Chapter 6

Approximation of algebraic numbers by rationals

Literature:
W.M. Schmidt, Diophantine approximation, Lecture Notes in Mathematics 785, Springer Verlag 1980, Chap.II, §§1,2, Chap. IV, §1
L.J. Mordell, Diophantine Equations, Pure and applied Mathematics series, vol. 30, Academic Press, 1969. reprint of the 1971 edition.

6.1 Liouville's Theorem and Roth's Theorem

We are interested in the problem how well a given real algebraic number can be approximated by rational numbers.

Recall that the height $H(\xi)$ of a rational number ξ is given by $H(\xi) := \max(|x|, |y|)$, where x, y are coprime integers such that $\xi = x/y$. In Homework exercise 10b, you were asked to prove the following inequality, which is a variation on a result of Liouville from 1844:

Theorem 6.1. *Let α be an algebraic number of degree $d \geqslant 1$. Then there is an effectively computable number $c(\alpha) > 0$ such that*

$$|\alpha - \xi| \geqslant c(\alpha) H(\xi)^{-d} \quad \textit{for every } \xi \in \mathbb{Q} \textit{ with } \xi \neq \alpha. \tag{6.1}$$

Here we may take $c(\alpha) = \text{den}(\alpha)^{-d} \cdot (1 + \overline{|\alpha|})^{1-d}$.

Let α be an algebraic number of degree $d \geqslant 2$. One of the central problems in Diophantine approximation is, to obtain improvements of (6.1) with in the right-hand side $H(\xi)^{-\kappa}$ with $\kappa < d$ instead of $H(\xi)^{-d}$. More precisely, the problem is, whether there exist $\kappa < d$ and a constant $c(\alpha, \kappa) > 0$ depending only on α, κ, such that

$$(6.2) \qquad |\xi - \alpha| \geqslant c(\alpha, \kappa) H(\xi)^{-\kappa} \text{ for every } \xi \in \mathbb{Q}.$$

Recall that by Dirichlet's Theorem, there exist infinitely many pairs of integers x, y such that $\left|\frac{x}{y} - \alpha\right| \leqslant |y|^{-2}$, $y \neq 0$. For such solutions we have $|x| \leqslant (|\alpha| + 1) \cdot |y|$. Hence, writing $\xi = \frac{x}{y}$ we infer that there is a constant $c_1(\alpha) > 0$ such that

$$|\xi - \alpha| \leqslant c_1(\alpha) H(\xi)^{-2} \text{ for infinitely many } \xi \in \mathbb{Q}.$$

This shows that there can not exist an inequality of the shape (6.2) with $\kappa < 2$. In particular, for rational or quadratic algebraic numbers α, Theorem 6.1 gives the best possible result in terms of the exponent on $H(\xi)$.

Now let α be a real algebraic number of degree $d \geqslant 3$. In 1909, the Norwegian mathematician A. Thue made an important breakthrough by showing that for every $\kappa > \frac{d}{2} + 1$ there exists a constant $c(\alpha, \kappa) > 0$ such that (6.2) holds. In 1921, C.L. Siegel proved the same for every $\kappa \geqslant 2\sqrt{d}$. In 1949, A.O. Gel'fond and independently Freeman Dyson (the famous physicist) improved this to $\kappa > \sqrt{2d}$. Finally, in 1955, K.F. Roth proved the following result, for which he was awarded the Fields medal.

Theorem 6.2 (Roth, 1955)**.** *Let α be a real algebraic number of degree $\geqslant 3$. Then for every $\kappa > 2$ there exists a constant $c(\alpha, \kappa) > 0$ such that*

$$(6.2) \qquad |\xi - \alpha| \geqslant c(\alpha, \kappa) H(\xi)^{-\kappa} \textit{ for every } \xi \in \mathbb{Q}.$$

As mentioned before, Roth's Theorem is valid also if α is a rational or quadratic number (with the proviso that $\xi \neq \alpha$ if $\alpha \in \mathbb{Q}$) but then it is weaker than (6.1). Further, Roth's Theorem holds true also for complex, non-real algebraic numbers α; then we have in fact $|\xi - \alpha| \geqslant |\mathrm{Im}\, \alpha|$ for $\xi \in \mathbb{Q}$, i.e., (6.2) holds even with $\kappa = 0$.

Exercise 6.1. *Let α be a real algebraic number of degree $\geqslant 3$. Prove that the following three assertions are equivalent:*

(i) for every $\kappa > 2$ there is a constant $c(\alpha, \kappa) > 0$ with (6.2)*;*

(ii) for every $\kappa > 2$, the inequality

$$|\xi - \alpha| \leqslant H(\xi)^{-\kappa} \quad \textit{in } \xi \in \mathbb{Q} \tag{6.3}$$

has only finitely many solutions;
(iii) for every $\kappa > 2$, $C > 0$, the inequality

$$|\xi - \alpha| \leqslant CH(\xi)^{-\kappa} \quad \textit{in } \xi \in \mathbb{Q} \tag{6.4}$$

has only finitely many solutions.

It should be noted that Theorem 6.1 is *effective*, i.e., the constant $c(\alpha)$ in (6.1) can be computed. In contrast, the results of Thue, Siegel, Gel'fond, Dyson and Roth mentioned above are *ineffective*, i.e., with their methods of proof one can prove only the *existence* of a constant $c(\alpha, \kappa) > 0$ as in (6.2), but one can not compute such a constant. Equivalently, the methods of proof of Thue ,..., Roth show that the inequalities (6.3), (6.4) have only finitely many solutions, but they do not provide a method to determine these solutions.

Thue used his result on the approximation of algebraic numbers stated above, to prove his famous theorem that if F is a binary form in $\mathbb{Z}[X, Y]$ such that $F(X, 1)$ has at least three distinct roots and m is a non-zero integer, then the equation

$$F(x, y) = m \quad \text{in } x, y \in \mathbb{Z}$$

has at most finitely many solutions.

We prove a more general result. A binary form $F(X, Y) \in \mathbb{Z}[X, Y]$ is called *square-free* if it is not divisible in $\mathbb{C}[X, Y]$ by $(\alpha X + \beta Y)^2$ for some $\alpha, \beta \in \mathbb{C}$, not both 0.

Theorem 6.3. *Let $F(X, Y) \in \mathbb{Z}[X, Y]$ be a square-free binary form of degree $d \geqslant 3$. Then for every $\kappa > 2$ there is a constant $c(F, \kappa) > 0$ such that for every pair of integers (x, y) with $F(x, y) \neq 0$ we have*

$$|F(x, y)| \geqslant c(F, \kappa) \max(|x|, |y|)^{d-\kappa}. \tag{6.5}$$

If F is a binary form of degree $d \leqslant 2$ the theorem holds true as well but then it is trivial since $|F(x, y)|$ is a positive integer, hence $\geqslant 1$.

