STRASSMANN'S THEOREM AND AN APPLICATION

KEITH CONRAD

1. Introduction

Let $\{a_m\}$ be the sequence defined by the linear recursion

$$a_m = 2a_{m-1} - 3a_{m-2} \tag{1.1}$$

with initial conditions $a_0 = 1$, $a_1 = 1$. Here are the values of a_m for $m = 0, 1, \ldots, 14$.

m	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
a_m	1	1	−1	−5	−7	1	23	43	17	−95	−241	−197	329	1249	1511

One feature suggested by the data is that a_m is always odd. It is easy to prove this by induction from the fact that a_0 and a_1 are both odd, since the recursion reduced mod 2 shows $a_m \equiv a_{m-2} \bmod 2$.

The data also suggest that $|a_m| \to \infty$ as $m \to \infty$, and (seeing how $|a_m|$ starts growing) $a_m = \pm 1$ only for the times we see it happening in the table: for $m = 0, 1, 2$, and 5. This all turns out to be true, and while it sounds like a problem in real analysis, it will explained by p-adic analysis!

A natural way to study a_m is with an explicit formula for the sequence. Using complex numbers, such a formula is

$$a_m = \frac{(1+\sqrt{-2})^m}{2} + \frac{(1-\sqrt{-2})^m}{2}. \tag{1.2}$$

(To verify this formula, check the right side satisfies the recursion (1.1) and has value 1 at $m = 0$ and 1.) This shows the integer a_m is the real part of the complex number $(1+\sqrt{-2})^m$, and that is the context in which the equation $a_m = \pm 1$ first came to my attention [1]. Determining when $a_m = \pm 1$ is equivalent to finding all integers x such that $1 + 2x^2$ is a power of 3; see Appendix A for that, which shows understanding the values of $\{a_m\}$ has applications to number theory.

In $\mathbf{C}$ we have $|1 \pm \sqrt{-2}| = \sqrt{1+2} = \sqrt{3} > 1$, so the absolute value of both terms in (1.2) tends to ∞ with m. This is not sufficient to conclude $|a_m| \to \infty$ as $m \to \infty$ because the two terms in (1.2) have the same magnitude. We need to rule out the possibility of a massive cancellation for some large m that makes a_m small.

Let's write the condition "$|a_m| \to \infty$ as $m \to \infty$" in another way: since each a_m is an integer, saying $|a_m|$ tends to ∞ as $m \to \infty$ is equivalent to saying for each $c \in \mathbf{Z}$ that the equation $a_m = c$ is satisfied for only finitely many m. Here is our goal.

Theorem 1.1. *For each integer c, the equation $a_m = c$ holds for only finitely many integers m. In particular, $a_m = 1$ if and only if $m = 0, 1,$ or 5 and $a_m = -1$ if and only if $m = 2$.*

To make progress on Theorem 1.1, the key idea is to interpret (1.2) not in $\mathbf{C}$, but in some $\mathbf{Q}_p$ containing a square root of -2. Using the right side of (1.2) in $\mathbf{Q}_p$ we will see how to extend a_m from being a function of the integral parameter m to being a *locally p-adic analytic* function of m: there are finitely many p-adic power series, for a suitable prime p, whose values at the nonnegative integers m are the sequence $\{a_m\}$. This will let us think about the equation $a_m = c$ as a special case of the equation $f(x) = c$ where f is one of finitely many p-adic power series and $x \in \mathbf{Z}_p$. We will prove qualitative and quantitative theorems about zeros of p-adic power series that will tell us each equation $f(x) = c$ has a finite number of solutions in $\mathbf{Z}_p$ and at most how many such solutions there can be. If the upper bound on the number of solutions in $\mathbf{Z}_p$ is accounted for by the known $m \geq 0$ for which $a_m = c$, we will have provably found all $m \geq 0$ for which $a_m = c$.

2. Zeros of a p-adic analytic function

Theorem 2.1. *Let $f(x)$ be a power series with coefficients in $\mathbf{Q}_p$ that converges on $\mathbf{Z}_p$ and is not identically zero. The zeros of f in $\mathbf{Z}_p$ are isolated: for each $\alpha \in \mathbf{Z}_p$ at which $f(\alpha) = 0$ there is an $r > 0$ such that $f(x) \neq 0$ for $0 < |x - \alpha|_p < r$.*

This theorem is analogous to a property of real power series: each real zero of a real power series has an open interval around it in which there are no other real zeros.

Proof. We can recenter the power series at α: $f(x) = \sum_{n\geq 0} a_n(x-\alpha)^n$ on $\mathbf{Z}_p$ with $a_0 = f(\alpha) = 0$. Some a_n is not 0, since otherwise f would be identically zero on $\mathbf{Z}_p$. Let $a_N \neq 0$ with $N \geq 1$ minimal, so $f(x) = \sum_{n\geq N} a_n(x-\alpha)^n = (x-\alpha)^N g(x)$, where $g(x) = \sum_{n\geq N} a_n(x-\alpha)^{n-N}$. The power series $g(x)$ converges for each $x \in \mathbf{Z}_p$: this is obvious at $x = \alpha$, and for $x \neq \alpha$ in $\mathbf{Z}_p$ we have $|a_n(x-\alpha)^n|_p \to 0$ when $n \to \infty$, so $|a_n(x-\alpha)^{n-N}|_p = |a_n(x-\alpha)^n|_p/|x-\alpha|_p^N \to 0$ when $n \to \infty$.

Although g was constructed as a power series centered at α, since $0 \in \mathbf{Z}_p$ we can recenter g at 0 and the new series still converges on $\mathbf{Z}_p$. Since $g(\alpha) = a_N$ and a power series on $\mathbf{Z}_p$ is continuous, $\lim_{x\to\alpha} g(x) = a_N \neq 0$. Therefore there is a small $r > 0$ such that $|x-\alpha|_p < r \Longrightarrow g(x) \neq 0$. Then $0 < |x-\alpha|_p < r \Longrightarrow f(x) = (x-\alpha)^N g(x) \neq 0$. □

Corollary 2.2. *For a sequence $c_n \in \mathbf{Q}_p$ such that the series $f(x) = \sum_{n\geq 0} c_n x^n$ converges on $\mathbf{Z}_p$, if the coefficients are not all zero then f has only finitely many zeros in $\mathbf{Z}_p$.*

Proof. We will prove the contrapositive. Suppose f has infinitely many zeros $x_1, x_2, \ldots$ in $\mathbf{Z}_p$. Since $\mathbf{Z}_p$ is compact, this sequence has a convergent subsequence, say $x_{n_i} \to x \in \mathbf{Z}_p$. Then $f(x) = \lim_{i\to\infty} f(x_{n_i}) = \lim_{i\to\infty} 0 = 0$, and the zero x is not isolated since it is a limit of the zeros x_{n_i}. Theorem 2.1 implies f is identically 0, so all of its coefficients are 0. □

3. Turning $\{a_m\}$ into the values of a p-adic power series

In the formula (1.2) we would like to extend integer powers $(1+\sqrt{-2})^m$ and $(1-\sqrt{-2})^m$ to p-adic integer powers $(1+\sqrt{-2})^x$ and $(1-\sqrt{-2})^x$, where $x \in \mathbf{Z}_p$. This can't be done directly, because there is a restriction on the base b to be sure a power sequence $\{b^m\}$ extends to a p-adic power function b^x that is a p-adic power series in x: we want

$$|b-1|_p \leq \begin{cases} 1/p, & \text{if } p \neq 2, \\ 1/4, & \text{if } p = 2. \end{cases} \tag{3.1}$$

Under this condition, b^x has a power series representation

$$b^x = e^{x\log b} = \sum_{n\geq 0} \frac{(x\log b)^n}{n!} = \sum_{n\geq 0} \frac{(\log b)^n}{n!} x^n$$

that converges for all $x \in \mathbf{Z}_p$ since $|(\log b)^n/n!|_p = |b-1|_p^n/|n!|_p \to 0$ as $n \to \infty$.

Even if $\mathbf{Z}_p$ contains a square root of -2, $1+\sqrt{-2}$ and $1-\sqrt{-2}$ can't both satisfy (3.1) in the role of b.

