p-adic numbers in the real world

Wiskundigen kennen allerlei soorten getallen – en veel daarvan worden ook in de natuurkunde gebruikt. Heel fascinerend zijn de zogeheten p-adische getallen. In deel 3 van deze serie legt Lizzy Rieth uit hoe de structuur van p-adische getallen ook terug te vinden is in de wereld om ons heen.

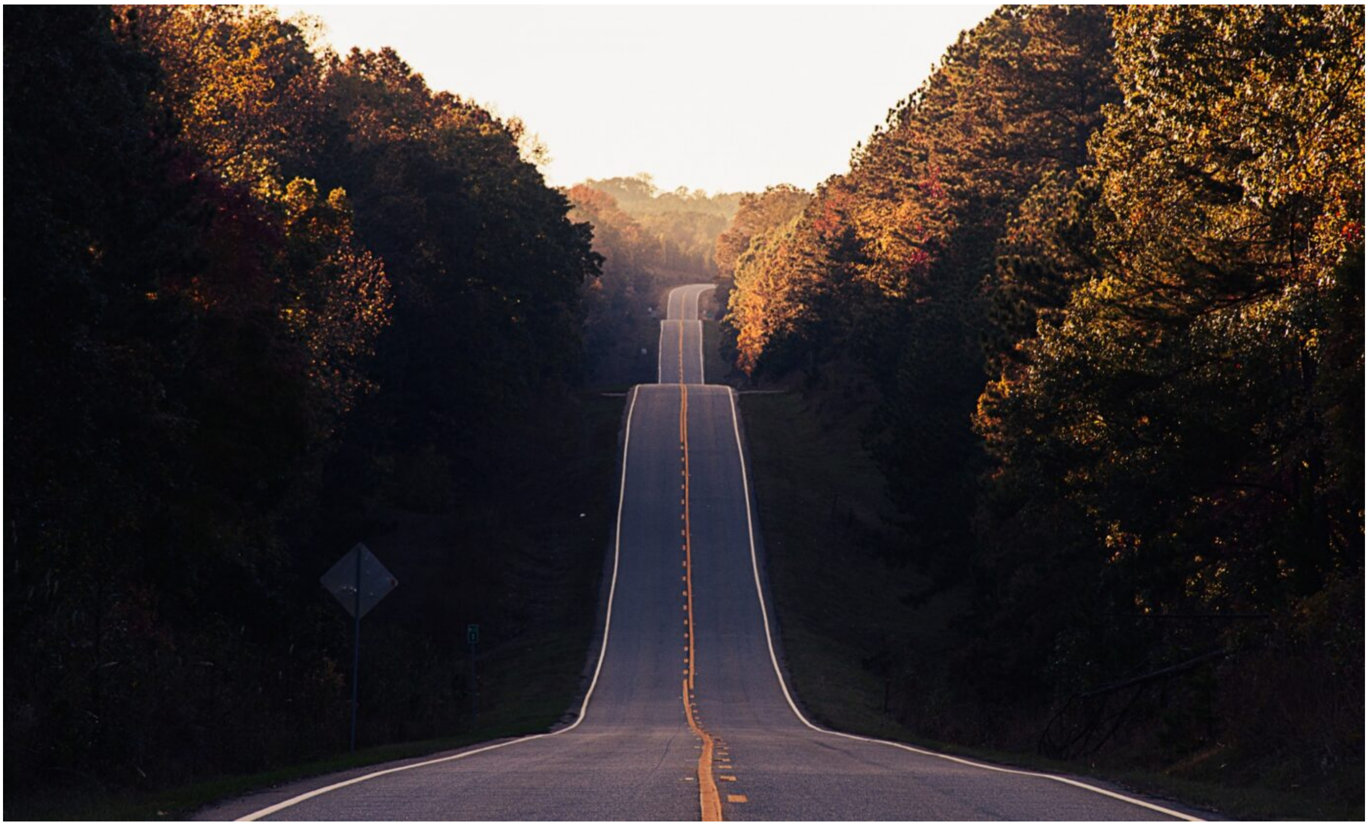

Figure 1. Distance matters. We often think of numbers as situated on a real line – think of measuring distances on a road. However, if we change the shape of the road – and therefore: change the distance concept – the notion of numbers also changes. The p-adic numbers are what you get if you do a very rigorous change of this type, much less smooth than the relatively minor deformation of the road in this picture. [Image source: Public domain]

In the first article of this series, we began by discussing natural numbers, integers, rational numbers and real numbers. We saw that rational numbers, which can be written as fractions of two integers, can be pictured as discrete points lying densely along the familiar number line. This implies that we can always find another rational number between any given two by rewriting fractions with increasingly large denominators. However, fractions do not fill the entire number line completely: there are numbers, like $\sqrt{2}$ and π, which cannot be written as fractions, although sequences of rational approximations can converge to them. For instance, we saw the following series of rational numbers, $1.4, 1.41, 1.414, \ldots$, which converges to the real value of $\sqrt{2}$ in the limit of infinitely many digits following the decimal point. By adding all such missing limits, we can “fill in the gaps” and obtain all the real numbers along the familiar continuous number line as a completion of the rational numbers. A crucial observation was that this completion process depends on what we mean by distance. Using the ordinary notion of distance to complete the rationals produces the real numbers, whereas using the *p-adic distance*, which measures how many times the difference between two numbers is divisible by a chosen prime, leads to the p-adic numbers.

