CERN 75-8
24 June 1975

ORGANISATION EUROPÉENNE POUR LA RECHERCHE NUCLÉAIRE
CERN EUROPEAN ORGANIZATION FOR NUCLEAR RESEARCH

PROBLEMS OF CLASSICAL DYNAMICAL SYSTEMS

W. Thirring

Lectures given in the
Academic Training Programme of CERN 1974-1975

G E N E V A
1975

CERN - Service d'Information scientifique - RD/32 - 2000 - juin 1975

CERN 75-8
24 June 1975

ORGANISATION EUROPÉENNE POUR LA RECHERCHE NUCLÉAIRE
CERN EUROPEAN ORGANIZATION FOR NUCLEAR RESEARCH

PROBLEMS OF CLASSICAL DYNAMICAL SYSTEMS

W. Thirring*)

Lectures given in the
Academic Training Programme of CERN 1974-1975

G E N E V A
1975

*) University of Vienna, Vienna, Austria.
Notes taken by P. Grassberger.

ABSTRACT

After a brief survey of Hamiltonian theory and of relevant notions of set theory and manifolds, these lecture notes present some general properties of orbits, paying special attention to integrable systems. This is followed by a discussion of the Kolmogorov-Arnol'd-Moser theorem, dealing with the stability of orbits under small perturbations, and its importance for ergodic theory. Ergodicity and mixing are then treated in detail. In particular, the ergodic theorem of von Neumann is derived, and a specific example is given of a (strongly) mixing system.

SIS/kw-jb-yp-el-hm

CONTENTS

1. INTRODUCTION

A subject such as classical mechanics is interesting because out of very simple general laws -- essentially the laws of Newtonian mechanics and their relativistic generalizations -- emerges an overwhelming richness of phenomena. I do not think I have to review the beginnings of classical mechanics, as I assume that everybody knows about this. But I will try to come to the more interesting situations which have a great amount of richness and complications. Usually one does not notice them because the problems one can solve are more or less, by definition, simple, and what is hidden in the unsolvable are just the interesting aspects. Nevertheless, we shall see that we can make some statements and at least get an idea of the complicated real situation.

Just to be definite, let me write down the laws of classical mechanics. They are differential equations of a particular type, derivable from a Lagrangian. In the non-relativistic version, the Lagrangian is the kinetic energy of the particles minus the potential energy,

$$L = \sum_i \frac{m_i}{2} \dot{q}_i^2 - V(q_i) \ , \tag{1}$$

where the coordinates q_i and potential $V(q_i)$ will be specified later.

If you treat electromagnetic interaction of a particle relativistically, you have a Lagrangian which looks similar,

$$L = \frac{m}{2} \dot{q}^2 - e \, \dot{q} A(q) \ , \tag{2}$$

but now q, $\dot{q}$, and the vector potential $A(q)$ are four-vectors, and the dot means derivation with respect to *proper* time.

Finally, if you want to treat gravitation relativistically, you have again a Lagrangian which for a single particle is of the form

$$L = \frac{m}{2} \dot{q}^\alpha \dot{q}^\beta g_{\alpha\beta}(q) \ . \tag{3}$$

The interaction now is contained in the $g_{\alpha\beta}(q)$.

The structure in all three cases is very similar. The equations of motion are the variational equations. Of course, the above problems can also be stated in Hamiltonian form, with the Hamiltonians*)

$$H = \sum_i \frac{1}{2m_i} p_i^2 + V(q_i) \qquad \text{(non-relativistic)} \tag{4}$$

$$= \frac{1}{2m} \left[p + e\, A(q)\right]^2 \qquad \text{(relativistic electromagnetism)} \tag{5}$$

$$= \frac{1}{2m} p^\alpha p^\beta g_{\alpha\beta}(q) \qquad \text{(relativistic grativation)} \ . \tag{6}$$

---

*) Notice that, in the second and third examples, H is not the generator of time translations, but of translations in proper time. Accordingly, the value it takes is not the energy but the mass of the particle.

These are the standard formulations. To discuss the coordinates in detail, let us go to the non-relativistic Kepler problem. There, a natural choice for the q_i is the Cartesian coordinates $\vec{x}$, and the potential is

$$V(\vec{x}) = \frac{\alpha}{|\vec{x}|}, \qquad \alpha = \text{const.} \tag{7}$$

There is a small difficulty. You may ask what happens if the particle hits right at the centre, where the equations of motion become singular, and you do not know what really happens. You might say that singularities do not really exist in nature, so the centre is indeed smooth and the particle goes straight through. Another argument might be that you do have a point mass at the centre, and when you hit closer and closer, the orbit looks more and more like a reflection. (Though a reflection seems strange for a purely attractive force, you might have seen this in movies where Kepler motion is filmed.)

Perhaps the best is, however, to say that the problem is somewhat academic. If you apply Kepler motion to stars, neither the one nor the other will happen, but a terrible catastrophe will be the result. So we have to exclude the origin explicitly and to say that the equations do not apply in all space R^3, but only in $R^3 - \{0\}$. At this point, and for future reference, it will be useful to recapitulate some modern mathematical notions.

2. A MATHEMATICAL INTERLUDE

I assume that you are familiar with the notations of set theory, so I will not spend much time on this: R means always the set of real numbers, R^3 the space of three dimensions, $\{0\}$ means the origin, and $R^3 - \{0\}$ means the three-dimensional space with the origin taken out.

Perhpas not so familiar as set theoretical notions are the topological ones we shall need. First we need the concept of open and closed sets. To start with the simplest case, *open intervals* in R are intervals which do not include their end points. *Open sets* are just arbitrary unions of open intervals, and thus arbitrary unions of open sets are again open sets. If you generalize this to R^3, you can similarly define open cubes, and arbitrary unions of them form open sets. *Closed* intervals are defined as complements of open intervals, or equivalently as intervals including the end points. *Closed sets* are the complements of open sets, and arbitrary intersections of them are again closed sets. According to this, a single point is a closed set because it is the intersection of all sets containing it, and $R^3 - \{0\}$, being the complement of a closed set, is an open set.

Another notion we need is the denseness of a set. Consider a set A. The union of all open sets contained in A, which by definition is again open, is called the *interior* of A, denoted by int(A):

$$\text{int}(A) = \cup B\,; \quad B \subseteq A\,, \quad B \text{ open}\ .$$

Similarly, the intersection of all closed sets containing A is again closed. It is called the *closure* of A, and is written $\bar{A}$;

$$\bar{A} = \cap B\,; \quad B \supseteq A\,, \quad B \text{ closed}\ .$$

A set A is called *dense* in another set -- say in M -- if $\bar{A} = M$, that is if the closure of A is the whole set under consideration. If A is dense, it means that there are no gaps left between the points contained in A, but it does not mean that A makes up almost the whole set M, or even a big part of it. A good example is the rational numbers which are dense in R (there are no open intervals which do not contain a rational number). But the complement of them -- the irrational numbers -- is also dense.

If a set is not dense, this does not yet say very much. A more useful statement about a set is to say it is *nowhere dense*, which means that even its closure has an empty interior,

$$\mathrm{int}(\bar{A}) = 0 \ .$$

A finite number of points, for example, are nowhere dense in R.

We shall come back to these notions later, as they will arise. What was important above was to consider the equations of motion not on the whole of space but in an open subset where they make sense. In this open set there are still orbits which at some later time hit the excluded point, but we shall see that in simple cases we can take out some further part of phase space which contains these orbits. In the remaining part the time evolution is defined for all times, without any catastrophe ever occurring. The importance of this is that time translations then form a one-parameter group, which will be used in Section 6. In the example of Kepler motion, we had to take out only a single point of R^3, and correspondingly a small portion of phase space. In the other two examples, the problem is somewhat more serious. In the relativistic Coulomb problem, i.e. Eq. (2) with the vector potential given by $A_\mu = [A_0(\vec{x}), \vec{0}]$, $A_0(\vec{x}) = \alpha/|\vec{x}|$, a particle hitting the centre at a sufficiently small impact parameter will spiral into it in a finite time, and so we have to take out all such orbits. Similarly, in the case of relativistic gravitation, we have the possibility of black holes, and thus we have to exclude a sizeable part of phase space.

To make the mathematical setting somewhat wider, we have to introduce also the concept of a *differentiable manifold*, which is a generalization of the concept of a smooth surface. A differentiable manifold looks *locally* like a Euclidean space of a certain dimension -- which is then called the dimension of the manifold -- but it may be different globally. A circle is, for example, a one-dimensional manifold, and a sphere a two-dimensional one. Open subsets of R^n are other examples of manifolds.

Why should one consider this notion? We shall see that the motions which emerge are characterized by certain constants, thus the orbits will not fill all of the space but will be constrained to satisfy conditions

$$f_i(q) = 0 \ ; \quad i = 1,\dots, k; \quad q \in R^n \ .$$

If the $f_i(q)$ are independent and sufficiently differentiable functions of the coordinates, this defines just a manifold of n-k dimensions in n-dimensional space. The unit circle is for instance defined by $f(x) \equiv x_1^2 + x_2^2 - 1 = 0$. It is advisable to study the equations of motion directly on such manifolds. In particular, in the generally relativistic case the space to start with will not even be part of Euclidean space, but something which locally -- as seen by a short-sighted observer -- looks like Euclidean space. Globally it might be completely different.

The mathematical definition of a manifold is that locally we can always introduce coordinates which map it into some open subset of Euclidean space, although these coordinates cannot be used globally. To cover other parts of the manifold, we might be obliged to use different mappings. There are a great many technicalities that go along with this, but we do not have the time to go into them now. To keep the course at an elementary level, it will be necessary to sacrifice some precision and elegance, and only the general idea can be described*).

