Appendix 6A. Polynomial solutions of Diophantine equations

6.6. Fermat's Last Theorem in $\mathbb{C}[t]$

The notation $\mathbb{C}[t]$ denotes polynomials whose coefficients are complex numbers. In section 6.1 we saw that all integer solutions to $x^2+y^2=z^2$ are derived from letting t be a rational number in the polynomial solution

$$(t^2-1)^2+(2t)^2 \;=\; (t^2+1)^2.$$

We now prove that there are no "genuine" polynomial solutions to Fermat's equation

$$x^p+y^p=z^p \tag{6.6.1}$$

with exponent p larger than 2 (where by *genuine* we mean that $(x(t),y(t),z(t))$ is not a polynomial multiple of a solution of (6.6.1) in complex numbers).

Proposition 6.6.1. *There are no genuine polynomial solutions $x(t),y(t),z(t)\in\mathbb{C}[t]$ to $x(t)^p+y(t)^p=z(t)^p$ with $p\geq 3$.*

Proof. Assume that there is a solution with x, y, and z all non-zero to (6.6.1) where $p\geq 3$. We may assume that x, y, and z have no common (polynomial) factor or else we can divide out by that factor (and that they are pairwise coprime by the same argument as in section 6.1). Our first step will be to differentiate (6.6.1) to get

$$px^{p-1}x'+py^{p-1}y'=pz^{p-1}z'$$

and after dividing out the common factor p, this leaves us with

$$x^{p-1}x'+y^{p-1}y'=z^{p-1}z'. \tag{6.6.2}$$

We now have two linear equations (6.6.1) and (6.6.2) (thinking of x^{p-1}, y^{p-1}, and z^{p-1} as our variables), which suggests we use linear algebra to eliminate a variable: Multiply (6.6.1) by y' and (6.6.2) by y, and subtract, to get

$$x^{p-1}(xy' - yx') = x^{p-1}(xy' - yx') + y^{p-1}(yy' - yy') = z^{p-1}(zy' - yz').$$

Therefore x^{p-1} divides $z^{p-1}(zy' - yz')$, but since x and z have no common factors, this implies that

$$x^{p-1} \text{ divides } zy' - yz'. \tag{6.6.3}$$

This is a little surprising, for if $zy' - yz'$ is non-zero, then a high power of x divides $zy' - yz'$, something that does not seem consistent with (6.6.1).

Now, if $zy' - yz' = 0$, then $(y/z)' = 0$ and so y is a constant multiple of z, contradicting our statement that y and z have no common factor. Therefore (6.6.3) implies, taking degrees of both sides, that

$$(p-1)\ \deg(x) \le \deg(zy' - yz') \le \deg(y) + \deg(z) - 1,$$

since $\deg(y') = \deg(y) - 1$ and $\deg(z') = \deg(z) - 1$. Adding $\deg(x)$ to both sides gives

$$p\ \deg(x) < \deg(x) + \deg(y) + \deg(z). \tag{6.6.4}$$

The right side of (6.6.4) is symmetric in x, y, and z. The left side is a function of x simply because of the order in which we chose to do things above. We could just as easily have derived the same statement with y or z in place of x on the left side of (6.6.4), so that

$$p\ \deg(y) < \deg(x) + \deg(y) + \deg(z)$$
$$\text{and} \quad p\ \deg(z) < \deg(x) + \deg(y) + \deg(z).$$

Adding these last three equations together and then dividing out by $\deg(x) + \deg(y) + \deg(z)$ implies

$$p < 3,$$

and so Fermat's Last Theorem is proved, at least for polynomials. □

That Fermat's Last Theorem is not difficult to prove for polynomials is an old result, going back certainly as far as Liouville in 1851.

Exercise 6.6.1. Prove that all solutions to $x(t)^2 + y(t)^2 = z(t)^2$ in polynomials are a scalar multiple of some solution of the form $(r(t)^2 - s(t)^2)^2 + (2r(t)s(t))^2 = (r(t)^2 + s(t)^2)^2$.

