

## ACCURATE STRATEGIES FOR SMALL DIVISOR PROBLEMS

R. DE LA LLAVE AND DAVID RANA

### I. INTRODUCTION

Many interesting problems in mechanics are close to systems that can be solved exactly. In such systems it is natural to consider perturbative expansions about the known system. Frequently it is possible to compute all of the terms in the expansion recursively, which can lead to expressions involving denominators that become arbitrarily small. The convergence of these expansions is difficult to establish and sometimes even false. In the 1960s, Kolmogorov, Arnold and Moser developed a systematic method, known as K.A.M. theory, to deal with these small divisor problems.

An unfortunate limitation of the K.A.M. theory is that the range of validity of the perturbation expansions that are established are very small in comparison to physically relevant values (see [Mo] in this regard). We have considered the problem of systematically improving the values yielded by K.A.M. theory. In two examples discussed below, we have proven lower bounds for the range of validity which are 90% of values for which the results are known to be false.

Two classical theorems whose proofs present all of the essential complications of K.A.M. are

**Theorem I.1.** *Given a family  $f_\varepsilon$  of analytic functions on  $\mathbf{C}$ ,*

$$f_\varepsilon(z) = az + \frac{1}{\varepsilon} \hat{f}(\varepsilon z)$$

where  $\hat{f}(z) = \mathcal{O}(z^2)$ , and, for some  $C$ ,  $\nu > 0$ ,

$$|a| = 1, \quad |a^n - 1|^{-1} \leq Cn^\nu \quad \forall n > 0.$$

---

Received by the editors March 8, 1989.

1980 *Mathematics Subject Classification* (1985 Revision). Primary 39-04, 39B99, 70K50, 58F27; Secondary 65J15, 30D05, 58F30, 58F10.

The first author was supported in part by grant DMS 85-04984.

The second author was supported in part by grant DMS 88-02796.

Then  $\exists \varepsilon_0 > 0$ , s.t. if  $|\varepsilon| < \varepsilon_0$  there is a conformal change of variables  $\psi$  on the unit disk for which  $\psi(0) = 0$ ,  $\psi'(0) = 1$  and

$$(I.1) \quad \psi^{-1} \circ f_\varepsilon \circ \psi - f_0 = 0.$$

**Theorem I.2.** Let  $F_\varepsilon$  be a  $C^k$  family of  $C^k$  diffeomorphisms of  $\mathbf{R} \times \mathbf{S}^1$ ,

$$F_\varepsilon(x, \theta) = (f(x, \theta), \phi(x, \theta))$$

for which

- (i) the area  $dx \wedge d\theta$  is preserved,
- (ii)  $\phi_x(x, \theta) \geq c > 0$ ,
- (iii)  $\int (f(x, \theta) - x) d\theta = 0$ ,
- (iv)  $F_0(x, \theta) = (x, (\theta + x) \bmod 1)$ .

Let  $\omega \in \mathbf{R}$  satisfy

$$|n\omega - m|^{-1} \leq Cn^\nu n, \quad m \in \mathbf{Z}, n \neq 0, \quad C, \nu > 0, \quad 3\nu < k.$$

Then  $\exists \varepsilon_0 > 0$  s.t. if  $|\varepsilon| < \varepsilon_0$  there is a  $C^{k-3\nu}$  map  $K : \mathbf{S}^1 \mapsto \mathbf{R} \times \mathbf{S}^1$  whose image is a topologically nontrivial circle and for which

$$(I.2) \quad F_\varepsilon \circ K - K \circ T_\omega = 0$$

with  $T_\omega(\theta) = (\theta + \omega) \bmod 1$ .

Applying most versions of these theorems one can find in the literature to specific families, e.g.,

$$(I.3) \quad f_\varepsilon(z) = e^{2\pi i \gamma} z + \varepsilon z^2 \quad \gamma = \frac{\sqrt{5} - 1}{2}$$

or

$$(I.4) \quad F_\varepsilon(x, \theta) = \left( x + \frac{\varepsilon}{2\pi} \sin(2\pi\theta), \left( x + \theta + \frac{\varepsilon}{2\pi} \sin(2\pi\theta) \right) \bmod 1 \right)$$

we find that the  $\varepsilon$  for which the smallness conditions of the theorems (K.A.M. bounds) are met come out several orders of magnitude smaller than those relevant for physical situations. See [Mo] for a discussion of K.A.M. bounds in realistic applications.

We consider the question of improving the K.A.M. bounds. We find that systematic optimizations of some K.A.M. type proofs yield bounds converging to optimal. Other proofs have intrinsic limitations that do not allow such optimization.