Proof. We prove the inequality only for pairs of integers (x,y) with $|y| \geqslant |x|$. Then the inequality can be deduced for pairs (x,y) with $|x| > |y|$ by interchanging x, y and repeating the argument below.

Next, we restrict to the case that $|y| \geqslant |x|$ and F is not divisible by Y. If F is divisible by Y we have $F = Y \cdot F_1$ where $F_1 \in \mathbb{Z}[X,Y]$ is a square-free binary form of degree $d-1 \geqslant 2$ which is not divisible by Y. Then if the inequality holds for F_1 and with $d-1$ instead of d, it follows automatically for F.

So assume that F is a square-free binary form of degree $d \geqslant 2$ that is not divisible by Y. Then $F(X,Y) = a_0X^d + a_1X^{d-1}Y + \cdots + a_dY^d$ with $a_0 \neq 0$, and so,

$$F(X,Y) = a_0(X - \alpha_1 Y) \cdots (X - \alpha_d Y) \text{ with } \alpha_1, \ldots, \alpha_d \text{ distinct.}$$

Let (x,y) be a pair of integers with $F(x,y) \neq 0$ and $|y| \geqslant |x|$. Then $y \neq 0$. Let $\xi := x/y$. Notice that $|y| = \max(|x|,|y|) \geqslant H(\xi)$ (with equality if $\gcd(x,y) = 1$). Let i be the index with

$$|\xi - \alpha_i| = \min_{j=1,\ldots,d} |\xi - \alpha_j|.$$

Let $\kappa > 2$. Theorem 6.2 says that if α_i is real algebraic then there is a constant $c(\alpha_i, \kappa) > 0$ such that

$$|\xi - \alpha_i| \geqslant c(\alpha_i,\kappa) H(\xi)^{-\kappa} \geqslant c(\alpha_i,\kappa) \max(|x|,|y|)^{-\kappa};$$

as has been observed above this is true as well if α_i is not real. For $j \neq i$ we have

$$|\alpha_i - \alpha_j| \leqslant |\alpha_i - \xi| + |\xi - \alpha_j| \leqslant 2|\xi - \alpha_j|,$$

implying

$$|\xi - \alpha_j| \geqslant \frac{1}{2}|\alpha_i - \alpha_j|.$$

Hence

$$\begin{aligned} |F(x,y)| &= |y|^d \cdot |a_0| \prod_{j=1}^{d} |\xi - \alpha_j| = \max(|x|,|y|)^d \cdot |a_0| \prod_{j=1}^{d} |\xi - \alpha_j| \\ &\geqslant c(\alpha_i,\kappa)|a_0| \prod_{j \neq i} \left(\tfrac{1}{2}|\alpha_i - \alpha_j|\right) \cdot \max(|x|,|y|)^{d-\kappa}. \end{aligned}$$

□

We deduce Thue's Theorem.

Corollary 6.4. *Let $F(X,Y)$ be a binary form in $\mathbb{Z}[X,Y]$ such that $F(X,1)$ has at least three distinct roots. Further, let m be a non-zero integer. Then the equation*

$$F(x,y) = m \quad \text{in } x,y \in \mathbb{Z}$$

has at most finitely many solutions.

Proof. We first make a reduction to the case that $F(X,Y)$ is square-free, by showing that F is divisible in $\mathbb{Z}[X,Y]$ by a square-free binary form $F^* \in \mathbb{Z}[X,Y]$ of degree $\geqslant 3$.

We can factor the polynomial $F(X,1)$ as $cg_1(X)^{k_1}\cdots g_t(X)^{k_t}$ where c is a non-zero integer and $g_1(X),\dots,g_t(X)$ are irreducible polynomials in $\mathbb{Z}[X]$ none of which is a constant multiple of the others. Let $f^*(X) := g_1(X)\cdots g_t(X)$. Then $f^* \in \mathbb{Z}[X]$, and $\deg f^* =: d \geqslant 3$ since $F(X,1)$ has at least three zeros in $\mathbb{C}$. We have $F(X,1) = f^*(X)g(X)$ with $g \in \mathbb{Z}[X]$. Put $F^*(X,Y) = Y^d f(X/Y)$ and $G(X,Y) := Y^{\deg F - d} g(X/Y)$. Then $F = F^*G$ with $G \in \mathbb{Z}[X,Y]$. The polynomial f^* has degree $d \geqslant 3$ and d distinct zeros, and it divides $F(X,1)$ in $\mathbb{Z}[X]$. Hence F^* is square-free, F^* has degree $d \geqslant 3$ and F^* divides F in $\mathbb{Z}[X,Y]$.

Let x,y be integers with $F(x,y) = m$. Then $F^*(x,y)$ divides m. Take κ with $2 < \kappa < d$. Then by Theorem 6.3,

$$|m| \geqslant |F^*(x,y)| \geqslant c(F^*,\kappa)\max(|x|,|y|)^{d-\kappa},$$

implying that $|x|,|y|$ are bounded. $\square$

The *total degree* of a polynomial $G = \sum_{\mathbf{i}} a_{\mathbf{i}} X_1^{i_1}\cdots X_r^{i_r}$, notation $\operatorname{totdeg} G$, is the maximum of all quantities $i_1+\cdots+i_r$, taken over all tuples $\mathbf{i} = (i_1,\dots,i_r)$ with $a_{\mathbf{i}} \neq 0$. For instance, $3X_1^7X_2^5X_3^2 - 2X_1X_2^{12}X_3^2$ has total degree 15.

Exercise 6.2. *Let $F \in \mathbb{Z}[X,Y]$ be a square-free binary form of degree $d \geqslant 4$, and let $G \in \mathbb{Z}[X,Y]$ be a polynomial of total degree $\leqslant d-3$. Prove that there are only finitely many pairs $(x,y) \in \mathbb{Z}^2$ with $F(x,y) = G(x,y)$ and $F(x,y) \neq 0$.*

As mentioned before, the proof of Roth's Theorem is ineffective, and an effective proof of Roth's Theorem seems to be very far away. There are however effective improvements of Liouville's inequality, i.e., inequalities of the shape

$$|\xi - \alpha| \geqslant c(\alpha,\kappa)H(\xi)^{-\kappa} \quad \text{for } \xi \in \mathbb{Q}$$

where α is algebraic of degree $d \geqslant 3$ and $\kappa < d$ (but very close to d) and with some explicit expression for $c(\alpha, \kappa)$. We mention the following result of the Russian mathematician Fel'dman, obtained using lower bounds for linear forms in logarithms.

