Annals of Mathematics, **157** (2003), 557–574

Convergence or generic divergence of the Birkhoff normal form

By Ricardo Pérez-Marco

Abstract

We prove that the Birkhoff normal form of hamiltonian flows at a nonresonant singular point with given quadratic part is always convergent or generically divergent. The same result is proved for the normalization mapping and any formal first integral.

Introduction

In this article we study analytic (**R** or **C**-analytic) hamiltonian flows

$$\begin{aligned} \dot{x}_k &= +\frac{\partial H}{\partial y_k}, \\ \dot{y}_k &= -\frac{\partial H}{\partial x_k}, \end{aligned}$$

where $x_k, y_k \in \mathbf{C}$ (resp. $\mathbf{R}$), $k = 1, 2, \ldots n$, and H is an analytic hamiltonian with power series expansion at 0 beginning with quadratic terms (so that 0 is a singular point of the analytic vector field). We shall restrict our attention to those H having nonresonant quadratic parts: If $(\lambda_1, \ldots, \lambda_{2n})$ are the eigenvalues of the matrix JQ where $\frac{1}{2}(x,y)Q(x,y)^t$ is the quadratic part of H with $\lambda_{n+1} = -\lambda_1, \ldots \lambda_{2n} = -\lambda_n$, there is no relation of the form

$$i_1\lambda_1 + \ldots + i_n\lambda_n = 0$$

with integral coefficients $i_1, \ldots, i_n$ except for the trivial case $i_1 = \ldots = i_n = 0$. Due to some confusion in some of the literature on the distinction between the problem of convergence of the Birkhoff normal form and Birkhoff transformation, we start with a brief historical overview.

The normal form of a hamiltonian flow near a singular point has been studied since the origins of mechanics. The long time evolution of the system near the equilibrium position is better controlled in variables oscullating those of the normal form that corresponds to a completely integrable system. This idea is at the base of many computations in celestial mechanics. Its importance, both practical and theoretical, cannot be overestimated. One can consult the reference memoir "Les méthodes nouvelles de la mécanique céleste" by H. Poincaré ([Po]) to get an idea of the central place that the perturbative approach played in the XIXth century. A highlight was the discovery of Neptune by U. Le Verrier (and J. C. Adams) based on perturbative analysis of the orbit of Uranus. The theory of perturbations can be traced back to the origins of mechanics in the "Principia" of I. Newton (as noted by F. R. Moulton in [Mou] in the historical notes at the ends of Chapters IX and X).

Assuming that the eigenvalues of the quadratic part of H present no resonances, we have a simple, formal, normal form. This result goes back to C. E. Delaunay [De] and A. Lindstedt [Li] (also see [Po], [Si2]). Nowadays this normal form is named after Birkhoff. The Birkhoff normal form is the starting point of most of the studies of stability near the equilibrium point: the first studies by E. T. Whittaker [Wh], T. M. Cherry [Ch], G. D. Birkhoff [Bi1], [Bi2], and C. L. Siegel [Si1], [Si2], K.A.M. theory ([Ko], [Ar], [Mo]), Nehoroshev's diffusion estimates [Ne], … .

The dream of an analytic conjugacy to the normal form (uniform on the quadratic part of H) was quickly dissipated after the work of H. Poincaré ([Po, vol.I, chapitre V]). Poincaré's divergence theorem is the starting point of his difficult proof of the nonexistence of nontrivial local first integrals in the three body problem for some particular configuration of masses.

Research then focused on understanding the divergence of the conjugation mapping (normalization mapping) with a fixed nonresonant quadratic part for H. The normal form is unique. The normalization mapping is not unique, but appropriate normalizations determine it uniquely. Different results showed with increasing strength that the normalization mapping was generically divergent. We refer to the book of C. L. Siegel and J. Moser ([Si-Mo, Chap. 30]) for an overview. The strongest result on divergence was proved by Siegel in 1954 ([Si2]) and showed the generic divergence of the normalization, the quadratic part of the hamiltonian being fixed but otherwise arbitrary. A. D. Bruno [Br] and H. Rüssman [Ru2], [Ru3] proved the convergence of the normalization when the Birkhoff normal form for the hamiltonian is quadratic and the eigenvalues satisfy Bruno's arithmetic condition (other proofs can be found in [El2], [E-V]).

Despite this progress, the most natural question remains untouched. The question is not the convergence or divergence of the normalizing map, but actually the convergence or divergence of the Birkhoff normal form itself. If in the first place the Birkhoff normal form is diverging, then there is no point in trying to conjugate to the normal form. Also, in this case, the normalization is necessarily diverging.

Very surprisingly, there seems to be no significant result on this fundamental question. It appears to be a very hard question. The author first learned about it from H. Eliasson. The references in the literature are scarce. H. Eliasson points out in the introduction of his article [El1] that

> "*...if the normal form itself is convergent or divergent is not known...*",

and he points out in [El2],

> "*...Generically* (...) *the formal transformation is divergent* (*if the normal form itself also is generically divergent is not known*)."

These are the only citations in the literature that the author is aware of (despite the title of [It] what is really proved there is the convergence of the normalization). On the other hand, one frequently finds in some literature the wrong claim "Birkhoff normal form is generically diverging" in place of the "Birkhoff transformation is generically diverging"... .

More surprisingly, not a single example is known of an analytic hamiltonian having a divergent Birkhoff normal form. The main result in this article is that the existence of a single example with divergent Birkhoff normal form forces generic divergence. To be more precise we need to introduce the notion of a pluripolar subset of $\mathbf{C}^n$. This is the $-\infty$ locus of plurisubharmonic functions in $\mathbf{C}^n$. This notion generalizes to higher dimension the notion of a polar set in dimension 1 (that is, a set with logarithmic capacity 0). An important property, as in dimension 1, is that a pluripolar set $E \subset \mathbf{C}^n$ is Lebesgue and Baire thin; i.e., E has zero Lebesgue measure and is of the first category (a countable union of nowhere dense sets). A pluripolar set is small in all senses. For example, in dimension 1 it has Hausdorff dimension 0. In higher dimension n there are even smooth arcs which are not pluripolar.

