} is zero. The probability current that is associated with the probability density $Q(\vec{p}, t)$ is given by (1.14), and for the function $f(\vec{r}, \vec{p}, t)$ its value is

$$\begin{aligned} \vec{K}(\vec{p}, t) &= \int d^3r \vec{F}(\vec{r}, t) f(\vec{r}, \vec{p}, t) \\ &= \int d^3r \vec{F}(\vec{r}, t) (\nabla_r \times \nabla_p) \cdot \hat{n}(\vec{r}, t) g(\vec{r}, \vec{p}, t) \end{aligned} \quad (1.27)$$

which in general does not have zero divergence. This appears as the inconsistency with the fact that the appropriate $Q(\vec{p}, t)$ is identically zero, but it can be resolved with the more elaborate analysis. It will be made later, when the concept of the spin will be introduced, however, at this stage it should only be mentioned that introducing the function $\hat{n}(\vec{r}, t)$ assumes that the criterion (1.10) for deriving the Liouville equation should be generalized with the inhomogeneous term. The physical reason behind it will be discussed later, but again a brief note about it: additional degree of freedom in the equations of motion suggests that in the phase space two trajectories can cross at a particular instant in time, in which case phase space density may not be conserved, hence inhomogeneous term in the Liouville equation. Two trajectories can cross because the appropriate values of the additional degree of freedom are different.

It was shown that by adding $f(\vec{r}, \vec{p}, t)$ to the phase space density the total angular momentum of the particle is altered. It is of interest to see what happens to its total energy. By definition that part of the total energy that is associated with the function $f(\vec{r}, \vec{p}, t)$ is given by

$$E_{\text{tot}} = \int d^3r d^3p \left[\frac{p^2}{2m} + V(r) \right] (\nabla_r \times \nabla_p) \cdot \hat{n}(\vec{r}, t) g(\vec{r}, \vec{p}, t)$$

and after partial integration in the variables \vec{p} the result is

$$E_{\text{tot}} = - \int d^3r d^3p \frac{\vec{p}}{m} (\hat{n}_0 \times \nabla_r) g(\vec{r}, \vec{p}, t) = 0.$$

Therefore by adding the function $f(\vec{r}, \vec{p}, t)$ to the phase space density $\rho(\vec{r}, \vec{p}, t)$ the total energy of the particle is not altered.

1.2.2 Solving Liouville equation

1.2.2.1 Exact solution

After the initial phase space density is determined the next task is to solve equation (1.11). There are several ways it can be done, and they all utilize a very important property of the phase space density. It says that $\rho(\vec{r}_0, \vec{p}_0)$ is constant along a trajectory that starts at the point (\vec{r}_0, \vec{p}_0) and it is solution of the equations (1.1). If trajectory is written as a function $\vec{r}(t, \vec{r}_0, \vec{p}_0)$ then this property is given as a formal expression

$$\rho[\vec{r}(t, \vec{r}_0, \vec{p}_0), \vec{p}(t, \vec{r}_0, \vec{p}_0)] = \rho_0(\vec{r}_0, \vec{p}_0)$$

which is not very useful for calculating phase space density as a function of time because it is not the solution that is required. Solution normally assumes that phase space density is given as a function of arbitrary phase point (\vec{r}, \vec{p}) at time t , while that expression gives it as a function of initial points. Therefore in solution of the Liouville equation the phase point (\vec{r}, \vec{p}) should be explicitly present while the initial point (\vec{r}_0, \vec{p}_0) from which it is obtained should be calculated. In other words if the point (\vec{r}, \vec{p}) in the phase space is fixed then (\vec{r}_0, \vec{p}_0) should be found which is the initial condition that is associated with it. This is inverse problem of dynamics and can be easily solved by using the time symmetry of the set of equations (1.1). If the relationship $\vec{r} = \vec{f}(t, \vec{r}_0, \vec{p}_0)$ is known then the inverse is given as $\vec{r}_0 = \vec{f}(-t, \vec{r}, \vec{p})$, i.e. the initial and the final phase space points are interchanged and t is replaced by $-t$.³ The phase space density is then

$$\rho(\vec{r}, \vec{p}, t) = \rho_0[\vec{r}_0(-t, \vec{r}, \vec{p}), \vec{p}_0(-t, \vec{r}, \vec{p})]. \quad (1.28)$$

