ALFRED J. VAN DER POORTEN
Computing the lower bound for linear forms in logarithms
Séminaire Delange-Pisot-Poitou. Théorie des nombres, tome 17, n° 1 (1975-1976), exp. n° 17, p. 1-9
<http://www.numdam.org/item?id=SDPP_1975-1976__17_1_A16_0>
Denote by $\alpha_1, \alpha_2, \ldots, \alpha_n$, $n > 2$, non-zero algebraic numbers of degree respectively not exceeding $D$ and heights respectively not exceeding $A_1 \ldots, A_n$ where $A_1 \leq \cdots \leq A_{n-1} = A' \leq A_n = A$. We shall write $\Omega' = (\log A_1) \ldots (\log A_{n-1})$ and $\Omega = \Omega' \log A$ (It is convenient to suppose that $\log A_1 \geq 2$, say, and $\log \Omega' \geq 1$). Further by $\beta_0, \beta_1, \ldots, \beta_n$ ($\beta_n \neq 0$), we denote algebraic numbers of degree not exceeding $D$ and heights not exceeding $B$; in the event that these coefficients be rational integers we denote by $b_1, \ldots, b_n$ ($b_n \neq 0$) rational integers of height not exceeding $B$ (It is convenient to suppose that, say, $\log B > 1$ but the results cited are quite trivial unless $B$ be considerably larger). The above notation is used throughout in the sequel without further comment.

It is intended to briefly describe recent techniques whereby one obtains sharp lower bounds for expressions of the shape $\Lambda = b_1 \log \alpha_1 + \cdots + b_n \log \alpha_n$. It is a trivial matter (see lemma 6 for details) to conclude that

$$\Lambda = 0 \quad \text{or} \quad |\Lambda| > (Dh)^{-4nDB} > \exp(-4nb^2(\log A)B).$$

On the other hand, any sharpening of the trivial inequality (1) in the variable $B$ is non-trivial, and may have dramatic implications for a variety of problems in number theory; see, for example, the book BAKER [8], and the surveys TILDEMAN [27], [28]. We will not consider applications here.

1. A brief comment on past results.

For the case $n = 2$, SEL'FOND derived non-trivial estimates of the following shape: there is a $B_0 = B_0(A_1, A_2, \epsilon, D)$ such that for $B > B_0$ and every $\epsilon > 0$

$$\frac{\log \alpha_2}{\log \alpha_1} - \beta_1 > \exp(-(\log B)^{\kappa + \epsilon}).$$

Eventually this result was sharpened to allow $\kappa = 2$ (for details see, say, the book, SEL'FOND [14]). Notice that although $B_0$ is computable, it is not explicitly specified in terms of $A_1, A_2,$ and $D$.

In 1966, and then subsequent papers, BAKER ([1], [2], [3], [4]) revolutionised the theory by extending Sel'don's method to the case of $n > 2$ logarithms, obtaining bounds of the shape

$$|\beta_0 + \beta_1 \log \alpha_1 + \cdots + \beta_n \log \alpha_n| > C \exp(-(\log B)^{\kappa})$$
for $\lambda > n + 1$, $C = C(A_1, \ldots, A_n, D, n, \lambda) > 0$ effectively computable,
under the condition that either $\log \alpha_1, \ldots, \log \alpha_n$ or $\beta_1, \ldots, \beta_n$ be linearly independent
over the rationals, or $\beta_0 \neq 0$. FEL'DMAN [13] sharpened this result to obtain the bound

$$\exp(-C(\log A)2n^2 + 4n \log B)$$

for $C = C(n, D) > 0$ effectively computable, provided that $\log \alpha_1, \ldots, \log \alpha_n$ be linearly independent
over $\mathbb{Q}$. This result is best possible in $B$.

