The Kolmogorov Theorem

Travis Westura

May 3, 2015

Abstract

This paper gives a proof of the Kolmogorov Theorem on the conservation of invariant tori. We follow the approach given by Hubbard and Ilyashenko in . Their proof is similar to the one given by Bennettin, Galgani, Giorgilli, and Strelcyn in , which itself resembles Kolmogorov's original argument.

1 Introduction

But before moving on, let's take a look at our main goal:

Theorem 1.1 (The Kolmogorov Theorem)**.** *Let $\rho, \gamma > 0$ be given, and let $h(\boldsymbol{q}, \boldsymbol{p}) = h_0(\boldsymbol{p}) + h_1(\boldsymbol{q}, \boldsymbol{p})$ 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*

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

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 $o(|\boldsymbol{p}|^2)$, such that if $\|h_1\|_\rho \leq \epsilon$, there exists a symplectic mapping $\Phi : A_{\rho_*} \to A_\rho$ such that if we set $(\boldsymbol{q}, \boldsymbol{p}) = \Phi(\boldsymbol{Q}, \boldsymbol{P})$ and $H = h \circ \Phi$, we have

$$H(\boldsymbol{Q}, \boldsymbol{P}) = A + \omega \boldsymbol{P} + R(\boldsymbol{Q}, \boldsymbol{P}), \tag{1}$$

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

We will return to this statement after building up some intuition for it.

2 A Motivating Example

To start understanding the theorem, let's consider an example to which we would like to apply it. While staring up at the night sky, one might be driven to wonder, "Why do the planets orbit around the sun and not just crash into the sun or fly off by themselves in their own directions?" Kolmogorov's Theorem gives us a method to attempt to answer this question.

Let's start by discussing a simplified model of our solar system. In this model we will assume that planets have zero mass. This assumption is reasonable, as the masses of the planets are extremely small compared to that of the sun. Thus we may, at least for the moment, consider these masses to be negligible.

A system with n bodies, each with a mass m_i and a position $\boldsymbol{x}_i$ satisfies Netwon's second law: $\boldsymbol{F} = m\boldsymbol{a}$. For each i, we have

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

Here, $\mathrm{G} \approx 6.62 \cdot 10^{-11} m^3/(kg\, s^2)$ is the universal gravitational constant. On the left hand side we have each planet's mass and its acceleration. On the right hand side we have the force that acts on this mass. Note that the force is inversely proportional to the square of the distances between the forces. The top of the fraction contains the vector $\boldsymbol{x}_j - \boldsymbol{x}_i$ that has length $|\boldsymbol{x}_j - \boldsymbol{x}_i|$. To cancel out this length, we divide by $|\boldsymbol{x}_j - \boldsymbol{x}_i|^3$. This gives us the desired inverse proportionality to the square of the distance.

In our model of the solar system we take the sun to be given by the index 0 and the planets given by indices from 1 to 8 or 9, depending on the reader's opinions of Pluto. Now we will examine what happens as the masses of the planets, the m_j for $j = 1, \ldots, n$, tend to zero.

Rewriting the previous equations, we have

$$\ddot{\boldsymbol{x}}_0 = \mathrm{G} \sum_{j=1}^{n} m_j \frac{\boldsymbol{x}_j - \boldsymbol{x}_0}{|\boldsymbol{x}_j = \boldsymbol{x}_i|^3},$$
$$\ddot{\boldsymbol{x}}_i = \mathrm{G} m_0 \frac{\boldsymbol{x}_0 - \boldsymbol{x}_i}{|\boldsymbol{x}_0 - \boldsymbol{x}_i|^3} + \mathrm{G} \sum_{\substack{j=1,\ldots,n \\ j \neq i}} m_j \frac{\boldsymbol{x}_j - \boldsymbol{x}_i}{|\boldsymbol{x}_j - \boldsymbol{x}_i|^3}.$$

Now we keep the mass of the sun m_0 constant but let the masses of the planets go to zero. These equations then become

$$\ddot{\boldsymbol{x}}_0 = 0,$$
$$\ddot{\boldsymbol{x}}_j = \mathrm{G} \frac{\boldsymbol{x}_0 - \boldsymbol{x}_i}{|\boldsymbol{x}_0 - \boldsymbol{x}_i|^3}.$$

