ON THE ELIMINATION OF THE IRRATIONALITY CONDITION AND BIRKHOFF'S CONCEPT OF COMPLETE STABILITY

By Jurgen Moser

1. Introduction

Consider a real system of differential equations

$$\begin{aligned} \dot{x}_\nu &= f_\nu(x, y) = \alpha_\nu y_\nu + \cdots \\ \dot{y}_\nu &= g_\nu(x, y) = -\alpha_\nu x_\nu + \cdots \end{aligned} \qquad (\nu = 1, 2, \cdots, n) \tag{1}$$

in the neighborhood of the equilibrium $x = y = 0$. If α_ν are real numbers, satisfying $\alpha_\nu \pm \alpha_\mu \neq 0$ then the solutions of the linearized system represent n independent oscillations which can be represented in terms of $e^{i\alpha_\nu t}$, $e^{-i\alpha_\nu t}$. In particular, all solutions of the linearized system are bounded and almost periodic. It is well known that in general the behavior of the nonlinear system is quite different from that of the linearized system, as solutions might approach the equilibrium for $t \to \infty$, for instance, or also leave every neighborhood of it. To investigate this behavior one has devised many methods to describe the solutions of (1) by infinite series of exponentials. It is known from Birkhoff and Cherry ([1], [2], [3]) that these series become trigonometrical series (i.e., the exponents are purely imaginary) if one restricts (1) to Hamiltonian systems, i.e., systems for which

$$\begin{aligned} f_\nu &= \frac{\partial H}{\partial y_\nu} \\ g_\nu &= \frac{\partial H}{\partial x_\nu} \end{aligned} \tag{2}$$

where $H(x, y)$ is a real function. In the following we are going to make the same restriction. This result shows that the solutions of Hamiltonian systems have a strong stability behavior which Birkhoff described by the term "complete stability".

The expansion methods lead to the well known difficulty of "small divisors" and the requirement that the numbers α_ν are independent over the rationals. This assumption entails (a) the divergence of the series involved, and (b) the exclusion of a dense set of α vectors by infinitely many conditions

$$(g, \alpha) = \sum_{\nu=1}^{n} g_\nu \alpha_\nu \neq 0 \quad \text{for} \quad g \neq 0 \tag{3}$$

where $g = (g_1 \cdots, g_n)$ represents any vector with integer coefficients. The first of these two handicaps is less important since one knows that divergent series can be used for asymptotic representations of solutions. The second (i.e., (b)), however, makes the method completely useless for applications, since condition (3) cannot be verified in case one knows the coefficients of (1) only to a certain degree of accuracy. It is true that the set of excluded α is countable and therefore is exceptional; however, the asymptotic series will not permit a uniform description for varying values of α_ν. It is the aim of this paper to develop a method which leads to a description of the solutions near the equilibrium by asymptotic series where irrationality conditions like (3) are discarded and replaced by only finitely many conditions

$$(g, \alpha) \neq 0 \quad \text{for} \quad |g| \leqq 6 \tag{4}$$

where $|g| = \sum_{\nu=1}^{n} |g_\nu|$. The precise formulation of the result will be stated in Theorem 1 below. The elimination of infinitely many conditions of (3) is possible at the expense of losing a set of solutions which have an arbitrarily small measure compared to the solutions admitted. In other words, instead of excluding systems of differential equations violating (3), we suggest the exclusion of a minority of solutions and the description of the majority by asymptotic series.

This idea is closely related to and originates partly in a paper of A. N. Kolmogorov on the stability of conditionally periodic solutions of analytic Hamiltonian systems [4, 5]. (It seems, however, that a complete proof of his statement has not been published.) While Kolmogorov's work refers to perturbation of an integrable Hamiltonian system the present paper deals with the solutions of a fixed Hamiltonian system in the neighborhood of an equilibrium. The same methods can be used to discuss solutions of a Hamiltonian system near a periodic solution. In this form this paper can be considered an extension to higher degrees of freedom of a previous investigation [6] in which the slowness of escape of a solution from a periodic solution was studied.

The detailed proofs are omitted and will be published elsewhere.

