Kolmogorov's Theorem

Travis Westura

Cornell University

April 20, 2015

Our goal is to understand the following theorem.

Kolmogorov's Theorem

Let $\rho, \gamma > 0$ *be given, and let* $h(\mathbf{q}, \mathbf{p}) = h_0(\mathbf{p}) + h_1(\mathbf{q}, \mathbf{p})$ *be a Hamiltonian, with* $h_0, h_1 \in \mathcal{A}_\rho$ *and* $\|h\|_\rho \leq 1$. *Suppose the Taylor polynomial of* h_0 *is*

$$h_0(\mathbf{p}) = a + \omega\mathbf{p} + \frac{1}{2}\mathbf{p} \cdot C\mathbf{p} + o(|\mathbf{p}|^2),$$

with $\omega \in \Omega_\gamma$ *and* C *is symmetric and invertible. Then for any* $\rho_* \leq \rho$, *there exists* $\epsilon > 0$, *which depends on* C *and* γ, *but not on the remainder term in* $o(|\mathbf{p}|^2)$, *such that if* $\|h_1\|_\rho \leq \epsilon$, *there exists a symplectic mapping* $\Phi : A_{\rho_*} \to A_\rho$ *such that if we set* $(\mathbf{q}, \mathbf{p}) = \Phi(\mathbf{Q}, \mathbf{P})$ *and* $H = h \circ \Phi$, *we have*

$$H(\mathbf{Q}, \mathbf{P}) = A + \omega\mathbf{P} + R(\mathbf{Q}, \mathbf{P}),$$

with $R(\mathbf{Q}, \mathbf{P}) \in O(|\mathbf{P}|^2)$.

Kolmogorov's Theorem

The **KAM** theory is named after

Andreï **K**olmogorov,

Vladimir **A**rnold,

Jürgen **M**oser.

Kolmogorov gave an original proof in the 1950s but never published it. Arnold gave a proof of the theorem in 1963 and Moser published a different but related result in 1962.

A Motivating Example

The solar system when planets have zero mass.
A system with n masses m_i and positions $\mathbf{x}_i$ satisfies Newton's $\mathbf{F} = m\mathbf{a}$ law:

$$m_i \ddot{\mathbf{x}}_i = \sum_{j \neq i} G m_i m_j \frac{\mathbf{x}_j - \mathbf{x}_i}{|\mathbf{x}_j - \mathbf{x}_i|^3}.$$

We take the sun to be given by index 0 and the plants to be given by the indices $1, \ldots n$.

A Motivating Example

Take the sun to be at the center of the solar system ($\mathbf{x}_0 = 0$) and setting the masses of the planets to zero ($m_i = 0$ for $i \neq 0$).
We obtain $\mathbf{x}_0'' = 0$ and for $i \neq 0$

$$\mathbf{x}_i'' = -Gm_0 \frac{\mathbf{x}_i}{|\mathbf{x}_i|^3}.$$

Hamiltonian Mechanics

Named for William Rowan Hamilton (1805-1865).
A reformulation of classical Newtonian mechanics.
Relies on a symplectic structure.

Hamiltonian Mechanics

Let (X, σ) be a symplectic manifold. That is, X is a differentiable manifold and σ is a nowhere vanishing 2-form such that $d\sigma = 0$. A function H on a X has a *symplectic gradient* denoted $\nabla_\sigma H$. This gradient is the unique vector field such that for any vector field ξ, we have

$$\sigma(\xi, \nabla_\sigma H) = dH(\xi).$$

We will consider a *Hamiltonian differential equation* that makes use of this symplectic gradient:

$$\dot{\mathbf{x}} = (\nabla_\sigma H)(\mathbf{x}).$$

Hamiltonian Mechanics

Let's compare the symplectic gradient and the gradient ∇f of a function $f : \mathbb{R}^n \to \mathbb{R}$ from multivariable calculus.

The gradient is the unique vector field such that

$$df(\xi) = \langle \xi, \nabla f \rangle.$$

We could also consider the "gradient differential equation"

$$\dot{\mathbf{x}} = \nabla f(\mathbf{x}).$$

Since f increases along its solutions, they can never return to their starting point.

