# CONVERGENCE OF THE VON ZEIPEL PROCEDURE*

RICHARD B. BARRAR
*Dept. of Mathematics, University of Oregon, Eugene, Ore., U.S.A.*

(Received 24 February, 1970)

In the last fifteen years notable progress has been made in celestial mechanics, centered around the contributions of Siegel, Kolmogorov, Moser and Arnold. Perhaps the key paper in a whole series of brilliant papers by these distinguished mathematicians was the four page announcement by Kolmogorov [5]. Moser and Arnold generalized Kolmogorov's theorems, and extended them in various fashion with appropriate proofs, but no proof has been given following Kolmogorov's original outline. The present paper is devoted to doing this. Moreover, it is felt that as an introduction to this chain of ideas, the Kolmogorov approach is the least complicated.

Before proceeding, let us give a brief history of the problem under consideration. Given a Hamiltonian:

$$H(p_1, \ldots, p_n, q_1, \ldots, q_n) = H_0(p_1, \ldots, p_n) + \mu H_1(p_1 \ldots p_n, q_1, \ldots q_n\mu)$$

that is real analytic in all its variables and periodic in the $q$'s of period $2\pi$, one of the basic techniques in celestial mechanics is to make a canonical change of variable $(p, q) \leftrightarrow (P, Q)$ by means of a generating function of the form

$$W(P, q) = \sum P_i q_i + \sum \mu^n S_n(P, q)$$

such that in terms of the new variables $(P, Q)$ the Hamiltonian is only a function of the $P$'s. Let us define this as a von Zeipel transformation.

In [10, Vol. II, Par. 125] Poincaré showed that formally a von Zeipel transformation was always possible, at least at points $p$ where the $n$ numbers $\lambda_i = \partial H_0/\partial p_i$ were rationally independent. In this case the Hamiltonian is called nondegenerate. However Poincaré felt that the resulting formal expansion would never converge, but was rather an asymptotic expansion. The central idea of Kolmogorov (which we state more precisely below) is that if the $\lambda_i$ satisfy inequalities of the form:

$$(*) \qquad |\sum h_i \lambda_i| \geqslant \varepsilon / (\sum |h_i|)^n$$

for all integers $h_i$ not identically zero then the von Zeipel transformation actually will converge, not everywhere, but at points where (*) holds.

Inequalities of the form (*) were first used by Siegel [12, 13] in similar investigations. He has given a simple proof [11, p. 165, 166] of the fact that in the sense of measure theory almost all points $\lambda$ satisfy (*) for some $\varepsilon > 0$. Arnold [2] has given proofs of results that extend those of Kolmogorov to degenerate Hamiltonians,

* Portions of this paper were presented at the Seventh Annual Summer Institute in Dynamical Astronomy held at Purdue University, June 19–July 14, 1967.

*Celestial Mechanics* **2** (1970) 494–504. *All Rights Reserved*

hence which are applicable to the main problems of celestial mechanics. Kolmogorov and Arnold deal only with analytic Hamiltonians, while Moser [6, 8] has given proofs that are applicable to functions which are assumed to have only a finite number of derivatives.

Following Kolmogorov, we establish our main theorem below by taking the limit of an infinite number of canonical transformations. This is not how we defined a von Zeipel transformation above. However, from recent results of Moser [9] it follows that if the Kolmogorov procedure converges, so does the usual von Zeipel series expansion.

In summary, consider any point $p_i$ where (*) is satisfied, translate it to the origin and perform a von Zeipel transformation, then the results of this paper imply that (if $\mu$ is small enough) the von Zeipel transformation will converge at $P_i=0$.

The main theorem that we wish to prove is:

THEOREM 1. If $H(p,q)=H(p_1 \dots p_m, q_1 \dots q_m)$ is analytic, and periodic of period $2\pi$ in the $q$'s, and can be written in the form:

$$H(p, q) = \eta + \sum \lambda_i p_i + A(q) + \sum B_i(q)\, p_i + \sum C_{ij}(q)\, p_i p_j + D(p, q) \quad (1)$$

($\eta$ a constant, and $D(p, q) \equiv$ terms of third and higher order in the $p_i$) satisfying the hypotheses

(H1) $$|\sum \lambda_i h_i| \geqslant \frac{\varepsilon}{(\|h\|)^s}, \text{ for integer } h_i \text{ with } \|h\| = \sum_{i=1}^{m} |h_i| > 0,\ s \text{ an integer} \geqslant m.$$