Theorem 6.5 (Fel'dman, 1971)**.** *Let α be a real algebraic number of degree $d \geqslant 3$. Then there exist effectively computable numbers $c_1(\alpha)$, $c_2(\alpha) > 0$ depending on α such that*

$$|\xi - \alpha| \geqslant c_1(\alpha) H(\xi)^{-d+c_2(\alpha)} \quad \textit{for } \xi \in \mathbb{Q}. \tag{6.6}$$

The proof is too complicated to be given here, but we can give a brief sketch. The hard core is the following effective result on Thue equations, given by Fel'dman. The proof is by making explicit the arguments in the previous chapter.

Lemma 6.6. *Let $F \in \mathbb{Z}[X, Y]$ be a binary form such that $F(X, 1)$ has at least three zeros in $\mathbb{C}$. Then there are effectively computable numbers A, B depending only on F, such that for every non-zero integer m and every solution $(x, y) \in \mathbb{Z}^2$ of $F(x, y) = m$ we have*

$$\max(|x|, |y|) \leqslant A|m|^B.$$

Proof of Theorem 6.5 (assuming Lemma 6.6). Let

$$f(X) = a_0 X^d + a_1 X^{d-1} + \cdots + a_d = a_0(X - \alpha^{(1)}) \cdots (X - \alpha^{(d)})$$

be the primitive minimal polynomial of α and $F(X, Y) := Y^d f(X/Y)$. Then $F(X, Y)$ is a binary form in $\mathbb{Z}[X, Y]$. Let $\xi = x/y$ with $x, y \in \mathbb{Z}$ coprime. Then $f(\xi) \neq 0$ and this implies $m := F(x, y) \neq 0$. By Lemma 6.6 we have

$$|F(x, y)| = |m| \geqslant \big(\max(|x|, |y|)/A\big)^{1/B} = \big(H(\xi)/A\big)^{1/B}, \tag{6.7}$$

where A, B are effectively computable positive numbers depending on F, hence α. It remains to estimate from below $|\xi - \alpha|$ in terms of $|F(x, y)|$ and $H(\xi)$.

Assume $\alpha^{(1)} = \alpha$. Notice that $F(x, y) = a_0 \prod_{i=1}^{d} (x - \alpha^{(i)} y)$. We estimate the factors as follows:

$$\begin{aligned}
&|x - \alpha^{(1)} y| = |\xi - \alpha| \cdot |y| \leqslant |\xi - \alpha| \cdot H(\xi), \\
&|x - \alpha^{(i)} y| \leqslant |x| + |\alpha^{(i)}| \cdot |y| \leqslant (1 + |\alpha^{(i)}|) H(\xi) \ (i = 2, \ldots, d).
\end{aligned}$$

Thus,

$$\begin{aligned}
|F(x,y)| &= |a_0|\prod_{i=1}^{d}\left|x-\alpha^{(i)}y\right| \\
&\leqslant |a_0|\cdot|\xi-\alpha|\cdot H(\xi)^d\prod_{i=2}^{d}(1+|\alpha^{(i)}|) \\
&= |\xi-\alpha|\cdot C(\alpha)H(\xi)^d,
\end{aligned}$$

say, where $C(\alpha)$ is effectively computable. Hence

$$|\xi-\alpha| \geqslant |F(x,y)|\cdot C(\alpha)^{-1}H(\xi)^{-d}.$$

Combined with (6.7) this gives

$$|\xi-\alpha| \geqslant A^{-1/B}C(\alpha)^{-1}H(\xi)^{-d+1/B} = c_1(\alpha)H(\xi)^{-d+c_2(\alpha)}$$

with $c_1(\alpha)=A^{-1/B}C(\alpha)^{-1}$, $c_2(\alpha)=1/B$. □

The quantities $c_1(\alpha), c_2(\alpha)$ are very small numbers for which one can find an explicit expression by going through the proof. For instance, Bugeaud proved in 1998, that (6.6) holds with

$$\begin{aligned}
c_1(\alpha) &= \exp\Big(-10^{27d}d^{16d}H^{d-1}\big(\log(edH)\big)^{d-1}\Big), \\
c_2(\alpha) &= \Big(10^{27d}d^{16d}H^{d-1}\big(\log(edH)\big)^{d-1}\Big)^{-1}
\end{aligned}$$

where d is the degree of α and $H=H(\alpha)$ its height.

One can obtain better results for certain special classes of algebraic numbers using other methods. M. Bennett obtained good effective improvements of Liouville's inequality for various numbers of the shape $\sqrt[m]{a}$ where m is a positive integer and a a positive rational number. For instance he showed that

$$\left|\xi-\sqrt[3]{2}\right| \geqslant \frac{1}{4}H(\xi)^{-2.45} \quad \text{for } \xi\in\mathbb{Q}. \tag{6.8}$$

Exercise 6.3. *Using* (6.8), *compute explicit constants* A, B *such that the following holds:*

for any solution $x,y\in\mathbb{Z}$ *of* $x^3-2y^3=m$ *we have* $\max(|x|,|y|)\leqslant A|m|^B$.

Hint. *Go through the proof of Theorem 6.3 and compute a constant* c *such that* $|x^3-2y^3|\geqslant c\max(|x|,|y|)^{3-2.45}$ *for all* $x,y\in\mathbb{Z}$. *Notice that we may have* $|x|>|y|$.

The techniques used by Thue,..., Roth cannot be used in general to solve Diophantine equations, but together with suitable refinements, they allow to give explicit upper bounds for the *number* of solutions of Diophantine equations. For instance we have:

Theorem 6.7 (Bombieri, Schmidt, 1986)**.** *Let $F(X,Y)$ be a binary form in $\mathbb{Z}[X,Y]$ such that $F(X,1)$ has precisely $d \geqslant 3$ distinct roots. Then the equation*

$$F(x,y) = 1 \quad \textit{in } x, y \in \mathbb{Z}$$

has at most $c \cdot d$ solutions where c is a positive constant not depending on d or F.

The importance of the result is that the bound is uniform, i.e. for *all* binary forms F as in the theorem, we get the upper bound cd. It is possible to compute c explicitly. Bombieri and Schmidt showed that for binary forms F that are irreducible over $\mathbb{Q}$ and for which d is sufficiently large, the constant c can be taken equal to 430. Probably the constant c can be improved, but the dependence on d is optimal. For instance, let $F(X,Y) = (X - a_1 Y)\cdots(X - a_d Y) + Y^d$, where $a_1, \ldots, a_d$ are distinct integers. Then the equation $F(x,y) = 1$ has the d solutions $(a_1, 1), \ldots, (a_d, 1)$.

M. Bennett proved the following remarkable result:

Theorem 6.8 (Bennett, 2002)**.** *Let d be an integer with $d \geqslant 3$ and let a, b be positive integers. Then the equation*

$$|ax^d - by^d| = 1$$

has at most one solution in positive integers x, y.

