

Dear Dr. Geuze,

Thank you for the question. Here is the situation as I remember it long ago in 1985. I should stress that these are my personal recollections, and I have not tried to reconcile them with anyone else's. It is possible that Oesterlé or Szpiro or Mason may have differing points of view.

I attended a talk by Oesterlé in the Max Planck Institute Bonn, and he wrote on the board a conjecture relating the discriminant and conductor of elliptic curves defined over the rational field. It is very likely that the name of Szpiro was mentioned. For the special elliptic curves defined by

$$y^2 = x(x - a)(x + b) \tag{1}$$

with non-zero coprime rational integers  $a, b$  this conjecture amounted to an inequality

$$(abc)^2 \leq C(S(abc))^\lambda \tag{2}$$

where  $c$  is defined by  $a + b + c = 0$ ,  $S(n)$  is the "squarefree kernel" defined as the product of all positive primes dividing  $n$  (not counted with multiplicity), and  $\lambda, C$  are unspecified constants independent of  $a$  and  $b$ . I am fairly sure that he wrote also (1) and (2) on the board, or at least something very similar and essentially equivalent.

As I recall it, this conjecture was put forward as a conjecture about elliptic curves (which are somewhat sophisticated concepts), with (2) merely a special case. I clearly remember thinking that in fact (2) has nothing to do with elliptic curves: given any non-zero  $a, b, c$  with  $a + b + c = 0$ , one can define an elliptic curve by (1). I do not claim this thought as any special insight; and it must have also occurred to many others.

So I immediately started to forget about the elliptic curve aspect, and began to consider (2) by itself.

The square occurs on the left-hand side of (2) because it is more or less the discriminant. Of course it is irrelevant if the exponent  $\lambda$  on the right is not specified. I remember thinking that for similar reasons one could replace  $abc$  by its maximum  $\max\{|a|, |b|, |c|\}$ . It was fairly automatic also to think of the analogue where  $\mathbf{Z}$  is replaced by a polynomial ring such as  $\mathbf{C}[t]$ . Then  $\max\{|a|, |b|, |c|\}$  gets replaced by  $\max\{\deg a, \deg b, \deg c\}$ , and  $S(abc)$  - well, not quite by the product of all "primes"  $t - \tau$  dividing  $abc$  - but simply their number. That is, the number of zeroes  $Z(abc)$  of the polynomial  $abc$ , not counted with multiplicity.

Because the degree function is additive, as opposed to the multiplicative absolute value on  $\mathbf{Z}$ , the exponent comes downstairs, and the analogue should be

$$\max\{\deg a, \deg b, \deg c\} \leq \lambda Z(abc) + C.$$

I next thought of a (then) recent inequality of Richard Mason which was much more precise, and I remember going to the library and verifying it, as I vaguely recall in some volume of Springer Lecture Notes. But just now I checked my memory on this point; it was indeed Volume 1068 (the “Noordwijkerhout” proceedings), on page 156 to be precise. There you see (in a slightly different notation)

$$\max\{\deg a, \deg b, \deg c\} \leq Z(abc) - 1 \tag{3}$$

for any non-zero coprime polynomials  $a, b, c$ , not all constants, in  $\mathbf{C}[t]$  with  $a + b + c = 0$ . Thus not only is  $\lambda = 1$ , but  $C = -1$  works in our favour.

So next I tried to return to the situation over  $\mathbf{Z}$ . The literal analogue of (3) (ignoring the favourable  $-1$ ) would be

$$\max\{|a|, |b|, |c|\} \leq S(abc);$$

however I knew that one must allow an extra bit of legroom to accommodate counterexamples coming somehow from the archimedean nature of  $\mathbf{Z}$ . This means taking  $\lambda > 1$  but arbitrarily close to 1; and, as we are now back to exponents, the natural end conjecture becomes

$$\max\{|a|, |b|, |c|\} \leq C(S(abc))^\lambda \tag{4}$$

for any  $\lambda > 1$ ; however the value of  $C$  must now be allowed to depend on  $\lambda$ . This may seem rather close to (2) but now the precise nature of the left-hand side and the exponent have become more critical.

I announced (4) in July 1985 at a London conference in honour of Roth’s sixtieth birthday. The proceedings never made it into publication, but the “Open Problems” survive, and below I reproduce my contribution.

#### PROBLEM OF D.W. MASSER (After Oersterlé [sic])

Disprove (or prove) that for every  $\varepsilon > 0$  there exists  $C(\varepsilon)$  such that

$$\max\{|a|, |b|, |c|\} \leq C(\varepsilon) \left( \prod_{p|abc} p \right)^{1+\varepsilon}$$

for all coprime integers  $a, b, c$  with  $a + b + c = 0$ .

