Some remarks on the Classical KAM Theorem, following Pöschel

Abed Bounemoura

CNRS - PSL Research University
(Université Paris-Dauphine and Observatoire de Paris)

June 8, 2020

Abstract

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 [Pö01] 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, which has been recently pointed out and corrected in the PhD thesis [Kou19]. 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 [Rü11]), Theorem A of [Pö01] 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, rather than super-linear, speed of convergence.

The second one concerns the application of Theorem A to an ε-perturbation of a non-degenerate integrable Hamiltonian system. This gives persistence of a set of positive measure of analytic invariant quasi-periodic tori with fixed diophantine frequencies, such that

\footnote{1}{The choices of $h_0$ and $K_0$, page 23 in [Pö01], violate the condition $h_0 \leq \alpha (2K_0^\nu)^{-1}$.}

\footnote{2}{We would like to quote here the paper [Rü89]: “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.”}
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 $\varepsilon/\alpha$, where $\alpha$ is the constant of the Diophantine vector. This is not a new result, as this was already proved in [Vil08] using a refinement of Kolmogorov approach (for an individual torus).

So let us recall the main result of [Pöss01], 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

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

where $|.|$ is the sup norm for vectors, and let $|.|_{r,s,h}$ the sup norm for functions defined on $D_{r,s} \times O_h$ and $|.|_{r,s,h}$ the Lipschitz semi-norm with respect to $\omega$. Let $N(I, \omega) = e(\omega) + \omega \cdot I$, which can be seen as a family $N_\omega$ 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 $\Phi_0(\omega, \theta) = (0, \theta)$ defines, for each $\omega \in \Omega$, a Lagrangian torus invariant by the Hamiltonian flow of $N_\omega$ and quasi-periodic of frequency $\omega$.

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

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

where $\nu = \tau + 1$ and $\gamma$ is a small constant depending only on $n$ and $\tau$. 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 quasi-periodic of frequency $\omega$. Moreover, $\Phi$ is real-analytic on $T_\alpha = \{ \theta \mid |\Im(\theta)| < s/2 \}$ for each $\omega$ and

$$\begin{align*}
|W(\Phi - \Phi_0)|, \alpha s^\nu|W(\Phi - \Phi_0)|_L & \leq c(\alpha rs^\nu)^{-1}|P|_{r,s,h} \\
|\varphi - \text{Id}|, \alpha s^\nu|\varphi - \text{Id}|_L & \leq cr^{-1}|P|_{r,s,h}
\end{align*}$$

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

As expressed in (1.2), the map $(\Phi, \varphi)$ is Lipschitz regular with respect to $\omega \in \Omega_\alpha$, and its Lipschitz norm (suitably weighted) is close to the one of $(\Phi_0, \text{Id})$; this is all what is needed to transfer the positive measure in parameter space $\omega \in \Omega_\alpha$ to a positive measure of quasi-periodic solutions in phase space. One course one may ask whether $(\Phi, \varphi)$ is more regular with respect to $\omega \in \Omega_\alpha$ (since $\Omega_\alpha$ 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

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

for some $h > 0$ and some $\gamma(l) > 0$, $(\Phi, \varphi)$ is of class $C^l$ with respect to $\omega$: 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 \((\Phi, \varphi)\) 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 [Pö01] for instance) but then the exponent \(\nu\) in (1.1) has to be deteriorate: given any \(\mu > \nu\), we have that \((\Phi, \varphi)\) is smooth with respect to \(\omega\) 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 [Pop04]) showed that one can even go further and obtain some Gevrey smoothness of \((\Phi, \varphi)\) under a stronger smallness condition; without going into these rather technical issues, let us just say that \((\Phi, \varphi)\) can be shown to be Gevrey with exponent \(1 + \mu\) provided the polynomially small threshold \(s^\nu\) 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 non-degenerate 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 [Pö01], one can write \(H\) as in Theorem A with

\[
P = P_f + P_h, \quad |P_f| \leq \varepsilon, \quad |P_h| \leq Mr^2
\]

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 \(\varepsilon\) and Theorem A can be applied; yet with such a choice it is obvious that because of the estimates for \(\varphi\) 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 [Pö01], 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

\[
|P| \lesssim \alpha rs^\nu
\]

is then obviously implied by the conditions

\[
|P_f| \lesssim \alpha s^\nu, \quad r \lesssim \alpha s^\nu
\]

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

\[
|P|_{r,s,h} \leq \gamma \alpha rs^\nu, \quad r \leq \delta M^{-1}\alpha s^\nu, \quad \alpha s^\nu \leq h \tag{1.3}
\]
where \( \nu = \tau + 1 \), \( \gamma \) and \( \delta \) are small constants depending only on \( n \) and \( \tau \). 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 quasi-periodic of frequency \( \omega \). 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 [Vil08]; alternatively, a slight modification in the proof in [BF19] yields the same result.

