Mason-Stothers Theorem and the ABC Conjecture

Rohan Mukherjee

May 21, 2023

Abstract

In this paper we look at the interesting Mason-Stothers theorem, its history, and how this theorem relates closely to the ABC conjecture.

Throughout this paper, we will work towards understanding and proving:

Theorem 1 (Mason-Stothers)**.** *Let $a(t), b(t)$, and $c(t)$ be relatively prime polynomials over a field such that $a + b = c$ and such that not all of them have vanishing derivative. Then* $\max\{\deg(a), \deg(b), \deg(c)\} \leq \deg(\mathrm{rad}(abc)) - 1$.

1 Required background knowledge

This theorem is pretty abstract. First, to even understand what it is saying, we need to define everything. We will quickly see that polynomial rings over a field share many of the same properties as the integers, which is also why this theorem is closely related to the ABC conjecture, but that will come later on. First,

Definition 1. A nonempty set R with two operations $+ : R \times R \to R$, and $\cdot : R \times R \to R$ is a *ring* if it has the following properties:

(i) $(a + b) + c = a + (b + c)$.

(ii) $a + b = b + a$.

(iii) There is an element $0 \in R$ so that $0 + a = a + 0 = a$ for every $a \in R$.

(iv) For every $a \in R$, there exists an $x \in R$ so that $a + x = 0$.

(v) $a \cdot (b \cdot c) = a(bc) = (ab)c$.

(vi) $a(b + c) = ab + ac$, and $(a + b)c = ac + bc$.

R is said to be commutative if it also satisfies $ab = ba$ for every $a, b \in R$. R is said to have identity if there is an element $1 \in R$ so that $1a = a1 = a$ for every $a \in R$.

Example. $\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$ are all commutative rings with identity with the usual addition and multiplication.

Example. The trivial ring $\{0\}$ is a commutative ring with identity, with the operations being $0 + 0 = 0$ and $0 \cdot 0 = 0$. It's additive and multiplicative identities are both 0.

The trivial ring can often lead to problems, so definitions tend to exclude it. The underlying structure of rings is not strong enough for us to state the theorem in. For example, what if there was two polynomials so that they are both nonzero, but their product is? Then we would get a situation where we are comparing the degrees of nonzero things to the degree of something 0, which is not desirable. So we have a name for rings where this doesn't happen:

Definition 2. An integral domain is a commutative ring R with identity $1 \neq 0$ (This condition excludes the trivial ring) so that if $a, b \neq 0$, then $ab \neq 0$. Equivalently, if $ab = 0$, then either $a = 0$ or $b = 0$.

Finally, we may give the definition of a field.

Definition 3. A field is a commutative ring k with identity $1 \neq 0$ so that given any $a \neq 0 \in k$, there exists an $x \in k$ so that $xa = ax = 1$.

Theorem 2. *Additive and multiplicative inverses are unique, and every field is an integral domain.*

Proof. Let R be a ring and $a \in R$. Suppose that b, c satisfy $a + b = a + c = 0$. Adding b to the left of both equations, we get $(b + a) + b = (b + a) + c$. Since $b + a = 0$, we get that

$b = c$, as claimed. Similarly, suppose k is a field, and $a \neq 0 \in k$. Let b, c satisfy $ab = ac = 1$. Multiplying by b on the left side, we get that $(ba)b = (ba)c$, and since $ba = ab = 1$, we see that $b = 1 \cdot b = 1 \cdot c = c$, as claimed. We write the unique additive inverse as $-a$, and the unique multiplicative inverse as a^{-1}. First, given any $a \in R$ where R is any ring, $a \cdot 0 = a \cdot (0 + 0) = a \cdot 0 + a \cdot 0$, which says that $a \cdot 0 = 0$. A similar argument works when 0 is on the left of a. Suppose there were $a, b \neq 0 \in k$ so that $ab = 0$. Since $a \neq 0$, it has a multiplicative inverse, so we see that $b = (a^{-1} \cdot a)b = a^{-1} \cdot 0 = 0$, a contradiction. □

Definition 4. Subring Let R be a ring. A set $\emptyset \neq S \subset R$ is called a subring of R if for every $x, y \in S$, $xy \in S$, and $x - y \in S$.

Now that we know what a field is, we need to discuss the "relatively prime" condition in our theorem. What does it mean for a polynomial to be relatively prime? First, it helps to understand that it means for *integers* to be relatively prime. We know that two integers are relatively prime if their greatest common divisor is 1. So maybe something similar works for polynomials. But first, we need a definition:

Definition 5. Let R be a commutative ring with identity. $R[x]$, the polynomial ring in x with coefficients from R, is the set of all formal sums $a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$, with $n \geq 0$, and each $a_i \in R$. If $a_n \neq 0$, n is called the degree of the polynomial, and a_n is it's leading coefficient. The polynomial is monic if $a_n = 1$. Addition is defined as follows:

$$\sum_{i=0}^{n} a_i x^i + \sum_{i=0}^{n} b_i x^i = \sum_{i=0}^{n} (a_i + b_i) x^i$$

where a_i or b_i can be zero so addition of polynomials of different degrees to be defined. Multiplication of polynomials is defined as

$$\left(\sum_{i=0}^{n} a_i x^i \right) \times \left(\sum_{i=0}^{m} b_i x^i \right) = \sum_{k=0}^{n+m} \left(\sum_{i=0}^{k} a_i b_{k-i} \right) x^k$$

Defined like this, $R[x]$ is a commutative ring with identity (the identity is the same as the identity from R). R can be identified inside $R[x]$ as the subring of all constant polynomials (Dummit & Foote, 2003).

We continue with one last theorem:

Theorem 3. *Let R be an integral domain, and let $f(x), g(x) \in R[x]$. Then* $\deg f(x)g(x) = \deg f(x) + \deg g(x)$.