This prohibits recurrence, whereas the Hamiltonian equations allow and almost impose recurrence (see the Poincaré recurrence theorem).

Hamiltonian Mechanics

For a simple example of a Hamiltonian system, consider $X = \mathbb{R}^{2n}$ with coordinates $(q_1, \ldots, q_n, p_1, \ldots, p_n)$. Take $\sigma = \sum_i dp_i \wedge dq_i$.

Our Hamiltonian differential equation is

$$\dot{q}_i = \frac{\partial H}{\partial p_i},$$

$$\dot{p}_i = -\frac{\partial H}{\partial q_i},$$

where these are called the *Hamiltonian equations of motion*.

Hamiltonian Mechanics

How does the Solar System fit this model? Consider the case of a single body of zero mass. (It is sufficient to consider the single body system, since the planets having zero mass means that they do not effect each other). For $\mathbf{x} \in \mathbb{R}^2$ the equation $\ddot{\mathbf{x}} = -\frac{\mathbf{x}}{|\mathbf{x}|^3}$ is Hamilton's equation for the manifold $X = \mathbb{R}^2 \times \mathbb{R}^2$ with points $(\mathbf{q}, \mathbf{p})$, standard symplectic form

$$\sigma = dp_1 \wedge dq_1 + dp_2 \wedge dq_2,$$

and Hamiltonian

$$H(\mathbf{q}, \mathbf{p}) = \frac{1}{2}(p_1^2 + p_2^2) - \frac{1}{\sqrt{q_1^2 + q_2^2}}.$$

Hamiltonian Mechanics

After some computation, we can obtain

$$\nabla_\sigma H = \left(\mathbf{p}, -\frac{1}{(q_1^2+q_2^2)^{\frac{3}{2}}}\mathbf{q}\right).$$

The Hamiltonian Differential Equation $\dot{\mathbf{x}} = \nabla_\sigma H(\mathbf{x})$ is

$$q_1' = p_1 \qquad p_1' = -\frac{q_1}{(q_1^2+q_2^2)^{\frac{3}{2}}},$$

$$q_2' = p_2 \qquad p_2' = -\frac{q_2}{(q_1^2+q_2^2)^{\frac{3}{2}}}.$$

From here we can recover $\ddot{\mathbf{x}} = -\dfrac{\mathbf{x}}{|\mathbf{x}|^3}$.

Hamiltonian Mechanics

The vector field $\nabla_\sigma f$ has a flow ϕ_f^t such that

> ϕ_f^t preserves f: $f \circ \phi_f^t = f$,
>
> ϕ_f^t preserves σ: $(\phi_f^t)^* \sigma = \sigma$.

Our proof of Kolmogorov's Theorem will involve the construction of a symplectic diffeomorphism Φ that is a composition of Hamiltonian flows.

We will need to construct Taylor polynomials of functions of the form $t \mapsto g \circ \phi_f^t$.
To do this we use the Poisson bracket.

Poisson Bracket

Let f and g be two functions on X. The Poisson bracket, denoted $\{f, g\}$, is defined by

$$\{f, g\} = \sigma(\nabla_\sigma f, \nabla_\sigma g) = df(\nabla_\sigma g) = -dg(\nabla_\sigma f).$$

Hamiltonian Mechanics

The gradient of the Poisson bracket corresponds to the Lie bracket:

$$\nabla_\sigma \{f, g\} = [\nabla_\sigma f, \nabla_\sigma g].$$

Proving this equality is a computation that is not difficult, but rather tedious.

Two function f and g commute if their Poisson bracket is zero: $\{f, g\} = 0$.