2. Formulation of the result

In the following the system (1) is assumed to be of Hamiltonian character with the Hamiltonian

$$H = \sum_{\nu=1}^{n} \frac{\alpha_\nu}{2} (x_\nu^2 + y_\nu^2) + \cdots$$

which is a real analytic function near $x = y = 0$. We introduce the n-vector $w = (w_1, \cdots, w_n)$ by

$$w_\nu = x_\nu^2 + y_\nu^2$$

and $\alpha = (\alpha_1, \cdots, \alpha_n)$. It is well known (see [1] or [2]) that under condition (4) one can introduce new coordinates x_ν', y_ν' in such a way that up to terms of order 7 the Hamiltonian depends on w only. Therefore, it is no loss of generality to assume that H has the form

$$H = \frac{1}{2} (\alpha, w) + \frac{1}{4} (w, \beta w) + \cdots \tag{5}$$

where $\beta = (\beta_{\nu\mu})$ is a symmetric n by n matrix. For the following the condition

$$(6) \qquad \det \beta \neq 0$$

will be of fundamental importance, which allows only nonlinear problems. We denote by x the $2n$-vector

$$x = (x_1, x_2, \cdots, x_{2n})$$

where

$$x_{\nu+n} = y_\nu \quad \text{for} \quad \nu = 1, 2, \cdots, n$$

and

$$|x| = \left(\sum_{\nu=1}^{2n} x_\nu^2\right)^{1/2}$$

$$dx = \prod_{\nu=1}^{2n} dx_\nu .$$

For sufficiently small ϵ we will study the solutions in the sphere

$$|x| < \epsilon$$

the volume of which is $\epsilon^{2n}\tau$ (τ denoting the volume of the unit sphere). For any measurable S in $|x| < \epsilon$ we will denote by

$$m(S) = \int_S dx$$

its $2n$-dimensional measure. The quotient

$$\frac{m(S)}{\epsilon^{2n}\tau}$$

will be called the "relative measure" of S.

THEOREM 1. *Let* (1) *be a Hamiltonian system satisfying the finitely many conditions* (4) *and* (6), *and* δ *in* $0 < \delta < 1$ *and an integer* $s > 6$ *be given. Then there exists a set* S *of relative measure* $< \delta$ *such that the solutions* $x(t)$ *with* $x(0) \in S$ *admit the asymptotic description*

$$(7) \qquad \left| x(t) - \sum_{\mu=1}^{m} a_\mu e^{i\gamma_\mu t} \right| < c\epsilon^s |t| \quad \text{for} \quad \epsilon^s |t| < c^{-2}, \quad \epsilon < \epsilon_0(s).$$

Here a_μ *are complex,* γ_μ *real constants depending on* x. *The constant* c *is independent of* ϵ *and the particular solution* $x(t)$ *but may depend on* δ, s, S *and the system.*

Estimate (7) allows a long time description of the majority of the solutions: For $|t| < \epsilon^{-s/2}$ the error in (7) becomes

$$< c\epsilon^{s/2}.$$

This property agrees with the concept of "complete stability" of Birkhoff [1], who required the above description for *all* solutions. For this purpose he was forced to impose the infinitely many conditions (3).

It is easy to show that for $n > 1$ the exclusion of a set of small measure is necessary. Let $h = (h_1, \cdots, h_n) \neq 0$ be a vector of non-negative integers, and assume

$$(h, \alpha) = (h, \beta h) = 0$$

which does not contradict the above conditions (4) and (6). The counterexample is provided by the Hamiltonian

$$H = \tfrac{1}{2}(\alpha, w) + \tfrac{1}{4}(w, \beta w) + \operatorname{Im} \prod_{\nu=1}^{n} (x_\nu + ix_{\nu+n})^{h_\nu}.$$

One verifies that on the invariant surface $\operatorname{Im} \prod_{\nu=1}^{n} (x_\nu + ix_{\nu+n})^{h_\nu} = 0$ there are solutions satisfying

$$w = h\lambda(t)$$

with the *a* scalar $\lambda(t)$:

$$\dot{\lambda} = 2 \prod_{\nu=1}^{n} h_\nu^{h_\nu/2} \lambda^{|h|/2}.$$

These solutions do not satisfy an inequality of the form (7) if $s > |h| - 2$. They form, however, a set of small relative measure.

