12.
SOME THEOREMS OF ANALYTICAL MECHANICS OF GREAT IMPORTANCE FOR QUANTUM THEORY

"Alcuni Teoremi di Meccanica Analitica Importanti per la Teoria dei Quanti,"
Nuovo Cimento **25**, 271–285, (1923).

§ 1.– Ehrenfest's principle of adiabatics,[1] as is known, states that, if a mechanical system is in a quantum orbit and its mechanism, forces or constraints, are changed in an infinitely slow way, the system remains in a quantum preferred orbit during the whole transformation. In order that this principle have a definite sense, it is obviously necessary that the final orbit of the system only depend on the final mechanism and not on any particular sequence of intermediate mechanisms followed during the transformation. Burgers[2] has shown that this is really the case, at least for those kinds of systems which up to now have been considered almost exclusively only in quantum theory, i.e., for systems which either admit a complete separation of variables or at least can be represented by means of angular coordinates.[3] In this case, their motion can always be considered as resulting from periodic motions, generally having as many periods as the number of degrees of freedom or in case of degeneracy, a lower number. But just at this moment, the study of the simplest atomic structures having been accomplished,[4] some problems which do not admit angular coordinates are presenting themselves with increasing urgency, first of all the three-body problem which occurs in the study of the Helium molecule and the simplified form of the four-body problem which occurs in the study of the hydrogen molecule. As is known, all the efforts made up to now to reduce the study of these systems to that of systems with angular coordinates have failed. Thus it seems to be desirable to investigate whether and how far is it possible to attempt an extension of the principle of adiabatics to general systems, hoping that it can give some information which can help in the search for rules suitable to determine the preferred orbits of these more general systems.

§ 2.– First of all we should fix a classification of the systems to be studied.

[1] P. Ehrenfest, *Ann. d. Phys.* **51**, p. 327, 1916.
[2] Burgers. Versl. Akad. van Wetensch. Amsterdam, 25 November 1916. – Ann. d. Phys. 52, p. 195; 1917. – Phil. Mag. 33, p. 514; 1917.
[3] See for instance Sommerfeld. "Atombau und Spektrallinen, III ed. Zusatz 7.
[4] They are the hydrogen atom and its various perturbations (Zeeman effect, Stark effect, and Feinstruktur) and the ion of the hydrogen molecule H_2^+, when nucleus rotations are not present.

Therefore we turn to the usual representation of the state of the system by means of a point in a $2f$-dimensional space Γ, which has $q_1, q_2, \ldots, q_f$ as the general coordinates of the system and $p_1, p_2, \ldots, p_f$ as their conjugate momenta. We have, through each point of this space, a trajectory which corresponds to the motion of the system having its initial position and velocity determined by the point itself. We shall assume the forces and the constraints of the system to be time-independent and the forces are derivable from a potential so that an integral associated with conservation of energy exists. We call E hypersurfaces the hypersurfaces energy = constant; through each point of Γ passes one of the E's on which (because of the energy integral) the trajectory through the point is located.

The so called quasi-ergodic[5] mechanical systems enjoy the property that the trajectory generally passes infinitely close to every point of E, so to densely fill a $(2f-1)$-dimensional manifold.

However, it may be that our system, besides the energy integral, admits some other uniform integral independent of time. In this case the manifold filled by the trajectory will obviously have a lower number of dimensions. Thus let us assume that our system has all together m uniform first integrals independent of time,

$$\Phi_1(p,q) = c_1; \Phi_2 = c_2; \ldots; \Phi_m = c_m$$

where c_i are arbitrary constants. We shall have, through each point of Γ, a $(2f-m)$-dimensional manifold G, the intersection of the m hypersurfaces $\Phi_i = c_i$; and the trajectory passing through that point will be entirely contained in G.

In general it will not be possible to find a submanifold within G which contains the entire trajectory; on the contrary, based on the analogy with quasi-ergodic systems, we shall assume for our systems that in general all of G is densely filled by the trajectory, i.e., that the trajectory passes infinitely close to all the points of G. In this way, the trajectory will come out characterized, at least in its statistical elements, only by the knowledge of the values $\Phi_1, \Phi_2, \ldots, \Phi_m$ corresponding to it.

