This discussion is adapted from two readily available sources.

Hensel lifting lemma – following Keith Conrad:
https://kconrad.math.uconn.edu/blurbs/gradnumthy/hensel.pdf

Ostrowski's Theorem – from Wikipedia following Shickof(2007)
https://en.m.wikipedia.org/wiki/Ostrowski%27s_theorem

Hensel's Lifting Lemma (1897)

If $f(X) \in \mathbf{Z}_p[X]$ *and* $a \in \mathbf{Z}_p$ *satisfies*

$$f(a) \equiv 0 \bmod p,\ f'(a) \neq 0 \bmod p$$

then there is a unique $\alpha \in \mathbf{Z}_p$ *such that* $f(\alpha) = 0$ *in* $\mathbf{Z}_p$ *and* $\alpha \equiv a \bmod p$.

It is almost as good to consider $f(X) \in \mathbf{Z}[X]$ (that is, a polynomial with integer coefficients). The generalization to $f(X) \in \mathbf{Z}_p[X]$ (that is, p-adic integer coefficients) seems fairly clear but more cumbersome. It seems likely that a slightly simpler approach like this one came first.

For polynomials, the Binomial Theorem gives us a first order Taylor series:

$$f(x+y) = f(x) + f'(x)y + g(x,y)y^2$$

With $a = a_0$ we have the starting case for an induction approach.

So assume we have

$$a_n = \sum_{k=0}^{n} b_k p^k$$

With $b_0 = a$ and $0 <= b_k < p$ and $n>0$

$$f(a_n) \equiv 0\ mod\ p^{n+1} \text{ and } f'(a_n) <> 0\ mod\ p^{n+1}$$

Then (mod p^{n+2})

$$f(a_n + r\,p^{n+1}) \equiv f(a_n) + f'(a_n)r\,p^{n+1} + g(a_n, r\,p^{n+1})r^2\,p^{2n+2}$$

$$\equiv s\,p^{n+1} + f'(a_n)r\,p^{n+1} + 0$$

This last item is zero (mod p^{n+2}) if

$$0 \equiv s + f'(a_n)r \text{ mod p}$$

Of course, mod p, $a_n = a$ so that

$$b_{n+1} \equiv \quad r \equiv -s[f'(a)]^{-1} \text{ mod p}$$

And we are done.

The p-adics may have started as formal power series with the Lifting Lemma providing an impetus to think of a way to have Taylor series in p converge.

By using the idea of a norm on the p-adics with the norm of p less than 1, convergence resulted.

In 1916, Ostrowski was a student of Hensel and characterized the possible norms on **Z** and **Q** as the usual distance between two numbers and the p-adics (one for each prime).

A norm is a map from a set to the real line. The key properties of a norm $|.|$ are these (given a,b in **Q)**:

$$|0| = 0$$

$$|1| = b-1$$

$$|ab| = |a| \ |b|$$

$$|a+b| \leq |a| + |b| \text{ (triangle inequality)}$$

With the norm of a product equal to the product of the norms, it should be clear that unique factorization in **Z** means that a norm on **Q** is fully defined by the norms of primes.

The proof of Ostrowski's Theorem is in two stages:

- If there is an integer with norm greater than 1, then we get a norm equivalent to the usual distance (the norm of a number is its absolute value).
- Otherwise, there is a prime with norm less than 1 which gives us the p-adics for that prime.

Suppose we have an integer n > 1 with norm greater than 1. Let b and k be positive integers greater than 1. Write n^k in base b. Then there is a positive integer m and non-negative integers c_J for j=0 to m with:

$$n^k = \sum_{j=0}^{m-1} c_j b^j$$

Each c_j is at most b-1. Now $n^k > b^{m-1}$ so, taking logs, m <= 1+k $\log_b(n)$ or m-1 <= k $\log_b(n)$.

From the triangle inequality

$$|c_j| \leq |1+1+\cdots+1| \leq c_j \leq b-1$$

Also, the norm of b is either less than or equal to 1 or greater than or equal to 1. so

$$|b|^i \leq \max(1, |b|^{m-1})$$

$$|n^k| = |n|^k = \left| \sum_{j=0}^{m-1} c_j b^j \ \ a \right| \leq m\,(b-1)\max(1,|b|^{m-1}) \leq (1+\mathrm{k}\log_b(\mathrm{n}))\,(b-1)\max\left(1,|b|^{k\,log_b\,n}\right)$$

Raising to the 1/k power, we have

$$|n| \le [\left(1 + \mathrm{k}\log_b(\mathrm{n})\right)(b-1)]^{1/k} \max\left(1, |b|^{log_b n}]\right)$$

Letting k go to infinity, the term with k goes to 1 and we have

$$1 < |n| \le \max\left(1, |b|^{log_b n}]\right) = |b|^{log_b n}$$

Also 1 < $|b|$.

From

$$|n| \le |b|^{log_b n}$$

Take $\log_n$ of both sides:

$$log_n|n| \le log_n(|b|^{log_b n}) = (log_b n) log_n|b| = log_n|b|(log_b n) = \left(log_b n^{log_n|b|}\right) = log_b|b|$$

As b and n are both positive integers greater than 1, we can reverse the process

$$log_b|b| \le log_n|n|$$

So that

$$log_b|b| = log_n|n|$$

Fix any b (such as 2) and $log_2|2|$ determines this norm. In particular, if $log_2|2| = 1$ then the norm is the usual absolute value.

For the second part, for all integer n, we must have $|n| \le 1$. If the norm is not trivial, there is n with $|n| < 1$. In particular, there is at least one prime with norm less than 1. We wish to show that there is only one. So, suppose p and q are distinct primes with norm less than 1.

(We will use Bezout's Theorem so a proof is given later.)

There is a positive integer k such that

$$|p^k| < \frac{1}{2} \; and \; |q^k| < \frac{1}{2}$$

Bezout's Theorem tells us that there are integers a and b such that

$$ap^k + bq^k = 1$$

Then

1= $|ap^k + bq^k| \le |ap^k| + |bq^k| = |a||p^k| + |b||q^k| \le |p^k| + |q^k| < \frac{1}{2} + \frac{1}{2} = 1$

1 < 1 gives us the necessary contradiction and establishes the second type of norm.

Bezout's Theorem has a number of generalizations but we need only the integer case.

Given (positive) integers a and b, there are integers x and y with $ax+by = GCD(a,b)$.

If a and b are relatively prime, then there are x and y with $ax+by = 1$.

Consider the set of values $ax+by$ as x and y vary. Let d be the smallest positive value in this set.

If d is not the GCD(a,b) then there are integers n and r with $0<r<d$ and

$a = nd + r$

Then $r = a - nd = a - n(ax+by) = a(1-nx) + b(-ny) < d$.

This contradicts the construction of d.