This is exact solution for the time evolution of the phase space density, but in practice it is often very difficult to implement. In relatively few instances the analytic expression for trajectory $\vec{r}(t, \vec{r}_0, \vec{p}_0)$ is available and so its inverse is not known. However, numerically the problem can be solved quite readily. By starting from some points (\vec{r}, \vec{p}) at time t one needs only to integrate the equations backward in time to $t = 0$. The resulting points are (\vec{r}_0, \vec{p}_0) , and the phase space density $\rho(\vec{r}, \vec{p}, t)$ is then equal to $\rho_0(\vec{r}_0, \vec{p}_0)$.

That formal rule how to calculate the phase space density describes its spirit, but in applications one should be careful how to implement it. For example when the force on the particle is explicitly time dependent the time reversal of trajectories are not obtained by formally reversing the

³This simple rule sometime does not apply, which will be pointed out in the examples where time evolution of phase space density is calculated.

arrow of time. Details of how this is achieved will be described in particular examples.

1.2.2.2 Numerical solution

Very often phase space density is not interesting to calculate but certain averages over it. For example, one would want to calculate the average of a function $A(\vec{r}, \vec{p})$ over momentum variables, whose exact expression is

$$a(\vec{r}, t) = \int d^3p A(\vec{r}, \vec{p}) \rho(\vec{r}, \vec{p}, t).$$

The previous, exact, method becomes tedious in application because it has two steps, one is obtaining entire function $\rho(\vec{r}, \vec{p}, t)$ and then it is integrated over \vec{p} . The intermediate step of calculating phase space density is not necessary, because the average $a(\vec{r}, t)$ can be calculated directly. The following is the numerical method that does that, and in general it will be assumed that the phase space density is also negative.

- (1) Random numbers for $(x_0, y_0, z_0, p_{0x}, p_{0y}, p_{0z})$ are generated with the weight $w = |\rho_0(\vec{r}_0, \vec{p}_0)|$, and the sign $sign[\rho_0(\vec{r}_0, \vec{p}_0)]$ of the phase space density at this point is determined. However, if the spherical coordinates are used for \vec{r}_0 and \vec{p}_0 then the weight function is

$$w = r_0^2 \sin(\theta_{0r}) p_0^2 \sin(\theta_{0p}) \rho_0(r_0, \theta_{0r}, \phi_{0r}, p_0, \theta_{0p}, \phi_{0r})$$

where the index refers to the variable for which the spherical angles are used. There are a number of numerical procedures that can be used for selecting random \vec{r}_0, \vec{p}_0 according to a prescribed weight or probability distribution. The most straightforward (not necessarily the most efficient) is the rejection method, and consists in choosing the interval within which (\vec{r}_0, \vec{p}_0) , or any other variables, are calculated. Maximum value M of the function $|\rho_0(\vec{r}_0, \vec{p}_0)|$ in this interval is determined. Random number for each of (\vec{r}_0, \vec{p}_0) and M (seven) parameters is generated with uniform weight in the interval of their definition. For the set (\vec{r}_0, \vec{p}_0) the function $|\rho_0(\vec{r}_0, \vec{p}_0)|$ is calculated, and if its value is greater than the generated random number in the interval M then this one is accepted as the correct one. Otherwise the whole set of (seven) random numbers is rejected, and the new is generated.

- (2) With the randomly chosen initial set (\vec{r}_0, \vec{p}_0) equation (1.1) is solved
- (3) The coordinate space is divided into boxes of the size $\Delta x \Delta y \Delta z$, and each box is labeled by the appropriate coordinates that contains

it. The coordinates \vec{r} of the trajectory at time t (appropriate momentum is \vec{p}) are therefore associated with this particular box, and $A(\vec{r}, \vec{p}) \text{ sign}[\rho_0(\vec{r}_0, \vec{p}_0)]$ is added to it. Also normalization parameter σ is calculated, which is obtained by adding $\text{sign}[\rho_0(\vec{r}_0, \vec{p}_0)]$ to it. After generating N randomly chosen trajectories the sum in the box with the coordinates (x_i, y_i, z_i) will have the value B_i and the average $a(\vec{r}, t)$ at this position will be approximately

$$a(\vec{r}_i, t) \approx \frac{B_i}{\sigma} \frac{1}{\Delta x \Delta y \Delta z}.$$

In the limit $N \rightarrow \infty$ and $\Delta x \Delta y \Delta z \rightarrow 0$ the procedure gives exact result.