Proof. If $f(x) = \sum_{i=0}^{n} a_i x^i$ (with $a_n \neq 0$), and $g(x) = \sum_{i=0}^{m} b_i x^i$ (with $b_m \neq 0$), then $f(x)g(x) = \sum_{k=0}^{n+m} \left(\sum_{i=0}^{k} a_i b_{k-i}\right) x^k$ by definition. The largest power of x in this sum is $n+m$, with coefficient $a_n b_m$. Since $a_n \neq 0$ and $b_m \neq 0$, $a_n b_m \neq 0$ (integral domain!), the degree of $f(x)g(x)$ is $n+m$. □

This next theorem highlights how polynomial rings over a field are similar to the integers.

Theorem 4 (Division Algorithm)**.** *Let a, b be integers with $b > 0$. Then there exists unique integers q, r so that*

$$a = bq + r \text{ and } 0 \leq r < b.$$

For example, $4 = 3 \cdot 1 + 1$, where $1 < 3$. Given a, b, one can find q, r using integer long division. This relates very closely to polynomial rings because...

Theorem 5 (Division Algorithm)**.** *Let k be a field, $f(x), g(x) \in k[x]$, with $g(x) \neq 0$. Then there exists unique polynomials $q(x), r(x)$ so that*

$$f(x) = g(x)q(x) + r(x) \text{ with } r(x) \equiv 0 \text{ or } \deg r(x) < \deg g(x)$$

The condition $r(x) \equiv 0$ has to be added because the 0 polynomial technically doesn't have degree. For example, let our field be $\mathbb{R}$. The constant polynomials $f(x) \equiv 1$, $g(x) \equiv 2$ certainly satisfy this condition–for $1 = 2 \cdot \frac{1}{2} + 0$, and as $\deg g(x) = 0$, if we didn't add this condition we would have to say $\deg 0 < 0$, which is weird. Anyways, this theorem is a (near) identical copy of the division algorithm for the integers. Once again you can find $q(x)$ and $r(x)$ using the division algorithm. As another example, in the ring $\mathbb{R}[x]$, if $f(x) = x^3 + 1$, and $g(x) = x^2 + 1$, $x^3 + 1 = (x^2 + 1)(x) + (-x + 1)$, and $\deg -x + 1 = 1 < \deg x^2 + 1 = 2$. The proof of this theorem is long and tedious, so it will be omitted, but one can find a proof on page 91 of (Hungerford, 2012). This fact, along with many others about polynomial rings in fields, is how Mason-Stothers is going to relate to the real ABC conjecture. We are close to defining what it means to be relatively prime as polynomials, but first we need to talk about divisibility.

Definition 6. Let k be a field, and $a(x), b(x) \in k[x]$ with $b(x)$ nonzero. $b(x)$ divides $a(x)$, written $b(x) \mid a(x)$, if $a(x) = b(x)c(x)$ for some $c(x) \in k[x]$.

Something about polynomial rings that is not in the integers is the following theorem:

Theorem 6. *Let k be a field and $a(x), b(x) \in k[x]$ with $b(x)$ nonzero.*

(1) If $b(x) \mid a(x)$, then $cb(x) \mid a(x)$ for every $0 \neq c \in k$.

(2) If $d(x) \mid a(x)$, then $\deg d(x) \leq \deg a(x)$

In the integers, if $4 \mid 8$ then, even though $3 \neq 0$, $4 \cdot 3 \nmid 8$, so this is a bizarre property. Due to this property, we have a new definition:

Definition 7. $f(x)$ is an associate of $g(x)$ in $k[x]$ if $f(x) = c \cdot g(x)$ for some $0 \neq c \in k$.

This definition is relevant because if one recalls the fundamental theorem of arithmetic, factorization was defined "uniquely up to reordering", but in polynomial rings we have to define it "uniquely up to both reordering and associates". Finally, gcd!

Definition 8. Let k be a field and $a(x), b(x) \in k[x]$ not both zero. The greatest common divisor of $a(x)$ and $b(x)$ is $d(x)$ given $d(x)$ is monic, and

(1) $d(x) \mid a(x)$, $d(x) \mid b(x)$,

(2) If $c(x) \mid a(x)$, and $c(x) \mid b(x)$, then $\deg c(x) \leq \deg d(x)$.

Here the "greatest" means in the sense of degree, not in the sense of absolute value like from integers. Also, monic is required so the gcd is unique (Or else, all associates of the monic gcd would also be gcds!). Finally,

Definition 9. Let k be a field. $a(x), b(x) \in k[x]$ are said to be relatively prime if $\gcd(a(x), b(x)) = 1$.

This is precisely the same definition for integers. Similar theorems come over, such as

Theorem 7 (Bezout's Identity for Polynomials)**.** *Let k be a field, and $a(x), b(x) \in k[x]$ not both zero. Then there is a unique greatest common divisor $d(x)$ of $a(x)$ and $b(x)$. Furthermore, there are polynomials $u(x)$, $v(x)$ so that $d(x) = a(x)u(x) + b(x)v(x)$.*