(H2) $$\det c_{ij}(0) \neq 0. \quad c_{ij}(0) = \frac{1}{(2\pi)^m} \int_0^{2\pi} \cdots \int_0^{2\pi} C_{ij}(q)\, dq_1 \dots dq_m,$$

then for sufficiently small $A(q)$ and $B_i(q)$, there is a solution of the Hamiltonian equations of the form

$$q_i = \bar{Q}_i + f_i(\bar{Q})$$
$$p_i = g_i(\bar{Q})$$

with $\bar{Q}=(\bar{Q}_1, \dots, \bar{Q}_n)$, $\bar{Q}_i=\lambda_i(t-\tau_i)$, $\tau_i$ a constant, and $f_i$ and $g_i$ periodic of period $2\pi$ in the $\bar{Q}_i$'s and analytic.

Theorem 1 follows from the following result:

THEOREM 2 (KOLMOGOROV THEOREM). If $H(p, q)$ of Theorem 1 is analytic for $|p_i| \leqslant r_0$, and $|\mathrm{Im}\, q_i| \leqslant \varrho_0$, then for sufficiently small $A(q)$ and $B_i(q)$ in this region, there is a canonical change of variables $(p, q) \leftrightarrow (P, Q)$ with

$$p_i = P_i + g_i(Q) + \sum P_j g_{ij}(Q) \quad (2)$$
$$q_i = f_i(Q) + Q_i$$

such that $f_i$, $g_i$, $g_{ij}$, which are periodic of period $2\pi$ in the $Q$'s, are defined and analytic for $|\mathrm{Im}\, Q_i| < 3\varrho_0/4$. Further, the mapping (2) maps $|P_i| < 3r_0/4$, $|\mathrm{Im}\, Q_1| < 3\varrho_0/4$ into

$|p_i| \leqslant r_0$, $|\text{Im}\, q_i| \leqslant \varrho_0$. In terms of $P$, $Q$ the Hamiltonian is analytic in the former region and is of the form

$$H = \sum \lambda_i P_i + \sum \tilde{C}_{ij}(Q)\, P_i P_j + 0(P^3). \tag{3}$$

(That is, the terms $A(q)$ and $B_i(q)$ no longer appear, hence $P_i = 0$, $Q_i = \lambda_i(t - \tau_i)$ is a solution of the canonical equations.)

We now outline the method suggested by Kolmogorov to prove Theorem 2. The transformation specified is the limit of an infinite sequence of transformations, in each of which the terms $A(q)$ and $B_i(q)$ get smaller and smaller.

Starting with the Hamiltonian in the form (1), let us introduce the first change of variables $(p, q) \leftrightarrow (P, Q)$ defined by the generating function

$$S(P, q) = \sum_{i=1}^{m} (P_i + \xi_i)\, q_i + X(q) + \sum P_i Y_i(q), \quad (\xi_i \text{ a constant}).$$

Then

$$p_i = \frac{\partial S}{\partial q_i} = (P_i + \xi_i) + \frac{\partial X}{\partial q_i} + \sum P_j \frac{\partial Y_j(q)}{\partial q_i}, \tag{4}$$

and

$$Q_i = \frac{\partial S}{\partial P_i} = q_i + Y_i(q). \tag{5}$$

Substituting the expression (4) for $p_i$ in the original Hamiltonian and noting from Equation (5) that $q_i = F_i(Q)$ for some $F_i$, we find

$$\begin{aligned} H = \eta + \sum \lambda_i \xi_j + a(0) + \sum \lambda_i P_i + A^*(q) + \sum P_i B_i^*(q) + A^{(1)}(Q) \\ + \sum P_i B_i^{(1)}(Q) + \sum C_{ij}^{(1)}(Q) P_i P_j + D^{(1)}(P, Q) \end{aligned}$$

where $D^{(1)}(P, Q)$ signifies terms of third and higher power in the $P_i$. In the above we set

$$A^*(q) = \sum \lambda_i \frac{\partial X}{\partial q_i} + A(q) - a(0)$$

$$B_i^*(q) = \sum_j \lambda_j \frac{\partial Y_i}{\partial q_j} + \sum_j C_{ij}(q) \left[ \xi_j + \frac{\partial X}{\partial q_j} \right] + B_i(q),$$

and we use the notation $z_j = e^{iq_j}$, $A(q) = \sum a(k)\, e^{i(k,q)} = \sum a(k)\, z^k$, with

$$a(k)\, e^{i(k,q)} = a_{k_1 \dots k_m}\, e^{i(k_1 q_1 \dots + k_m q_m)} = a_{k_1 \dots k_m} z_1^{k_1} \dots z_m^{k_m}, \text{ etc}.$$