This implies that their flows commute:

$$\phi_f(s) \circ \phi_g(t) = \phi_g(t) \circ \phi_f(s).$$

Hamiltonian Mechanics

In the example involving the Hamiltonian equations of motion, we can write a Poisson bracket as

$$\{f, g\} = \sum_{i=1}^{n} \left(\frac{\partial f}{\partial q_i} \frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i} \frac{\partial g}{\partial q_i} \right).$$

The Poisson bracket will allow us to write Taylor polynomials:

$$f \circ \phi_g^t = f + t\{f, g\} + \frac{t^2}{2}\{\{f, g\}, g\} + \{\{\{f, g\}, g\}, g\} + \cdots .$$

Hamiltonian Mechanics

Denote the torus by $\mathbb{T} = \mathbb{R}/\mathbb{Z}$. A *totally integrable system* is a symplectic manifold $X = \mathbb{T}^n \times \mathbb{R}^n$ with variables $(\mathbf{q} \in \mathbb{T}^n, \mathbf{p} \in \mathbb{R}^n)$, symplectic form $\sigma = \sum_i dp_i \wedge dq_i$, and Hamiltonian function $H(\mathbf{p})$ that depends only on $\mathbf{p}$.

Irrationality

We want to make formal the notion of "suffiently irrational."
(Later we will see that this will help us to avoid dividing by zero).
A number θ is irrational if for all pairs $p, q \in \mathbb{Z}$, we have
$\left|\theta - \frac{p}{q}\right| \neq 0$.

So for θ to be "very" irrational, we want $\left|\theta - \frac{p}{q}\right|$ to be "very" different from 0, or "very big."
But that is not quite possible, as we can always approximate an irrational number by rational numbers.

Irrationality

Instead we want to consider what happens with small divisors.

We want $\left|\theta - \frac{p}{q}\right|$ to be "small" *only if* the denominator is "big."

Irrationality

We start with *diophantine conditions*.

Diophantine of Exponent d

A number θ is Diophantine of exponent d if there exists a constant $\gamma > 0$ such that for all coprime integers p and q we have

$$\left|\theta - \frac{p}{q}\right| > \frac{\gamma}{|q|^d}.$$

Irrationality

Note that it is a stronger condition to be Diophantine of a smaller exponent.

Diophantine Numbers with Full Measure

For all $\epsilon > 0$, the set of Diophantine numbers of exponent $2 + \epsilon$ is of full measure.

Irrationality

Diophantine Numbers with Full Measure.

Consider numbers in $\mathbb{R}/\mathbb{Z}$.
For all positive integers q, there are at most q elements of $\mathbb{Q}/\mathbb{Z}$ with denominator q in reduced form. Thus for all $\gamma > 0$,

$$\left|\left\{\theta \in \mathbb{R}/\mathbb{Z} : \left|\theta - \frac{p}{q}\right| < \frac{\gamma}{|q|^{2+\epsilon}}\right\}\right| < 2\gamma \sum_{q=1}^{\infty} \frac{1}{q^{1+\epsilon}}.$$

The intersection of these sets as $\gamma \to 0$ has measure 0, and this intersection is the complement of the set of Diophantine numbers of exponent $2 + \epsilon$. QED

Irrationality

We generalize "sufficiently irrational" from numbers to vectors.

Diophantine Vector in $\mathbb{R}^n$

Let $\omega = (\omega_1, \ldots, \omega_n)$. We say ω is Diophantine if there exists $\gamma > 0$ such that for all vectors with integer coefficients $(k_1, \ldots, k_n)$, we have

$$|k_1\omega_1 + \cdots + k_n\omega_n| \geq \frac{\gamma}{(k_1^2 + \cdots + k_n^2)^{\frac{n}{2}}}.$$

Let Ω_γ^n be the subset of such $\omega \in \mathbb{R}^n$.

We could also rewrite this condition as a dot product

$$|\mathbf{k} \cdot \omega| \leq \frac{\gamma}{|\mathbf{k}|^n}.$$

In our example of the solar system, the ω_i represent the frequencies of the planets' orbits.

Irrationality

How common are the vectors ω? Actually very common, as if we pick components ω_i at random, we will almost surely select such a vector.

Diophantine Vectors are of Full Measure

The union $\Omega = \bigcup_{\gamma > 0} \Omega_\gamma$ *is of full measure.*

Irrationality

Diophantine Vectors are of Full Measure.

Consider $S_{\mathbf{k},\gamma}$, the region around the hyperplane orthogonal to $\mathbf{k}$ of thickness $\dfrac{2\gamma}{|\mathbf{k}|^{n+1}}$.

The part within the unit cube Q has measure at most $\dfrac{M\gamma}{|\mathbf{k}|^{n+1}}$, with M the constant giving the max $(n-1)$-dim measure of a hyperplane intersected with Q.