3. Description of the method

The proof of the above statement is based on the following approach: We are trying to introduce new variables

$$\xi = (\xi_1, \xi_2, \cdots, \xi_{2n}), \quad \xi_{\nu+n} = \eta_\nu$$

by a canonical coordinate transformation

$$\xi = \varphi(x) \tag{8}$$

in such a way that the new system has a Hamiltonian $\Gamma = \Gamma(\xi_1, \cdots, \xi_{2n})$ of a particular simple form, namely, that it depends on

$$\omega_\nu = \xi_\nu^2 + \eta_\nu^2 \quad (\nu = 1, 2, \cdots, n)$$

only. The solutions of such a system are given by trigonometrical expressions.

This method is identical with Birkhoff's in [1]. The difference lies in the fact that Birkhoff used power series expansions for $\varphi(x)$ while we suggest rational series expansion with denominators. In both cases one is led to divergent series; therefore, all series are considered as formal series. We consider the given system as a perturbation of the nonlinear system with the Hamiltonian

$$F = \tfrac{1}{2}(\alpha, w) + \tfrac{1}{4}(w, \beta w) \tag{9}$$

while in the previous approaches the given system was compared with the linear system.

To specify the group of formal transformations (8) which we admit we introduce for each g with integer components the term

$$z_g = \frac{|g|}{(g, \alpha) + (g, \beta w)}$$

which represents a small divisor. In particular, if $(g, \alpha) = 0$ the expression z_g^{-1} is a homogeneous quadratic polynomial in x. The series $\varphi(x)$ will be represented as power series in x and the infinitely many variables z_g. Let $k = (k_1, \cdots, k_{2n})$ be a vector of integers which are non-negative, which we indicate by $k \geqq 0$. We define

$$| k | = \sum_{\nu=1}^{2n} k_\nu .$$

By m we denote a vector with infinitely many integer components $m_g \geqq 0$ (g lattice points) and assume

$$| m | = \sum_g m_g < \infty$$

so that only finitely many m_g are different from zero. We define

$$x^k z^m = \prod_{\nu=1}^{2n} x_\nu^{k_\nu} \prod_g z_g^{m_g}$$

and associate with such a term the order

$$\sigma = | k | - 2 | m |.$$

By $T[\sigma]$ we denote a finite[1] sum of terms

$$\sum a_{km} x^k z^m$$

with

$$\begin{aligned} k \geqq 0, \quad m \geqq 0, \\ | k | - 2 | m | = \sigma \geqq 2 \\ | m | \leqq A (\sigma - 1) - 1 \end{aligned} \tag{11}$$

where A is a fixed number $A \geqq 1$. The class of series $\varphi(x)$ of the form

$$\varphi(x) = x + \sum_{\sigma \geq 2} T[\sigma]$$

will be denoted by $\mathfrak{F}_A$. It will not cause confusion if the formal coordinate transformation (8) will also be said to belong to $\mathfrak{F}_A$ if all components $\varphi_\nu(x)$ belong to $\mathfrak{F}_A$.

These series have several properties in common with power series.

[1] I.e., we assume that $| g |$ is bounded by some γ_σ.

LEMMA 1. *The formal transformations* (8) *which belong to* $\mathfrak{F}_A$ *form a group under formal substitution and inversion.*

It is to be noted that the representation of a series φ in $\mathfrak{F}_A$ is not unique, in so far as a term

$$z_g = \frac{|g|}{(g, \alpha) + (g, \beta w)}$$

with $(g, \alpha) \neq 0$ can be expanded into a power series. This could be avoided by introducing only those z_g for which $(g, \alpha) = 0$. We are not going to do this since one would introduce a nonuniformity in the validity of the series this way. The above statement holds for any representation.

We will call such a transformation (8) canonical if formally the identity

$$\sum_{\nu=1}^{n} d\xi_{\nu+n} \wedge d\xi_\nu = \sum_{\nu=1}^{n} dx_{\nu+n} \wedge dx_\nu$$

holds, which is equivalent to the familiar formulae for the Lagrange brackets.