Example 3.1. In $\mathbf{Z}_3$ there is a square root of -2 since $-2 \equiv 1 \bmod 3$. Explicitly, we can take

$$\sqrt{-2} = 1 + 3 + 2\cdot 3^2 + 2\cdot 3^5 + \cdots,$$

so

$$1+\sqrt{-2} = 2 + 3 + 2\cdot 3^2 + \cdots, \quad 1-\sqrt{-2} = 2\cdot 3 + 2\cdot 3^3 + \cdots.$$

Neither $1+\sqrt{-2}$ nor $1-\sqrt{-2}$ is in $1+3\mathbf{Z}_3$: one is in $2+3\mathbf{Z}_3$ and the other is in $3\mathbf{Z}_3$.

Example 3.2. In $\mathbf{Z}_{11}$ there is a square root of -2 since $-2 \equiv 9 \bmod 11$. We can take $\sqrt{-2} \equiv 3 \bmod 11$, so $1+\sqrt{-2} \equiv 4 \bmod 11$ and $1-\sqrt{-2} \equiv -2 \equiv 9 \bmod 11$. More explicitly,

$$\sqrt{-2} = 3 + 9\cdot 11 + 4\cdot 11^2 + \cdots,$$

so

$$1+\sqrt{-2} = 4 + 9\cdot 11 + 4\cdot 11^2 + \cdots, \quad 1-\sqrt{-2} = 9 + 11 + 6\cdot 11^2 + \cdots.$$

Both $1+\sqrt{-2}$ and $1-\sqrt{-2}$ are in $\mathbf{Z}_{11}^\times$, but neither is in $1+11\mathbf{Z}_{11}$.

Unless a p-adic integer b is p-adically close to 1, the power sequence $\{b^m\}$ is not the values at $0,1,2,3,\ldots$ of a p-adic power series. However, if $b \in \mathbf{Z}_p^\times$ then the sequence $\{b^m\}$ is the values at nonnegative integers of a *finite number* of p-adic power series.

Theorem 3.3. *Let $b \in \mathbf{Z}_p^\times$. If $p \neq 2$ then for each $r \in \{0,1,\ldots,p-2\}$ there are power series $f_r(x)$ converging on $\mathbf{Z}_p$ such that $f_r(k) = b^{(p-1)k+r}$ for all integers $k \geq 0$. If $p = 2$ then for $r = 0$ and 1 there are 2-adic power series $f_r(x)$ converging on $\mathbf{Z}_2$ such that $f_r(k) = b^{2k+r}$ for all integers $k \geq 0$.*

Proof. For $0 \leq r \leq p-2$ and $k \geq 0$,

$$b^{(p-1)k+r} = b^r(b^{p-1})^k.$$

Since $b \not\equiv 0 \bmod p$, by Fermat's little theorem $b^{p-1} \equiv 1 \bmod p$. Thus $|b^{p-1}-1|_p \leq 1/p$, so when $p \neq 2$ we can extend integer powers of b^{p-1} to p-adic integer powers: for $0 \leq r \leq p-2$ define the power series

$$f_r(x) = b^r(b^{p-1})^x = b^r e^{x\log(b^{p-1})} = b^r \sum_{n\geq 0} \frac{(\log b^{p-1})^n}{n!}x^n.$$

(Do *not* rewrite $\log b^{p-1}$ as $(p-1)\log b$ if $b \not\equiv 1 \bmod p$ since otherwise b is not in the domain of convergence of the p-adic logarithm series.) Each power series f_r converges on $\mathbf{Z}_p$ since its coefficients tend to 0, and for nonnegative integers k we have

$$f_r(k) = b^r(b^{p-1})^k = b^{(p-1)k+r}.$$

For $p = 2$ we have $b \equiv 1 \bmod 2 \implies b^2 \equiv 1 \bmod 4$, so $|b^2-1|_2 \leq 1/4$. (In fact, $|b^2-1|_2 \leq 1/8$.) Therefore we can take 2-adic integer powers of b^2 and define for $r=0$ and 1 the power series

$$f_r(x) = b^r(b^2)^x = b^r e^{x\log(b^2)} = b^r \sum_{n\geq 0} \frac{(\log b^2)^n}{n!}x^n.$$

This power series converges on $\mathbf{Z}_2$, and for integers $k \geq 0$ we have

$$f_r(k) = b^r(b^2)^k = b^{2k+r}.$$

□

We used b^{p-1} for $p \neq 2$ and b^2 for $p = 2$ to have a power of b that we know is congruent to 1 mod p (or 1 mod 4, if $p = 2$). This led to $p-1$ power series for $p \neq 2$ (or 2 power series if $p = 2$) whose values on $\mathbf{Z}_p$ include all values of b^m. If a smaller power of b is congruent to 1 mod p then we can use fewer power series in Theorem 3.3.

Example 3.4. For $b \in \mathbf{Z}_7^\times$, we have $b^6 \equiv 1 \bmod 7$ and Theorem 3.3 says for $r = 0, 1, \ldots, 5$ that there are 7-adic power series $f_r(x)$ converging on $\mathbf{Z}_7$ such that $f_r(k) = b^{6k+r}$ for integers $k \geq 0$.

If $b \equiv 2 \bmod 7$ then $b^3 \equiv 1 \bmod 7$, so we can take 7-adic integer powers of b^3, not just b^6. The sequence $\{b^m\}$ lies among the values of just three 7-adic power series: for $0 \leq r \leq 2$ set $f_r(x) = b^r(b^3)^x = b^r \sum_{n\geq 0}((\log b^3)^n/n!)x^n$. These series converge on $\mathbf{Z}_7$ and $f_r(k) = b^{3k+r}$ for integers $k \geq 0$.

Example 3.5. If $b \equiv 1 \bmod p$ for $p \neq 2$ or $b \equiv 1 \bmod 4$ then we only need a single p-adic power series to include all nonnegative integral powers of b: $f(x) = b^x = \sum_{n\geq 0}((\log b)^n/n!)x^n$ is a power series converging on $\mathbf{Z}_p$ and $f(k) = b^k$ for integers $k \geq 0$.

Example 3.6. Why do we require $|b|_p = 1$ in Theorem 3.3? If $|b|_p < 1$ and $b \neq 0$ then Theorem 3.3 breaks down: for no arithmetic progression $\{Mk + r\}_{k\geq 0}$, where $M \geq 1$ and $r \in \{0, \ldots, M-1\}$, can $b^{Mk+r} = f(k)$ for a p-adic power series $f(x)$. Indeed, since p-adic power series are continuous, $f(p^t) \to f(0)$ as $t \to \infty$ while $b^{Mp^t+r} \to 0$ as $t \to \infty$ since $|b^{Mp^t+r}|_p = |b^r|_p|b|_p^{Mp^t} \leq |b|_p^{p^t} \to 0$. Therefore we need $f(0) = 0$, so $b^r = 0$, which is false.

The underlying problem here is that every p-adic integer is the p-adic limit of integers that are large in the ordinary sense, and when $|b|_p < 1$ the number b^m has to be very small when m is very large in the ordinary sense. If $|b|_p = 1$ then at least $|b^m|_p = 1$ all the time.

