Journal of Statistical Physics, Vol. 45, Nos. 5/6, 1986

Improved KAM Estimates for the Siegel Radius

C. Liverani[1] **and G. Turchetti**[2]

Received August 6, 1985; revision received August 27, 1986

For the Siegel center problem we explore the possibility of improving the KAM estimates, with a view to possible extensions to Hamiltonian systems. The use of a suitable norm and explicit perturbative computations allow estimates to within a factor 2 of the Siegel radius for the quadratic map.

KEY WORDS: Dynamical systems; KAM estimates; Siegel center problem.

1. INTRODUCTION

The "stochastic transition" is a crucial aspect of nonlinear mechanics for many of its applications.(1) However, one must rely upon empirical procedures, such as the direct inspection of the orbits and the computation of Liapunov indices,(2) in order to obtain any quantitative information about it.

No rigorous and computationally accurate tool is available. The KAM(3) theorem indeed allows one to determine a bound for the perturbation strengths for which invariant tori are preserved, but the values currently obtained are many orders of magnitude below the expected one. As a consequence, it is of interest to see whether the KAM estimates can be improved until they reach the correct order of magnitude.

There are several ways of improving the KAM estimates: the use of more suitable norms, the numerical evaluation of a finite number of steps in the iterative procedure, and the introduction of perturbative information.

In view of the technical difficulties involved, we investigate the simplest small denominator problem to which the KAM method applies, i.e., the

---

[1] Department of Mathematics, Rutgers University, New Brunswick, New Jersey.
[2] Dipartimento di Fisica, Università di Bologna, and Istituto Nazionale di Fisica Nucleare, Sezione di Bologna, Bologna, Italy.

0022-4715/86/1200-1071$05.00/0

822/45/5-6-20

Siegel center problem.[4] In this case there are orbits conjugated to a rotation within an open disc (Siegel domain) of radius r_s. When the winding number ω is the golden mean, the classical estimate of r_s given by Moser is five orders of magnitude too small for a quadratic nonlinearity.

An estimate of the correct order of magnitude was given by de la Llave,[5] who adapted to this problem a method proposed by Herman.[6] However, this method applies only to constant type winding numbers.

In a previous note[7] we showed that, with suitable changes in Moser's proof, which include geometric convergence rather then superconvergence, and with the explicit calculation of a finite number of iterations, the estimate of the Siegel radius for ω equal to the golden mean with a quadratic perturbation differs by one order of magnitude from the expected value. However, this result was obtained by neglecting the truncation errors in the perturbative computation and consequently is not rigorous.

In the present work we modify the technique to include the truncation errors while keeping the previous improvements. The only reason the numerical part of the paper cannot be considered a "proof" is because of the possibility of machine roundoff errors. This problem could have been overcome through the introduction of the "interval arithmetic,"[11] but we felt the effort disproportionate to the goals of the present work and we contended ourselves with a careful study of the numerical stability of our results. Comparison is made with the value of r_s computed by transforming conformally the innermost trajectory of the critical points of the map.

2. THE SIEGEL CENTER PROBLEM

Suppose we are given a function $f(z)$ analytic in a neighborhood of the origin with $f(0) = f'(0) = 0$ and consider the mapping F defined by

$$z' = F(z) \equiv az + f(z) \tag{2.1}$$

where a has unit modulus

$$a = e^{2\pi i\omega}, \qquad \omega \in \mathbf{R} \tag{2.2}$$

In order to analyze the stability of the origin, we look for an analytic function Ψ conjugating F with its linear part, namely

$$z = \Psi(\zeta) \equiv \zeta + \psi(\zeta), \qquad \psi(0) = 0, \qquad \psi'(0) = 0 \tag{2.3}$$

such that

$$\zeta' = a\zeta \tag{2.4}$$

The function ψ satisfies the functional equation

$$\psi(a\zeta) - a\psi(\zeta) = f(\zeta + \psi(\zeta)) \tag{2.5}$$

A theorem by Siegel states that a unique analytic solution of (2.5) exists provided that ω is a diophantine number, that is, provided that $\gamma > 0$, $\mu \geqslant 1$ exist such that

$$|a^n - 1|^{-1} \leqslant \gamma n^\mu \tag{2.6}$$

In order to estimate the radius of convergence of $\psi(\xi)$, we follow the KAM procedure. Rather than solving (2.5), we solve the linearized homologic equation

$$\varphi(a\zeta) - a\varphi(\zeta) = f(\zeta) \tag{2.7}$$

The mapping $F_1 = \Phi^{-1} F \Phi$, where $\phi(\zeta) = \zeta + \varphi(\zeta)$, is

$$\zeta' = F_1(\zeta) = a\zeta + f_1(\zeta) \tag{2.8}$$

where f_1 satisfies the functional equation

$$f_1(\zeta) = \varphi(a\zeta) - \varphi(a\zeta + f_1(\zeta)) + f(\zeta + \varphi(\zeta)) - f(\zeta) \tag{2.9}$$

The procedure is then iterated and a sequence of maps F_n analytic in discs of radius r_n such that $F_n(\zeta) \to a\zeta$ and $r_n \to r_\infty \leqslant r_s$ as $n \to \infty$ is obtained.