Proof. This proof is nearly identical for the integers. Let

$$S = \{a(x)m(x) + b(x)n(x) \mid m(x), n(x) \in k[x]\}.$$

S contains nonzero polynomials, such as $a(x) \cdot 1 + b(x) \cdot 0$, or $a(x) \cdot 0 + b(x) \cdot 1$ (Note: at least one has to be nonzero by hypothesis). So, the set of degrees of polynomials in S is a nonempty set of nonnegative integers, and therefore has a smallest element. So there is a polynomial $w(x)$ of smallest degree in S. If t is the leading coefficient of $w(x)$, $d(x) = t^{-1}w(x)$ is a *monic* polynomial of smallest degree in S (it is still in S, since we can just move the constant into the factors of $a(x), b(x)$). By the division algorithm, $a(x) = d(x) \cdot q(x) + r(x)$, where $\deg r(x) < \deg d(x)$ or $r(x) = 0$. We notice that, since $d(x) = a(x)n(x) + b(x)m(x)$ for some $n(x), m(x) \in k[x]$, $r(x) = a(x) - d(x) = a(x)(1 - n(x)) + b(x)m(x)$, which is in S. Since d is a polynomial of smallest positive degree in S, it cannot be the case that $\deg r(x) < \deg d(x)$, and so $r(x) \equiv 0$, which shows that $d(x) \mid a(x)$. A similar argument shows $d(x) \mid b(x)$. If $c(x) \mid a(x)$ and $c(x) \mid b(x)$, then $t(x)c(x) = a(x)$ and $y(x)c(x) = b(x)$, so we see that $d(x) = c(x)t(x)n(x) + c(x)y(x)m(x) = c(x)(t(x)n(x) + y(x)m(x))$, which shows that $c(x) \mid d(x)$. Let $w(x)$ be another greatest common divisor of $a(x)$ and $b(x)$. Since $w(x)$ is a common divisor of a, b, $\deg w(x) \leq \deg d(x)$. Since $d(x)$ is a common divisor of a, b, $\deg d(x) \leq \deg w(x)$ (since $w(x)$ is also a gcd). So $\deg d(x) = \deg w(x)$. The above shows that $w(x) \mid d(x)$, i.e. that $w(x) \cdot p(x) = d(x)$. Then $\deg p(x) = 0$, i.e. $p(x) = p$ for some $p \in k$. Since both are monic, we see that $p = 1$, which shows that the gcd is unique. □

We get an important consequence, whose proof is copied from the integers,

Theorem 8. *Let k be a field, and $a(x), b(x), c(x) \in k[x]$. If $a(x) \mid b(x)c(x)$, and $a(x), b(x)$ are relatively prime, then $a(x) \mid c(x)$.*

Proof. Since $a(x), b(x)$ are relatively prime, there are $u(x), v(x)$ so that $a(x)u(x) + b(x)v(x) = 1$. Multiplying both sides by $c(x)$ gives us $a(x)u(x)c(x) + c(x)b(x)v(x) = c(x)$. Since $a(x) \mid c(x)b(x)$, $c(x)b(x) = d(x)a(x)$ for some $d(x)$. This tells us that $a(x)(u(x)c(x) + d(x)) = a(x)u(x)c(x) + a(x)d(x) = c(x)$, i.e. that $a(x) \mid c(x)$. □

We are almost able to define the radical:

Definition 10. Let k be a field. A nonconstant polynomial $p(x) \in k[x]$ is said to be **irreducible** if its only divisors are it's associates and the nonzero constant polynomials. A nonconstant polynomial that is not irreducible is said to be **reducible**.

Finally, we define the radical:

Definition 11. Let k be a field, and $f(x) \in k[x]$. $\text{rad}(f)$ is the product of distinct irreducible divisors of f, where $f(x), g(x) \in k[x]$ are distinct if they are not associates.

For example, in $\mathbb{C}[x]$, $\text{rad}((x-1)^2(x+1)) = (x-1)(x+1)$. One notes in this case that $\deg \text{rad}((x-1)^2(x+1)) = 2$, which is the number of distinct roots of $(x-1)^2(x+1)$. This is true in general, although I could not find a proof of this fact ((Lang, 2005) mentions the idea of this proof, but does not give it in full). My attempt at proving this goes at follows: if we work over $\mathbb{C}[x]$, it is clear enough that if we have complex numbers $\alpha_1, \ldots, \alpha_n \in \mathbb{C}$, where $\alpha_i \neq \alpha_j$ for $i \neq j$, then letting $p(x) = \prod_{i=1}^{n}(x - \alpha_i)$, $\deg(p(x)) = n$, which is indeed the number of distinct roots of $p(x)$. Then given an arbitrary $q(x) \in \mathbb{C}[x]$, if q has a root of multiplicity $n \geq 2$ at $x = \beta$, then $q(x) = (x-\beta)^n \cdot r(x)$ for some $r(x)$ with $(x-\beta) \nmid r(x)$. $\text{rad}(q(x))$ will now be $(x-\beta)\text{rad}(r(x))$ since these polynomials have distinct roots. So any factor containing a multiple root will have the multiple root reduced to a single root–therefore counting the number of distinct roots. Note in $\mathbb{C}[x]$ that $\deg(f(x))$ counts the number of roots, multiplicity included.

Definition 12 (Algebraically closed)**.** A field k is algebraically closed if every non-constant polynomial in $k[x]$ has a root in k.

Which gives:

Corollary 1. *A polynomial $p(x) \in k[x]$ where k is algebraically closed of degree $n \geq 1$ has n roots, multiplicity included.*

Proof. Every polynomial of degree 1 is of the form $x - \alpha$. Clearly α is a root of this equation, establishing the base case. Suppose that every polynomial of degree n has exactly n roots in k, for some $n \geq 1$. Let $q \in k[x]$ be a polynomial of degree $n+1$. By the definition of algebraically closed, q has a root in k, say β. By the division algorithm, we can write

$q(x) = (x - \beta)p(x) + r$ for some $r \in k$. Evaluating both sides at β shows that $r = 0$. $p(x)$ is a polynomial of degree n, since k is an integral domain, which has n roots. So q has $n + 1$ roots, which completes the proof. □

As an example, proving that $\mathbb{C}$ is algebraically closed was on the first homework. We are ready to see all 3 Mason-Stothers theorems, with 2 proofs, one of which I have omitted because it's proof is beyond the scope of both my understanding and this paper.

2 Statement and proof of the theorem

The original theorem in Stothers's paper states:

Theorem 9 (Stothers). *Let $k(f)$ be the leading coefficient of f, $d(f)$ the degree of f, and finally $c(f)$ the number of distinct zeros of f. Suppose that $p, q \in \mathbb{C}[z]$ with $k(p) = k(q)$, $d(p) = d(q) > 0$, and p, q relatively prime. Then*

$$c(p) + c(q) + c(p - q) \geq d(p) + 1 \tag{1}$$

with equality if and only if the non-trivial branch points of

$$R(z, w) = (1 - w)p(z) + wq(z) = 0 \tag{2}$$

lie over $w \in \{0, 1, \infty\}$.