Corollary 3.7. *For b_1 and b_2 in $\mathbf{Z}_p^\times$ and c_1 and c_2 in $\mathbf{Z}_p$, the numbers $c_1b_1^m + c_2b_2^m$ for integers $m \geq 0$ are the values of finitely many p-adic power series at nonnegative integers.*

Proof. Assume $p \neq 2$. Then $b_1^{p-1} \equiv 1 \bmod p$ and $b_2^{p-1} \equiv 1 \bmod p$. For $0 \leq r \leq p-2$ and $x \in \mathbf{Z}_p$ set

$$\begin{aligned} f_r(x) &= c_1b_1^r(b_1^{p-1})^x + c_2b_2^r(b_2^{p-1})^x \\ &= \sum_{n\geq 0} \frac{c_1b_1^r(\log b_1^{p-1})^n + c_2b_2^r(\log b_2^{p-1})^n}{n!}x^n. \end{aligned}$$

This power series converges on $\mathbf{Z}_p$, and for integers $k \geq 0$

$$f_r(k) = c_1b_1^r(b_1^{p-1})^k + c_2b_2^r(b_2^{p-1})^k = c_1b_1^{(p-1)k+r} + c_2b_2^{(p-1)k+r}.$$

If $p = 2$ then $b_1^2 \equiv 1 \bmod 4$ and $b_2^2 \equiv 1 \bmod 4$, so for $r = 0$ or 1 and $x \in \mathbf{Z}_2$, define

$$\begin{aligned} f_r(x) &= c_1b_1^r(b_1^2)^x + c_2b_2^r(b_2^2)^x \\ &= \sum_{n\geq 0} \frac{c_1b_1^r(\log b_1^2)^n + c_2b_2^r(\log b_2^2)^n}{n!}x^n. \end{aligned}$$

This series converges on $\mathbf{Z}_2$, and for integers $k \geq 0$

$$f_r(k) = c_1b_1^r(b_1^2)^k + c_2b_2^r(b_2^2)^k = c_1b_1^{2k+r} + c_2b_2^{2k+r}.$$

□

Corollary 3.7 extends to a linear combination of the powers of more than two p-adic units. We stick to two units for concreteness, as it will be sufficient for our intended application.

For $p \neq 2$, if $b_1^M \equiv 1 \bmod p$ and $b_2^M \equiv 1 \bmod p$ for some $M < p-1$ then the sequence $\{c_1 b_1^m + c_2 b_2^m\}_{m \geq 0}$ lies among the values of M power series in Corollary 3.7 instead of $p-1$ power series. In particular, if $b_1 \equiv 1 \bmod p$ and $b_2 \equiv 1 \bmod p$ then the sequence $\{c_1 b_1^m + c_2 b_2^m\}_{m \geq 0}$ lies among the values of a single power series converging on $\mathbf{Z}_p$.

Example 3.8. In $\mathbf{Z}_7$, if $b_1 \equiv 2 \bmod 7$ and $b_2 \equiv 4 \bmod 7$ then $b_1^3 \equiv 1 \bmod 7$ and $b_2^3 \equiv 1 \bmod 7$, so the numbers $c_1 b_1^m + c_2 b_2^m$ for $m \geq 0$ can be broken up into three sequences

$$c_1 b_1^r (b_1^3)^k + c_2 b_2^r (b_2^3)^k$$

for $r = 0, 1, 2$ and $k \geq 0$, which each extend to a 7-adic power series converging on $\mathbf{Z}_7$:

$$c_1 b_1^r (b_1^3)^x + c_2 b_2^r (b_2^3)^x = \sum_{n \geq 0} \frac{c_1 b_1^r (\log b_1^3)^n + c_2 b_2^r (\log b_2^3)^n}{n!} x^n.$$

Let's return to our original sequence of interest a_m in (1.1), which has an explicit formula in terms of powers of $1+\sqrt{-2}$ and $1-\sqrt{-2}$ in (1.2). Although there are square roots of -2 in $\mathbf{Z}_3$, one lying in $1+3\mathbf{Z}_3$ and one lying in $2+3\mathbf{Z}_3$, there is not a formula for a_m using 3-adic power series: taking $\sqrt{-2} \equiv 1 \bmod 3$, there are problems with powers of $1-\sqrt{-2}$ since $|1-\sqrt{-2}|_3 < 1$, and if we had chosen $\sqrt{-2} \equiv 2 \bmod 3$ then we'd have problems with powers of $1+\sqrt{-2}$ for a similar reason.

The next prime after $p = 3$ where -2 has square roots in $\mathbf{Z}_p$ is 11, so let's turn it up to 11. We saw in Example 3.2 that we can choose $\sqrt{-2} \equiv 3 \bmod 11$, so $1+\sqrt{-2} \equiv 4 \bmod 11$ and $1-\sqrt{-2} \equiv -2 \equiv 9 \bmod 11$. Since $4^5 \equiv 1 \bmod 11$ and $9^5 \equiv 1 \bmod 11$, both $(1+\sqrt{-2})^5$ and $(1-\sqrt{-2})^5$ lie in $1+11\mathbf{Z}_{11}$. Explicitly,

$$(1+\sqrt{-2})^5 = 1 - 11\sqrt{-2}, \quad (1-\sqrt{-2})^5 = 1 + 11\sqrt{-2}.$$

For $r = 0, 1, 2, 3, 4$, and $x \in \mathbf{Z}_{11}$, define

$$\begin{aligned} f_r(x) &= \frac{(1+\sqrt{-2})^r((1+\sqrt{-2})^5)^x + (1-\sqrt{-2})^r((1-\sqrt{-2})^5)^x}{2} \\ &= \frac{(1+\sqrt{-2})^r}{2}(1-11\sqrt{-2})^x + \frac{(1-\sqrt{-2})^r}{2}(1+11\sqrt{-2})^x \\ &= \sum_{n \geq 0} \left(\frac{(1+\sqrt{-2})^r}{2} \frac{(\log(1-11\sqrt{-2}))^n}{n!} + \frac{(1-\sqrt{-2})^r}{2} \frac{(\log(1+11\sqrt{-2}))^n}{n!} \right) x^n. \end{aligned} \tag{3.2}$$

For integers $k \geq 0$,

$$f_r(k) = \frac{(1+\sqrt{-2})^{5k+r}}{2} + \frac{(1-\sqrt{-2})^{5k+r}}{2} = a_{5k+r}.$$

This way of looking at the sequence $\{a_m\}$, as the values at nonnegative integers of five 11-adic power series, leads to a solution of the qualitative problem about values of a_m.

Theorem 3.9. *The sequence $\{a_m\}$ in (1.2) with initial conditions $a_0 = a_1 = 1$ has $|a_m| \to \infty$ as $m \to \infty$.*

Proof. We will show for each $c \in \mathbf{Z}$ that the equation $a_m = c$ is satisfied for only finitely many integers $m \geq 0$ by showing a more general property in the 11-adic integers: for each $c \in \mathbf{Z}_{11}$ and $r \in \{0, 1, \ldots, 4\}$ the equations $f_r(x) = c$, where f_r is defined by (3.2), each have only finitely many solutions x in $\mathbf{Z}_{11}$. To prove that, we will show each f_r is a nonconstant power series, since that makes the power series $f_r(x)-c$ nonconstant and thus it has finitely many zeros in $\mathbf{Z}_{11}$ by Corollary 2.2.

To check each of the five power series f_r in (3.2) is not constant, we could compute the linear coefficient of f_r and check it is not 0 (and if it were 0, we could then check the quadratic coefficient is not 0, and so on). But we will do something simpler: compare $f_r(0) = a_r$ and $f_r(1) = a_{5+r}$ for $0 \leq r \leq 4$. If they are not equal then f_r is not a constant series. We already saw these values in the table at the start of Section 1. Here they are again, in a more suitable form for us now.

r	0	1	2	3	4
a_r	1	1	-1	-5	-7
a_{5+r}	1	23	43	17	-95

We see $a_r \neq a_{5+r}$ when r is 1, 2, 3, and 4, so f_r is not constant, but at $r = 0$ we have $a_0 = a_5 = 1$. That is not a problem: just compute one more value: $f_0(2) = a_{5\cdot 2} = a_{10} = -241$. So f_0 is not constant either. ☐

To bound how often $a_m = \pm 1$, we will bound how often $f_r(x) = 1$ and $f_r(x) = -1$ in $\mathbf{Z}_{11}$ for $0 \leq r \leq 4$ in (3.2). This is equivalent to bounding the number of 11-adic integer zeros of $f_r(x) - 1$ and $f_r(x) + 1$, which can be thought of as a quantitative refinement of Corollary 2.2. To do this we will use a theorem from p-adic analysis called Strassmann's theorem.