The method is approximate but it is straightforward, easily implemented, and quite reliable. However, the result for the average $a(\vec{r}, t)$ is not a smooth function of the coordinates, it has fluctuations that originate in the random number generator. The rule is that the amplitude of fluctuations diminishes as $1/\sqrt{N}$, which is not a very fast convergence. In the case when the value of the average $a(\vec{r}, t)$ is of the order or smaller than the amplitude of these fluctuations the method becomes very difficult to implement, and as the rule its reliability becomes difficult to check.

1.2.2.3 Perturbation solution

In some cases perturbation method may be used, because it has certain advantages over the other ones. Its main idea is the following: if for a force \vec{F}_0 solution of equations (1.1) is known then when additional weak force \vec{F}_1 is applied the solution is expansion in the power series of the small parameter that measures it. More precisely, if $\vec{F} = -\nabla V_0 - \eta \nabla V_1$, where η is parameter that measures order of magnitude of the additional force, then solution of equations (1.1) is written as expansion

$$\vec{r} = \vec{r}_0 + \eta \vec{r}_1 + \dots$$

where only the leading correction is retained. The vector $\vec{r}(t)$ is solution of the equation

$$\begin{aligned} m\ddot{\vec{r}}_0 + m\eta\ddot{\vec{r}}_1 &= -\nabla V_0(\vec{r}_0 + \eta \vec{r}_1) - \eta \nabla V_1(\vec{r}_0 + \eta \vec{r}_1) \\ &\approx -\nabla_{r_0} V_0(\vec{r}_0) - \eta \nabla_{r_0} [(\vec{r}_1 \cdot \nabla_{r_0}) V_{r_0}(\vec{r}_0) + V_1(\vec{r}_0)] \end{aligned}$$

where the index of the gradient means that derivatives are taken with respect to the components of the vector \vec{r}_0 . By comparing the same order of magnitudes on both sides the equation for the first order correction to the unperturbed trajectory is

$$m\ddot{\vec{r}}_1 = -\nabla_{\vec{r}_0} [(\vec{r}_1 \cdot \nabla_{\vec{r}_0}) V_0(\vec{r}_0) + V_1(\vec{r}_0)]$$

for which the initial conditions are $\vec{r}_1(0) = \dot{\vec{r}}_1(0) = 0$, i.e. the exact and the unperturbed trajectories have the same initial conditions $\vec{r}(0) = \vec{r}_0(0) = \vec{r}_i$ and $\dot{\vec{r}}(0) = \dot{\vec{r}}_0(0) = \frac{1}{m}\vec{p}_i$. The phase space density, and this is of the main interest, is given as expansion

$$\rho(\vec{r}, \vec{p}, t) = \rho_i [\vec{r}_i(\vec{r}, \vec{p}, -t), \vec{p}_i(\vec{r}, \vec{p}, -t)] \approx \rho_0(\vec{r}_0, \vec{p}_0, t) + \eta \rho_1(\vec{r}_0, \vec{p}_0, t)$$

where ρ_i designates initial phase space density. Correction to the unperturbed phase space density is obtained by replacing the exact trajectory by its expansion in the power series of η , in which case

$$\begin{aligned} \rho_1(\vec{r}_0, \vec{p}_0, t) = & [\vec{r}_1 \cdot \nabla_{\vec{r}_0} \vec{r}_i + \vec{p}_1 \cdot \nabla_{\vec{p}_0} \vec{r}_i] \cdot \nabla_{\vec{r}_0} \rho_0 \\ & + [\vec{r}_1 \cdot \nabla_{\vec{r}_0} \vec{p}_i + \vec{p}_1 \cdot \nabla_{\vec{p}_0} \vec{p}_i] \cdot \nabla_{\vec{p}_0} \rho_0. \end{aligned}$$

This correction has only meaning if the initial phase space density does not changes very much, otherwise it would be more accurate to calculate the phase space density from the perturbed trajectory.