The sum $\displaystyle\sum_{\mathbf{k}\in\mathbb{Z}^n-\{0\}} \frac{1}{|\mathbf{k}|^{n+1}}$ is finite, so the volume of $\displaystyle\bigcup_{\mathbf{k}\in\mathbb{Z}^n-\{0\}} S_{\mathbf{k},\gamma} \cap Q$ is at most a constant times γ.

The intersection of these sets has measure 0, and Ω is the complement, similar to the previous proof. QED

Torus

We view the solar system as an n dimensional torus.
Give the planets initial positions $\mathbf{a} = (a_1, \ldots, a_n)$.
Then the motion $t \mapsto \mathbf{a} + t\omega = (a_1 + t\omega_1, \ldots, a_n + t\omega_n)$ is called a *linear flow on* $(\mathbb{R}/\mathbb{Z})^n$ *in the direction* ω.
The trajectory dense on the torus if and only if the trajectory ω is irrational.
A motion $\mathbf{x}(t)$ of a perturbed system is "the same" as the motion of an unperturbed system $\mathbf{x}_1(t)$ dense on a torus T_1 means that $\mathbf{x}(t)$ is dense on a torus T and that there exists a homeomorphism $\Phi : T \to T_1$ such that $\Phi(\mathbf{x}(t)) = \mathbf{x}_1(t)$.

Torus

KAM for the Solar System

Let $\mathbf{x}_1(t)$ *be a motion of the zero-masses system for which the frequency vector* ω *is Diophantine. (Then* $\mathbf{x}_1(t)$ *is dense on the corresponding torus* T_1*).*
Then there exists $\epsilon > 0$ *such that, if the planets have masses* $m_i < \epsilon$*, then there exists a trajectory* $\mathbf{x}(t)$ *of the system that is dense on a torus* T*, and there exists a homeomorphism* $\Phi : T \to T_1$ *such that* $\Phi(\mathbf{x}(t)) = \mathbf{x}_1(t)$*.*
The probability of being on such a trajectory goes to 1 *as* ϵ *goes to* 0*.*

Another Look at the Theorem

Kolmogorov's Theorem

Let $\rho, \gamma > 0$ be given, and let $h(\mathbf{q}, \mathbf{p}) = h_0(\mathbf{p}) + h_1(\mathbf{q}, \mathbf{p})$ be a Hamiltonian, with $h_0, h_1 \in \mathcal{A}_\rho$ and $\|h\|_\rho \leq 1$. Suppose the Taylor polynomial of h_0 is

$$h_0(\mathbf{p}) = a + \omega\mathbf{p} + \frac{1}{2}\mathbf{p} \cdot C\mathbf{p} + o(|\mathbf{p}|^2),$$

with $\omega \in \Omega_\gamma$ and C is symmetric and invertible. Then for any $\rho_* \leq \rho$, there exists $\epsilon > 0$, which depends on C and γ, but not on the remainder term in $o(|\mathbf{p}|^2)$, such that if $\|h_1\|_\rho \leq \epsilon$, there exists a symplectic mapping $\Phi : A_{\rho_*} \to A_\rho$ such that if we set $(\mathbf{q}, \mathbf{p}) = \Phi(\mathbf{Q}, \mathbf{P})$ and $H = h \circ \Phi$, we have

$$H(\mathbf{Q}, \mathbf{P}) = A + \omega\mathbf{P} + R(\mathbf{Q}, \mathbf{P}),$$

with $R(\mathbf{Q}, \mathbf{P}) \in O(|\mathbf{P}|^2)$.

Analytic Functions

We will be solving a system of equations iteratively, using an analog of Newton's method.

Newton's method for solving an equation $f(\mathbf{x}) = 0$ consists of choosing a starting guess $\mathbf{x}_0$ and then defining

$$\mathbf{x}_{i+1} = \mathbf{x}_i - [Df(\mathbf{x}_i)]^{-1} f(\mathbf{x}_i).$$

But we might pick a "bad" initial guess for which Newton's method does not converge to a solution.

For example, taking $f(x) = x^3 - x + \frac{\sqrt{2}}{2}$ and the initial guess $x_0 = 0$, Newton's method will oscillate between 0 and $\frac{\sqrt{2}}{2}$.