For instance, the equation $(a+1)x^d - ay^d = 1$ has $(1,1)$ as its only solution in positive integers. In his proof, Bennett uses various techniques (good lower bounds for linear forms in two logarithms, Diophantine approximation techniques based on so-called hypergeometric functions, and heavy computations).

We finish with some exercises related to the *abc*-conjecture, formulated by Masser and Oesterlé in 1985.
The *radical* $\mathrm{rad}(N)$ of a non-zero integer N is the product of the primes dividing N. For instance, $\mathrm{rad}(\pm 2^3 5^7 11^8) = 2 \cdot 5 \cdot 11$.

abc-conjecture. *For every $\varepsilon > 0$ there is a constant $C(\varepsilon) > 0$ such that for all positive integers a, b, c with $a + b = c$, $\gcd(a,b,c) = 1$ we have*

$$c \leqslant C(\varepsilon)\mathrm{rad}(abc)^{1+\varepsilon}.$$

The abc-conjecture has many striking consequences. As an example we deduce a weaker version of Fermat's Last Theorem. Let x, y, z be positive coprime integers and $n \geqslant 4$. Assume that $x^n + y^n = z^n$. Apply the abc-conjecture with $a = x^n, b = y^n, c = z^n$. Then $\mathrm{rad}(abc) \leqslant xyz \leqslant z^3$. Together with the abc-conjecture this implies $z^n \leqslant C(\varepsilon)(z^3)^{1+\varepsilon}$. Taking $\varepsilon < \frac{1}{4}$ it follows that z and n are bounded.

Exercise 6.4. *(i) Assuming the abc-conjecture, prove that the Fermat-Catalan equation*

$$x^m + y^n = z^k$$

has only finitely many solutions in positive integers x, y, z, m, n, k with $x > 1, y > 1, z > 1$, $\gcd(x, y, z) = 1$ and $\frac{1}{m} + \frac{1}{n} + \frac{1}{k} < 1$.
(ii) Does this assertion remain true if we drop the condition $\gcd(x, y, z) = 1$?
(iii) Determine the triples of positive integers (m, n, k) such that $\frac{1}{m} + \frac{1}{n} + \frac{1}{k} \geqslant 1$.

Remark. At the moment, 10 solutions of the Fermat-Catalan equation are known (see the Wikipedia page on the Fermat-Catalan equation), and of each of which at least one of m, n, k equals 2. Beal offered $\$10^6$ for a correct proof that the Fermat-Catalan equation has no solutions in integers x, y, z, m, n, k with $x, y, z > 1$ and $m, n, k > 2$.

Exercise 6.5. *Assuming the abc-conjecture, prove that for every $\varepsilon > 0$, the inequality*

$$|x^m - y^n| \leqslant \big(\max(x^m, y^n)\big)^{\frac{1}{m}+\frac{1}{n}-\varepsilon}$$

has only finitely many solutions in integers x, y, m, n with $x > 1, y > 1$, $\gcd(x, y) = 1$ and $m \geqslant 3, n \geqslant 2$.

Granville and Langevin proved independently that the abc-conjecture is equivalent to the following:

Granville-Langevin conjecture. *Let $F(X, Y) \in \mathbb{Z}[X, Y]$ be a square-free binary form of degree $d \geqslant 3$. Then for every $\kappa > 2$ there is a constant $C(F, \kappa) > 0$ such that*

$$\mathrm{rad}(F(x, y)) \geqslant C(F, \kappa) \max(|x|, |y|)^{d-\kappa} \quad \textit{for every } x, y \in \mathbb{Z} \textit{ with } \gcd(x, y) = 1,\ F(x, y) \neq 0.$$

Exercise 6.6. *(i) Prove that the Granville-Langevin conjecture implies the abc-conjecture (the converse is also true but this is much harder to prove).*

(ii) Prove that the Granville-Langevin conjecture implies Roth's Theorem.

(iii) An integer $n \neq 0$ *is called* powerful *if every prime in the prime factorization of* n *occurs with exponent at least* 2*. In other words,* n *is powerful if it can be expressed as* $\pm a^2 b^3$ *for certain positive integers* a, b *not both equal to* 1*.*
Let $F(X,Y) \in \mathbb{Z}[X,Y]$ *be a square-free binary form of degree at least* 5*. Assuming the Granville-Langevin conjecture, prove that there are only finitely many pairs of integers* x, y *with* $\gcd(x,y) = 1$ *such that* $F(x,y)$ *is powerful.*

(iv) Assuming the Granville-Langevin conjecture, prove the following. Let $f(X) \in \mathbb{Z}[X]$ *be a polynomial of degree* $d \geqslant 2$ *with* d *distinct zeros in* $\mathbb{C}$*. Then for every* $\varepsilon > 0$ *there is a constant* $C'(f,\varepsilon) > 0$ *such that*

$$\mathrm{rad}(f(x)) \geqslant C'(f,\varepsilon)|x|^{d-1-\varepsilon} \text{ for all } x \in \mathbb{Z} \text{ with } f(x) \neq 0.$$

Hint. *Construct from* f *a binary form* F *of degree* $d+1$*.*

(v) Deduce the following conjecture of Schinzel: if f *is any square-free polynomial in* $\mathbb{Z}[X]$ *of degree* $\geqslant 3$*, then there are only finitely many integers* x *such that* $f(x)$ *is powerful.*

In 2012, the Japanese mathematician Shinichi Mochizuki published four papers, together consisting of about 500 pages, easily traceable on internet, in which he developed a new theory based on totally new mathematics, "Interuniversal Teichmüller theory," and as a consequence of this, in the last of the four papers, deduced the abc-conjecture. At present, some people are still working through these papers and trying to understand them, but up to now there has not been an official confirmation whether they contain a correct proof of the abc-conjecture or not.

It should be mentioned here that the argument with which the Granville-Langevin conjecture is deduced from the abc-conjecture, is constructive. That is, any effective version of the abc-conjecture, with the constant $C(\varepsilon)$ effectively computable in terms of ε, would imply an effective version of the Granville-Langevin conjecture, with $C(F,\kappa)$ effectively computable in terms of F and κ, and thus by Exercise 6.6 (ii), an effective version of Roth's theorem.

6.2 Thue's approximation theorem

We intend to prove the following result of Thue:

Theorem 6.9. *Let α be a real algebraic number of degree $d \geqslant 3$ and $\kappa > \frac{d}{2} + 1$. Then the inequality*

$$|\xi - \alpha| \leqslant H(\xi)^{-\kappa} \quad \text{in } \xi \in \mathbb{Q} \tag{6.9}$$

has only finitely many solutions.