So you see that the problem was rather to disprove it, not prove it. The latter would be hopeless, and at the time the analogy between  $\mathbf{Z}$  and  $\mathbf{C}[t]$  was recognized as a useful guide to the truth, but not absolutely precisely reliable. And indeed shortly afterwards Stewart and Tijdeman showed that in some refined technical sense the  $\varepsilon$  cannot **tend to zero too quickly**.

Anyway, given this analogy, perhaps your question can be transferred back to the polynomial context: how did Mason find his result (3)? It may seem equally surprising in that context.

Generally speaking, polynomials are much easier to handle than integers, because one can meaningfully differentiate. Thus it has been known for ages that the Fermat equation  $x^n + y^n = z^n$  has no non-trivial polynomial solutions when  $n \geq 3$ . For example, one can divide by  $z^n$  with the effect that  $z$  becomes 1 but  $x$  and  $y$  become rational functions in  $\mathbf{C}(t)$ . Then differentiate  $x^n + y^n = 1$  with respect to  $t$  (using Newton's dots), and solve the resulting pair of equations for  $x^n$  and  $y^n$  in terms of  $\frac{\dot{x}}{x}$  and  $\frac{\dot{y}}{y}$ . Now in these logarithmic derivatives the multiplicities have disappeared, and therefore  $\frac{\dot{x}}{x}$  and  $\frac{\dot{y}}{y}$  have comparatively small degrees. On the other hand the degrees of  $x^n$  and  $y^n$  are comparatively large if  $n \geq 3$ . This leads quickly to a contradiction. Of course Newton himself could have done all this; and we already have the main ideas behind the proof of (3) and all its subsequent generalizations.

More recently than Fermat, a 1965 problem of Davenport on the degree of  $f^3 - g^2$  could be solved by using  $a = f^3, b = -g^2$  in (3). Note that the resulting equation

$$g^2 = f^3 + c \tag{5}$$

looks like an elliptic curve (1)!

And it is not only polynomials that can be differentiated; also entire functions (Borel) and meromorphic functions (Cartan-Nevanlinna) too. Especially with Nevanlinna the above simple proof could be generalized to give far-reaching generalizations of the famous Picard Theorem that a non-constant entire function takes every value with at most one exception.

So there came early into existence a simple and powerful strategy for solving certain problems involving functions, which can be expressed: "When in doubt, differentiate".

In fact Mason came to (3) not directly through polynomials or meromorphic functions, but through the theory of linear forms in logarithms invented by his research supervisor

Alan Baker, who used it among other things to solve equations like

$$x^3 - 3xy^2 + y^3 = 1 \tag{6}$$

in integers  $x$  and  $y$ . An old idea is to factorize the left-hand side of (6) as  $(x - \theta y)(x - \phi y)(x - \psi y)$  with algebraic numbers  $\theta, \phi, \psi$ ; here they all happen to lie in a cubic extension  $K$  of  $\mathbf{Q}$ . It follows that  $a = x - \theta y$ ,  $b = x - \phi y$ ,  $c = x - \psi y$  are units in  $K$ . They may not satisfy  $a + b + c = 0$  but something quite like it involving some extra fixed coefficients independent of the unknowns  $x$  and  $y$ . The unit group of  $K$  has two generators, say  $\xi$  and  $\eta$ . This already means that the prime factors (in  $K$ ) of  $a, b, c$  are severely restricted, and in fact their number is bounded independently of  $x$  and  $y$ . So the  $K$ -analogue of  $S(abc)$  in (4) is also bounded, showing that (4) is liable to provide useful information. Unfortunately we cannot prove (4) even over  $\mathbf{Z}$ , let alone over  $K$ ! What Baker did was to rewrite (after some additional arguments) the equation like  $a + b + c = 0$  as  $\xi^r \eta^s = 1 + \epsilon$ , where  $r$  and  $s$  are in  $\mathbf{Z}$  and  $\epsilon$  is very small. He then took logarithms to give  $r \log \xi + s \log \eta =$  very small (actually there should be an extra multiple of  $2\pi i$  here, but never mind), to which his theory of linear forms in logarithms could be applied. The end result (already proved long ago with other methods) is that (6) has at most finitely many solutions  $x$  and  $y$  in  $\mathbf{Z}$ .

Mason's original project was to get the same results over  $\mathbf{C}[t]$ . I imagine that he started out by trying to extend the preceding argument to  $\mathbf{C}[t]$ . A lot goes through, but then there is the problem of extending the theory of linear forms in logarithms. A lot of work! At some point Mason must have realized that one can simply differentiate when one gets to the step involving (4). This now has the form (3), so one can use the dot argument above. The right-hand side of (3) is bounded, and this leads to the finiteness of the solution set, all without taking any logarithms.

So Mason's contribution, apart from differentiating, was to think of (3) freed from the idea that the right-hand side is bounded. So the right-hand side is unrestricted, and this finally led to the simple formulation. It also fitted well with the Davenport Problem, in which nothing is said about the prime factors of  $f$  and  $g$  in (5).

Yours sincerely,

David Masser (18th February 2006).