In order to carry out the basic estimates, we introduce the norms

$$\|f\|_r^\infty = \operatorname*{Max}_{|z| \leqslant r} |f(z)| \tag{2.10}$$

and

$$\|f\|_r = \sum_n |f_n| \, r^n \tag{2.11}$$

where f_n are the coefficients of the Taylor expansion of $f(z)$, namely $f(z) = \sum_n f_n z^n$. The following inequality obviously holds:

$$\|f\|_r^\infty \leqslant \|f\|_r \tag{2.12}$$

We assume that $f(z)$ is analytic in the disc $|z| \leqslant R$ and define $\delta(r)$ according to

$$\|f'\|_r = \sum_{n \geqslant 2} n \, |f_n| \, r^{n-1} \leqslant \delta(r), \qquad r \leqslant R \tag{2.13}$$

The estimates for the norm of φ, solution of Eq. (2.7), and its derivative, in a disc of radius $r(1-\vartheta)$ where $0<\vartheta<1$, are readily obtained

$$\|\varphi\|_{r(1-\vartheta)}=\sum_{n\geqslant 2}\frac{|f_n|}{|a^n-a|}r^n(1-\vartheta)^n\leqslant\gamma\,\delta r\,\operatorname*{Max}_{n\geqslant 2}(1-\vartheta)^n\frac{(n-1)^\mu}{n}$$

$$\leqslant\gamma\,\delta r\begin{cases}(1-\vartheta)^{\bar n}\dfrac{(\bar n-1)^\mu}{\bar n}, & \vartheta<1-e^{1/2-\mu}\\ \dfrac{1}{2}(1-\vartheta)^2, & \vartheta\geqslant 1-e^{1/2-\mu}\end{cases}\tag{2.14}$$

where

$$\bar n=\frac{1}{2}\left\{1-\frac{\mu-1}{\log(1-\vartheta)}+\left[\left(1-\frac{\mu-1}{\log(1-\vartheta)}\right)^2-\frac{4}{\log(1-\vartheta)}\right]^{1/2}\right\}$$

and

$$\|\varphi'\|_{r(1-\vartheta)}\leqslant\gamma\delta\operatorname*{Max}_{n\geqslant 2}(1-\vartheta)^{n-1}(n-1)^\mu$$

$$\leqslant\gamma\delta\begin{cases}e^{-\mu}\mu^\mu/[-\log(1-\vartheta)]^\mu, & \vartheta<1-e^{-\mu}\\ 1-\vartheta, & \vartheta\geqslant 1-e^{-\mu}\end{cases}\tag{2.15}$$

For the purpose of analytic manipulation it is convenient to simplify the estimates (2.14) and (2.15) even with some loss of accuracy. The "rough estimates" read

$$\|\varphi\|_{r(1-\vartheta)}\leqslant\gamma\,\delta r\left(\frac{\mu-1}{e\vartheta}\right)^{\mu-1}\tag{2.16}$$

$$\|\varphi'\|_{r(1-\vartheta)}\leqslant\gamma\delta\left(\frac{\mu}{e\vartheta}\right)^{\mu}\tag{2.17}$$

In the previous formulas δ stands for $\delta(r)$ everywhere.

3. SINGLE STEP ESTIMATES

In order to obtain the basic estimates on $f_1(\zeta)$ we first quote the following lemma.

Lemma 1. If $f(\zeta)$ is analytic in a neighborhood of the disc $|\zeta|<\rho+M$ and $\varphi(\zeta)$ and $\psi(\zeta)$ are analytic for $|\zeta|<\rho$ with $\|\varphi\|_\rho<M$ and $\|\psi\|_\rho<M$, then

$$\|f(\zeta+\varphi(\zeta))-f(\zeta+\psi(\zeta))\|_\rho\leqslant\|f'\|_{\rho+M}\,\|\varphi-\psi\|_\rho\tag{3.1}$$

The proof is given in Appendix A.

The main result is then summarized by the following theorem.

Theorem 1. If $f(\zeta)$ is analytic in a disc of radius R and $\|f'\|_r \leqslant \delta(r)$ for $r < R$, then $f_1(\zeta) = \phi^{-1}F\phi(\zeta) - \zeta a$ exists and is analytic in any disc of radius $r(1-\sigma-\eta)$, where σ and η are constrained by

$$\|\varphi\|_{r(1-\sigma)}/r \leqslant \sigma, \qquad \|\varphi'\|_{r(1-\sigma)} < 1 \tag{3.2}$$

and

$$\eta \geqslant \eta_0 = \frac{\|\varphi\|_{r(1-\sigma)}/r}{1 - \|\varphi'\|_{r(1-\sigma)}}\, \delta(r(1-\sigma) + \|\varphi\|_{r(1-\sigma)}) \tag{3.3}$$

provided that

$$\sigma < 1 - \eta \tag{3.4}$$

The estimate on $f'(\zeta)$ holds in a disc of radius $r_1 = r(1-\sigma-\eta-\hat{\sigma})$, where $\hat{\sigma}$ is arbitrary, provided that $\sigma + \eta + \hat{\sigma} < 1$, and reads

$$\|f_1'\|_{r_1} < \eta/2\hat{\sigma} \tag{3.5}$$

Proof. Let $B_r(M)$ denote the set of functions $f(\zeta)$ analytic in a disc $|\zeta| < r$, such that $\|f\|_r \leqslant M$. The hypothesis of the theorem can be written $f' \in B_r(\delta)$, $\varphi \in B_{r(1-\sigma)}(r\sigma)$, $\varphi' \in B_{r(1-\sigma)}(\tau)$ for $\tau < 1$.

We first show that the solution of the functional equation (2.9), $\phi \circ F_1 = F \circ \phi$, exists and is unique for $f_1 \in B_{r(1-\sigma-\eta)}(r\eta)$, provided that η satisfies (3.3) and (3.4). Consider the mapping

$$Th(\zeta) = \varphi(a\zeta) - \varphi(a\zeta + h(\zeta)) + f(\zeta + \varphi(\zeta)) - f(\zeta)$$

and require that it map $B_{r(1-\sigma-\eta)}(r\eta)$ into itself. Using (3.1) with $\psi = 0$, we find

