DIOPHANTINE APPROXIMATION

Jan-Hendrik Evertse

e-mail evertse@math.leidenuniv.nl

Fall 2021

Chapter 1

Introduction

We give a brief overview of the contents of this course.

1.1 Geometry of numbers

One of the central theorems in the geometry of numbers is Minkowski's first convex body theorem. Before discussing this, we prove a predecessor, due to Dirichlet.

Theorem 1.1 (Dirichlet, 1842)**.** *Let $\alpha \in \mathbb{R}$. Then for every integer $Q \geqslant 2$ there are integers x, y, not both 0, such that $|x - \alpha y| \leqslant Q^{-1}$, $0 < y \leqslant Q$ and $\gcd(x, y) = 1$.*

Proof. The proof is based on *Dirichlet's box principle*: if n boxes contain altogether at least $n + 1$ objects, then one of the boxes must contain at least two objects.

The largest integer smaller or equal than a given real number α is denoted by $[\alpha]$. Partition the interval $[0, 1]$ into Q subintervals of length $1/Q$. Consider the $Q + 1$ numbers 1 and $\alpha - [\alpha], \ldots, Q\alpha - [Q\alpha]$. By the box principle, two among these numbers must lie in the same subinterval of length $1/Q$. So we either have $|(k\alpha - [k\alpha]) - (l\alpha - [l\alpha])| \leqslant 1/Q$ for two different integers $k, l \in \{1, \ldots, Q\}$, or $|(k\alpha - [k\alpha]) - 1| \leqslant 1/Q$ for some integer $k \in \{1, \ldots, Q\}$. In both cases, we find integers x, y with $0 < |y| \leqslant Q$ and $|x - \alpha y| \leqslant 1/Q$. By dividing x, y by their greatest common divisor, and changing sign if necessary, we get integers x, y as in Theorem 1.1. □

Corollary 1.2. *Let $\alpha \in \mathbb{R} \setminus \mathbb{Q}$. Then there are infinitely many pairs (x, y) such that*

$$|\alpha - \frac{x}{y}| \leqslant y^{-2}, \quad x, y \in \mathbb{Z},\ y \geqslant 0,\ \gcd(x, y) = 1. \tag{1.1}$$

Proof. By Theorem 1.1, for every integer $Q \geqslant 2$ there is a pair of integers (x_Q, y_Q) with $|x_Q - \alpha y_Q| \leqslant Q^{-1}$, $0 < y_Q \leqslant Q$, and $\gcd(x_Q, y_Q) = 1$. This pair clearly satisfies

$$|\alpha - \frac{x_Q}{y_Q}| \leqslant Q^{-1} y_Q^{-1} \leqslant y_Q^{-2}.$$

We are done if we have shown that if we let $Q \to \infty$, then (x_Q, y_Q) runs through infinitely many different pairs of integers. Assume this is false. Then there is a fixed pair (x_0, y_0) such that $(x_Q, y_Q) = (x_0, y_0)$ for arbitrarily large Q. But then,

$$|x_0 - \alpha y_0| \leqslant Q^{-1}$$

for arbitrarily large Q, and thus, $x_0 - \alpha y_0 = 0$, i.e., $\alpha = x_0/y_0$. This is against our assumption $\alpha \in \mathbb{R} \setminus \mathbb{Q}$. Corollary 1.2 follows. □

Remark. There is an alternative (and earlier) proof of Corollary 1.2 using the theory of continued fractions, see Exercise 1.6 below or for instance the classic *G.H. Hardy, E.M. Wright, An Introduction to the Theory of Numbers.*

Exercise 1.1. *Is Corollary 1.2 true or false for $\alpha \in \mathbb{Q}$?*

We now formulate Minkowski's first convex body theorem. A set $C \subset \mathbb{R}^n$ is called *convex* if for every $\mathbf{x}, \mathbf{y} \in C$, the line segment connecting them, that is $\{(1-t)\mathbf{x} + t\mathbf{y} : 0 \leqslant t \leqslant 1\}$, is also contained in C.

A *central symmetric convex body* in $\mathbb{R}^n$ is a closed, bounded, convex subset of $\mathbb{R}^n$ that contains $\mathbf{0}$ as an interior point and is symmetric about $\mathbf{0}$, i.e., for every $\mathbf{x} \in C$ we have $-\mathbf{x} \in C$.