Subsequent results also display the dependence on each of $\alpha_1, \ldots, \alpha_n$ in the case of rational integer coefficients; we write $A = b_1 \log \alpha_1 + \ldots + b_n \log \alpha_n$. Then BAKER [5] showed that

$$\Lambda = 0 \text{ or } |\Lambda| > \exp(-C \log A \log B),$$

with $C = C(n, D) > 0$, effectively computable, and TILDEMAN [29] refined this result to obtain $C = C(n, D)(\log A)^2n^2 + 7n$. BAKER [7] also proved that

$$\Lambda = 0 \text{ or } |\Lambda| > \exp(-C(n, D)\Omega \log \Omega \log B),$$

and recently, by combining the ideas of these papers, van der POORTEN [19] obtained

$$C(n, D) = (2^5(n + 1)D)^{6(n+1)+1}.$$ 

It is the above results we propose to discuss. For comments on results of a somewhat different shape see say [27], also [8]. Similarly, comment on recent results for forms with algebraic coefficients may be found in [27], or see BAKER [6], STAUK [26], SHOREY [24]. Much sharper results are possible if all of $\alpha_1, \ldots, \alpha_n$ are close to 1; see, say SHOREY [23].

There are $p$-adic analogues of the results mentioned. See SCHINZEL [22], for the case $n = 2$, BRUENER [11], COATES [12], SPRINDŽUK [25] and, in particular, KAUPMAN [16], who gives a result possible in $B$ in the general case of algebraic coefficients. Recently, van der POORTEN [19] obtained the analogue of (2), namely, if $b_n \neq 0 \pmod{p}$,

$$(3) \alpha_1 \cdots \alpha_n = 1 \text{ or } |\alpha_1 \cdots \alpha_n - 1|_p > p^{-C(p,n,D)\Omega \log \Omega \log \log \log B},$$

where $p$ is a prime ideal of the field $K$. Subject to the same independence condition mentioned above, it is shown in van der POORTEN [21] that one may take in (3)

$$C(p, n, D) = (\text{Norm } p - 1)(2^5(n + 1)D)^{6(n+1)+2}.$$
The explicit good dependence on $\delta$ is particularly striking.

2. Outline of the proof of Baker's inequality.

A description of the basic principles underlying the proofs of the propositions mentioned may be found in TIJDEHMAN [27], and, in more precise detail, in the book, BAKER [8]. I will attempt to emphasise only those aspects that have led to the recent refinements in the inequalities.

2.1. To show that, say, $A = 0$ or $|A| > \exp(-C \Omega' \log \Omega' \log A \log B)$, one commences by supposing, contrary to what one wishes to prove, that there do exist rational integers $b_1, \ldots, b_n$ ($b_n \neq 0$) of heights not exceeding $B$ such that

$$0 < |b_1 \log \alpha_1 + \ldots + b_n \log \alpha_n| < \exp(-C \Omega' \log \Omega' \log A \log B).$$

One then constructs an appropriate exponential polynomial in $n$ variables

$$\phi(z_0, \ldots, z_{n-1}) = \sum_{\mu_0=0}^{L_0} \sum_{\lambda_1=0}^{L_1} \ldots \sum_{\lambda_n=0}^{L_n} q(\mu_0, \lambda_1, \ldots, \lambda_n) \mu_0 \gamma_1 z_1 \ldots \gamma_{n-1} z_{n-1},$$

where $\gamma_r = \lambda_r - b_r \lambda_n/b_r$ $(1 < r < n)$, with rational integer coefficients $q(\mu_0, \lambda_1, \ldots, \lambda_n)$ such that $\phi$ and many of its partial derivatives are very small at many points $z_0 = \ldots = z_{n-1} = 1, 2, 3, \ldots, h$. Actually, this vital opening step introduces the grossest inefficiency into the argument, because we are actually only interested in total derivatives of the function $F(z) = \phi(z, z, \ldots, z)$, but we have no way of suitably estimating these derivatives except by way of partial derivatives of $\phi$ along the diagonal; the problem is the appearance of powers of logarithms, and we have no way of coping with these until we obtain much deeper results on the algebraic independence of logarithms.