As a simple example, let us again discuss the circle or, equivalently, the use of polar coordinates in the plane. (The freedom to use other coordinates besides the Cartesian ones corresponds exactly to the fact that one is working on a manifold. A manifold is not characterized by a single coordinate system but by a whole equivalence class of coordinate systems.) To describe the position on the circle, we will naturally introduce an angle. But this is not a permissible global coordinate since, as we go around the circle, it would jump by 2π. Coordinates should of course be smooth functions. But we are free to use the angle for almost all of the circle, and to use another angle for the rest.

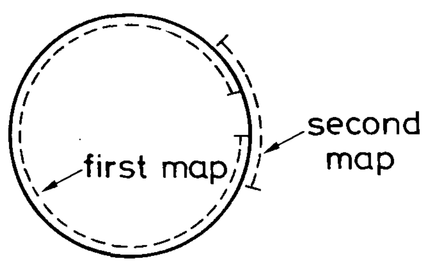

Fig. 1

3. CONSTANTS OF MOTION AND GENERAL PROPERTIES OF ORBITS

Let us come back to the equations of motion, which we write in the Hamiltonian form: the time derivative of an arbitrary differentiable function is given by the Poisson bracket with the Hamiltonian,

$$\begin{aligned} \dot{f}(p,q) &= \{H(p,q),\, f(p,q)\} \\ &\underset{\text{def.}}{=} \sum_i \left(H_{,pi}\, f_{,qi} - H_{,qi}\, f_{,pi}\right) . \end{aligned} \tag{8}$$

According to what we said above, this function can be considered as a new coordinate which is as good as any other -- and in particular as a Cartesian coordinate.

If we can find an f such that $\dot{f} = 0$, we see that the orbit is restricted to a submanifold defined by

$$f(p,q) = \text{const.}$$

One constant that is immediately seen is the Hamiltonian itself, the corresponding manifold being called the energy shell**). An interesting question is, of course, how many constants exist, and whether they determine the orbit. If there are n degrees of freedom, the phase space has dimension 2n, and the energy shell is 2n-1 dimensional. If we can find a sufficient number of constants of the motion, we can maybe reduce this to a one-dimensional manifold, which we would call the orbit.

---

*) Technical details can be found in lecture notes: W. Thirring, Vorlesungen über Mathematische Physik, T5: Punktmechanik. (Inst. f. Theoretische Physik, Universität Wien, 1974)

**) Or mass shell (see footnote on p. 1).

Let us look at the Kepler problem, given by Eq. (7), to see whether this works. Here $n = 3$, so the phase space is six-dimensional. Conservation of energy and angular momentum $\vec{L}$ restricts this to a two-dimensional submanifold, which in our case corresponds to a motion in a plane perpendicular to $\vec{L}$ with the momentum uniquely determined by the position. As we can easily check, a further constant quantity is the Lenz vector:

$$\vec{F} = [\vec{p} \wedge \vec{L}] + m\alpha \frac{\vec{x}}{r} . \tag{9}$$

Now it seems we have seven constants (two too many). This just means that they are not all independent, and we find indeed two relations,

$$\vec{F} \cdot \vec{L} = 0$$
$$\vec{F}^2 = 2m\vec{L}^2 H + m^2\alpha^2 .$$

Thus we have exactly five independent constants, and the orbit is uniquely defined. The actual Kepler ellipses are obtained by noticing that Eq. (9) implies

$$\vec{F} \cdot \vec{x} = \vec{L}^2 + m\alpha r ,$$

and therefore

$$r = \frac{\vec{L}^2}{F \cos \theta - m\alpha} . \tag{10}$$

At this point some confusion might arise. On the one hand, it is well known that in a system of 2n first-order differential equations there are always 2n integration constants, and 2n-1 of them are time-independent constants of the motion. In fact, *locally* we can always introduce a coordinate system $(y_1, \ldots, y_{2n})$ such that the time evolution becomes trivial: when the time changes from zero to t, $(y_1, \ldots, y_{2n})$ changes:

$$[y_1(0), y_2, \ldots, y_{2n}] \to [y_1(0) + t, y_2, \ldots, y_{2n}] . \tag{11}$$

On the other hand, we learn in statistical mechanics that there is in general no constant of the motion except for the energy, which seems to be a contradiction of the above.

The point is that one has to distinguish between *local* and *global* quantities. Coordinates satisfying Eq. (11) can in general only be found locally, i.e. the constants change in general if we go from one mapping to the other. In order to define a submanifold, we need, however, quantities which are globally defined, independently of any map. I will use the term *"constant of motion"* only for *global* constants, since they are the ones that really matter.

Let us take the example of a two-dimensional harmonic oscillator, defined by a Hamiltonian:

$$H = \tfrac{1}{2}(p_1^2 + \omega_1^2 q_1^2 + p_2^2 + \omega_2^2 q_2^2) . \tag{12}$$

We assume the potential to be anisotropic, so the frequencies ω_1 and ω_2 may be different. Two constants of motion are clearly given by the two energies, or equivalently by the functions

$$K_i = q_i^2 + \frac{p_i^2}{\omega_i^2} ; \quad i = 1,2 . \tag{13}$$

Using these and the angle variables $\theta_{1,2}$, the solution can be written

$$(q_i, p_i) = \sqrt{K_i}\,(\sin 2\pi\theta_i,\ \omega_i \cdot \cos 2\pi\theta_i) \ . \tag{14}$$

The time evolution in the new (local) coordinates (k_i, θ_i) is simply

$$(K_i, \theta_i) \rightarrow \left(K_i,\ \theta_i + \frac{\omega_i t}{2\pi}\right) ,$$

i.e. the θ_i change linearly with time.

Phase space in this example is four-dimensional, and we have already found two (global) constants. The maximum we might hope to get is three constants. Whether we can indeed find a third depends on whether the ω_i are rationally dependent or not. The θ_i are local variables, defined only modulo 1, and in the (θ_1, θ_2) - plane the motion appears as shown in the figure. The fact that the θ_i are defined only modulo 1 means that after the orbit reaches one of the boundaries, it will reappear on the other side. A local constant is easily found to be

$$\theta_1\omega_2 - \theta_2\omega_1 \ ,$$

but each time one of the θ_i jumps by one unit, this quantity will in general also jump, so it is not a global constant. If, however, the frequencies are rationally dependent, that is if $\omega_1 g_2 = \omega_2 g_1$ with some integers g_i, a global constant is for instance given by

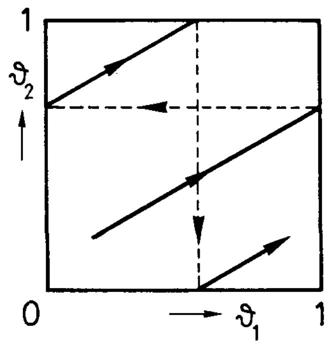

Fig. 2

$$K_3 = \sin 2\pi(\theta_1 \cdot g_2 - \theta_2 \cdot g_1) \ . \tag{15}$$

In this case the orbit is a one-dimensional manifold, which means in the above figure that after some time the orbit will close.

If the frequencies are not rationally dependent, we can of course not construct K_3 in the above way. Indeed, it turns out that we cannot find a third global constant at all, since we can show that the orbit is dense in the square $(0,1) \times (0,1)$. According to the definition of denseness, this means that its closure is the square, and that the orbit comes arbitrarily close to any point in it.

You know that historically this had some significance since originally Boltzmann, in his ergodic theory, assumed that an orbit may even go through each point. In fact there are the so-called Peano curves that have this pathological feature, but it is impossible for differentiable curves. In particular, an orbit always has measure zero. Nevertheless it can be dense.

Let us now prove that in our example the orbit is dense in the square. To do this, we consider a fixed value of θ_1. We have to show that the corresponding values of θ_2 are dense in (0,1) or, equivalently, that there is no open interval which does not contain a point of the orbit. Two different times t and t' which correspond to the same θ_1 must obey $\omega_1(t' - t) = 2\pi g_1$, where g_1 is some integer. The corresponding difference of θ_2 is

$$\theta_2(t') - \theta_2(t) = \frac{1}{\omega_1}(\omega_2 g_1 - \omega_1 g_2), \quad g_2 = \text{integer} \tag{16}$$

which is different from zero by our assumption of rational independence. So there are infinitely many different values of θ_2 corresponding to the same value of θ_1. Since all these values are on the compact interval (0,1), they must have an accumulation point, i.e. for any ε one can find two times t_1, t_2 such that

$$|\theta_2(t_1) - \theta_2(t_2)| < \varepsilon .$$

Owing to the time translation invariance, θ_2 will advance after a second elapse of the same time difference t_2-t_1 by the same small amount, and so on. Eventually, the values of θ_2 will fill the whole interval with points whose distance is closer than ε, which just means that they are dense. Q.E.D.

This example suggests the following as a general rule: if an orbit is restricted to a certain manifold, it either fills it densely or there are further constants of the motion, defining a submanifold which is then densely filled. Unfortunately, this is not true in general, which comes from the fact that the topological notion of denseness is not quite adequate to this problem. Later it will be replaced by measure-theoretic concepts.

Let me illustrate this by a one-dimensional potential of a shape which nowadays is a favourite among field theoreticians. Since the energy shell has already only one dimension, one would expect that the orbit fills the energy shell densely unless there is a further constant of motion. For $E < V_0$, the energy shell consists of two disconnected pieces, and each orbit covers only one of these halves. Now one can construct a constant of the motion for this, namely a suitably "smoothed" step function, which is +1 on the right part and -1 on the left. So this is not yet a counterexample. But if $E = V_0$, this no longer works. At this energy there are three different orbits: one where the particle is always on the right of x_0, one where it stays on the left, and one where it sits just at x_0, on top of the bump of the potential. In this case, no separating constant can be constructed. In particular, if we wanted to use the same trick as for $E < V_0$, we would have to use a true step function. This is, however, not permissible since a constant of motion should always be a differential function.

