A PROOF OF THE INVARIANT TORUS THEOREM OF KOLMOGOROV

JACQUES FÉJOZ

Abstract. A variant of Kolmogorov's initial proof is given, in terms of a group of symplectic transformations and of an elementary fixed point theorem.

Let $\mathcal{H}$ be the space of Hamiltonians which are real analytic in neighborhoods of $\mathrm{T}^n_0 := \mathbb{T}^n \times \{0\}$ in $\mathbb{T}^n \times \mathbb{R}^n = \{(\theta, r)\}$ ($\mathbb{T}^n = \mathbb{R}^n/\mathbb{Z}^n$). The vector field associated with $H \in \mathcal{H}$ is

$$\dot\theta = \partial_r H, \quad \dot r = -\partial_\theta H.$$

For $\alpha \in \mathbb{R}^n$, let $\mathcal{K}^\alpha$ be the subspace of Hamiltonians $K \in \mathcal{H}$ such that $K|_{\mathrm{T}^n_0}$ is constant (i.e. T^n_0 is invariant) and $\vec{K}|_{\mathrm{T}^n_0} = \alpha$:

$$\mathcal{K}^\alpha = \{K \in \mathcal{H},\ \exists c \in \mathbb{R},\ K(\theta, r) = c + \alpha \cdot r + O(r^2)\}, \quad \alpha \cdot r = \alpha_1 r_1 + \cdots + \alpha_n r_n,$$

where $O(r^2)$ are terms of the second order in r, which depend on θ.

Let also $\mathcal{G}$ be the space of symplectic transformations G real analytic in neighborhoods of T^n_0 in $\mathbb{T}^n \times \mathbb{R}^n$, of the following form:

$$G(\theta, r) = (\varphi(\theta), (r + \rho(\theta)) \cdot \varphi'(\theta)^{-1}),$$

where φ is an analytic transformation of $\mathbb{T}^n$ fixing the origin (meant to straighten the flow on an invariant torus), and ρ is a closed 1-form on $\mathbb{T}^n$ (or an irrotational vector field, meant to straighten an invariant torus). To be more precise, $\rho(\theta)\, d\theta$ is a closed 1-form on $\mathbb{T}^n$.

We fix $\alpha \in \mathbb{R}^n$ Diophantine ($0 < \gamma \ll 1$ and $\tau > n - 1$; see [6]):

$$|k \cdot \alpha| \geq \gamma |k|^{-\tau} \quad (\forall k \in \mathbb{Z}^n \setminus \{0\}), \quad |k| = |k_1| + \cdots + |k_n|$$

and

$$K^o(\theta, r) = c^o + \alpha \cdot r + Q^o(\theta) \cdot r^2 + O(r^3) \in \mathcal{K}^\alpha$$

such that the average of the quadratic form valued function Q^o is non-degenerate:

$$\det \int_{\mathbb{T}^n} Q^o(\theta)\, d\theta \neq 0.$$

Theorem 1 (Kolmogorov [1, 4, 5])**.** *For every $H \in \mathcal{H}$ close to K^o, there exists $(K, G) \in \mathcal{K}^\alpha \times \mathcal{G}$ close to (K^o, id) such that $H = K \circ G$ in some neighborhood of $G^{-1}(\mathrm{T}^n_0)$.*

See [3, 6, 9] and references therein for background. Here we present Kolmogorov's initial proof in its simplest form, but in the concise functional language of [2]. A beautiful paper claims to give the "shortest complete KAM proof for perturbations of integrable vector fields available so far" [7]. In fact, the only significant difference of this paper with Kolmogorov's induction is that at each step of the induction, Rüssmann [8] and Pöschel [7] optimize the error by taking a well chosen polynomial approximation of the right hand side of the cohomological equation; the convergence is slower and the range of convergence probably larger, but, as for the length of the proof, readers will judge by themselves.

