A Whirlwind Tour of the p-adic Numbers

Patrick O'Melveny

November 23, 2021

1 Introduction

1.1 What's so special about $\mathbb{R}$ anyways?

We often treat our standard collection of number systems as a progression of extra features tacked on to the natural numbers. We can, with our gift of hindsight, see quite a logical progression from natural numbers to integers to rationals. Indeed, the move from rationals to the reals seems fairly straightforward. We, however, are interested in the step from $\mathbb{Q}$ to $\mathbb{R}$.

While there are many reasons you may invent the reals, an analyst may stumble upon one in particular. Consider the sequence

$$a_n := \left(1 + \frac{1}{n}\right)^n.$$

Each element of our sequence is quite safely a rational number, and a little work will get you that it is Cauchy.[1] Unsettlingly though, its limit doesn't seem to exist, at least not within the realm of rationals. We say, since we have at least one Cauchy sequence whose limit does not exist, that $\mathbb{Q}$ is not complete. Then $\mathbb{R}$ could be thought of as the completion; i.e. it is everything in $\mathbb{Q}$ and all possible limits of Cauchy sequences.[2]

However, we've taken something for granted here. The property of being complete really only makes sense in a metric space, and in a metric space we need a metric. Likely, you filled in the correct metric without even thinking about it. It seems uncontroversial to assume the distance from point $a \in \mathbb{Q}$ to point $b \in \mathbb{Q}$ is

$$|a - b|.$$

But then it would be more accurate to say $\mathbb{R}$ the completion of $\mathbb{Q}$ with respect to the Euclidean metric.

[1] Hint: 3 is a safe upper bound for this sequence.

[2] It's a little more involved to make this rigorous. There are equivalence classes involved.

This leads to a very interesting question, what happens if we pick a different choice of metric on $\mathbb{Q}$? Are there any limitations to these metrics? Will we always recreate $\mathbb{R}$, or will we have to deal with intractably many different completions, each with wildly different properties? Can we still do analysis on these new fields?

Amazingly, there is actually only one other family of completions of $\mathbb{Q}$, which we call the p-adic numbers. We will provide a brief survey into the very interesting and occasionally strange world of p-adics, how they're formed, and how analysis works with them. First let us review some basic concepts around metric spaces.

2 Norms and their Metrics, Both Old and New

2.1 A Brief Review

To begin to understand how the p-adic numbers arise we need a different sense of "nearness", and to that end we must recall the basic ideas of a norm[1].

Definition 2.1 (norm)**.** A norm on a field F is a map $||\cdot|| : F \to \mathbb{R}_+$ such that for any $x, y \in F$

i. $||x|| = 0$ if and only if $x = 0$,

ii. $||xy|| = ||x||\,||y||$,

iii. and $||x + y|| \leq ||x|| + ||y||$ (the triangle inequality).

From a norm, it is always possible to define a sense of distance, which we call metric. For any norm $||\cdot||$ defined on the field F, we can make a corresponding metric function

$$\begin{aligned} d : F \times F &\to \mathbb{R}_+ \\ (x, y) &\mapsto ||x - y||. \end{aligned}$$

A metric, along with the set it is defined on gives us a metric space defined[6] as

Definition 2.2 (metric space)**.** A metric space is a set X together with a function $d : X \times X \to \mathbb{R}_+$ which satisfies the following properties for all $x, y, z \in X$:

i. $d(x, y) \geq 0$ with $d(x, y) = 0$ if and only if $x = y$ (positive definiteness),

ii. $d(x, y) = d(y, x)$ (symmetry),

iii. $d(x, y) \leq d(x, z) + d(z, y)$ (the triangle inequality).

2.1.1 The "Standard" Absolute Value

The norm on $\mathbb{Q}$ we used implicitly in the introduction is our good old friend, the absolute value function

$$|q| = \begin{cases} q, & q \geq 0 \\ -q, & q < 0 \end{cases}.$$

For reasons that will become clear momentarily, going forward we will notate the absolute value of a rational number as

$$|q|_\infty$$

though just keep in mind it's just a notational convenience and not really anything to do with infinity.

The absolute value gives rise to the Euclidean metric on $\mathbb{Q}$, which, as we've stated above, we can use to extend $\mathbb{Q}$ to $\mathbb{R}$. This is likely the most intuitive sense of distance on $\mathbb{Q}$, since it coincides so well with how we experience distance as beings inhabiting three dimensional space.[3] But what kind of mathematicians would we be if we stop while things are still intuitive? Could we come up with some kind of other norm on $\mathbb{Q}$?

