Hamiltonian Systems
MSRI Publications
Volume **72**, 2024

Some remarks on the classical KAM theorem, following Pöschel

ABED BOUNEMOURA

We propose a slight correction and a slight improvement on the main result contained in "A lecture on Classical KAM Theorem" by J. Pöschel.

1. Introduction and results

The paper [5] contains a very nice exposition of the classical KAM theorem which has been very influential. It is our purpose in this short and non-self-contained note to add two remarks to this remarkable paper.

The first one concerns a technical mistake in the proof of the main abstract statement Theorem A,[1] which has been recently pointed out and corrected in the PhD thesis [3]. Yet a correction of this mistake, following Pöschel arguments, leads to a final statement which is both less elegant and quantitatively weaker. We would like to explain how, by modifying slightly the arguments using ideas due to Rüssmann (see for instance [7]), Theorem A of [5] can be proved without any changes. The aforementioned modifications consist of replacing the crude Fourier truncation by a more refined polynomial approximation, and then set an iterative scheme with a linear,[2] rather than super-linear, speed of convergence.

The second one concerns the application of Theorem A to an ε-perturbation of a nondegenerate integrable Hamiltonian system. This gives persistence of a set of positive measure of analytic invariant quasiperiodic tori with fixed diophantine frequencies, such that each torus in this set is at a distance of order $\sqrt{\varepsilon}$ to its associated unperturbed invariant torus. By using a more adapted version of Theorem A, we can actually show that the distance is of order ε/α, where α is the constant of the Diophantine vector. This is not a new result, as this was already

MSC2020: 37J40.
Keywords: classical KAM theorem.

[1] The choices of h_0 and K_0, page 23 in [5], violate the condition $h_0 \leq \alpha(2K_0^{\nu})^{-1}$.

[2] We would like to quote here the paper [6]: "It has often been said that the rapid convergence of the Newton iteration is necessary for compensating the influence of small divisors. But a deeper analysis shows that this is not true. The Newton method compensates not only the influence of small divisors but also many bad estimates veiling the true structure of the problems."

proved in [9] using a refinement of Kolmogorov approach (for an individual torus).

So let us recall the main result of [5], keeping the same notations. For a given domain $\Omega \subseteq \mathbb{R}^n$, consider a subset $\Omega_\alpha \subseteq \Omega$ of Diophantine vectors with constant $\alpha > 0$ and exponent $\tau \geq n-1$. Given $0 < r, s, h \leq 1$, define

$$D_{r,s} = \{I \mid |I| < r\} \times \{\theta \mid |\mathrm{Im}(\theta)| < s\} \subseteq \mathbb{C}^n \times \mathbb{C}^n, \quad O_h = \{\omega \mid |\omega - \Omega_\alpha| < h\} \subseteq \mathbb{C}^n$$

where $|\cdot|$ is the sup norm for vectors, and let $|\cdot|_{r,s,h}$ the sup norm for functions defined on $D_{r,s} \times O_h$ and $|\cdot|_L$ the Lipschitz seminorm with respect to ω. Let $N(I, \omega) = e(\omega) + \omega \cdot I$, which can be seen as a family N_ω of linear integrable Hamiltonian depending on parameters $\omega \in \Omega$; the family of embedding $\Phi_0 : \mathbb{T}^n \times \Omega \to \mathbb{R}^n \times \mathbb{T}^n$ defined by $\Phi_0(\theta, \omega) = (0, \theta)$ defines, for each $\omega \in \Omega$, a Lagrangian torus invariant by the Hamiltonian flow of N_ω and quasiperiodic of frequency ω.