In order to talk about generic properties we define a natural Baire space. We consider the Fréchet space $\mathcal{H}$ of Hamiltonians holomorphic in the unit ball, endowed with the topology of uniform convergence on compact subsets of the unit ball. We choose this natural complete metric space as a working setting. The proof goes through other richer or poorer topologies. The meaning of the result is then different. A generic set contains a dense G_δ. Thus if the topology is richer, then it is easier to be open, so to be G_δ but harder to be dense. The opposite happens for poorer topologies. Similar results hold for $\mathbf{C}$-analytic and $\mathbf{R}$-analytic hamiltonians.

We can now state:

THEOREM 1. *Consider the subspace of $\mathcal{H}_Q \subset \mathcal{H}$ of analytic hamiltonians*

$$H = \sum_{l=2}^{+\infty} H_l$$

with fixed nonresonant quadratic part H_2 given by the symmetric matrix Q.

If there exists one hamiltonian $H_0 \in \mathcal{H}_Q$ with divergent Birkhoff normal form (resp. normalization), then a generic hamiltonian in $\mathcal{H}_Q$ has divergent Birkhoff normal form (resp. normalization).

More precisely, all hamiltonians in any complex (resp. real) affine finite-dimensional subspace V of $\mathcal{H}_Q$ have a convergent Birkhoff normal form (or normalization), or only an exceptional pluripolar (resp. of Lebesgue measure 0) subset of hamiltonians in V has this property.

Observe that the second scenario holds for all affine subspaces containing H_0. The result obtained in the real analytic case is stronger than stated. When V is a real-dimensional affine line, the exceptional set has zero capacity in the complexification of V. So the exceptional set has even Hausdorff dimension zero. A popular particular case worth pointing out is the case of the perturbed hamiltonian $H_0 + \varepsilon H_1$ where both H_0 and H_1 are independent of ε and H_1 is a perturbation of order 3 or more. Then these hamiltonians are all integrable, or the set of values of $\varepsilon \in \mathbf{C}$ yielding integrable hamiltonians has 0 capacity in $\mathbf{C}$.

The important issue that remains unsettled is thus the existence of hamiltonians with diverging Birkhoff normal form for any nonresonant quadratic part. The prevalent opinion among specialists is that there is generic divergence for all nonresonant quadratic parts. This feeling is probably motivated by the divergence results on the normalization, which, it is worth noting, are independent of the quadratic part. The author knows no reason against the convergence of Birkhoff normal forms, in particular when the eigenvalues of the quadratic part of H enjoy good arithmetic properties. If we fix the quadratic part of the hamiltonian, the answer may depend on the arithmetic of its eigenvalues.[1]

On the other hand, by standard methods of small divisors, it is not difficult to exhibit hamiltonians with diverging normalizations using Liouville eigenvalues for the quadratic part. Combining this construction with the previous theorem, one recovers with a simple proof Siegel's result ([Si2]) on the generic divergence of the normalization mapping for some fixed quadratic parts.

Note that fixing the quadratic part of the hamiltonian makes the problem much harder, not allowing one to take any advantage of the arithmetic of the eigenvalues. One can find in the literature results without fixing the quadratic part ([Po, vol. I, Ch. V], [Koz]). One may ask about the reason for studying

[1]After the appearance of the preprint version of this paper, L. Stolovitch announced the proof of this result in [Sto3]. Unfortunately the manuscript of L. Stolovitch is erroneous, as I pointed out to the author. After thinking more about the problem, I saw that there may be reasons to indicate that the Birkhoff normal form could be diverging independently of the arithmetic nature of the quadratic part. Also A. Jorba has shown to me numerical evidence that points to the divergence of the Birkhoff normal form.

hamiltonians with fixed quadratic part. Note that for systems with particles, the masses enter directly into the quadratic part of the hamiltonian through the kinetic energy. Thus if one, for example, wants to show the nonintegrability of a given system *with given masses* then families of hamiltonians with fixed quadratic part arise naturally. One can cite at this juncture the strict criticism of A. Wintner of Poincaré's proof of nonintegrability of the three body problem ([Wi, p. 241]):

> *...Poincaré has established a result which concerns the nonexistence of additional integrals Nevertheless, his result, as well as its formal refinement obtained by Painlevé, is not satisfactory* (...) *In fact, these negative results do not deal with the case of fixed, but rather with unspecified, values of the masses m_i* (...) *Clearly, these assumptions in themselves do not allow any dynamical interpretation, since a dynamical system is determined by a fixed set of positive numbers m_i...*

Without sharing this strict view, one cannot deny some point in Wintner's criticism.

The problem of convergence of the Birkhoff normal form arises also in geometric quantification, in the so-called EBK, for Einstein-Brillouin-Keller, quantification. Bohr-Sommerfeld semi-classical quantification provides a set of rules to obtain the energy levels of the quantification of some classical system. A. Einstein [Ein] studied, in a somewhat forgotten article, which systems admit a Bohr-Sommerfeld quantification procedure. He pointed out the link to complete integrability. Later J. B. Keller [Kel] rediscovered the Einstein article and extended the procedure to non-completely integrable systems. For the hamiltonian systems considered here, if K denotes the Birkhoff normal form, the discrete energy levels of the quantified system should be approximated by

$$E(l_1, \dots, l_n) = K((l_1 + 1/2)h, \dots, (l_n + 1/2)h)$$

where h is Planck's constant and $l_1, \dots, l_n$ are positive integers. This corresponds to the quantification of the actions. Thus the above implicitly assumes that the Birkhoff normal form is convergent and has infinite radius of convergence. In practice the normal form must be truncated at some appropriate order, and the general interpretation should be in terms of asymptotic expansions. But the convergence of the Birkhoff normal form may be the correct condition to ensure EBK quantification. For more on this topic we refer the reader to M. C. Gutzwiller's book [Gu].