$$\begin{aligned}
\|Th\|_{r(1-\sigma-\eta)} &\leqslant \|\varphi'\|_{r(1-\sigma-\eta)+\|h\|_{r(1-\sigma-\eta)}} \|h\|_{r(1-\sigma-\eta)} \\
&\quad + \|f'\|_{r(1-\sigma-\eta)+\|\varphi\|_{r(1-\sigma-\eta)}} \|\varphi\|_{r(1-\sigma-\eta)} \\
&\leqslant \|\varphi'\|_{r(1-\sigma)} r\eta + \|f'\|_{r(1-\sigma)+\|\varphi\|_{r(1-\sigma)}} \|\varphi\|_{r(1-\sigma)} \\
&\leqslant \|\varphi'\|_{r(1-\sigma)} r\eta + \delta(r(1-\sigma) + \|\varphi\|_{r(1-\sigma)}) \|\varphi\|_{r(1-\sigma)} \leqslant r\eta
\end{aligned}$$

if (3.3) holds. The condition for T to be a contraction in $B_{r(1-\sigma-\eta)}(r\eta)$ is given by

$$\begin{aligned}
\|T(h_1 - h_2)\|_{r(1-\sigma-\eta)} &= \|\varphi(a\zeta + h_1(\zeta)) - \varphi(a\zeta + h_2(\zeta))\|_{r(1-\sigma-\eta)} \\
&\leqslant \|\varphi'\|_{r(1-\sigma-\eta)+\max\{\|h_1\|_{r(1-\sigma-\eta)}, \|h_2\|_{r(1-\sigma-\eta)}\}} \\
&\quad \times \|h_1 - h_2\|_{r(1-\sigma-\eta)} \leqslant \|\varphi'\|_{r(1-\sigma)} \|h_1 - h_2\|_{r(1-\sigma-\eta)}
\end{aligned}$$

and agrees with (3.2).

As a consequence, if the conditions of the theorem are satisfied, T has a unique fixed point in $B_{r(1-\sigma-\eta)}(r\eta)$, which is the solution of $\phi F_1 = F\phi$.

Denoting by $C[r]$ the disc $|\zeta| \leqslant r$, we now prove that $F_1 = \phi^{-1} \circ F \circ \phi$ in the disc $C[r(1-\sigma-\eta)]$. Let us first recall that $\phi(\zeta) = \zeta + \varphi(\zeta)$ is bijective in $C[r(1-\sigma)]$, where $\|\varphi'\|_{r(1-\sigma)} < 1$. It is sufficient to observe that

$$\begin{aligned} |z_1 - z_2| = |\phi(\zeta_1) - \phi(\zeta_2)| &\geqslant |\zeta_1 - \zeta_2| - \left| \int_{\zeta_1}^{\zeta_2} \varphi'(\zeta)\, d\zeta \right| \\ &\geqslant (1 - \|\varphi'\|^{\infty}_{r(1-\sigma)})\, |\zeta_1 - \zeta_2| \geqslant (1 - \|\varphi'\|_{r(1-\sigma)})\, |\zeta_1 - \zeta_2| \end{aligned}$$

Since $\varphi' \in B_{r(1-\sigma)}(\tau)$ for some $\tau < 1$, then $\phi(C[r(1-\sigma)])$ is within the domain of ϕ^{-1}. We can also notice that $\phi(C[r(1-\sigma-\eta)]) \subseteq C[r]$ where F is defined; $F_1(C[r(1-\sigma-\eta)]) \subseteq C[r(1-\sigma)]$, where ϕ is defined so that the compositions $F \circ \phi$ and $\phi \circ F_1$ are allowed. Moreover,

$$F \circ \phi(C[r(1-\sigma-\eta)]) = \phi \circ F_1(C[r(1-\sigma-\eta)]) \subseteq \phi(C[r(1-\sigma)])$$

that is, the image of $C[r(1-\sigma-\eta])$ is within the domain of ϕ^{-1} so that $\phi^{-1} \circ F \circ \phi$ is well defined there.

The last step consists in estimating the norm of $f_1'(\zeta)$ in a smaller disc $C[r_1]$ where $r_1 = r(1-\sigma-\eta-\hat{\sigma})$, where $\hat{\sigma}$ is arbitrary, provided that $\sigma + \eta + \hat{\sigma} < 1$. Defining $x = \hat{\sigma}/(1-\sigma-\eta)$, we have

$$\begin{aligned} \|f_1'\|_{r_1} &= \sum_{n \geqslant 2} n\, |f_{1,n}|\, r^{n-1} (1-\sigma-\eta-\hat{\sigma})^{n-1} \\ &= \frac{\|f_1\|_{r(1-\sigma-\eta)}}{r\hat{\sigma}}\, x \operatorname*{Max}_{n \geqslant 2} n(1-x)^{n-1} \\ &\leqslant \frac{\|f_1\|_{r(1-\sigma-\eta)}}{r\hat{\sigma}} \begin{cases} \dfrac{e^{-1}}{1-x} \dfrac{x}{-\log(1-x)}, & e^{-1/2} < 1-x < 1 \\ 2x(1-x), & 0 < 1-x < e^{-1/2} \end{cases} \end{aligned}$$

Since the maximum of the function defined by the curly brackets is 1/2, the final estimate (3.5) follows.

Remark 1. Let us observe that the condition (3.2) on the domains can be conveniently satisfied by choosing

$$\sigma = (\gamma\delta)^{1/\mu}\, \mu e^{1/\mu - 1} \tag{3.6}$$

Indeed, by using the estimates (2.16) and (2.17), we obtain

$$\frac{\|\varphi\|_{r(1-\sigma)}}{r} \leqslant \sigma \frac{1}{\mu} \left(\frac{\mu - 1}{\mu} \right)^{\mu - 1} \leqslant \sigma$$

and

$$\|\varphi'\|_{r(1-\sigma)} \leqslant e^{-1} < 1$$

As far as condition (3.4) is concerned, it is easy to see that if

$$\delta < \gamma^{-1} e^{-2} (e/\mu)^{\mu}$$

then (3.4) follows from (3.2) and (3.3) once the choice (3.6) is made.

