

## KAM à la R

JÜRGEN PÖSCHEL

In [4] Rüssmann proposed – quoting from his abstract – “a new variant of the KAM-theory, containing an artificial parameter  $q$ ,  $0 < q < 1$ , which makes the steps of the KAM-iteration infinitely small in the limit  $q \nearrow 1$ . ... The new technique of estimation differs completely from all what has appeared about KAM-theory in the literature up to date. Only Kolmogorov’s idea of local linearization and Moser’s modifying terms are left. The basic idea is to use the polynomial structure in order to transfer, at least partially, the whole KAM-procedure outside of the original domain of definition of the given dynamical system.”

It is the purpose of this note to make this scheme accessible in an even simpler setting, namely for analytic perturbations of constant vector fields on a torus. As a side effect the result may be the shortest complete KAM proof for perturbations of integrable vector fields available so far.

### 1 Result

Let  $N$  denote a constant vector field on the  $n$ -torus  $\mathbb{T}^n = \mathbb{R}^n / 2\pi\mathbb{Z}^n$  describing uniform rotational motions with frequencies  $\omega = (\omega_1, \dots, \omega_n)$ . Putting  $N$  into normal form, we have  $N = \omega$ . A small perturbation  $X = N + P$  usually destroys this simple flow, due to frequency drifts and the effect of resonances. If, however, the frequencies  $\omega$  are strongly nonresonant, the perturbation  $P$  is sufficiently smooth and small, and if we are allowed to add a small correctional  $n$ -vector to adjust frequencies, then  $X$  is conjugate to  $\omega$ . This is the content of the classical KAM theorem with modifying terms for this model problem, as introduced by Moser [2].

The precise setting is the following. We consider  $N$  as a vector field in normal form depending on the frequencies  $\omega$  as parameters. These vary in some neighbourhood of a fixed compact set  $\Omega \subset \mathbb{R}^n$  consisting of strongly nonresonant frequencies. That is, each  $\omega \in \Omega$  satisfies

$$(1) \quad |\langle k, \omega \rangle| \geq \frac{\alpha}{\Delta(|k|)}, \quad 0 \neq k \in \mathbb{Z}^n,$$

with some  $\alpha > 0$  and some *Rüssmann approximation function*  $\Delta$ . These are continuous, increasing, unbounded functions  $\Delta: [1, \infty) \rightarrow [1, \infty)$  such that  $\Delta(1) = 1$  and

$$\int_1^\infty \frac{\log \Delta(t)}{t^2} dt < \infty.$$

The perturbation  $P$  is assumed to be analytic in the angles  $\theta \in \mathbb{T}^n$  and may depend analytically on the parameters  $\omega$  as well. The complex domains are

$$D_s = \{\theta : |\operatorname{Im} \theta| < s\}, \quad \Omega_h = \{z : |z - \Omega| < h\},$$

where  $|\cdot|$  denotes the *max*-norm for complex vectors, while it denotes the *sum*-norm for integer vectors. To simplify matters considerably, we employ the *weighted norms*

$$(2) \quad |P|_{s,h} = |P|_{s,\Omega_h} = \sup_{\omega \in \Omega_h} \sum_{k \in \mathbb{Z}^n} |p_k(\omega)| e^{|k|s}, \quad P = \sum_{k \in \mathbb{Z}^n} p_k(\omega) e^{i\langle k, \theta \rangle}.$$

Finally, with any approximation function  $\Delta$  we associate another such function  $\Lambda$  by setting  $\Lambda(t) = t\Delta(t)$ .

**KAM Theorem** *Suppose  $X = N + P$  is real analytic on  $D_s \times \Omega_h$  with*

$$(3) \quad |P|_{s,h} = \varepsilon < \frac{h}{16} \leq \frac{\alpha}{32\Lambda(\tau)},$$

where  $\tau$  is so large that

$$r := 8 \int_\tau^\infty \frac{\log \Lambda(t)}{t^2} dt < \frac{s}{2}.$$

Then there exists a real map  $\varphi: \Omega \rightarrow \Omega_h$ , and for each  $\omega \in \Omega$  a real analytic diffeomorphism  $\Phi_\omega$  of the  $n$ -torus, such that

$$\Phi_\omega^*(\varphi(\omega) + P) = \omega.$$

Moreover,  $|\varphi - \operatorname{id}|_\Omega \leq \varepsilon$  and  $|\Phi - \operatorname{id}|_{s-2r,\Omega} \leq \Lambda(\tau)\alpha^{-1}\varepsilon$ .