We prove a second theorem on the divergence of first integrals. The classical approach to integrability of hamiltonian systems is based on *first integrals*. A first integral P is a convergent power series in the $2n$ variables $x_1, \dots, y_n$

such that

$$\{P, H\} = 0$$

where the Poisson bracket is defined by

$$\{P, H\} = \sum_{k=1}^{n} \left(\frac{\partial P}{\partial x_k} \frac{\partial H}{\partial y_k} - \frac{\partial P}{\partial y_k} \frac{\partial H}{\partial x_k} \right) .$$

The equation $\{P, H\} = 0$ is equivalent to $\dot{P} = 0$, that is to the conservation of P. By E. Noether's theorem, symmetries of the hamiltonian generate first integrals. Two first integrals, P_1 and P_2, are in involution (or functionally independent) if their Poisson bracket vanishes

$$\{P_1, P_2\} = 0 .$$

At a nonsingular point of the hamiltonian, Liouville's theorem shows that the hamiltonian system is integrable by quadratures if there exist n first integrals in involution. The case of a nonresonant singular point as considered here is more involved. It has been shown by H. Rüssman [Ru1] for $n = 2$ and in general by J. Vey [Ve] and H. Ito [It] that the existence of n first integrals in involution forces the convergence of the normalization to Birkhoff normal form (H. Eliasson settled the analogue of Vey's theorem in the C^∞ case [El1], [El3]). L. Stolovitch gave a unified approach to Bruno's theorem cited before and Vey's and Ito's theorems ([St1], [St2]). Once all symmetries of a system have been used to find first integrals in involution, the natural question is are there any others. Multiple approaches to nonintegrability have been developed starting from H. Poincaré. We refer to [Koz] for an overview of classical methods. R. de la Llave has proved that Poincaré's conditions are necessary and sufficient for uniform integrability ([Ll]; see also the paper by G. Gallavotti [Ga]). We refer to [Mor] for an account of recent methods of S. L. Ziglin, J. Morales Ruiz and J.-P. Ramis. In the smooth nonanalytic setting we refer to the work of R. C. Robinson ([Rob]).

It is natural to define the *degree of integrability* of a hamiltonian as the maximal number $1 \leq \iota(H) \leq n$ of functionally independent first integrals in involution. When the normalization is convergent, $\iota(H) = n$, so the study of convergent first integrals can be seen as a refinement of the study of the convergence of the normalization.

THEOREM 2. *In the space $\mathcal{H}_Q$, with a hamiltonian $H_0 \in \mathcal{H}_Q$, there is a generic hamiltonian $H \in \mathcal{H}_Q$, such that*

$$\iota(H) \leq \iota(H_0) .$$

More precisely, let P be a universal formal first integral. In any complex (*resp.*

real) affine finite-dimensional subspace V of $\mathcal{H}_Q$ all hamiltonians $H \in V$ have converging $P(H)$, or only an exceptional pluripolar (resp. Lebesgue measure zero) set in V has this property.

We give in Section 1 a precise definition of a universal formal first integral. This theorem reduces the proof of the generic divergence of a given formal first integral in a family of hamiltonians, to the divergence for one hamiltonian. Also, given a family V, the minimum degree of integrability in V,

$$\iota_V = \min_{H \in V} \iota(H)$$

is attained for a generic $H \in V$.

The families V in Theorems 1 and 2 can be more general than finite-dimensional affine subspaces. The same proof gives the results for example when V is parametrized polynomially by $\mathbf{C}^m$. It is interesting to note how in these theorems the complexification of the problem sheds new light on the real analytic case.

The main idea of this article has also been applied to other problems of small divisors ([PM1], [PM2]).

Acknowledgements. The author is grateful to A. Chenciner, H. Eliasson and A. Jorba for conversations on the subject. A. Jorba showed to the author numerical evidence on the divergence of Birkhoff normal forms. H. Eliasson provided corrections to the original version. Many thanks also to N. Sibony, E. Bedford and N. Levenberg for pointing out that the correct smallness condition to be used is pluripolarity and not Γ-capacity 0 which was used in the first version. The author thanks the referee for his careful reading, suggestions and corrections.

1. The Birkhoff normal form and first integrals

a) *The Birkhoff normal form.* We review briefly in this section the construction of the Birkhoff normal form following [Si-Mo]. We need to pay particular attention to the polynomial dependence of the transformation and Birkhoff normal form on the original coefficients of the hamiltonian function. More precisely, it is important for our purposes to keep track of the degrees of the polynomial dependence. We use the sub-index notation for partial derivatives.