\section{Sketch of proof}

In this section, we will sketch the proof of Theorem A and Theorem B; actually, we will simply indicate the modifications with respect to [Pöss01] and we will use the same convention for implicit constants depending only on \( n \) and \( \tau \).

\begin{proposition}
Let \( H = N + P \), and suppose that \( |P|_{s,r,h} \leq \varepsilon \) with
\begin{align}
\begin{cases}
\varepsilon < \alpha \eta^2 r \sigma^\nu, \\
\varepsilon < hr, \\
h \leq \alpha (2K^\nu)^{-1}, \quad K = \sigma^{-1} \log(n\eta^{-2})
\end{cases}
\end{align}
where \( 0 < \eta < 1/8 \) and \( 0 < \sigma < s/5 \). 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
\[ |P_+| \leq 9\eta^2 \varepsilon \]
and
\begin{align}
\begin{cases}
|W(\Phi - \text{Id})|, \ |W(D\Phi - \text{Id})W^{-1}| \ll (\alpha r \sigma^\nu)^{-1} \varepsilon \\
|\phi - \text{Id}|, \ h|D\varphi - \text{Id}|_L \ll r^{-1} \varepsilon
\end{cases}
\end{align}
uniformly on \( D_{\eta r, s-5\sigma} \times O_h \) and \( O_{h/4} \), with \( W = \text{Diag}(r^{-1} \text{Id}, \sigma^{-1} \text{Id}) \).

The above proposition is a variant of the KAM step of [Pöss01], which we already used in [Bou19]. The only difference is that in [Pöss01], instead of (2.1) the following conditions are imposed
\begin{align}
\begin{cases}
\varepsilon < \alpha \eta^2 r \sigma^\nu, \\
\varepsilon < hr, \\
h \leq \alpha (2K^\nu)^{-1}
\end{cases}
\end{align}
with a free parameter \( K \in \mathbb{N}^* \), leading to the following estimate
\[ |P_+| \ll (\varepsilon (r \sigma^\nu)^{-1} + \eta^2 + K^\nu e^{-K\sigma})\varepsilon. \]
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 $\theta$ of degree $K$; to obtain such an approximation, in [Pöß01] the author simply truncates the Taylor expansion in $I$ and the Fourier expansion in $\theta$ to obtain the following approximation error

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

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 [Rüs01] (choosing, in the latter reference, $\beta_1 = \cdots = \beta_n = 1/2$ and $\delta^{1/2} = 2\eta$ for $\delta \leq 1/4$); with the choice$^3$ of $K$ as in (2.1), 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 $\theta$) and a simpler error

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

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 [Pöß01].

Now, at variance with [Pöß01], we will use Proposition 2.1 in an iterative scheme with a linear speed of convergence as $\eta$ 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 $K_i = \sigma_i^{-1}\log(m\eta^2) = \sigma_i^{-1}$, we set

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

and the condition $\alpha\nu \leq h$ implies in particular than $h_0 \leq h$. 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 of (2.1) amount to $\varepsilon_i < \alpha r_i \sigma_i\nu$ which, in view of our choice of $\eta$, 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 $\varepsilon \leq \gamma \alpha r s^\nu$. Once we can iterate Proposition 2.1 infinitely many times, the convergence proof and the final estimates follow exactly as in [Pöß01], since the sequences $\varepsilon_i(h_ir_i)^{-1}$ and $\varepsilon_i(h_i^2r_i)^{-1}$ decrease geometrically, again by our choice of $\eta$. This concludes the sketch of proof.

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

$^3$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 [Rüs05] for instance.
Proposition 2.2. Let $H = N + P + Q$, suppose that $|P|_{s,r,h} \leq \varepsilon$, $|Q|_{r,h} \leq Mr^2$ with $Q$ integrable and at least quadratic in $I$ and

\[
\begin{aligned}
\varepsilon &< \alpha \eta^2 r \sigma^\nu, \\
r &< M^{-1} \alpha \eta^2 \sigma^\nu, \\
\varepsilon &< hr, \\
h &\leq \alpha (2K^\nu)^{-1}, \quad K = n\sigma^{-1} \log(\eta^{-2})
\end{aligned}
\tag{2.6}
\]

where $0 < \eta < 1/4$ and $0 < \sigma < s/5$. 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 $\{., .\}$ denotes the Poisson bracket and $[.]$ 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 [Pö01] (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

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

As compared to [Pö01], there is an extra term in $P_+$ coming from $Q$, whose contribution is easily bounded by the simple Poisson bracket

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

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.

Acknowledgements. The author have benefited from partial funding from the ANR project Beyond KAM.
References