First, I will show that this is a special case of the aforementioned Mason-Stothers theorem. Any divisor of both p, and $p - q$ will also divide $p - q - p = -q$, so any common divisor of $p - q, p$ is also a common divisor of p, q. The gcd of p, q is just 1, so indeed, $\gcd(p - q, p) = 1$. Similarly, $\gcd(p - q, q) = 1$. One also notes that $p - q + q = p$, so we can take $a = p - q$, $b = q$, and $c = p$ and rewrite our theorem as follows: $c(p) + c(q) + c(p - q) \geq d(p) + 1$. Finally, since p, q have the same leading coefficient and degree, $p - q$ has degree strictly less than $p - q$. Since $d(p) = d(q)$, $\max\{\deg(a), \deg(b), \deg(c)\} = \deg(b) = \deg(p)$. Similarly, since p, q, and $p - q$ are all pairwise coprime, $c(p) + c(q) + c(p - q) = c(pq(p - q)) = c(abc)$ (we may combine them because each of p, q, and $p - q$ have distinct roots from each other). So this theorem is truly a special case of the real one.

The proof of this theorem is very complicated, involving the heavy tools of algebraic geometry. So, I won't be discussing the proof here, but Stothers's original paper has some

really funny and fascinating things about it. For example, the following definition is due to (STOTHERS, 1981):

Definition 13. A pair $[p, q]$ of complex polynomials is special (of degree n) if p, q satisfy the hypothesis of Theorem 1.1 (the previous theorem) and give equality in (2) (and $d(p) = d(q) = n$).

He uses this definition to cite equality in his next theorem (Theorem 1.2). As a consequence of Theorem 1.2, he makes this definition (STOTHERS, 1981):

Definition 14. A pair $[p, q]$ of complex polynomials is *extra-special* (of degree n) if they give equality in (4) (and each has degree n).

(Equality in (4) has to do with Theorem 1.2). Interestingly enough, he uses Theorem 1.1 to prove something called Davenports theorem, which I will also prove a special case of later on, after discussing Mason's proof of the Mason-Stothers theorem. By using the tools of algebraic geometry, he is essentially slaughtering this theorem, as it can be proved using nothing but elementary calculus. A great reason to do so, however, is that he establishes equality in his theorems, which is not something you can get from just using calculus. We shall let $n_0(f)$ = the number of distinct roots of f, i.e. that $n_0(f) = \deg \operatorname{rad}(f)$. Mason's version of the theorem is the one from before:

Theorem 10. *(Lang, 2005) Let $a(t), b(t), c(t)$ be relatively prime polynomials such that $a+b = c$. Then*

$$\max\{\deg(a), \deg(b), \deg(c)\} \leq n_0(abc) - 1$$

Algebra by Lang gives the proof, but not in its entirety. It also has a lot of exercises that talk about consequences of Mason's theorem, including the aforementioned Davenports theorem. The proof goes as follows:

Proof. (Lang, 2005) If we let $f = a/c$, $g = b/c$, we get that $f + g = 1$, where f, g are rational functions. Differentiating gives us $f' + g' = 0$, and multiplying by 1 gives us $f'/f \cdot f + g'/g \cdot g = 0$. Rearranging gives us $(f'/f)/(-g'/g) = g/f = (b/c)/(a/c) = b/a$. Since $a(t) \in \mathbb{C}[t]$, it can be written as the product of linear factors, i.e. $a(t) = c_1 \prod (t - \alpha_i)^{m_i}$.

Similarly, $b(t) = c_2 \prod (t-\beta_j)^{n_j}$, and finally $c(t) = c_3 \prod (t-\gamma_k)^{r_k}$. Next, we need something called the **logarithmic derivative**. Let $L(x) = x'/x$. I claim that $L(xy) = L(x) + L(y)$. Indeed, $L(xy) = (xy)'/xy = (yx' + xy')/xy = x'/x + y'/y = L(x) + L(y)$. Side note: If you have ever wondered what "$\sin(x)/x$ has no elementary anti-derivative" actually means, it turns out that this logarithmic derivative is crucial helpful in proving this fact (See (Wikipedia contributors, 2023)). Anyways, from this we know that for any function $R(x) = c\prod (x-\alpha_i)^{m_i}$, we see that $R'/R = L(R) = L(c) + \sum L((x-\alpha_i)^{m_i})$. Furthermore, notice that $L((x-\alpha_i)^{m_i}) = \underbrace{L(x-\alpha_i) + \cdots + L(x-\alpha_i)}_{m_i \text{ times}} = m_i L(x-\alpha_i)$. So, $\sum L((x-\alpha_i)^{m_i}) = \sum m_i L(x-\alpha_i)$. Finally, $L(x-\alpha_i) = \frac{\mathrm{d}}{\mathrm{d}x}(x-\alpha_i)/(x-\alpha_i) = 1/(x-\alpha_i)$. We conclude that

$$R'/R = \sum \frac{m_i}{x-\alpha_i}$$

One also notes that $L(f^{-1}) = \frac{\mathrm{d}}{\mathrm{d}x} f^{-1}/f^{-1} = (-1/f^2 \cdot f')/f^{-1} = -f'/f = -L(f)$. We see that

$$f'/f = (a/c)'/(a/c) = L(a/c) = L(a) + L(c^{-1}) = L(a) - L(c) = \sum \frac{m_i}{t-\alpha_i} - \sum \frac{r_k}{t-\gamma_k}$$