We can even simplify these equations a bit further. Since $\ddot{\boldsymbol{x}}_0 = 0$, the body with mass m_0 travels in a straight line with constant speed. Thus we can work in a heliocentric system of coordinates with the sun at the center of our solar system with position $\boldsymbol{x}_0 = 0$. We then have

$$\ddot{\boldsymbol{x}}_0 = 0,$$
$$\ddot{\boldsymbol{x}}_j = -\mathrm{G}\, m_0 \frac{\boldsymbol{x}_i}{|\boldsymbol{x}_i|^3}.$$

In particular we can note that this system is stable. Over time, it will never diverge far from its present state.

3 Hamiltonian Mechanics

4 Irrationality

In our statement of Kolmogorov's Theorem, we included the hypothesis that $\omega \in \Omega_\gamma$. We now define this notation and begin to explain its importance. The set Ω_γ consists of vectors that are "sufficiently irrational," a notion that we need to make more precise.

Let's first consider the definition of an irrational number. If a real number θ is irrational, then for all pairs if integers p and q, with q positive, we have the following

$$\left| \theta - \frac{p}{q} \right| \neq 0.$$

This equation tells us simply that there does not exist and rational number $\frac{p}{q}$ that equals our irrational number θ. Here, the "not equal to zero" part of the equation will be stressed, as an expression similar to the one on the left hand side will later appear as the denominator of a fraction (see section 7). As dividing by zero can be rather troublesome, we wish to avoid it. This condition of irrationality is the tool we use to do so: if θ is irrational, then the left side will not be zero, so we can divide by it without any problems.

Our condition of "sufficiently irrational" will mean that $\left|\theta - \dfrac{p}{q}\right|$ is "sufficiently nonzero," or since we are using an absolute value, "sufficiently big." However, as we learn in our introductory courses in real analysis, the rationals are dense in the reals, and every real number, specifically every irrational number θ, may be approximated arbitrarily closely by the rationals. More precisely, given any real $\epsilon > 0$, there exists a rational number $\dfrac{p}{q}$ such that $\left|\theta - \dfrac{p}{q}\right| < \epsilon$. Thus trying to coerce $\left|\theta - \dfrac{p}{q}\right|$ to be big is quite impossible.

Unsatisfied with our answer, let's instead consider a different question. Instead of wanting $\left|\theta - \dfrac{p}{q}\right|$ to be "big,", we ask that it is small *only if the denominator is big.* This is the beginning of the theory of Diophantine approximation.

The numbers that we seek will satisfy the following definition.

Definition 4.1 (Diophantine Number 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

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

From this definition we see that it is a stronger requirement for a number to be Diophantine of a smaller exponent. For all irrational numbers θ there exist arbitrarily large q and p prime to q such that

$$\left|\theta - \frac{p}{q}\right| < \frac{1}{\sqrt{5}q^2}.$$

We see that no number is Diophantine of any exponent smaller than 2. And the number that are Diophantine of exponent exactly 2 are precisely the numbers whose continued fractions have bounded entries. These numbers form a set of measure zero.

But what about exponents greater than 2, that is, of the form $2+\epsilon$ for $\epsilon > 0$? In the sense of Lebesgue measure, these numbers are quite abundant, as for any $\epsilon > 0$ they form a set of full measure.

Proposition 4.2 (Diophantine Numbers with Full Measure)**.** *For all $\epsilon > 0$, the set of Diophantine numbers of exponent $2+\epsilon$ is of full measure.*

Proof. We consider numbers in $\mathbb{R}/\mathbb{Z}$. Given any positive integer q, there are at most q elements of $\mathbb{Q}/\mathbb{Z}$ that, in reduced form, have denominator q. Hence for any constant γ, we consider the set

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

The length of this set is at most $\dfrac{2\gamma}{q^{1+\epsilon}}$. Summing over all q, we see that the set of numbers θ for with there exists q such that

$$\left|\theta - \frac{p}{q}\right| < \frac{\gamma}{2^{2+\epsilon}}$$

has length strictly less than

$$2\gamma \sum_{q=1}^{\infty} \frac{1}{q^{1+\epsilon}}.$$

Take the intersection over all these sets as $\gamma \to 0$ and note that this intersection has measure 0. But this set is the complement of the set of Diophantine numbers of exponent $2+\epsilon$. Hence the claim holds. QED

Having this definition for numbers, we want to extend "irrationality" to vectors. As an example, we can consider the solar system with the vector $\omega = (\omega_1, \ldots, \omega_n)$, where each ω_i represents the frequency of the ith planet's orbit. Our statement of Kolmogorov's theorem will require that such a vector be irrational according to the following definition.