Analytic Functions

In order to guarantee convergence, we use Kantorovich's theorem.
This theorem requires a bound on the second derivative.
We want an analogous condition for our problem.

Analytic Functions

We need a way to measure "size," that is, we need to choose a norm for our functions.
Let $X \subseteq \mathbb{C}^k$ be compact. [1]
Let the caligraphic letter $\mathcal{X}$ be the Banach algebra of continuous functions on X that are analytic in the interior and have the sup norm

$$\|f\|_X = \sum_{\mathbf{x} \in X} |f(\mathbf{x})|.$$

We use the absolute value as the standard Euclidean norm on $\mathbb{C}^n$.

[1] Or as Alex might like to say, let X be a compact.

Analytic Functions

We consider three regions:

$$B_\rho = \{\mathbf{p} \in \mathbb{C} : |\mathbf{p}| \leq \rho\},$$
$$C_\rho = \{\mathbf{q} \in \mathbb{C}^n/\mathbb{Z}^n : |\operatorname{Im}(\mathbf{q})| \leq \rho\},$$
$$A_\rho = C_\rho \times B_\rho = \{(\mathbf{q}, \mathbf{p}) \in \mathbb{C}^n/\mathbb{Z}^n \times \mathbb{C}^n : |\mathbf{p}| \leq \rho, |\operatorname{Im}(\mathbf{q})| \leq \rho\}.$$

Denote by $\mathcal{B}_\rho$, $\mathcal{C}_\rho$, and $\mathcal{A}_\rho$ the corresponding Banach algebras. We can expand element of $\mathcal{B}_\rho$ as power series, and elements of $\mathcal{C}_\rho$ as Fourier series:

$$f(\mathbf{z}) = \sum_{\mathbf{k} \in \mathbb{Z}^n} f_{\mathbf{k}} e^{2\pi i\, \mathbf{k}\cdot\mathbf{z}}.$$

Analytic Functions

We can bound derivatives of analytic functions on balls in terms of the values of the function itself.

Cauchy's Inequalities on Balls

If $f \in \mathcal{B}_\rho$, then

$$\|Df\|_{\rho-\delta} \leq \frac{1}{\delta} \|f\|_\rho \,,$$

$$\left\|D^2 f\right\|_{\rho-\delta} \leq \frac{4}{\delta^2} \|f\|_\rho \,.$$

As a corollary, the case $\delta = \rho$ bounds derivatives at the center of balls.

$$|Df(0)| \leq \frac{1}{\rho} \|f\|_\rho \,,$$

$$|D^2 f(0)| \leq \frac{4}{\rho^2} \|f\|_\rho \,.$$

Analytic Functions

Cauchy's Inequalities on Balls.

Take $\mathbf{z} \in B_{\rho-\delta}$ and $\mathbf{u} \in \mathbb{C}^n$. Since $B_\delta(\mathbf{z}) \subseteq B_\rho$, the function $g : t \mapsto f(\mathbf{z} + t\delta\mathbf{u})$ is defined on the unit disc, and the normal Cauchy inequality implies that

$$\delta|(Df(\mathbf{z}))\mathbf{u}| = |g'(0)| \leq \|g\|_1 \leq \|f\|_\rho .$$

Applying the argument twice gives the result for the second derivative:

$$|D^2 f(\mathbf{z})(\mathbf{u}, \mathbf{v})| \leq \frac{2}{\delta} \|Df(\mathbf{z})(\mathbf{u})\|_{\rho-\frac{\delta}{2}} |\mathbf{v}| \leq \frac{4}{\delta^2} \|f\|_\rho |\mathbf{u}||\mathbf{v}|.$$

QED

Smell of the Proof

Recall the final equation in Kolmogorov's theorem:

$$H(\mathbf{Q}, \mathbf{P}) = A + \omega \mathbf{P} + R(\mathbf{Q}, \mathbf{P}).$$

This is an equation for a diffeomorphism Φ. We solve for this diffeomorphism. To do so we would like to use Newton's method, but using it is not quite sufficient. However we can still do something with a similar flavor.