And similarly,

$$g'/g = \sum \frac{n_j}{t-\beta_j} - \sum \frac{r_k}{t-\gamma_k}$$

Therefore,

$$\frac{b}{a} = -\frac{f'/f}{g'/g} = -\frac{\sum \frac{m_i}{t-\alpha_i} - \sum \frac{r_k}{t-\gamma_k}}{\sum \frac{n_j}{t-\beta_j} - \sum \frac{r_k}{t-\gamma_k}}$$

One notices that $N_0 = \prod (t-\alpha_i) \prod (t-\beta_j) \prod (t-\gamma_k)$ is a common denominator for this big fraction (for example, multiplying $1/(t-\beta_j) - 1/(t-\gamma_k)$ by $(t-\beta_j)(t-\gamma_k)$ certainly yields a polynomial), and therefore $b/a = -(N_0 f'/f)/(N_0 g'/g)$ where the RHS is a ratio of polynomials. N_0 is the product of the distinct factors of a, b, c, and therefore has degree $n_0(abc)$. Since $\deg(f') = \deg(f) - 1$, and since we canceled out all fractions, $N_0 \cdot f'/f$ has degree at most $n_0(abc) - 1$ (we are multiplying something of degree $n_0(abc)$ by something of degree at most -1). Since b, a are relatively prime, b/a is a reduced fraction. We conclude that b/a is a fraction in which the numerator and denominator are minimal in degree. In particular, since $N_0 f'/f/N_0 g'/g$ is also a ratio of polynomials that equals b/a, it's numerator has degree at least the degree of b (if it were less, it would contradict b, a being minimal), and similarly with denominator at least the degree of a. We conclude that $\deg a \leq n_0(abc) - 1$,

and $\deg(b) \leq n_0(abc) - 1$. Since $a + b = c$ we know that $\deg(c) \leq \max\{\deg(a), \deg(b)\}$, hence, $\deg(c) \leq n_0(abc) - 1$ as well. Hence, $\max\{\deg(a), \deg(b), \deg(c)\} \leq n_0(abc) - 1$. □

Although it may not currently seem it, Mason's proof is widely simpler than Stothers. The most important idea in the proof of this theorem is the logarithmic derivative–and, it's incredibly properties! However, as I was saying earlier, nothing in this proof can give us conditions for equality. For some historical context, Mason's proof was discovered independently of Stothers, around 3 years after Stothers discovered it. I do think that Stothers's proof is a little bit like YouTuber Michael Penn's Overkill series. In fact, I think that Stothers use of algebraic geometry is so overkill, I will now give you an even simpler proof of Mason-Stothers, due to Synder.

Lemma 1. *Let $f \in k[x]$ be nonzero. Then* $\deg(f) \leq \deg(\gcd(f, f')) + n_0(f)$

A proof of this lemma is given in (*Synder*,2000). The intuition for this theorem is that if f has a multiple root α, then f' has a root at α with multiplicity the multiplicity of $f - 1$, so noting that $\gcd(f, f')$ is going to count the number of multiple roots of f minus 1 for each root, which is why you have to add back $n_0(f)$ in the end. This intuition is so good in fact that equality holds if $k = \mathbb{C}$.

For a possibly nonobvious detail: if $a, b \mid c$, and $\gcd(a, b) = 1$, then $ab \mid c$, and also if $c \mid a$ and $c \mid b$, then $c \mid \gcd(a, b)$.

Proof. Proof of the first detail: Since $\gcd(a, b) = 1$, by Bezout's identity there are s, t so that $as + bt = 1$. Multiplying by c gives $cas + bct = c$. Since $b \mid c$, $c = bk$ for some k. Similarly, $c = al$ for some l. Plugging this in gives $baks + ablt = c$, which says $ab \mid c$. For the second detail, we know by Bezout that there are some s, t so that $as + bt = \gcd(a, b)$. If c is a common divisor of a, b then $a = ck$ and $b = cl$ for some k, l. Plugging this in gives $cks + clt = \gcd(a, b)$, i.e. that $c \mid \gcd(a, b)$. One should note that I never specified where a, b, s, t, k, l came from, since indeed, this proof works in both the integers AND in polynomial rings. □

The last thing is that $(x - \alpha_i)^{m_i}$, $(x - \alpha_j)^{m_j}$ are relatively prime if $\alpha_i \neq \alpha_j$. Here is his proof:

Proof. If $a + b = c$, then $a' + b' = c'$, which tells us that $a'a + b'a = c'a$. Also, $a'a + a'b = a'c$. Subtracting gives $a'b - ab' = a'c - ac'$ (Yes, this is the Wronskian!). Since $\gcd(a', a) \mid a'$ and $\gcd(a', a) \mid a$, we see that $\gcd(a', a) \mid a'b - ab'$. Similarly, $\gcd(b, b') \mid a'b - ab'$, and finally using the equality in the second line gives $\gcd(c, c') \mid a'c - ac' = a'b - ab'$. Since a, b are relatively prime, a, c and b, c must also be relatively prime, since any divisor of a, c will also divide $c - a = a + b - a = b$, and similar for b, c. This also means that $\gcd(a, a')$, $\gcd(b, b')$ are relatively