Higher order corrections become increasingly more complicated to calculate, but one can question their usefulness. Calculation of the first order, on the other hand, can be useful in some cases, in the analysis of the basic properties of the phase space densities.

When the force on the particle is weak the assumption can be made that the phase space density is the sum $\rho = \rho^{(0)} + \rho^{(1)}$, where $\rho^{(1)}$ is small and of the same order as the force. The main effort is to calculate this small term directly, without calculating the term $\rho^{(0)}$ whose fluctuations may have significant impact on the accuracy of the final results. A general idea how this can be done was developed earlier, but for the purpose of true numerical calculations it is not convenient. Better method for this purpose is developed here, and the starting point is the general expression for the time evolution of the phase space density

$$\rho(\vec{r}, \vec{p}, x_4) = \rho_0 \left[\vec{f}(\vec{r}, \vec{p}, -x_4), \vec{g}(\vec{r}, \vec{p}, -x_4) \right].$$

As described earlier if at time x_4 one wants the phase space density at the phase space point (\vec{r}, \vec{p}) then one needs to solve the equations of motion

backward in time to the initial instant $x_4 = 0$. The initial instant for this solution is at time x_4 and the initial condition is the pair (\vec{r}, \vec{p}) . The solution for the position vector is $\vec{f}(\vec{r}, \vec{p}, -x_4)$ and for the momentum is $\vec{g}(\vec{r}, \vec{p}, -x_4)$, and the end phase space point at $x_4 = 0$ is (\vec{r}_0, \vec{p}_0) . If the force on the particle is weak then the anticipation is that the solution will be slightly different from the one when no force is applied, at least this may be correct for a relatively brief interval of time. Based on this observation one can write the solution of the equations of motion in the form

$$\vec{f}(\vec{r}, \vec{p}, -x_4) = \vec{f}_0(\vec{r}, \vec{p}, -x_4) + \vec{f}_1(\vec{r}, \vec{p}, -x_4)$$

where

$$\vec{f}_0(\vec{r}, \vec{p}, -x_4) = \vec{r} - \frac{\vec{p}}{p_4} x_4, \quad \vec{f}_1(\vec{r}, \vec{p}, -x_4) = \vec{f}(\vec{r}, \vec{p}, -x_4) - \left(\vec{r} - \frac{\vec{p}}{p_4} x_4 \right)$$

where \vec{f}_1 is small. Similarly one writes for the solution for the momentum \vec{g} . The phase space density is now expanded as

$$\begin{aligned} \rho(\vec{r}, \vec{p}, x_4) &= \rho_0 \left(\vec{f}_0 + \vec{f}_1, \vec{g}_0 + \vec{g}_1 \right) \\ &\approx \rho_0 \left(\vec{f}_0, \vec{g}_0 \right) + \vec{f}_1 \cdot \nabla_{f_0} \rho_0 \left(\vec{f}_0, \vec{g}_0 \right) + \vec{g}_1 \cdot \nabla_{g_0} \rho_0 \left(\vec{f}_0, \vec{g}_0 \right) \end{aligned}$$

from where $\rho^{(0)} = \rho_0 \left(\vec{f}_0, \vec{g}_0 \right)$ and the small term $\rho^{(1)}$ is

$$\rho^{(1)}(\vec{r}, \vec{p}, x_4) = \vec{f}_1 \cdot \nabla_{f_0} \rho_0 \left(\vec{f}_0, \vec{g}_0 \right) + \vec{g}_1 \cdot \nabla_{g_0} \rho_0 \left(\vec{f}_0, \vec{g}_0 \right)$$

Similarly the higher order correction to the phase space density is calculated.