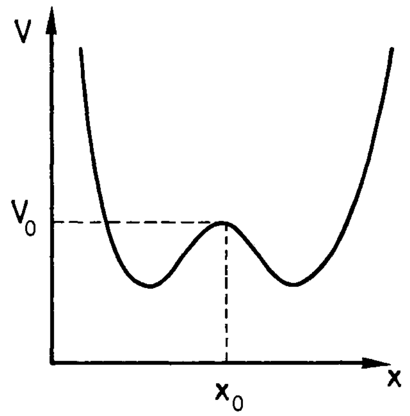

Fig. 3

Although this counter-example shows that the above suggestion cannot be exactly true, what we have learnt from the harmonic oscillator is rather typical of what happens in general. Most of the solvable problems are of similar type, and the interesting problems are just those which go beyond this.

4. INTEGRABLE SYSTEMS

Those systems for which the pair of harmonic oscillators is the prototype are called "integrable systems". A Hamiltonian system of n degrees of freedom in general is called integrable if there exist n (global) constants of the motion K_i, which satisfy

$$\{K_i, K_j\} = 0 \quad i,j = 1,2, \dots n \ . \tag{17}$$

In our example, these were just the K_i defined by Eq. (13). Besides them, we had two other variables, namely the angles θ_i defined by Eq. (14). One easily checks that

$$\{\theta_i, \theta_j\} = 0 \tag{18}$$

and, apart from constants which could have been eliminated by redefining the θ_i,

$$\{K_i, \theta_j\} = \delta_{ij} \ . \tag{19}$$

Indeed, this is a general property. There is a theorem by Liouville which states that in any integrable system with n degrees of freedom, one can always find n variables θ_i in addition to the K_i such that the Poisson brackets (17)-(19) hold. The K_i are called action variables, the θ_i angle variables.

One has to stress that the θ_i are only local variables, i.e. they are not defined everywhere on the manifold, and they are not uniquely determined by Eqs. (18) and (19). For any arbitrary differentiable function $f(K) = f(K_1, \dots, K_n)$,

$$\theta_i + \frac{\partial f(K)}{\partial K_i}$$

are also suitable angle variables, and this is exactly the whole amount of non-uniqueness. The proof of the theorem is not very difficult, but I do not want to go into details.

Let us now use the fact that the K's are constants, to deduce the time evolution of the θ_i. Since the K_i at two different times are the same, the corresponding θ_i must be equal modulo the above arbitrariness,

$$\theta_i(t) = \theta_i(0) + \frac{\partial f(K,t)}{\partial K_i} \ .$$

Because of the group property of the time development, $f(K,t)$ must be a linear function of time, which we call $t \cdot H(K)$, and we write

$$\theta_i(t) = \theta_i(0) + t\,\frac{\partial H(K)}{\partial K_i} \ . \tag{20}$$

Obviously, the $H(K)$ introduced in this way is the Hamiltonian expressed by the K_i, and we have in general exactly what we had found in our example, namely that the angle variables vary linearly in time.

So far, the θ_i are only locally defined. There is another theorem, due to Arnold, that the manifold described by the θ_i (keeping all K_i constant) is the product

$$T^{n-r} \times R^r, \qquad 0 \le r \le n$$

of a torus in n-r dimensions times the real numbers to the power r. This means, in practice, that n-r variables are cyclic, i.e. are angles varying over a finite interval, while the others correspond to linear motion. Trivial examples are:

i) one free particle -- in this example, the K_i form the momentum vector, and the θ_i the position. There is no cyclic motion, thus $n = r = 3$;

ii) the pair of harmonic oscillators considered above -- here both θ_i are angles, and thus $r = 0$, $n = 2$.

Non-trivial examples for the above are the relativistic motion in a constant electromagnetic field or in the field of a plane wave, and the betatron. As another example, let us discuss the *relativistic Kepler problem* in somewhat more detail.

In the notation used in the introduction, we consider Eq. (5) with a static vector potential $A_\mu(\vec{x}) = (A_0(\vec{x}),\vec{0})$ with

$$A_0(\vec{x}) = \frac{\alpha}{|\vec{x}|}\,, \qquad \alpha = \text{const.}$$

In Eq. (5), the Hamiltonian for relativistic motion was written in extended phase space, in which both space and time variables are functions of proper time. With the above vector potential, p_0 (the energy) is constant and the dependence of time on proper time is such that the latter can be trivially eliminated from the equations of motion. The spatial part of the motion is described by a Hamiltonian,

$$H = \frac{\vec{p}^2}{2m} + \frac{\alpha}{r}\frac{p_0}{m} - \frac{\alpha^2}{2mr^2}\,, \tag{21}$$

which looks very similar to the non-relativistic Hamiltonian except for the last term and for the constant factor p_0/m in the term proportional to $1/r$.

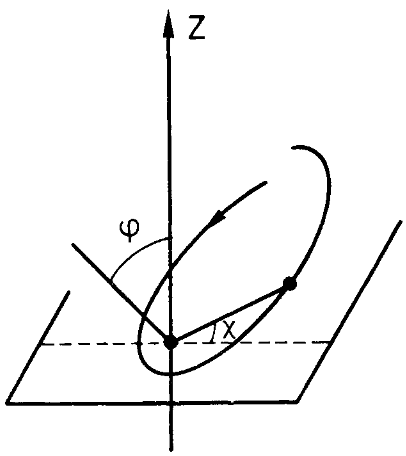

Fig. 4

It is clear that angular momentum is conserved and the motion will be restricted to a plane. Thus a first pair of action and angle variables is the third component of angular momentum L_3, and the angle ϕ between the normal to the plane of motion and the z-axis. A second action variable is

$$K_2 = |\vec{L}|\,,$$

the absolute value of angular momentum. Its conjugate variable is χ, the angle describing the actual position of the particle in the plane of motion. For the radial motion, one usually takes

$$K_3 = -\sqrt{\vec{L}^2 - \alpha^2} + \frac{\alpha p_0}{\sqrt{2m|H|}}\,. \tag{22}$$

The conjugate variable is a somewhat complicated function $f(r,p_r)$ of the radius and the radial momentum, which we will not bother to write up. It is clear that the Poisson brackets between the K_i vanish, and we can check in detail that the conjugate variables are as stated above.

The recipe now is to express the Hamiltonian in terms of the K_i, and its derivatives with respect to the K will give the frequencies of the corresponding motions. Inverting the above, we find

$$H = -\frac{\alpha^2 p_0^2}{2m\left(K_3 + \sqrt{K_2^2 - \alpha^2}\right)^2} \quad . \tag{23}$$

This does not depend on $K_1 = L_3$, so the corresponding angle ϕ is fixed, as was already clear from the beginning. The derivative $\partial H/\partial K_2$ gives the frequency ω_2 of revolutions, while the radial oscillation is governed by another frequency $\omega_3 = \partial H/\partial K_3$. The manifold of solutions is thus described by three cyclic variables and is therefore restricted to a torus T^3, with however one of the frequencies equal to zero.

In general, the two other frequencies ω_2 and ω_3 will be rationally independent, and fill densely a two-dimensional region. Of course, as you change the constants continuously, the frequencies will change between rationally independent and rationally dependent. In the latter case, you will get closed orbits, but infinitely close to this there will be orbits which fill the shaded ring densely.

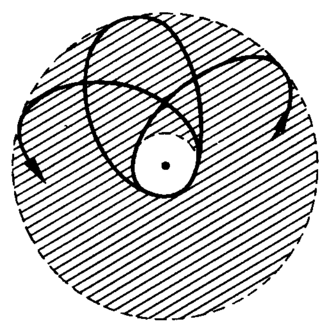

Fig. 5

5. PERTURBATIONS AND THE KAM THEOREM

The above are those cases which are usually found in text-books because they are solvable. But already in slightly more complicated cases -- as three bodies with gravitational interactions -- one is in a situation where one cannot find the constants of the motion exactly. Often, however, one can consider problems as approximately integrable, with some slight perturbation.

Consider, for example, a three-body problem where the distance between the first two particles is small compared to their distance from the third. The sum of the three potentials can be approximated by an interaction of the third particle with the centre of gravity of the first two on the one hand, and the interaction between particles 1 and 2 on the other hand. In this approximation, the Hamiltonian splits into two separate parts, and the exact problem is reduced to two separate Kepler problems with a slight disturbance.

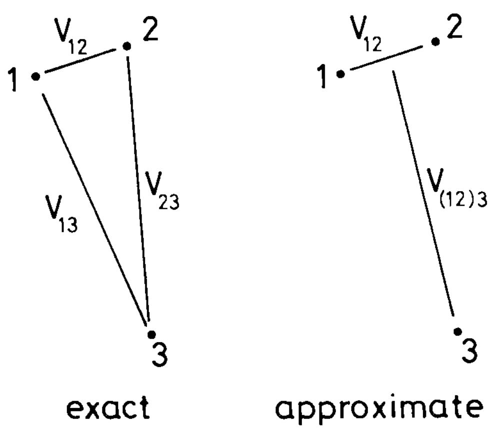

Fig. 6

So assume you have a Hamiltonian

$$H = H_0(K_i) + \varepsilon H'(K_i, \theta_i) \tag{24}$$

which consists of a piece H_0 depending only on the K_i, and of a piece $\varepsilon H'$ which depends also on the angle variables θ_i. The usual way of proceeding is to search for other K_i's which are such that expressed in them we have again an integrable system. So one makes a canonical transformation

$$K_i \to \bar{K}_i \ ,$$

$$\theta_i \to \bar{\theta}_i$$

in order to eliminate the part in the Hamiltonian depending on the $\bar{\theta}_i$. In general, this can be done only successively, getting first

$$H = \bar{H}_0(\bar{K}_i) + \varepsilon^2 H''(\bar{K}_i,\bar{\theta}_i) \ ,$$

and suppressing the perturbation term in each step by a factor ε.

