SÉMINAIRE DE THÉORIE DES NOMBRES DE BORDEAUX

W. J. ELLISON

Recipes for Solving Diophantine Problems by Baker's Method

Séminaire de théorie des nombres de Bordeaux (1970-1971), exp. n° 9, p. 1-10

<http://www.numdam.org/item?id=STNB_1970-1971____A9_0>

Séminaire de Théorie des Nombres
Année 1970-1971 - exposé n° 11

10 Décembre 1970

RECIPES FOR SOLVING DIOPHANTINE PROBLEMS BY BAKER'S METHOD

by

W. J. ELLISON

-:-:-:-

§. I. - In recent years Alan Baker has proved a remarkable series of theorems about the maximum magnitude of the integral solutions to a wide class of diophantine problems. A typical example is the following theorem.

THEOREM. If $k \neq 0$ is an integer, then all integral solutions (x, y) of the diophantine equation $y^2 = x^3 + k$ satisfy the inequality

$$\max \{|x|, |y|\} \leq \exp (10^{10} |k|^{10^4}) . \qquad (1)$$

Thus, in principle, a constructive algorithm for find all the integral solutions of the above equation for a given value of k would be : "Try all possible values of (x, y) which satisfy (1) ".

Needless to say, this could never be done in practice for any given value of k .

Here is another example of one of Baker's theorems.

THEOREM. If d is a positive integer such that:

(1) The field $Q(\sqrt{-d})$ has class number 2

and (2) The field has even discriminant, $d \leq 10^{500}$.

Again, "in principle", one can use this result to find all complex quadratic number fields with class number 2 and even discriminant. by testing each value of $d \leq 10^{500}$. But of course this simple procedure cannot be carried out in practice.

Whenever Baker's method is applied to a diophantine problem one always reduces the problem to that of finding all the integral solutions $\{b_1, \ldots, b_n\}$ of an inequality of the type

$$0 < |b_1 \log \alpha_1 + \ldots + b_n \log \alpha_n| < e^{-\delta H} , \qquad (2)$$

where $H = \max\{|b_1|, \ldots, |b_n|\}$ and $\{\alpha_1, \ldots, \alpha_n\}$ are given algebraic numbers and $\delta > 0$ is a given real number.

There will be a simple relationship between the integral solution of (3) and the integral solutions of the original diophantine problem.

Baker [ ] has given an explicit upper bound for the possible magnitude of H in (2). His theorem is a follows.

THEOREM. For $n \geq 2$ let $\alpha_1, \ldots, \alpha_n$ be non-zero algebraic numbers whose heights and degrees do not exceed A and d respectively, where $A \geq 4$, $d \geq 4$. Furthermore, suppose that $0 < \delta \leq 1$. If rational integers $\{b_1, \ldots, b_n\}$ exist such that

$$0 < |b_1 \log \alpha_1 + \ldots + b_n \log \alpha_n| < e^{-\delta H} , \qquad (3)$$

where $H \geq \max\{|b_1|, \ldots, |b_n|\}$, then

$$0 \leq H \leq (4^{n^2} \delta^{-1} d^{2n} \log A)^{(2n+1)^2} . \qquad (4)$$

So, in solving a specific diophantine problem by Baker's method, the real problem is to find all the integral solutions of an inequality of type (3) where H is in the range (4). This is a non-trivial problem. Even in the

simplest case when $n = 2$ and α_1, α_2 are quadratic irrationals the upper bound for H given by (4) is about 10^{300} (which is greater than the cube of the number of atomic particles in the observable Universe !!).

To-day I wish to show you that for any specific diophantine problem to which Baker's method is applicable one can always find all the integral solutions of the corresponding inequality of type (4) . Of course one does need an electronic computer, but the computation time needed is quite small ; it is only a matter of minutes rather than hours or days, ...

§. II. - From now on I am going to consider the following problem. "Given real numbers $\theta_1, \ldots, \theta_n$ and $C > 1$, $K > 1$ find all integers $\{b_1, \ldots, b_n\}$ which satisfy

$$|b_1\theta_1 + \ldots + b_n\theta_n| < K^{-H} ,$$

where $H = \max\{|b_1|, \ldots, |b_n|\}$ and H lies in the range

$$0 \leq H \leq C . \text{ "}$$

To fix ideas about the numerical quantities involved in a specific problem one can expect K to always be about 2 , for $n = 2$, C is about 10^{300} ; for $n = 3$, C is about 10^{700} ; for $n = 4$, C is about 10^{2000} .

The simplest case is $n = 2$, $\theta_1 = \theta$, $\theta_2 = 1$. (This is the case which arose when Baker proved that all complex quadratic fields $Q(\sqrt{-d})$ with class number 2 and even discriminant satisfy $d \leq 10^{500}$).