The construction depends on noticing that

$$\alpha_1 z_1 \ldots \alpha_n z_n - \alpha_1 \ldots \alpha_{n-1} = \alpha_1 \ldots \alpha_n (1 - (\alpha_1 \ldots \alpha_n)^{b_1 \ldots b_n - \lambda_n z/b_n}),$$

whence by the assumption (3) this difference is extremely small. Hence if one wants partial derivatives of $\phi$ to be small on the diagonal, then it is good enough to arrange that

$$\sum_{\mu_0=0}^{L_0} \sum_{\lambda_1=0}^{L_1} \ldots \sum_{\lambda_n=0}^{L_n} q(\mu_0, \lambda_1, \ldots, \lambda_n) \frac{\mu_0}{(\mu_0 - m_0)^{m_0}} \gamma_1 \ldots \gamma_{n-1} \alpha_1 \ldots \alpha_n = 0$$

for, say, $z = 1, 2, \ldots, h$ and all non-negative integers $m_0, \ldots, m_{n-1}$ satisfying $m_0 + \ldots + m_{n-1} < M$. One notices that the quantities in (4) are linear forms in the unknowns $q$ with algebraic coefficients. Taking into account that each equation is in effect $D^n$ equations with rational coefficients, we arrange that we have at least twice as many unknowns, namely $(L_0 + 1)(L_1 + 1) \ldots (L_n + 1)$ unknowns $q(\mu_0, \lambda_1, \ldots, \lambda_n)$, as the number of equations to be satisfied, namely roughly $hD^n n^n$ equations, and then the box principle allows us to solve the system (4) so as to obtain integers $q(\mu_0, \lambda_1, \ldots, \lambda_n)$, not all zero, of the same order of size as the coefficients $(\mu_0/(\mu_0 - m_0)^m \gamma_1 \ldots \gamma_{n-1} \alpha_1 \ldots \alpha_n)$. 

17-03
2.2 : Unfortunately if one is seeking a bound sharp in $A$ and $B$, then the parameters one has to choose force the said coefficients to be too large. Some stratagems are required to deal with this problem, and it is these which materially increase the apparent complexity of the argument in recent versions [5], [7], [20], [29] of the proof of Baker's inequality. Firstly, $\mu_0! z^{\mu_0}/(\mu_0 - m_0)!$ is threateningly large, and ruins the sharpness in $B$. To deal with this write

$$\Delta(x ; h) = (x + 1)(x + 2) \ldots (x + h)/h! ; \quad \Delta(x ; 0) = 1$$

and further for any integers $\lambda > 0$, $m_0 > 0$ denote by $\Delta(x ; h, \lambda, m_0)$ the $m_0$-th derivative of $(\Delta(x ; h))^\lambda/m_0!$

Furthermore, notice that instead of the partial derivative

$$\left(\partial^{m_0 + \ldots + m_{n-1}}/\partial z_0 \ldots \partial z_{n-1}\right)$$

it is as convenient to always consider

$$(1/m_0!) m_1! \ldots m_{n-1}! \left(\partial^{m_0 + \ldots + m_{n-1}}/\partial z_0 \ldots \partial z_{n-1}\right)$$

Finally observe that the polynomials $\Delta(x + \lambda_{-1} ; h)\lambda_0$, $\lambda_{-1} = 0, 1, \ldots, h - 1$; $\lambda_0 = 0, 1, \ldots, L_0$ are linearly independent over $K$ and generate the powers $1, x, \ldots, x^0$ if $M_0 \leq h(L_0 + 1)$.

Hence we can, throughout the argument, replace (writing $L_1 + 1 = h$)

$$\sum_{\mu_0 = 0}^{M_0} q(\mu_0, \lambda_1, \ldots, \lambda_n) \frac{\mu_0!}{(\mu_0 - m_0)!} z^{\mu_0 - m_0}$$

by

$$\sum_{\lambda_{-1}}^{L_1} \sum_{\lambda_0}^{L_0} p(\lambda_{-1}, \lambda_0, \ldots, \lambda_n) \Delta(z + \lambda_{-1} ; h, \lambda_0, m_0).$$

This works efficiently by virtue of a lemma of Tijdeman ([29], lemma T1) which tells us that if $q$ and $qx$ are positive integers then $q^{2h(x)}(\nu(h))^\mu_0 \Delta(x, h, \lambda, m_0)$ is a positive integer (where $\nu(h)$ is the lowest common multiple of $1, 2, \ldots, h$), and $\Delta(x ; h, \lambda, m_0) \leq 4^\lambda(x+h)$, $\nu(h) \leq 4^h$ (actually even $e^{1.02h}$). This lemma is quite critical in the most recent arguments.