Theorem A. *Let $H = N + P$. Suppose P is real-analytic on $D_{r,s} \times O_h$ with*

$$|P|_{r,s,h} \leq \gamma \alpha r s^\nu, \quad \alpha s^\nu \leq h \tag{1-1}$$

where $\nu = \tau + 1$ and γ is a small constant depending only on n and τ. Then there exist a Lipschitz map $\varphi : \Omega_\alpha \to \Omega$ and a Lipschitz family of real-analytic Lagrangian embedding $\Phi : \mathbb{T}^n \times \Omega_\alpha \to \mathbb{R}^n \times \mathbb{T}^n$ that defines, for each $\omega \in \Omega_\alpha$, a Lagrangian torus invariant by the Hamiltonian flow of $H_{\varphi(\omega)}$ and quasiperiodic of frequency ω. Moreover, Φ is real-analytic on $T_* = \{\theta \mid |\mathrm{Im}(\theta)| < s/2\}$ for each ω and

$$\begin{cases} |W(\Phi - \Phi_0)|, & \alpha s^\nu |W(\Phi - \Phi_0)|_L \leq c(\alpha r s^\nu)^{-1} |P|_{r,s,h}, \\ |\varphi - \mathrm{Id}|, & \alpha s^\nu |\varphi - \mathrm{Id}|_L \leq c r^{-1} |P|_{r,s,h}, \end{cases} \tag{1-2}$$

uniformly on $T_* \times \Omega_\alpha$ and Ω_α respectively, where c is a large constant depending only on n and τ, and $W = \mathrm{Diag}(r^{-1}\,\mathrm{Id}, s^{-1}\,\mathrm{Id})$.

As expressed in (1-2), the map (Φ, φ) is Lipschitz regular with respect to $\omega \in \Omega_\alpha$, and its Lipschitz norm (suitably weighted) is close to the one of (Φ_0, Id); this is all what is needed to transfer the positive measure in parameter space $\omega \in \Omega_\alpha$ to a positive measure of quasiperiodic solutions in phase space. One course one may ask whether (Φ, φ) is more regular with respect to $\omega \in \Omega_\alpha$ (since Ω_α is a closed set, smoothness has to be understood in the sense of Whitney). In fact, the sketch of proof we will give below implies the following: given any $l \in [1, +\infty[$, provided (1-1) is replaced by

$$|P|_{r,s,h} \leq \gamma(l) \alpha r s^\nu$$

for some $h > 0$ and some $\gamma(l) > 0$, (Φ, φ) is of class C^l with respect to ω: we simply chose $l = 1$ in Theorem A to obtain Lipschitz regularity. However, as

$l \to +\infty$, $\gamma(l) \to 0$ and thus we cannot conclude that (Φ, φ) is smooth. In order to reach such a statement, one can replace the linear scheme of convergence by the usual super-linear scheme (as described in [5] for instance) but then the exponent ν in (1-1) has to be deteriorate: given any $\mu > \nu$, we have that (Φ, φ) is smooth with respect to ω provided (1-1) is replaced by

$$|P|_{r,s,h} \leq \gamma(\mu, \nu)\alpha r s^{\mu}$$

for some $h > 0$ and some $\gamma(\mu, \nu) > 0$: again $\gamma(\mu, \nu) \to 0$ as $\mu \to \nu$. Popov (see [4]) showed that one can even go further and obtain some Gevrey smoothness of (Φ, φ) under a stronger smallness condition; without going into these rather technical issues, let us just say that (Φ, φ) can be shown to be Gevrey with exponent $1+\mu$ provided the polynomially small threshold s^{ν} in (1-2) is replace by a super-exponentially small threshold of order $\exp(-c(1/s)^{a})$ with $a = a(\mu, \nu) = \nu/(\mu - \nu)$. This is probably the best smoothness one can achieve in general.

Next we consider a small perturbation of a nondegenerate integrable Hamiltonian, that is a real-analytic Hamiltonian of the form

$$H(q, p) = h(p) + f(q, p), \quad |f| \leq \varepsilon$$

where $|f|$ is the sup norm on a proper complex domain. Introducing frequencies as independent parameters as in [5], one can write H as in Theorem A with

$$P = P_f + P_h, \quad |P_f| \leq \varepsilon, \quad |P_h| \leq M r^2$$

where M is a bound on the Hessian of h. At that point, the best choice for r seems to be $r \simeq \sqrt{\varepsilon}$ so that the size of P is of order ε and Theorem A can be applied; yet with such a choice it is obvious that because of the estimates for φ in (1-2), the distance between the perturbed and unperturbed torus will be of order $\varepsilon/r \simeq \sqrt{\varepsilon}$. Such an argument, used in [5], do not take into account the fact that the term P_h is actually integrable and at least quadratic in I (that is, $P_h(0, \omega) = 0$ and $\nabla_I P_h(0, \omega) = 0$): this is an important point, as the size of P_h will effectively enter into the conditions (1-1) but not in the estimates (1-2), simply because P_h do not get involved in the approximation procedure nor contribute to the linearized equations one need to solve at each step of the iteration. Then, taking into account the estimate for P_h (which itself is a consequence of the fact that it is at least quadratic in I), the requirement