The inequality which we are to consider is

$$|\theta b_1 + b_2| < K^{-H} \quad \text{with} \quad 0 \leq H \leq C \quad (\sim 10^{300}) . \quad (1)$$

We recall a classic lemma of Legendre about continued fractions.

LEMMA. (1) <u>If</u> θ <u>is a real number and</u> p/q <u>is a rational approximation to</u> θ <u>which satisfies</u>

$$|\theta - \frac{p}{q}| < \frac{1}{2q^2} ,$$

<u>then</u> p/q <u>occurs as a convergent in the continued fraction expansion of</u> θ .

(2) If p_n/q_n is a convergent in the continued fraction expansion of θ and a_n is the corresponding partial quotient, then the following inequalites hold

$$\frac{1}{(a_{n+1}+2)q_n^2} < \left|\theta - \frac{p_n}{q_n}\right| < \frac{1}{a_{n+1}q_n^2} .$$

We now write inequality (1) as

$$\left|\theta + \frac{b_2}{b_1}\right| < \frac{K^{-H}}{|b_1|} \le \frac{K^{-|b_1|}}{|b_1|} .$$

If $|b_1|$ is sufficiently large, say $|b_1| \ge C_o$, we have

$$2K^{|b_1|} > |b_1|$$

and so

$$\frac{K^{-|b_1|}}{|b_1|} < \frac{1}{2|b_1|^2} .$$

A possible value for C_o is $\frac{\log C}{\log K}$, but in a numerical case one can usually take C_o to be smaller than this quantity.

Thus, for values of $\{b_1, b_2\}$ which satisfy (1) and $C_o \le |b_1| \le C$ we know, by Legendre's lemma, that b_2/b_1 must occur as a convergent in the continued fraction expansion of θ and that the corresponding partial quotients must satisfy

$$\frac{1}{(a_{n+1}+2)b_1^2} < \frac{K^{-|b_1|}}{|b_1|}$$

or

$$a_{n+1} > \frac{K^{|b_1|}}{|b_1|} - 2 \ge \frac{K^{C_o}}{C_o} - 2 ,$$

since the function $\frac{K^x}{x}$ is increasing for $x > (\log K)^{-1}$ and we are assuming $|b_1| \ge C_o = \frac{\log C}{\log K}$.

(In a numerical case the lower bound for a_{n+1} will be quite large, about 10^{10}).

So in order to check the range $C_o \le |b_1| \le C$ for possible solutions of (1) we merely evaluate the continued fraction expansion of θ until the denominators of the partial convergents exceed C. Then we check to see if any large partial quotients have occurred. If all the partial quotients are less than

$(\frac{K}{C_o}^{C_o} - 2)$ then there are no solutions of the inequality (1) with $|b_1|$ in the range $C_o \le |b_1| \le C$. However, if some partial quotient does exceed $(\frac{K}{C_o}^{C_o} - 2)$ then we must test the corresponding partial convergent p_n/q_n to see if the inequality is indeed satisfied.

For values of $|b_1|$ in the range $0 \le |b_1| \le C_o$ one can test the inequality directly, since C_o is only about 300.

It was in this way that Baker's method was used to find all complex quadratic fields with class number 2 and even discriminant. The computation took about 90 seconds.

One can extend this reduction technique to the n variable case, but the amount of work which the computer is forced to do increases very quickly. A possible reduction method when $n = 3$ is as follows.

We wish to find all integral solutions to the following inequality

$$0 < |b_1\theta_1 + b_2\theta_2 + b_3| < K^{-H} \tag{1}$$

with

$$0 \le H \le C \quad (\sim 10^{700}) \ . \tag{2}$$

Let $(p_1/q\ ,\ p_2/q)$ be rational approximations to (θ_1, θ_2) which satisfy

$$|\theta_i - \frac{p_i}{q}| = |\omega_i| < \frac{1}{2qC} \quad \text{for} \quad 1 \le i \le 2 \ .$$

(By Dirichlet's theorem such approximations do exist if we allow $q > 4C^2$ and there are several practical computational methods of finding them, though none of them are very efficient).