In the second part of this series, we considered the notion of p-adic distance within the broader setting of metric spaces. In particular, we compared the triangle inequality of familiar Euclidean and curved geometries with the stronger ultrametric inequality obeyed by p-adic distances. As a consequence, we found that triangles in ultrametric spaces can only be equilateral or isosceles. This strongly constrains how points can be arranged: at successive distance scales, they naturally fall into clusters within clusters, thus often forming repeating, self-similar patterns. Ultrametric spaces can therefore be pictured as branching trees, or as nested clusters of points, rather than a smooth line or plane. In this picture, two points are considered close when they remain grouped together through many successive layers. Likewise, their distance is determined by the level at which their branches separate, rather than the physical proximity in a continuous space.

Now, you might wonder why mathematicians would bother coming up with such alternative number systems. It turns out that some structures found in nature are organised more like ultrametric spaces than like the smooth geometries familiar from everyday life. In such systems, closeness does not depend on ordinary spatial separation. Rather, separations between two elements are measured in terms of the number of shared levels in a hierarchical structure, which reflects precisely the kind of organisational principle captured by p-adic and ultrametric mathematics! In today’s article, we shall explore this idea further and encounter concrete examples from different sciences.

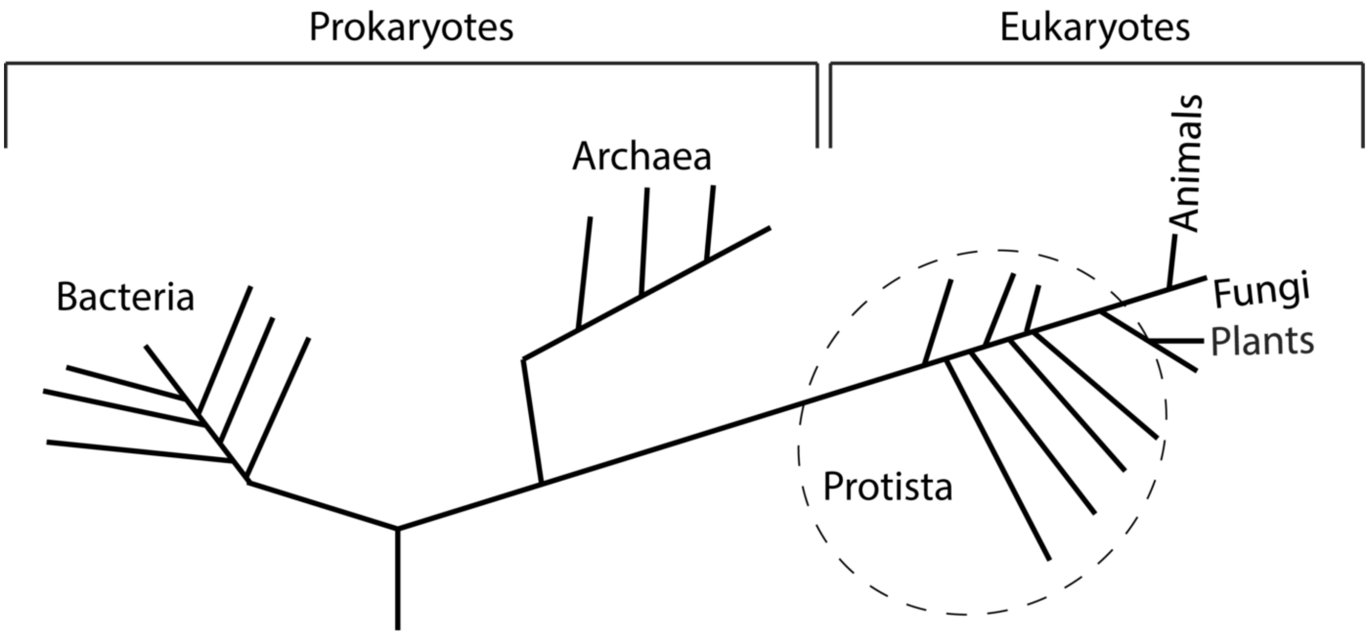

Figure 2. Genetic trees as ultrametric spaces. When systems are organised into nested, tree-like hierarchies, alternative notions of distance, such as the p-adic norm, can provide a more natural way of describing their structure. Spaces with this particular kind of geometry are known as ultrametric spaces. One example arises in biology, where ultrametric distances can describe evolutionary relationships: organisms are considered closer when they share a relatively recent common ancestor, like animals and fungi. Conversely, organisms are considered more distant if their lineages separated further back in time, as is the case when comparing animals to bacteria Figure: Sarah Greenwood, CC BY-SA 4.0, via Wikimedia Commons.