We obtain Φ as a limit of Φ_i, where

$$\Phi_i = \phi_i \circ \phi_{i-1} \circ \cdots \circ \phi_1.$$

Here, ϕ_i is the Hamiltonian flow for a Hamiltonian function g_i. This g_i is the unknown for which we solve.

Smell of the Proof

At the ith step we have a Hamiltonian $\tilde{h} = \Phi_i^* h$.
We expand $\tilde{h}$ up to order 2 in $\mathbf{p}$, and the coefficients will be Fourier series in $\mathbf{q}$.
We write $\tilde{h} = \tilde{h}_0 + \tilde{h}_1$:

- $\tilde{h}_1$ has the terms constant or linear in $\mathbf{p}$, except the constant in $\mathbf{q}$,
- $\tilde{h}_0$ is everything else.

We would like to eliminate $\tilde{h}_1$, but we won't be able to do so by solving a linear equation.
Instead we will solve a linear equation for a function g such that $\phi_g^* \tilde{h}$ is "better" than $\tilde{h}$, in some reasonable sense of "better."

Smell of the Proof

We expand $\phi_g^* \tilde{h}$ to first order in g:

$$\begin{aligned}\phi_g^* \tilde{h} &= \tilde{h} + \{g, \tilde{h}\} + o(|g|) \\ &= \tilde{h}_0 + \tilde{h}_1 + \{g, \tilde{h}_0\} + \{g, \tilde{h}_1\} + o(|g|).\end{aligned}$$

We want to eliminate the terms, other than the term constant in $\mathbf{q}$, that are not $O(|\mathbf{p}|)^2$.

Applying the standard Newton's method would necessitate solving the equation

$$\tilde{h}_1 + \{g, \tilde{h}_0\} + \{g, \tilde{h}_1\} \in o(|\mathbf{p}|).$$

But we will do something a bit different.

Smell of the Proof

We can assume that anything we want is small as long as the choice is justified by the resulting inequalities.
So suppose that $\{g, \tilde{h}_1\}$ is of order 2 since g and $\tilde{h}_1$ are both small.
Then the linear equation that we need to solve is

$$\tilde{h}_1 + \{g, \tilde{h}_0\} \in o(|\mathbf{p}|).$$

This is a *Diophantine differential equation*, and our ability to solve it will rely on the Diophantine vectors that we saw previously.

Diophantine Differential Equations

Let $g \in \mathcal{C}_\rho$. We are going to be solving linear equations of the form

$$\mathrm{D}f(\omega) = \sum_{i=1}^{n} \omega_i \frac{\partial f}{\partial q_i} = g,$$

with $f \in \mathcal{C}_{\rho'}$ for some $\rho' < \rho$.

Write the functions f and g as Fourier Series:

$$f(\mathbf{q}) = \sum_{\mathbf{k} \in \mathbb{Z}^n} f_{\mathbf{k}} e^{2\pi i\, \mathbf{k}\cdot\mathbf{q}}, \qquad g(\mathbf{q}) = \sum_{\mathbf{k} \in \mathbb{Z}^n} g_{\mathbf{k}} e^{2\pi i\, \mathbf{k}\cdot\mathbf{q}}.$$

The solution is given by

$$f_{\mathbf{k}} = \frac{1}{2\pi i\,(\mathbf{k}\cdot\omega)} g_{\mathbf{k}}.$$

We need g_0 to be zero. Then f_0 is arbitrary, and otherwise the series for f is unique.

Diophantine Differential Equations

The convergence properties of f depend on the Diophantine properties of ω.

We cannot divide by zero when computing the Fourier coefficients $f_{\mathbf{k}}$.

Even though we have convergence for ρ, we might not have boundedness. So we need some $\rho' < \rho$ where we have boundedness, but large enough so that the limit is nonempty.

Diophantine Differential Equations

We will use the following tool for choosing ρ'.

Proposition

If $g \in \mathcal{C}_\rho$ and $\omega \in \Omega_\gamma$, then for all $\delta \in (0, \rho)$ we have

$$\|f\|_{\rho-\delta} \leq \frac{\kappa_n}{\gamma \delta^{2n}} \|g\|_\rho ,$$

$$\|Df\|_{\rho-\delta} \leq \frac{\kappa_n}{\gamma \delta^{2n+1}} \|g\|_\rho ,$$

where κ_n is a constant that depends only on n.