In (4) and (5) $X$, $\xi_i$, $Y_i$ are assumed to be small quantities. They occur in $A^*(q)$ and $B_i^*(q)$ at most to the first order, and in $A^{(1)}(Q)$ and $B_i^{(1)}(Q)$ at least to the second order.

The procedure suggested by Kolmogorov is to choose $X$, $\xi_i$, $Y_i$ such that $A^*(q) = B_i^*(q) = 0$. This is done as follows:

Setting $X(q) = \sum' x(k)\, z^k$, $Y_i(q) = \sum' y_i(k)\, z^k$, with $\sum'$ meaning that the term $k = 0$ is absent (i.e., setting $x(0)=y_i(0)=0$), we have from $A^*(q)=0$:

$$\sqrt{-1}(k,\lambda)\,x(k)+a(k)=0,\ \text{with}\,(k,\lambda)=\sum k_i\lambda_i, \quad \text{which determines } x(k). \tag{6}$$

Setting

$$E_i(q)=\sum e_i(k)\,z^k=\sum_j C_{ij}(q)\frac{\partial X}{\partial q_j} \tag{7}$$

we have from $B_i^*(q)=0$:

$$c_{ij}(0)\,\xi_j+e_i(0)+b_i(0)=0,\ \text{which determines } \xi_j; \tag{8}$$

then

$$\sqrt{-1}(k,\lambda)\,y_i(k)+e_i(k)+b_i(k)+c_{ij}(k)\,\xi_j=0, \quad \text{which determines } y_i(k). \tag{9}$$

Note that hypothesis (H1) and (H2) insure that (6), (8) and (9) have unique solutions. Since we assume all quantities $X$, $\xi_i$, $Y_i$ are small, it follows that $A^{(1)}$, $B_i^{(1)}$ are small of the second order, and thus by repeated application of this scheme, we look for convergence. This is what we now establish.

LEMMA 1. If the $\lambda_i$ satisfy hypothesis (H1) of Theorem 1, and if $T=\sum' t(k)z^k$, and

$$S=-\sum{}'\frac{t(k)\,z^k}{\sqrt{-1}(k,\lambda)},$$

i.e., if $S$ satisfies the equation

$$\sum\lambda_i\frac{\partial S}{\partial q_i}=\sqrt{-1}\sum_{i=1}^{m}\lambda_i z_i\frac{\partial S}{\partial z_i}=T,$$

then $S$ satisfies the following estimates for some constants $C$:

$$\|S\|_{S(\varrho-h)}\leqslant\frac{C}{\varepsilon(h)^{s+m}}\|T\|_{S(\varrho)}$$

$$\left\|z_i\frac{\partial S}{\partial z_i}\right\|_{S(\varrho-h)}\leqslant\frac{C}{\varepsilon(h)^{s+m+1}}\|T\|_{S(\varrho)},$$

with $0<h<\rho$, where we use the notation:

$$S(\varrho)=\{z_i\,|e^{-\varrho}\leqslant|z_i|\leqslant e^{\varrho}\}=\{q_i|\,|\mathrm{Im}\,q_i|\leqslant\varrho\}$$

$$\|F\|_{S(\varrho)}=\sup_{z\in S(\varrho)}|F(z)|.$$

PROOF: The mapping $z_j=e^{iq_j}$ takes the strip $|\mathrm{Im}\,q_i|\leqslant\varrho$ into the annulus $e^{-\varrho}\leqslant|z_i|\leqslant e^{\varrho}$. In the following we work in the annulus.