$$|P| \lesssim \alpha r s^{\nu}$$

is then obviously implied by the conditions

$$|P_f| \lesssim \alpha r s^{\nu}, \quad r \lesssim \alpha s^{\nu}$$

and thus we can state the following theorem (with a change of notations).

Theorem B. *Let $H = N + P + Q$. Suppose P, Q are real-analytic on $D_{r,s} \times O_h$, Q is integrable and at least quadratic in I with $|Q|_{r,h} \leq Mr^2$ and*

$$|P|_{r,s,h} \leq \gamma \alpha r s^{\nu}, \quad r \leq \delta M^{-1} \alpha s^{\nu}, \quad \alpha s^{\nu} \leq h \tag{1-3}$$

where $\nu = \tau + 1$, γ and δ are small constants depending only on n and τ. Then there exist a Lipschitz map $\varphi : \Omega_\alpha \to \Omega$ and a Lipschitz family of real-analytic Lagrangian embedding $\Phi : \mathbb{T}^n \times \Omega_\alpha \to \mathbb{R}^n \times \mathbb{T}^n$ that defines, for each $\omega \in \Omega_\alpha$, a Lagrangian torus invariant by the Hamiltonian flow of $H_{\varphi(\omega)}$ and quasiperiodic of frequency ω. Moreover, the estimates (1-2) holds true.

We may now choose r as large as possible, namely $r \simeq \alpha s^{\nu}$, and obtain as a consequence that the distance between perturbed and unperturbed torus is of order $\varepsilon(\alpha s^{\nu})^{-1}$. As we already said, this fact was proved in [9]; alternatively, a slight modification in the proof in [2] yields the same result.

2. Sketch of proof

In this section, we will sketch the proof of Theorems A and B; actually, we will simply indicate the modifications with respect to [5] and we will use the same convention for implicit constants depending only on n and τ.

Proposition 2.1. *Let $H = N + P$, and suppose that $|P|_{s,r,h} \leq \varepsilon$ with*

$$\begin{cases} \varepsilon \cdot\!< \alpha \eta^2 r \sigma^{\nu}, \\ \varepsilon \cdot\!< hr, \\ h \leq \alpha (2K^{\nu})^{-1}, \quad K =\cdot\, \sigma^{-1} \log(n\eta^{-2}) \end{cases} \tag{2-1}$$

where $0 < \eta < \frac{1}{8}$ and $0 < \sigma < \frac{s}{5}$. Then there exists a real-analytic transformation

$$\mathcal{F} = (\Phi, \varphi) : D_{\eta r, s-5\sigma} \times O_{h/4} \to D_{r,s} \times O_h$$

such that $H \circ \mathcal{F} = N_+ + P^+$ with

$$|P_+| \leq 9\eta^2 \varepsilon \tag{2-2}$$

and

$$\begin{cases} |W(\Phi - \mathrm{Id})|, \quad |W(D\Phi - \mathrm{Id})W^{-1}| \cdot\!< (\alpha r \sigma^{\nu})^{-1} \varepsilon, \\ |\phi - \mathrm{Id}|, \qquad h|D\varphi - \mathrm{Id}|_L \cdot\!< r^{-1} \varepsilon, \end{cases} \tag{2-3}$$

uniformly on $D_{\eta r, s-5\sigma} \times O_h$ and $O_{h/4}$, with $W = \mathrm{Diag}(r^{-1}\,\mathrm{Id}, \sigma^{-1}\,\mathrm{Id})$.

