NOTES ON FERMAT'S LAST THEOREM VIII

ceNTRe
Macquarie University
NSW 2109 Australia

A. J. van der Poorten

Centre for Number Theory Research, Macquarie University

It seems perfectly reasonable to say that if something is true locally everywhere then it is true globally, and conversely. The only question is what on earth one might mean in saying such a thing in the context of diophantine problems. Let me make some suggestions, starting with the Riemann ζ-function $\zeta(s) = \sum n^{-s}$. By the way, here s is a complex variable, traditionally given as $s = \sigma + it$. Of course the series converges absolutely only in the right plane $\sigma > 1$. We owe to Euler the 'factorisation'

$$\sum_{n=1}^{\infty} \frac{1}{n^s} = \prod_p \Big(1 - \frac{1}{p^s}\Big)^{-1}.$$

Indeed, the typical Euler factor is $1+p^{-s}+p^{-2s}+p^{-3s}+\cdots$, and the product representation amounts to the unique factorisation theorem whereby each positive integer has a unique representation as a product of primes. In this case we might hope to obtain local information, that is, about the primes, from global knowledge, of the positive integers. For example, because $1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} \sim \log n$, we find, after taking a logarithm, that $\sum_{p<x} p^{-1} \sim \log\log x$. Since $\log\log x = \int_e^x (t\log t)^{-1}\,dt$, we might deduce that the m-th prime is $\sim m\log m$ and that, therefore, the number of primes less than x is $\sim x/\log x$.

But I have something more down-to-earth in mind. Suppose I am wondering whether there are rational points on the elliptic curve $x^4 - 17 = 2y^2$. I will probably set $x = X/Z$, $y = Y/Z$ and multiply by Z^4 to obtain a homogeneous polynomial in the *projective* co-ordinates X, Y and Z — this also has the advantage of giving a glimpse of ∞: just consider the line $Z = 0$. But, whatever, after failing to notice a solution we might well choose to consider the problem mod p — with the intention of employing the evident principle that if there is no solution mod m there is certainly none *globally*; that is, there then cannot be a rational solution. However, in the present example there is a solution mod p for all p. That doesn't require an infinite amount of work; in fact it suffices to check by trial and error at $p = 2$ and 17, the other *places* being quite automatic. Of course we do want to deal with an arbitrary modulus m, so we also need the powers of the primes. However, once one has a solution mod p it is easy to see whether or not there is a solution mod p^n for all n. Let me illustrate that by showing that $x^2 \equiv 2$

1991 *Mathematics Subject Classification.* 11D41, 11-01.
Work supported in part by grants from the Australian Research Council and a research agreement with Digital Equipment Corporation.

(mod 7^n) has a solution for all positive integers n, given that $3^2 \equiv 2 \pmod 7$. It will give a feeling of generality if I set $f(x) = x^2 - 2$ and $p = 7$.

Suppose we already know that $f(x_{n-1}) \equiv 0 \pmod{p^n}$. We want $x_n = x_{n-1} + xp^n$ so that $f(x_{n-1} + xp^n) \equiv 0 \pmod{p^{n+1}}$. But the Taylor expansion

$$f(x_n) = f(x_{n-1} + xp^n) = f(x_{n-1}) + xp^n f'(x_{n-1}) + \cdots ,$$

where the $\cdots$ indicate terms in yet higher powers of p, tells us that

$$0 \equiv p^{-n} f(x_{n-1}) + xf'(x_{n-1}) \equiv p^{-n} f(x_{n-1}) + xf'(x_0) \pmod p;$$

recall that $p^n \mid f(x_{n-1})$ and $x_{n-1} \equiv x_0 \pmod p$. So we have only to solve a simple linear congruence in x to obtain x_n. This argument is known as Hensel's lemma. Mind you, there is a little more work to be done if x_0 happens to be a singular point, that is if $f'(x_0) \equiv 0$, but ultimately things settle down to just a sequence of linear problems as above. So it's almost painless to decide whether there is a solution mod p^∞, so to speak. In our example we obtain

$$x_\infty = 3 + 1\cdot 7 + 2\cdot 7^2 + 6\cdot 7^3 + 1\cdot 7^4 + 2\cdot 7^5 + 1\cdot 7^6 + 2\cdot 7^7 + 4\cdot 7^8 + \cdots .$$

It is a well known principle both of magic and of mathematics that once one knows the true name of an object one controls it. Thus we say that x_∞ is a 7*-adic integer* and observe that we can do arithmetic with objects such as it much as if they were absolutely convergent power series in powers of 7, provided we remember to 'carry' to cope with coefficients outside the range $\{0, 1, 2, 3, 4, 5, 6\}$.