2.2 The p-adic Absolute Value

Let us choose a prime p. All the following[4] holds for any choice of p, but requires p to be fixed. Let $x \in \mathbb{Q}$, then we define $\upsilon_p(x)$ such that

$$x = p^{\upsilon_p(x)} x'$$

where x' is coprime[5] to p. That is to say $\upsilon_p(x)$ is the highest power of p that still divides x. We will write $\upsilon_p(x)$ simply as $\upsilon(x)$ when p is clear from context.

Now we can define the p-adic absolute value $|\cdot|_p$ on $\mathbb{Q}$,

$$|x|_p = \begin{cases} p^{-\upsilon(x)} & x \neq 0 \\ 0 & x = 0 \end{cases}.$$

We claim $|\cdot|_p$ constitutes a norm.[4]

Proof. We must check the properties of a norm. First we trivially have $|0|_p = 0$ by definition. Likewise, if $x \neq 0$ then $\frac{1}{p^{\upsilon(x)}} \neq 0$.

[3] This probably needs a citation from either a physicist or a philosopher.

[4] Exceptions may apply for $p = 2$. Though we will note those explicitly as needed.

[5] A rational, x, is coprime to prime p if $x = \frac{a}{b}$ and a is coprime to p and b is coprime to p

If we have x, y in Q, then we know $x = p^{\upsilon(x)}x'$ and $y = p^{\upsilon(y)}y'$. Then

$$\begin{aligned} xy &= p^{\upsilon(x)}x'p^{\upsilon(y)}y' \\ &= p^{\upsilon(x)+\upsilon(y)}(xy)', \end{aligned}$$

since the product $x'y'$ couldn't suddenly contain an additional factor of p. From there we can confirm

$$\begin{aligned} |x|_p|y|_p &= p^{-\upsilon(x)}p^{-\upsilon(y)} \\ &= p^{-(\upsilon(x)+\upsilon(y))} \\ &= |xy|_p \end{aligned}$$

Finally, we must turn to the triangle inequality. Assume we have $x = p^{\upsilon(x)}\frac{m}{n}$ and $y = p^{\upsilon(y)}\frac{k}{l}$. Without loss of generality, assume $\min(\upsilon(x), \upsilon(y)) = \upsilon(x)$. We have

$$x + y = p^{\upsilon(x)}\frac{m}{n} + p^{\upsilon(y)}\frac{k}{l} = p^{\upsilon(x)}\frac{ml + nkp^{\upsilon(y)-\upsilon(x)}}{nl}. \tag{1}$$

Since n, l don't have p as a factor, the denominator nl likewise contains no factor of p. It is possible that $ml + nkp^{\upsilon(y)-\upsilon(x)}$ introduces a factor of p, however. So we have then $\upsilon(x + y) \geq \min(\upsilon(x), \upsilon(y)) = \upsilon(x)$ and therefore

$$|x + y|_p = p^{-\upsilon(x+y)} \leq \max\left(p^{\upsilon(x)}, p^{\upsilon(y)}\right) = \max(|x|_p, |y|_p)$$

Since, trivially,

$$\max(|x|_p, |y|_p) \leq |x|_p + |y|_p$$

we have shown the triangle inequality. □

So our p-adic absolute value does indeed constitute a norm. But, looking again at our proof of the triangle inequality, it is interesting we showed something stronger than just the triangle inequality. In fact we have that $|x+y|_p \leq \max(|x|_p, |y|_p)$. This general property is not unique to just the p-adic absolute value and is important enough to have it's own name.

2.2.1 Non-Archimedean Norms and Ultrametrics

We call a norm like the p-adic absolute value non-Archimedean.[1]

Definition 2.3 (non-Archimedean)**.** A norm is called non-Archimedean if $||x+y|| \leq \max(||x||, ||y||)$ always holds.

As you might hope, the metric induced by a non-Archimedean norm inherits this stronger triangle inequality and is an ultrametric.

Definition 2.4 (ultrametric space[6])**.** A metric space (X, d) is an ultrametric space if it satisfies the property

$$d(x, y) \leq \max(d(x, z), d(z, y))$$

for all $x, y, z \in X$.

In fact, the p-adic absolute value is even slightly stronger still, as if $|x|_p \neq |y|_p$ then $|x + y| = \max(|x|_p, |y|_p)$.