Theorem 1.3 (Minkowski's first convex body theorem, 1896)**.** *Let $C \subset \mathbb{R}^n$ be a central symmetric convex body of volume (n-dimensional measure)* $\operatorname{vol} C \geqslant 2^n$. *Then C contains a point $\mathbf{x} \in \mathbb{Z}^n$ with $\mathbf{x} \neq \mathbf{0}$.*

We show that Theorem 1.3 implies Theorem 1.1. Let Q be an integer with $Q \geqslant 2$. We apply Minkowski's convex body theorem to

$$C_Q := \left\{(x, y) \in \mathbb{R}^2 : |x - \alpha y| \leqslant Q^{-1},\ |y| \leqslant Q\right\},$$

which is easily seen to be a central symmetric convex body, with area (2-dimensional measure) 2^2. We infer that there is a non-zero point $(x, y) \in C_Q \cap \mathbb{Z}^2$. If we divide (x, y) by their gcd, and replace (x, y) by $(-x, -y)$ if needed, we get a point $(x, y) \in C_Q \cap \mathbb{Z}^2$ with with $\gcd(x, y) = 1$ and $y \geqslant 0$. If $y = 0$ then $x = \pm 1$. The point $(\pm 1, 0)$ does not belong to C_Q since $Q > 1$. Hence $y > 0$. □

We state without proof the following more general result of Dirichlet, which can be proved either in a similar way as Theorem 1.1, or can be deduced from Minkowski's first convex body theorem.

Theorem 1.4 (Dirichlet, 1842)**.** *(i) Let $\alpha_1, \ldots, \alpha_n$ be $n \geqslant 1$ real numbers. Then for every real $Q > 1$ there is a tuple $(x_1, \ldots, x_n, y) \in \mathbb{Z}^{n+1}$ such that*

$$|x_i - \alpha_i y| \leqslant Q^{-1} \text{ for } i = 1, \ldots, n, \quad 0 < y \leqslant Q^n.$$

(ii) Assume in addition that $\alpha_1, \ldots, \alpha_n$ are not all rational numbers. Then there are infinitely many tuples $(x_1, \ldots, x_n, y) \in \mathbb{Z}^{n+1}$ such that

$$|\alpha_1 - \frac{x_1}{y}| \leqslant y^{-1-1/n}, \ldots, |\alpha_n - \frac{x_n}{y}| \leqslant y^{-1-1/n}, \quad y > 0, \ \gcd(x_1, \ldots, x_n, y) = 1.$$

In the next chapter on the Geometry of Numbers, we discuss a far-reaching generalization of Minkowski's Theorem, and some further applications.

1.2 Approximation of algebraic numbers by rational numbers

In general, for given $\alpha \in \mathbb{R}$ one may ask whether Corollary 1.2 remains true if we replace y^{-2} by a smaller function, say $y^{-2-\delta}$ with $\delta > 0$. That is, we may ask whether the inequality

$$|\alpha - x/y| \leqslant y^{-2-\delta} \text{ in } x, y \in \mathbb{Z} \text{ with } y > 0, \ \gcd(x, y) = 1 \tag{1.2}$$

has infinitely many solutions. The set of α for which this holds is very rare, since by a special case of a theorem of the Russian mathematician Khintchine (1927), for every $\delta > 0$, the set of $\alpha \in \mathbb{R}$ such that inequality (1.2) has infinitely many solutions has Lebesgue measure 0. But it is certainly possible to construct numbers α for which (1.2) has infinitely many solutions, as is shown by the exercise below.

Exercise 1.2. *Let a be an integer $\geqslant 3$ and put $\alpha := \sum_{n=0}^{\infty} 10^{-a^{2n}}$. Then the inequality*

$$|\alpha - x/y| \leqslant y^{-a}$$

has infinitely many solutions in integers x, y with $y > 0$, $\gcd(x, y) = 1$.

The number α constructed in this exercise seems very superficial. One may wonder, whether there are "reasonable" numbers α for which (1.2) has infinitely many solutions for some $\delta > 0$. A famous and difficult theorem by K.F. Roth (1955), states that this is not the case if α is algebraic. Recall that a number α is called *algebraic* if there is a non-zero polynomial $P \in \mathbb{Q}[X]$ with $P(\alpha) = 0$.

Theorem 1.5 (Roth, 1955)**.** *Let $\alpha \in \mathbb{R}$ be an algebraic number and let $\delta > 0$. Then the inequality*

$$|\alpha - x/y| \leqslant y^{-2-\delta} \quad \text{in } x, y \in \mathbb{Z} \text{ with } y > 0, \gcd(x, y) = 1$$

has only finitely many solutions.

The proof of this result is too long to be included in this course. We will give some applications of this result to Diophantine equations. Further, we will deduce a weaker version of Theorem 1.5.