Remark 2. One could state Theorem 1 by using the norm $\|\cdot\|^{\infty}$ rather than $\|\cdot\|$. After replacing $\|\cdot\|$ with $\|\cdot\|^{\infty}$ in (3.2) and (3.3), one need only change (3.5) into

$$\|f'\|_{r_1}^{\infty} \leqslant \frac{\mu}{2\hat{\sigma} - \hat{\sigma}^2/(1-\sigma-\eta)} \leqslant \frac{\eta}{\hat{\sigma}}$$

Since $\|\cdot\|^{\infty} \leqslant \|\cdot\|$, one might think that this formulation would give better results. This is not true, because the estimates of $\|\cdot\|^{\infty}$ are much worse with respect to $\|\cdot\|$. Compare, for instance, for a quadratic irrational winding number ($\mu = 1$) the norms of φ and φ'. From Ref. 7 we have

$$\|\varphi\|_{r(1-\sigma)}^{\infty} \leqslant r\gamma\delta \frac{(1-\sigma)^2}{\sigma}, \qquad \|\varphi'\|_{r(1-\sigma)}^{\infty} \leqslant \gamma\delta \frac{1-\sigma}{\sigma^2}$$

whereas (2.16) and (2.17) give

$$\|\varphi\|_{r(1-\sigma)} \leqslant r\gamma\delta, \qquad \|\varphi'\|_{r(1-\sigma)} \leqslant \gamma\delta/e\sigma$$

4. ITERATIONS

Given a disc of radius r_0 included in the region where $f(z)$ is analytic with a derivative whose norm (2.11) is bounded by $\delta_0 \equiv \delta(r_0)$, we can define $f_1(z)$ and bound its derivative by δ_1 according to Theorem 1 in a disc of radius $r_1 = r_0(1 - \sigma_0 - \eta_0 - \hat{\sigma}_0)$.

The process can be iterated and at the step $n+1$ we define $f_{n+1}(z)$ and bound its derivative by δ_{n+1} in a disc of radius $r_{n+1} = r_n(1 - \sigma_n - \eta_n - \hat{\sigma}_n)$.

One must then prove that r_n converges to a positive limit r_∞ as $n \to \infty$ and that δ_n converges to zero.

The sequence of the function $\Psi_n(z) = \Phi_1 \circ \Phi_2 \cdots \circ \Phi_n(z)$ must be shown to converge in $|z| < r_\infty$ to an analytic function $\Psi(z)$ solution of (2.5).

Analytic estimates of r_∞ can be obtained by using the "rough" estimates (2.16) and (2.17) for the norms of φ and φ'. Sharper estimated can be obtained by iterating the process numerically until an order N such that $\delta_N < \varepsilon$, for some given ε, and then using only the analytic estimate for the remaining (infinite) sequence of iterations.

The further improvements obtained by introducing perturbative information will be discussed in Section 5.

Theorem 2. If $f(z)$ is analytic for $|z| < R$ and $\|f'\|_{r_0} < \delta(r_0)$ for $r_0 < R$, then the map (2.1), subject to condition (2.6), is conjugated to its linear part (2.4) in a disc of radius r_∞ given by

$$r_\infty = r_0(1 - \sigma_0 - \eta_0 - \hat{\sigma}_0)^{1/(1-\chi^{1/\mu})} \tag{4.1}$$

where

$$\sigma_0 = (\gamma\delta_0)^{1/\mu}\,\alpha, \qquad \eta_0 = (\gamma\delta_0)^{1/\mu}\,\delta_0\beta, \qquad \hat{\sigma}_0 = \frac{\eta_0}{2\chi\delta_0} = (\gamma\delta_0)^{1/\mu}\frac{\beta}{2\chi} \tag{4.2}$$

and α and β are two constants defined by

$$\alpha = \mu e^{1/\mu - 1} \tag{4.3}$$

$$\beta = \frac{e^{1/\mu - 1}}{1 - e^{-1}}\left(\frac{\mu - 1}{\mu}\right)^{\mu - 1} \tag{4.4}$$

for any $r_0 \in\,]0, R[$ and $\chi \in\,]0, 1[$ such that $\sigma_0 + \eta_0 + \hat{\sigma}_0 < 1$. The maximum of r_∞ with respect to r_0 and χ provides the best estimate of the conjugacy radius.

The sequence $\Psi_n(\zeta) = \Phi_1 \circ \Phi_2 \cdots \circ \Phi_n(\zeta)$ has a limit $\Psi(z)$ for $n \to \infty$ analytic in $|\zeta| < r_\infty$.

Proof. According to Theorem 1, supposing that $f_n(\zeta)$ is analytic in $|\zeta| < r_n$, with $\|f'\|_{r_n} < \delta_n$, then $f_{n+1}(\zeta)$ will be analytic in $|\zeta| < r_{n+1} = r_n(1 - \sigma_n - \eta_n - \hat{\sigma}_n)$, with $\|f'_{n+1}\|_{r_{n+1}} < \eta_n/2\hat{\sigma}_n = \delta_{n+1}$.

We choose σ_n according to (3.6),

$$\sigma_n = (\gamma\delta_n)^{1/\mu}\,\mu e^{1/\mu - 1}$$

and η_n equal to an upper bound of the rhs of (3.3) obtained by replacing the square brackets with r and $\|\varphi_n\|_{r_n(1-\sigma_n)}$ and $\|\varphi'_n\|_{r_n(1-\sigma_n)}$ according to (2.16) and (2.17), so that

$$\eta_n = \delta_n(\gamma\delta_n)^{1/\mu}\frac{e^{1/\mu - 1}}{1 - e^{-1}}\left(\frac{\mu - 1}{\mu}\right)^{\mu - 1}$$

Choosing for δ_n a geometric convergence, we can write

$$\delta_{n+1} = \chi\delta_n$$

for $0<\chi<1$, so that the previous equations lead to

$$\begin{aligned}\delta_n&=\chi^n\delta_0, & \sigma_n&=\chi^{n/\mu}\sigma_0\\ \eta_n&=\chi^{n(1+1/\mu)}\eta_0, & \hat{\sigma}_n&=\frac{\eta_n}{2\chi\delta_n}=\chi^{n/\mu}\hat{\sigma}_0\end{aligned} \tag{4.5}$$

where σ_0, η_0, and $\hat{\sigma}_0$ are given by (4.2)–(4.4).