6.7. $a + b = c$ in $\mathbb{C}[t]$

We now intend to extend the idea in our proof of Fermat's Last Theorem for polynomials to as wide a range of questions as possible. It takes a certain genius to generalize to something far simpler than the original. But what could possibly be more simply stated, yet more general, than Fermat's Last Theorem? It was Richard C. Mason (1983) who gave us that insight: *Look for solutions to*

$$a + b = c.$$

We will just follow through the above proof of Fermat's Last Theorem for polynomials (Proposition 6.6.1) and see where it leads: Start by assuming, with no loss

of generality, that a, b, and c are all non-zero polynomials without common factors (or else all three share the common factor and we can divide it out). Then we differentiate to get

$$a'+b'=c'.$$

Next we need to do linear algebra. It is not quite so obvious how to proceed analogously, but what we do learn in a linear algebra course is to put our coefficients in a matrix and solutions follow if the determinant is non-zero. This suggests defining

$$\Delta(t):=\begin{vmatrix} a(t) & b(t) \\ a'(t) & b'(t)\end{vmatrix}.$$

Then if we add the first column to the second, we get

$$\Delta(t)=\begin{vmatrix} a(t) & c(t) \\ a'(t) & c'(t)\end{vmatrix},$$

and similarly

$$\Delta(t)=\begin{vmatrix} c(t) & b(t) \\ c'(t) & b'(t)\end{vmatrix}$$

by adding the second column to the first, a beautiful symmetry.

We note that $\Delta(t)\neq 0$, or else $ab'-a'b=0$ so b is a scalar multiple of a (with the same argument as above), contradicting our hypothesis.

To find the appropriate analogy to (6.6.3), we consider the power to which the factors of a (as well as b and c) divide our determinant: Let α be a root of $a(t)$, and suppose that $(t-\alpha)^e$ is the highest power of $(t-\alpha)$ which divides $a(t)$ (we write $(t-\alpha)^e\|a(t)$). Now we can write $a(t)=U(t)(t-\alpha)^e$ where $U(t)$ is a polynomial that is not divisible by $(t-\alpha)$, so that $a'(t)=(t-\alpha)^{e-1}V(t)$ where $V(t):=U'(t)(t-\alpha)+eU(t)$. Now $(t-\alpha,V(t))=(t-\alpha,eU(t))=1$, and so $(t-\alpha)^{e-1}\|a'(t)$. Therefore

$$\Delta(t)=a(t)b'(t)-a'(t)b(t)=(t-\alpha)^{e-1}W(t)$$

where $W(t):=U(t)(t-\alpha)b'(t)-V(t)b(t)$ and $(t-\alpha,W(t))=(t-\alpha,V(t)b(t))=1$ as $t-\alpha$ does not divide $b(t)$ or $V(t)$. Therefore we have proved that

$$(t-\alpha)^{e-1}\|\Delta(t).$$

This implies that $(t-\alpha)^e$ divides $\Delta(t)(t-\alpha)$. Multiplying all such $(t-\alpha)^e$ together we obtain (since they are pairwise coprime) that

$$a(t) \text{ divides } \Delta(t)\prod_{a(\alpha)=0}(t-\alpha).$$

In fact $a(t)$ only appears on the left side of this equation because we studied the linear factors of a; analogous statements for $b(t)$ and $c(t)$ are also true, and since $a(t),b(t),c(t)$ have no common roots, we can combine those statements to read

$$a(t)b(t)c(t) \quad \text{divides} \quad \Delta(t)\prod_{(abc)(\alpha)=0}(t-\alpha). \tag{6.7.1}$$

The next step is to take the degrees of both sides of (6.7.1). The degree of $\prod_{(abc)(\alpha)=0}(t-\alpha)$ is precisely the total number of distinct roots of $a(t)b(t)c(t)$.

Therefore

$$\text{degree}(a) + \text{degree}(b) + \text{degree}(c) \leq \text{degree}(\Delta) + \#\{\alpha \in \mathbb{C} : \ (abc)(\alpha) = 0\}.$$

Now, using the three different representations of Δ above, we have

$$\text{degree}(\Delta) \leq \begin{cases} \text{degree}(a) + \text{degree}(b) - 1, \\ \text{degree}(a) + \text{degree}(c) - 1, \\ \text{degree}(c) + \text{degree}(b) - 1. \end{cases}$$

Inserting all this into the previous inequality we get

$$\text{degree}(a), \text{degree}(b), \text{degree}(c) < \#\{\alpha \in \mathbb{C} : \ (abc)(\alpha) = 0\}.$$

Put another way, this result can be read as:

Theorem 6.3 (The *abc* Theorem for Polynomials)**.** *If $a(t), b(t), c(t) \in \mathbb{C}[t]$ do not have any common roots and provide a genuine polynomial solution to $a(t)+b(t) = c(t)$, then the maximum of the degrees of $a(t), b(t), c(t)$ is less than the number of distinct roots of $a(t)b(t)c(t) = 0$.*

This is a "best possible" result in that we can find infinitely many examples where there is exactly one more zero of $a(t)b(t)c(t) = 0$ than the largest of the degrees, for example the familiar identity

$$(2t)^2 + (t^2 - 1)^2 = (t^2 + 1)^2;$$

or the rather less interesting

$$t^n + 1 = (t^n + 1).$$

Exercise 6.7.1. Let a, b, and c be given non-zero integers, and suppose $n, p, q, r > 1$.

(a) Prove that there are no genuine polynomial solutions $x(t), y(t), z(t)$ to $ax^n + by^n = cz^n$ with $n \geq 3$.

(b) Prove that if there is a genuine polynomial solution $x(t), y(t), z(t)$ to $ax^p + by^q = cz^r$ in which x, y, and z have no common root, then $\frac{1}{p} + \frac{1}{q} + \frac{1}{r} > 1$.

(c) Deduce in (b) that this implies that at least one of p, q, and r must equal 2.

(d) One can find solutions in (b) if one allows common factors, for example $x^3 + y^3 = z^4$ where $x = t(t^3 + 1)$ and $y = z = t^3 + 1$. Generalize this construction to as many other sets of exponents p, q, r as you can. (Try to go beyond the construction in exercise 6.5.8.)

Exercise 6.7.2. Let a and b be given non-zero integers, $p, q > 1$, and $x(t), y(t) \in \mathbb{C}[t]$. Let D be the maximum of the degrees of x^p and y^q, and assume that $ax^p + by^q \neq 0$.

(a) Prove that the degree of $ax^p + by^q$ is $> D(1 - \frac{1}{p} - \frac{1}{q})$.

(b)† Prove that if $g = (p, q) > 1$, then the degree of $ax^p + by^q$ is $\geq D/g$.

(c) Deduce that the degree of $ax^p + by^q$ is always $> D/6$.
(This is "best possible" in the case $(t^2 + 2)^3 - (t^3 + 3t)^2 = 3t^2 + 8$.)

Appendices. The extended version of chapter 6 has the following additional appendices:

Appendix 6B. *No Pythagorean triangle of square area via Euclidean geometry* presents another proof (due to a student, Stephanie Chan, in 2017) of this theorem of Fermat, now via clever geometric manipulations.

Appendix 6C. *Can a binomial coefficient be a square?* addresses and resolves the question of whether a binomial coefficient can be a square.

Appendix 6B. No Pythagorean triangle of square area via Euclidean geometry

In this appendix we use Euclidean geometry to show that there is no integer-sided right-angled triangle whose area is a square,[6] rather than the algebraic methods of section 6.3. In this proof one sees algebra and geometry working together, foreshadowing a theme one frequently encounters as one studies advanced mathematics.

An algebraic proof, by descent

We will suppose that there are integer-sided right-angled triangles with square area and establish a contradiction. We take the integer-sided right-angled triangle ABC whose area is a square with smallest hypotenuse. By (6.1.1) its sides have lengths

$$AB = 2MN, \quad AC = N^2 - M^2, \quad \text{and} \quad BC = M^2 + N^2,$$

where M and N are coprime positive integers of different parities. The area of this triangle, $MN(N-M)(N+M)$, is a square by hypothesis. We now prove that the factors M, N, $N-M$, and $N+M$ are pairwise coprime:

We have $(M, N \pm M) = (N \pm M, M) = (N, M) = 1$. Finally let $g = (N-M, N+M)$. Then g is odd since $N-M$ is odd. Moreover g divides both $(N+M)-(N-M) = 2M$ and $(N+M)+(N-M) = 2N$, so that g divides $(2M, 2N) = 2(M, N) = 2$. The only odd positive integer g dividing 2 is $g = 1$.

We have proved that the product $MN(N-M)(N+M)$ is a square and that the factors M, N, $N-M$, and $N+M$ are pairwise coprime. Since they are all

[6] This proof is due to a student, Stephanie Chan, working with me in London in 2017.