We can write (1) as

$$0 < |b_1\frac{p_1}{q} + b_1\omega_1 + b_2\frac{p_2}{q} + b_2\omega_2 + b_3| < K^{-H}$$

or

$$0 < |b_1p_1 + b_2p_2 + b_3q + b_1q\omega_1 + b_2q\omega_2| < qK^{-H}.$$

If C_o is such that $qK^{-H} < 1$ (i.e. $C_o > \frac{\log q}{\log K}$ will do), then if $\{b_1, b_2, b_3\}$ satisfy (1) with, H in the range $C_o \le H \le C$ we must have

$$b_1p_1 + b_2p_2 + b_3q = 0 \quad \text{and} \quad |b_1\omega_1 + b_2\omega_2| < K^{-H} \ ,$$

since $b_1 p_1 + b_2 p_2 + b_3 q$ is an integer and

$$|b_1 q\omega_1 + b_2 q\omega_2| \le |b_1 q\omega_1| + |b_2 q\omega_2| < \frac{Cq}{2Cq} + \frac{Cq}{2Cq} = 1 .$$

We can now find all the integral solutions to $|b_1\omega_1 + b_2\omega_2| < K^{-H}$ by using the previons reduction process and hence we find all the integral solutions to (1) which satisfy $C_o \le H \le C$. For values of $H \le C_o$ it is very easy to do a direct search.

The above reduction process does work, but it would be desirable to find a more efficient method. So far, the above technique has not been used to solve any specific diophantine problems.

When one uses Baker's method to find all the integral solutions of an equation of the form $f(x, y) = k$, where $f(x, y)$ is a binary form of degree n with integral coefficients, then the diophantine inequality which one obtains is inhomogenous. It is of the form

$$0 < |b_1 \log \alpha_1 + \ldots + b_{n-1} \log \alpha_{n-1} - \log \alpha_n| < e^{-\delta H} ,$$

where

$$0 \le H \le C .$$

The first non-trivial case is $n = 3$ and the following lemma can be used to reduce the upper bound for H from C to about $\log C$.

LEMMA. Suppose that θ, β, $K > 1$ are given real numbers and that C, $B > 6$ are given positive integers. Let p, q be integers such that

$$1 \le q \le BC \text{ and } |q\theta - p| \le \frac{2}{BC} .$$

Then if $\|q\beta\| \ge \frac{3}{B}$ there are no solutions in integers b_1, b_2 of the inequality

$$0 < |b_1\theta + b_2\beta| < K^{-H}$$

with H satisfying

$$\frac{\log(B^2 C)}{\log K} \le H \le C .$$

Proof. - Let $\theta - \frac{p}{q} = \omega$, where $|\omega| \le \frac{2}{qBC}$. We have

$$0 < |b_1 q\theta + b_2 q - q\beta| < q K^{-H} \le BCK^{-H}$$

or $$0 < |b_1 p + b_2 q + b_1 q\omega - q\beta| < BC \cdot K^{-H} .$$

Because $\|q\beta\| \geq \frac{3}{B}$ and since $|b_1 q\omega| \leq \frac{C \cdot q}{BCq} = \frac{2}{B}$ it follows that

$$\|b_1 q\omega - q\beta\| \geq \frac{1}{B} .$$

This implies that

$$\frac{1}{B} \leq |b_1 p + b_2 q + b_1 q\omega - q\beta| < BCK^{-H}$$

i.e. : $$1 \leq B^2 C \cdot K^{-H}$$

or $$H \leq \frac{\log(B^2 C)}{\log K} .$$

In order to apply this lemma in a numerical case, one is given θ, β and one can compute p, q with the required property by the usual continued fraction algorithm. Once one has found q it is then a very simple matter to test whether or not $\|\beta q\| \geq \frac{3}{B}$. If the test is satisfied we have a very much smaller upper bound for H. We can either apply the lemma again or do a direct test of all the possible values of (b_1, b_2). In numerical cases one can usually reduce the upper bound for H, which is initially about 10^{700}, to about 20 by two or three applications of the lemma.

However, if $\|q\beta\| \leq \frac{3}{B}$ then all is not lost, for we do know that if

$$H > \frac{\log(B^2 C)}{\log K} ,$$

then $BCK^{-H} < 1$ and that $|b_1 q\omega| < \frac{1}{2}$ and $q\beta = c + \varepsilon$, where c is an integer and $|\varepsilon| \leq \frac{3}{B}$. And we can conclude that $b_1 p + b_2 q - c = 0$, since $b_1 p + b_2 q - c$ is an integer with absolute value less than 1.

So in this exceptional case we have

$$b_1 p \equiv -c \pmod{q} .$$

This congruence has two solutions (mod q) with $0 < |b_1| \leq q$ which we can find quite easily. But we know that $|b_1| \leq C$ and if $q > C$ it means that these two solutions of the congruence are the only possible values of b_1 which satisfy

$$\frac{\log(B^2 C)}{\log K} \leq |b_1| \leq C$$

and

$$|b_1\theta + b_2 - \beta| < K^{-H} .$$

We must test these two values of b_1 to see if, in fact, they do satisfy the two inequalities.