Next we will prove this proposition.

Diophantine Differential Equations

For $\mathbf{y} \in \mathbb{R}^n$ with $|\mathbf{y}| \leq \rho$ (in particular $\mathbf{y} = \rho \frac{\mathbf{k}}{|\mathbf{k}|}$), the function $\mathbf{q} \mapsto g(\mathbf{q} - i\mathbf{y})$ is continuous and periodic in $\mathbf{q}$ of period 1.

This function can be written

$$g(\mathbf{q} - i\mathbf{y}) = \sum_{\mathbf{k} \in \mathbb{Z}^n} g_{\mathbf{k}} e^{2\pi i \mathbf{k} \cdot (\mathbf{q} - i\mathbf{y})} = \sum_{\mathbf{k} \in \mathbb{Z}^n} \left(g_{\mathbf{k}} e^{2\pi \mathbf{k} \cdot \mathbf{y}} \right) e^{2\pi i \mathbf{k} \cdot \mathbf{q}}.$$

By Parseval's theorem, we have

$$\|g\|_\rho^2 \geq \int_{\mathbb{T}^n} |g(\mathbf{q} - i\mathbf{y})|^2 |d^n\mathbf{q}| = \sum_{\mathbf{k} \in \mathbb{Z}^n} |g_{\mathbf{k}}|^2 e^{4\pi \mathbf{k} \cdot \mathbf{y}}.$$

Since the series has positive numbers, we have

$$\|g\|_\rho^2 \geq |g_{\mathbf{k}}|^2 e^{4\pi\rho|\mathbf{k}|}.$$

Diophantine Differential Equations

Using the Diophantine-ness of ω and the Fourier coefficients $f_{\mathbf{k}}$, we have

$$|f_{\mathbf{k}}| \leq \frac{1}{2\pi\gamma} \|g\|_{\rho} |\mathbf{k}|^{n} e^{-2\pi|\mathbf{k}|\rho}.$$

Next we split the proposition's proof into two parts: the f case and the Df case.

Diophantine Differential Equations

For $|\mathbf{q}| \leq \rho - \delta$, we can write

$$\left| f_{\mathbf{k}} e^{2\pi i(\mathbf{k}\cdot\mathbf{q})} \right| \leq \sum_{\mathbf{k}\in\mathbb{Z}^n} \frac{|\mathbf{k}|^n}{2\pi\gamma} \|g\|_\rho \, e^{-2\pi\rho|\mathbf{k}|(\rho-\delta)}$$

$$= \frac{\|g\|_\rho}{2\pi\gamma} \sum_{\mathbf{k}\in\mathbb{Z}^n} |\mathbf{k}|^n e^{-2\pi|\mathbf{k}|\delta}.$$

We then rewrite the sum on the right:

$$\sum_{\mathbf{k}\in\mathbb{Z}^n} |\mathbf{k}|^n e^{-2\pi|\mathbf{k}|\delta} = \frac{1}{(2\pi\delta)^{2n}} \left((2\pi\delta)^n \sum_{\mathbf{k}'\in 2\pi\delta\mathbb{Z}^n} |\mathbf{k}'|^n e^{-|\mathbf{k}'|} \right).$$

$$\sum_{\mathbf{k}\in\mathbb{Z}^n} |\mathbf{k}|^n e^{-2\pi|\mathbf{k}|\delta} = \frac{1}{(2\pi\delta)^{2n}} \left((2\pi\delta)^n \sum_{\mathbf{k}'\in 2\pi\delta\mathbb{Z}^n} |\mathbf{k}'|^n e^{-|\mathbf{k}'|} \right).$$

As $\delta \to 0$, the expression inside the parentheses approaches $\displaystyle\int_{\mathbb{R}^n} |\mathbf{x}|^n e^{-|\mathbf{x}|} |d^n\mathbf{x}|$.