This was attempted long ago, mainly by Poincaré. He found that this iteration procedure does not converge in general. This led him to suggest that the whole idea is wrong, and after a slight disturbance one will in general have no global constants of the motion (except the energy). In such a case, the orbit would fill the energy shell completely, not just some submanifold. But there were some assumptions involved in his proof of non-convergence, and Weierstrass, for instance, pointed out that there might be circumstances where it does not work. Nevertheless it was widely accepted, and was used later by Fermi to show that any small perturbation renders a system ergodic (we shall come to this notion later). Immediately after Fermi's paper, in the middle of the twenties, somebody pointed out that this was not correct. But his opinions were brushed aside by claims that this concerned only some irrelevant mathematical pathologies.

The correct answer clearly depends on what is a pathology, i.e. what is likely to happen and what is unlikely. To get a good idea of this, we have to introduce another notion, namely that of the *Liouville measure*. I suppose you know Liouville's theorem which tells you that one has a natural volume element in phase space.

$$dq_1 \cdots dq_n \ dp_1 \cdots dp_n \ , \tag{25}$$

which is invariant under canonical transformations. Thus it does not depend on what variables one uses to write it, as long as they are canonically conjugate. It is reasonable to assume that this measure is also natural in the sense that processes which have a big volume in phase space are likely to happen, and vice versa.

We will not go into the details of the proof of invariance, though it is not difficult. In the modern language we have used, there is indeed hardly anything to prove, as the definitions were arranged such that the invariance is automatic. Canonical transformations were per definition such as to leave the two-form $dq \wedge dp$ invariant, and $dq_1 \ldots dq_n \, dp_1 \ldots dp_n$ is just the n^{th} power of this and is thus also invariant.

In particular, Liouville's theorem tells us that the Liouville measure is constant in time, since the time evolution is a particular canonical transformation.

Using the Liouville measure to estimate what is a pathology and what is likely to happen, the question raised by Poincaré was finally settled by Kolmogorov in his famous address to the mathematical congress in Amsterdam in 1954. Later on, his proof was improved by Arnol'd and Moser.

Kolmogorov showed the following (KAM theorem): if $\varepsilon = 0$, the motion is constrained to tori (assume there are no variables describing linear motion), the revolution frequencies ω_i depending on the K_i. In general, these frequencies are different from each other. Now turn on the perturbation. Those tori for which the frequencies are sufficiently irrational relative to each other, are not completely destroyed but are only deformed. Those tori, however, whose frequencies are not sufficiently irrational, are completely destroyed and nothing is left of them. The concept of sufficient irrationality means the following: frequencies ω_i are rationally dependent if there exist some integers g_i (not all equal zero) such that

$$\sum_i g_i \omega_i = 0 \ .$$

If the ω_i are irrational with respect to each other, this is not true, but the sum can become arbitrarily small by taking some of the g_i large enough. Sufficient irrationality finally means that this sum is not arbitrarily small, but

$$\left| \sum_i g_i \omega_i \right| \geq f(g_i) \quad \forall g_i \ , \tag{26}$$

with a certain well-defined function $f(g_i)$.

In addition, if this condition is satisfied, the above iteration scheme converges and yields the correct constants of motion.

The most remarkable thing found is that Eq. (26) -- and thus the stability of the tori -- holds on a closed subset I of phase space (essentially a set of closed surfaces) which is nowhere dense but nevertheless has a measure of order of magnitude $1 - \mathcal{O}(\varepsilon)$. Here, the total measure of phase space is normalized to unity, so the complement I^c has only a measure of magnitude $\mathcal{O}(\varepsilon)$, though it is dense. On this complement I^c the motion is rather chaotic, and therefore the components of I^c are called zones of instability.

Thus, if denseness were the correct property for estimating what is probable, Fermi would have been right. But what really counts in this respect is the Liouville measure, and he was essentially wrong. For small perturbances the stable parts fill amost all of phase space. Presumably, this is the only wrong paper by Fermi, but it was copied in many textbooks. In particular, the above means that the famous "speck of dust" does not work in statistical mechanics.

You see that the whole depends on an apparent clash between topological and measure-theoretical concepts. To see in some more detail how a paradoxical situation such as the above can happen, let us consider a simple example. Take the real numbers, and among them the rational ones. Though the latter are dense they are countable, and we numerate them $q_1, q_2, \ldots$. Take an open interval $(q_1 - \varepsilon/2, q_1 + \varepsilon/2)$ with length ε centred at q_1, a similar interval $(q_2 - \varepsilon/4, q_2 + \varepsilon/4)$ with length $\varepsilon/2$ around q_2, one with length $\varepsilon/4$ around q_3, and so on. There is a general theorem which tells us that the measure Ω of a union of sets is never larger than the sum of the individual measures,

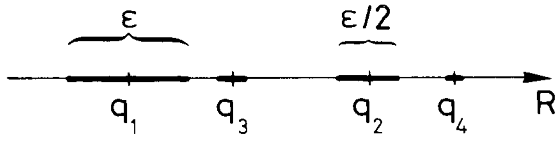

Fig. 7

$$\Omega\left(\bigcup_i A_i\right) \leq \sum_i \Omega(A_i) \ , \tag{27}$$

which in our case means that the total length of a union of intervals is always smaller or equal to the sum of the individual lengths. (This is of course very plausible: if the intervals do not overlap, both sides are equal, whereas overlapping intervals contribute twice to the right-hand side and only once to the left.) Therefore the union of the above intervals has a measure which is less than $\varepsilon + \varepsilon/2 + \varepsilon/4 + \ldots = 2\varepsilon$. Since it is the union of open intervals it is an open set, and it is dense since it contains all rational numbers. Its complement is nowhere dense since any two of its points are separated by one of the above intervals, but it has nevertheless a big measure.

One might suspect that this example is a mathematical pathology with no relevance for physics, but indeed we have seen essentially what happens in physics.

The proof of the KAM theorem is too complicated to be given here. It is just a careful estimate of the convergence of the iterative procedure discussed at the beginning of this section. The physical idea is, however, not too difficult to understand: if there are two frequencies which are rationally dependent, an arbitrarily small perturbation will lead to a resonance behaviour between them in which the two modes influence each other strongly and destroy the regular motion. Since we consider in general non-linear systems, the same will happen if the ratio of the frequencies is irrational but sufficiently close to a ratio of small integers. If this is not the case, which means that the frequencies are sufficiently irrational, no such internal reasonances will build up, and a small perturbation will lead to only slightly perturbed orbits. The difficulty of the proof consits essentially in finding out precisely what is sufficiently irrational.

6. ERGODICITY

Up to now, our main concern was the study of individual orbits. The KAM theorem showed us, however, that this is not a very meaningful problem: infinitely close to each stable orbit there are unstable ones, and since it is impossible to prepare a system with infinite accuracy, one does not really know what will go on. The only thing one can really do is to give the probability that the system is in a particular point in phase space, at some given time. In a quantum mechanical language you would say that the probability distribution $\rho(q,p)$ describes the state of the system. The observables are functions $f(q,p)$ of the coordinates, which may be the coordinates themselves. If you want their expectation values, you have to integrate over phase space:

$$\bar{f} = \int d\Omega \; \rho(q,p) \; f(q,p) \; , \tag{28}$$

with $d\Omega = dq_1, \ldots, dp_n$.

The time evolution is described by

$$\bar{f}_t = \int d\Omega \; \rho(q,p) \; f(q_t,p_t) \; . \tag{29}$$

Because the Liouville measure $d\Omega$ is invariant under time translations, we can replace this also by

$$\bar{f}_t = \int d\Omega \; \rho(q_{-t},p_{-t}) \; f(q,p) \; . \tag{30}$$

Obviously, Eq. (29) corresponds to the Heisenberg picture and Eq. (30) to the Schrödinger one.

If you consider a sharp orbit, as we had done before, the density $\rho(q,p)$ degenerates to a δ-function. By the study of general $\rho(q,p)$, we will be led to the notions of ergodicity and mixing of a system. Ergodicity is a purely measure theoretic notion which is related but not identical to the topological statement that the orbits are dense on the energy shell. It is slightly stronger as it implies that almost all orbits are dense, while the converse is not true: density of almost all orbits does not imply ergodicity. We know already that we cannot hope that *all* orbits are dense, as we know that there are always periodic orbits interspersed between the dense ones. Even for very complicated systems it is rather simple to find individual periodic orbits.

It is amusing to see that we can find such periodic (therefore not dense) orbits even in the n-body problem with gravitational interactions, so let me discuss this in some detail. Consider the special case of n particles with potentials

$$V = -\sum_{i<k} \frac{m_i m_k}{r_{ik}}$$

moving in a plane. Because of this constraint, we can unify the x and y coordinates to complex variables

$$Z_j(t) = x_j(t) + iy_j(t) \; .$$

To simplify the problem further, let us assume that the time dependence of all $z_j(t)$ is given in terms of a single complex function $z(t)$ by

$$z_j(t) = z(t) \cdot \xi_j \; ,$$

with ξ_j being time-independent complex numbers. If you insert this into the equations of motion, you find

$$\ddot{z}(t) \cdot \xi_i = \sum_{j \neq i} m_j \frac{\xi_j - \xi_i}{|\xi_j - \xi_i|^3} \cdot \frac{z}{|z|^3} \; .$$

This, however, can be separated into two equations

$$-\omega^2 \xi_i = \sum_{j \neq i} m_j \frac{\xi_j - \xi_i}{|\xi_j - \xi_i|^3} \tag{31}$$

and

$$\ddot{z} = -\omega^2 \frac{z}{|z|^3} \; .$$

The latter is exactly the equation describing a planar Kepler problem, for which we of course know the solution, and thus the whole problem is reduced to finding complex numbers which satisfy Eq. (31). This is not very difficult; for three particles a solution is just an equilateral triangle. The whole motion consists of a rotation and periodic dilatation of this triangle, with a common frequency ω.