Therefore we call these values *characteristics of the trajectory.*

Then a quasi-ergodic system has only one characteristic, its energy.

A system with its energy independent of time, which admits separation of variables, has in general as many characteristics as degrees of freedom, corresponding to the f a constants of Jacobi's complete integral; a higher number can only occur in the case of degeneracy, i.e., when linear relations with integer coefficients between the fundamental frequencies exist. Let us consider, for instance, the motion of a point in a plane acted on by a force proportional to the distance from two orthogonal straight lines. If the two attraction coefficients are not commensurate, the point describes an open Lissajous curve in the plane.

And in the four-dimensional space Γ the representative point densely fills a two-dimensional surface G. Therefore the system has two characteristics; for them we can

[5]The author recently demonstrated that ordinary mechanical systems are in general quasi-ergodic, so that this is the most common case.

take the energies of the projections of the motion onto the two orthogonal straight lines. If instead the attraction coefficients are commensurate, the Lissajous curve degenerates into a closed curve and G becomes one-dimensional; this corresponds to three characteristics.

§ 3.– Now suppose we can change arbitrarily the forces, or the constraints of the system, i.e., what all together we shall call the *mechanism* of the system, with a happy naming due to P. Hertz,[6]

If we change the mechanism in an infinitely slow way, we have what is said to be an adiabatic transformation; and in § 5, we shall easily find a system of differential equations which shows how the characteristics of the system change when the guiding parameter μ of the mechanism changes adiabatically. But, as we have already mentioned, one can speak of the application of Ehrenfest's principle to a definite system only if the values that its characteristics take at the end of an adiabatic transformation only depend on the final mechanism and not on the intermediate workings passed through during the transformation. To study this question, we shall assume afterwards that the mechanism, rather than depending on only one parameter, depends on two parameters λ and μ. The dependence of the characteristics on λ and μ, instead of being a system of ordinary differential equations, will then be obviously expressed by a system of total differential equations; then the conditions for having the final values of the characteristics not depending on the path followed during the transformation in the λ, μ plane coincide with the integrability conditions for this system. We shall demonstrate that these conditions, for the quasi-ergodic system, are really satisfied. Instead, for the systems having more than one characteristic, they are not satisfied in general although important classes of exceptions exist.

§ 4.– Before passing to the study of adiabatic transformations it is convenient to consider some formulas which are useful for calculating the probability that, at any instant, the representative point is in G. Then for uniformity of notation, we call $x_1, x_2, ..., x_{2f}$ the coordinates of Γinstead of $p_1, \ldots, p_f$, $q_i, \ldots, q_f$ as above. Our problem can now be formulated in this way: calculate the probability that, at a certain instant, $x_1, x_2, \ldots, x_{2f-m}$ have values between x_1 and $x_1 + dx_1$, x_2 and $x_2 + dx_2, \ldots, x_{2f-m}$ and $x_{2f-m} + dx_{2f-m}$, while the remaining m x's obviously take the values necessary to maintain the representative point in G.

Now we know from statistical mechanics that Liouville's theorem states that the necessary condition for having a stationary distribution of the points in the space Γ is that their density in Γ should have a constant value on any G. A volume element of Γ can be written $dx_1\, dx_2 \ldots dx_{2f}$, but also, taking as new variables $x_1, x_2, \ldots, x_{2f-m}, \Phi_1, \Phi_2, \ldots, \Phi_m$ as $\frac{1}{D} dx_1\, dx_2 \ldots dx_{2f-m}\, d\Phi_1\, d\Phi_2 \ldots d\Phi_m$, where

[6]P. Herz. Ann. d. Phys. 33, p. 225, 537; 1910. Weber, Gans. Repertorium der Physik I, 2; 1916. We refer to these articles for any explanations regarding the statistical part of the text.

D is the functional determinant $\frac{\partial(\Phi_1,\ldots,\Phi_m)}{\partial(x_{2f-m+1},\ldots,x_{2f})}$. And since during the motion $d\Phi_1, d\Phi_2, \ldots, d\Phi_m$ obviously remain constant, the aforesaid volume element comes out to be proportional to $\frac{1}{D}dx_1 \ldots dx_{2f-m}$. Therefore also the desired probability is proportional to this expression; and since the total probability is obviously = 1, we finally find that the desired probability is given by

$$\frac{\frac{d\sigma}{D}}{\int \frac{d\sigma}{D}}, \tag{1}$$

where for short we set $d\sigma = dx_1\, dx_2 \ldots dx_{2f-m}$ and the integral is extended over all values of $x_1, x_2, \ldots, x_{2f-m}$ corresponding to the points of G.