To keep things as simple as possible, we do not aim to optimize our constants, nor do we address regularity questions with respect to  $\omega$ .

The above smallness condition does not depend explicitly on the dimension  $n$  of the problem. However, this dimension enters implicitly through the small divisor conditions (1) and Dirichlet's lemma which states that for nonresonant vectors  $\omega$ ,

$$\min_{0 < |k| \leq K} |\langle k, \omega \rangle| \leq \frac{|\omega|}{K^{n-1}}.$$

Hence the approximation function  $\Delta$  has to grow at a rate depending on  $n$  in order to obtain admissible frequencies. A typical example is  $\Delta(t) = t^\nu$  with  $\nu > n - 1$ .

## 2 Outline

We prove the theorem by an iterative process of successive coordinate transformations proposed by Kolmogorov [1]. However, at variance with the crustimoney proseedcake [5, Chapter IV], we use a scheme of estimates proposed by Rüssmann, which does not rely on superlinear convergence speeds, but aims to decrease the size of the perturbation just a tiny bit at each step.

To this end, we split  $P$  into an 'infrared' part  $\tilde{P}$  and an 'ultraviolet' part  $\hat{P}$ . However - and this is the new twist introduced by Rüssmann -  $\hat{P}$  also contains fractions of the Fourier coefficients of  $P$  of *low order*. As a result,  $\tilde{P}$  will be bounded on a *larger domain*, with even a *better bound* than  $P$  itself.

The term  $\tilde{P}$  is then handled as usual. We write the coordinate transformation  $\Phi$  as the time-1-map of the flow  $F_t$  of a vector field  $F$ , which solves the homological equation  $[F, N] = \tilde{P} - P_0$ , where  $P_0$  denotes the mean value of  $\tilde{P}$  and  $[\cdot, \cdot]$  the Lie bracket of two vector fields. We obtain

$$\begin{aligned} \Phi^*(N + \tilde{P}) &= F_t^*(N + \tilde{P}) \Big|_{t=1} \\ &= N + [N, F] + \int_0^1 (1-t) F_t^* [[N, F], F] dt \\ &\quad + \tilde{P} + \int_0^1 F_t^* [\tilde{P}, F] dt. \end{aligned}$$

Using the homological equation for  $[N, F]$ , the result is

$$\Phi^*(N + \tilde{P}) = N + P_0 + \int_0^1 F_t^* [(t\tilde{P} + (1-t)P_0), F] dt.$$

The new constant vector field is then

$$N_+ = N + P_0 = \omega + p_0(\omega) = \omega_+.$$

A change of parameters  $\varphi: \omega_+ \mapsto \omega$  transforms this back into standard normal form. Taking into account the discarded ultraviolet term  $\hat{P}$ ,

$$P_+ = F_1^* \hat{P} + \int_0^1 F_t^* [P_t, F] dt, \quad P_t = t\tilde{P} + (1-t)P_0,$$

is then the new perturbation of the new vector field  $N_+$ .

In carrying out the pertinent estimates, it is extremely convenient to use weighted norms such as (2) instead of sup-norms as in [4]. This way, estimates on larger domains such as (5) and small divisor estimates such as (6) are immediate and do not require results from harmonic analysis. Indeed, it could be argued that sup-norms should be converted to weighted norms before entering the KAM machinery - once inside this machinery they are rather clumsy, inefficient, and produce a lot of unwieldy constants.

### 3 Step lemma

Before proceeding to the details we note that by a proper scaling of time and hence of the vector fields we can assume that the small divisor conditions (1) hold with the normalized value  $\alpha = 2$ .

In the following everything will be *real analytic* without explicitly saying so.

**Step Lemma** *Let  $0 < \sigma < s/2$  and  $\tau \geq 1$ . Set  $a = 1 - e^{-\tau\sigma} > 0$  and assume that*

$$(4) \quad |P|_{s,h} \leq \varepsilon < \min \left\{ \frac{h}{2a}, \frac{1}{2\Lambda(\tau)} \right\}, \quad h \leq \frac{1}{\Lambda(\tau)}.$$

*Then there exist parameter and coordinate transformations  $\varphi: \Omega_{h-2a\varepsilon} \rightarrow \Omega_h$  and  $\Phi: D_{s-2\sigma} \rightarrow D_s$  which together transform  $X = N + P$  into  $X_+ = N + P_+$  with*

$$|P_+|_{s-2\sigma, h-2a\varepsilon} \leq q\varepsilon,$$

where

$$q = (1 - a + a^2b)(1 + b)e^a, \quad b = \Lambda(\tau)\varepsilon.$$

Moreover,  $|\varphi - id|_{h-2a\varepsilon} \leq a\varepsilon$  and  $|\Phi - id|_{s-2\sigma, h-2a\varepsilon} \leq \Lambda(\tau)\sigma\varepsilon$ .