Since σ_n, η_n, and $\hat{\sigma}_n$ are decreasing sequences, $\sigma_0+\eta_0+\hat{\sigma}_0<1$ implies $\sigma_n+\eta_n+\hat{\sigma}_n<1$ for any $n\geqslant 1$. In order to obtain r_∞, we observe that $\eta_n\leqslant\chi^{n/\mu}\eta_0$, so we can write

$$\begin{aligned}r_\infty&=r_0\prod_{n=0}^{\infty}(1-\sigma_n-\hat{\sigma}_n-\eta_n)\\ &\geqslant r_0\exp\left\{\sum_{n=0}^{\infty}\log[1-\chi^{n/\mu}(\sigma_0+\eta_0+\hat{\sigma}_0)]\right\}\\ &=r_0\exp\left\{-\sum_{k=1}^{\infty}\frac{1}{k}\frac{(\sigma_0+\eta_0+\hat{\sigma}_0)^k}{1-\chi^{k/\mu}}\right\}\\ &\geqslant r_0(1-\sigma_0-\eta_0-\hat{\sigma}_0)^{\frac{1}{1-\chi^{1/\mu}}}\end{aligned}$$

Finally, we consider the sequence $\Psi_n=\Phi_1\circ\Phi_2\cdots\circ\Phi_n$, where

$$z\equiv\zeta_0=\Phi_1(\zeta_1)=\zeta_1+\varphi_1(\zeta_1),\qquad \zeta_{n-1}=\Phi_n(\zeta_n)=\zeta_n+\varphi_n(\zeta_n)$$

and notice that Φ_n is defined in $|\zeta_n|\leqslant r_n$ and the image of this disc is within a disc of radius r_{n-1}, since

$$\|\Phi_n\|_{r_n}\leqslant r_n+\|\varphi_n\|_{r_n}\leqslant r_{n-1}(1-\sigma_{n-1})+\|\varphi_n\|_{r_{n-1}(1-\sigma_{n-1})}\leqslant r_{n-1}$$

As a consequence, the composition is correctly defined and

$$\zeta_0=\Phi_1\circ\Phi_2\circ\cdots\circ\Phi_n(\zeta_n)=\sum_{j=1}^{n}\varphi_j(\zeta_j)+\zeta_n$$

implies

$$\begin{aligned}\|\Phi_1\circ\cdots\circ\Phi_n-\zeta_n\|_{r_n}&\leqslant\sum_{j=1}^{n}\|\varphi_j\|_{r_j}\leqslant\sum_{j=1}^{n}\|\varphi_j\|_{r_{j-1}(1-\sigma_{j-1})}\\ &\leqslant\left(\frac{\mu-1}{e}\right)^{\mu-1}\gamma\sum_{j=1}^{n}r_{j-1}\frac{\delta_{j-1}}{\sigma_{j-1}^{\mu-1}}\leqslant r_0\sigma_0\frac{(\mu-1)^{\mu-1}}{\mu^\mu}\frac{1}{1-\chi^{1/\mu}}\end{aligned}$$

where (2.16) has been used, replacing r_{j-1} with the upper bound r_0 and δ_{j-1} and σ_{j-1} with (4.5), and taking (3.6) into account.

5. NUMERICAL RESULTS

For the quadratic map with a winding number equal to the golden mean ω we compare the previous estimates with the results obtained by transforming conformally the trajectory of the critical point.

Indeed, a theorem proved by Herman[12] states that for a quadratic polynomial map the Siegel disc is conjugated with a domain D whose boundary is the trajectory of a critical point. The conformal mapping from D to the Siegel disc is $\zeta = \Psi^{-1}(z)$, where Ψ is defined by (2.3) and (2.5), and bounds to the measure πr_s^2 of $\Psi^{-1}(D)$ can be obtained with polynomial approximations to $\Psi^{-1}(z)$, as described in Appendix B.

The approximations do rapidly converge and the result $r_s = 0.326$ agrees with the radius of convergence of the series $\Psi(\zeta)$ obtained by using the Hadamard criterion. The classical estimate by Moser, explicitly derived in Ref. 7, gives for the quadratic map the following lower bound to the Siegel radius:

$$r_\infty = 40^{-\mu-2}/5\gamma\Gamma(\mu+1) \tag{5.1}$$

and for the golden mean ($\mu = 1$, $\gamma = 0.53646$) one has $r_\infty = 5.8 \times 10^{-6}$.

The best result obtained with the new estimate given by (4.1), Theorem 2, is $r_\infty = 0.0486$ and corresponds to $r_0 = 0.1$, $\chi = 0.425$.

A significant improvement is obtained if we perform explicitly a finite number L of the iterations specified by Theorem 2, using for the norms $\|\phi\|_{r(1-\sigma)}$ and $\|\phi'\|_{r(1-\sigma)}$ the sharper estimates (2.14) and (2.15).

Table I. Values of the Norms of the Taylor Series for φ, φ′, and f' Truncated at Order $N = 128$ and Computed for the First Two Iterations, for Some Values of r [a]

Iteration	r	$\|\varphi^{(N)}\|_r$	$\|\varphi'^{(N)}\|_r$	$\|f'\|_r$
1	0.05	0.134×10^{-2}	0.536×10^{-1}	0.878×10^{-2}
	0.1	0.536×10^{-2}	0.107	0.385×10^{-1}
	0.15	0.121×10^{-1}	0.161	0.965×10^{-1}
	0.2	0.215×10^{-1}	0.215	0.189
	0.25	0.355×10^{-1}	0.268	0.332
2	0.05	0.109×10^{-3}	0.676×10^{-2}	0.917×10^{-4}
	0.1	0.960×10^{-3}	0.306×10^{-1}	0.173×10^{-2}
	0.15	0.355×10^{-2}	0.778×10^{-1}	0.106×10^{-1}
	0.2	0.927×10^{-2}	0.157	0.416×10^{-1}
	0.25	0.200×10^{-1}	0.282	0.132

[a] The results obtained for $N = 32$ are the same within the quoted accuracy.