So we can find periodic orbits even in very complicated systems, but our point of view will be to forget about this as they have only measure zero. What will interest us are those things which are more likely to happen according to our measure.

As we have already said, the observables are functions $f(q,p)$ of the variables. In the following, we will assume that they are square integrable with respect to the Liouville measure, i.e. we assume

$$\int d\Omega \, |f(q,p)|^2 < \infty \; .$$

Instead of integrating here over the whole of phase space, we shall sometimes restrict the integration to some invariant region C of phase space, in particular to the energy shell. These square integrable functions over C form a Hilbert space,

$$\mathfrak{H}_C = \left\{ f : \int_C d\Omega \, |f(q,p)|^2 < \infty \right\} , \tag{32}$$

with

$$\langle g|f\rangle = \int_C d\Omega \; g^*(q,p) \; f(q,p) \; .$$

The time evolution is obviously given by a linear invertible operator $U(t)$,

$$|f_t\rangle = U(t)|f_0\rangle , \qquad f \in \mathfrak{H}_C \; .$$

Since we have assumed that C is invariant, the norm of f must be constant,

$$\langle f_t|f_t\rangle = \langle f_0|f_0\rangle \; ,$$

which implies that $U(t)$ is a unitary operator.

This simple observation allows us already to prove the so-called

Ergodic theorem *(von Neumann):* there exists the (strong) limit

$$\lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} dt \, |f_t\rangle$$

(the so-called Cesaro average), and the limiting vector is obtained by projecting onto those components of $|f\rangle$ which are invariant under time translations.

In order to prove this, we have to introduce the spectral representation of a unitary operator. You know that unitary matrices can always be diagonalized, and the diagonal elements are then all of unit absolute magnitude. In infinite dimensional Hilbert space there is an analogon: the Hilbert space $\mathfrak{H}_C$ can always be written as an orthogonal sum

$$\mathfrak{H}_C = \bigoplus_j \mathfrak{H}_j$$

of other Hilbert spaces, in each of which there exists a resolution of the identity $1 = \int d\mu_j(h) \, P_h$ such that $U(t)$ applied to P_h is just a phase factor e^{ith}. More specifically, let me call $|f_j\rangle$ the component of $|f\rangle$ in $\mathfrak{H}_j$, so that

$$\langle f|f\rangle = \sum_j \langle f_j|f_j\rangle$$

and (h is some function of q and p depending on the dynamics)

$$\langle f_j|f_j\rangle = \int d\mu_j(h) \, |f_j(h)|^2 \; . \tag{33}$$

Then, U(t) acts on $|f_j\rangle$ according to

$$[U(t)f_j](h) = e^{iht}f_j(h) \ .$$

Using this, we notice first that

$$\frac{1}{2T}\int_{-T}^{T} dt[U(t)f_j](h) = \frac{\sin hT}{hT} f_j(h) \ . \tag{34}$$

The function sin hT/hT is frequently encountered in physics, and one knows that it converges everywhere pointwise to a function $P_0(h)$ which is

$$P_0(h) = \lim_{T\to 0} \frac{\sin hT}{hT} = \begin{cases} 1 \text{ for } h = 0 \\ 0 \text{ otherwise} \end{cases} . \tag{35}$$

As a next step, let us define P_0 as the operator which projects onto time-invariant states, which means formally

$$\langle g_j|P_0 f_j\rangle = \int d\mu_j(h)\, g_j^*(h)\, P_0(h)\, f_j(h) \ . \tag{36}$$

Using Eqs. (33)-(36), we get

$$\lim_{T\to\infty} \left\| \frac{1}{2T}\int_{-T}^{T} dt|f_j\rangle - P_0|f_j\rangle \right\|^2 = \lim_{T\to\infty} \int d\mu_j(h)|f_j(h)|^2 \left\{\frac{\sin hT}{hT} - P_0(h)\right\}^2 .$$

Now $|f_j(h)|^2$ was integrable by assumption, and the curly bracket converges to zero pointwise. There is a famous theorem due to Lebesgue that in such a case you can interchange the integral and the limit, which means that the Cesaro limit for each $|f_j\rangle$ is equal to $P_0|f_j\rangle$, and consequently

$$\lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} dt|f_t\rangle = P_0|f\rangle, \qquad \text{Q.E.D.} \tag{37}$$

Notice that we do not know yet whether $P_0|f\rangle$ is different from zero. From Eq. (36) we see that this can only be the case if (at least) one of the $d\mu_j(h)$ has a δ-function at $h = 0$, corresponding to an eigenvector of U(t) with eigenvalue 1. Whether such an eigenvector exists is in general not known. However, if the subset C of phase space is compact, i.e. has a finite volume, the vector $|1\rangle$ which is defined by $f(q,p) \equiv 1$ is square integrable and time invariant, so P_0 projects at least onto this vector.

Let us make some remarks:

1) The existence of an invariant mean for functions which depend on time can also be shown in a more abstract way. Consider the time axis, i.e. the real numbers, as an additive Abelian group. There is a general theorem that Abelian groups are amenable, which means that you can always define an invariant mean $\eta(f)$ for functions depending on the group parameter. An invariant mean by definition is a positive linear functional, i.e.

$$\eta(f_1 + f_2) = \eta(f_1) + \eta(f_2) \ ,$$

$$\eta(\alpha f) = \alpha\eta(f) \ ,$$

$$f \geq 0 \Rightarrow \eta(f) \geq 0 \ ,$$

it is normalized to $\eta(1) = 1$, and it is invariant:

$$\eta(f_t) = \eta(f) \; .$$

The von Neumann ergodic theorem tells us that, for dynamical systems, the Cesaro limit exists and is one such mean. For functions which do not describe dynamical systems, such as, for example, $f_t \sim \sin(\ln t)$, the Cesaro average does not converge but the general theorem just mentioned tells us that some other average does exist.

2) In view of this, more important than an existence proof such as the one given above would be a uniqueness theorem. We would of course like the outcome of an averaging procedure not to depend on the detailed prescription how to take this average. For mechanical systems such a theorem can indeed be easily proven as follows:

As we have said, for each function $f(q_t,p_t)$ we can find a mean $\eta[f(q_t,p_t)]$. It defines an operator $\tilde{\eta}$ on the Hilbert space $\mathfrak{H}_C$ by $(\tilde{\eta}f)(q,p) = \eta[f(q_t,p_t)]$. Let me first show that $P_0\tilde{\eta} = P_0$. Take any two square integrable functions f and $g \in \mathfrak{H}_C$. Since $\langle g|P_0\tilde{\eta}f\rangle$ is an absolutely convergent integral, we can interchange the integration and the averaging so that

$$\langle g|P_0\tilde{\eta}f\rangle = \langle P_0 g|\tilde{\eta}f\rangle = \eta\Big(\langle P_0 g|U(t)f\rangle\Big) \; .$$

Here, we have also used that P_0 is hermilean and $f_t = U(t)f$. Using furthermore that $|P_0 g\rangle$ is time-independent, we can also write this as

$$\eta\Big(\langle U^*(t)P_0 g|f\rangle\Big) = \eta\Big(\langle P_0 g|f\rangle\Big) = \langle g|P_0 f\rangle \; .$$

Since this is true for all f and g in our Hilbert space, this means that $P_0\tilde{\eta} = P_0$. On the other hand, since $\tilde{\eta}|f\rangle$ is time-independent, one has $P_0\tilde{\eta}|f\rangle = \tilde{\eta}|f\rangle$ for all $|f\rangle$, which means $P_0\tilde{\eta} = \tilde{\eta}$. Taken together, these two results show that $\tilde{\eta} = P_0$, i.e. independently of how the time averaging is defined, η is given by the projection operator onto the invariant states.

We are now essentially in the position to define what is ergodic. Before we do so, let us recapitulate our definitions:

i) Previously, we had said that a *state* is given by a density $\rho(p,q)$ in phase space. More generally, a state is a positive measure $d\rho$ in phase space. I assume that I can restrict myself to a finite invariant region C in phase space on which $d\rho$ will be normalized to unity,

$$\int_C d\rho = 1 \; .$$

ii) *Observables* are square integrable functions $f(q,p)$ with respect to the Liouville measure. Square integrability is not a serious restriction but it allows us to consider f as a member of a Hilbert space with $\langle g|f\rangle = \int d\rho\, g^* f$. The expectation value of f is

$$\bar{f} = \int_C d\rho \cdot f = \langle 1|f\rangle \; .$$

iii) A state is called *invariant*, if all expectation values are independent of time, more specifically if for all f

$$\int d\rho\, |f_t|^2 = \int d\rho\, |f_0|^2 \; .$$

A specific example is the Liouville measure $d\rho = d\Omega$; another is the energy shell $d\rho = \int d\Omega\, \delta(H - E)$.