To see what is actually going on here it is best first to remember that we barely understand the real numbers $\mathbb{R}$. We would be satisfied to say that the reals are the set of all decimals and proud to pop out with the statement that $\mathbb{R}$ *is the field of limit points of convergent sequences of rationals*; which it is. If we were asked to clarify this claim we'd remark that a sequence $(a_h) = \{a_0, a_1, a_2, \dots\}$ of rationals is convergent if it is Cauchy, that is, if $|a_n - a_m|$ is arbitrarily small as soon as m and n are sufficiently large; we'd show that the termwise difference and product of Cauchy sequences is again Cauchy; and we'd agree to identify such sequences if they have the same limit — that is, if they differ by a *null sequence*: one for which $|a_n|$ is as small as we like whenever n is large enough. In a trice we would have taken a leap in sophistication and would hear ourselves saying that $\mathbb{R}$ is just *the ring of Cauchy sequences of rationals modulo its maximal ideal of null sequences*.

The joker in this explanation is that we could give it without telling anyone what we meant by 'small'; that is, what we mean by the absolute value $|\quad|$. Well, yes, we'd have used that $|\quad|$ is a positive definite map on the rationals preserving multiplication — $|a| \ge 0$ and $= 0$ if and only if $a = 0$; and $|a \cdot b| = |a| \cdot |b|$. And we will have needed the *triangle inequality*: $|a + b| \le |a| + |b|$. But that's it.

So a listener could say, "Aha! If I were to pick my favourite prime p then I could take a rational r/s and apply unique factorisation to write $r/s = \prod_p p^{v_p}$ with integers v_p, mostly zero but otherwise positive or negative, and set $|r/s| = p^{-v_p}$." On overcoming our surprise we'd have to admit that there is indeed nothing wrong with believing that 'small' means 'divisible by a high power of p'. After making her write $|\ \ |_p$ instead of just $|\ \ |$ to avoid confusion (we might agree to write $|\ \ |_\infty$ for the absolute value we thought we had meant, so as not to make too special a claim for it) we'd find that all was well. She has indeed nominated an absolute value and her *completion* $\mathbb{Q}_p$ — the p-adic rationals — has as much right to respect as does the completion $\mathbb{R}$.

Is $\mathbb{R}$ special? Well. physical reality, as we, and Archimedes, believe it, has no infinitesimals. In $\mathbb{R}$, given $a \neq 0$ no matter how close to 0, and b no matter how large, there is always an integer n so that $|na|_\infty > |b|_\infty$. We say that $\mathbb{R}$, and thence the absolute value $|\ \ |_\infty$, are *archimedean*. On the other hand, the p-adic absolute values satisfy the *ultrametric* inequality $|a+b|_p \leq \max\bigl(|a|_p, |b|_p\bigr)$, a stronger form of the triangle inequality. That entails $|na|_p \leq |a|_p$ for all integers n. We say that the fields $\mathbb{Q}_p$, and thence the absolute values $|\ \ |_p$, are *nonarchimedean*. By the way, *pace* Archimedes, in recent years the theoretical physicists have learned about the p-adic fields and have begun to wonder whether they may not help in modelling just what happens in the nuclei of atoms, for instance. So much for reality.

The x_∞ of our example is an element of the field $\mathbb{Q}_7$. It is in fact a 7-adic *integer*, that is, it satisfies $|x_\infty|_7 \leq 1$. Just as $\mathbb{Q}$ is the field of quotients of elements of $\mathbb{Z}$, so the fields $\mathbb{Q}_p$ are the respective quotient fields of the domains $\mathbb{Z}_p$. The elements of $\mathbb{Q}_p$ can be represented as sums $\sum_{n=-m} a_n p^n$ just as the elements of $\mathbb{R}$ are often represented as decimals $\sum_{n=-m} b_n 10^{-n}$. All's well, because just as 10^{-n} gets small at exponential rate in $\mathbb{R}$, so p^n becomes small in $\mathbb{Q}_p$.

The bottom line is that checking diophantine equations mod m can usefully be viewed as asking for solutions in the fields $\mathbb{Q}_p$ of p-adic rationals. That's true, by the Chinese Remainder Theorem, whereby any congruence modulo $m = \prod p^{\operatorname{ord}_p m}$ is the same thing as a collection of congruences modulo the respective $p^{\operatorname{ord}_p m}$. One talks about looking for solutions *at* p; these are the solutions *locally*. But we'd better not forget about ∞. The diophantine equation $x^2 + y^2 = -1$ has solutions at all p, but it really has no rational solutions, because it has no solutions in $\mathbb{R}$. Thus, when we speak about *locally everywhere* we mean at all p *and* ∞. The absolute values are not totally unrelated. For nonzero x, unique factorisation gives the 'product formula' $\sum_p \log|x|_p = -\log|x|_\infty$.

With all that said, is it true that if a diophantine equation has solutions locally everywhere — thus in all $\mathbb{Q}_p$ and in $\mathbb{R}$ — then it has *global* solutions — thus in $\mathbb{Q}$? Sadly, no. There is a theorem of Hasse-Minkowski showing that a *quadratic form* — a homogeneous polynomial expression of degree 2 — takes a given rational value if and only if it has a solution locally everywhere. But this *Hasse principle* does not apply generally. In fact I obnoxiously chose the example equation $x^4 - 17 = 2y^2$ to display an equation with solutions locally everywhere but with no rational solutions. Nonetheless, our view of diophantine equations is that they would like to obey the Hasse principle. We will want to study the obstruction to their doing so.