Define complex extensions $\mathbb{T}^n_{\mathbb{C}} = \mathbb{C}^n/\mathbb{Z}^n$ and $\mathrm{T}^n_{\mathbb{C}} = \mathbb{T}^n_{\mathbb{C}} \times \mathbb{C}^n$, and neighborhoods ($0 < s < 1$)

$$\mathbb{T}^n_s = \{\theta \in \mathbb{T}^n_{\mathbb{C}},\ \max_{1 \leq j \leq n} |\mathrm{Im}\, \theta_j| \leq s\} \quad \text{and} \quad \mathrm{T}^n_s = \{(\theta, r) \in \mathrm{T}^n_{\mathbb{C}},\ \max_{1 \leq j \leq n} \max(|\mathrm{Im}\, \theta_j|, |r_j|) \leq s\}.$$

For $U = \mathbb{T}^n_s$ or T^n_s, we will denote by $\mathcal{A}(U)$ the space of real holomorphic maps from the interior of U to $\mathbb{C}$ which extend continuously on U, endowed with the supremum norm $|\cdot|_s$.

• Let $\mathcal{H}_s = \mathcal{A}(\mathrm{T}^n_s)$, such that $\mathcal{H} = \bigcup_s \mathcal{H}_s$.

There exist $s_0 < s$ and $\epsilon_0 > 0$ such that $K^o \in \mathcal{H}_s$ and, for all $H \in \mathcal{H}_{s_0}$ with $|H - K^o|_{s_0} \le \epsilon_0$,

$$\left|\det \int_{\mathbb{T}^n} \frac{\partial^2 H}{\partial r^2}(\theta, 0)\, d\theta\right| \ge \frac{1}{2}\left|\det \int_{\mathbb{T}^n} \frac{\partial^2 K^o}{\partial r^2}(\theta, 0)\, d\theta\right| \ne 0. \tag{1}$$

Hereafter we assume that s is always $\ge s_0$. Let

$$\mathcal{K}^\alpha_s = \{K \in \mathcal{H}_s \cap \mathcal{K}^\alpha,\ |K - K^o|_{s_0} \le \epsilon_0\};$$

the vector space $\vec{\mathcal{K}}_s$ directing $\mathcal{K}^\alpha_s$ identifies with $\mathbb{R} \times O(r^2)$.

• Let $\mathcal{D}_s$ be the space of real holomorphic invertible transformations $\varphi : \mathbb{T}^n_s \to \varphi(\mathbb{T}^n_s) \subset \mathbb{T}^n_{\mathbb{C}}$ with $\varphi(0) = 0$, and $\mathcal{Z}_s$ be the space of real holomorphic closed 1-forms on $\mathbb{T}^n_s$ (seen as maps $\mathbb{T}^n_s \to \mathbb{C}^n$). Elements of $\mathcal{G}_s = \mathcal{Z}_s \times \mathcal{D}_s$ define symplectic transformations of the phase space,

$$G : \mathrm{T}^n_s \to \mathrm{T}^n_{\mathbb{C}}, \quad (\theta, r) \mapsto (\varphi(\theta), (\rho(\theta) + r) \cdot \varphi'(\theta)^{-1}), \quad G = (\rho, \varphi), \tag{2}$$

and of $\mathcal{H}_s$, $K \mapsto K \circ G$ (the latter Hamiltonian being defined on $G^{-1}(\mathrm{T}^n_s)$).

• Let $\mathfrak{d}_s := \{\dot\varphi \in \mathcal{A}(\mathbb{T}^n_s)^n,\ \dot\varphi(0) = 0\}$ with norm $|\dot\varphi|_s := \max_{\theta \in \mathbb{T}^n_s} \max_{1 \le j \le n} |\dot\varphi_j(\theta)|$, be the space of vector fields on $\mathbb{T}^n_s$ which vanish at 0. Similarly, let $|\dot\rho|_s = \max_{\theta \in \mathbb{T}^n_s} \max_{1 \le j \le n} |\dot\rho_j(\theta)|$ on $\mathcal{Z}_s$. An element $\dot G = (\dot\rho, \dot\varphi)$ of the sum $\mathfrak{g}_s = \mathcal{Z}_s \oplus \mathfrak{d}_s$ (with norm $|(\dot\rho, \dot\varphi)|_s = \max(|\dot\rho|_s, |\dot\varphi|_s)$) identifies with the locally Hamiltonian vector field

$$\dot G : \mathrm{T}^n_s \to \mathbb{C}^{2n}, \quad (\theta, r) \mapsto (\dot\varphi(\theta), \dot\rho(\theta) - r \cdot \dot\varphi'(\theta)). \tag{3}$$

Constants $\gamma_i, \tau_i, c_i, t_i$ below do not depend on s or σ.