THEOREM 2. *For a Hamiltonian system* (1) *there exists a formal canonical transformation* (8) *of class* $\mathfrak{F}_4$ *such that the new Hamiltonian* Γ *depends on* $\omega_1, \omega_2, \cdots, \omega_n$ *only and is of the form*

$$\begin{aligned}\Gamma &= \tfrac{1}{2}(\alpha, \omega) + \tfrac{1}{4}(\omega, \beta\omega) + \cdots \\ &= F(\omega) + \sum_{\sigma=2|k|-2|m|} \omega^k \zeta^m\end{aligned}$$

where

$$\zeta_g = \frac{|g|}{(g, \alpha) + (g, \beta\omega)}.$$

For the indices one can prove the inequalities $|m| \leqq 4\sigma$; $|g| \leqq 9\sigma$ *which impose a bound on the number of terms of order* σ.

Remark. Writing the transformation (8) of Theorem 2 in inverted form

$$x = \psi(\xi) = \xi + \sum_{\sigma=|k|-2|m|\geq 2} k_{km}\xi^k\zeta^m$$

one can show that the terms of order σ also satisfy

$$|m| \leqq 4\sigma$$
$$|g| \leqq 9\sigma.$$

The proof follows the usual pattern of comparison of coefficients where the small divisors $((g, \alpha) + (g, \beta\omega))/|g|$ are introduced as new variables ζ_g^{-1} each time they occur.

4. Asymptotic description of the solutions

In order to utilize Theorem 2 for an asymptotic description of the solutions of (1) we break off the series ψ and Γ at terms of order $\sigma = s + 3$, $\sigma = s$ respectively; the series so obtained are denoted by $\psi^{(s+3)}$, $\Gamma^{(s)}$. The solutions of

$$\dot{\xi}_\nu = \Gamma^{(s)}_{\xi_{\nu+n}}, \quad \dot{\xi}_{\nu+n} = -\Gamma^{(s)}_{\xi_\nu}$$

have n integrals, $\omega_1, \cdots, \omega_n$ and the solutions are explicitly

$$\xi_\nu + i\xi_{\nu+n} = \omega_\nu^{\frac{1}{2}} e^{-it2\Gamma_{\omega_\nu}} \tag{13}$$

provided $(g, \alpha) + (g, \beta\omega) \neq 0$ where t can be replaced by $t - t_\nu$. In order to provide that these expressions are not too small we require now

$$|(g, \alpha + \beta\omega)| \geqq \delta' |\omega| |g|^{-n-1} \tag{14}$$

for all g,[2] or

$$|\zeta_g| \leqq \frac{|g|^{n+2}}{|\omega|\delta'} \tag{14'}$$

which is a requirement on ω. As is well known from the theory of diophantine approximations, this condition (14) will exclude from $|\omega| < \epsilon$[3] only a set of a relative measure which is small for small δ'. One computes for the relative measure

$$K_n \frac{\operatorname{Max} |\beta^n_{\nu\mu}|}{|\det \beta|} \delta' < \delta$$

if δ' is sufficiently small.

Condition (14′) guarantees that a term of order σ in

$$|\xi^k \zeta^m| \leqq M\epsilon^{|k|-2m} = M\epsilon^\sigma$$

tends to zero like ϵ^σ as $\epsilon \to 0$.

From this fact it can be derived that

$$|x(t) - \psi^{(s+3)}(\xi)| \leqq k\epsilon^s t \tag{15}$$

where the ξ_ν have to be replaced by the exponentials (13).

This statement contains Theorem 1 and even allows to specify the exceptional set as a set of tori.

It is to be noted that the condition (14) excludes in particular all periodic solutions near the equilibrium, which were constructed in [3]. They form, in general, only a set of measure zero.

5. Nature of the expansion

The above series will certainly diverge in general if s tends to infinity. Therefore, one can only expect asymptotic statements. These asymptotic expansions, however, cannot be considered significant, if the forms of order σ grow too rapidly as $\sigma \to \infty$. In particular, since the terms of order σ contained in a func-

[2] It would suffice to make this requirement for $|g| \leq 9\sigma$.