Our basic tool will be Siegel's Lemma, which we recall here. We consider systems of linear equations

$$\begin{array}{ccccccc} a_{11}x_1 & + & \cdots & + & a_{1N}x_N & = 0 \\ \vdots & & & & & \vdots \\ a_{M1}x_1 & + & \cdots & + & a_{MN}x_N & = 0 \end{array} \tag{6.10}$$

with coefficients a_{ij} from the ring of integers O_K of a number field K.

Proposition 6.10. *Let K be an algebraic number field of degree d, let M, N be integers with $N > dM > 0$, let A be a real $\geqslant 1$, and suppose that*

$$a_{ij} \in O_K, \quad \overline{|a_{ij}|} \leqslant A \quad \text{for } i = 1, \ldots, M,\ j = 1, \ldots, N.$$

Then (6.10) *has a non-zero solution $\mathbf{x} = (x_1, \ldots, x_N) \in \mathbb{Z}^N$ such that*

$$\max_{1 \leqslant i \leqslant N} |x_i| \leqslant (3NA)^{dM/(N-dM)}. \tag{6.11}$$

Proof. This is Corollary 4.24 from Chapter 4. □

We introduce some notation. The norm of a polynomial $P = \sum_{i=0}^{D} p_i X^i \in \mathbb{C}[X]$ is given by

$$\|P\| := \sum_{i=0}^{D} |p_i|.$$

It is not difficult to check that

$$|P(\alpha)| \leqslant \|P\| \cdot \max(1, |\alpha|)^{\deg P} \quad \text{for } P \in \mathbb{C}[X],\ \alpha \in \mathbb{C}, \tag{6.12}$$

$$\|P + Q\| \leqslant \|P\| + \|Q\|, \quad \|PQ\| \leqslant \|P\| \cdot \|Q\| \quad \text{for } P, Q \in \mathbb{C}[X]. \tag{6.13}$$

From these properties it can be deduced that if $P \in \mathbb{C}[X]$, $\alpha \in \mathbb{C}$, then for the polynomial $\widetilde{P}(X) := P(X+\alpha)$ we have

$$\|\widetilde{P}\| \leqslant \|P\| \cdot (1+|\alpha|)^{\deg P}. \tag{6.14}$$

Exercise 6.7. *Prove* (6.12)–(6.14).

The *k-th divided derivative* of a polynomial $P \in \mathbb{C}[X]$ is defined by $P^{((k))} := P^{(k)}/k!$. Thus, if $P = \sum_{i=0}^{D} p_i X^i$, then

$$P^{((k))} = \sum_{i=0}^{D} \binom{i}{k} p_i X^{i-k} \quad \text{with } \binom{a}{b} := 0 \text{ if } b > a.$$

Notice that if $P \in \mathbb{Z}[X]$ then also $P^{((k))} \in \mathbb{Z}[X]$. Further, since each binomial coefficient $\binom{i}{k}$ can be estimated from above by $2^i \leqslant 2^{\deg P}$, we have

$$\|P^{((k))}\| \leqslant 2^{\deg P}\|P\|. \tag{6.15}$$

Lastly, we have the product rule

$$(PQ)^{((k))} = \sum_{j=0}^{k} P^{((k-j))}Q^{((j))} \text{ for } P, Q \in \mathbb{C}[X]. \tag{6.16}$$

The advantage of using divided derivatives over derivatives is that their coefficients are much smaller, while the divided derivatives of a polynomial with integer coefficients are still integral.

A brief outline of the proof of Theorem 6.9. We give a brief, informal outline of the proof, ignoring technicalities. More details and explanation are given later. We follow the usual procedure to assume that (6.9) has infinitely many solutions, and to construct a non-zero integer of absolute value < 1.

The first step of the proof is to take, for any positive integer r, non-zero polynomials $P_r, Q_r \in \mathbb{Z}[X]$ of degree as small as possible such that $P_r - \alpha Q_r$ is divisible by $(X-\alpha)^r$. Using Siegel's Lemma, one can prove the existence of such P_r, Q_r of degree at most $m := [(\frac{1}{2}d + \varepsilon)r]$ for any $\varepsilon > 0$, where $[x]$ denotes the largest integer $\leqslant x$.

To see this, view the coefficients of P_r, Q_r as a system of $2m+2$ unknowns. The condition $P_r - \alpha Q_r$ divisible by $(X-\alpha)^r$ is equivalent to α being a zero of the k-th (divided) derivative of $P_r - \alpha Q_r$, for $k = 0, \dots, r-1$, i.e.,

$$P_r^{((k))}(\alpha) - \alpha Q_r^{((k))}(\alpha) = 0 \ \text{ for } k = 0, \dots, r-1.$$

By expanding this, we get a system of r linear equations with coefficients in $K := \mathbb{Q}(\alpha)$ in the $2m+2$ unknown coefficients of P_r, Q_r. Now $[K : \mathbb{Q}] = d$ and $2m+2 > dr$, hence by Proposition 6.10, this system has a non-trivial solution in integers.

In the second step, we take two solutions of (6.9), say $\xi_1 = x_1/y_1$, $\xi_2 = x_2/y_2$ with $x_i, y_i \in \mathbb{Z}$, $\gcd(x_i, y_i) = 1$, $y_i > 0$ for $i = 1, 2$, and consider the number

$$A_r := y_1^{[(\frac{1}{2}+\varepsilon)dr]} y_2 (P_r(\xi_1) - \xi_2 Q_r(\xi_1)).$$

This is an integer since $P_r(\xi_1)$, $Q_r(\xi_1)$ are rational numbers with denominators dividing $y_1^m = y_1^{[(\frac{1}{2}+\varepsilon)dr]}$. We want to show that we can choose solutions ξ_1, ξ_2 and r such that $A_r \neq 0$ and $|A_r| < 1$, thus obtaining a contradiction.

To prove $|A_r| < 1$, we write

$$P_r - \alpha Q_r = V_r \cdot (X-\alpha)^r \text{ with } V_r \in \mathbb{C}[X]$$

and obtain

$$A_r = y_1^{[(\frac{1}{2}+\varepsilon)dr]} y_2 \Big(V_r(\xi_1)(\xi_1 - \alpha)^r - (\xi_2 - \alpha) Q_r(\xi_1) \Big).$$

To keep our discussion informal, we ignore ε and the terms $|V_r(\xi_1)|$, $|Q_r(\xi_1)|$ and are sloppy with constants. We choose $r = \log H(\xi_2)/\log H(\xi_1)$ (being again sloppy and assuming that the latter is an integer). Then $y_1 \leqslant H(\xi_1)$, $y_2 \leqslant H(\xi_1)^r$, $|\xi_1 - \alpha| \leqslant H(\xi_1)^{-\kappa}$, $|\xi_2 - \alpha| \leqslant H(\xi_1)^{-\kappa r}$. This leads to the 'estimate'

$$|A_r| \ '\!\leqslant' \ H(\xi_1)^{(dr/2)+r-\kappa r} = H(\xi_1)^{r((d/2)+1-\kappa)}.$$

Since the exponent on $H(\xi_1)$ is negative we get $|A_r| \ '\!<' \ 1$. Of course, we do have to take into account ε and estimates for $|V_r(\xi_1)|$, $|Q_r(\xi_1)|$. Further, the quantity $\log H(\xi_2)/\log H(\xi_1)$ need not be an integer and thus, in general we can not choose r equal to this quantity but only close to it. But with some modifications in the above argument, we can deduce in a correct manner that $|A_r| < 1$, provided we assume that $H(\xi_1)$ and $\log H(\xi_2)/\log H(\xi_1)$ are sufficiently large. This is allowed thanks to our assumption that (6.9) has infinitely many solutions.