By Hölder's inequality, and Hypothesis (H1) of Theorem 1 we obtain

$$\left|\frac{1}{(k,\lambda)}\right| \leqslant \frac{\left(\sum_{i=1}^{m} |k_i|\right)^s}{\varepsilon} \leqslant \frac{m^{s-1}}{\varepsilon} \sum_{i=1}^{m} |k_i|^s .$$

If $l(x)=x/1-x$, and $M=\|T\|_{S(\varrho)}$, we obtain by Cauchy's inequality

$$T \ll M \prod_{i=1}^{m} \left(l(z_i/e^{\varrho}) + l(e^{-\varrho}/z_i)\right) = M(z) = \sum{}' \, m(k)\, z^k ,$$

where $\ll$ means majorized by.
Thus

$$S \ll \sum{}' \frac{m(k)\, z^k}{|(k,\lambda)|} \ll \frac{m^{s-1}}{\varepsilon} \sum{}' \, m(k)\,[(|k_1|^s + \cdots + |k_m|^s)]\, z^k ;$$

then since

$$\left(x\frac{\mathrm{d}}{\mathrm{d}x}\right)^s x^k = (k)^s x^k :$$

$$S \ll \frac{m^{s-1}}{\varepsilon} \sum_{j=1}^{m} \left(\pm z_j \frac{\partial}{\partial z_j}\right)^s M(z)$$

$$z_i \frac{\partial S}{\partial z_i} \ll \frac{m^{s-1}}{\varepsilon} \left(\pm z_i \frac{\partial}{\partial z_i}\right) \sum_{j=1}^{m} \left(\pm z_j \frac{\partial}{\partial z_j}\right)^s M(z)$$

and the result follows since

$$\left(x\frac{\mathrm{d}}{\mathrm{d}x}\right)^s (x/1-x) = P(x)/(1-x)^{s+1} .$$

LEMMA 2. Let

$$\varepsilon_0 = \max(\|A\|_{S(\varrho_0)}, \|B_i\|_{S(\varrho_0)}),$$
$$\varepsilon_1 = \max(\|A^{(1)}\|_{S(\varrho_1)}, \|B_i^{(1)}\|_{S(\varrho_1)}),$$
$$r_1 = r_0 - 2h, \quad \varrho_1 = \varrho_0 - 4h, \quad \text{with} \quad 0 < h < r_0/2, \quad 0 < h < \varrho_0/4,$$

and introduce the following additional notation:

$$S(r,\varrho) = \{(p_i, q_i) \,\big|\, |p_i| \leqslant r;\ |\mathrm{Im}\, q_i| \leqslant \varrho\},\ \|F(p,q)\|_{S(r,\varrho)} = \sup_{(p,q)\in S(r,\varrho)} |F(p,q)| .$$

Let

$$\max |c_{ij}^{-1}(0)| \leqslant 2N,\ \|D(p,q)\|_{S(r_0,\varrho_0)} \leqslant 2M,\ \|C_{ij}\|_{S(\varrho_0)} \leqslant 2M ,$$

where $c_{ij}^{-i}(0)$ is the inverse of the matrix $c_{ij}(0)$ defined in (H2) of Theorem 1, and let the $\lambda_i$ satisfy (H1) of Theorem 1, and using the notation of that hypothesis let $h<\varepsilon$,

$t=s+m+1$, and finally let $1 \geqslant r_0 \geqslant \frac{1}{2}$. Then there are constants $C_i$ depending only on $N$, $M$ and $m$, such that if

$$\text{(H3)} \qquad \varepsilon_0 \leqslant \frac{h^{2t+3}}{C_0}$$

then the following assertions hold:

(A1) The transformation (4), (5) may be written:

$$q_i = Q_i + f_i(Q) = F_i(Q),$$
$$p_i = P_i + g_i(Q) + \sum g_{ij}(Q) P_j = G_i(P, Q),$$

with $F_i$, $G_i$ being analytic for $Q \varepsilon S(\varrho_1)$,

$$\|f_i\|_{S(\varrho_1)} \leqslant 2h, \; \|g_i\|_{S(\varrho_1)} \leqslant 2h/m+2, \; \|g_{ij}\|_{S(\varrho_1)} \leqslant h/m+2,$$

thus $F_i$, $G_i$ maps $S(r_1, \varrho_1)$ into $S(r_0-h, \varrho_0-2h)$.

$$\text{(A2)} \qquad |\varepsilon_1| \leqslant \frac{C_1}{h^{3t+3}} |\varepsilon_0|^2 .$$

(A3). The Hamiltonian is defined for $(P, Q) \varepsilon S(r_1, \varrho_1)$ and the following estimates hold:

$$\text{(a)} \qquad \|C_{ij}^{(1)}\|_{S(\varrho_1)} \leqslant \|C_{ij}\|_{S(\varrho_0)} + C_2 h ,$$
$$\text{(b)} \qquad \|D^{(1)}(P, Q)\|_{S(r_1, \varrho_1)} \leqslant \|D(p, q)\|_{S(r_0, \varrho_0)} + C_3 h ,$$
$$\text{(c)} \qquad |c_{ij}^{(1)}(0) - c_{ij}(0)| \leqslant \left(\frac{C_4}{\varrho_1} + C_5\right) h .$$

PROOF: In the following, $C$ will stand for a constant that depends on $N$, $M$ and $m$. Its exact value may change from equation to equation.