There is an analytic hamiltonian ($\mathbf{R}$ or $\mathbf{C}$ analytic)

$$H(x,y) = \sum_{l=2}^{+\infty} H_l(x,y)$$

where H_l is the homogeneous part of degree l in the real or complex variables $x_1, \ldots, x_n, y_1, \ldots, y_n$. We can assume, by means of a preliminary linear change of variables, that H_2 is already in diagonal form ([Bi1, §III.7]):

$$H_2(x,y)=\sum_{k=1}^{n}\lambda_k x_k y_k \ .$$

We look for a simpler normal form of the system

$$\begin{aligned}\dot{x}_k &= H_{y_k},\\ \dot{y}_k &= -H_{x_k}\end{aligned}$$

and consider symplectic transformations that leave unchanged the hamiltonian character of the system of differential equations. The new variables (ξ,η) are related to the old ones (x,y) by the canonical transformation

$$\begin{aligned}x_k &= \varphi_k(\xi,\eta)=\xi_k+\sum_{l=2}^{+\infty}\varphi_{kl}(\xi,\eta),\\ y_k &= \psi_k(\xi,\eta)=\eta_k+\sum_{l=2}^{+\infty}\psi_{kl}(\xi,\eta)\end{aligned}$$

where φ_{kl} and ψ_{kl} are the homogeneous parts of degree l. These canonical transformations are defined by a generating function

$$v(x,\eta)=\sum_{l=2}^{+\infty}v_l(x,\eta)$$

where v_l is the homogeneous part of degree l, and $v_2(x,\eta)=\sum_{k=1}^{+\infty}x_k\eta_k$. Then the canonical transformation is defined by the equations

$$\begin{aligned}\xi_k &= v_{\eta_k}(x,\eta)=x_k+\sum_{l=3}^{+\infty}v_{l,\eta_k}(x,\eta),\\ y_k &= v_{x_k}(x,\eta)=\eta_k+\sum_{l=3}^{+\infty}v_{l,x_k}(x,\eta).\end{aligned}$$

Thus

$$\begin{aligned}x_k &=\xi_k-\textstyle\sum_{l=3}^{+\infty}v_{l\eta_k}(\varphi(\xi,\eta),\eta),\\ y_k &=\eta_k+\textstyle\sum_{l=3}^{+\infty}v_{lx_k}(\varphi(\xi,\eta),\eta),\end{aligned}$$

and

$$\begin{aligned}\varphi_{kl}(\xi,\eta) &=-v_{l+1,\eta_k}(\xi,\eta)-\left\{\textstyle\sum_{j=3}^{l}v_{j,\eta_k}(\varphi(\xi,\eta),\eta)\right\}_l,\\ \psi_{kl}(\xi,\eta) &=v_{l+1,x_k}(\xi,\eta)+\left\{\textstyle\sum_{j=3}^{l}v_{j,x_k}(\varphi(\xi,\eta),\eta)\right\}_l,\end{aligned}$$

where $\{.\}_l$ indicates the l homogeneous part of the expression within brackets. From these expressions we have that the coefficients of φ_{kl} and ψ_{kl} are polynomials with integer coefficients on the coefficients of $v_3,\dots,v_l,v_{l+1}$.

To each coefficient of v_l we assign a degree $l-2$ (next, we will choose a canonical transformation so that the coefficients of the v_l's are polynomials of the coefficients of H of degree $l-2$ at most). By induction, we show that the degree of φ_{kl} is at most $l-1$. For $l=2$ it is clear. Then by induction, the degree of the coefficients of the homogeneous part of degree l of a homogeneous monomial

$$\prod_{k=1}^{n}(\varphi_k(\xi,\eta))^{\alpha_k}\eta_k^{\beta_k}$$

of total degree j ($\sum\alpha_k+\sum\beta_k=j$) is at most $l-j$. Thus the degree of the coefficients of the homogeneous part of degree l of

$$v_{j,\eta_k}(\varphi(\xi,\eta),\eta)$$

is at most $(j-2)+(l-j+1)=l-1$, and this finishes the induction. The same discussion applies to ψ and the coefficient ψ_{kl} has degree $l-1$.

Now the canonical transformation generated by v transforms the differential system into

$$\begin{aligned}\dot{\xi}_k &= K_{\eta_k}\\ \dot{\eta}_k &= -K_{\xi_k}\end{aligned}$$

where

$$K(\xi,\eta)=\sum_{l=2}^{+\infty}H_l(\varphi(\xi,\eta),\psi(\xi,\eta))=\sum_{l=2}^{+\infty}K_l(\xi,\eta)$$

and K_l is the l-homogeneous part.

Our aim is to construct a canonical transformation which gives a hamiltonian K only depending on power series of the products $\omega_k=\xi_k\eta_k$. The coefficients of v are constructed by induction on the degree l of the homogeneous part. Assume that the choices for $v_3,\dots,v_{l-1}$ have been done so that the new hamiltonian has monomials of degree $\le l-1$ only depending on the ω_k's. We consider a monomial of degree l

$$P=\prod_{k=1}^{n}\xi_k^{\alpha_k}\eta_k^{\beta_k}\ .$$

We want to choose the coefficient γ of P in $v_l(\varphi(\xi,\eta),\eta)$ such that the new hamiltonian does not contain the monomial P. Note that

$$K_l(\xi,\eta)=\sum_{k=1}^{n}\lambda_k\left(\xi_k v_{lx_k}(\varphi(\xi,\eta),\eta)-\eta_k v_{l\eta_k}(\varphi(\xi,\eta),\eta)\right)+A$$

where the first term comes from the expansion of $H_2(\phi(\xi,\eta),\psi(\xi,\eta))$ and the second term A collects everything coming from higher order. The coefficients in the expression A are polynomials in the coefficients of $v_3,\dots,v_{l-1}$ and linear functions in the coefficients of $H_3,\dots H_l$.

By induction we prove at the same time that the coefficients of v_l are polynomials of degree $l-2$ of the coefficients of $H_3, \ldots, H_l$, and also the coefficients of K_l are polynomials of degree $l-2$ of the coefficients of $H_3, \ldots, H_l$. Assuming the induction hypothesis, we have as before that the right-hand side in the above formula for K_l is a polynomial of degree $\leq l-2$ of the coefficients of $H_3, \ldots, H_l$.