After fixing the number L of explicit iterations with the condition $\delta_L < 2 \times 10^{-3}$ we still determine r_∞ by using (4.1) Theorem 2, where r_0 and δ_0 are replaced by r_L and δ_L. The best result is then $r_\infty = 0.102$ and corresponds to the choice $r_0 = 0.4$ and $\chi = 0.275$.

The last improvement we have made consists in performing some steps of the KAM iteration keeping information about the coefficients of the functions involved, not only about their norms. This technique was first used by Porzio[10] in a Hamiltonian context. If we denote by $f^{(N)}$ the truncation of the power series up to order N, the following estimates can be used in the place of (2.14), (2.15), and (3.5) at the ith step of the KAM iteration when $\mu = 1$:

$$\|\phi_i\|_{r_{i-1}(1-\sigma_{i-1})} \leqslant \|\phi_i^{(N)}\|_{r_{i-1}(1-\sigma_{i-1})} + r_{i-1}\gamma\delta_{i-1}(1-\sigma_{i-1})^N$$

$$\sigma_{i-1} \geqslant 1 - e^{-1/N} \approx 1/N \tag{5.2}$$

$$\|\phi_i'\|_{r_{i-1}(1-\sigma_{i-1})} \leqslant \|\phi_i'^{(N)}\|_{r_{i-1}(1-\sigma_{i-1})} + \gamma\delta_{i-1} N(1-\sigma_{i-1})^N$$

and

$$\|f_i'\|_{r_i} \leqslant \delta_i = \|f_i'^{(N)}\|_{r_i} + \frac{\eta_{i-1}}{\sigma_{i-1}}(N+1)\, x_{i-1}(1-x_{i-1})^N$$

$$x_{i-1} \geqslant 1 - e^{-1/(N+1)} \approx 1/N \tag{5.3}$$

where η_i is defined by (3.3) and $x_i = \hat{\sigma}_i/(1 - \sigma_i - \eta_i)$.

The truncated part can be computed with the help of a computer according to the procedure illustrated in Appendix C. As i increases, σ_i and $\hat{\sigma}_i$ decrease and the remainders grow with respect to the perturbative part; however, for the first two iterations they are almost negligible. The estimates (5.2) and (5.3) have been used for the first two steps with $N = 64$ and $N = 62$, respectively. The computation of the coefficients of the Taylor expansion for ϕ_i and f_i, $i = 1, 2$, is the only place where numerical methods are relevant for our results. Since we did not use "interval arithmetic techniques," roundoff errors may have affected, in principle, our computations. However, we first performed these computations on a HP-1000 machine with eight significant digits, then checked them on a CRAY and an IBM with 14 and 28 significant digits, respectively (the program was written in Fortran). As a result of the agreement obtained in this procedure, we feel that roundoff errors are completely negligible with regard to the quantities in which we are interested. Using as parameters $r_0 = 0.4$, $\hat{\sigma}_0 = \hat{\sigma}_1 = 0.05$, $\sigma_0 = 0.2$, $\sigma_1 = 0.07$, and $\chi = 0.3$, which are used in (4.1) to compute r_∞ starting from r_2 and δ_2, we obtain the value $r_\infty = 0.18$, about one-half the Siegel radius.

For $\mu > 1$ the new estimate is still considerably better than the classical one given by (5.1), but when $\mu \to \infty$, both behave as $\Gamma(\mu)^{-1}$.

6. CONCLUSION

We have succeeded in improving the KAM estimates for the Siegel problem, obtaining the correct order of magnitude for the Siegel conjugacy radius when the winding number has strong enough diophantine properties and the perturbation is a polynomial of degree 2. The basic ingredients of this improvement are the choice of a new norm, a modified strategy in the estimates of the domains, and a geometrically convergent rather than superconvergent iteration process. The choice of the norm (2.11) rather then (2.10) is particularly relevant, as can be seen by comparing the estimates for the solution of the homologic equation where the gain of $\|\phi\|_{r(1-\sigma)}$ compared to $\|\phi\|^{\infty}_{r(1-\sigma)}$ is essentially a factor $1/\sigma$. Finally the explicit numerical computation of a finite number of steps and the explicit introduction of perturbative information allows one to approach, in the case explicitly considered, the Siegel radius within a factor 2.

The present strategy can be extended to area-preserving maps and to Hamiltonian systems.

APPENDIX A

The proof of Lemma 1 is given.

Lemma 1. If $f(\zeta)$ is analytic in a neighborhood of the disc $|\zeta| < \rho + M$ and $\varphi(\zeta)$ and $\psi(\zeta)$ are analytic for $|\zeta| < \rho$ with $\|\varphi\|_\rho < M$ and $\|\psi\|_\rho < M$, then

$$\|f(\zeta + \varphi(\zeta)) - f(\zeta + \psi(\zeta))\|_\rho \leqslant \|f'\|_{\rho+M} \|\varphi - \psi\|_\rho$$