Solving for $X(q)$ from (6), we find from Lemma 1, that

$$X_1^* = \max \left\| \frac{\partial X}{\partial q_i} \right\|_{S(\varrho_0 - h)} \leqslant \frac{C_x \varepsilon_0}{h^{t+1}} .$$

Thus, recalling the definitions and equation for $E_i$ and $\xi_i$((4), (7), (8)), we find:

$$\|E_i\|_{S(\varrho_0 - h)} \leqslant \frac{C_e \varepsilon_0}{h^{t+1}}$$
$$|\xi_i| \leqslant \frac{C_\xi \varepsilon_0}{h^{t+1}} .$$

Having found $X$, $E_i$, $\xi_i$ in $S(\varrho_0 - h)$, we can solve (9) for $Y_i$ in $S(\varrho_0 - 2h)$. Lemma 1 implies

$$Y^* = \max \|Y_i\|_{S(\varrho_0 - 2h)} \leqslant \frac{C_y \varepsilon_0}{h^{2t+1}} ; \quad Y_1^* = \max \left\| \frac{\partial Y_i}{\partial q_j} \right\|_{S(\varrho_0 - 2h)} \leqslant \frac{C_{y1} \varepsilon_0}{h^{2t+2}} .$$

We now show, that if $Y^* < 2h$, that transformation (5) has an inverse $q_i = Q_i + f_i(Q)$ with $\| f_i \|_{S(\varrho_0 - 4h)} < 2h$, mapping $S(\varrho_1)$ into $S(\varrho_0 - 2h)$. We carry out the proof assuming two degrees of freedom. The proof readily generalizes to $n$ dimensions.

In the Equation (5) involving $Y_1$, we consider $q_2$ fixed. Rouche's theorem applied for $|\mathrm{Im}\, q_1| \leqslant \varrho_0 - 2h$, asserts that every point $Q_1$ with $|\mathrm{Im}\, Q_1| \leqslant \varrho_0 - 4h$ is the image of one and only one point $q_1$. The derivative of the transformation cannot vanish, otherwise the transformation would not be locally one to one. (See H. Cartan [4] p. 174). Hence by the implicit function theorem for analytic functions:

$$(*) \qquad q_1 = Q_1 + g_1(Q_1, q_2).$$

Substituting this expression for $q_1$ into the Equation (5) involving $Y_2$, and considering $Q_1$ fixed, we find as above

$$q_2 = Q_2 + f_2(Q_1, Q_2).$$

Substituting this expression for $q_2$ into the equation (*) establishes the equations $q_i = Q_1 + f_i(Q)$.

Since $f_i(Q) = -Y_i(q)$, it follows that $\| f_i \|_{S(\varrho_0 - 4h)} < 2h$. Further if $q_i$ and $Q_i$ satisfy (5) so do $q_i + 2\pi n$ and $Q_1 + 2\pi n$; from this it follows that the $f_i(Q)$ are periodic.

The estimates for $X_1^*$, $Y_1^*$, $\xi_i$ apply in $S(\varrho_0 - 2h)$. Hence it follows from transformation (4), that the assertion (A1) of the Lemma will follow if $C_0 = \max\,(C_x, C_\xi, C_y, C_{y1}, m+2)$. We further remark that with this choice, $X_1^*$, $Y_1^*$, $\xi_i$ are all less than $h/m+2$.