4. Strassmann's theorem

By Corollary 2.2, a nonzero series $f(x) = \sum_{n\geq 0} a_n x^n$ with $a_n \in \mathbf{Q}_p$ that converges on $\mathbf{Z}_p$ has finitely many zeros in $\mathbf{Z}_p$. We want to bound the number of those zeros. The series $\sum_{n\geq 0} a_n x^n$ converges on $\mathbf{Z}_p$ if and only if $a_n \to 0$. If $a_n \to 0$ and the a_n's are not all 0 then the numbers $|a_n|_p$ have a positive maximum and there is a last time the maximum occurs. The largest index for a coefficient of maximal absolute value is denoted $N(f)$. That is,

$$N(f) = \max\{N \geq 0 : |a_n|_p \leq |a_N|_p \text{ for all } n \geq 0\}.$$

For the power series f whose coefficients are all 0, $N(f)$ is not defined.

Theorem 4.1 (Strassmann)**.** *Let $f(x) = \sum_{n\geq 0} a_n x^n$ where $a_n \in \mathbf{Q}_p$ and $a_n \to 0$. If the a_n's are not all zero then the number of solutions to $f(x) = 0$ in $\mathbf{Z}_p$ is at most $N(f)$.*

We can apply this theorem to polynomials, which are power series with finitely many terms.

Example 4.2. Over $\mathbf{Q}_p$, $f(X) = 1+pX+X^2+pX^5$ has $N = 2$, so it has at most 2 zeros in $\mathbf{Z}_p$. The actual number of zeros of $f(X)$ in $\mathbf{Z}_p$ is 0 when $p = 2$ ($a \in \mathbf{Z}_2 \Rightarrow f(a) \equiv 1, 2 \bmod 4$) and $p = 3$ ($a \in \mathbf{Z}_3 \Rightarrow f(a) \equiv 1, 2 \bmod 3$) and 2 when $p = 5$ (use Hensel's lemma for $f(X)$ with $a = 2$ and $a = 3$).

Example 4.3. Over $\mathbf{Q}_p$, $1 + X + pX^2$ has $N = 1$ and thus at most 1 zero in $\mathbf{Z}_p$. In fact there is a zero in $\mathbf{Z}_p$, as you can check with the quadratic formula; a second zero is in $\mathbf{Q}_p$ but outside of $\mathbf{Z}_p$.

Example 4.4. Over $\mathbf{Q}_p$, $X^n - p$ has $N = n$ and no roots in $\mathbf{Z}_p$ (or $\mathbf{Q}_p$) for $n \geq 2$. This illustrates that the bound in Strassmann's theorem is only an upper bound on the number of roots in $\mathbf{Z}_p$, not a formula in general for the number of roots in $\mathbf{Z}_p$.

Strassmann's theorem can be regarded as an analogue for p-adic power series of bounding the number of roots of a polynomial over a field by the degree of the polynomial. In the polynomial theorem the key idea is to factor out $x-\alpha$ if α is a root, which lowers the degree of the polynomial by one, and the proof of Strassmann's theorem will have a step just like this where the value of $N(f)$ drops by one after removing a factor corresponding to a root (if one exists). When dealing with power series rather than polynomials we have to be a little more careful at the factoring step due to convergence issues.

Proof. We use induction on $N(f)$.

When $N(f) = 0$, $|a_n|_p < |a_0|_p$ for all $n \geq 1$, so $a_0 \neq 0$ and $\max_{n\geq 1} |a_n|_p < |a_0|_p$ because the a_n's tend to 0. For $x \in \mathbf{Z}_p$,

$$\left|\sum_{n\geq 1} a_n x^n\right|_p \leq \max_{n\geq 1} |a_n x^n|_p \leq \max_{n\geq 1} |a_n|_p < |a_0|_p,$$

so by the strong triangle inequality $|f(x)|_p = |a_0 + \sum_{n\geq 1} a_n x^n|_p = |a_0|_p > 0$. Thus f has no zero in $\mathbf{Z}_p$.

Now suppose $N \geq 1$ and the theorem is proved for all power series $g(x)$ with coefficients in $\mathbf{Q}_p$ converging on $\mathbf{Z}_p$ with $N(g) < N$. If $N(f) = N$ and f has no zeros in $\mathbf{Z}_p$ then we are done since $0 < N$. If f has a zero $\alpha \in \mathbf{Z}_p$ then by the same reasoning as in the proof of Theorem 2.1 we can write

$$f(x) = (x-\alpha)g(x) \tag{4.1}$$

where g is a power series centered at 0 that converges on $\mathbf{Z}_p$. By (4.1), for $x \in \mathbf{Z}_p$ we have $f(x) = 0$ if and only if $x = \alpha$ or $g(x) = 0$. We will show $N(g) = N(f) - 1 = N - 1$, so by induction g has at most $N-1$ zeros in $\mathbf{Z}_p$, and therefore the number of zeros of f in $\mathbf{Z}_p$ is at most $1 + (N-1) = N$.

Writing $g(x) = \sum_{n\geq 0} b_n x^n$, to show $N(g) = N-1$ means showing

$$|b_n|_p \leq |b_{N-1}|_p \text{ for all } n, \quad |b_n|_p < |b_{N-1}|_p \text{ for } n \geq N. \tag{4.2}$$

While doing this we will also show $|b_{N-1}|_p = |a_N|_p$.

If $\alpha = 0$ then $f(x) = xg(x)$, so $b_n = a_{n+1}$ for all n, and then (4.2) and $|b_{N-1}|_p = |a_N|_p$ are clear. If $\alpha \neq 0$, substituting the power series representations $f(x) = \sum_{n\geq 0} a_n x^n$ and $g(x) = \sum_{n\geq 0} b_n x^n$ into (4.1) and equating coefficients of like powers of x on both sides, we get

$$a_0 = -\alpha b_0, \quad a_n = b_{n-1} - b_n\alpha \text{ for } n \geq 1.$$

Replacing n by $n+1$ in this recursion,

$$\begin{aligned} b_n &= a_{n+1} + b_{n+1}\alpha \\ &= a_{n+1} + (a_{n+2} + b_{n+2}\alpha)\alpha \\ &= a_{n+1} + a_{n+2}\alpha + b_{n+2}\alpha^2 \\ &= a_{n+1} + a_{n+2}\alpha + (a_{n+3} + b_{n+3}\alpha)\alpha^2 \\ &= a_{n+1} + a_{n+2}\alpha + a_{n+3}\alpha^2 + b_{n+3}\alpha^3. \end{aligned}$$

Repeating this, for any $m \geq 1$

$$b_n = \sum_{k=1}^{m} a_{n+k}\alpha^{k-1} + b_{n+m}\alpha^{m-1}.$$

Since $\alpha \neq 0$, $|a_{n+k}\alpha^{k-1}|_p = |a_{n+k}\alpha_p^{n+k}|/|\alpha|_p^{n+1} \to 0$ as $k \to \infty$ since the power series for f centered at 0 converges at α, and similarly $|b_{n+m}\alpha^{m-1}|_p \to 0$ as $m \to \infty$ since the power series for g centered at 0 converges at α. Therefore

$$b_n = \sum_{k\geq 1} a_{n+k}\alpha^{k-1}, \tag{4.3}$$

so for all n

$$|b_n|_p \leq \max_{k\geq 1} |a_{n+k}|_p = \max_{k\geq n+1} |a_k|_p \leq |a_N|_p.$$

If $k \geq N+1$ then $|a_k|_p < |a_N|_p$ by the definition of N, so $n \geq N \Longrightarrow |b_n|_p < |a_N|_p$. Also $b_{N-1} = a_N + \sum_{k\geq 2} a_{N-1+k}\alpha^{k-1}$ where $|a_{N-1+k}\alpha^{k-1}|_p \leq |a_{N-1+k}|_p < |a_N|_p$ for $k \geq 2$, so $|b_{N-1}|_p = |a_N|_p$. Thus $|b_n|_p$ is maximized for the last time at $n = N-1$, so $N(g) = N-1$. $\square$

Remark 4.5. The number of roots of a polynomial over a field need not equal its degree, but equality does occur in degree 1: $ax+b=0$ if and only if $x = -b/a$ (if $a \neq 0$). Similarly, if $N(f) = 1$ in Strassmann's theorem then there really is a root of $f(x)$ in $\mathbf{Z}_p$. This can be proved using a version of Hensel's lemma for power series.