The above proposition is a variant of the KAM step of [5], which we already used in [1]. The only difference is that in [5], instead of (2-1) the following conditions are imposed

$$\begin{cases} \varepsilon \cdot\!< \alpha \eta r \sigma^{\nu}, \\ \varepsilon \cdot\!< hr, \\ h \leq \alpha (2K^{\nu})^{-1}, \end{cases} \tag{2-4}$$

with a free parameter $K \in \mathbb{N}^*$, leading to the following estimate

$$|P_+| \lessdot (\varepsilon(r\sigma^\nu)^{-1} + \eta^2 + K^n e^{-K\sigma})\varepsilon. \tag{2-5}$$

instead of (2-2). The last two terms in the estimate (2-5) comes from the approximation of P by a Hamiltonian R which is affine in I and a trigonometric polynomial in θ of degree K; to obtain such an approximation, in [5] the author simply truncates the Taylor expansion in I and the Fourier expansion in θ to obtain the following approximation error

$$|P - R|_{s-\sigma,2\eta r,h} \lessdot (\eta^2 + K^n e^{-K\sigma}).$$

Yet we can use a more refined approximation result, which allows to get rid of the factor K^n in the above estimate. More precisely, we use Theorem 7.2 of [7] (choosing, in the latter reference, $\beta_1 = \cdots = \beta_n = \frac{1}{2}$ and $\delta^{1/2} = 2\eta$ for $\delta \le \frac{1}{4}$); with the choice of K as in (2-1),[3] this gives another approximation $\tilde{R}$ (which is nothing but a weighted truncation, both in the Taylor and Fourier series, which is affine in I and of degree bounded by K in θ) and a simpler error

$$|P - \tilde{R}|_{s-\sigma,2\eta r,h} \le 8\eta^2.$$

As for the first term in the estimate (2-5), it can be easily bounded by $\eta^2\varepsilon$ in view of the first part of (2-1) which is stronger than the first part of (2-4) required in [5].

Now, at variance with [5], we will use Proposition 2.1 in an iterative scheme with a linear speed of convergence as η will be chosen to be a small but fixed constant: for convenience, let us set

$$\eta = 10^{-1}4^{-\nu}, \quad \kappa = 9\eta^2.$$

Next, we define for $i \in \mathbb{N}$,

$$\sigma_0 = s/20, \quad \sigma_i = 2^{-i}\sigma_0, \quad s_0 = s, \quad s_{i+1} = s_i - 5\sigma_i$$

so that s_i converges to $s/2$. Then, for $K_i =\cdot\, \sigma_i^{-1}\log(n\eta^2) =\cdot\, \sigma_i^{-1}$, we set

$$h_i = \alpha(2K_i^\nu)^{-1} = 2^{-i\nu}h_0, \quad h_i \cdot\!= \alpha\sigma_i^\nu$$

and the condition $\alpha s^\nu \le h$ implies in particular than $h_0 \le h$. Finally, we put

$$\varepsilon_i = \kappa^i\varepsilon, \quad r_i = \eta^i r$$

and we verify that Proposition 2.1 can be applied infinitely many times: the third condition of (2-1) holds true by definition, whereas the first two conditions

[3] There is a constant depending only on n that we left implicit in the definition of K, which depends on the precise choice of norms for real and integer vectors, see [8] for instance.

of (2-1) amount to $\varepsilon_i \cdot\!< \alpha r_i \sigma_i^\nu$ which, in view of our choice of η, holds true for all $i \in \mathbb{N}$ provided it holds true for $i = 0$; for $i = 0$ the condition is satisfied in view of the threshold $\varepsilon \le \gamma \alpha r s^\nu$. Once we can iterate Proposition 2.1 infinitely many times, the convergence proof and the final estimates follow exactly as in [5], since the sequences $\varepsilon_i (h_i r_i)^{-1}$and $\varepsilon_i (h_i^2 r_i)^{-1}$ decrease geometrically, again by our choice of η. This concludes the sketch of proof.

To prove Theorem B, one needs the following variant of Proposition 2.1.