What remains is to show that $A_r \neq 0$. Unfortunately, it is not all clear how to do this. In fact, r depends on ξ_1 and ξ_2 and we may have the bad luck that with our particular choice of r, the quantity A_r just becomes 0. Instead, we prove that for any two distinct solutions ξ_1, ξ_2 of (6.9) and any positive integer r, there is a not too large value $k_0 = k_0(r, \varepsilon)$ depending on r and ε but independent of ξ_1, ξ_2, such that $P_r^{((k))}(\xi_1) \neq \xi_2 Q_r^{((k))}(\xi_1)$ for some $k \leqslant k_0$. Then

$$A_{r,k} := y_1^{[(\frac{1}{2}+\varepsilon)dr]} y_2 \big(P_r^{((k))}(\xi_1) - \xi_2 Q_r^{((k))}(\xi_1)\big)$$

is a non-zero integer. Similarly as above we prove that if $H(\xi_1)$ and $\log H(\xi_2)/\log H(\xi_1)$ are sufficiently large, then $|A_{r,k}| < 1$ for all $k \leqslant k_0$ and obtain a contradiction. $\square$

The precise proof of Theorem 6.9. We need a parameter ε with $0 < \varepsilon < \frac{1}{2}$. Later, ε will be chosen depending on d, κ. Further, r will be a positive integer, to be chosen later.

We start with the construction of the polynomials P_r, Q_r.

Lemma 6.11. *For every positive integer r there exist polynomials $P_r, Q_r \in \mathbb{Z}[X]$ of degree at most $[(\frac{1}{2}+\varepsilon)dr]$, not both equal to 0, with the following properties:*

(6.17) $$P_r - \alpha Q_r \text{ is divisible by } (X-\alpha)^r,$$

(6.18) $$\|P_r\| \leqslant C_1^r,\ \ \|Q_r\| \leqslant C_1^r,$$

where C_1 is an effectively computable number, depending only on α, ε.

Proof. Let $K = \mathbb{Q}(\alpha)$. Put $m := [(\frac{1}{2}+\varepsilon)dr]$. Write

$$P_r = \sum_{i=0}^{m} p_i X^i, \quad Q_r = \sum_{i=0}^{m} q_i X^i,$$

where p_i, q_i are unknowns, taken from the integers. The condition to be satisfied is

$$P_r^{((k))}(\alpha) - \alpha Q_r^{((k))}(\alpha) = 0 \quad (k = 0, \ldots, r-1).$$

Let b be a denominator of α, i.e., $b \in \mathbb{Z}_{>0}$, $b\alpha \in O_K$. By expanding the above expressions and multiplying with b^m we obtain

$$\sum_{i=0}^{m} \binom{i}{k} b^m \alpha^i p_i - \sum_{i=0}^{m} \binom{i}{k} b^m \alpha^{i+1} q_i = 0 \quad (k = 0, \ldots, r-1),$$

which is a system of r linear equations in $2m+2 > (1+2\varepsilon)dr$ unknowns with coefficients in O_K. Thus, the number of unknowns is larger than $[K:\mathbb{Q}]$ times the number of equations, and the condition of Proposition 6.10 is satisfied. As a consequence, the above system has a non-trivial solution $(p_0,\dots,p_m,q_0,\dots,q_m) \in \mathbb{Z}^{2m+2}$ such that

$$\max(\max_i |p_i|, \max_i |q_i|) \leqslant (3(2m+2)A)^{\frac{dr}{2m+2-dr}} \leqslant (3(2+2\varepsilon)drA)^{1/2\varepsilon},$$

where (with $0 \leqslant i \leqslant m$, $0 \leqslant k \leqslant r$),

$$\begin{aligned} A &= \max\Big(\max_{i,k} \tbinom{i}{k} b^m \overline{|\alpha|}^i,\ \max_{k,i} \tbinom{i}{k} b^m \overline{|\alpha|}^{i+1} \Big) \\ &\leqslant 2^m b^m \max(1, \overline{|\alpha|})^{m+1} \leqslant \big(2b \max(1, \overline{|\alpha|})\big)^{(2+2\varepsilon)dr}. \end{aligned}$$

Then using $3(2+2\varepsilon)dr \leqslant 3^{(2+2\varepsilon)dr}$, we see that Lemma 6.11 holds with
$C_1 = \big(6b \cdot \max(1, \overline{|\alpha|})\big)^{d(1+\varepsilon^{-1})}$. □

We now take two solutions ξ_1, ξ_2 of (6.9) and show that for every r there is a not too large k such that $P_r^{((k))}(\xi_1) - \xi_2 Q_r^{((k))}(\xi_1) \neq 0$. We start with a simple lemma.

Lemma 6.12. *Let $F \in \mathbb{Q}[X]$, β an algebraic number such that $(X-\beta)^m$ divides F, and $f \in \mathbb{Q}[X]$ the minimal polynomial of β. Then f^m divides F.*

Proof. Recall that if $g \in \mathbb{Q}[X]$ is a polynomial with $g(\beta) = 0$ then f divides g. Further, β is not a multiple root of f. So f divides F, and by induction, f divides F/f^i for $i = 0, \dots, m-1$. Hence f^m divides F. □

Lemma 6.13. *Let ξ_1, ξ_2 be two rational numbers, and r a positive integer. Then there is $k \leqslant d(2\varepsilon r + 1)$ such that $P_r^{((k))}(\xi_1) \neq \xi_2 Q_r^{((k))}(\xi_1)$.*

Proof. The proof rests upon an analysis of the polynomial

$$F := P_r Q_r' - P_r' Q_r.$$

We first show that F is not identically 0. Assume the contrary. At least one of P_r, Q_r, say Q_r, is not identically 0. Then $(P_r/Q_r)' = 0$ hence P_r/Q_r is identically equal to some constant $c \in \mathbb{Q}$. But then, $Q_r = (c-\alpha)^{-1}\big(P_r - \alpha Q_r\big)$ is divisible by $(X-\alpha)^r$ and so, in view of Lemma 6.12, by f^r, where f is the minimal polynomial of α. But this is impossible, since by our assumption $\varepsilon < \frac{1}{2}$ we have $r \deg f = rd > (\frac{1}{2} + \varepsilon)dr \geqslant \deg Q_r$.