To complete the proof, we write out the expressions for the terms in $H(P, Q)$:

$$A^{(1)}(Q) = \sum C_{ij}(q)\left[\xi_i + \frac{\partial X}{\partial q_i}(q)\right]\left[\xi_j + \frac{\partial X}{\partial q_j}(q)\right] + D_0(q)$$

$$B_i^{(1)}(Q) = \sum C_{kj}(q)\left[\xi_k + \frac{\partial X}{\partial q_k}(q)\right]\left[\frac{\partial Y_i}{\partial q_j}\right] + D_i(q)$$

$$C_{ij}^{(1)}(Q) = C_{ij}(q) + \sum C_{kl}(q)\frac{\partial Y_i}{\partial q_k}(q)\frac{\partial Y_j}{\partial q_l}(q) + D_{ij}(q)$$

$$D^{(1)}(P, Q) = D\left(P_i + \xi_i + \frac{\partial X}{\partial q_i}(q) + \sum P_j \frac{\partial Y_j}{\partial q_i}(q), q\right) - D_0(q) - \sum D_i(q) P_i - \sum D_{ij}(q) P_i P_j$$

with

$$D_0(q) = D\left(\xi_i + \frac{\partial X}{\partial q_i}(q), q\right)$$

$$D_i(q) = \sum \frac{\partial D}{\partial p_k}\left(\xi_j + \frac{\partial X}{\partial q_j}(q), q\right)\left[\delta_{ik} + \frac{\partial Y_i}{\partial q_k}\right]$$

and

$$D_{ij}(q) = \sum \frac{\partial^2 D}{\partial p_k \partial p_k}\left(\xi_j + \frac{\partial X}{\partial q_j}(q), q\right)\left[\delta_{ik} + \frac{\partial Y_i}{\partial q_k}\right]\left[\delta_{jl} + \frac{\partial Y_j}{\partial q_l}\right],$$

where in all terms on the right we substitute $q_i = Q_i + f_i(Q)$.

When $Q \varepsilon S(\varrho_1)$, assertion A1 of the Lemma implies $q \varepsilon S(\varrho_0 - 2h)$, and hence the estimates for $X_1^*$, $\xi_i$, $Y_1^*$ apply.

To estimate the terms $D_0(p, q)$, $D_i(p, q)$, $D_{ij}(p, q)$, we note that if $D(p, q)$ is of the third order in $p$, and if $1 \geqslant r_0 \geqslant \frac{1}{2}$, with $|p| \leqslant r_0/2$, $|p| = \max(|p_1| \ldots |p_m|)$ it readily follows from Cauchy's inequality, and then from Schwarz's lemma for several complex variables, that:

$$\|D(p, q)\|_{S(\varrho_0)} \leqslant C_{01} |p|^3$$

$$\left\| \frac{\partial D}{\partial p_i}(p, q) \right\|_{S(\varrho_0)} \leqslant C_{02} |p|^2$$

$$\left\| \frac{\partial^2 D}{\partial p_i \partial p_j}(p, q) \right\|_{S(\varrho_0)} \leqslant C_{03} |p|.$$

Thus assertion (A2) readily follows. If we recall that $X_1^*$, $\xi_i$, $Y_1^*$ are all less than $h$, assertion (A3), parts (a) and (b), also follow. For part (c), we need the further estimate:

$$\begin{aligned}
\overline{C_{ij}(q)} &= \frac{1}{(2\pi)^m} \int_0^{2\pi} \cdots \int_0^{2\pi} C_{ij}(q) \, dQ_1 \ldots dQ_m \\
&= \frac{1}{(2\pi)^m} \int_0^{2\pi} \cdots \int_0^{2\pi} C_{ij}(Q_k + f_k(Q)) \, dQ_1 \ldots dQ_m \\
&= \frac{1}{(2\pi)^m} \int_0^{2\pi} \cdots \int_0^{2\pi} C_{ij}(Q) \, dQ_1 \ldots dQ_m \\
&\qquad + \sum_l \frac{1}{(2\pi)^m} \int_0^{2\pi} \cdots \int_0^{2\pi} \frac{\partial C_{ij}}{\partial Q_l}(Q_k + \theta f_k(Q)) \, f_k \, dQ_1 \ldots dQ_m,
\end{aligned}$$

where we integrate over real $Q$. Thus estimating the integrand in the last term we find:

$$|\overline{C_{ij}(q)} - c_{ij}(0)| \leqslant \frac{C}{\varrho_0 - 2h} 2h \leqslant \frac{C}{\varrho_1} 2h.$$

This, together with our previous estimates, establishes the Lemma.