5. Proof of Theorem 1.1 using $\mathbf{Q}_{11}$

The formula for a_m in (1.2) uses a square root of -2. Since $-2 \equiv 9 \bmod 11$, -2 has a square root in $\mathbf{Z}_{11}$ that is congruent to $3 \bmod 11$. *Define* $\sqrt{-2}$ to be that 11-adic integer:

$$\sqrt{-2} = 3 + 9\cdot 11 + 4\cdot 11^2 + 11^3 + \cdots.$$

Step 1: Estimate values of the p-adic logarithm on $1 + p\mathbf{Z}_p$.

We will show for odd p (the case of interest is $p = 11$) and $y \in p\mathbf{Z}_p$ that $|\log(1+y)|_p = |y|_p$ and $\log(1+y) \equiv y \bmod p^2$.

Since $\log(1+y) = \sum_{n\geq 1}(-1)^{n-1}y^n/n$ it suffices, for both the desired equation and congruence, to check when $n \geq 2$ and $|y|_p \leq 1/p$ that $|y^n/n|_p < |y|_p$, or equivalently that $1/p < |n|_p^{1/(n-1)}$. This is clear if $|n|_p = 1$, and if $|n|_p < 1$ set $n = p^r m$ for $r \geq 1$ and $p \nmid m$. Then

$$|n|_p^{1/(n-1)} = \frac{1}{p^{r/(p^r m-1)}} > \frac{1}{p^{r/(p^r-1)}} \overset{?}{>} \frac{1}{p} \Longleftrightarrow 1 > \frac{r}{p^r-1} \Longleftrightarrow p^r - 1 \overset{\checkmark}{>} r \text{ (since } p > 2).$$

Step 2: Make the numbers a_m into values of several 11-adic power series.

We seek j such that $|(1+\sqrt{-2})^j - 1|_{11} \leq 1/11$ and $|(1-\sqrt{-2})^j - 1|_{11} \leq 1/11$. Use $j = 5$:

$$(1+\sqrt{-2})^5 = 1 - 11\sqrt{-2}, \quad (1-\sqrt{-2})^5 = 1 + 11\sqrt{-2}.$$

Therefore if we write $m = 5k + r$ where $k \geq 0$ and $0 \leq r \leq 4$, we have

$$\begin{aligned} a_{5k+r} &= \frac{(1+\sqrt{-2})^r}{2}((1+\sqrt{-2})^5)^k + \frac{(1-\sqrt{-2})^r}{2}((1-\sqrt{-2})^5)^k \\ &= \frac{(1+\sqrt{-2})^r}{2}(1-11\sqrt{-2})^k + \frac{(1-\sqrt{-2})^r}{2}(1+11\sqrt{-2})^k. \end{aligned}$$

This formula suggests looking at the 11-adic analytic functions

$$f_r(x) = \frac{(1+\sqrt{-2})^r}{2}(1-11\sqrt{-2})^x + \frac{(1-\sqrt{-2})^r}{2}(1+11\sqrt{-2})^x$$

where $0 \leq r \leq 4$ and $x \in \mathbf{Z}_{11}$. For integers $k \geq 0$,

$$\boxed{f_r(k) = a_{5k+r}.} \tag{5.1}$$

Do not forget this! In terms of the 11-adic exponential series,

$$\begin{aligned} f_r(x) &= \frac{(1+\sqrt{-2})^r}{2}e^{x\log(1-11\sqrt{-2})} + \frac{(1-\sqrt{-2})^r}{2}e^{x\log(1+11\sqrt{-2})} \\ &= \sum_{n\geq 0} c_{r,n}x^n, \end{aligned}$$

where $c_{r,n} = \dfrac{(1+\sqrt{-2})^r}{2}\dfrac{(\log(1-11\sqrt{-2}))^n}{n!} + \dfrac{(1-\sqrt{-2})^r}{2}\dfrac{(\log(1+11\sqrt{-2}))^n}{n!}$ in $\mathbf{Q}_{11}$. We have $|\log(1\pm 11\sqrt{-2})|_{11} = |11|_{11}$ by Step 1, so from $|11^n/n!|_{11} \leq 1$ we get $c_{r,n} \in \mathbf{Z}_{11}$.

Step 3: Estimate how quickly the coefficients of f_r tend to 0.

Theorem 5.1. *For $0 \leq r \leq 4$ and $n \geq 1$, $|c_{r,n}|_{11} \leq 1/11^{(9n+1)/10} \leq 1/11$. In particular, $|c_{r,n}|_{11} \leq 1/11$ for $n \geq 1$, $|c_{r,n}|_{11} \leq 1/11^2$ for $n \geq 2$, and $|c_{r,n}|_{11} \leq 1/11^3$ for $n \geq 3$.*

Proof. Since $(1+\sqrt{-2})^r/2$ and $(1-\sqrt{-2})^r/2$ are in $\mathbf{Z}_{11}^\times$,

$$\begin{aligned} |c_{r,n}|_{11} &\leq \max\left(\left|\frac{(\log(1-11\sqrt{-2}))^n}{n!}\right|_{11}, \left|\frac{(\log(1+11\sqrt{-2}))^n}{n!}\right|_{11}\right) \\ &= \max\left(\frac{|11|_{11}^n}{|n!|_{11}}, \frac{|11|_{11}^n}{|n!|_{11}}\right) \text{ by Step 1} \\ &= \frac{(1/11)^n}{(1/11)^{(n-s_{11}(n))/(11-1)}} \\ &= \frac{1}{11^{9n/10+s_{11}(n)/10}} \\ &\leq \frac{1}{11^{9n/10+1/10}} \text{ since } n \geq 1. \end{aligned}$$

For $n \geq 1$ we have $9n/10 + 1/10 \geq 1$, for $n \geq 2$ we have $9n/10 + 1/10 \geq 1.9$, and for $n \geq 3$ we have $9n/10 + 1/10 \geq 2.8$, Since $\mathrm{ord}_{11}(c_{n,r})$ is an integer (or ∞), if $\mathrm{ord}_{11}(c_{r,n}) \geq 1.9$ then $\mathrm{ord}_{11}(c_{r,n}) \geq 2$ and if $\mathrm{ord}_{11}(c_{r,n}) \geq 2.8$ then $\mathrm{ord}_{11}(c_{r,n}) \geq 3$. ☐

Step 4: Finishing the proof of Theorem 1.1.

We want to show $a_m = 1$ only when $m = 0$, 1, and 5, and $a_m = -1$ only when $m = 2$. The following table writes these m as $5k + r$: 1 arises twice when $r = 0$ (at $k = 0, 1$) and once when $r = 1$ (at $k = 0$), and -1 arises once when $r = 2$ (at $k = 0$).

$5k+r$	k	r	a_{5k+r}
0	0	0	1
1	0	1	1
2	0	2	-1
5	1	0	1

Since $a_{5k+r} = f_r(k)$, we want to show the only zeros of $f_r(x) - 1$ and $f_r(x) + 1$ in $\mathbf{Z}_{11}$ are as described in the following table, where k is replaced with the 11-adic integer variable x.

r	Zeros of $f_r(x) - 1$	Zeros of $f_r(x) + 1$
0	$x = 0, 1$	None
1	$x = 0$	None
2	None	$x = 0$
3	None	None
4	None	None

The indicated zeros for $f_0(x)-1$, $f_1(x)-1$, and $f_2(x)+1$ follow from (5.1). We will show for Strassmann's theorem that $f_0(x)-1$ has $N=2$, $f_1(x)-1$ and $f_2(x)+1$ have $N=1$, and other $f_r(x)\pm 1$ have $N=0$, so the upper bound on zeros is reached by the known zeros.

Adding and subtracting 1 to $f_r(x)$ affects the constant term but no other coefficients:

$$f_r(x)\pm 1 = (c_{r,0}\pm 1) + \sum_{n\geq 1} c_{r,n}x^n = (a_r \pm 1) + \sum_{n\geq 1} c_{r,n}x^n.$$

Let's first take care of the series where no zeros are expected.