Now in a dynamical system it might be that if you start in a specific region of the energy shell, the orbits do not leave this region, while orbits from outside do not enter it. This would mean that there is an invariant measure $d\rho_1$ restricted to this region, and another measure $d\rho_2$ concentrated outside it. More generally, $\alpha d\rho_1 + (1 - \alpha)d\rho_2$ (with $0 < \alpha < 1$) would also be an invariant state. Such a situation is surely not what we would call ergodic, which suggests the

Definition: An invariant *state* $d\rho$ is called *ergodic* (or maximal invariant), if it cannot be decomposed into two invariant states $d\rho_1$ and $d\rho_2$, that is

$$d\rho \neq \alpha d\rho_1 + (1 - \alpha)\, d\rho_2\,; \quad \text{for any } 0 < \alpha < 1, \quad d\rho_1 \neq d\rho_2 \; . \tag{38}$$

Similarly, a *dynamical system* is called ergodic if the energy shell $\int d\Omega\, \delta(H - E)$ is an ergodic state. This agrees with our general understanding that for an ergodic system there should be no invariant subsets on the energy shell.

How is this notion of ergodicity related to the ergodic theorem? Assume that there exists a continuous function $f(q,p)$ which is constant in time but not constant over the support of a particular state $d\rho$. The function $\bar{f}$ defined by

$$\bar{f}(q,p) = \inf\left[1, |f(q,p)|\right]$$

is also invariant in time and -- perhaps after having rescaled f -- not identical to 1. In this case, we can write the state $d\rho$ as a sum of two invariant states

$$d\rho = \tfrac{1}{2}(1 + \bar{f})\, d\rho + \tfrac{1}{2}(1 - \bar{f})\, d\rho \;,$$

which means that $d\rho$ is not ergodic. In particular, taking $d\rho$ to be the energy shell, we see that a system is not ergodic if there exists a continuous function which is constant in time but not constant on the energy shell.

Let us see how this applies to our example of the two-dimensional harmonic oscillator, or equivalently to the motion on the two-dimensional torus. In this model, we had two action variables K_1, K_2 which were constant in time, and two angle variables θ_1, θ_2 whose time evolution was linear on the torus,

$$\theta_i(t) = \theta_i(0) + t\omega_i/2\pi \quad (\mathrm{mod}\ 1) \; . \tag{39}$$

First of all, since there are two independent constants of the motion, the system cannot be ergodic. But we can replace the energy shell by the region of phase space characterized by constant K_i, and can ask whether the state

$$d\rho = d\theta_1\, d\theta_2\,; \quad 0 < \theta_i < 1$$

is ergodic. We already know that for rationally dependent frequencies ω_i there exists a third constant of the motion K_3, so we expect $d\rho$ to be ergodic only if the frequencies are rationally independent. Let us verify this in detail. The unitary operator $U(t)$ describing the time development is, according to Eq. (39), given by

$$U(t)\; f(\theta_1,\theta_2) = e^{it(\omega_1 p_1 + \omega_2 p_2)}\; f(\theta_1,\theta_2) \;,$$

where the operators p_i are displacement operators

$$p_i = \frac{1}{i}\frac{\partial}{\partial\theta_i}\,; \quad i = 1,2\ .$$

The observables $f(\theta_1,\theta_2)$ have to be periodic,

$$f(\theta_1,\theta_2) = f(\theta_1 + n,\ \theta_2 + m)\,; \quad n,m \text{ integer}$$

since the θ_i are only defined modulo 1. The problem of finding the eigenvalues of $U(t)$ is now an elementary problem encountered, for example, in quantum mechanics, and we find that all eigenvalues can be written as

$$\exp\left[2\pi i(\omega_1 g_1 + \omega_2 g_2)t\right]\ ,$$

with g_1 and g_2 being integer. For $g_1 = g_2 = 0$, one has of course always a unit eigenvalue corresponding to the eigenfunction $f = \text{const}$. Whether there are further eigenfunctions with unit eigenvalue depends on the ω_i. If the ω_i are rationally dependent, one can find such $g_i \neq 0$ that $\omega_1 g_1 + \omega_2 g_2 = 0$, and the state $d\theta_1 d\theta_2$ is not ergodic. If the ω_i are, however, rationally independent, there are no such further eigenfunctions, and it is ergodic.

7. MIXING

In the above example the whole energy shell was not ergodic (for ω_i irrational) but only the states in which K_1 and K_2 are fixed. Besides this, the motion had another regularity which we could describe by saying that the states do not get mixed. By this we mean the following: assume that your state is at time $t = 0$ confined to some domain $\mathcal{D}$ of phase space. As time proceeds, this domain will be moved to any other place in phase space such that the *time-average* becomes uniformly distributed. But for any *fixed* time, the state will

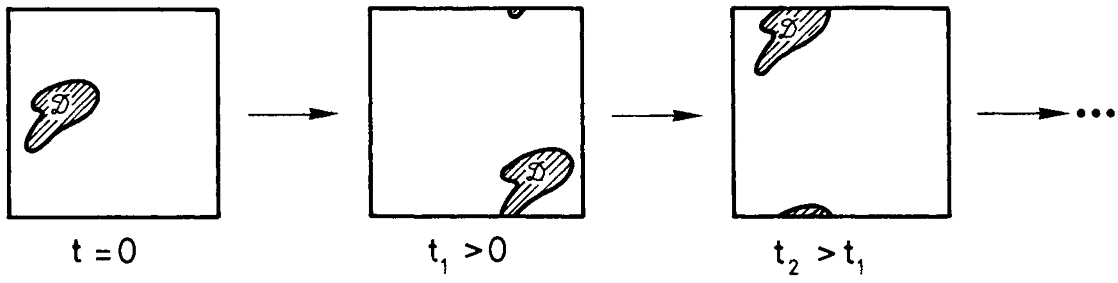

Fig. 8

not spread out and mix with other states, originally concentrated in other domains. It is, however, just the latter which would be more interesting for statistical physics, as it would tell us that after sufficiently long time all information referring to the initial state is lost.

To be more precise, assume that the measures $d\mu_i(h)$ appearing in the spectral representation of $U(t)$,

$$\langle g|U(t)f\rangle = \sum_j \int d\mu_j(h)\, e^{iht}\, g_j(h)\, f_j(h)\ , \tag{40}$$

are of the form

$$d\mu_j(h) = \begin{cases} \phi_1(h)\ dh + c_1\delta(h)\ dh & \text{for } j = 1 \\ \phi_j(h)\ dh & \text{otherwise} \end{cases} \tag{41}$$

with integrable functions $\phi_j(h)$. That is, U(t) has only one non-degenerate discrete eigenvalue at unity, while the rest of the spectrum is continuous. Now consider the limit of Eq. (40) when $t \to \infty$. Owing to the Riemann-Lebesgue lemma, the integrals

$$\int dh\ e^{iht}\ \phi_j(h)\, g_j(h)\, f_j(h)$$

vanish when $t \to \infty$. The only piece remaining then is the contribution of the δ-function, and it projects onto the only time-independent vector which must be the vector $|1\rangle$ constant over phase space. So we can write

$$\lim_{t\to\infty} \langle g|U(t)f\rangle = \langle g|1\rangle\ \langle 1|f\rangle \tag{42}$$

or simply

$$U(t) \underset{t\to\infty}{\rightharpoonup} P_0 \tag{43}$$

($\rightharpoonup$ means "converges weakly"). Another way to write Eq. (42) is

$$\int d\rho g f_t \underset{t\to\infty}{\longrightarrow} \int d\rho g \cdot \int d\rho f \ , \tag{44}$$

which shows that if you measure the product of two observables at sufficiently different times, your expectation value will be just the product of the two separate expectation values, i.e. all correlations will be lost if you wait long enough between two measurements. Systems which obey Eq. (42) are called *mixing systems* *). This notation is suggested by taking for f and g the characteristic functions of two domains A and B. Then

$$\int d\rho f = \mu(A)$$

is the measure of domain A, and similarly for B. Equation (42) reads in this case

$$\mu(A \cap B_t) \underset{t\to\infty}{\longrightarrow} \mu(A) \cdot \mu(B) \ ,$$

i.e. after a long enough lapse of time different parts of phase space will be completely mixed through.

As we had already said, the two-dimensional harmonic oscillator is not mixing. Formally, this follows from the fact that all eigenvalues of U(t) were discrete there, while mixing requires a continuous spectrum of U(t) except for a single non-degenerate unit eigenvalue.

8. AN EXAMPLE OF A MIXING SYSTEM (following R. Jost, private communication)

There are actually not many systems for which the mixing property has been proven exactly. One would hope to get it in a system resembling the two-dimensional torus, but with θ_1 and θ_2 varying over a region with oblique boundaries.

---

*) Sometimes called "strong mixing".

Here we shall study a similar example, namely the free motion on the pseudosphere $x_0^2 - x_1^2 - x_3^2 = 1$. In a cosmological setting, it is the motion of a particle in a de Sitter universe. More exactly, we shall work on a pseudosphere which is made periodic in a similar way to that in which the (θ_1, θ_2)-plane was made periodic by identifying θ_i with $\theta_i + 1$.

An advantage of this model is that we can use powerful group theoretic methods. Notice that we started with the study of differential equations, later used the techniques of Hilbert space, and now we shall see that even the theory of group representations becomes a useful tool.