An elementary aside. Suppose that r/s is a rational zero of the polynomial $f(x) = a_0x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n \in \mathbb{Z}[x]$. If $p \mid r$ then this zero is 0 $(\mathrm{mod}\ p)$ so 0 must be a zero of $f(x)$ $(\mathrm{mod}\ p)$. That is, $a_n \equiv 0 \pmod{p}$, or $p \mid a_n$. This beginning of a local argument of course eventually entails that $r \mid a_n$. Now consider the so-called *reciprocal* polynomial $x^n f(x^{-1})$. We see that it has s/r as a zero and conclude that $s \mid a_0$.

An instructive *degenerate* case of this discussion deals with the case of a zero 0 of $f(x)$. Since anything whatsoever divides 0, anything whatsoever divides a_n, showing that of course $a_n = 0$. Now the preceding argument would have us say that the reciprocal polynomial has ∞ as a zero. Indeed, by rights it should be a polynomial of degree n. It displays its zero ∞ by actually being of lower degree.

A generalisation is Gauß's lemma which asserts that if a polynomial $f(x)$ with integer coefficients has a factorisation as a product of polynomials with rational coefficients then it already has a factorisation into a product of polynomials with integer coefficients. To see this we write $df(x) = g(x)h(x)$, where the nonzero integer d is a multiplier required so that g and h have integer coefficients. If we fear that $p \mid d$ we look at the identity modulo p, obtaining $0 \equiv g(x)h(x) \pmod p$. But, with p prime, this entails that either $g(x) \equiv 0$ or $h(x) \equiv 0 \pmod p$. So we may divide through by p, in particular replacing d by d/p. Then by descent we may conclude that d has no prime divisors, thus $d = 1$ as we wished to show.

The terminology 'global' and 'local', and the talk of 'at' p is sometimes jolting. Its genesis can be readily illustrated. Suppose we think about the arithmetic of functions $f(z)$ defined on $\mathbb{C}$. It seems reasonable to consider the polynomials as the integers and the *rational functions* — quotients of polynomials — as the rationals. What are the irreducibles? That's easy, they're just the linear polynomials $z - \alpha$, one for each point $\alpha \in \mathbb{C}$. So that's why we say 'at' a prime! To say that $z - \alpha$ divides some polynomial f exactly $m = \operatorname{ord}_\alpha f$ times, is to say that f has a zero of order $\operatorname{ord}_\alpha f$ at α; or, if you prefer, that the Taylor expansion of f at α has $(m!)^{-1} f^{(m)}(\alpha)(z-\alpha)^m$ as its first nonzero term.

And just as the numbers have a special 'prime' corresponding to the absolute value $|\ \ |_\infty$, so there is a place in $\mathbb{C}$, the point* at ∞. The fundamental theorem of algebra that, in $\mathbb{C}$, a polynomial f of degree $\deg f$ has exactly $\deg f$ zeros counted according to multiplicity, becomes the formula $\sum_{\alpha\in\mathbb{C}} \operatorname{ord}_\alpha f = -\operatorname{ord}_\infty f$.

Geometry with an ultrametric is great fun. Every triangle is isosceles and each point inside a circle is its centre. However, we're interested in arithmetic and what pleases us is that in $\mathbb{Q}_p$ the integers $\mathbb{Z}$ no longer discretely keep their distance one from the other. That allows one to look at difficult functions like the Riemann ζ-function in a new light. After all, we are a bit blasé to take a function like n^{-s} for granted. An honest integral power, like n^{-2k}, is fine; it just means n^{-2} multiplied by itself k times. But to suddenly generalise that to an analytic function $e^{-s\log n}$ is a drastic step. In that spirit, one might well feel that except for the perfectly decent rational numbers $\zeta(2k)/\pi^{2k}$ all the other $\zeta(s)$ should be disregarded. But, it was nice to have an analytic function. Now a miracle comes to the rescue. Kummer's congruences: if $k \equiv k' \pmod{p^n}$ then $(1-p^{k-1})B_k/k \equiv (1-p^{k'-1})B_{k'}/k' \pmod{p^{n+1}}$, amount exactly to our being able to view the numbers $(-1)^k(1-p^{2k-1})(2k-1)!\zeta(2k)/2^{2k-1}\pi^{2k}$ as the values of a p-adic analytic function of a p-adic variable k. This is the p-adic ζ-function.

---

* Plainly $\mathbb{C}$ is not a plane, but a sphere. Just put a globe on the origin of the plane and draw a line from the North Pole N to each point in $\mathbb{C}$. Now identify each point in $\mathbb{C}$ with the point on the globe cut by its line. Clearly N corresponds to ∞, and we may speak of *the* point at ∞.