Proof. Let us introduce the Banach spaces

$$B_\rho = \{f\colon C(\rho) \to \mathbb{C}/f \text{ analytic, } \|f\|_\rho < \infty\}$$

where $C(\rho) = \{\zeta \in \mathbb{C}/|\zeta| < \rho\}$. Now one first notices that $\forall h, g \in B_\rho$, $hg \in B_\rho$, and, moreover, $\|hg\|_\rho \leqslant \|h\|_\rho \|g\|_\rho$ (this result follows immediately by direct series manipulation). The second thing to notice is that for each $f \in B_{\rho+M}$ and $h \in B_\rho$: $\|h\|_\rho < M$ the function $f(\zeta + h(\zeta))$ is an element of B_ρ, just because it is analytic in $C(\rho)$ and $\|f(\zeta + h(\zeta))\|_\rho \leqslant \|f\|_{\rho+M}$. From the preceding considerations, calling $\Gamma_r = \{\zeta \in \mathbb{C}/|\zeta| = r\}$, it follows that

$$f(\zeta + \varphi(\zeta)) - f(\zeta + \psi(\zeta))$$
$$= \frac{1}{2\pi i} \int_{\Gamma_{\rho+M}} \left(\frac{f(z)}{z - \zeta - \varphi(\zeta)} - \frac{f(z)}{z - \zeta - \psi(\zeta)} \right) dz$$

$$=\frac{\varphi(\zeta)-\psi(\zeta)}{2\pi i}\int_{\Gamma_{\rho+M}}\frac{f(z)}{[z-\zeta-\varphi(\zeta)][z-\zeta-\psi(\zeta)]}\,dz$$

$$=\frac{\varphi(\zeta)-\psi(\zeta)}{2\pi i}\sum_{n=0}^{\infty}\sum_{j=0}^{n}\left[\int_{\Gamma_{\rho+M}}\frac{f(z)}{z^{n+2}}\,dz\right][\zeta+\varphi(\zeta)]^{j}[\zeta+\psi(\zeta)]^{n-j}$$

$$=[\varphi(\zeta)-\psi(\zeta)]\left\{\sum_{n=0}^{\infty}\sum_{j=0}^{n}f_{n+1}(\zeta+\varphi(\zeta))^{j}(\zeta+\psi(\zeta))^{n-j}\right\}$$

Now the next step is to show that the function in curly brackets is an element of B_ρ; to do so, it is enough to prove that the series are absolutely convergent in the $\|\cdot\|_\rho$ sense; indeed,

$$\sum_{n=0}^{\infty}\sum_{j=0}^{n}|f_{n+1}|\;\|[\zeta+\varphi(\zeta)]^{j}\|_\rho\;\|[\zeta+\psi(\zeta)]^{n-j}\|_\rho$$

$$\leqslant\sum_{n=0}^{\infty}\sum_{j=0}^{n}|f_{n+1}|\;\|\zeta+\varphi(\zeta)\|_\rho^{j}\;\|\zeta+\psi(\zeta)\|_\rho^{n-j}$$

$$\leqslant\sum_{n=0}^{\infty}(n+1)\,|f_{n+1}|(\rho+M)^{n}=\|f'\|_{\rho+M}$$

As a consequence of all this, we get

$$\|f(\zeta+\varphi(\zeta))-f(\zeta+\psi(\zeta))\|_\rho$$

$$\leqslant\left\|\sum_{n=0}^{\infty}\sum_{j=0}^{n}f_{n+1}(\zeta+\varphi(\zeta))^{j}(\zeta+\psi(\zeta))^{n-j}\right\|_\rho\|\varphi-\psi\|_\rho$$

$$\leqslant\|f'\|_{\rho+M}\,\|\varphi-\psi\|_\rho \quad \blacksquare$$

APPENDIX B

We recall a result on the conformal mappings.[9] Let B be a simply connected set, with $0\in B$, and Φ an analytic function, univalent on B, such that $\Phi(0)=0$, $\Phi'(0)=1$, which maps B into the disc $C[r]$. The functional $I(f)$,

$$I(f)=\int_B|f'(z)|^2\,d\mu(z) \tag{B.1}$$

defined on all the functions $f(z)$ analytic in D such that $f(0)=0$ and $f'(0)=1$ is minimum for $f=\Phi$ and $I(\Phi)=\pi r^2$.

As a consequence, if $\mathscr{P}_n$ is the space of polynomials of degree n, one has

$$\pi r^2 = \lim_{n \to \infty} \min_{f \in \mathscr{P}_n} I(f) \tag{B.2}$$

Letting

$$\Gamma_{jk} = \int_B z^{*j} z^k \, d\mu(z), \qquad \gamma_j = \Gamma_{j,0} \tag{B.3}$$

and $\Gamma_{00} = \mu(B)$ be the measure of B, it is not hard to show that

$$\min_{f \in \mathscr{P}_n} I(f) = \pi r_n^2 = \mu(B) - \gamma^+ \Gamma \gamma \tag{B.4}$$

where Γ and γ denote the $(n-1) \times (n-1)$ matrix and $n-1$ vector whose components are Γ_j and γ_j for $j, k = 1,..., n-1$.

Indeed, if $f \in \mathscr{P}_n$, one can write

$$f(z) = z + \sum_{k=1}^{n-1} \frac{a_n}{n+1} z^{n+1}$$

so that

$$I(f) \equiv J(a, a^+) = \Gamma_{00} + a^+ \Gamma a + a^+ \gamma + \gamma^+ a \tag{B.5}$$

and the minimum of $I(f)$ for $f \in \mathscr{P}_n$ is the minimum of the bilinear functional $J(a, a^+)$ for $a, a^+ \in C^{n-1}$. The unique stationary point of J is found to occur for $a = -\Gamma^{-1}\gamma$, $a^+ = -\gamma^+ \Gamma^{-1}$.

Since the Siegel disc, according to Herman's theorem, is the image of the connected set whose boundary is the trajectory of the critical point, the inverse $\Psi(z)$ of such a mapping is univalent and $\Psi(0) = 0$ and $\Psi'(0) = 1$, so that the previous theorem applies with $\Phi = \Psi^{-1}$.

In order to evaluate the integrals on B numerically, one uses a trapezoidal rule on the trajectory of the critical point defined by a finite number M of iterations of the map.