[3] $|\omega| = \sum_{\nu=1}^{n} \omega_\nu$.

tion of $\mathfrak{F}_4$ depend on many variables z_g it is important to have an estimate on the number of variables z_g and the number of terms which have the order σ.

From the estimate $|g| \leqq 9\sigma = G$ follows that the number of variables z_g in a term of order σ is at most

$$(2G + 1)^n \leqq (18\sigma + 1)^n.$$

The number of power products $x^k z^m$ with $|k| - 2|m| = \sigma$ therefore will be

$$< (c\sigma)^{4n\sigma}.$$

This shows that *the quality of the asymptotic series gets worse as the number n of degrees of freedom increases.*

It is also to be mentioned that the Theorem 1 and 2 are valid if the Hamiltonian H is only $(s + 1)$ times differentiable near $x = 0$, from which we conclude that the asymptotic expansion is valid to the same order up to which H is differentiable. This seems significant from the point of view that most Hamiltonian systems considered in celestial mechanics are analytic.

Finally all estimates (7) and (15) above can be made uniform over systems for which there exists a constant such that the terms of order s in H can be estimated by M^s and which satisfy

$$(4') \qquad |(g, \alpha)| > M^{-1} \quad \text{for} \quad |g| \leqq 6$$

$$(6') \qquad \begin{gathered} |\det \beta| > M^{-1} \\ |\beta_{\nu\mu}| < M. \end{gathered}$$

As the α vary only the exceptional set of excluded initial values varies.

6. Conclusion

We are concerned with the theory of analytic Hamiltonian systems near an equilibrium with purely imaginary eigenvalues. It is known that the formal expansions devised by Birkhoff may diverge, and that the divergence is even the general situation. A precise statement in this direction was formulated and proven by C. L. Siegel [7]. This result is related to the nonexistence of integrals near the equilibrium, which are not functions of H.

This negative result leaves the question of the stability of the equilibrium unanswered. The well known stability criterion of Dirichlet and Liapounoff which requires the construction of a positive definite integral, however, is recognized to be useless unless H is positive definite. There are many examples in applications where H is not definite. The most famous one, perhaps, refers to the periodic solutions of the 3 body problems where the three mass points move on an equilateral triangle. (See [8], in particular §477.)

The aim of this note is to show that under general assumption the majority of the solutions have a strong stability behavior which is described by asymptotic series. A stronger structure of the solutions should, maybe, not even be expected.

The same method can with few formal modifications be applied to the stability problem of periodic solutions in which case the Hamiltonian is usually not a definite function near the periodic reference orbit.

MASSACHUSETTS INSTITUTE OF TECHNOLOGY, CAMBRIDGE

BIBLIOGRAPHY

[1] G. D. BIRKHOFF, *Dynamical Systems*, Amer. Math. Soc. Coll. Publ., Vol. 9, New York, 1927.

[2] T. M. CHERRY, *On the solutions of Hamiltonian systems of differential equations in the neighborhood of a singular point*, Proc. London Math. Soc., Ser. 2, Vol. 27 (1926), 151–170.

[3] T. M. CHERRY, *On periodic solutions of Hamiltonian systems of differential equations*, Phil. Trans. Royal Soc. London, Ser. A, Vol. 227 (1927), 137–221.

[4] A. N. KOLMOGOROV, *Théorie générale des systèmes dynamiques et mécanique classique*, Proc. Int. Congress of Math., Amsterdam, 1 (1954) 315–333.

[5] A. N. KOLMOGOROV, Dokl. Akad. Nauk, 98 (1954), 527–530.

[6] J. MOSER, *Stabilitätsverhalten kanonischer Differentialgleichungssysteme*, Nachr. Akad. Wiss. Göttingen, Math.-Phys. Abt., 6 (1955), 87–120.

[7] C. L. SIEGEL, *Über die Existenz einer Normalform anaytischer Hamiltonscher Differentialgleichungen in der Nähe einer Gleichgewichtslösung.*, Math. Ann., 128 (1954), 144–170.

[8] A. WINTNER, The analytic foundations of celestial mechanics, Princeton, 1947.