*Proof* Let  $P = \tilde{P} + \hat{P}$  with

$$\hat{P} = \sum_{|k| \geq \tau} p_k e^{i(k,\theta)} + (1-a) \sum_{|k| < \tau} p_k e^{|k|\sigma} e^{i(k,\theta)}.$$

Clearly, in view of  $e^{-\tau\sigma} = 1 - a$ ,

$$|\hat{P}|_{s-\sigma} \leq (1-a) |P|_s,$$

where we dropped  $\langle h \rangle$  from the notation since this parameter stays fixed until the very last paragraph of this section. On the other hand, the polynomial rest,

$$\tilde{P} = \sum_{|k| < \tau} \tilde{p}_k e^{i\langle k, \theta \rangle}, \quad \tilde{p}_k = (1 - (1-a)e^{k|\sigma}) p_k,$$

is bounded on a *larger* domain. Indeed, with  $\tilde{\sigma} = \sigma(1-a)/a$ ,

$$(5) \quad |\tilde{P}|_{s+\tilde{\sigma}} \leq \sup_{0 \leq t \leq \tau} (1 - (1-a)e^{t\sigma}) e^{t\tilde{\sigma}} \sum_{|k| < \tau} |p_k| e^{|k|s} \leq a |P|_s,$$

as the function under the sup is monotonically decreasing for  $0 \leq t \leq \tau$  and equals  $a$  at  $t = 0$ .

The linearized equation  $[F, N] = \tilde{P} - P_0$  is solved as usual by

$$F = \sum_{0 < |k| < \tau} \frac{\tilde{P}_k}{i\langle k, \omega \rangle} e^{i\langle k, \theta \rangle}.$$

For any  $\omega \in \Omega_h$  there is a vector  $\tilde{\omega} \in \Omega$  with  $|\omega - \tilde{\omega}| < h \leq 1/\Lambda(\tau)$  by (4) and hence, in view of  $\Lambda(\tau) = \tau\Delta(\tau)$ ,

$$|\langle k, \omega - \tilde{\omega} \rangle| \leq |k| |\omega - \tilde{\omega}| \leq \tau h \leq \frac{\tau}{\Lambda(\tau)} = \frac{1}{\Delta(\tau)}.$$

As  $\tilde{\omega}$  satisfies (1) with  $\alpha = 2$ , all relevant divisors thus admit the lower bound

$$|\langle k, \omega \rangle| \geq |\langle k, \tilde{\omega} \rangle| - |\langle k, \omega - \tilde{\omega} \rangle| \geq \frac{2}{\Delta(\tau)} - \frac{1}{\Delta(\tau)} = \frac{1}{\Delta(\tau)}.$$

So with (5) and  $a = 1 - e^{-\tau\sigma} \leq \tau\sigma$  we get

$$(6) \quad |F|_{s+\tilde{\sigma}} \leq \Delta(\tau) |\tilde{P}|_{s+\tilde{\sigma}} \leq \Delta(\tau) a \varepsilon \leq \Lambda(\tau) \sigma \varepsilon.$$

In particular,  $|F|_{s+\tilde{\sigma}} \leq \sigma$  by hypothesis (4), so the vector field  $F$  generates a flow  $F_t$  that for  $0 \leq t \leq 1$  satisfies

$$F_t : D_{s-2\sigma} \rightarrow D_{s-\sigma}, \quad |F_t - id|_{s-2\sigma} \leq \Lambda(\tau) \sigma \varepsilon.$$

To estimate  $P_+$  we note that  $|P_t|_{s+\tilde{\sigma}} \leq t |\tilde{P}|_{s+\tilde{\sigma}} + (1-t) |P_0|_{s+\tilde{\sigma}} \leq a \varepsilon$  and

$$(s + \tilde{\sigma}) - (s - \sigma) = \frac{1-a}{a} \sigma + \sigma = \frac{\sigma}{a}.$$