We now solve for the integers $p(\lambda) = p(\lambda_1, \ldots, \lambda_n)$, and these are of a suitable size relative to $B$.

Secondly, the quantities $\gamma_i$ are too large relative to $A$. In the case of algebraic coefficients $\beta_1, \ldots, \beta_m$ in the original form, there is no solution to this difficulty, and we remain condemned to $(\log A)^{1+c}$ in the bound (see Baker [6]), at least for the present. But when we have integer coefficients $b_1, \ldots, b_n$ then the quantities $\Delta(b_n \gamma_i ; m_i)$ are integers which are not too large (in effect an entire $m_i!$ is eaten by this stratagem), but which strategically generate the required quantities $\gamma_i/m_i!$. It has become standard to write

$$n(z ; m_0, \ldots, m_{n-1}) = \Delta(z + \lambda_1, h, \lambda_0, m_0) \prod_{r=1}^{n-1} \Delta((b_n \lambda_r - b_r \lambda_n) ; m_r)$$
whence the system (4), which we are to solve, becomes

\begin{equation}
(5) \quad g(z; m_0, \ldots, m_{n-1}) = \sum_{\lambda_1=0}^{L-1} \cdots \sum_{\lambda_n=0}^{L_n} p(\lambda) A(z; m_0, \ldots, m_{n-1}) \lambda_1 z^{\lambda_1} \cdots \lambda_n z^{\lambda_n} = 0.
\end{equation}

Having solved the system (5) suitably, we obtain by virtue of (3) that the quantities

\begin{equation}
(6) \quad f(z; m_0, \ldots, m_{n-1}) = \sum_{\lambda_1=0}^{L-1} \cdots \sum_{\lambda_n=0}^{L_n} p(\lambda) A(z; m_0, \ldots, m_{n-1}) \gamma_1 z^{\gamma_1} \cdots \gamma_n z^{\gamma_n}
\end{equation}

are all very small, and this eventually gives us that the appropriate partial derivatives of \( \phi \) are suitably small.

2.3: Now, we are actually not interested in \( \phi \) at all, but only in its total derivatives along the diagonal. We easily see that along the diagonal the partial derivatives \( z_1, \ldots, z_1 \) for \( m_0 + \ldots + m_{n-1} < M \) are generated by \( f(z; m_0, \ldots, m_{n-1}) \) with \( m_0 + \ldots + m_{n-1} < N \); having made this observation (for details see lemma 8 of [19], or lemma 4 of [20]) we need no longer consider \( \phi \) but will refer only to the functions \( f \).

We now come to the first extrapolation argument. Here the basic idea is to trade depth (number of derivatives) for length (number of points). One notices that \( f(z; m_0, \ldots, m_{n-1}) \) small for \( z = 1, \ldots, R \) and \( m_0 + \ldots + m_{n-1} < S \) implies that \( f_m(z; m_0, \ldots, m_{n-1}) \) is small for \( z = 1, \ldots, R \),

\[
m = 0, 1, \ldots, S - S', m_0 + \ldots + m_{n-1} < S'
\]

(here \( f_m(z) = \frac{1}{m!} f^{(m)}(z) \), and the argument goes by way of \( \psi \), as alluded to above). The extrapolation argument then shows that this implies that \( f(z; m_0, \ldots, m_{n-1}) \) is small for \( z = 1, \ldots, R' (> R) \) and \( m_0 + \ldots + m_{n-1} < S' \) (\( < S \)); whence \( g(z; m_0, \ldots, m_{n-1}) \) is small for these ranges of the parameters, so small indeed that these algebraic numbers necessarily vanish. This completes an inductive step, and one repeats the inductive step an appropriate number of times.