We consider a three-dimensional Minkowski space with a scalar product

$$(x,y) = x_1 y_1 + x_2 y_2 - x_0 y_0 \ .$$

The pseudosphere is defined as the surface satisfying

$$(x,x) = -1$$

with $x_0 > 0$. A particle moving freely on it is described by the Lagrangian

$$L = \tfrac{1}{2}(\dot{x},\dot{x}) \ ,$$

where the dot denotes derivative with respect to a parameter called proper time s. If you derive from this the equations of motion, you have to remember the constraint $(x,x) = -1$, and thus you get

$$\ddot{x}_i = \lambda x_i \ , \quad i = 0,1,2$$

where λ is a Lagrangian multiplier. In the usual way you then find

$$\frac{1}{2}\frac{d}{ds}(\dot{x},\dot{x}) = (\dot{x},\ddot{x}) = \lambda(\dot{x},x) = \frac{1}{2}\lambda\frac{d}{ds}(x,x) = 0 \ .$$

Thus one has already three constants of the motion:

$$(x,x) = -1 \ , \tag{45}$$

$$(\dot{x},x) = 0 \ , \tag{46}$$

and

$$(\dot{x},\dot{x}) = 1 \ . \tag{47}$$

The last equation is a normalization for s. Equations (45)-(47) define our three-dimensional energy shell (or, strictly speaking, the "mass shell"). The phase-space measure on this energy shell is

$$d\Omega = d^3x \ d^3\dot{x} \ \delta[(\dot{x},x)] \ \delta[(x,x)+1] \ \delta[(\dot{x},\dot{x})-1] \ \theta(x_0) \ . \tag{48}$$

Further constants of the motion are obtained from angular momentum conservation. One checks immediately that

$$\ell_i = \varepsilon_{ikm} x_k \dot{x}_m \ , \quad i = 0,1,2$$

are three constants, but they are not independent. There is the relation

$$(\ell,\ell) = (x,x)\cdot(\dot{x},\dot{x}) = -1$$

among them, so you have a one-dimensional manifold left for the orbit. Its projection into phase space is simply given by

$$(\ell, x) = 0 \ ,$$

which is the intersection of the hyperboloid with a plane through the origin.

Equations (45)-(47) are obviously invariant under transformations belonging to the three-dimensional Lorentz group SO(2,1), and the same is true for the measure $d\Omega$ defined in Eq. (48). Even more, you can reach every point on the energy shell by applying a Lorentz transformation to the point $\{x_0, \dot{x}_0\} = \{(1,0,0),(0,1,0)\}$, i.e.

$$\{x, \dot{x}\} = \{M(1,0,0), \quad M(0,1,0)\} \ , \qquad (49)$$

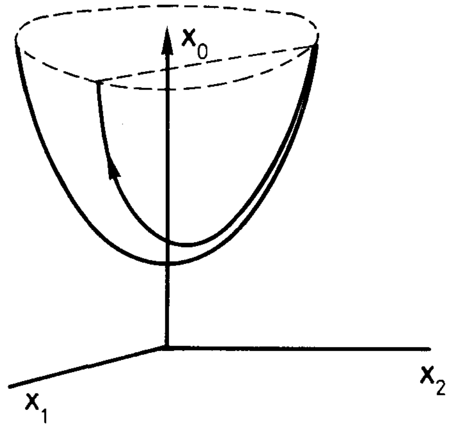

Fig. 9

for every $\{x,\dot{x}\} \in \Omega$ and with M denoting a special Lorentz transformation. This is perhaps most easily seen by looking at the inverse transformation: every point $\{x,\dot{x}\}$ can be translated into $\{(1,0,0),(0,1,0)\}$ by first transforming x into (1,0,0), and applying after this a rotation such that $\dot{x}$ goes over to (0,1,0). Here we use, of course, that x and $\dot{x}$ are orthogonal. In fact, Eq. (49) is an isomorphic mapping (i.e. a one-to-one correspondence) between the group SO(2,1) and the energy shell.

Since in course of the time development the energy shell is mapped onto itself, time evolution can be represented by a one-parameter subgroup of the Lorentz group SO(2,1). Take, for instance,

$$M(s) = \begin{pmatrix} \text{ch } s & \text{sh } s & 0 \\ \text{sh } s & \text{ch } s & 0 \\ 0 & 0 & 1 \end{pmatrix} . \qquad (50)$$

This is a particular orbit where you cut the hyperboloid with the plane $x_2 = 0$. Any other orbit can be obtained from this by an s-independent Lorentz transformation.

Another fact we shall use is that the group SO(2,1) is isomorphic to the group SL(2,R)/{1,-1}, that is the group of special linear transformations of real 2 × 2 matrices, divided by its centre {1,-1}. This correspondence is well known from the four-dimensional Lorentz group: take a matrix

$$\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \qquad (51)$$

in SL(2,R), i.e. such that $\alpha\delta - \beta\gamma = 1$. The corresponding element of SO(2,1) is defined by the transformation $\{x,\dot{x}\} \to \{x',\dot{x}'\}$ with

$$\begin{pmatrix} x_0' + x_2', & x_1' \\ x_1', & x_0' - x_2' \end{pmatrix} = \begin{pmatrix} \alpha\beta \\ \gamma\delta \end{pmatrix} \begin{pmatrix} x_0 + x_2, & x_1 \\ x_1, & x_0 - x_2 \end{pmatrix} \begin{pmatrix} \alpha\gamma \\ \beta\delta \end{pmatrix} . \qquad (52)$$

The fact that $SL(2,R)/\{1,-1\}$ is isomorphic to $SO(2,1)$, instead of $SL(2,R)$ itself, just means that you should not distinguish between the matrix in Eq. (51) and its negative - both give the same Lorentz transformation.

The problem we have considered so far is not yet the one we wanted: the orbit is a one-dimensional submanifold of the energy shell, so there is no ergodicity. Furthermore, the volume of the energy shell is infinite. Indeed, up to now the problem is very similar to the free motion of a particle in the plane. In the two-dimensional harmonic oscillator, we had a motion which looked like a free motion in the (θ_1,θ_2)-plane, except that we had taken the θ_i as real numbers *modulo 1*. In a more abstract language, the θ_i were not just real numbers but elements of the factor group R/Z, where Z is the (additive) group of integers. By doing this, we had distroyed rotation invariance, so that the constant K_3 [see Eq. (15)] was in general no longer a global constant, and the motion could become ergodic. Also, the volume of the energy shell became finite, as we restricted ourselves to $0 \leq \theta_i < 1$.

We want to do now essentially the same. Instead of considering the motion on the whole pseudosphere, we want to work on the factor

$$\Omega_0 = \Omega/Z = SO(2,1)/Z = SL(2,R)/\{1,-1\}/Z \ ,$$

where Z is now some discrete subgroup of Ω. In a less abstract language, this means that we shall identify two points $\{x,\dot{x}\}$ and $\{x',\dot{x}'\}$ on the energy shell, if $\{x',\dot{x}'\}$ can be written as

$$\{x',\dot{x}'\} = \{Mx, M\dot{x}\}$$

with any transformation $M \in Z$. Otherwise stated, we restrict ourselves to periodic observables which are invariant under transformations in Z.

A discrete subgroup Z of $SL(2,R)$ is simply given by the matrices

$$\begin{pmatrix}\alpha & \beta\\ \gamma & \delta\end{pmatrix}; \quad \alpha\delta - \gamma\beta = 1$$

with *integer* α, β, γ, and δ *). Integer matrices would not work if we had started with the motion on the sphere instead of the pseudosphere, where the discrete subgroups are the cristallographic groups.

Our next problem is now to construct a fundamental region, corresponding to the square $0 \leq \theta_1, \theta_2 \leq 1$ in our old example. By fundamental region we mean a region in Ω which contains exactly one point of each equivalence class $\{Zx, Z\dot{x}\}$, except on its boundary where to each point corresponds another point on the other side of the boundary. To find it, let us look in detail how the simplest elements of Z act on the pseudosphere -- or rather on its projection onto the (x_1,x_2)-plane. Besides the unit element, the basic elements of Z are

$$A = \begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}$$

and

$$B = \begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix} .$$

*) Notice that this is not an *invariant* subgroup, so Ω_0 is no factor *group*. This is however not necessary.

According to Eq. (52), the matrix B gives a reflection $(x_1,x_2) \to (-x_1,-x_2)$. So the fundamental region can be restricted to the upper half plane, part of its boundary being formed by the x_1-axis. On the other hand, take the curves h_+ and h_- defined as the images of the positive x_2-axis under the transformations

$$\begin{pmatrix} 1 & \pm\frac{1}{2} \\ 0 & 1 \end{pmatrix} ,$$

i.e.

$$h_\pm = \left\{ x_1' \,\middle|\, \begin{pmatrix} x_0' + x_2', & x_1' \\ x_1', & x_0' - x_2' \end{pmatrix} = \begin{pmatrix} 1 & \pm\frac{1}{2} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \sqrt{1+x_2^2} + x_2, & 0 \\ 0, & \sqrt{1+x_2^2} - x_2 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \pm\frac{1}{2} & 1 \end{pmatrix} \right\} .$$

They can also be written as

$$x_2' = \pm \frac{1}{4}\left(\frac{1}{x_1'} - 3x_1'\right); \quad x_2' > 0 .$$

Since

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & -\frac{1}{2} \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & \frac{1}{2} \\ 0 & 1 \end{pmatrix} ,$$

the matrix A maps the curve h_- onto h_+. A fundamental region is then given by the boundaries h_-, h_+, and the part of the x_1 axis between them. This is shown in Fig. 10, where also the action of the matrices A and B is indicated. Indeed, we have not shown that there is no other element in Z which maps any two interior points of this region one onto the other -- in which case the fundamental region would be part of the above. A more careful study shows that this does not happen.