Before leaving this subject, we still want to deduce a formula that will be useful in the case of quasi-ergodic systems. In this case G is a hypersurface, and we assume for the sake of simplicity that it should be closed, and such that it is intersected in only one point by each radius vector coming out from an origin within it. This because a more general approach, even though it is not essentially different, would cause rather complicated calculations. We refer the space Γ to polar coordinates, by characterizing each point in terms of its radius vector and the intersection of this vector with the unit hypersphere having the origin as center. We denote by H the only characteristic, i.e., the energy. In accordance with what was said above, the probability that at a certain instant the representative point lies within an element of solid angle $d\omega$ is proportional to the hypervolume comprised between the two hypersurfaces $H(x_1, \ldots, x_{2f}) = H$, and $H(x_1, \ldots, x_{2f}) = H + dH$, and the solid angle $d\omega$. This volume, except for the constant factor dH, is evidently $\frac{r^{2f-1}d\omega}{H_r}$, where $H_r = \frac{\partial H}{\partial r}$. Since the total probability must be =1, we find that the desired probability is given by

$$\frac{r^{2f-1}\frac{d\omega}{H_r}}{\int r^{2f-1}\frac{d\omega}{H_r}}, \tag{2}$$

where the integral is extended over the entire unit sphere.

§ 5.– In this section we assume the mechanism of our system to be a function of a parameter μ and we aim to study how the characteristics change when this parameter changes adiabatically. Since the mechanism depends on the parameter μ, in general also the characteristics $\Phi_1, \Phi_2, \ldots, \Phi_m$ will depend on μ, besides on the p's and q's. Then, if at a certain instant the parameter μ changes by $d\mu$, the characteristic Φ_i will correspondingly undergo the change $\frac{\partial\Phi_i}{\partial\mu}\partial\mu$. Since we are in the presence of an adiabatic change, to have the effective change of Φ_i, we must consider the average of this expression which, according to the results of the previous section, will be

$$d\mu \frac{\int \frac{\partial\Phi_i}{\partial\mu}\frac{d\sigma}{D}}{\int \frac{d\sigma}{D}} \tag{3}$$

which is a function only of μ and $\Phi_1, \ldots, \Phi_m$. The dependence of the characteristics on μ in an adiabatic transformation will then be expressed by the system of ordinary differential equations:

$$\frac{d\Phi_1}{d\mu} = \frac{\int \frac{\partial \Phi_1}{\partial \mu} \frac{d\sigma}{D}}{\int \frac{d\sigma}{D}}\,;\ \frac{d\Phi_2}{d\mu} = \frac{\int \frac{\partial \Phi_2}{\partial \mu} \frac{d\sigma}{D}}{\int \frac{d\sigma}{D}}\,;\ \ldots\,;\ \frac{d\Phi_m}{d\mu} = \frac{\int \frac{\partial \Phi_m}{\partial \mu} \frac{d\sigma}{D}}{\int \frac{d\sigma}{D}}\,. \tag{4}$$

If we know the values of the Φ's, for instance at $\mu = 0$, the integration of this system gives us their values for any μ. In the particular case of quasi-ergodic systems, the system (4) reduces to the single equation:

$$\frac{dH}{d\mu} = \frac{\int \frac{H_\mu}{H_r} r^{2f-1} d\omega}{\int \frac{d\omega}{H_r} r^{2f-1}}\,, \tag{5}$$

where $H_\mu = \dfrac{\partial H}{\partial \mu}$.