2.3 Are There Any More Completions of $\mathbb{Q}$ to Find?

Since we've been launched off on this whole journey by looking at a different norm on $\mathbb{Q}$ than the one we're all accustomed too, it seems worth considering that there might be even more norms out there.

However, as hinted at in the introduction, that is not actually the case. First, some definitions

Definition 2.5. Two metric spaces d_1 and d_2) on X are said to be equivalent if they induce the same topology on X.

Definition 2.6. Two norms are said to be equivalent if their induced metrics are equivalent.

Brushing aside what it means for two topologies are equivalent, armed with these two definitions are able to take on the concise but surprising theorem by Ostrowski.[3]

Theorem 2.1 (Ostrowski, 1918)**.** *Every non-trivial norm $||\cdot||$ on $\mathbb{Q}$ is equivalent to $|\cdot|_p$ or $|\cdot|_\infty$.*

. A nice proof of this theorem can be found in [2].

While this still leaves an infinite family of norms, since there's one for each prime p, this theorem is quite surprising. There's no metric we can define on $\mathbb{Q}$ that won't be equivalent to one we now know, no matter how hard we look. This just means it's even more important to understand our new norm. Our next step is to finally use our norm to define a completion.

3 Completing $\mathbb{Q}$, the p-adic Numbers

3.1 Construction of $\mathbb{Q}_p$

We will, as promised, complete $\mathbb{Q}$ with respect to Cauchy sequences. Recall a sequence $\{x_n\}_{n=1}^{\infty}$ is Cauchy in a norm induced metric space if for any $\varepsilon > 0$ there exists some

[6] An ultrametric space can also called be called a non-Archimedean space, making it's connection to the non-Archimedean norm a bit more explicit. Though not all ultrametric spaces arise from non-Archimedean norms.

$N \in \mathbb{N}$ such that $m, n \geq N$ implies $||x_n - x_m|| < \varepsilon$. This is mostly just to say that the metric is what defines what sequences are Cauchy. The sequence

$$a_n := 5^n$$

is not Cauchy with respect to the Euclidean metric, but most certainly is with respect to the metric induced by $|\cdot|_5$.

Let R be the set of all Cauchy sequences in $\mathbb{Q}$ with respect to the metric induced by $|\cdot|_p$. We can endow R with a ring structure using the operations of $\mathbb{Q}$. Given $\{a_n\}, \{b_n\} \in R$, let

$$\{a_n\} + \{b_n\} = \{a_n + b_n\}$$

and

$$\{a_n\}\{b_n\} = \{a_n b_n\}.$$

We have then that $(R, +, \times)$ is a commutative ring. Let

$$M = \big\{\{x_n\} \in R \;\big|\; \lim_{n\to\infty} x_n = 0\big\},$$

that is, the set of all Cauchy sequences that go to 0. Then M is a maximal ideal[2], and so the quotient R/M is a field. Specifically it is a field where every element is an equivalence class of Cauchy sequences where $\{a_n\} \sim \{b_n\}$ if and only if $\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n$.

This field we have constructed is our sought after $\mathbb{Q}_p$, the completion of $\mathbb{Q}$ with respect to the norm $|\cdot|_p$.

3.2 Elements of $\mathbb{Q}_p$

Given how we have defined $\mathbb{Q}_p$ thinking of the constituent elements can be a bit difficult at first. They're equivalence classes of Cauchy sequences in a metric space that has an entirely different sense of distance. But, much like we aren't don't usually need to stress about the underlying equivalence classes of the reals or rationals, now that we have constructed $\mathbb{Q}_p$ we can mostly forget about how we did it. Instead we can think about elements of $\mathbb{Q}_p$ as being just (p-adic) numbers. To that end, consider the following theorem[4]

Theorem 3.1. *Every p-adic number $x \in \mathbb{Q}_p$, $x \neq$) has a unique canonical form*

$$x = p^{v(x)} \sum_{i=0}^{\infty} x_i p^i$$

where $v(x) \in \mathbb{Z}$ and x_i are integers such that $0 \leq x_i \leq p - 1$.

This gives us two important ideas to latch on to. First, this canonical form gives us a guaranteed to exist choice of a Cauchy sequence to represent an equivalence class in $\mathbb{Q}_p$; this is because

$$x_n := p^{v(x)} \sum_{i=0}^{n} x_i p^i$$

is a Cauchy sequence that converges to x. Second and more helpful is it gives us a sense of how we could have a more concrete notational representation of the p-adic numbers.