prime. The intuition is that we cannot pick up a common factor by removing factors. For if $d \mid \gcd(a, a')$ and $d \mid \gcd(b, b')$, then $d \mid a$ and $d \mid b$, which shows that $d = 1$. By the relatively prime intuition that I gave above, and using that all three factors divide $a'b - b'a$, we see that $\gcd(a, a') \gcd(b, b') \gcd(c, c') \mid a'b - b'a$. If $a'b - b'a = 0$, then $ab' = a'b$, and so $a \mid a'b$. Since a, b are relatively prime, $a \mid a'$ (The proof of this fact follows by Bezout, since there are some s, t so that $as + bt = 1$, multiplying by a' gives $aa's + a'bt = a'$, and since $a \mid a'b$, $a'b = ca$ for some c, then we have that $aa's + act = a'$, which says that $a \mid a'$. All of these number theory facts follow from Bezout.) But we know that $\deg(a') < \deg(a)$ (by the power rule!), unless $a' = 0$. So $a' = 0$. Similarly, b', c' would be 0 (using the right equation, for example for showing $c' = 0$ one would have to use $a'b - b'a = a'c - ac' = 0$). But then a', b', c' would all be 0, contradicting the assumption that a, b, c are not constant. So $a'b - b'a$ is nonzero. We recall that $\deg(a'b - b'a) \leq \max\{\deg(a'b, b'a)\}$. In $\mathbb{C}[x]$, $\deg(a') = \deg(a) - 1$, and similarly $\deg(b') = \deg(b) - 1$ (This result does NOT hold in any algebraically closed field, since things can go wrong with characteristic. For example, in $(\mathbb{Z}/2)[x]$, the derivative of x^2 is $2x$ which is 0, so in that case the degree would be less than the degree of x^2 - 1). In any case, $\max\{\deg(a'b, b'a)\} \leq \deg(a) + \deg(b) - 1$. Also, note that $\deg(\gcd(a, a') \gcd(b, b') \gcd(c, c')) = \deg(\gcd(a, a')) + \deg(\gcd(b, b')) + \deg(\gcd(c, c'))$. Finally, if $b = ca$, then $\deg(b) = \deg(c) + \deg(a)$, and since $\deg(c) \geq 0$, we see that $\deg(b) \geq \deg(a)$. Using all of these facts gives $\deg(\gcd(a, a') \gcd(b, b') \gcd(c, c')) = \deg(\gcd(a, a')) + \deg(\gcd(b, b')) + \deg(\gcd(c, c')) \leq \deg(a) + \deg(b) - 1$. Moving all of these terms to the RHS and adding $\deg(c)$ to both sides gives $\deg(c) \leq \deg(a) - \deg(\gcd(a, a')) + \deg(b) - \deg(\gcd(b, b')) + \deg(c) - \deg(\gcd(c, c')) - 1$. Looking back to the lemma, we see that $\deg(f) - \deg(\gcd(f, f')) \leq n_0(f)$. Applying this 3 times gives that the RHS is $\leq n_0(a) + n_0(b) + n_0(c) - 1$. Since all a, b, c have distinct factors, we see that $n_0(abc) = n_0(a) + n_0(b) + n_0(c)$, which completes the proof. □

This proof is entirely elementary–using just a simple lemma, and some basic number theory facts. Despite this, I actually found this one a little harder to understand than Mason's proof, probably because I really liked the use of the logarithmic derivative. Synder was an undergraduate at Harvard when he discovered this proof, which is truly amazing. The original paper that I referenced has in something similar to an abstract saying that he plans to continue doing math for grad school and a career. Wondering how this went, I looked this up and found that he got a Ph.D. from UC Berkeley, and is now a professor at Indiana University Bloomington. To date he has supervised 3 students. I wanted to go look at some of his original research, but it has to do with things outside of my current knowledge, like low-dimensional topology. I find it incredible however that he found such a simple proof of Mason-Stothers. Now we will move into applications of Mason-Stothers, and the ABC conjecture. First, we can look at Davenport's theorem:

Theorem 11 (Davenport's theorem)**.** *Let f, g be non-constant polynomials such that $f^3 - g^2 \neq 0$. Then* $\deg(f^3 - g^2) \geq \frac{1}{2}\deg f - 1$.

Here is a proof of the special case where f, g are relatively prime:

Proof. Assuming that f, g are relatively prime, it follows that $\gcd(f^3 - g^2, f) = 1$, and that $\gcd(f^3 - g^2, g) = 1$, since any divisor of f and $f^3 - g^2$ will also divide $f^3 - g^2 - f^2 \cdot f = g^2$. If $\gcd(f, g^2) \neq 1$, there would be some linear polynomial $x - \alpha$ dividing both f and g^2. Since $x - \alpha$ is irreducible, it follows that $x - \alpha \mid g$, a contradiction. Similarly, $\gcd(g, f^3 - g^2) = 1$. We may now apply Mason-Stothers:

$$\begin{aligned} 3\deg(f) &\leq \max\{3\deg(f), 2\deg(g), \deg(f^3 - g^2)\} \\ &= \max\{\deg(f^3), \deg(-g^2), \deg(f^3 - g^2)\} \\ &\leq n_0(f^3 \cdot g^2 \cdot (f^3 - g^2)) - 1 \end{aligned}$$

Note now that $n_0(f^3) = n_0(f) \leq \deg(f)$, since n_0 removes all multiple roots, and it can be strictly less since f could still have some multiple roots. Similarly, $n_0(g^2) \leq \deg(g)$. Also, $n_0(ab) \leq n_0(a) + n_0(b)$, since a, b could have non-distinct roots. We see that

$$n_0(f^3 \cdot g^2 \cdot (f^3 - g^2)) - 1 \leq \deg(f) + \deg(g) + \deg(f^3 - g^2) - 1$$

And, comparing the LHS to the RHS, we get that $2\deg(f) \leq \deg(g) + \deg(f^3 - g^2) - 1$. Similarly, since $2\deg(g) \leq \max\{3\deg(f), 2\deg(g), \deg(f^3 - g^2)\}$, we see that $2\deg(g) \leq \deg(f) + \deg(g) + \deg(f^3 - g^2) - 1$, i.e. that $\deg(g) \leq \deg(f) + \deg(f^3 - g^2) - 1$. These together give us

$$2\deg(f) \leq \deg(f) + 2\deg(f^3 - g^2) - 2$$

Which completes the proof in the relatively prime case. □

One again notes that we have no conditions on equality. However, in Stothers's original paper, he is able to use his original condition for equality to conclude that equality holds in Davenports theorem iff $[f^3, g^2]$ is an extra special pair, and fg has distinct zeros (no multiple roots). He also goes on to have definitions for special and extra special groups, which I find fun. Later on, he defines something called "legitimate pairs". Sadly, I was not able to find a definition for extra-extra special pairs. The rest of Stothers paper has to do with some high-level group theory far outside the scope of my knowledge. Mason-Stothers

also gives an insanely quick proof of Fermat's last theorem for polynomials: if $x(t)^n, y(t)^n$ are relatively prime and $x(t)^n + y(t)^n = z(t)^n$, then by Mason-Stothers

$$n \deg(x) = \deg(x^n) \leq n_0(x^n y^n z^n) - 1 \leq \deg(x) + \deg(y) + \deg(z) - 1$$