Now we have

$$\sum_{k=1}^{n} \lambda_k (\xi_k P_{\xi_k} - \eta_k P_{\eta_k}) = \left(\sum_{k=1}^{n} \lambda_k (\alpha_k - \beta_k) \right) P.$$

Thus if $\lambda = \sum_{k=1}^{n} \lambda_k (\alpha_k - \beta_k) \neq 0$, and

$$\gamma = -\frac{1}{\lambda} \{A\}_P$$

(where brackets indicate that we extract the P monomial) the new hamiltonian will not contain the monomial P. Note that by the nonresonance condition, $\lambda = 0$ only happens when $\alpha_k = \beta_k$ for $k = 1, \ldots, n$. In that way we determine all coefficients of v_l except those of the monomials which are a product of ω_k's. Note also that by induction these coefficients are polynomials on the coefficients of $H_3, \ldots, H_l$ of degree $\leq l-2$.

In order to determine the coefficients of v_l for the remaining monomials one takes the normalization that no product of powers of ω_k's appears in

$$\Phi = \sum_{k=1}^{n} (\xi_k y_k - \eta_k x_k)$$

when expressed in (ξ, η) variables. One checks that this determines uniquely v and thus the canonical transformation that transforms the hamiltonian into its Birkhoff normal form. When H is real analytic, it is easy to check ([Si-Mo]) that the previous construction yields a real formal canonical transformation and a real Birkhoff normal form. We summarize this discussion in the following proposition.

Proposition 1.1. *Given a hamiltonion flow*

$$\begin{aligned} \dot{x}_k &= H_{y_k} \\ \dot{y}_k &= -H_{x_k} \end{aligned}$$

with $H(x, y) = \sum_{l=2}^{+\infty} H_l(x, y)$ *and with nonresonant quadratic part* H_2*, there exists a unique formal canonical transformation defined by a formal generating series*

$$v(x, \eta) = \sum_{l=2}^{+\infty} v_l(x, \eta)$$

such that in the new variables (ξ_k, η_k) *the differential system takes the form*

$$\begin{aligned} \dot{\xi}_k &= K_{\eta_k} \\ \dot{\eta}_k &= -K_{\xi_k} \end{aligned}$$

where the new hamiltonian K *is a formal power series in the products* $\omega_k = \xi_k \eta_k$, *and the expression*

$$\Phi = \sum_{k=1}^{n} (\xi_k y_k - \eta_k x_k)$$

contains no product of the ω_k *in the* (ξ, η) *variables. Moreover, the coefficients of the homogeneous part of* K *of degree* l *and of* v_l *are polynomials of degree* $l-2$ *in the coefficients of* $H_3, \ldots, H_l$.

b) *First integrals.* We review some classical facts about first integrals (see [Si1]).

If the normalization is converging, then all expressions

$$\omega_k = \xi_k \eta_k$$

are first integrals since

$$\{\omega_k, K\} = \eta_k K_{\eta_k} - \eta_k K_{\xi_k} = \xi_k \eta_k (K' - K') = 0 \ .$$

Expressing ω_k in terms of the initial variables (x, y) we get n formal first integrals

$$P_k(x, y) = \xi_k(x, y) \eta_k(x, y) \ .$$

Observe that

$$\eta_k = y_k - \sum_{l=3}^{+\infty} v_{l, x_k}(x, \eta) \ .$$

So if

$$\eta_k(x, y) = y_k + \sum_{l=2}^{+\infty} \eta_{kl}(x, y)$$

where η_{kl} is the l-homogeneous part of η, then by induction the coefficients of η_{kl} are polynomial on the coefficients of $H_3, \ldots, H_{l+1}$ of degree $l-1$.

We reach the same conclusion for ξ_k using the fact that

$$\xi_k(x, y) = v_{\eta_k}(x, \eta) = x_k + \sum_{l=3}^{+\infty} v_{l, \eta_k}(x, \eta) \ .$$

Now, we have the following formal lemma ([Si1, Lemma 1]):

LEMMA 1.2. *Any formal integral* P *can be represented as a formal power series in the* n *first integrals* $\omega_1, \ldots, \omega_n$.

Thus we can identify the set of formal first integrals with the formal power series in n variables.

Definition 1.3. A *universal formal first integral* $P(H) = F(\omega_1, \ldots, \omega_n)$ where F is *a formal power series in* n *variables.*

COROLLARY 1.4. *Any universal formal first integral* $P(H)$ *has coefficients that are monomials of degree* l *depending polynomially on the coefficients of* $H_3, \ldots, H_{l+1}$ *with degree* $\leq l-1$.

2. Proof of the theorems

a) *Potential and pluripotential theory.* Since the present article presents a possible interest to researchers working in hamiltonian dynamics who are maybe less familiar with potential theory, following the suggestion of the referee we give some background material.

We refer to [Ra] (or [Tsu] for a more encyclopedic exposition) for potential theory in dimension 1. We refer to [Kli, Ch. 5] for proofs and supplementary material on pluripotential theory.

We recall that a function u of one complex variable defined in an open subset $U \subset \mathbf{C}$ is subharmonic if u is upper semi-continuous and satisfies the local sub-mean inequality; i.e., given $z \in U$, there exists $\rho > 0$ such that for any $0 \leq r < \rho$,

$$u(z) \leq \frac{1}{2\pi} \int_0^1 u(z + re^{2\pi i t})\ dt\ .$$

A function u defined on an open subset of $\mathbf{C}^m$ is plurisubharmonic if the restriction of u to any complex line is subharmonic.