§ 6.– Now we want to study in which cases the final values of the characteristics are independent of the way followed in passing adiabatically from the initial mechanism to the final one. Therefore we shall represent the mechanism of the system as a function of two parameters, λ and μ . If one alters adiabatically these two parameters by $d\lambda$ and $d\mu$ respectively, the same conclusion of the preceding section shows that the corresponding change of the characteristic is:

$$d\Phi_i = \frac{\int \frac{\partial \Phi_i}{\partial \lambda} \frac{d\sigma}{D}}{\int \frac{d\sigma}{D}} d\lambda + \frac{\int \frac{\partial \Phi_i}{\partial \mu} \frac{d\sigma}{D}}{\int \frac{d\sigma}{D}} d\mu \qquad (i = 1, 2, \ldots, m)\,. \tag{6}$$

The coefficients of $d\lambda$ and $d\mu$ are evidently functions only of λ and $\Phi_1, \ldots, \Phi_m$, so the m equations (6) represent a system of total differential equations; if it turns out to be completely integrable, the final values of the Φ's will be effectively independent of the way followed during the transformation, otherwise not. We want to demonstrate that, in the case of quasi-ergodic systems, the condition of complete integrability is satisfied. In fact, for these systems, system (6) reduces to only a single total differential equation constructed analogously to (5)

$$dH = L d\lambda + M d\mu\,, \tag{7}$$

where we set

$$L = \frac{\int \frac{r^{2f-1} H_\lambda d\omega}{H_r}}{\int \frac{r^{2f-1} d\omega}{H_r}}\,, \quad M = \frac{\int \frac{r^{2f-1} H_\mu d\omega}{H_r}}{\int \frac{r^{2f-1} d\omega}{H_r}} \tag{8}$$

so L and M represent two functions of λ, μ and H. As we know, for obtaining the complete integrability of (7), it is necessary and sufficient that the total derivatives of L with respect to μ and of M with respect to λ be equal. Therefore it must be that

$$\frac{\partial L}{\partial \mu} + M \frac{\partial L}{\partial H} = \frac{\partial M}{\partial \lambda} + L \frac{\partial M}{\partial H}\,. \tag{9}$$

To demonstrate that this equality is really satisfied, let us begin by calculating the left hand side. Therefore, let us imagine giving independent variations δH and $\delta\mu$ to H and μ, leaving λ unchanged; we will then have

$$\delta L = \frac{\partial L}{\partial H}\delta H + \frac{\partial L}{\delta \mu}\delta\mu\,. \tag{10}$$

On the other hand, from the first of (8), we find that:

$$\delta L = \frac{1}{\left(\int \frac{r^{2f-1}d\omega}{H_r}\right)^2}\Bigg\{ \left(\int \frac{r^{2f-1}d\omega}{H_r}\right)\delta\int \frac{r^{2f-1}H_\lambda d\omega}{H_r} - \left(\int \frac{r^{2f-1}H_\lambda d\omega}{H_r}\right)\delta\int \frac{r^{2f-1}d\omega}{H_r^2}\Bigg\}\,. \tag{11}$$

In the calculation of the two variations of the integrals within the curly brackets, we can of course interchange the symbols δ and $\int$, since the limits of integration do not change since it is extended over the whole unit hypersphere. Thus we have:

$$\delta\int \frac{r^{2f-1}d\omega}{H_r} = (2f-1)\int \frac{r^{2f-2}\delta r d\omega}{H_r} - \int \frac{r^{2f-1}\delta H_r d\omega}{H_r^2}\,. \tag{12}$$

On the other hand, from the invariance on the unit sphere, one has:

$$\delta H = H_r\delta r + H_\mu\delta\mu$$

from which it follows that

$$\delta r = \frac{\delta H}{H_r} - \frac{H_\mu}{H_r}\delta\mu$$

and also

$$\delta H_r = H_{rr}\delta r + H_{r\mu}\delta\mu = \frac{H_{rr}}{H_r}\delta H + \left(H_{r\mu} - \frac{H_{rr}H_\mu}{H_r}\right)\delta\mu\,.$$

By substituting these expressions for δr, δH_r into (12), one finds:

$$\delta\int \frac{r^{2f-1}d\omega}{H_r} = \delta H\Bigg\{(2f-1)\int \frac{r^{2f-2}d\omega}{H_r^2} - \int \frac{r^{2f-1}H_{rr}d\omega}{H_r^3}\Bigg\} - \delta\mu\Bigg\{(2f-1)\int \frac{r^{2f-2}H_\mu d\omega}{H_r^2} + \int \frac{r^{2f-1}d\omega}{H_r^2}\left(H_{\mu r} - \frac{H_\mu H_{rr}}{H_r}\right)\Bigg\}\,.$$