Lemma 0. *If $\dot G \in \mathfrak{g}_{s+\sigma}$ and $|\dot G|_{s+\sigma} \le \gamma_0 \sigma^2$, then $\exp \dot G \in \mathcal{G}_s$ and $|\exp \dot G - id|_s \le c_0 \sigma^{-1} |\dot G|_{s+\sigma}$.*

Proof. Let $\chi_s \in \mathcal{A}(\mathrm{T}^n_s)^{2n}$, with norm $\|v\|_s = \max_{\theta \in \mathrm{T}^n_s} \max_{1 \le j \le 2n} |v_j(\theta)|$. Let $\dot G \in \mathfrak{g}_{s+\sigma}$ with $|\dot G|_{s+\sigma} \le \gamma_0 \sigma^2$, $\gamma_0 := (36n)^{-1}$. Using definition (3) and Cauchy's inequality, we see that if $\delta := \sigma/3$,

$$\|\dot G\|_{s+2\delta} = \max\left(|\dot\varphi|_{s+2\delta}, |\dot\rho + r \cdot \dot\varphi'(\theta)|_{s+2\delta}\right) \le 2n\delta^{-1} |\dot G|_{s+3\delta} \le \delta/2.$$

Let $I_s = [0,1] \times i[-s, s]$ and $\mathcal{F} := \left\{f \in \mathcal{A}(I_s \times \mathbb{T}^n_s)^{2n},\ \forall (t,\theta) \in I_s \times \mathbb{T}^n_s,\ |f(t,\theta)|_s \le \delta\right\}$. By Cauchy's inequality, the Lipschitz constant of the Picard operator

$$P : \mathcal{F} \to \mathcal{F}, \quad f \mapsto Pf, \quad (Pf)(t, \theta) = \int_0^t \dot G(\theta + f(s, \theta))\, ds$$

is $\le 1/2$. Hence, P possesses a unique fixed point $f \in \mathcal{F}$, such that $f(1, \cdot) = \exp(\dot G) - \mathrm{id}$ and $|f(1, \cdot)|_s \le \|\dot G\|_{s+\delta} \le c_0 \sigma^{-1} |\dot G|_{s+\sigma}$, $c_0 = 6n$.

Also, $\exp \dot G \in \mathcal{G}_s$ because at all times the curve $\exp(t\dot G)$ is tangent to $\mathcal{G}_s$ (another proof uses the method of the variation of constants). ☐

Lemma 1 (Cohomological equation)**.** *For all $(K, \dot H) \in \mathcal{K}^\alpha_{s+\sigma} \times \mathcal{H}_{s+\sigma}$, there exists a unique $(\dot K, \dot G) \in \vec{\mathcal{K}}_s \times \mathfrak{g}_s$ such that*

$$\dot K + K' \cdot \dot G = \dot H \quad \textit{and} \quad \max(|\dot K|_s, |\dot G|_s) \le c_1 \sigma^{-t_1} (1 + |K|_{s+\sigma})\, |\dot H|_{s+\sigma}.$$

Proof. We want to solve the linear (cohomological) equation $\dot K + K' \cdot \dot G = \dot H$. Write

$$\begin{cases} K(\theta, r) = c + \alpha \cdot r + Q(\theta) \cdot r^2 + O(r^3) \\ \dot K(\theta, r) = \dot c + \dot K_2(\theta, r), & \dot c \in \mathbb{R}, \quad \dot K_2 \in O(r^2) \\ \dot G(\theta, r) = (\dot\varphi(\theta), R + S'(\theta) - r \cdot \dot\varphi'(\theta)), & \dot\varphi \in \chi_s, \quad \dot R \in \mathbb{R}^n, \quad \dot S \in \mathcal{A}(\mathbb{T}^n_s). \end{cases}$$

Expanding the equation in powers of r yields

$$\left(\dot c + (\dot R + \dot S') \cdot \alpha\right) + r \cdot \left(-\dot\varphi' \cdot \alpha + 2Q \cdot (\dot R + \dot S')\right) + \dot K_2 = \dot H =: \dot H_0 + \dot H_1 \cdot r + O(r^2), \tag{4}$$

where the term $O(r^2)$ on the right hand side does not depend on $\dot K_2$.