Basic examples of subharmonic functions in $\mathbf{C}$ are $\log|f|$ where f is an entire function or the potential of a finite Borel measure μ with compact support

$$p_\mu(z) = \int \log|z-w|\ d\mu(w)\ .$$

The energy of such a measure is defined as

$$I(\mu) = \int p_\mu(z)\ d\mu(z) = \int\int \log|z-w|\ d\mu(z)\ d\mu(w)\ .$$

The logarithmic capacity (or capacity in short) of a subset $E \in \mathbf{C}$ is then defined by

$$\mathrm{cap}(E) = \sup_\mu e^{I(\mu)}$$

where the supremum is taken over Borel probability measures with compact support in E. This supremum is attained for compact sets E (there exists an *equilibrium measure* maximizing the supremum). Capacity is a useful notion that enjoys good set function properties (see for example [Ra, Ch. 5]).

A *polar set* is a set with 0 capacity. It is easy to prove from the definition that a countable union of polar sets is polar. Closed polar sets in $\mathbf{C}$ are the locus where subharmonic functions are $-\infty$ ([Ra, 3.5.4]). Polar sets in $\mathbf{C}$ have area 0. The notion of *thinness* at a point ([Ra, 3.8]), or *barriers*, can be used as a geometric characterisation of polarity. In particular they imply that closed polar sets are totally disconnected.

In higher dimension different notions of capacity are possible (see [Ce]). Pluripolarity is defined via plurisubharmonic functions generalizing a possible definition in dimension 1.

A set $E \subset \mathbf{C}^m$ is locally pluripolar if for each $z \in E$ there is a neighborhood U of z and a plurisubharmonic function u defined on U such that

$$E \cap U \subset u^{-1}(-\infty) \ .$$

A set $E \subset \mathbf{C}^m$ is pluripolar if it is locally pluripolar. There is a nontrivial theorem (Josefson's theorem, [Kli, Th. 4.7.4]) that in such a case there exists a global plurisubharmonic function u such that $E \subset u^{-1}(-\infty)$. Pluripolar sets in C^m have $2m$-dimensional Lebesgue measure 0. For $m = 1$ this is the usual notion of polar set, that is the same as having zero logarithmic capacity.

We consider the set $\mathcal{L}$ of plurisubharmonic functions u defined in $\mathbf{C}^m$ and of minimal growth, i.e. $u(z) - \log ||z||$ is bounded above when $||z|| \to \infty$. Given a subset $E \subset \mathbf{C}^m$, we define

$$V_E(z) = \sup\{u(z); u \in \mathcal{L}, u_{/E} \leq 0\} \ .$$

The upper semi-continuous regularization V_E^* of V_E is called the pluri-subharmonic Green function of E. This function V_E^* is either plurisubharmonic or identically $+\infty$. We are in the former case when E is not pluripolar, when V_E^* has logarithmic growth at ∞; that is, $V_E^*(z) - \log ||z||$ is bounded above when $z \to \infty$.

The Bernstein-Walsh lemma. The following is a classical lemma in potential theory and approximation theory (see [Ra, p. 156] for the one-dimensional version and some applications). It plays a crucial role in the proof of Theorems 1 and 2.

LEMMA (Bernstein-Walsh). *If $E \subset \mathbf{C}^m$ is not pluripolar, and P is a polynomial of degree d, then for $z \in \mathbf{C}^m$,*

$$|P(z)| \leq ||P||_{C^0(E)} \ e^{dV_E(z)} \ .$$

Proof. The plurisubharmonic function

$$u(z) = \frac{1}{\deg P} \log \left(\frac{P(z)}{||P||_{C^0(E)}} \right)$$

has minimal growth and $u \leq 0$ on E; thus $u \leq V_E$. □

For later reference, we note that a countable union of pluripolar sets is pluripolar.

b) *Proof of Theorem* 1. The result about the divergence of the normalization mapping follows the same lines as the case of the Birkhoff normal form. The convergence or divergence of the normalizing transformation is equivalent to the convergence or divergence of the generating function. Then the proof proceeds in the same way as below if we use the polynomial dependence of the generating function on the coefficients of H (Prop. 1.1).

For the elementary construction of hamiltonians with divergent normalization mentioned at the end of the introduction, we refer the reader to the end of Section 30 of [Si-Mo], and to Siegel's article [Si1].

We consider the problem of convergence or divergence of the Birkhoff normal form. The first assertion of the theorem follows from the second. Actually, consider the set $F_n \subset \mathcal{H}_Q$ of hamiltonians in $\mathcal{H}_Q$ having a converging Birkhoff normal form with radius of convergence $\geq 1/n$, and bounded by 1 in the open ball of center 0 and radius $1/n$. We claim that this set F_n is closed. To prove this, consider a sequence of hamiltonians (H_i) in F_n converging uniformly on compacts sets to $H \in \mathcal{H}_Q$. Denote (K_{H_i}) the corresponding sequence of Birkhoff normal forms. They are all bounded by 1 in the open ball of center 0 and radius $1/n$; thus they form a normal family. Moreover, any limit point of the sequence (K_{H_i}) must be K_H because of the coefficient convergence (the coefficients of the H_i converge to those of H, and so the ones of K_{H_i} converge to those of K_H). Thus K_H has radius of convergence $\geq 1/n$ and is bounded by 1 in the open ball of center 0 and radius $1/n$. Now,

$$F = \bigcup_{n \geq 1} F_n$$

is the set of all hamiltonians in $\mathcal{H}_Q$ having a convergent Birkhoff normal form (so this set is an F_σ-set). Moreover, the open set $\mathcal{H}_Q - F_n$ is dense. Otherwise let H_1 be a hamiltonian in the interior of F_n. Considering the complex (resp. real) affine subspace $V = \{(1-t)H_0 + tH_1; t \in \mathbf{C}(\text{resp.}\mathbf{R})\} \subset \mathcal{H}_Q$ we have, according to the second assertion in Theorem 1, that the set of hamiltonians with converging Birkhoff normal form must have capacity zero (resp. Lebesgue measure 0). But on the other hand it contains a neighborhood of 1. Contradiction. Note that for this argument we only needed to use the second part of the theorem for a one-dimensional subspace V (thus the reader only interested in this first part, only needs classical potential theory and not pluripotential theory in higher dimension).