We now prove our lemma. Assume that there exists an integer $t \geqslant 1$ such that

$$P_r^{((k))}(\xi_1) = \xi_2 Q_r^{((k))}(\xi_1) \quad \text{for } k = 0, \ldots, t$$

(if not, we are done). By eliminating ξ_2 we obtain

$$P_r^{((k))}(\xi_1)Q_r^{((l))}(\xi_1) - P_r^{((l))}(\xi_1)Q_r^{((k))}(\xi_1) = 0 \text{ for } k, l \leqslant t.$$

For each $k \geqslant 0$, $F^{((k))}$ is a linear combination of $P_r^{((l))}Q_r^{((m))} - P_r^{((m))}Q_r^{((l))}$, $0 \leqslant l, m \leqslant k+1$. Hence $F^{((k))}(\xi_1) = 0$ for $k \leqslant t-1$, and therefore, F is divisible by $(X - \xi_1)^t$.

By construction, $P_r - \alpha Q_r$ is divisible by $(X - \alpha)^r$, hence $P_r' - \alpha Q_r'$ is divisible by $(X - \alpha)^{r-1}$. So, using

$$F = P_r(Q_r' - \alpha P_r') - P_r'(Q_r - \alpha P_r)$$

we see that F is divisible by $(X - \alpha)^{r-1}$. But $F \in \mathbb{Q}[X]$ hence by Lemma 6.12 it is divisible by f^{r-1}. So F is in fact divisible by $(X - \xi_1)^t f^{r-1}$. Since

$$\deg F \leqslant \max(\deg P_r + \deg Q_r', \deg P_r' + \deg Q_r) \leqslant (1 + 2\varepsilon)dr - 1, \qquad \deg f = d,$$

it follows that

$$t \leqslant (1 + 2\varepsilon)dr - 1 - d(r - 1) = d(2\varepsilon r + 1) - 1.$$

This proves our lemma. □

Take two solutions ξ_1, ξ_2 of (6.9). Write $\xi_i = x_i/y_i$ with $x_i, y_i \in \mathbb{Z}$, $\gcd(x_i, y_i) = 1$ and $y_i > 0$ for $i = 1, 2$. For integers $r > 0$, $k \geqslant 0$ consider the number

$$A_{r,k} := y_1^{[(\frac{1}{2}+\varepsilon)dr]} y_2 \Big(P_r^{((k))}(\xi_1) - \xi_2 Q_r^{((k))}(\xi_1) \Big).$$

This is clearly an integer, and by Lemma 6.13 there is $k < d(2\varepsilon r + 1)$ such that $A_{r,k} \neq 0$. We proceed to prove that $|A_{r,k}| < 1$ for appropriate ξ_1, ξ_2 and r.

We note that the polynomial $P_r^{((k))} - \alpha Q_r^{((k))}$ is divisible by $(X - \alpha)^{r-k}$, that is,

$$P_r^{((k))} - \alpha Q_r^{((k))} = V_r \cdot (X - \alpha)^{r-k} \text{ with } V_r \in \mathbb{C}[X].$$

This gives for $A_{r,k}$ the expression

$$A_{r,k} = y_1^{[(\frac{1}{2}+\varepsilon)dr]} y_2 \Big(V_r(\xi_1)(\xi_1-\alpha)^{r-k} - (\xi_2-\alpha) Q_r^{\{k\}}(\xi_1) \Big). \tag{6.19}$$

We first estimate $V_r(\xi_1)$ and $Q_r(\xi_1)$.

Lemma 6.14. *There is an effectively computable number C_2 depending only on α, κ and ε such that*

$$|V_r(\xi_1)| \leqslant C_2^r, \quad |Q_r(\xi_1)| \leqslant C_2^r.$$

Proof. We use (6.12)–(6.15). Define $\widetilde{P}(X) := P_r^{((k))}(X+\alpha)$, $\widetilde{Q}(X) := Q_r^{((k))}(X+\alpha)$, $\widetilde{V}(X) := V_r(X+\alpha)$ and $\widetilde{\xi} := \xi_1 - \alpha$. Then

$$\begin{aligned} \|\widetilde{P}\| &\leqslant \|P^{((k))}\|(1+|\alpha|)^{(\frac{1}{2}+\varepsilon)dr} \leqslant \big(2(1+|\alpha|)\big)^{(\frac{1}{2}+\varepsilon)dr}\|P\| \\ &\leqslant \big(2(1+|\alpha|)\big)^{(\frac{1}{2}+\varepsilon)dr} C_1^r \end{aligned}$$

and likewise $\|\widetilde{Q}\| \leqslant \big(2(1+|\alpha|)\big)^{(\frac{1}{2}+\varepsilon)dr} C_1^r$. Since $\widetilde{P} - \alpha\widetilde{Q} = X^{r-k}\widetilde{V}$, the polynomial $\widetilde{V}$ has the same coefficients as $\widetilde{P} - \alpha\widetilde{Q}$, and thus,

$$\|\widetilde{V}\| \leqslant \|\widetilde{P}\| + |\alpha| \cdot \|\widetilde{Q}\| \leqslant (1+|\alpha|)\big(2(1+|\alpha|)\big)^{(\frac{1}{2}+\varepsilon)dr} C_1^r.$$

Together with $|\widetilde{\xi}| = |\xi_1 - \alpha| \leqslant 1$, this leads to

$$|Q_r^{((k))}(\xi_1)| = |\widetilde{Q}(\widetilde{\xi})| \leqslant \|\widetilde{Q}\| \leqslant C_2^r, \quad |V(\xi_1)| = |\widetilde{V}(\widetilde{\xi})| \leqslant \|\widetilde{V}\| \leqslant C_2^r,$$

with $C_2 := 2^{(\frac{1}{2}+\varepsilon)d}(1+|\alpha|)^{1+(\frac{1}{2}+\varepsilon)d} C_1$. □

Proof of Theorem 6.9. Let ξ_1, ξ_2 be two solutions of (6.9) with $H(\xi_2) > e \cdot H(\xi_1)$ and define the integer r by

$$r \leqslant \frac{\log H(\xi_2)}{\log H(\xi_1)} < r+1.$$

Then $r \geqslant 1$. Next choose $\varepsilon > 0$ such that

$$\kappa = \big(\tfrac{1}{2} + (2\kappa+2)\varepsilon\big)d + 1.$$

This certainly implies our earlier condition $0 < \varepsilon < \frac{1}{2}$. Let k be an integer with $0 \leqslant k < d(2\varepsilon r + 1)$. Then with our choice of r we have

$$y_1 \leqslant H(\xi_1), \quad y_2 \leqslant H(\xi_1)^{r+1}, \quad |\xi_1 - \alpha| \leqslant H(\xi_1)^{-\kappa}, \quad |\xi_2 - \alpha| \leqslant H(\xi_1)^{-\kappa r}.$$