Theorem 5.2. *The series $f_2(x)-1$, $f_3(x)-1$, $f_4(x)-1$, $f_0(x)+1$, $f_1(x)+1$, $f_3(x)+1$, and $f_4(x)+1$ all have no zeros in $\mathbf{Z}_{11}$.*

Proof. To prove an 11-adic power series has no zeros in $\mathbf{Z}_{11}$ with Strassmann's theorem, we want to show $N=0$: the constant term of $f_r(x)\pm 1$ has larger absolute value than every other coefficient. The table below lists the constant term $f_r(0)\pm 1 = a_r \pm 1$.

r	0	1	2	3	4
$f_r(0)$	1	1	-1	-5	-7
$f_r(0)-1$	0	0	-2	-6	-8
$f_r(0)+1$	2	2	0	-4	-6

Thus $f_2(x)-1$, $f_3(x)-1$, $f_4(x)-1$, $f_0(x)+1$, $f_1(x)+1$, $f_3(x)+1$, and $f_4(x)+1$ have constant terms in $\mathbf{Z}_{11}^\times$. The higher-degree coefficients are the same as those of $f_r(x)$, namely $c_{r,n}$ for $n\geq 1$. Those coefficients are in $11\mathbf{Z}_{11}$ by Theorem 5.1, so $f_2(x)-1$, $f_3(x)-1$, $f_4(x)-1$, $f_0(x)+1$, $f_1(x)+1$, $f_3(x)+1$, and $f_4(x)+1$ all have $N=0$. □

It remains to handle $f_0(x)-1$, $f_1(x)-1$, and $f_2(x)+1$.

Theorem 5.3. *The only zeros of $f_0(x)-1$ in $\mathbf{Z}_{11}$ are $x=0$ and $x=1$.*

Proof. The constant term of $f_0(x)-1$ is 0. For the linear and quadratic coefficients we will show $|c_{0,1}|_{11}=1/121$ and $|c_{0,2}|_{11}=1/121$. For $n\geq 3$, Theorem 5.1 tells us $|c_{0,n}|_{11}<1/121$, so $f_0(x)-1$ would have $N=2$ and that upper bound on the zeros in $\mathbf{Z}_{11}$ is already accounted for by the two zeros we know (corresponding to $a_0=1$ and $a_5=1$).

The linear coefficient of $f_0(x)-1$ is

$$c_{0,1} = \frac{1}{2}\log(1-11\sqrt{-2}) + \frac{1}{2}\log(1+11\sqrt{-2}) = \frac{1}{2}\log(1+2\cdot 11^2),$$

so $|c_{0,1}|_{11} = |2\cdot 11^2|_{11} = 1/121$. The quadratic coefficient of $f_0(x)-1$ is

$$\begin{aligned} c_{0,2} &= \frac{1}{2}\frac{(\log(1-11\sqrt{-2}))^2}{2} + \frac{1}{2}\frac{(\log(1+11\sqrt{-2}))^2}{2} \\ &= \frac{(\log(1-11\sqrt{-2}))^2 + (\log(1+11\sqrt{-2}))^2}{4} \\ &= \frac{1}{4}\left(\Big(\underbrace{\log(1-11\sqrt{-2})+\log(1+11\sqrt{-2})}_{\log((1-11\sqrt{-2})(1+11\sqrt{-2}))}\Big)^2 - 2\log(1-11\sqrt{-2})\log(1+11\sqrt{-2})\right). \end{aligned}$$

Since $(1-11\sqrt{-2})(1+11\sqrt{-2}) = 1+242$, the squared term has absolute value $|242|_{11}^2 = 1/11^4$, while by Step 1 $|\log(1-11\sqrt{-2})\log(1+11\sqrt{-2})|_{11} = (1/11)(1/11) = 1/11^2$, so by the strong triangle inequality $|c_{0,2}|_{11} = |1/4|_{11}(1/11^2) = 1/121$.

Another way to show $|c_{0,2}|_{11} = 1/121$ is to compute $c_{0,2} \bmod 11^3$. By Step 1, $\log(1 \pm 11\sqrt{-2}) \equiv \pm 11\sqrt{-2} \bmod 11^2$. Thus $\log(1 \pm 11\sqrt{-2}) = \pm 11\sqrt{-2} + 11^2 x_\pm$ with $x_\pm \in \mathbf{Z}_{11}$, so

$$(\log(1 \pm 11\sqrt{-2}))^2 = -2 \cdot 11^2 + 11^3(\text{11-adic integer}) \equiv -2 \cdot 11^2 \bmod 11^3$$

for both choices of sign. Therefore

$$c_{0,2} \equiv \frac{1}{4}(-2 \cdot 11^2) + \frac{1}{4}(-2 \cdot 11^2) \bmod 11^3 \equiv -11^2 \bmod 11^3,$$

so $|c_{0,2}|_{11} = 1/121$.

□

Theorem 5.4. *The only zero of $f_1(x) - 1$ in $\mathbf{Z}_{11}$ is $x = 0$.*

Proof. The constant term of $f_1(x) - 1$ is 0. We will prove $|c_{1,1}|_{11} = 1/11$. By Theorem 5.1, $|c_{1,n}|_{11} < 1/11$ for $n \geq 2$, so $f_1(x) - 1$ would have $N = 1$ and thus its known zero at $x = 0$ (corresponding to $a_1 = 1$) is its only zero in $\mathbf{Z}_{11}$.

The linear coefficient of $f_1(x) - 1$ is

$$c_{1,1} = \frac{1 + \sqrt{-2}}{2} \log(1 - 11\sqrt{-2}) + \frac{1 - \sqrt{-2}}{2} \log(1 + 11\sqrt{-2}).$$

Using the congruence mod p^2 in Step 1 at $p = 11$,

$$c_{1,1} \equiv \frac{1 + \sqrt{-2}}{2}(-11\sqrt{-2}) + \frac{1 - \sqrt{-2}}{2}(11\sqrt{-2}) \equiv 22 \bmod 11^2 \Longrightarrow |c_{1,1}|_{11} = \frac{1}{11}.$$

□

Theorem 5.5. *The only zero of $f_2(x) + 1$ in $\mathbf{Z}_{11}$ is $x = 0$.*

Proof. The constant term of $f_2(x) + 1$ is 0. We will prove $|c_{2,1}|_{11} = 1/11$, which suffices by the same reasoning as in the proof of the previous theorem. Since

$$\begin{aligned}
c_{2,1} &= \frac{(1 + \sqrt{-2})^2}{2} \log(1 - 11\sqrt{-2}) + \frac{(1 - \sqrt{-2})^2}{2} \log(1 + 11\sqrt{-2}) \\
&= \frac{-1 + 2\sqrt{-2}}{2} \log(1 - 11\sqrt{-2}) + \frac{-1 - 2\sqrt{-2}}{2} \log(1 + 11\sqrt{-2}) \\
&\equiv \frac{-1 + 2\sqrt{-2}}{2}(-11\sqrt{-2}) + \frac{-1 - 2\sqrt{-2}}{2}(11\sqrt{-2}) \bmod 11^2 \text{ by Step 1} \\
&\equiv 4 \cdot 11 \bmod 11^2,
\end{aligned}$$

we get $|c_{2,1}|_{11} = 1/11$. □

6. Further values of a_m

The method used to determine all $m \geq 0$ for which $a_m = \pm 1$ can be applied to other values in the sequence $\{a_m\}$. The values of a_m for $0 \leq m \leq 10$ besides ± 1 are

$$a_3 = -5, \quad a_4 = -7, \quad a_6 = 23, \quad a_7 = 43, \quad a_8 = 17, \quad a_9 = -95, \quad a_{10} = -241. \tag{6.1}$$

To prove these values occur exactly once in the sequence, let's write out what each $f_r(x)$ looks like. The constant term of $f_r(x)$ is $f_r(0) = a_r$, so

$$
\begin{aligned}
f_0(x) &= 1 + \sum_{n\geq 1} c_{0,n}x^n, \\
f_1(x) &= 1 + \sum_{n\geq 1} c_{1,n}x^n, \\
f_2(x) &= -1 + \sum_{n\geq 1} c_{2,n}x^n, \\
f_3(x) &= -5 + \sum_{n\geq 1} c_{3,n}x^n, \\
f_4(x) &= -7 + \sum_{n\geq 1} c_{4,n}x^n.
\end{aligned}
$$

We already showed in Theorems 5.3, 5.4, and 5.5 that $|c_{0,1}|_{11} = 1/121$, $|c_{1,1}|_{11} = 1/11$, and $|c_{2,1}|_{11} = 1/11$. It is left to the reader to check that $|c_{3,1}|_{11} = 1/11$ and $|c_{4,1}|_{11} = 1/11$. For $n \geq 2$, $|c_{r,n}|_{11} \leq 1/121$ by Theorem 5.1.