The real analytic result follows from the $\mathbf{C}$-analytic one by the observation that the intersection of a pluripolar set in $\mathbf{C}^n$ with $\mathbf{R}^n \subset \mathbf{C}^n$ has Lebesgue measure 0 (see [Ro, Lemma 2.2.7, p. 90]).

We consider a complex finite-dimensional affine subspace V of $\mathcal{H}$, $V \approx \mathbf{C}^m$. We can parametrize linearly the coefficients of hamiltonians $H \in V$ with a complex parameter $t \in \mathbf{C}^m$, and we denote H_t the corresponding hamiltonian in V. Note that the coefficients of H_t are linear functions of t.

We assume that the Birkhoff normal form of hamiltonians H_t corresponding to a set of values $t \in F \subset \mathbf{C}^m$, with F not pluripolar, are converging. We want to prove that all the other hamiltonians in V have converging Birkhoff normal form.

Now

$$F = \bigcup_{n \geq 1} F_n$$

where F_n is the set of parameters $t \in \mathbf{C}^m$ such that the hamiltonian H_t has a Birkhoff normal form K_t with radius of convergence larger or equal to $1/n$ where K_t is bounded by 1 in this polydisk of radius $1/n$. Now, if F is not pluripolar, we have for some $n \geq 1$ that F_n is not pluripolar (and this set is also closed). If we denote

$$K_t(\xi, \eta) = \sum_i K_i(t)(\xi, \eta)^i \ ,$$

then, according to Proposition 1.1, the coefficients $K_i(t)$ depend polynomially on t with degree $\leq |i| - 2$ (for $|i| \geq 3$). Now, by Cauchy inequality, there exists $\rho_0 > 0$ such that for all $t \in F_n$,

$$\varphi(t) = \sup_i |K_i(t)| \rho_0^{-|i|} < +\infty \ .$$

The function φ is lower semicontinuous, and

$$F_n = \bigcup_m L_m$$

where $L_m = \{z \in F_n ; \varphi(t) \leq m\}$ is closed [2]. For some m, L_m has positive capacity. Finally we found a compact set $C \subset L_m$ of positive capacity such that there exists $\rho_1 > 0$ such that for any $t \in C$ and and all i,

$$|K_i(t)| \leq \rho_1^{|i|} \ .$$

[2]The closedness is not really necessary but we prefer to work with closed sets.

Using the Bernstein-Walsh lemma and Proposition 1.1 we get that for any compact set $C_0 \subset \mathbf{C}^m$, for $|i| \geq 3$,

$$||K_i||_{C^0(C_0)} \leq e^{(|i|-2)\max_{t\in C_0} V_C(t)} \rho_1^{|i|} \leq \rho^{|i|-2} \rho_1^{|i|} \ ,$$

for some constant ρ depending only on C_0. Thus K_t is converging for any $t \in \mathbf{C}^m$.

Remark. Note that in the case of convergence, the proof gives an explicit lower bound for the radius of convergence for all $t \in \mathbf{C}^m$. More precisely, using the sub-exponential growth at ∞ of V_C we get that there exists a constant $C_1 > 0$ such that

$$R(K_t) \geq \frac{C_1}{1 + ||t||} \ .$$

c) *Proof of Theorem* 2. The proof of Theorem 2 follows exactly the same lines as the proof of Theorem 1 once the polynomial dependence of universal formal first integrals has been proved (Corollary 1.4).

UCLA, Los Angeles, CA
Université Paris-Sud, Orsay, France
E-mail address: ricardo@math.ucla.edu

References

[Ar] V. I. Arnold, Proof of a theorem of A. N. Kolmogorov on the preservation of conditionally periodic motions under a small perturbation of the Hamiltonian, *Usp. Mat. Nauk*. **18** (1963), 13–40.

[Bi1] G. D. Birkhoff, *Dynamical Systems*, A. M. S., Providence, RI, 1927.

[Bi2] ———, Surface transformations and their dynamical applications, *Acta Math.* **43** (1922), 1–119.

[Br] A. D. Brjuno, Analytical form of differential equations I, II, *Trans. Mosc. Math. Soc.* **25** (1971), 119–262, **26** (1972), 199–239.

[Ce] U. Cegrell, Capacities in complex analysis, in *Aspects of Mathematics* **E14**, Friedr. Vieweg & Sohn, Braunschweig, 1988.

[Ch] T. M. Cherry, On the solution of Hamiltonian systems of differential equations in the neighborhood of a singular point, *Proc. London Math. Soc.* **27** (1928), 151–170.

[De] C. E. Delaunay, Théorie du mouvement de la lune, *Paris Mem. Prés.* **28** (1860), **29** (1867).

[Ein] A. Einstein, *Verch. Dtsch. Phys. Ges.* **19:82** (1917).

[El1] L. H. Eliasson, Normal forms for Hamiltonian systems with Poisson commuting integrals-elliptic case, *Comment. Math. Helv.* **65** (1990), 4–35.

[El2] ———, Hamiltonian systems with linear form near an invariant torus, in *Non-linear Dynamics* (Bologna, 1988), World Sci. Publ., Teaneck, NJ, 1989, 11–29.

[El3] ———, Hamiltonian systems with Poisson commuting integrals, Ph.D. thesis, Univ. Stockholm, 1984.