Likewise, one may ask whether Theorem 1.4 is best possible. In case that $\alpha_1, \ldots, \alpha_n$ are all real algebraic numbers we have the following famous result of W.M. Schmidt. A set of numbers $\{\alpha_1, \ldots, \alpha_n\}$ in $\mathbb{C}$ is said to be *linearly independent over* $\mathbb{Q}$ if

$$\{(x_1, \ldots, x_n) \in \mathbb{Q}^n : x_1\alpha_1 + \cdots + x_n\alpha_n = 0\} = \{(0, \ldots, 0)\}.$$

Theorem 1.6 (Schmidt, 1971)**.** *Let $\alpha_1, \ldots, \alpha_n$ be algebraic numbers in $\mathbb{R}$ such that $\{1, \alpha_1, \ldots, \alpha_n\}$ is linearly independent over $\mathbb{Q}$. Further, let $\delta > 0$. Then there are only finitely many tuples $(x_1, \ldots, x_n, y) \in \mathbb{Z}^{n+1}$ such that*

$$y > 0, \ \gcd(x_1, \ldots, x_n, y) = 1,$$
$$|\alpha_1 - x_1/y| \leqslant y^{-1-(1/n)-\delta}, \ldots, |\alpha_n - x_n/y| \leqslant y^{-1-(1/n)-\delta}.$$

Theorem 1.1 is a consequence of a far more general, very central result in Diophantine approximation, the *Subspace Theorem*. This theorem is too difficult to be stated in this introduction, but we will discuss it later in this course. The Subspace Theorem has many consequences, in particular to Diophantine equations and inequalities, but also to other areas in number theory.

1.3 Transcendence

Recall that a number $\alpha \in \mathbb{C}$ is transcendental (over $\mathbb{Q}$) if it is not algebraic, i.e., there is no non-zero $P \in \mathbb{Q}[X]$ with $P(\alpha) = 0$. The following counting argument implies that transcendental numbers exist.

Theorem 1.7. *(i) $\mathbb{R}$ is uncountable.*
(ii) The set of algebraic numbers in $\mathbb{C}$ is countable.

Proof. (i) We use Cantor's diagonal argument. It suffices to prove that the open interval $]0,1[= \{x \in \mathbb{R} : 0 < x < 1\}$ is uncountable. We have to prove that $]0,1[\setminus T \neq \emptyset$ for any countable subset T of $]0,1[$. Take an arbitrary countable subset T, and represent its numbers by their decimal expansions. The assumption that T is countable means that its elements can be arranged in a sequence, say

$$\begin{aligned} &0.x_{11}x_{12}x_{13}\ldots \\ &0.x_{21}x_{22}x_{23}\ldots \\ &0.x_{31}x_{32}x_{33}\ldots \\ &\vdots \end{aligned}$$

There is $0.y_1y_2y_3\ldots \in]0,1[$ with $y_i \neq x_{ii}$ for all $i \geqslant 1$, and this number clearly does not belong to T.

(ii) A number $\alpha \in \mathbb{C}$ is algebraic if there is non-zero $F \in \mathbb{Z}[X]$ such that $F(\alpha) = 0$. For $F = a_0X^r + a_1X^{r-1} + \cdots + a_r \in \mathbb{Z}[X]$ we put $S(F) := \max(r, |a_0|, \ldots, |a_r|)$. Note that for given value of S there are only finitely many $F \in \mathbb{Z}[X]$ with $S(F) = S$ and each of these F has only finitely many zeros in $\mathbb{C}$. Now order the algebraic numbers in a sequence as follows: first take all algebraic numbers which are zeros of polynomials $F \in \mathbb{Z}[X]$ with $S(F) = 1$, then take the algebraic numbers not considered so far that are zeros of polynomials $F \in \mathbb{Z}[X]$ with $S(F) = 2$, and so on. In this way, we eventually obtain all algebraic numbers, and thus they can be arranged in a sequence. □

It is of course a much more interesting (and difficult) problem whether numbers "from nature" such as e and π are transcendental. In 1873, Hermite proved that e is transcendental, and in 1882 Lindemann did the same for π. In fact, Lindemann proved the following result, which covers both. Here we define $e^z := \sum_{n=0}^{\infty} z^n/n!$ for $z \in \mathbb{C}$.

Theorem 1.8. *Let $\alpha \in \mathbb{C}$ be a non-zero algebraic number. Then e^α is transcendental.*

To deduce from this that π is transcendental, assume that it is algebraic. Then πi would be algebraic, while $e^{\pi i} = -1$ is not transcendental, contradicting Lindemann's Theorem.