Proposition 2.2. *Let $H = N + P + Q$, suppose that $|P|_{s,r,h} \le \varepsilon$, $|Q|_{r,h} \le M r^2$ with Q integrable and at least quadratic in I and*

$$\begin{cases} \varepsilon \cdot\!< \alpha \eta^2 r \sigma^\nu, \\ r \cdot\!< M^{-1} \alpha \eta^2 \sigma^\nu, \\ \varepsilon \cdot\!< hr, \\ h \le \alpha (2K^\nu)^{-1}, \qquad K = n\sigma^{-1} \log(\eta^{-2}), \end{cases} \tag{2-6}$$

where $0 < \eta < \frac{1}{4}$ and $0 < \sigma < \frac{s}{5}$. Then there exists a real-analytic transformation

$$\mathcal{F} = (\Phi, \varphi) : D_{\eta r, s-5\sigma} \times O_{h/4} \to D_{r,s} \times O_h$$

such that $H \circ \mathcal{F} = N_+ + P_+ + Q$ with the estimates (2-2) and (2-3).

Let $\tilde{R}$ be the approximation of P; if $\{\,\cdot\,,\cdot\,\}$ denotes the Poisson bracket and $[\,\cdot\,]$ averaging over the angles, we solve the equation

$$\{F, N\} = \tilde{R} + Q - [\tilde{R} + Q]$$

which, since Q is integrable, is exactly the equation

$$\{F, N\} = \tilde{R} - [\tilde{R}]$$

that is solved in [5] (with, of course, R instead of $\tilde{R}$ as we explained above). This justifies that the transformation in Proposition 2.2 is the same as in Proposition 2.1, and in particular it satisfy the estimates (2-2). The only difference is that the new Hamiltonian writes

$$H \circ \mathcal{F} = N_+ + P_+ + Q, \quad N_+ = N + [\tilde{R}]$$

with

$$P_+ = \int_0^1 \{(1-t)[\tilde{R}] + t\tilde{R} + Q, F\} \circ X_F^t dt + (P - \tilde{R}) \circ X_F^1.$$

As compared to [5], there is an extra term in P_+ coming from Q, whose contribution is easily bounded by the simple Poisson bracket

$$|\{Q, F\}| \lessdot Mr(\alpha\sigma^\nu)^{-1}\varepsilon$$

and, in view of the extra condition we imposed in (2-6), one can easily arrange the estimate (2-3). This justifies Proposition 2.2, and the iteration leading to Theorem B is exactly the same as the one leading to Theorem A.

References

[1] A. Bounemoura, "Positive measure of KAM tori for finitely differentiable Hamiltonians", *J. Éc. polytech. Math.* **7** (2020), 1113–1132.

[2] A. Bounemoura and J. Féjoz, "KAM, α-Gevrey regularity and the α-Bruno–Rüssmann condition", *Ann. Sc. Norm. Super. Pisa Cl. Sci.* (5) **19**:4 (2019), 1225–1279.

[3] C. E. Koudjinan, "Quantitative KAM normal forms and sharp measure estimates", preprint, 2019. arXiv 1904.13062

[4] G. Popov, "KAM theorem for Gevrey Hamiltonians", *Ergodic Theory Dynam. Systems* **24**:5 (2004), 1753–1786.

[5] J. Pöschel, "A lecture on the classical KAM theorem", pp. 707–732 in *Smooth ergodic theory and its applications* (Seattle, WA, 1999), edited by A. Katok et al., Proc. Sympos. Pure Math. **69**, Amer. Math. Soc., Providence, RI, 2001.

[6] H. Rüssmann, "Nondegeneracy in the perturbation theory of integrable dynamical systems", pp. 5–18 in *Number theory and dynamical systems* (York, 1987), edited by M. M. Dodson and J. A. G. Vickers, London Math. Soc. Lecture Note Ser. **134**, Cambridge Univ. Press, 1989.

[7] H. Rüssmann, "Invariant tori in non-degenerate nearly integrable Hamiltonian systems", *Regul. Chaotic Dyn.* **6**:2 (2001), 119–204.

[8] H. Rüssmann, "Invariant tori in non-degenerate nearly integrable Hamiltonian systems", *Regul. Chaotic Dyn.* **6**:2 (2001), 119–204.

[9] J. Villanueva, "Kolmogorov theorem revisited", *J. Differential Equations* **244**:9 (2008), 2251–2276.

bounemoura@ceremade.dauphine.fr

Université Paris Dauphine, CEREMADE, Paris, France