The correction to the phase space density is given directly in terms of the small quantities and therefore it is expected to give accurate results with relatively simple effort. However, in practical terms $\rho^{(1)}(\vec{r}, \vec{p}, x_4)$ is calculated from the following recipe. First it is transformed into the form

$$\begin{aligned} &\rho^{(1)}(\vec{r}, \vec{p}, x_4) \tag{1.29} \\ &= \left(\vec{f}_1 \cdot \frac{\nabla_{\vec{r}_0} \rho_0(\vec{r}_0, \vec{p}_0)}{\rho_0(\vec{r}_0, \vec{p}_0)} + \vec{g}_1 \cdot \frac{\nabla_{\vec{p}_0} \rho_0(\vec{r}_0, \vec{p}_0)}{\rho_0(\vec{r}_0, \vec{p}_0)} \right) \rho_0(\vec{r}_0, \vec{p}_0) \\ &= \Delta \rho_0(\vec{r}_0, \vec{p}_0; \vec{r}, \vec{p}, x_4) \rho_0(\vec{r}_0, \vec{p}_0) \end{aligned}$$

where the right side is again written in terms of the initial values (\vec{r}_0, \vec{p}_0) at $x_4 = 0$ for the pair of points (\vec{r}, \vec{p}) , and the corrections to the trajectory

are

$$\left(\vec{f}_1, \vec{g}_1 \right) = \left[\vec{r}_0 - \left(\vec{r} - \frac{\vec{p}}{p_4} x_4 \right), \vec{p}_0 - \vec{p} \right].$$

In practical terms the correction to the phase space density is calculated by choosing N sets of random pairs (\vec{r}_0, \vec{p}_0) from the phase space distribution $\rho_0(\vec{r}_0, \vec{p}_0)$, and for each one of them the equations of motion are solved. At time x_4 the solution is the pair (\vec{r}, \vec{p}) from which the small quantity $\Delta\rho_0$ is calculated and added to the box in the phase space within which the trajectory is found. In the end each box is normalized by the factor $1/(N\Delta_r^3\Delta_p^3)$, where Δ_r and Δ_p are the linear sizes of boxes in the coordinate and momentum directions, respectively. The procedure is the same as the numerical recipe that was described in Section 1.2.2, except that instead of adding 1 to the box one chooses $\Delta\rho_0$. Calculating correction to the average quantity, say the momentum, which is defined as

$$\langle \vec{p} \rangle = \int d^3r d^3p \vec{p} \rho^{(1)}(\vec{r}, \vec{p}, x_4)$$

follows the same logic. For each randomly selected pair (\vec{r}_0, \vec{p}_0) one calculates $\vec{p} \Delta\rho_0(\vec{r}_0, \vec{p}_0; \vec{r}, \vec{p}, x_4)$ and sums over all N number of them. The final result should be properly normalized and the result is the average momentum at time x_4 . The method has obviously difficulties when $\rho_0(\vec{r}_0, \vec{p}_0)$ is zero for some points in the phase space, but such cases should be examined individually.

One illustration how perturbation method works is given in Example 1.

1.2.3 Analysis of final results

Finally, analysis of the phase space density $\rho(\vec{r}, \vec{p}, t)$ should be made. There are no rules what to do about this and each example should be treated on its own, which also includes asking the question what is the relevant information that is extracted from the final phase space density. Very often it is not the phase space density that is of interest, mainly because it is not simple to measure it. The other quantities are of interest, such as the average coordinate of the particle or its velocity. The probability density for coordinates $P(\vec{r}, t)$ is of theoretical interest because it says where the particle is expected to be. However, none of those quantities are relevant when the electron is analyzed, say in the Hydrogen atom. In such a case other quantities become relevant, such as the radiation that is produced by its non uniform motion. It will be shown that the important quantity in

such a case is the probability density $P(\vec{r}, t)$ and the probability current, which play the role of the charge density and the charge current, respectively. Intensity of radiation from the electron is analyzed as the function of its frequency, which is called the spectrum, and in its essence it is given by

$$I(\omega) \sim \int d^3r d^3p dt e^{i\omega t} \rho(\vec{r}, \vec{p}, t).$$

Therefore the question that will often be asked is how to make this information useful to extract the knowledge about the phase space density.