APPENDIX C

In order to compute the Taylor series of the KAM iterates $f_N(z)$ where $f_0(z) \equiv f(z)$, we first observe that

$$f_N(z) = z^{2^N + 1} \sum_{n=0}^{\infty} \beta_n z^n \tag{C.1}$$

Defining

$$\varphi_{N+1}(z) = z^{2^N+1} \sum_{n=0}^{\infty} \eta_n z^n, \qquad \eta_n = \frac{\beta_n}{a^{n+2^N+1} - a} \tag{C.2}$$

and

$$f_{N+1}(z) = z^{2^{N+1}+1} \sum_{n=0}^{\infty} \gamma_n z^n \tag{C.3}$$

we provide an explicit algorithm to compute the γ_n, given the β_n. The functional equation (2.9) now reads

$$\begin{aligned} f_{N+1}(z) = \varphi_{N+1}(z) - \varphi_{N+1}(az + f_{N+1}(z)) \\ + f_N(z + \varphi_{N+1}(z)) - f_N(z) \end{aligned} \tag{C.4}$$

and its series expansion involves powers of $f_{N+1}(z)$ and $\varphi_{N+1}(z)$. It is convenient to write

$$\varphi^j_{N+1}(z) = z^{j(2^N+1)} \sum_{n=0}^{\infty} \eta^{(j)}_n z^n \tag{C.5}$$

and

$$f^j_{N+1}(z) = z^{j(2^{N+1}+1)} \sum_{n=0}^{\infty} \gamma^{(j)}_n z^n \tag{C.6}$$

where

$$\eta^{(j)}_l = \sum_{r=0}^{l} \eta^{(j-1)}_r \eta_{l-r} \tag{C.7}$$

and a similar relation holds for the $\gamma^{(j)}_n$.

The recurrence from β_n to γ_n is obtained after some tedious but straightforward series manipulations and reads

$$\begin{aligned} \gamma_n = & - \sum_{l=0}^{n-2^N} \sum_{j=1}^{j_{\max}} \binom{k+2^N+1}{j} a^{k+2^N+1-j} \eta_k \gamma^{(j)}_l \\ & + \sum_{l=0}^{n} \sum_{j=1}^{j'_{\max}} \binom{k'+2^N+1}{j} \beta_{k'} \eta^{(j)}_l \end{aligned} \tag{C.8}$$

where

$$k = n - 2^{N+1} j - l + 2^N, \qquad k' = n - 2^N j - l + 2^N \tag{C.9}$$

and

$$
\begin{aligned}
j_{\max}(n, l, N) &= \min\left\{\left[\frac{n-l+2^N}{2^{N+1}}\right], \left[\frac{n-l+2^{N+1}+1}{2^{N+1}+1}\right]\right\} \\
j'_{\max}(n, l, N) &= \min\left\{\left[\frac{n-l+2^N}{2^N}\right], \left[\frac{n-l+2^{N+1}+1}{2^N+1}\right]\right\}
\end{aligned}
\tag{C.10}
$$

The recurrence is well defined if, after computing γ_n, we evaluate $\gamma_l^{(j)}$ for $0 \leqslant l \leqslant n$ and $j = 0, 1,...$; indeed, γ_{n+1} involves only $\gamma_l^{(j)}$ with $l \leqslant n - 2^N \leqslant n-1$. If we decide to truncate all the series $f_N(z)$ at the order $2^{N_{\max}} + 1$ (then $f_{N_{\max}}(z)$ will consist of just a single term), the sum on the rhs of (C.3) will be truncated to $n_{\max} = 2^{N_{\max}} - 2^{N+1}$. The $\eta_l^{(j)}$ [see (C.7)] will be computed for $l \leqslant n_{\max}$, $j \leqslant j'_{\max}(n_{\max}, n, N)$, while for any $n \leqslant n_{\max}$, the $\gamma_l^{(j)}$ will be computed for $l \leqslant n$ and $j \leqslant j_{\max}(n_{\max}, n, N)$.

ACKNOWLEDGMENTS

This work was partially supported by grant DMR 81-14726-02. One of us (C. L.) thanks S. Berhanu for a helpful discussion about the content of Appendix A, A. Porzio for teaching him the perturbative technique to improve the KAM estimates, and R. de la Llave for very useful suggestions and discussions.

REFERENCES

1. G. Casati and J. Ford, *Lecture Notes in Physics*, No. 93 (Springer, 1979); E. G. D. Cohen, *Fundamental Problems in Statistical Mechanics* (North-Holland, 1980).
2. G. Benettin, L. Galgani, and A. Giorgilli, *Meccanica* **15**:9, 21 (1980).
3. V. I. Arnold, *Chapitres supplémentaires de la théorie des équations différentielles ordinaires* (Mir, Moscow, 1980), and references cited therein.
4. C. L. Siegel and J. K. Moser, *Lectures in Celestial Mechanics* (Springer, Berlin, 1971).
5. R. de la Llave, *J. Math. Phys.* **24**:2118 (1983).
6. M. Herman, *Astérisque* **103–104** (1983).
7. C. Liverani, G. Servizi, and G. Turchetti, *Lett. Nuovo Cimento* **39**:417 (1984).
8. N. S. Manton and M. Nauemberg, *Commun. Math. Phys.* **89**:555 (1983).
9. V. I. Smirnov, *A Course of Higher Mathematics* (Mir, Moscow), Vol. 3, p. 161.
10. A. Porzio, Stability bounds for the existence of KAM tori in a forced pendulum, Preprint, Universita' di Roma (1984), to be published in *Boll. UMI.*
11. O. E. Lanford III, A computer assisted proof of Feigenbaum's conjectures, *Bull. AMS* **47**:289 (1984); R. E. Moore, *Interval Analysis* (Prentice-Hall, 1966); *Methods and Applications of Interval Arithmetic* (SIAM, Philadelphia, 1979).
12. M. Herman, *Commun. Math. Phys.* **99**:593 (1985).