Perhaps I ought to mention that the condition $\|q\beta\| \geq \frac{3}{B}$ has never failed in any of the numerical calculations. We have had $C \sim 10^{700}$ and $B \sim 10^{20}$. The computed values of $\|q\beta\|$ have never been less than 0.01, usually they were 0.2 ; 0.1 , etc.

As before it is possible to extend this reduction technique to the n variable case. I shall only give the case $n = 4$.

LEMMA. Let $C, B > 10^4$ be given integer with $B > C^2$ and let $\theta_1, \theta_2, \beta, K > 1$ be given real numbers. If p_1, p_2, q are integers such that

$$1 \leq q \leq BC \; ; \; |\theta_i - \frac{p_i}{q}| < \frac{2}{(BC)^{3/2}} \qquad \text{for } i = 1, 2 ,$$

then if

$$\|q\beta\| > \frac{5}{B^{1/4}} \tag{1}$$

there are no solutions of the inequality

$$|b_1\theta_1 + b_2\theta_2 + b_3 - \beta| < K^{-H} \tag{2}$$

with H satisfying

$$\frac{\log (B^{5/4} C)}{\log K} \leq H \leq C . \tag{3}$$

Proof. - Let $\theta_i - \frac{p_i}{q} = \omega_i$, where $|\omega_i| \leq \frac{2}{(BC)^{3/2}}$ and substitute into (2). We obtain

$$|b_1 p_1 + b_2 p_2 + b_3 q + b_1 q\omega_1 + b_2 q\omega_2 - q\beta| < qK^{-H} \leq BCK^{-H} .$$

Now

$$|q b_i \omega_i| \leq \frac{BC^2 2}{(BC)^{3/2}} = 2(\frac{C}{B})^{1/2} < \frac{2}{B^{1/4}} \quad , \text{ since } B > M^2$$

and by hypothesis $\|q\beta\| > \frac{5}{B^{1/4}}$, so as before we have

$$\frac{1}{B^{1/4}} \leq (BC)K^{-H}$$

which implies that

$$H < \frac{\log (B^{5/4} C)}{\log K} .$$

If the inequality (1) does not hold then we have $q\beta = c + \varepsilon$, where c is an integer and $|\varepsilon| < 5/B^{1/4}$. If there is a solution of (2) with H in the range (3) then we certainly have

$$|q\omega_1| < \frac{1}{4} \quad \text{and} \quad |q\omega_2| < \frac{1}{4} .$$

This implies that

$$b_1 p_1 + b_2 p_2 + b_3 q - c = 0$$

and

$$|b_1 q\omega_1 + b_2 q\omega_2 - \varepsilon| < (B C) K^{-H} . \qquad (4)$$

We can now use our previous reduction process to find all the integral solutions of (4) with H in the range (3).

Again we remark that inequality (1) has never failed in any numerical case which has been tested.

§. III. - So for I have not spoken about the computing techniques which must be used in order to apply the above reduction methods. First of all one must have an electronic computer to-gether with efficient multiprecision arithmetic routines, for we are calculating with numbers which may be several thousand decimal digits long. Writing such a set of programs is fairly straightforward but rather labourious. I have such a program suitable for an I. B. M. 360/67 machine. If anybody wants a copy of it I will gladly send them a copy.

Once one has a multiprecision package the computation of the algebraic numbers $\{\alpha_1, \ldots, \alpha_n\}$, which are roots of polynomial equations, is simple and well known. One uses the Newton approximation method.

However we must also compute $\{\log \alpha_1, \ldots, \log \alpha_n\}$ and it seems that a very efficient way of doing this is less well known. The standard text-books on Numerical Analysis suggest either using a modified series expansion, of $\log (1+x)$ or using a continued fraction expansion, due to Thiele. The methods are very useful when one only need the logarithm correct to about 16 decimal

places. For higher accuracy they are both fairly inefficient.

The method which we adopted in our calculations was as follows. Given the real number α, we require $\log \alpha$ to several thousand decimal places, and we compute $\log \alpha$ by solving the equation $0 = f(x) = e^x - \alpha$ using the Newton approximation method. If one starts with a reasonable approximation to $\log \alpha$, say to 10 decimal places, and arranges to take advantage of the nice properties of e^x, then one can compute $\log \alpha_o$ to about 2000 decimal places in 8 iterations. This will only take about 15 seconds computation time.

In a similar manner one can compute $\tan^{-1}(\alpha)$, a function which is useful for computing the arguments of complex numbers.

-:-:-:-

W. J. ELLISON
U. E. R. de Mathématiques
et d' Informatique
Université de Bordeaux 1
351, cours de la Libération
33 - T A L E N C E