Lemma B.1 and the abbreviation  $b = \Lambda(\tau) \varepsilon$  thus yield

$$|[P_t, F]|_{s-\sigma} \leq \frac{a}{\sigma} |P_t|_{s+\tilde{\sigma}} |F|_{s+\tilde{\sigma}} \leq \Lambda(\tau) a^2 \varepsilon^2 = a^2 b \varepsilon.$$

In view of (6) and  $\varepsilon \leq 1/2\Lambda(\tau)$  by (4) we can apply Lemma B.2 with  $r = s - \sigma$  and  $\lambda = 1/a$  to obtain

$$\int_0^1 |F_t^* [P_t, F]|_{s-2\sigma} dt \leq (1+b)e^a |[P_t, F]|_{s-\sigma} \leq a^2 b(1+b)e^a \varepsilon.$$

Similarly,

$$|F_1^* \hat{P}|_{s-2\sigma} \leq (1+b)e^a |\hat{P}|_{s-\sigma} \leq (1-a)(1+b)e^a \varepsilon.$$

Both estimates together yield the stated estimate of  $P_+$ .

Finally,  $N_+ = N + P_0$  has frequencies  $\omega_+ = \omega + p_0(\omega)$ . As

$$|p_0|_h \leq |\tilde{P}|_{s+\tilde{\sigma}, h} \leq a\varepsilon < h/2$$

by hypothesis (4), the map  $\omega \mapsto \omega_+$  has an inverse

$$\varphi : \Omega_{h-2a\varepsilon} \rightarrow \Omega_{h-a\varepsilon}, \quad \omega = \varphi(\omega_+),$$

satisfying  $|\varphi - id|_{h-2a\varepsilon} \leq a\varepsilon$  by Lemma A.1. Then  $N_+$  is again in standard normal form, and the proof of the Step Lemma is complete. ■

#### 4 Iteration and convergence

Iterating the Step Lemma is simple. We can always choose  $0 < a < 1$  and  $0 < b \leq 1/2$  so that

$$q = (1 - a + a^2 b)(1 + b)e^a < 1,$$

and  $q$  may be as close to 1 as we wish. It then suffices to choose for  $\varepsilon$ ,  $h$  and  $\Lambda$  geometric sequences with the same base  $q$ , namely

$$\varepsilon_\nu = \varepsilon_0 q^\nu, \quad h_\nu = h_0 q^\nu, \quad \Lambda_\nu = \Lambda_0 q^{-\nu},$$

where we assume that  $\Lambda_0 \geq \Lambda(1) = \Delta(1)$ . Next, let  $\tau_\nu = \Lambda^{-1}(\Lambda_\nu)$  and define  $\sigma_\nu$  and  $s_\nu$  through

$$1 - a = e^{-\tau_\nu \sigma_\nu}, \quad s_{\nu+1} = s_\nu - 2\sigma_\nu.$$

As we will see in a moment, the  $s_\nu$  have a positive limit for  $\Lambda_0$  sufficiently large.

**Iterative Lemma** *Suppose that*

$$|P|_{s_0, h_0} \leq \varepsilon_0 < \min \left\{ (1-q) \frac{h_0}{2a}, \frac{b}{\Lambda_0} \right\}, \quad h_0 \leq \frac{1}{\Lambda_0},$$

with  $\Lambda_0$  sufficiently large. Then for each  $\nu \geq 1$  there exists a parameter and coordinate transformation

$$(\Phi_\nu, \varphi_\nu) : D_{s_\nu} \times \Omega_{h_\nu} \rightarrow D_{s_0} \times \Omega_{h_0}$$

which transforms  $N + P_0$  into  $N + P_\nu$  such that  $|P_\nu|_{s_\nu, h_\nu} \leq \varepsilon_\nu$ .

*Proof* This follows by applying the Step Lemma repeatedly and composing the resulting maps. Just note that  $\Lambda(\tau_\nu) = \Lambda_\nu$  and

$$\varepsilon_\nu \Lambda_\nu = \varepsilon_0 \Lambda_0 \leq b \leq 1/2, \quad \frac{h_\nu - 2a\varepsilon_\nu}{h_{\nu+1}} = \frac{h_0 - 2a\varepsilon_0}{qh_0} \geq 1,$$

for all  $\nu$  by construction and hypotheses. ■

*Proof of the KAM Theorem* Recall that we normalized  $\alpha = 2$ . Then the hypotheses of the Iterative Lemma are satisfied by  $P_0 = P$  with  $\varepsilon_0 = \varepsilon$ ,  $h_0 = h$ ,  $s_0 = s$  and the same  $\Lambda_0$  as above.