And similarly, $n \deg(y) \leq \deg(x) + \deg(y) + \deg(z) - 1$ and $n \deg(z) \leq \deg(x) + \deg(y) + \deg(z) - 1$ (we use that $a \leq \max\{a, b, c\}$ 3 times). Summing gives

$$n(\deg(x) + \deg(y) + \deg(z)) \leq 3(\deg(x) + \deg(y) + \deg(z)) - 3$$

And if $n \geq 3$, this shows that $0 \leq -3$, a contradiction. So $x(t)^n + y(t)^n = z(t)^n$ only has solutions for $n \leq 2$. The proof of Fermat's last theorem (for integers) is extremely complicated. Andrew Wiles, the man who proved it, won the Abel Prize, considered the nobel prize in mathematics, for doing so. A simpler version of FLT follows immediately from the real ABC conjecture. Here is (a specialized version of) the real ABC conjecture:

Theorem 12. *Let a, b be relatively prime positive integers and call $c = a + b$. Then for every $\varepsilon > 0$ there is some $C(\varepsilon) > 0$ so that $c \leq C(\varepsilon) n_0(abc)^{1+\varepsilon}$.*

Where if $a = p_1^{\alpha_1} \cdot p_2^{\alpha_2} \cdot \cdots \cdot p_n^{\alpha_n}$, then $n_0(a) = p \cdot p_2 \cdot \cdots \cdot p_n$ (reduce the power of all primes dividing it to have power 1).

This is a sort of restricted form of the abc conjecture that makes the next argument simpler, but the one you will find on Wikipedia has a, b not necessarily positive, and changes $c \leq$ to $\max\{|a|, |b|, |c|\} \leq$. (Lang, 2005) gives an excellent example of this: consider the equation $2^n + 1 = m$. This conjecture would say $c \leq C(\varepsilon) n_0(2^n \cdot 1 \cdot m)^{1+\varepsilon}$, which would say that $m \leq C(\varepsilon) n_0(m)^{1+\varepsilon}$. Taking m to be absolutely massive, say $> 10^{100}/C(\varepsilon)$, would give $10^{100} \leq n_0(m)^{1+\varepsilon}$. Since $n_0(m)$ is the product of primes to the first power, we see that m would have to be the product of large primes to the first power, since 10^{100} is huge. For example, consider $2^{100} + 1$. This has a factor of 3173389601, which is indeed quite big. One could go on. Also, a simpler but still incredibly difficult to prove version of Fermat's Last Theorem, who's solver won $1m for, can be proved in around 3 lines by the abc conjecture. We shall prove the Asymptotic FLT: For all n sufficiently large, $x^n + y^n = z^n$ has no solutions in relatively prime positive integers. Indeed, since for any $\varepsilon > 0$, there is some $C(\varepsilon) > 0$ so that given any a, b relatively prime, if $a^n + b^n = c^n$, we have that

$$c^n \leq C(\varepsilon) n_0(a^n b^n c^n)^{1+\varepsilon} \leq C(\varepsilon)(abc)^{1+\varepsilon}$$

Doing this for a and b and then multiplying these inequalities gives $(abc)^n \leq 3C(\varepsilon)(abc)^{3+3\varepsilon}$. Clearly $a = b = c = 1$ doesn't work. If it did have solutions for all n sufficiently large, then the LHS could be made arbitrarily large while the RHS remains fixed, which is indeed a contradiction. This is an incredibly simple proof of an incredibly complicated theorem, and while this is not the full theorem in it's entire generality, if one could prove that n being sufficiently large means $n \leq 4,000,000$, then you would have proven the full theorem (Fermat's last theorem had already been proven in the special case $n \in \{1, \ldots, 4,000,000\}$). The story behind Fermat's theorems is that he had a journal, and would write interesting conjectures that he thought of on the side notes, and then he wrote on the side notes that his proof was "a truly marvelous proof, which this margin is too narrow to contain." In this case, the general consensus is that Fermat didn't have a correct proof of his last theorem, since it's actual solution is so incredibly complicated (See (Wiles, 1995)). One might wonder why there are ε's floating around. It turns out it is necessary. First, we need Euler's theorem:

Theorem 13 (Euler). *Let $a, b \in \mathbb{Z}_{>0}$. If $\gcd(a, b) = 1$, then $a^{\varphi(b)} \equiv 1 \mod b$.*

One can find a proof of this theorem in any algebra textbook, as it can be easily proven with something called Langrange's theorem. $\varphi(b)$ is the order of the multiplicative group $(\mathbb{Z}/b)^\times$, and then it would follow immediately, since for any element $a \in G$, $|a| \mid |G|$. There are more elementary proofs, but this is a paper on algebra, so this is the one I will mention. This one is also really nice and short. All we need for our purposes is that $\varphi(p^n) = p^{n-1}(p-1)$ (it can be shown that φ is multiplicative, and how it acts on primes. $\varphi(b)$ counts the number of integers that are relatively prime with b and in $[1, b)$. For example, $\varphi(4) = 2$, since 1 and 3 are relatively prime to 4, but not 3. The idea here is that p^{n-1} aren't relatively prime with p^n, since given any list of l consecutive numbers, l/p of them are divisible by p (Every pth number in the list isn't divisible by p). Anyways, one notes now that $\varphi(2^n) = 2^{n-1}(2-1) = 2^{n-1}$, and as $\gcd(3, 2^n) = 1$, we see that $3^{2^{n-1}} \equiv 1 \mod 2^n$. Multiplying this by itself gives $3^{2^n} \equiv 1 \mod 2^n$, i.e. that $2^n \mid 3^{2^n} - 1$. Taking $a = 3^{2^n}$, $b = -1$, and $c = a + b$, we see that $|a| = 3^{2^n}$, and $n_0(abc) = n_0(3^{2^n} \cdot (3^{2^n} - 1)) \leq 3n_0\left(\frac{3^{2^n}-1}{2^{n-1}}\right) \leq \frac{3}{2^{n-1}} \cdot \left(3^{2^n} - 1\right) \leq \frac{3}{2^{n-1}} \cdot 3^{2^n}$. Then for any $C > 0$, $Cn_0(abc) < |a|$, since $Cn_0(abc) \leq \frac{3C}{2^{n-1}} \cdot 3^{2^n}$. We can now find n sufficiently large so that $\frac{3C}{2^{n-1}} < 1$, which shows that the ε is necessary. A paper by M. Waldschmidt (Waldschmidt, 2015) used this example to derive the following lemma:

Lemma 2. *There exists infinitely many triples (a, b, c) of positive integers where $a + b = c$ and a, b, c are coprime so that*

$$c > \frac{1}{6 \log 3} R \log R$$

where $R = n_0(abc)$.

He proved this using the counterexample to the ABC conjecture that I showed above. However, he took a different approach to showing that $2^{k+2} \mid 3^{2^k} - 1$, instead of using Euler's theorem, he used induction on k. I'm not entirely sure if you can use Euler's theorem for the stronger claim that 2^{k+2} divides $3^{2^k} - 1$, since $\varphi(2^{k+2}) = 2^{k+1}$. Then

$$1 \equiv 3^{2^{k+1}} \equiv \left(3^{2^k}\right)^2 \mod 3$$

Then $3^{2^k} \equiv 1 \mod 3$ or $3^{2^k} \equiv -1 \mod 3$. From here I do not think you can derive that it is in fact the first case, so I believe doing it by induction is actually necessary. Of course, this doesn't actually matter in the lemma, since you could say independent of which case it is you could use the same n_0 argument that I used above. The proof of the lemma is quite complicated, so I won't be going into it here, but the paper I referenced references another paper with said proof. I will conclude this paper with a discussion of the history of the ABC conjecture and it's attempted proof. The ABC conjecture arose from studying the Mason Stothers conjecture and its consequences, which is opposite to what I thought was true. I thought people conjectured a bunch of things about the integers, and then added some structure in to see if the proofs were easier–which they obviously were (ABC still remains unproven, for one thing). But that is not the case. The ABC conjecture was conjectured in 1985–4 years after Mason-Stothers was proven by Stothers in 1891. The paper I referenced above is 26 pages of applications of the ABC conjecture, most of which are still unsolved. The ABC conjecture is very powerful, but at the same time believed to be unprovable with current mathematical tools, similar to Fermat's last theorem (which was proven! Maybe ABC can be proven...) In 2012, Shinichi Mochizuki, a Japanese mathematician and professor at Kyoto University, published 4 papers detailing the development of "Inter-universal Teichmüller theory", which, in those papers, was used to prove ABC. My instructor, near the beginning of the year, told us about him, and joked that "Of the 12 people who can understand half of what he's saying, 6 believe the proof." This is actually true, since an error was found in the third paper, which mind you was probably 400 pages in. Mochizuki, and a few of his supporters claim that the "error" discovered wasn't actually an error–it

was just the reader not understanding the paper. Since, once again, 12 people can read this paper, at least one of which doesn't understand it fully, there isn't really a way to know who's right in this situation. However, I think it's pretty likely that ABC wasn't actually proven, since apparently Mochizuki resorted to personal attacks (Ad Hominem!) instead of justifying why his proof was correct (I talked briefly about this above). Here is an interesting fact: Fields medalist Richard Borcherds says that he is not qualified to answer if Mochizuki has proven ABC in this video. Richard Borcherds also has a lot of videos on all undergrad/grad algebra topics, like algebraic geometry, etc. so if you are interested in those subjects you should go check him out. It is amazing that one can watch lectures by a fields medalist online for free. On a happier note, Mochizuki's website is pretty awesome looking: see this. I personally am a big fan of the blue sky in the background, and all the nice colors. This concludes my discussion on Mason-Stothers and the ABC conjecture. I learned a lot, and I hope the reader enjoyed this paper.

References

Dummit, D., & Foote, R. (2003). *Abstract algebra*. Wiley. Retrieved from https://books.google.com/books?id=KJDBQgAACAAJ

Hungerford, T. (2012). *Abstract algebra: An introduction*. Cengage Learning. Retrieved from https://books.google.com/books?id=zRkLAAAAQBAJ

Lang, S. (2005). *Algebra*. Springer New York. Retrieved from https://books.google.com/books?id=Fge-BwqhqIYC

STOTHERS, W. W. (1981, 09). POLYNOMIAL IDENTITIES AND HAUPTMODULN. *The Quarterly Journal of Mathematics*, *32*(3), 349-370. Retrieved from https://doi.org/10.1093/qmath/32.3.349 doi: 10.1093/qmath/32.3.349

Synder, N. (2000). An alternate proof of mason's theorem. Retrieved from https://ems.press/journals/em/articles/609

Waldschmidt, M. (2015). Lecture on the abc conjecture and some of its consequences..

Wikipedia contributors. (2023). *Liouville's theorem (differential algebra) — Wikipedia, the free encyclopedia.* Retrieved from https://en.wikipedia.org/w/index.php?title=Liouville%27s_theorem_(differential_algebra)&oldid=1147605007 ([Online; accessed 1-May-2023])

Wiles, A. (1995). Modular elliptic curves and fermat's last theorem. *Annals of Mathematics*, *141*(3), 443–551. Retrieved 2023-04-30, from http://www.jstor.org/stable/2118559