PROOF of Theorem 2. Using the Kolmogorov Transformations we have taken the Hamiltonian in the region $(p, q) \, \varepsilon S(r_0, \varrho_0)$ and transformed it into a Hamiltonian of the same form in the region $(P, Q) \, \varepsilon S(r_0 - 2h, \varrho_0 - 4h)$. We can now keep repeating the process as many times as we wish. If we define $r_{n+1} = r_n - 2h_{n+1}$, $\varrho_{n+1} = \varrho_n - 4h_{n+1}$; at the $n$th stage we will have a Hamiltonian defined for $(P^{(n)}, Q^{(n)}) \, \varepsilon S(r_n, \varrho_n)$, with $\varepsilon_n = \max (\|A^{(n)}\|_{S(\varrho_n)}, \|B_i^{(n)}\|_{S(\varrho_n)})$.

We now wish to show that we can make Lemma 2 applicable at each stage of this process. To begin with, let us observe that the set of equations

$$x_{n+1} = (L)^{n+1} x_n^2 \quad n = 0, 1, 2, \ldots$$

has the solution:

$$x_n = \frac{(L^2 x_0)^{2^n}}{L^{n+2}}, \quad n = 1, 2, \ldots. \tag{10}$$

This can easily be verified by induction.

We apply this remark to Lemma 2 especially (H3) and (A2), which in the present notation are

$$\varepsilon_n \leqslant \frac{h_{n+1}^{2t+3}}{C_0}, \ \varepsilon_{n+1} \leqslant \frac{C_1}{h_{n+1}^{3t+3}} (\varepsilon_n)^2.$$

Thus, if we choose an $L > 1$ such that

$$\frac{\max(C_0, C_1)}{h_{n+1}^{3t+3}} \leqslant L^{n+1}, \tag{11}$$

and then $\varepsilon_0$ such that $(L^2 \varepsilon_0) = a < 1$, (10) and (A2) imply that $|\varepsilon_n| \leqslant a^{2^n}$ and (10) shows (H3) is satisfied since

$$\varepsilon_n < \frac{1}{L^{n+2}} < \frac{1}{L^{n+1}} \leqslant h_{n+1}^{3t+3} / C_0 .$$

To satisfy (11), we set $h_n = \delta / 2^n$ and $L \geqslant (2^{3t+3} / \delta^{3t+3}) \max(C_0, C_1)$ and $L > 1$.

We now determine $\delta > 0$, so that certain conditions are fulfilled. Assume that in the original Hamiltonian

$$|c_{ij}^{-1}(0)| \leqslant N, \ \|C_{ij}\|_{S(\varrho_0)} \leqslant M, \ \|D(p, q)\|_{S(r_0, \varrho_0)} \leqslant M .$$

Next we note that by the change of variables $p_i = R_0 p_i'$, $H = R_0 H'$, $q_i = q_i'$, we may always assume that our Hamiltonian at first converges for $|p_i| \leqslant 1$ or equivalently we may assume $r_0 = 1$.

Since $\sum_{n=1}^{\infty} h_n = \delta$, we now choose $\delta > 0$, but sufficiently small, to accomplish all of the following:

(a) $$r_N = r_0 - 2 \sum_{n=1}^{N} h_n \geqslant r_0 - 2\delta \geqslant 3r_0/4$$

(b) $$\varrho_N = \varrho_0 - 4 \sum_{n=1}^{N} h_n \geqslant \varrho_0 - 4\delta \geqslant 3\varrho_0/4$$

(c) $$\|C_{ij}^{(n)}\|_{S(\varrho_n)} \leqslant M + C_2 \delta \leqslant 2M$$

(d) $$\|D^{(n)}\|_{S(r_n, \varrho_n)} \leqslant M + C_3 \delta \leqslant 2M$$

(e) $$|c_{ij}(0) - c_{ij}^{(n)}(0)| \leqslant \left( \frac{2C_4}{\varrho_0} + C_5 \right) \delta .$$

Hence by continuity $\delta$ may be chosen so $|d_{ij}^{(n)}| \leqslant 2N$, with $d_{ij}^{(n)}$ the inverse matrix of $c_{ij}^{(n)}$ (0).

Thus we may continually apply Lemma 2, and have $\varepsilon_n \to 0$ in $S(3r_0/4, 3\varrho_0/4)$.
Finally, using the notation

$$(P^{(n)}, Q^{(n)}) = W^{(n)}(P^{(n+1)}, Q^{(n+1)})$$

for the mapping obtained at the $n$th stage of the iteration corresponding to the $F_i$, $G_i$ of assertion (A1) of Lemma 2, we now describe the limit of these mappings and show it fulfills the conditions of Theorem 2.