If $g \in \mathcal{A}(\mathbb{T}^n_{s+\sigma})$ has zero average, there is a unique function $f \in \mathcal{A}(\mathbb{T}^n_s)$ of zero average such that $L_\alpha f := f' \cdot \alpha = g$, and $|f|_s \le c\sigma^{-t}|g|_{s+\sigma}$, $c = c_{\gamma,\tau,n}$. Using the Diophantine condition and Cauchy's inequality to estimate Fourier coefficients, the unique formal solution indeed satisfies

$$|f|_s = \left| \sum_{k \in \mathbb{Z}^n \setminus \{0\}} \frac{g_k}{ik \cdot \alpha} e^{ik\cdot\theta} \right|_s \le \frac{|g|_{s+\sigma}}{\gamma} \sum_k |k|^\tau \, e^{-|k|\,\sigma},$$

and the wanted upper bound then follows from an elementary estimate [6].

Equation (4) is triangular in the unknowns and successively yields:

$$\begin{cases} \dot S & = L_\alpha^{-1}\left(\dot H_0 - \int_{\mathbb{T}^n} \dot H_0(\theta)\, d\theta\right) \\ \dot R & = \frac{1}{2}\left(\int_{\mathbb{T}^n} Q(\theta)\, d\theta\right)^{-1} \int_{\mathbb{T}^n} \left(\dot H_1(\theta) - 2Q(\theta) \cdot \dot S'(\theta)\right)\, d\theta \\ \dot\varphi & = \dot\varphi_1 - \dot\varphi_1(0), \quad \dot\varphi_1 = L_\alpha^{-1}\left(\dot H_1(\theta) - 2Q(\theta) \cdot (\dot R + \dot S'(\theta))\right) \\ \dot c & = \int_{\mathbb{T}^n} \dot H_0(\theta)\, d\theta - \dot R \cdot \alpha \\ \dot K_2 & = O(r^2). \end{cases}$$

The wanted estimate follows from Cauchy's inequality. □

Let us bound the discrepancy between the action of $\exp(-\dot G)$ and the infinitesimal action of $-\dot G$.

Lemma 2 (Quadratic error)**.** *For all $(K, \dot H) \in \mathcal{K}^\alpha_{s+\sigma} \times \mathcal{H}_{s+\sigma}$ such that $(1 + |K|_{s+\sigma})|\dot H|_{s+\sigma} \le \gamma_2\sigma^{\tau_2}$, if $(\dot K, \dot G) \in \vec{\mathcal{K}} \times \mathfrak{g}_s$ solves the equation $\dot K + K' \circ \dot G = \dot H$ (lemma 1), then $\exp \dot G \in \mathcal{G}_s$, $|\exp \dot G - id|_s \le \sigma$ and*

$$|(K + \dot H) \circ \exp(-\dot G) - (K + \dot K)|_s \le c_2\sigma^{-t_2}(1 + |K|_{s+\sigma})^2 |\dot H|^2_{s+\sigma}.$$

Proof. Set $\delta = \sigma/2$. Lemmas 0 and 1 show that, under the hypotheses for some constant γ_2 and for $\tau_2 = t_1 + 1$, we have $|\dot G|_{s+\delta} \le \gamma_0\delta^2$ and $|\exp \dot G - \mathrm{id}\,|_s \le \delta$.

Let $H = K + \dot H$. Taylor's formula says

$$\mathcal{H}_s \ni H \circ \exp(-\dot G) = H - H' \cdot \dot G + \left(\int_0^1 (1-t)\, H'' \circ \exp(-t\dot G)\, dt\right) \cdot \dot G^2$$

or, using the fact that $H = K + \dot K + K' \cdot \dot G$,

$$H \circ \exp(-\dot G) - (K + \dot K) = -(\dot K + K' \cdot \dot G)' \cdot \dot G + \left(\int_0^1 (1-t)\, H'' \circ \exp(-t\dot G)\, dt\right) \cdot \dot G^2.$$

The wanted estimate thus follows from the estimate of lemma 1 and Cauchy's inequality. □

End of the proof of theorem 1. Let $B_{s,\sigma} = \{(K, \dot H) \in \mathcal{K}^\alpha_{s+\alpha} \times \mathcal{H}_{s+\sigma},\ |K|_{s+\sigma} \le \epsilon_0,\ |\dot H|_{s+\sigma} \le (1+\epsilon_0)^{-1}\gamma_2\sigma^{\tau_2}\}$ (recall (1)).