## II. DEFINITIONS

The proofs of I.1 and I.2 involve infinitely many arbitrary choices such as analyticity losses, parameters entering in the choice of various norms, etc. At the expense of computational effort, these theorems can yield improving K.A.M. bounds by optimizing the choices.

**Definition II.1.** A strategy is an algorithm which, for a family  $\mathcal{F}_\varepsilon$  ( $f_\varepsilon$  of Theorem I.1 or  $F_\varepsilon$  of I.2), produces an  $\varepsilon_N > 0$  computed from  $\mathcal{F}_\varepsilon$  in  $\leq N$  algebraic operations and branches s.t. the conclusions of the theorems hold when  $|\varepsilon| \leq \varepsilon_N$ .

Some strategies are optimal in the following sense:

**Definition II.2.** A strategy is accurate when  $\lim_{N \rightarrow \infty} \varepsilon_N = \varepsilon^*$  and if  $\varepsilon^{**} > \varepsilon^* \exists \varepsilon$  in  $(\varepsilon^*, \varepsilon^{**})$  for which the theorem is false.

Analogous definitions of strategy and accuracy are made for statements of the form “if  $|\varepsilon| \geq \varepsilon_N$  then  $\mathcal{F}_\varepsilon$  does not verify the conclusions of the theorem.” We call such strategies accurate strategies for converse K.A.M. bounds. This concept of accuracy is very similar to the “finite computability hypothesis” of [MP] (see also [S]).

### III. THEORY

We have constructed accurate strategies for all of the classical small denominator problems in [Z]. Following Zehnder, we observe that each of these problems can be reformulated as a question on the existence of zeros of an appropriately chosen functional, e.g. equations I.1 and I.2. Then, the strategies are made up of three ingredients, a “constructive” implicit function theorem (I.F.T.) that asserts that, from an approximate zero satisfying certain conditions, one can construct an actual zero and two algorithms that systematically construct approximations and verify the conditions:

**Theorem III.1.** For certain choices of Banach spaces  $X_0, X_1$  and certain functionals  $\mathcal{F}_\varepsilon : X_0 \mapsto X_0$  there exist a computable function  $g : \mathbf{R}^n \times \mathbf{R}^+ \mapsto \mathbf{R}$  satisfying  $\lim_{\mu \rightarrow 0} g(M_1, \dots, M_n; \mu) = \infty \forall M_1, \dots, M_n$  and computable functionals  $\mathcal{L}_1, \dots, \mathcal{L}_n : X_0 \mapsto \mathbf{R}$  with the following property. Suppose that  $h^* \in X_0$  satisfies:

- (i)  $\|\mathcal{F}_\varepsilon(h^*)\|_{X_1} \leq \mu,$
- (ii)  $\mathcal{L}_1(h^*) \leq M_1, \dots, \mathcal{L}_n(h^*) \leq M_n,$
- (iii)  $g(M_1, \dots, M_n; \mu) \geq 0.$

Then,  $\exists h \in X_1$  s.t.  $\mathcal{F}_\varepsilon(h) = 0$  and  $\|h - h^*\|_{X_1} \leq \mathcal{O}(\mu).$

*Remarks.* (1) The point of this theorem is that a finite number of bounds on a specific function  $h^*$ , show that there is a true solution. When the bounds in (ii) are finite, it suffices that  $h^*$  satisfies the equation with sufficient approximation.

(2) Not all functionals admit I.F.T.’s of this type. For a review of functionals relevant for these problems see [Z].

(3) In the literature  $h^*$  is usually chosen such that  $\mathcal{F}_0(h^*) = 0$ . Then (iii) holds only for  $\varepsilon$  small.

**Lemma III.2.** *There is a sequence of finite dimensional Banach spaces,  $X^{(N)} \subset X_0$ , and two algorithms,  $\mathcal{A}$  and  $\mathcal{B}$  such that*

- (i)  $\mathcal{A}$  yields an  $h_N \in X^{(N)}$  in  $\leq N$  operations. If  $\mathcal{F}_\varepsilon$  has a zero in  $X_0$  then,  $\lim h_N = h$  exists in  $X_0$ , and  $\mathcal{F}_\varepsilon(h) = 0$ .
- (ii) For any  $h^*$  in some  $X^{(N)}$ ,  $\mathcal{B}$  yields a sequence of bounds:

$$(III.1) \quad \mathcal{L}_1(h^*) \leq M_{1, \hat{N}}, \dots, \mathcal{L}_n(h^*) \leq M_{n, \hat{N}} \quad \text{and} \quad \|\mathcal{F}_\varepsilon(h^*)\|_{X_1} \leq \mu_{\hat{N}}.$$