When the orbit reaches the boundary of the fundamental region, you apply just the appropriate element of Z -- either A, A^{-1}, or B -- to bring it back into it, as indicated in Fig. 11. It turns out that this leads to a rather irregular pattern, so almost every orbit fills the fundamental region densely. Indeed, the orbit will even be dense on the three-dimensional reduced energy shell Ω_0, the projection into configuration space of which is

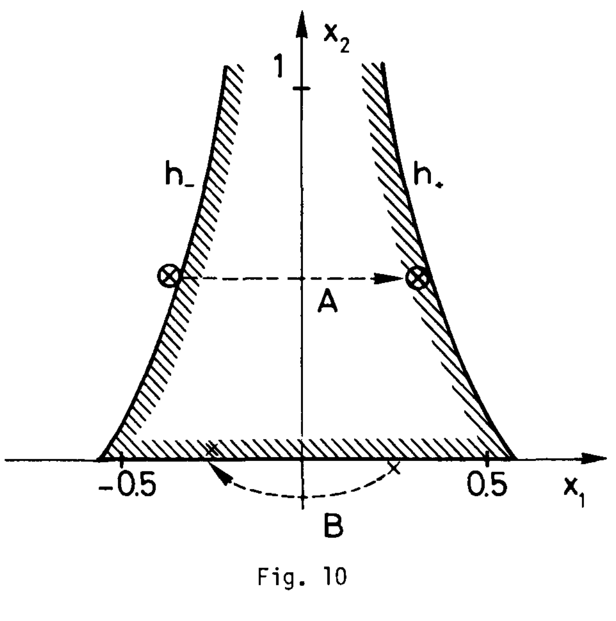

Fig. 10

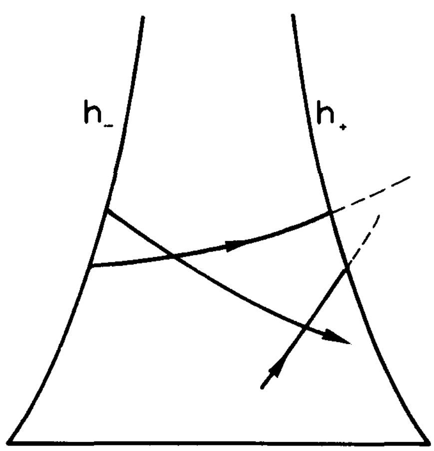

Fig. 11

just the fundamental region*). This will be a by-product of our result since mixing $\Rightarrow$ $\Rightarrow$ ergodicity $\Rightarrow$ almost every orbit is dense.

Another fact is that Ω_0 has a finite volume. It is very plausible, and it would not be too difficult to show it in detail**).

What we shall demonstrate is that the system has the mixing property discussed in the last section. To agree with the notation used there, let us again denote the (proper) time by t instead of s. As already mentioned, the motion is generated by a one-parameter subgroup of unitary transformations of SO(2,1), which we can write as

$$U(t) = e^{mt} \ .$$

Here, the antihermitean operator m is one of the three generators of the group. In analogy with the generators of the rotation group SO(3), the three generators m, m_+, and m_- of SO(2,1) form a Lie algebra with commutation relations

$$\begin{aligned} [m,m_\pm] &= \pm\, m_\pm \ , \\ [m_+,m_-] &= 2m \ . \end{aligned} \tag{53}$$

The only difference to the rotation group is a difference in sign, and the fact that $m_\pm$ are here defined as $m_1 \pm m_2$ without an i, so they are antihermitean. This slight difference is however crucial. We know that the rotation group has a discrete spectrum, while we want to show that the commutation relations (53) lead to a continuous spectrum, except for one discrete point.

As we said, m_+ is an antihermitean generator of the group, and the two elements m and m_+ alone form a Lie algebra as well as the pair $\{m,m_-\}$. Thus they generate two subgroups, the elements of which we can write

$$U_\pm(a,t) = e^{am_\pm}\, e^{tm} \ .$$

Using the commutation relations, we find indeed

$$U_\pm(a,t)\ U_\pm(a',t') = U_\pm(a + e^{\pm t'}a',\, t + t') \ . \tag{54}$$

If we set here a = 0, we get back the transformations U(t) describing time development

$$U(t) = U_+(0,t) = U_-(0,t) \ .$$

On the other hand, the transformations $\{U_+(a,o)\}$ and $\{U_-(a,o)\}$ together generate the whole group. Though they do not involve the generator m, this generator arises from the commutation relation $[m_+,m_-] = 2\,m$. A general element can e.g. be written as $U_+(a,o)\ U_-(a',o)\ U_+(a'',o)$.

---

*) Points in Ω_0 are essentially points in the fundamental region to which a direction is attached. The denseness of the orbit means that arbitrarily close to any point in the fundamental region is a part of the orbit the direction of which is arbitrarily close to any prescribed direction.

**) $\int d^3\dot{x}\ \delta[(\dot{x},x)]\ \delta[(\dot{x},\dot{x}) - 1] = F[(x,x)] < \infty$ and $F(-1)\cdot\int d^3x\ \delta[(x,x) + 1] < \infty$, where the integral extends over the fundamental region.

Let us now study the irreducible unitary representations of the group $\{U_+(a,t)\}$. [The same results hold, of course, also for the $U_-(a,t)$.] These representations are well known and indeed there are only two types of representations:

I : One-dimensional representations

$$U_+(a,t) = e^{i\lambda t}, \quad \lambda \in R \ .$$

These are of course not faithful, but they are faithful representations of the factor group $\{U_+(a,t)/U_+(a,o)\}$, which is a one-parameter Abelian group, isomorphic to the group of real numbers.

II: Infinite-dimensional representations which act as operators on square-integrable functions $\psi(x)$ according to

$$\left[U_+(a,t)\psi\right](x) = \exp\left(iae^x\right)\, \psi(x+t), \qquad \psi \in L_2(R, dx)$$

These are faithful, and one can explicitly check that the multiplication law agrees with Eq. (54).

The subgroup of time translations is represented in these two types simply by

$$\begin{aligned} \text{I:}\quad & U(t) = e^{i\lambda t} \\ \text{II:}\quad & \left[U(t)\psi\right](x) = \psi(x+t) \ . \end{aligned} \tag{55}$$

The Hilbert space of observables can then be written as a direct sum of two terms

$$\mathfrak{H} = \mathfrak{H}_I^+ \oplus \mathfrak{H}_{II}^+ \ ,$$

which are defined by the way $U_+(a,t)$ acts on them. Similarly, we can decompose it into

$$\mathfrak{H} = \mathfrak{H}_I^- \oplus \mathfrak{H}_{II}^- \ ,$$

according to the representations of $U_-(a,t)$. The time translations both on $\mathfrak{H}_{II}^+$ and on $\mathfrak{H}_{II}^-$ are simple shifts as shown in Eq. (55). This is exactly the same way as momentum is usually represented in wave mechanics, and we know that this spectrum is absolutely continuous and (Lebesgue-) integrable. Thus, according to what we found in the previous section, the only trouble for mixing could come from those functions which are in

$$\mathfrak{H}_I =: \quad \mathfrak{H}_I^+ \cap \mathfrak{H}_I^- \ .$$

But for any function $\psi \in \mathfrak{H}_I$, we have

$$U_\pm(a,0)\psi = \psi \ .$$

Since the whole group can be generated by the $U_\pm(a,o)$, we see that ψ is invariant under the whole group SO(2,1):

$$U(g)\psi = \psi \quad \text{for all} \quad g \in SO(2,1),\ \psi \in \mathfrak{H}_I \ .$$

Finally, because the group is isomorphic to the energy shell -- or because any point on the energy shell can be reached by some g -- this means that ψ is constant on the whole energy shell.

This argument shows that *if* there is a square-integrable eigenfunction ψ of U(t), it must be constant over phase space, and the mixing property holds. This argument is independent of whether we use the whole energy shell Ω or the part Ω_0 (whether we consider the Hilbert space of all observables or of those which are invariant under the discrete group Z). The reason for working on Ω_0 is from this point of view that it has a finite volume, so ψ = const *is* square integrable, and we do get mixing: if you start with any distribution and wait long enough, you will finally end up with an equilibrium distribution which is evenly distributed over phase space (the so-called "microcanonical distribution").

The above is the simplest known model which displays mixing. There is a physically more interesting but mathematically much more complicated model consisting of hard spheres confined to a box. In an unfortunately unpublished proof, Sinai has proven mixing for this.

One should, of course, not overestimate the importance of this for statistical mechanics, as quantum mechanics is relevant for real matter. It makes it however more plausible that the same will happen there, as the uncertainty principle implies that the states to start with are already more diffuse.

LITERATURE

General:

C.L. Siegel and J.K. Moser, Lectures on celestial mechanics (Springer, 1971).

R. Abraham, Foundations of mechanics (Benjamin, 1967).

V. Szebehely, Theory of orbits (Academic Press, 1967).

S. Sternberg, Celestial Mechanics, Parts I and II (Benjamin, 1969).

Sections 2-4:

J. Dieudonné, Eléments d'analyse (Gauthier-Villars, Paris, 1968).

E. Hlawka, Suppl. Acta Phys. Austriaca VII, 265 (1970).

H. Flanders, Differential forms (Academic Press, 1963).

S. Sternberg, Differential geometry (Prentice Hall, 1964).

R. Jost, Rev. Mod. Phys. 36, 572 (1964).

G. Mackey, Foundations of quantum mechanics (Benjamin, 1963).

J. Souriau, Structure des systèmes dynamiques (Dunod, 1970).

W. Thirring, Vorlesungen uber mathematische Physik T5: Punktmechanik; Inst. f. theoret. Physik, Universität Wien, 1974.

Sections 5-7:

V. Arnold, A. Avez: Problèmes ergodiques de la mécanique classique (Gauthier-Villars, Paris, 1968).

Y. Sinai, *in* The Boltzmann equation (eds. E.G.D. Cohen and W. Thirring) (Springer, 1973).

Section 8:

R. Jost, Suppl. to New Physics (Korean Physical Soc.) 7, 26 (1968).