Theorem 6.1. *We have $a_m = -5$ if and only if $m = 3$.*

Proof. For $r = 0, 1, 2, 4$ the series $f_r(x) + 5$ has constant term in $\mathbf{Z}_{11}^\times$ and higher-degree coefficients in $11\mathbf{Z}_{11}$, so $N(f_r + 5) = 0$ and thus $a_{5k+r} \neq -5$ for all $k \geq 0$. What if $r = 3$? The series $f_3(x) + 5$ has constant term 0, linear coefficient of absolute value 1/11 and $|c_{3,n}|_{11} \leq 1/121$ for $n \geq 2$, so $N(f_3 + 5) = 1$ and thus the only solution to $f_3(x) + 5 = 0$ in $\mathbf{Z}_{11}$ is $x = 0$. That proves $a_m = -5$ only for $m = 5 \cdot 0 + 3 = 3$. □

Theorem 6.2. *We have $a_m = 23$ if and only if $m = 6$.*

Proof. For $r = 2, 3, 4$, the series $f_r(x) - 23$ has constant term in $\mathbf{Z}_{11}^\times$ and higher-degree coefficients in $11\mathbf{Z}_{11}$, so none of these series has a zero in $\mathbf{Z}_{11}$. Both $f_0(x) - 23$ and $f_1(x) - 23$ have constant term $-22 \in 11\mathbf{Z}_{11}$. Since $|-22|_{11} = 1/11$, $|c_{0,1}|_{11} = 1/121$, and $|c_{0,n}|_{11} \leq 1/121$ for $n \geq 2$, $N(f_0 - 23) = 0$ and thus $f_0(x) - 23$ is nonvanishing on $\mathbf{Z}_{11}$. Since $|-22|_{11} = 1/11$, $|c_{1,1}|_{11} = 1/11$, and $|c_{1,n}|_{11} \leq 1/121$ for $n \geq 2$, $N(f_1 - 23) = 1$ and thus the zero of $f_1(x) - 23$ at $x = 1$ (corresponding to $a_6 = 23$) is its only zero in $\mathbf{Z}_{11}$. □

It is left as an exercise to the reader to show the values of a_m in (6.1) at $m = 4$, 7, 8, and 9 each occur only once among all $m \geq 0$.

While $a_{10} = -241$, showing $a_m = -241$ only at $m = 10$ doesn't work using $\mathbf{Q}_{11}$ because something new happens: two of the series $f_r(x) + 241$ have a root in $\mathbf{Z}_{11}$ that is not a nonnegative integer, so the Strassmann bound is too big. The reader can check $f_r(x) + 241$ has $N = 0$ for $r = 2, 3, 4$. At $r = 0$ and 1 we have

$$
\begin{aligned}
f_0(x) + 241 &= 242 + \sum_{n\geq 1} c_{0,n}x^n, \\
f_1(x) + 241 &= 242 + \sum_{n\geq 1} c_{1,n}x^n,
\end{aligned}
$$

and $|242|_{11} = 1/121$. The linear and quadratic coefficients of $f_0(x) + 241$ also have absolute value 1/121 (see the proof of Theorem 5.3), while $|c_{0,n}|_{11} < 1/121$ for $n \geq 3$ (Theorem

5.1), so $N(f_0 + 241) = 2$. In $f_1(x) + 241$, $|c_{1,1}|_{11} = 1/11$ and $|c_{1,n}|_{11} < 1/11$ for $n \geq 2$, so $N(f_1 + 241) = 1$.

By Strassmann's theorem, $f_0(x) + 241$ has at most two zeros in $\mathbf{Z}_{11}$ and $f_1(x) + 241$ has at most one zero in $\mathbf{Z}_{11}$. The zero corresponding to the value $a_{10} = -241$ is $x = 2$ for $f_0(x) + 241$ (since $10 = 5 \cdot 2 + 0$). Write $f_0(x) + 241 = (x - 2)g(x)$ where $g(x)$ is a power series converging on $\mathbf{Z}_{11}$. Then $N(g) = N(f_0 + 241) - 1 = 1$ by the proof of Strassmann's theorem, so $g(x)$ and $f_1(x) + 241$ both have $N = 1$. By Remark 4.5, $g(x)$ and $f_1(x) + 241$ each have have one root in $\mathbf{Z}_{11}$, so a_m can be -241 for at most two values of m other than 10. The roots of $g(x)$ and $f_1(x) + 241$ don't appear to be nonnegative integers (we estimate them in Appendix B), but it is numerically hard to prove rigorously that an 11-adic integer is not a nonnegative integer from an 11-adic approximation. In order to prove $a_m = -241$ only at $m = 10$ (thereby also proving the unique roots of $g(x)$ and $f_1(x) + 241$ in $\mathbf{Z}_{11}$ are not nonnegative integers) we give up on the prime 11 and seek to apply Strassmann's theorem to $\mathbf{Q}_p$ for some $p > 11$.

Theorem 6.3. *For $m \geq 0$, $a_m = -241$ if and only if $m = 10$.*

Proof. We want to find a prime $p > 3$ such that -2 has a square root in $\mathbf{Z}_p$. Then $|1 \pm \sqrt{-2}|_p = 1$ and for $r \in \{0, 1, \ldots, p-2\}$ and $k \geq 0$, $a_{(p-1)k+r} = g_r(k)$ where

$$\begin{aligned} g_r(x) &= \frac{(1+\sqrt{-2})^r}{2}((1+\sqrt{-2})^{p-1})^x + \frac{(1-\sqrt{-2})^r}{2}((1-\sqrt{-2})^{p-1})^x \\ &= \frac{(1+\sqrt{-2})^r}{2}e^{x\log((1+\sqrt{-2})^{p-1})} + \frac{(1-\sqrt{-2})^r}{2}e^{x\log((1-\sqrt{-2})^{p-1})} \\ &= a_r + \sum_{n\geq 1} d_{r,n}x^n \end{aligned}$$

is a p-adic power series converging on all $x \in \mathbf{Z}_p$, and

$$d_{r,n} = \frac{(1+\sqrt{-2})^r}{2}\frac{(\log((1+\sqrt{-2})^{p-1}))^n}{n!} + \frac{(1-\sqrt{-2})^r}{2}\frac{(\log((1-\sqrt{-2})^{p-1}))^n}{n!} \in p\mathbf{Z}_p$$

for $n \geq 1$. Thus $g_r(x) \equiv a_r \bmod p$ for all $x \in \mathbf{Z}_p$, so if $a_r \not\equiv -241 \bmod p$ then $g_r(x) + 241$ has no zero in $\mathbf{Z}_p$. We want to find p so that $g_{10}(x) + 241$ (which has constant term 0) has $N = 1$ and all other $g_r(x) + 241$ have $N = 0$. (The series $g_r(x)$ and its coefficients $d_{r,n}$ all depend on the choice of p, but we omit this dependence in the notation.)

The first few primes $p > 3$ such that -2 has a square root in $\mathbf{Z}_p$ are 11, 17, 19, and 41. We already saw $p = 11$ is not a good choice.

$\underline{p = 17}$: The only $r \in \{0, 1, \ldots, 15\}$ such that $a_r \equiv -241 \bmod 17$ is $r = 10$, but over $\mathbf{Q}_{17}$, $g_{10}(x) + 241 = d_{10,1}x + d_{10,2}x^2 + \cdots$ has $d_{10,1} \equiv 4 \cdot 17^2 + \cdots$, $d_{10,2} = 6 \cdot 17^2 + \cdots$ and $d_{10,n} \equiv 0 \bmod 17^3$ for $n \geq 3$, so $g_{10}(x) + 241$ has $N = 2$. This is not good.