In our course we will discuss also other transcendence results.

To give some flavour, we finish with a proof that e is irrational. We start with a simple but useful *irrationality criterion*.

Lemma 1.9. *Let $\alpha \in \mathbb{R}$. Assume there is a sequence of pairs of integers (x_n, y_n) with $y_n > 0$ such that $x_n/y_n \neq \alpha$ and $|x_n - \alpha y_n| \to 0$ as $n \to \infty$. Then $\alpha \notin \mathbb{Q}$.*

Proof. Assume that $\alpha \in \mathbb{Q}$, that is, $\alpha = a/b$ with $a, b \in \mathbb{Z}$, $b > 0$. Then for any pair of integers x, y with $y > 0$, $x/y \neq \alpha$ we have

$$|x - \alpha y| = \frac{|bx - ay|}{b} \geqslant \frac{1}{b}$$

since the numerator is a non-zero integer. Hence a sequence of pairs (x_n, y_n) as in the statement of the lemma cannot exist. □

Theorem 1.10. $e \notin \mathbb{Q}$.

Proof. We use the identity $e = \sum_{k=0}^{\infty} 1/k!$. We apply Lemma 1.9 with $y_n = n!$ and $x_n = n! \sum_{k=0}^{n} \frac{1}{k!}$. Then

$$|x_n - e y_n| = n! \sum_{k=n+1}^{\infty} \frac{1}{k!}$$

hence

$$\begin{aligned} 0 < |x_n - e y_n| &= \frac{1}{n+1} + \frac{1}{(n+1)(n+2)} + \frac{1}{(n+1)(n+2)(n+3)} + \cdots \\ &< \frac{1}{n+1} + \frac{1}{(n+1)^2} + \frac{1}{(n+1)^3} + \cdots \\ &= \frac{1}{n} \to 0 \quad \text{as } n \to \infty. \end{aligned}$$

This proves that $e \notin \mathbb{Q}$. □

1.4 Exercises

Exercise 1.3. *Hurwitz (1891) improved Corollary 1.2 as follows: if α is any irrational, real number, then there are infinitely pairs of integers $(x, y) \in \mathbb{Z}^2$ such that*

$$|\alpha - x/y| \leqslant \frac{1}{\sqrt{5}} \cdot y^{-2}, \quad y > 0, \ \gcd(x, y) = 1.$$

The proof uses the theory of continued fractions (see exercise 1.6 below). You are asked to prove that Hurwitz' theorem becomes false if $\sqrt{5}$ is replaced by any constant $A > \sqrt{5}$. More precisely, you have to prove that if $\alpha := \frac{1}{2}(1 + \sqrt{5})$ then for every $A > \sqrt{5}$ there are only finitely many pairs of integers (x, y) such that $|\frac{x}{y} - \alpha| \leqslant 1/Ay^2$.
(i) Let $\alpha := \frac{1}{2}(1 + \sqrt{5})$, $\alpha' := \frac{1}{2}(1 - \sqrt{5})$. Prove that for any two integers x, y with $y > 0$,

$$1 \leqslant |x^2 - xy - y^2| = y^2 |\frac{x}{y} - \alpha| \cdot |\frac{x}{y} - \alpha'|.$$

(ii) Let $A > \sqrt{5}$ and let x, y be integers with $y > 0$ such that $|\frac{x}{y} - \alpha| \leqslant 1/Ay^2$. Estimate the right-hand side of the inequality in (i) from above, and deduce that y and x are bounded.

Exercise 1.4. *Prove that* $\sin 1 = \displaystyle\sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!}$ *is irrational.*

Exercise 1.5. *Complete the following irrationality proof for π (attributed to Cartwright, 1945).*
Assume that $\pi = \frac{a}{b}$ with $a, b \in \mathbb{Z}_{>0}$, $\gcd(a, b) = 1$.
(i) Define

$$I_n := \int_{-1}^{1} (1 - x^2)^n \cos(\tfrac{1}{2}\pi x) dx \quad \text{for } n = 0, 1, 2, \ldots .$$

Prove that

$$I_0 = \frac{4}{\pi}, \ \ I_1 = \frac{32}{\pi^3}, \quad I_n = \frac{8n}{\pi^2}\Big((2n-1)I_{n-1} - (2n-2)I_{n-2}\Big) \ \text{ for } n \geqslant 2.$$

(ii) Prove that $\frac{a^{2n+1}}{n!} \cdot I_n \in \mathbb{Z}$ for $n \geqslant 0$.
(iii) Prove that $0 < I_n \leqslant 2$ for all n and $0 < \frac{a^{2n+1}}{n!} \cdot I_n < 1$ for all sufficiently large n, and deduce a contradiction.