A number of remarks are in order. Firstly it should be emphasised that in the extrapolation argument one makes explicit use of the analytic properties of the functions \( f \); so one's eventual result is a consequence of the underlying analytic situation. Secondly, in order to gain by the extrapolation one needs \( R' S'^n > RS^n \) (at any rate in the sharper recent arguments; it was customary to take \( S' = \frac{1}{2} S \) but as noted by SHOREY [24] and confirmed in [20], it seems more efficient to take \( S' \) closer to \( S \)).

2.4: One now uses the data obtained from the extrapolation to perform an interpolation, which allows one to conclude that for some suitable prime \( q \) one has \( f((z/q); m_0, \ldots, m_{n-1}) \) for \( z = 1, \ldots, hq \) and \( m_0 + \ldots + m_{n-1} < q^{-1} M \), is small, and indeed \( g((z/q); m_0, \ldots, m_{n-1}) = 0 \) for the indicated ranges of the parameters.

Until recently (see for example BAKER [5]) one supposed that \( q > L_n \). Then if
one has \([K(\alpha^1_1, ..., \alpha^1_n) : K(\alpha^1_1, ..., \alpha^{n-1}_n)] = q\) (K is the field \(K = \mathbb{Q}(\alpha, ..., \alpha_n)\), it is easy to see that \(g((z/q) ; m_0, ..., m_{n-1}) = 0\) implies, for \((z, q) = 1\)

\[
\sum_{\lambda_{-1}}^{L-1} \sum_{\lambda_n = 0}^{L-n-1} p(\lambda_{-1}, ..., \lambda_n) \mathcal{A}(z; m_0, ..., m_{n-1}) \alpha^n_1 \alpha_{n-1} = 0
\]

for each \(\lambda_n, 0 \leq \lambda_n \leq L_n\). However, with the destruction of \(L_n\), the equations (7) now unravel and imply that \(p(\lambda_{-1}, ..., \lambda_n) = 0\) for all \((n+1)\)-tuples \((\lambda_{-1}, ..., \lambda_n)\) which is contrary to the original construction. Hence the condition on the field extension cannot hold and one must have \(\alpha^{1/q} \in K(\alpha^1_1, ..., \alpha^{n-1}_n)\).

By arguments detailed in Baker and Stark [10] this allows one to set up an inductive chain in which one sequentially has the height \(A\) of \(\alpha_n\) being reduced until the result one wishes to prove is in fact implied by an earlier weaker result. The alternative possibility is that \(g(z)\) vanishes identically not by virtue of the vanishing of the \(p(\lambda)\) but by virtue of a multiplicative relation between \(\alpha_1, ..., \alpha_n\). In that case one reduces \(n\) inductively until for \(n = 1\) the result to be proved is trivially the case.

2.5: A recent innovation (Baker [7]) has been to take \(q\) far smaller, which permits a bound of reasonable quality in \(\Omega'\). Now the extrapolation and interpolation arguments detailed above become an inductive step in a process which reduces \(L_n\) stepwise from \(L_n/q\) to \(L_n/q^{n+1}\) provided that \([K(\alpha^{1/q}_1, ..., \alpha^{1/q}_n) : K] = q^n\) (the argument of [20] follows this process in detail). The induction stops as soon as \(N\) is such that \(q^N\) exceeds \(L_n\). This argument does cost the \(\log \Omega'\) which appears in the bound obtained, but this is efficient indeed compared to the cost of at least \(\bar{\Omega}'^{n+1}\) occasioned by the argument previously detailed. There is a complicated and ingenious argument of Baker [7] which copes with the possibility that \([K(\alpha^{1/q}_1, ..., \alpha^{1/q}_n) : K] < q^n\) but this argument seems inefficient as regards the variables \(n\) and \(D\), and for this reason the arguments of [20] and [21] have been left in their present incomplete state.