Consider a real number in expressed in base-10. For digits $0 \leq x_i \leq 9$ we would think of the formal sum

$$\sum_{i=k}^{-\infty} 10^k x_i$$

equivalent to its "decimal expression"

$$x_k x_{k-1} \dots x_1 x_0 . x_{-1} x_{-2} \dots .$$

Here we have finitely many digits to the left of the decimal point and infinitely many to the right.

Our canonical form looks very similar to a base-p expression of a real number. But, in contrast, our p-adic number can have only finitely many terms associated with negative powers of p, while infinitely many terms tied to positive powers. This gives rise to a representation of p-adic numbers that seems both familiar but slightly backwards.[7] They extend off to the left of the decimal point[8] while having only finitely many to the right.

3.2.1 A Very Small Example

Consider the field $\mathbb{Q}_3$. We in this system we would write 2 as

$$\dots 0002,$$

and $10.33333\dots$ becomes

$$\dots 000101.1$$

in the p-adics.

Our rather basic choices in example hide the fact that, just like most real numbers have infinitely long decimal expansions, most elements of $\mathbb{Q}_p$ would have infinitely many non-zero digits to the left.

[7] this way of writing p-adic numbers is not universally used. Some sources only use the canonical form sum. This is however fairly standard and has the benefit of being similar to the decimal numbers we know and love.

[8] It can't really be a decimal point anymore, of course, but it looks the same and functions similarly.

An unexpected development in $\mathbb{Q}_p$ is that negative signs are no longer necessary, all p-adic numbers really never do vary from the canonical form. Continuing our toy examples, we see that -1 in this p-adic notation is now

$$\ldots 99999.$$

Because

$$\ldots 9999. + \ldots 0001. = \ldots 0000.$$

we have that it really is the inverse of 1. In general, all additive inverses can be found via a process called p's compliment.

4 Where To Go From Here

4.1 What Are they Good For

We now have the basic idea's of p-adic numbers down, but you may want to know why any of this is useful.

The very short answer is, they've shown themselves to be quite powerful in the context of number theory. A large source of modern fame for the p-adics is their role in Andrew Wiles's proof of Fermat's Last Theorem.[7] Along the same lines of famous problems, they have many properties that make them ideal for studying zeta functions, as noted in [1].

While beyond the scope of this paper, there are extensions of Q_p to an algebraically closed field, much in the same way one goes from $\mathbb{R}$ to $\mathbb{C}$. This opens up even further avenues of research, and is needed to define more analogues of calculus on p-adics.[1]

Last, but not least, there seems to be a pervasive interest in using p-adic numbers in mathematical physics as in [4] and [5]. While there seems to be several approaches in how to combine them, there seem to be analogs in both quantum and classical physics. One particularly interesting approach is how the non-Archimedean properties of the p-) may provide a useful model to how distances in the quantum scale.[4]

The p-adic numbers seem to have at least a little presence in every field of mathematics and are still found in active research.[9] You can almost certainly find them in or near your preferred sub-discipline, and you never know when the might be the tool you need.

[9] A Google scholar search for the term, even when limited just to results from 2021 returns a huge amount of recent work utilising them.

References

[1] Neal Koblitz. p-adic Numbers, p-adic Analysis, and Zeta-Functions. Vol. 58. Springer-Verlag, 1977.

[2] M Ram Murty. Introduction to p-adic analytic number theory. Vol. 27. American Mathematical Soc., 2009.

[3] Alexander Ostrowski. "Über einige Lösungen der Funktionalgleichung ψ (x)· ψ (x)= ψ (xy)". In: Acta Mathematica 41.1 (1916), pp. 271–284.

[4] Vasilii Sergeevich Vladimirov, Igor Vasilievich Volovich, and Evgenii Igorevich Zelenov. p-adic Analysis and Mathematical Physics. World Scientific, 1994.

[5] Vasilii Sergeevich Vladimirov, Igor Vasilievich Volovich, and Evgenii Igorevich Zelenov. p-adic Analysis and Mathematical Physics. World Scientific, 1994.

[6] William R. Wade. An Introduction to Analysis. Fourth Edition. Pearson, 2018.

[7] Andrew Wiles. "Modular elliptic curves and Fermat's last theorem". In: Annals of mathematics (1995), pp. 443–551.