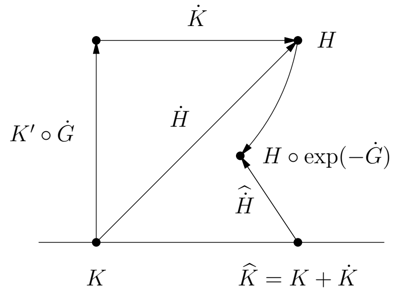
By lemmas 1 and 2, the map $\phi : B_{s,\sigma} \to \mathcal{K}^\alpha_s \times \mathcal{H}_s$,

$$\phi(K, \dot H) = (K + \dot K, (K + \dot H) \circ \exp(-\dot G) - (K + \dot K)),$$

satisfies, if $(\hat K, \widehat{\dot H}) = \phi(K, \dot H)$,

$$|\hat K - K|_s \le c_3 \sigma^{-t_3} |\dot H|_{s+\sigma}, \ |\widehat{\dot H}|_s \le c_3 \sigma^{-t_3} |\dot H|^2_{s+\sigma}.$$

Proposition 3 in the appendix applies and shows that if $H - K^o$ is small enough in $\mathcal{H}_{s+\sigma}$, the sequence $(K_j, \dot H_j) = \phi^j(K^o, H - K^o)$, $j \ge 0$, converges towards some $(K, 0)$ in $\mathcal{K}^\alpha_s \times \mathcal{H}_s$.



Let us keep track of the $\dot G_j$'s solving with the $\dot K_j$'s the successive linear equations $\dot K_j + K'_j \cdot \dot G_j = \dot H_j$ (lemma 1). At the limit,

$$K := K^o + \dot K_0 + \dot K_1 + \cdots = H \circ \exp(-\dot G_0) \circ \exp(-\dot G_1) \circ \cdots .$$

Moreover, lemma 1 shows that $|\dot G_j|_{s_{j+1}} \le c_4 \sigma_j^{-t_4} |\dot H_j|_{s_j}$, hence the transformations $\gamma_j := \exp(-\dot G_0) \circ \cdots \circ \exp(-\dot G_j)$, which satisfy

$$|\gamma_n - \mathrm{id}\,|_{s_{n+1}} \le |\dot G_0|_{s_1} + \cdots + |\dot G_n|_{s_{n+1}},$$

form a Cauchy sequence and have a limit $\gamma \in \mathcal{G}_s$. At the expense of decreasing $|H - K^o|_{s+\sigma}$, by the inverse function theorem, $G := \gamma^{-1}$ exists in $\mathcal{G}_{s-\delta}$ for some $0 < \delta < s$, so that $H = K \circ G$. □

Remark. The uniqueness property of lemma 1 and the estimate of lemma 2 show that if $\tilde G$ is in some small neighborhood of the identity in $\mathcal{G}$ and $K \circ \tilde G \in \mathcal{K}^\alpha$ then $\tilde G = \mathrm{id}$. The local uniqueness of the pair (K, G) such that $H = K \circ G$ follows directly.

Appendix. Quadratic convergence

Let $(E_s, |\cdot|_s)_{0<s<1}$ and $(F_s, |\cdot|_s)_{0<s<1}$ be two decreasing families of Banach spaces with increasing norms. On $E_s \times F_s$, set $|(x,y)|_s = \max(|x|_s, |y|_s)$. Fix $C, \gamma, \tau, c, t > 0$.

Let

$$\phi : B_{s,\sigma} := \{(x,y) \in E_{s+\sigma} \times F_{s+\sigma},\ |x|_{s+\sigma} \le C,\ |y|_{s+\sigma} \le \gamma\sigma^\tau\} \to E_s \times F_s$$

be maps such that if $(X, Y) = \phi(x, y)$,

$$|X - x|_s \le c\sigma^{-t} |y|_{s+\sigma} \quad \text{and} \quad |Y|_s \le c\sigma^{-t} |y|^2_{s+\sigma}.$$

In the proof of theorem 1, "$|x|_{s+\sigma} \le C$" allows us to bound the determinant of $\int_{\mathbb{T}^n} Q(\theta)\, d\theta$ away from 0, while "$|y|_{s+\sigma} \le \gamma\sigma^\tau$" ensures that $\exp \dot G$ is well defined.