2.6: The argument in the p-adic case is virtually identical with that appropriate to the complex case. The principal difference occurs in the extrapolation and interpolation steps where the analytic character of the functions considered is of relevance. Here one can invoke the Schnirelman integral which leaves the argument identical to the shape of the complex argument, but it seems both more appropriate and more efficient to use a technique specific to the p-adic case. The relevant ideas can be found in Schnirelman [22], but the technique as introduced in [18], used in the present context in [19], and, considerably refined, in [21] owes most to Mahler [17]. The interesting problem peculiar to the p-adic case which I want to mention here, is that \(\alpha^\xi = \exp(\xi \log \alpha)\) is a p-adic analytic function defined on the unit disc only if \(|\alpha - 1|_p < p^{-1/p-1}\). Now one notices that without loss of generality it may be supposed that \(\alpha_1, ..., \alpha_n\) are p-adic units, and
then one has \( \alpha_i^{p^{-1}} = 1 \pmod{p} \) \((1 \leq i \leq n)\), and this is the approach used in [9], BAKER and COATES [12]; that is, one replaces the \( \alpha_i \) by, respectively, \( \alpha_i^{p^{-1}} \). This approach is not however efficient as regards the quality of the bound relative to \( p \). A careful reading of HASSE [15] (I was alerted to this source by a conversation with Kurt MAHLER) reminds one that if \( \zeta \) is a primitive \( p \)-th root of unity, where \( p = N - 1 \), then for appropriate integers \( r_i \), \( 0 \leq r_i < p \), one already has \( \alpha_i^{p r_i} = 1 \pmod{p} \). A strategic use of this idea does seem to be efficient and yielded the rather good quality of the bound, relative to \( p \), of [21].

3. Concluding comments.

The intention of the preceding outline is to assist the reader in penetrating recent proofs of Baker's inequality. It is also more than just incidental that this report emphasises the rapid developments in the field. Notwithstanding apparent difficulties which seem to impede much further progress, it seems only reasonable to suppose that further developments will be forthcoming. In particular it is almost surely inefficient to prove general results which are to be applied in a variety of contexts; at this stage applications seem likely to be more suitably approached by means of inequalities tailored to the specific case.

Even though recent \( p \)-adic results appear to be of the same quality as the corresponding complex results it seems to me that there are grossly unsatisfactory aspects to the proofs presently employed. Put at its simplest, the proofs too closely follow the complex pattern. It is little wonder that there has been, by these techniques, no penetration at all into deeper, truly \( p \)-adic problems. A notorious instance is that (other than for experimental results) we seem to know very little about \( |2^{p-1} - 1|_p \). It is perhaps useful to remark that it is this kind of problem which is probably the obstruction to any attempt to remove the dependence on \( p \) from the bound in the \( p \)-adic case of Baker's inequality.

4. Recent news.

At the recent meeting on "Transcendence Theory and its Applications", at Cambridge, it was announced that A. BAKER had proved for the most general case that

\[
0 < |\beta_0 + \beta_1 \log \alpha_1 + \ldots + \beta_n \log \alpha_n| < \exp(-C(n,D)n!) \log n' \log A(\log B + \log n')
\]

with \( C(n, D) = (nD)^{on} \) (and \( c \) approximately 100), has no solution in algebraic numbers \( \beta_0, \ldots, \beta_n \) of height not exceeding \( B \) (The result supposes \( n > 4 \), \( D > 4 \)). P. L. CLJUSOUW and M. WALDSCHMIDT (Linear forms and simultaneous approximations) announced a similar result (which did not however explicitly compute the constant), and particularly interestingly indicated that their method avoided the inductive extrapolation argument which heretofore has been a feature of all the proofs. It would appear that their idea may be particularly efficient if \( n \).
is small, say $n = 2$ or $3$, and especially if $D = 1$; in general, this new approach probably is not very sharp in the parameters $n$ and $D$. The bound (8) does not depend upon any principles other than those described briefly above. I showed that there was no difficulty of obtaining on appropriate $p$-adic analogue of the result of Baker, above thus generalising and sharpening the result of KAUFMAN [16]. The results of Baker and of mine will appear in the proceedings of the conference mentioned above (probably in the Springer Lecture Notes) possibly not in the form or under the titles implied in this lecture.

REFERENCES


(Texte reçu le 2 février 1976)

Alfred J. van der Poorten
School of Mathematics
The University of New South Wales
Kensington (Sydney), NSW 2031
(Australie)