To be able to apply this lemma infinitely often we have to verify that the  $s_\nu$  tend to a positive limit. Indeed,

$$\sum_{\nu \geq 1} \frac{1}{\tau_\nu} \leq \int_0^\infty \frac{d\nu}{\Lambda^{-1}(\Lambda_0 q^{-\nu})} = \frac{1}{\log q^{-1}} \int_{\tau_0}^\infty \frac{d\Lambda(t)}{t\Lambda(t)}$$

by letting  $t = \Lambda^{-1}(\Lambda_0 q^{-\nu})$ . Integrating by parts and requiring  $\Lambda(\tau_0) \geq q^{-1}$  we get

$$\sum_{\nu \geq 0} \frac{1}{\tau_\nu} \leq \frac{1}{\log q^{-1}} \int_{\tau_0}^\infty \frac{\log \Lambda(t)}{t^2} dt.$$

It follows that

$$r := \sum_{\nu \geq 0} \sigma_\nu = \sum_{\nu \geq 0} \frac{\log(1-a)^{-1}}{\tau_\nu} \leq \frac{\log(1-a)}{\log q} \int_{\tau_0}^\infty \frac{\log \Lambda(t)}{t^2} dt.$$

Hence, by choosing  $\tau_0$  sufficiently large, we can achieve that  $r < s/2$  and thus  $s_\nu \searrow s - 2r > 0$ .

For the statement of the theorem we choose  $a = 1/2$  and  $b = 1/16$ , which results in

$$q \approx \frac{9}{10}, \quad \frac{\log(1-a)}{\log q} \leq 8.$$

In fact, 7 instead of 8 would also do, but we prefer powers of 2.

Now fix  $\omega \in \Omega$ , and consider the sequence of transformations provided by the Iterative Lemma for this  $\omega$ . On  $D_r$  the family  $(\Phi_\nu)$  is uniformly bounded,

as is the vector sequence  $(\varphi_\nu)$ . So by Montel's and Weierstrass' theorem, there is a convergent subsequence. As  $|P_\nu|_{S_\nu, h_\nu} \leq \varepsilon_\nu \rightarrow 0$  along any such subsequence, this subsequence transforms  $X = N + P$  at the limit parameter value  $\tilde{\omega}$  into the normal form  $N$  at  $\omega$ , that is, the constant vector field  $\omega$ . ■

#### A An inverse function theorem

**Lemma A.1** *Suppose  $f: \Omega_h \rightarrow \mathbb{C}^n$  is analytic and*

$$|f - id|_h \leq \varepsilon < h/2.$$

*Then  $f$  has an analytic inverse  $\varphi: \Omega_{h-2\varepsilon} \rightarrow \Omega_h$ , and  $|\varphi - id|_{h-2\varepsilon} \leq \varepsilon$ .*

*Proof* For any  $0 < k < h - 2\varepsilon$  we have

$$|Df - I|_{k+\varepsilon} \leq \frac{\varepsilon}{h - (k + \varepsilon)} < 1$$

by Cauchy's inequality. Therefore, the operator

$$T: \varphi \mapsto id - (f - id) \circ \varphi$$

defines a contraction on the space of analytic maps  $\varphi: \Omega_k \rightarrow \Omega_h$ ,  $|\varphi - id|_k \leq \varepsilon$ . Its unique fixed point  $\varphi$  is the analytic inverse to  $f$  on  $\Omega_k$ . Letting  $k \rightarrow h - 2\varepsilon$  we obtain the claim. ■

#### B Weighted norms

**Lemma B.1** *Let  $U$  and  $V$  be analytic vector fields on the torus  $\mathbb{T}^n$ . Then, for  $0 < r < \min\{u, v\}$ ,*

$$|[U, V]|_r \leq \frac{1}{e} \left( \frac{1}{u-r} + \frac{1}{v-r} \right) |U|_u |V|_v.$$