Lemma 3. *Given $s < s+\sigma$ and $(x,y) \in B_{s,\sigma}$ such that $|y|_{s+\sigma}$ is small, the sequence $(\phi^j(x,y))_{j\ge 0}$ exists and converges towards a fixed point $(\xi, 0)$ in $B_{s,0}$.*

Proof. It is convenient to first assume that the sequence is defined and $(x_j, y_j) := F^j(x,y) \in B_{s_j,\sigma_j}$, for $s_j := s + 2^{-j}\sigma$ and $\sigma_j := s_j - s_{j+1}$. We may assume $c \ge 2^{-t}$, so that $d_j := c\sigma_j^{-t} \ge 1$. By induction, and using the fact that $\sum 2^{-k} = \sum k 2^{-k} = 2$,

$$|y_j|_{s_j} \le d_{j-1} |y_{j-1}|^2_{s_{j-1}} \le \cdots \le |y|^{2^j}_{s+\sigma} \prod_{0 \le k \le j-1} d_k^{2^{k+1}} \le \Big(|y|_{s+\sigma} \prod_{k \ge 0} d_k^{2^{-k-1}}\Big)^{2^j} = \big(c(4\sigma)^{-t} |y|_{s+\sigma}\big)^{2^j}.$$

Given that $\sum_{n\ge 0} \mu^{2^n} \le 2\mu$ if $2\mu \le 1$, by induction we see that if $|y|_{s+\sigma}$ is small enough, (x_j, y_j) exists in B_{s_j,σ_j} for all $j \ge 0$, y_j converges to 0 in F_s and the series $x_j = x_0 + \sum_{0\le k\le j-1}(x_{k+1} - x_k)$ converges absolutely towards some $\xi \in E_s$ with $|\xi|_s \le C$. □

Thank you to the reviewer, Alain Albouy, Alain Chenciner, Ivar Ekeland, Andreas Knauf, Jessica Massetti and Éric Séré for their scrutiny and the improvements they have suggested.

The author has been supported by the French ANR (projets KamFaible ANR-07-BLAN-0361 and DynPDE ANR-10-BLAN-0102).

References

[1] L. Chierchia. A. N. Kolmogorov's 1954 paper on nearly-integrable Hamiltonian systems. A comment on: "On conservation of conditionally periodic motions for a small change in Hamilton's function". *Regul. Chaotic Dyn.*, 13(2):130–139, 2008.

[2] J. Féjoz and M. Garay. Un théorème sur les actions de groupes de dimension infinie. *C. R. Math. Acad. Sci. Paris*, 348(7-8):427–430, 2010.

[3] J. Hubbard and Y. Ilyashenko. A proof of Kolmogorov's theorem. *Discrete Contin. Dyn. Syst.*, 10(1-2):367–385, 2004. Partial differential equations and applications.

[4] A. N. Kolmogorov. On the conservation of conditionally periodic motions for a small change in Hamilton's function. *Dokl. Akad. Nauk SSSR (N.S.)*, 98:527–530, 1954.

[5] A. N. Kolmogorov. Preservation of conditionally periodic movements with small change in the Hamilton function. In *Stochastic behavior in classical and quantum Hamiltonian systems (Volta Memorial Conf., Como, 1977)*, volume 93 of *Lecture Notes in Phys.*, pages 51–56. Springer, Berlin, 1979.

[6] J. Pöschel. A lecture on the classical KAM theorem. In *Smooth ergodic theory and its applications (Seattle, WA, 1999)*, volume 69 of *Proc. Sympos. Pure Math.*, pages 707–732. Amer. Math. Soc., Providence, RI, 2001.

[7] J. Pöschel. KAM à la R. *Regul. Chaotic Dyn.*, 16(1-2):17–23, 2011.

[8] H. Rüssmann. KAM iteration with nearly infinitely small steps in dynamical systems of polynomial character. *Discrete Contin. Dyn. Syst. Ser. S*, 3(4):683–718, 2010.

[9] M. B. Sevryuk. The classical KAM theory at the dawn of the twenty-first century. *Mosc. Math. J.*, 3(3):1113–1144, 1201–1202, 2003. Dedicated to V. I. Arnold on the occasion of his 65th birthday.

Université Paris-Dauphine, CEREMADE (UMR 7534) & Observatoire de Paris, IMCCE (UMR 8028)

E-mail address: jacques.fejoz@dauphine.fr