[E-V] J. Ecalle and B. Vallet, Correction and linearization of resonant vector fields and diffeomorphisms, *Math. Z*. **229** (1998), 249–318.

[Ga] G. Gallavotti, A criterion for integrability for perturbed nonresonant harmonic oscillators. "Wick ordering" of the perturbations in classical mechanics and invariance of the frequency spectrum, *Commun. Math. Phys.* **87** (1982), 365–383.

[Gu] M. C. Gutzwiller, *Chaos in Classical and Quantum Mechanics*, *Interdisciplinary Applied Mathematics*, **1**, Springer-Verlag, New York, 1990.

[It] H. Ito, Convergence of Birkhoff normal forms for integrable systems, *Comment. Math. Helv*. **64** (1989), 412–461.

[Kel] J. B. Keller, Corrected Bohr-Sommerfeld quantum conditions for nonseparable systems, *Ann. Phys*. **4** (1958), 180–188.

[Kli] M. Klimek, *Pluripotential Theory*, *London Math. Society Monographs* **6**, Oxford Univ. Press, New York, 1991.

[Ko] A. N. Kolmogorov, Théorie générale des systèmes dynamiques et mécanique classique (Amsterdam, 1954), *Proc. Internat. Congress of Math.* **1** (1957), 315–333.

[Koz] V. V. Kozlov, Integrability and nonintegrability in Hamiltonian mechanics, *Usp. Mat. Nauk*. **38** (1983), 3–67; *Russian Math. Surveys* **38** (1983), 1–76.

[Li] A. Lindstedt, Beitrag zur Integration der Differentialgleichungen der Störungstheorie, *Abh. K. Akad. Wiss. St. Petersburg* **31** (1882).

[Ll] R. de la Llave, On necessary and sufficient conditions for uniform integrability of families of Hamiltonian systems, *Internat. Conf. on Dynamical Systems* (Montevideo, 1995), *Pitman Res. Notes Math. Ser*. **362**, Longman, Harlow, 1996, 76–109.

[Mor] J. J. Morales-Ruiz, *Differential Galois Theory and Non-Integrability of Hamiltonian Systems*, *Progr. in Math*. **179**, Birkhäuser Verlag, Basel, 1999.

[Mo] J. K. Moser, On invariant curves of area-preserving mappings of an annulus, *Nachr. Akad. Wiss. Göttingen Math.-Phys*. **1962** (1962), 1–20.

[Mou] R. F. Moulton, *An Introduction to Celestial Mechanics*, 2nd edition, The MacMillan Co., New York, 1923.

[Ne] N. N. Nehoroshev, The behavior of Hamiltonian systems that are close to integrable ones, *Funct. Anal. Appl*. **5** (1971), 82–83.

[PM1] R. Pérez-Marco, Total convergence or general divergence in small divisors, *Comm. Math. Phys.* **223** (2001), 451–464..

[PM2] ———, Linearization of holomorphic germs with resonant linear part, preprint, 2000.

[Po] H. Poincaré, *Les Méthodes Nouvelles de la Mécanique Céleste*, Paris, 1892.

[Ra] T. Ransford, Potential theory in the complex plane, *London Math. Soc. Student Texts* **28**, Cambridge Univ. Press, Cambridge, 1995.

[Rob] R. C. Robinson, Generic properties of conservative systems I, II, *Amer. J. Math.* **92** (1970), 562–603, 897–906.

[Ro] L. I. Ronkin, *Introduction to the Theory of Entire Functions of Several Variables*, *Transl. Math. Monographs* **44**, A. M. S., Providence, RI, 1974.

[Ru1] H. Rüssmann, Uber das Verhalten analytischer Hamiltonscher Differentialgleichungen in der Nähe einer Gleichgewichtslösung, *Math. Ann*. **154** (1964), 285–300.

[Ru2] ———, Über die Normalform analytischer Hamiltonscher Differentialgleichungen in der Nähe einer Gleichgewichtslösung, *Math. Ann*. **169** (1967) 55–72.

[Ru3] ———, On the convergence of power series transformations of analytic mappings near a fixed point, preprint I.H.E.S., Paris, 1977.

[Si1] C. L. Siegel, On the integrals of canonical systems, *Ann. of Math.* **42** (1941), 806–822.

[Si2] ———, Uber die Existenz einer Normalform analytischer Hamiltonscher Differentialgleichungen in der Nähe einer Gleichgewichtlösung, *Math. Ann*. **128** (1954), 144–170.

[Si-Mo] C. L. Siegel and J. Moser, *Lectures on Celestial Mechanics* **187**, Springer-Verlag, New York, 1971.

[St1] L. Stolovitch, Complète intégrabilité singulière, *C. R. Acad. Sci. Paris Sér. I Math.* **326** (1998), 733–736.

[St2] ———, Normalisation holomorphe d'algèbres de type Cartan de champs de vecteurs holomorphes singuliers, *C. R. Acad. Sci. Paris Sér. I Math.* **330** (2000), 121–124.

[St3] ———, Un phénomène de type KAM, non symplectique, pour les champs de vecteurs holomorphes singuliers, *C. R. Acad. Sci. Paris Sér. I Math.* **332** (2001), 545–550.

[Tsu] M. Tsuji, *Potential Theory in Modern Function Theory*, Maruzen Co., Ltd., Tokyo, 1959.

[Wh] E. T. Whittaker, On the solutions of dynamical problems in terms of trigonometric series, *Proc. London Math. Soc.* **11**, **34** (1902), 206–221.

[Wi] A. Wintner, *The Analytic Foundations of Celestial Mechanics*, *Princeton Math. Series* **5**, Princeton Univ. Press, Princeton, NJ, 1941.

(Received September 1, 2000)