In a similar way one finds:

$$\delta\int \frac{r^{2f-1}H_\lambda d\omega}{H_r} = \delta H\Bigg\{(2f-1)\int \frac{r^{2f-2}H_\lambda d\omega}{H_r^2} + \int \frac{r^{2f-1}H_{\lambda r}d\omega}{H_r^2} - \int \frac{r^{2f-1}H_\lambda H_{rr}}{H_r^3}d\omega\Bigg\} + \delta\mu\Bigg\{-(2f-1)\int \frac{r^{2f-2}H_\lambda H_\mu d\omega}{H_r^2} + \int \frac{r^{2f-1}d\omega}{H_r}\left(H_{\lambda r} - \frac{H_{\lambda r}H_\mu}{H_r}\right) - \int \frac{r^{2f-1}H_\lambda d\omega}{H_r^2}\left(H_{\mu r} - \frac{H_\mu H_{rr}}{H_r}\right)\Bigg\}\,.$$

By substituting these two last expressions into (11), and comparing with (10), one finally finds:

$$\frac{\partial L}{\partial H} = \frac{1}{\left(\int \frac{r^{2f-1}d\omega}{H_r}\right)^2}\Bigg[\left(\int \frac{r^{2f-1}d\omega}{H_r}\right)\Bigg\{(2f-1)\int \frac{r^{2f-2}H_\lambda d\omega}{H_r^2} + \int \frac{r^{2f-1}H_{\lambda r}d\omega}{H_r^2} - \int \frac{r^{2f-1}H_\lambda H_{rr}}{H_r^3}d\omega\Bigg\} - \left(\int \frac{r^{2f-1}H_\lambda d\omega}{H_r}\right)\Bigg\{(2f-1)\int \frac{r^{2f-2}d\omega}{H_r^2} - \int \frac{r^{2f-1}H_{rr}d\omega}{H_r^3}\Bigg\}\Bigg].$$

$$\frac{\partial L}{\partial \mu} = \frac{1}{\left(\int \frac{r^{2f-1}d\omega}{H_r}\right)^2}\Bigg[\left(\int \frac{r^{2f-1}d\omega}{H_r}\right)\Bigg\{-(2f-1)\int \frac{r^{2f-2}H_\lambda H_\mu d\omega}{H_r^2} + \int \frac{r^{2f-1}d\omega}{H_r}\left(H_{\lambda\mu} - \frac{H_{\lambda r}H_\mu}{H_r}\right) - \int \frac{r^{2f-1}H_\lambda d\omega}{H_r^2}\left(H_{\mu r} - \frac{H_\mu H_{rr}}{H_r}\right)\Bigg\} - \int \frac{r^{2f-1}H_\lambda d\omega}{H_r}\Bigg\{(2f-1)\int \frac{r^{2f-2}H_\mu d\omega}{H_r^2} + \int \frac{r^{2f-1}d\omega}{H_r^2}\left(H_{\mu r} - \frac{H_\mu H_{rr}}{H_r}\right)\Bigg\}\Bigg].$$

These two last equations, together with the second of (8), give us all the elements necessary to calculate the left hand side of (9). Once it has been calculated, it is immediate to recognize from its explicit expression that λ and μ appear symmetrically; thus (9) is verified.

Therefore we can conclude that for the quasi-ergodic systems, the value assumed by the energy at the end of an adiabatic transformation does not depend at all on the intermediate mechanisms of the transformation.

§ 7.– We now return to the systems with more than one characteristic. In order that the final characteristics be independent of the intermediate mechanisms of the transformation also for these systems, the conditions of complete integrability of the system (6) should be identically satisfied. But if through a calculation, obviously more complicated than that performed in the preceding section but not essentially different from it, we effectively build up these conditions, we find that in general they are not satisfied. Rather than report here this lengthy calculation, we prefer to show the argument through an example of a system with two characteristics. The example we choose is very similar to another one I have recently used in a note on the principle of adiabatics.