In the next exercise you are asked to prove some basic properties of continued fractions. The rational function $[x_0, \dots, x_n]$ in the variables $x_0, \dots, x_n$ is inductively defined by

$$[x_0] := x_0; \quad [x_0, x_1] := x_0 + \frac{1}{x_1}; \quad [x_0, \dots, x_n] := [x_0, \dots, x_{n-2}, x_{n-1} + \frac{1}{x_n}] \ (n \geqslant 2).$$

Thus,

$$[x_0, x_1, x_2] = x_0 + \frac{1}{x_1 + \frac{1}{x_2}}; \quad [x_0, x_1, x_2, x_3] = x_0 + \frac{1}{x_1 + \frac{1}{x_2 + \frac{1}{x_3}}}; \dots$$

For a real number α we define the sequences $\{a_n\}_{n \geqslant 0}$, $\{\alpha_n\}_{n \geqslant 1}$ inductively by

$$\begin{aligned} &\alpha = a_0 + \alpha_1, \ a_0 \in \mathbb{Z}, \ 0 \leqslant \alpha_1 < 1; \\ &\frac{1}{\alpha_n} = a_n + \alpha_{n+1}, \ a_n \in \mathbb{Z}, \ 0 \leqslant \alpha_{n+1} < 1 \quad (n = 1, \dots, n_0), \end{aligned}$$

where n_0 is the first index n with $\alpha_{n+1} = 0$. If no such index exists, we set $n_0 := \infty$. Further, we define the sequences $\{p_n\}_{n \geqslant -2}$, $\{q_n\}_{n \geqslant -2}$ by

$$\left.\begin{array}{llll} p_{-2} := 0, & p_{-1} := 1, & p_n := a_n p_{n-1} + p_{n-2} \\ q_{-2} := 1, & q_{-1} := 0, & q_n := a_n q_{n-1} + q_{n-2} \end{array}\right\} \quad \text{for } n = 0, \dots, n_0.$$

We call $a_0, a_1, \dots$ the *continued fractions*, and $\frac{p_0}{q_0}, \frac{p_1}{q_1}, \dots$ the *convergents* of α.

Exercise 1.6. *Prove that n_0 is finite if and only if $\alpha \in \mathbb{Q}$.*
Hint. *Supposing $\alpha \in \mathbb{Q}$, write $\alpha = \frac{r_0}{r_1}$ with $r_0, r_1 \in \mathbb{Z}$, $\gcd(r_0, r_1) = 1$ and $r_1 > 0$ and apply Euclid's algorithm to r_0 and r_1.*

Exercise 1.7. *Assume that $\alpha \in \mathbb{R} \setminus \mathbb{Q}$, i.e., $n_0 = \infty$, and prove the following:*
(i) $\alpha = [a_0, \dots, a_{n-1}, a_n + \alpha_{n+1}]$ for $n = 0, 1, 2, \dots$;
(ii) $\frac{p_n + x p_{n-1}}{q_n + x q_{n-1}} = [a_0, \dots, a_{n-1}, a_n + x]$ for $n = 0, 1, 2, \dots$ (identity of rational functions) and deduce that $\frac{p_n}{q_n} = [a_0, \dots, a_n]$ for $n = 0, 1, 2, \dots$;
(iii) $p_n q_{n-1} - p_{n-1} q_n = (-1)^{n-1}$ for $n = -1, 0, 1, 2, \dots$;
(iv) $(-1)^n (q_n \alpha - p_n) = \frac{1}{(a_{n+1} + \alpha_{n+2}) q_n + q_{n-1}}$ for $n = 0, 1, 2, \dots$;
(v) $\frac{1}{q_{n+2}} < (-1)^n (q_n \alpha - p_n) < \frac{1}{q_{n+1}}$ for $n = 0, 1, 2, \dots$.
Deduce that $(-1)^n (q_n \alpha - p_n)$ strictly decreases to 0 as $n \to \infty$.

Notice that the convergents p_n/q_n of α give infinitely many fractions x/y with $y > 0$, $\gcd(x, y) = 1$, and $|\alpha - x/y| < 1/y^2$. In fact (vi) implies

$$\frac{p_0}{q_0} < \frac{p_2}{q_2} < \frac{p_4}{q_4} \cdots < \alpha < \cdots < \frac{p_3}{q_3} < \frac{p_1}{q_1}.$$