Using the expression (6.19) for $A_{r,k}$ together with these inequalities and Lemma 6.14 we deduce

$$\begin{aligned} |A_{r,k}| &\leqslant |y_1^{[(\frac{1}{2}+\varepsilon)dr]} y_2| \cdot \Big(|V_r(\xi_1)| \cdot |\xi_1 - \alpha|^{r-k} + |Q_r^{\{k\}}(\xi_1)| \cdot |\xi_2 - \alpha| \Big) \\ &\leqslant C_2^r \cdot \Big(|y_1^{(\frac{1}{2}+\varepsilon)dr} y_2| \cdot |\xi_1 - \alpha|^{r-k} + |y_1^{(\frac{1}{2}+\varepsilon)dr} y_2| \cdot |\xi_2 - \alpha| \Big) \\ &\leqslant C_2^r (H(\xi_1)^u + H(\xi_1)^v) \end{aligned}$$

where

$$u = (\tfrac{1}{2} + \varepsilon)dr + r + 1 - \kappa(r - k), \quad v = (\tfrac{1}{2} + \varepsilon)dr + r + 1 - \kappa r.$$

With our choice $\kappa = (\frac{1}{2} + (2\kappa + 2)\varepsilon)d + 1$ for ε and the estimate $k < d(2\varepsilon r + 1)$ we deduce for u and v the upper bounds

$$\begin{aligned} u &\leqslant r((\tfrac{1}{2} + \varepsilon)d + 1 - \kappa + 2\kappa\varepsilon d) + 1 + \kappa d = -\varepsilon dr + 1 + \kappa d, \\ v &\leqslant r((\tfrac{1}{2} + \varepsilon)d + 1 - \kappa) + 1 \leqslant -\varepsilon dr + 1. \end{aligned}$$

So altogether,

$$|A_{r,k}| \leqslant 2C_2^r \cdot H(\xi_1)^{-\varepsilon dr + 1 + \kappa d}.$$

The right-hand side becomes smaller than 1 if $H(\xi_1)$ and r are sufficiently large, and for the latter we have to assume that $\log H(\xi_2)/\log H(\xi_1)$ is sufficiently large. More precisely, let us choose solutions ξ_1, ξ_2 of (6.9) such that

$$H(\xi_1) \geqslant (2C_2)^{2/\varepsilon d}, \quad \frac{\log H(\xi_2)}{\log H(\xi_1)} \geqslant 1 + \frac{2(1 + \kappa d)}{d\varepsilon}; \tag{6.20}$$

this is possible since we assumed that (6.9) has infinitely many solutions. With this choice we have

$$r \geqslant \frac{2(1 + \kappa d)}{d\varepsilon}$$

and so $-\varepsilon dr + 1 + \kappa d \leqslant -\frac{1}{2}\varepsilon dr$. Then thanks to our assumption for $H(\xi_1)$ we obtain

$$|A_{r,k}| \leqslant 2C_2^r H(\xi_1)^{-\varepsilon dr/2} < 1,$$

as required. On the other hand, by Lemma 6.13, there is $k \leqslant d(2\varepsilon r + 1)$ such that $A_{r,k}$ is a non-zero integer. This gives the contradiction we want. So (6.9) cannot have infinitely many solutions. □

Remark. To obtain a contradiction, we did not need the assumption that (6.9) has infinitely many solutions, but merely that there are solutions ξ_1, ξ_2 of (6.9) that satisfy (6.20). In other words, solutions ξ_1, ξ_2 of (6.9) satisfying (6.20) cannot exist. The constant C_2 is effectively computable. So in fact we can prove the following sharpening of Theorem 6.9:

Theorem 6.15. *Let α be an algebraic number of degree d and $\kappa > \frac{1}{2}d + 1$. There are effectively computable positive numbers C, λ depending on α, κ, such that if ξ_1 is a solution of*

$$|\xi - \alpha| \leqslant H(\xi)^{-\kappa} \quad \text{in } \xi \in \mathbb{Q} \tag{6.9}$$

with $H(\xi_1) \geqslant C$, then for any other solution ξ of (6.9) *we have $H(\xi) \leqslant H(\xi_1)^{\lambda}$.*

It should be noted that Theorem 6.15 would give an effective proof of Thue's Theorem in case we were extremely lucky and knew a solution ξ_1 of (6.9) with $H(\xi_1) \geqslant C$. However, to find such a solution seems quite hopeless, since the constant C is very large. It is very likely that such a solution ξ_1 does not even exist. However, there are variations on Thue's method, which work only for special algebraic numbers α of the shape $\sqrt[d]{a}$ with $a \in \mathbb{Q}$, where the constant C is much smaller and where a solution ξ_1 of (6.9) with $H(\xi_1) \geqslant C$ is known. For such α one can derive very strong effective approximation results, for instance Bennett's estimate (6.8) mentioned in Section 6.1.

On the other hand Theorem 6.15 can be used to estimate the *number* of solutions of (6.9). This is worked out in the exercise below.

Exercise 6.8. *(i) Let ξ_1, ξ_2 be distinct rational numbers. Prove that*

$$|\xi_1 - \xi_2| \geqslant \big(H(\xi_1)H(\xi_2)\big)^{-1}.$$

(ii) Let α be a real number, and $\kappa > 2$, and consider the inequality

$$|\xi - \alpha| \leqslant H(\xi)^{-\kappa} \text{ in } \xi \in \mathbb{Q} \text{ with } \xi > \alpha. \tag{6.21}$$

Prove that if ξ_1, ξ_2 are two distinct solutions of (6.21) *with $H(\xi_2) \geqslant H(\xi_1)$, then*

$$H(\xi_2) \geqslant H(\xi_1)^{\kappa - 1}.$$

(So there are large gaps between the solutions of (6.21)*; we call such an inequality a* gap principle.*)*

Hint. *Estimate from above* $|\xi_1 - \xi_2|$.

(iii) Let $A \geqslant 2$, $c > 1$. *Prove that the number of solutions* ξ *of* (6.21) *with* $A \leqslant H(\xi) < A^c$ *is bounded above by* $1 + \frac{\log c}{\log(\kappa - 1)}$.

(iv) Let α *be a real algebraic number of degree* $d \geqslant 3$ *and* $\kappa > \frac{d}{2} + 1$. *Compute an explicit upper bound for the number of solutions of* (6.21) *in terms of the constants* C *and* λ *from Theorem 6.15.*