Definition 4.3 (Diophantine Vector)**.** 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 $\boldsymbol{k} = (k_1, \ldots, k_n)$, 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 rewrite the condition in the definition as

$$\boldsymbol{k} \cdot \omega \geq \frac{\gamma}{|\boldsymbol{k}|^n}. \tag{2}$$

Again, we want to examine how common it is for such vectors to occur. We do not want just exceptional motions to be preserved, but rather we wish that most motions are preserved, and that we should not need to look hard to find such vectors. In the case of numbers, we have that satisfying answer that Diophantine numbers have full measure. We establish an analogous result for vectors, following a similar proof to the one we have just seen.

Proposition 4.4 (Diophantine Vectors are of Full Measure)**.** *The union*

$$\Omega = \bigcup_{\gamma > 0} \Omega_\gamma$$

is of full measure.

Proof. Consider the region $S_{\boldsymbol{k},\gamma}$, in which

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

is a region around the hyperplane orthogonal to $\boldsymbol{k}$ and with thickness $\frac{2\gamma}{|\boldsymbol{k}|^{n+1}}$. Denote the unit cube by Q. The part of $S_{\boldsymbol{k},\gamma}$ within Q has measure at most $\frac{M\gamma}{|\boldsymbol{k}|^{n+1}}$, where M denotes the constant giving the maximal $(n-1)$-dimensional measure of the intersection of Q with a hyperplane. Now consider the sum

$$\sum_{k \in \mathbb{Z}^n \setminus \{0\}} \frac{1}{|\boldsymbol{k}|^{n+1}}.$$

This sum is finite, so the volume of

$$\bigcup_{\boldsymbol{k} \in \mathbb{Z}^n \setminus \{0\}} S_{\boldsymbol{k},\gamma} \cap Q$$

is bounded by some constant times γ. As before, we now consider the intersection of these sets as $\gamma \to 0$:

$$\bigcap_{\gamma > 0} \bigcup_{\boldsymbol{k} \in \mathbb{Z}^n \setminus \{0\}} S_{\boldsymbol{k},\gamma} \cap Q.$$

This intersection has measure 0, and this set is the complement of our desired set Ω. Thus Ω has full measure.

Note that this proof is very similar to that of proposition 4.2, and reduces to the case $\epsilon = 1$ if we take $n = 2$. QED

We can now understand the condition $\omega \in \Omega_\gamma$ from our statement of Kolmogorov's Therorem. This requirement means that the vector ω must be "suitably irrational," and such vectors are rather "common" in the sense of Lebesgue.

5 KAM and the Solar System

6 Analytic Functions

As this is a paper for a complex analysis class, we begin this section by noting that the complex analysis is located here.

Consider the function $f(x) = x^3 - x + \frac{\sqrt{2}}{2}$.

7 Main Idea of the Proof

8 Diophantine Differential Equations

9 Solving Equations

10 Conclusion

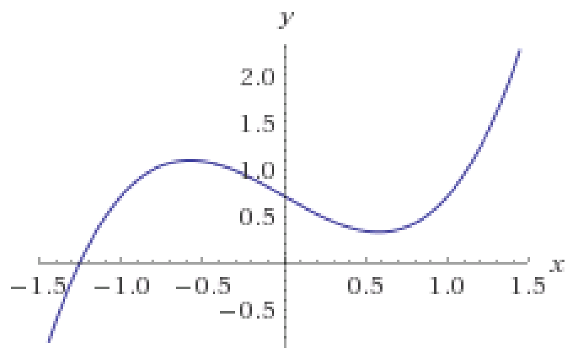

Figure 1: $f(x) = x^3 - x + \frac{\sqrt{2}}{2}$