Since the interval is convergent there exists κ_n' such that for $\delta \leq 1$, it holds that

$$\sum_{\mathbf{k}\in\mathbb{Z}^n} |\mathbf{k}|^n e^{-2\pi|\mathbf{k}|\delta} \leq \frac{\kappa_n'}{(2\pi\delta)^{2n}}.$$

This is the "f" part of the proof.

The "Df" part of the proof is similar.
Use

$$Df(\mathbf{q})(\mathbf{u}) = 2\pi i \sum (\mathbf{k}\cdot\mathbf{u}) f_{\mathbf{k}} e^{2\pi i(\mathbf{k}\cdot\mathbf{q})}.$$

We obtain

$$|Df(\mathbf{q})(\mathbf{u})| \leq \frac{|\mathbf{u}| \, \|g\|_{\rho} \, \kappa_n''}{\gamma(2\pi\gamma)^{n+1}}.$$

To complete the proof of the proposition, take

$$\kappa_n = \max\left\{ \frac{\kappa_n'}{(2\pi)^{2n}}, \frac{\kappa_n''}{(2\pi)^{2n+1}} \right\}.$$

The Equations to Solve

Let's return to $h_1 + \{g, h_0\} \in O(|\mathbf{p}|^2)$ (the Diophantine differential equation that we had before, but without the tildes and with $o(|\mathbf{p}|)$ replaced by $O(|\mathbf{p}|^2)$, which is equivalent since our functions are analytic).

The unknown is g, which we take to be degree 1 in $\mathbf{p}$:

$$g = \lambda \mathbf{q} + X(\mathbf{q}) + \sum_{i=1}^{n} Y_i(\mathbf{q}) p_i$$

Note that only the linear terms of g contribute to the linear terms of $\{g, h_0\}$.

The unknowns are: λ, X, Y_i.

We can expand X and Y_i as Fourier series, since they are functions of $\mathbf{q}$ only.

The Equations to Solve

Let $a \in \mathbb{R}$, $\omega \in \Omega_\gamma$, $R(\mathbf{q},\mathbf{p}) \in O(|\mathbf{p}|^3)$, and $\overline{A} = 0$ be the average of A on the torus $\mathbf{p} = 0$, we can write

$$h_0(\mathbf{q},\mathbf{p}) = a + \omega\mathbf{p} + \frac{1}{2}\mathbf{p}\cdot C(\mathbf{q})\mathbf{p} + R(\mathbf{q},\mathbf{p}),$$
$$h_1(\mathbf{q},\mathbf{p}) = A(\mathbf{q}) + B(\mathbf{q})\mathbf{p}.$$

We then have

$$\begin{aligned}(h_1 + \{g, h_0\})(\mathbf{q},\mathbf{p}) \\ &= \omega\cdot\lambda + A(\mathbf{q}) + DX(\mathbf{q})(\omega) \\ &\quad + \Big(B(\mathbf{q}) + \big(\lambda + DX(\mathbf{q})\big)C(\mathbf{q}) + \omega DY(\mathbf{q})\Big)\cdot\mathbf{p} + O(|\mathbf{p}|^2).\end{aligned}$$

The Equations to Solve

We need to solve the equations

$$DX(\mathbf{q})(\omega) = -A(\mathbf{q})$$
$$DY(\mathbf{q})(\omega) = -B(\mathbf{q}) - \big(\lambda + DX(\mathbf{q})\big)\, C(\mathbf{q})$$

for X and Y_i, $i = 1, \dots, n$.

These are Diophantine differential equations.

Since we have the hypothesis $\overline{A} = 0$, we can solve the first equation and find $X \in \mathcal{C}_{\rho'}$ for all $\rho' < \rho$.

We then substitute into the second equation.

The Equations to Solve

We need to solve the equations

$$DX(\mathbf{q})(\omega) = -A(\mathbf{q})$$
$$DY(\mathbf{q})(\omega) = -B(\mathbf{q}) - \big(\lambda + DX(\mathbf{q})\big)\, C(\mathbf{q})$$

for X and Y_i, $i = 1, \ldots, n$.

We can then determine λ since the averages of the right-hand sides in the n are all zero.

Then we can solve the second equation and find $Y_i \in \mathcal{C}_{\rho''}$ for all $\rho'' < \rho'$.

Iterating the process requires choosing ρ' and ρ'' "carefully."