We define the mapping

$$\tilde{W}^{(n)}(P, Q) = W^{(1)} \text{o} W^{(2)} \text{o} \ldots \text{o} W^{(n)}(P, Q).$$

Since $W^{(n)}$: $S(r_{n+1}, \varrho_{n+1}) \to S(r_n, \varrho_n)$, the domain of each $\tilde{W}^{(n)}$ includes $S(3r_0/4, 3\varrho_0/4)$, and its range is in $S(r_0, \varrho_0)$. Thus the analytic functions composing the mappings $\tilde{W}^{(n)}(P, Q)$ are uniformly bounded on $S(3r_0/4, 3\varrho_0/4)$ and a subsequence of the $\tilde{W}^{(n)}$ converges to a limiting mapping $W(P, Q)$. It can be shown that $W(P, Q)$ is unique, by using (A1) to show that the sequence $\tilde{W}^{(n)}(P, Q)$ is Cauchy however the present proof does not require this.

Siegel [11, p. 9–10] shows that a transformation is canonical if and only if its Jacobian $\sigma$ satisfies the equation

$$\sigma' J \sigma = J \tag{12}$$

where $\sigma'$ is the transpose of $\sigma$, and

$$J = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} \text{ where } I \text{ is the unit matrix of degree } n.$$

Since the $W^{(n)}$ are canonical it follows from (12) that the $\tilde{W}^{(n)}$ and $W$ are also canonical. We let $W$ be the transformation (2) of Theorem 2. By substituting (2) into (1), the Hamiltonian is brought to the form (3). This completes the proof of Theorem 2.

## Acknowledgement

This research was supported in part by System Development Corporation and in part by NSF Grant GP-9384.

## References

[1] V. I. Arnold: 1963, 'Proof of a Theorem of A. N. Kolmogorov of the Invariance of Quasi-Periodic Motions Under Small Perturbations of the Hamiltonian', *Russ. Math. Surv.* **18**, 9–36. Transl. of *Usp. Mat. Nauk.* **18**.

[2] V. I. Arnold: 1963, 'Small Denominators and Problems of Stability of Motion in Classical and Celestial Mechanics', *Russ. Math. Surv.* **18**, 85–193. Transl. of *Usp. Mat. Nauk.* **18**.

[3] R. B. Barrar: 1966, 'A Proof of the Convergence of the Poincaré-Von Zeipel Procedure in Celestial Mechanics', *Am. J. Math.* **88**, 206–220.

[4] H. Cartan: 1963, *Elementary Theory of Analytic Functions of One or Several Complex Variables,* Addison-Wesley.

[5] A. N. Kolmogorov: 1954, 'The Conservation of Conditionally Periodic Motions with a Small Change in the Hamiltonian', *Dokl. Akad. Nauk SSSR* **98**, 527–530.

[6] J. Moser: 1962, 'On Invariant Curves of Area-Preserving Mappings of an Annulus', *Nachr. Akad. Wiss. Göttingen* No. **1**.
[7] J. Moser: 1966, 'A Rapidly Convergent Iteration Method and Non-Linear Partial Differential Equations: I', *Ann. Scuola Norm. Sup. Pisa* **20**, 265–315.
[8] J. Moser: 1966, 'A Rapidly Convergent Iteration Method and Non-Linear Differential Equations: II', *Ann. Scuola Norm. Sup. Pisa*, **20**, 499–533.
[9] J. Moser: 1967, 'Convergent Series Expansions for Quasi-Periodic Motions', *Math. Ann.* **169**, 136–176.
[10] H. Poincaré: 1957, *Méthodes nouvelles de la mécanique céleste*, 3 volumes, Dover Reprint.
[11] C. L. Siegel: 1956, *Vorlesungen über Himmelsmechanik*, Springer-Verlag.
[12] C. L. Siegel: 1942, 'Iteration of Analytic Functions', *Ann. Math.* **43**, 607–612.
[13] C. L. Siegel: 1952, 'Über die Normalform analytischer Differentialgleichungen in der Nähe einer Gleichgewichtslösung', *Nachr. Akad. Wiss. Göttingen Math.-Phys. Kl.* Nr. 5, 21–30.