The  $M_{i, \hat{N}}, \mu_{\hat{N}}$  can be computed in less than  $\hat{N}$  operations and

$$(III.2) \quad \lim_{\hat{N} \rightarrow \infty} M_{i, \hat{N}} = \mathcal{L}_i(h^*) \quad 1 \leq i \leq n, \quad \text{and} \quad \lim_{\hat{N} \rightarrow \infty} \mu_{\hat{N}} = \|\mathcal{F}_\varepsilon(h^*)\|_{X_1}.$$

There are many ways of generating approximate solutions  $h_N$ , but (i) requires that the approximations converge whenever there is a true solution. Expansions in powers of  $\varepsilon$  can fail to satisfy (i) if there are complex singularities in  $\varepsilon$  of small modulus. This method is used in [CC] to find “good” numerical values for I.2.

For the problems we consider, there are different equivalent functionals admitting constructive I.F.T.’s, each employing slightly different approximation schemes. Certain “standard” proofs use functionals as in I.1 since solutions can be built from an infinite sequence of coordinate changes; but then bounds III.1 are difficult as approximations of both  $\psi$  and  $\psi^{-1}$  are necessary.

Arnold uses I.1 with truncated Taylor series to approximate  $\psi$  and  $\psi^{-1}$  [A]. As the maximal domain of  $\psi$  is a disk [R], but that of  $\psi^{-1}$  is generally more complicated, the truncations converge to  $\psi$  but not to  $\psi^{-1}$ , on their whole domain. Therefore III.2 fails for  $\mu_{\hat{N}}$ . This explains why [LT] only obtain K.A.M. bounds 65% of optimal for  $f_\varepsilon$  as in I.3, even though they ignored computer roundoff error.

We have constructed accurate strategies for I.1 using pointwise interpolation of  $\psi^{-1}$  [R] but find that it is more efficient to consider instead

$$(III.3) \quad \mathcal{F}_\varepsilon(\psi) = f_\varepsilon \circ \psi - \psi \circ f_0 = 0$$

which is equivalent to I.1, and involves no functional inverses.

It is important to choose functionals whose solutions have domains convenient for approximations, e.g. disks or strips in  $\mathbb{C}$ ,

for which we use as  $X^{(N)}$  order  $N$  polynomials or Fourier polynomials. For example, I.1 has solutions defined on disks, but the equivalent equation for  $\psi^{-1} = \phi$ :

$$\phi \circ f_\varepsilon - f_0 \circ \phi = 0$$

has solutions whose maximal domains are complicated sets and which therefore require complicated approximants.

Solutions of III.3 are univalent on their maximal domains of definition [R]. When the domain contains the unit disk the Taylor coefficients satisfy a priori bounds such as the Bierbach-de Branges bounds. Since the ( $\varepsilon$  dependent) Taylor coefficients are computed by power matching we find upper bounds to  $\varepsilon$  for which Theorem I.1 holds. These converse K.A.M. bounds can be shown to be accurate [R].

We also use the above a priori bounds to find limits on the number of operations required to reach a fixed precision:

**Theorem III.3.** *For equation III.3 with  $\hat{f}(z)$  a polynomial and any  $0 < p < 1$ ,  $\exists N_0 \geq 0$  depending on  $f_\varepsilon$  and  $p$  such that for  $N \geq N_0$  and*

$$\varepsilon \leq \varepsilon^* e^{-(1/N^p)}$$

where  $\varepsilon^*$  is the optimal value, the assumptions of the I.F.T. are verified with  $h^*$  the order  $N$  Taylor expansion of  $\psi$ .

#### IV. COMPUTER IMPLEMENTATION

Since a strategy involves an algorithm it is natural to use a computer for an implementation. Unfortunately, computers are equipped with an approximate form of arithmetic, so their results cannot be used as ingredients of a proof. This limitation is overcome by constructing a computer arithmetic on intervals that is guaranteed to generate upper and lower bounds of algebraic operations [M].

Of course, in the process of performing a computation these upper and lower bounds may stray quite far from the real results because of numerical instability of uncontrolled growth of round-off error.

In the implementation of the algorithms of Lemma III.2, we minimize roundoff error by carefully choosing the spaces  $X^{(N)}$  and analyzing the growth of intervals to eliminate instabilities in the computer procedures. Certain implementations based on interpolation are prone to roundoff error. For example, in an interesting paper, [BZ] construct an accurate strategy based on interpolation that they implement with approximate arithmetic. They find that