From an origin O, we draw in a plane two orthogonal axes x, y. Then we take in the first quadrant two points P, Q and draw the perpendiculars from them to axes (PA, PB, QC, QD). We shall assume that P lies inside the rectangle OCQD. Now let us suppose that inside the concave polygon APBDQCA a mass point is moving not acted on by forces and elastically bouncing off the walls of the polygon. The absolute values u, v of the components of the velocity of the point along the axes x, y clearly remain constant during the motion, so the system has two characteristics. Suppose we keep the point Q (of coordinates a, b) fixed but are free to move the

point P (of coordinates λ, μ). In this way we shall have accomplished a mechanical system with two characteristics u, v and depending on two parameters λ, μ. By easy arguments, analogous to the ones carried out in the note quoted above, one finds that, changing adiabatically the position of the point P, then u and v change following the rule:

$$d\log u = \frac{2\mu\, d\lambda}{ab - \lambda\mu}\,; \quad d\log v = \frac{2\lambda\, d\mu}{ab - \lambda\mu}\,.$$

Obviously neither of these two equations is completely integrable; therefore the values that u and v take at the end of a transformation also depend on the path followed by the point P. Therefore in general, it is not possible to apply Ehrenfest's principle to systems with more characteristics.

§ 8.– However, some important classes of exceptions to this rule exist. We aim to study them in this section. The first one, and also the most important, is that of the systems with angular coordinates. Of these systems, according to Burgers' theorems, we not only know that Ehrenfest's principle can be applied (in the sense that it leads in any case to definite final conditions) but also that for them the aforesaid principle proves be *verified, by experience* as a logical consequence of Sommerfeld's conditions which are supported by all the theory and the experience made on the hydrogen atom.

Another remarkable class of exceptions to the conclusions of § 7 is the following: Let us assume that of the m characteristics of our system only one, the energy, depends explicitly on the parameters λ, μ of the mechanism. I say that for these systems, at the end of every adiabatic transformation, the energy takes a value independent of the intermediate mechanisms, while the other characteristics in fact remain unchanged.

The fact that all the characteristics apart from the energy remain unchanged comes out clearly from the circumstance that, since they do not contain the parameters explicitly, they remain unchanged in all the elementary processes of the transformation; the same conclusion can be drawn from system (6) since, if Φ_i is one of these characteristics, one has by hypothesis $\dfrac{\partial\Phi_i}{\partial\lambda} = \dfrac{\partial\Phi_i}{\partial\mu} = 0$.

For demonstrating that the final value of the energy does not depend on the path followed during the transformation in the plane of λ, μ, one could put forward a consideration analogous to that of § 6. But it is easier to remark that, on the basis of the hypothesis, by means of a canonical transformation *independent of the parameters*, one can try to transform the characteristics independent of the parameters into coordinates of Γ. After this, the considerations of § 6 can be repeated word for word and the constant characteristics simply stand for constant parameters. Systems of this kind occur very frequently in applications; for instance, of this kind are all the systems which have, as the only uniform integrals besides the energy (and not dependent on the energy), some integral of the conservation of

momentum, or angular momentum, since the latter are always independent of the parameters of the mechanism.

§ 9.– As regards a possible application of these remarks to the theory of quanta, we remark the following: On the basis of our conclusions, the possibility of an extension of Ehrenfest's principle is ruled out, save for the exceptions mentioned above. Instead, for quasi-ergodic systems, or the exceptions studied in § 8, such an application is not a priori ruled out, though obviously it is not possible now to foresee if experience will confirm its results. All the same, one might try to see if, going in this direction, some useful information on the rules for the determination of the quantum orbits of the systems without angular coordinates might be obtained. Of course, Ehrenfest's principle by itself, even if experience should confirm it in this more general application, is not sufficient for the determination of such rules. It only allows us, when we know the selected orbits of a certain system, to deduce the orbits for all the systems which can be obtained from it by means of an adiabatic transformation. Therefore perhaps it might be useful, apart from the complexity of calculations, for finding the quantitative relations between the spark spectra, for instance of the alkaline metals, and the arc spectra of the noble gases. In fact, the systems which emit these spectra only differ in the charge of the nucleus and thus can be easily transformed from one into the other.

Göttingen, April 1923.