$\underline{p = 19}$: There are two $r \in \{0, 1, \ldots, 17\}$ such that $a_r \equiv -241 \bmod 19$: $r = 10$ and $r = 12$. Over $\mathbf{Q}_{19}$, $g_{10}(x) + 241$ and $g_{12}(x) + 241$ both have $N = 1$. This is not good.

$\underline{p = 41}$: The only $r \in \{0, 1, \ldots, 39\}$ such that $a_r \equiv -241 \bmod 41$ is $r = 10$. Over $\mathbf{Q}_{41}$ the series $g_{10}(x) + 241$ has constant term 0, linear coefficient $d_{10,1} = 40 \cdot 41 + 16 \cdot 41^2 + \cdots$, and $d_{10,n} \equiv 0 \bmod 41^2$ for $n \geq 2$, so $g_{10}(x) + 241$ has $N = 1$. Thus $x = 0$ is the only zero of $g_{10}(x) + 241$ in $\mathbf{Z}_{41}$. Therefore $a_m = 10$ only for $m = 10$ by working in $\mathbf{Q}_{41}$.

□

Appendix A. Relation to a Diophantine equation

Theorem A.1. *The $m \geq 0$ such that $a_m = \pm 1$ are also the $m \geq 0$ such that $3^m = 1 + 2x^2$ for some integer x.*

The solutions are $(m, x) = (0, 0), (1, \pm 1), (2, \pm 2)$, and $(5, \pm 11)$.

Proof. We will study the equation by working in $\mathbf{Z}[\sqrt{-2}]$, which like $\mathbf{Z}$ has unique factorization and its only units are ± 1. We will assume the reader knows enough number theory to understand how to work in such rings (norms, primes, and relatively prime elements).

In $\mathbf{Z}[\sqrt{-2}]$ both sides of the equation $3^m = 1 + 2x^2$ decompose:

$$((1 + \sqrt{-2})(1 - \sqrt{-2}))^m = (1 + x\sqrt{-2})(1 - x\sqrt{-2}).$$

On the left side, $1+\sqrt{-2}$ and $1-\sqrt{-2}$ are both prime elements of $\mathbf{Z}[\sqrt{-2}]$ since their norms equal 3, which is a prime number. On the right side, the numbers $1 + x\sqrt{-2}$ and $1 - x\sqrt{-2}$ are relatively prime: if δ is a common divisor then δ divides their sum 2, which has prime factorization in $\mathbf{Z}[\sqrt{-2}]$ equal to $-(\sqrt{-2})^2$, so δ is ± 1 or $\pm\sqrt{-2}$. Thus $\mathrm{N}(\delta)$ is 1 or 2. Also δ^2 divides $(1 + x\sqrt{-2})(1 - x\sqrt{-2}) = 1 + 2x^2 = 3^m$, so taking norms shows $\mathrm{N}(\delta)^2$ divides $\mathrm{N}(3^m) = 9^m$. Thus the integer $\mathrm{N}(\delta)$ is a power of 3, so $\mathrm{N}(\delta) = 1$, which means $\delta = \pm 1$.

Since $1+x\sqrt{-2}$ and $1-x\sqrt{-2}$ are relatively prime in $\mathbf{Z}[\sqrt{-2}]$, the only way their product can equal $(1 + \sqrt{-2})^m(1 - \sqrt{-2})^m$ is if

$$1 + x\sqrt{-2} = \pm(1 + \sqrt{-2})^m \text{ or } \pm(1 - \sqrt{-2})^m. \tag{A.1}$$

This is equivalent to saying $(1 + \sqrt{-2})^m$ has real part ± 1. Since the real part is the average of a complex number and its complex conjugate, (A.1) holds for some integer x and some nonnegative integer m if and only if

$$\frac{(1 + \sqrt{-2})^m}{2} + \frac{(1 - \sqrt{-2})^m}{2} = \pm 1,$$

which in light of (1.2) is equivalent to saying $a_m = \pm 1$. □

Appendix B. Estimating roots of $f_0(x) + 241$ and $f_1(x) + 241$ in $\mathbf{Z}_{11}$

We will show how to compute $f_0(x)+241 \equiv (x-2)g(x) \bmod 11^6$ and $f_1(x)+241 \bmod 11^6$ in order to estimate their roots in $\mathbf{Z}_{11}$. Both series have constant term 242. For $n \geq 1$, the coefficient of x^n in $f_0(x) + 241$ is

$$c_{0,n} = \frac{1}{2}\frac{(\log(1 - 11\sqrt{2}))^n + (\log(1 + 11\sqrt{-2}))^n}{n!}.$$

To estimate $c_{0,n}$ we estimate $\log(1+\sqrt{-2})$ and $\log(1-\sqrt{-2})$. If $|x|_{11} = 1/11$ then $|x^k/k|_{11} \leq 1/11^6$ for all $k \geq 6$, so $\log(1 + x) \equiv \sum_{k=1}^{5}(-1)^{k-1}x^k/k \bmod 11^6$. Using this together with the estimate $\sqrt{-2} \equiv 3 + 9 \cdot 11 + 4 \cdot 11^2 + 11^3 + 4 \cdot 11^4 + 4 \cdot 11^5 \bmod 11^6$, we have

$$\begin{aligned}
\log(1 - 11\sqrt{-2}) &\equiv 8 \cdot 11 + 2 \cdot 11^2 + 8 \cdot 11^3 + 3 \cdot 11^4 + 8 \cdot 11^5 \bmod 11^6, \\
\log(1 + 11\sqrt{-2}) &\equiv 3 \cdot 11 + 10 \cdot 11^2 + 2 \cdot 11^3 + 5 \cdot 11^4 + 2 \cdot 11^5 \bmod 11^6.
\end{aligned}$$

Recall from Section 6 that $c_{0,n} \equiv 0 \bmod 11^2$ for all $n \geq 1$, so we use the above to compute

$$\begin{aligned}
\frac{f_0(x) + 241}{11^2} &= 2 + (1 + 10 \cdot 11^2 + 10 \cdot 11^3)x + (10 + 10 \cdot 11 + 10 \cdot 11^2 + 11^3)x^2 + \\
&\qquad (10 \cdot 11^2 + 10 \cdot 11^3)x^3 + (2 \cdot 11^2 + 9 \cdot 11^3)x^4 \bmod 11^4.
\end{aligned}$$

This polynomial has two roots modulo 11^4: 2 and $10 + 10 \cdot 11 + 5 \cdot 11^2 + 3 \cdot 11^3$. (Since $(f_0(x) + 241)/11^2 \equiv 2 + x - x^2 \equiv -(x-2)(x+1) \bmod 11$, a version of Hensel's lemma for power series implies there are roots in $\mathbf{Z}_{11}$ that reduce to 2 and $-1 \bmod 11$.)

For $n \geq 1$, the coefficient of x^n in $f_1(x) + 241$ is

$$c_{1,n} = \frac{1+\sqrt{-2}}{2}\frac{(\log(1-11\sqrt{2}))^n}{n!} + \frac{1-\sqrt{-2}}{2}\frac{(\log(1+11\sqrt{-2}))^n}{n!}.$$

Since $c_{1,n} \equiv 0 \bmod 11$ for all $n \geq 1$, the above estimates let us compute

$$\begin{aligned}\frac{f_1(x)+241}{11} &= 2 \cdot 11 + (2 + 11 + 6 \cdot 11^2 + 2 \cdot 11^3 + 2 \cdot 11^4)x + (10 \cdot 11 + 11^2 + 6 \cdot 11^4)x^2 + \\ &\quad (3 \cdot 11^2 + 6 \cdot 11^3 + 9 \cdot 11^4)x^3 + (2 \cdot 11^3 + 11^4)x^4 + 3 \cdot 11^4 x^5 \bmod 11^5.\end{aligned}$$

There is a unique root modulo 11^5: $10 \cdot 11 + 5 \cdot 11^2 + 9 \cdot 11^4$. (Since $(f_1(x) + 241)/11 \equiv 2x \bmod 11$, Hensel's lemma for power series implies there is one root in $\mathbf{Z}_{11}$ that is congruent to $0 \bmod 11$.)

References

[1] http://math.stackexchange.com/questions/873147/finding-non-negative-integers-m-such-that-1-sqrt-2m-has-real-part/873529