*Proof* We have  $U = \sum_k u_k e_k$  and  $V = \sum_k v_k e_k$  with  $e_k = e^{i(k, \theta)}$ . Therefore,

$$DU \cdot V = \sum_k i u_k \langle k, V \rangle e_k = \sum_{k,l} i u_k \langle k, v_l \rangle e_{k+l} = \sum_{k,l} i u_k \langle k, v_{l-k} \rangle e_l$$

and thus

$$\begin{aligned} |DU \cdot V|_r &\leq \sum_{k,l} |k| |u_k| |v_{l-k}| e^{llr} \\ &\leq \sum_{k,l} |k| |u_k| e^{|k|r} |v_{l-k}| e^{|l-k|r} \end{aligned}$$



$$\begin{aligned} &\leq \sup_{t \geq 0} t e^{-(u-r)t} \left( \sum_k |u_k| e^{k|u|} \right) \left( \sum_l |v_l| e^{l|v|} \right) \\ &\leq \frac{1}{e(u-r)} |U|_u |V|_v. \end{aligned}$$

Exchanging the roles of  $U$  and  $V$  we get an analogous estimate for  $|DV \cdot U|_r$  which proves the claim. ■

**Lemma B.2** *Suppose the vector fields  $F$  and  $V$  are analytic on the torus  $\mathbb{T}^n$ . If  $b = \sigma^{-1} |F|_{r+\lambda\sigma} \leq 1/2$  with  $0 < \sigma < r$  and  $\lambda > 0$ , then*

$$|F_t^* V|_{r-\sigma} \leq (1 + bt) e^{1/\lambda} |V|_r, \quad 0 \leq t \leq 1.$$

*Proof* We have the Lie series expansion

$$F_t^* V = \sum_{n \geq 0} \frac{1}{n!} V_n t^n$$

with  $V_0 = V$  and  $V_n = [V_{n-1}, F]$  for  $n \geq 1$ . Let  $|\cdot|_i = |\cdot|_{r-i\sigma/n}$  for  $0 \leq i \leq n$ . Then, by the preceding lemma,

$$\begin{aligned} |V_n|_{r-\sigma} &= |V_n|_n = |[V_{n-1}, F]|_n \\ &\leq \left( \frac{n}{e\sigma} + \frac{1}{e\lambda\sigma} \right) |V_{n-1}|_{n-1} |F|_{r+\lambda\sigma} \\ &= \frac{n}{e\sigma} \left( 1 + \frac{1}{\lambda n} \right) |V_{n-1}|_{n-1} |F|_{r+\lambda\sigma}. \end{aligned}$$

Applying this estimate  $n$  times, we get

$$|V_n|_{r-\sigma} \leq \left( \frac{n}{e\sigma} \right)^n e^{1/\lambda} |V|_r |F|_{r+\lambda\sigma}^n.$$

Summing up and replacing  $\sigma^{-1} |F|_{r+\lambda\sigma}$  by  $b$  we obtain

$$|F_t^* V|_{r-\sigma} \leq |V|_r e^{1/\lambda} \sum_{n \geq 0} \frac{1}{n!} \left( \frac{nb t}{e} \right)^n.$$

With  $n! \geq n^n / e^{n-1}$  for  $n \geq 1$  and  $0 \leq bt \leq 1/2$  the last sum is bounded by

$$1 + \sum_{n \geq 1} \frac{(bt)^n}{e} \leq 1 + bt. \quad \blacksquare$$

If  $V$  depends on parameters  $\omega$  in a point set  $\Pi$ , we define the norm  $|V|_{s,\Pi}$  as in (2). All the preceding estimates are uniform with respect to such parameters, so the results extend to this case.

*Acknowledgement.* It is a pleasure to thank the referees for pointing out a number of imprecisions, in particular a slight misquotation of Winnie-the-Pooh and a mistaken application of my non-existent French.

### References

- [1] A. N. KOLMOGOROV, On the conservation of conditionally periodic motions for a small change in Hamilton's function. *Dokl. Akad. Nauk SSSR* **98** (1954) 527-530 [Russian]. English translation in *Lecture Notes in Physics* **93**, Springer, 1979, 51-56.
- [2] J. MOSER, Convergent series expansions for quasi-periodic motions. *Math. Ann.* **169** (1967) 136-176.
- [3] J. PÖSCHEL, A lecture on the classical KAM theorem. *Proc. Symp. Pure Math.* **69** (2001) 707-732.
- [4] H. RÜSSMANN, KAM-iteration with nearly infinitely small steps in dynamical systems of polynomial character. *Discr. Contin. Dynam. Syst. Ser. S* **3** (2010), 683-718.
- [5] A. A. MILNE, *Winnie-the-Pooh*. Dutton, New York.

*Institut für Analysis, Dynamik und Optimierung  
Universität Stuttgart, Pfaffenwaldring 57, D-70569 Stuttgart  
poschel@mathematik.uni-stuttgart.de*