i) Visualizing an ultrametric space

A beautiful way to visualise the space of 2-adic numbers is through the **Bruhat–Tits tree**, a branching graph that represents the 2-adic numbers. Remember, from the previous article, that any 2-adic number has an infinite binary expansion

$$a_0+a_1 2+a_2 2^2+a_3 2^3+\ldots \text{ with } a_i \in\{0,1\}.$$

The 2-adic numbers can then be thought of as the endpoints at infinity, or “leaves” of the tree. Each path towards an endpoint records increasingly precise information about a number. At each branching step, the path chooses between the next digit being 0 or 1. Note that none of the nodes actually represent individual digits. Instead, each node represents a family of numbers that share the digits chosen so far. Thus, two numbers are considered 2-

adically close when their paths remain together for many steps before separating. This means that they have many initial binary digits in common, or equivalently, that their difference is divisible by a large power of 2. Numbers, i.e., “leaves”, that share a longer path are therefore closer, which naturally organises the 2-adic numbers into hierarchical nested clusters.

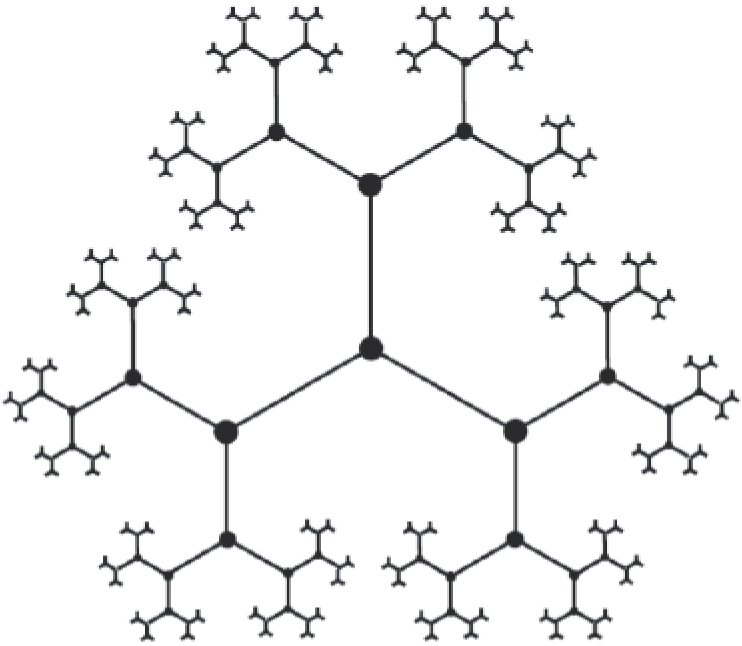

Figure 3. The Bruhat-Tits tree. The branching structure of the Bruhat–Tits tree is a geometric representation of the 2-adic numbers. Two numbers are close when their paths through the tree remain together for many branching steps before separating.

On a more technical note, it is important to stress that the Bruhat–Tits tree is an infinite structure and the graph shown in Figure 3 represents only a small truncated part of it. The complete tree has no natural centre and continues indefinitely in every direction. To create a finite visualisation, we therefore have to choose one node as a starting point, or “root”, and arrange the remaining nodes in levels around it. Thus, the outermost nodes in the diagram should not be regarded as true endpoints. Instead, each represents a cluster of possible boundary values that can only be distinguished at the chosen level of detail. In this sense, truncating the tree is like choosing a limited resolution, or working with only a finite number of digits of a 2-adic expansion. If the tree were extended, each cluster would further divide into smaller clusters, according to the self-similar branching structure. A true particular 2-adic number is described by following a path through infinitely many levels, even when its expansion eventually consists entirely of zeros.

Moreover, for p=2, every node of the Bruhat–Tits tree is connected to three others, resulting in a self-similar branching structure with fractal-like properties. Once a root has been chosen, this is easy to picture for all nodes away from the centre: one edge leads back towards the root, while the other two lead forwards, corresponding to the two possible binary digits, 0 or 1. The root itself has no parent, so all three of its edges appear to point outwards. These three “initial branches” in fact represent three broad directions. The first two contain the 2-adic integers: one branch contains those whose first digit is 0 corresponding to the even 2-adic integers, while the other contains those whose first digit is 1, corresponding to the odd ones. The third branch contains the 2-adic numbers lying outside the 2-adic integers, together with a special point denoted by \infty. These numbers still have perfectly valid 2-adic expansions, but their expansions involve negative powers of 2. Thus, the boundary of the complete Bruhat–Tits tree represents not only the 2-adic integers (\(\mathbb{Z}_2 \)), but a larger space known as the projective 2-adic line.

The fractal-like structure of the p-adic integers can also be visualised using nested circles, as shown in Figure 4, inspired by Heiko Knospe’s beautiful plots for (p=3,5,7) on this website. As in the previous article, let’s consider the 5-adic integers as an example. At the first level, the integers are divided into the five residue classes of the integers modulo 5, also denoted as \(\mathbb Z/5\mathbb Z \). A residue class modulo 5 consists of all integers with the same remainder upon division by 5. Dividing any integer by 5 can yield results with 5 distinct possible remainders; 0,1,2,3,4; which can be used to label the distinct residue classes. For example, at level one, the light green circle labelled “0” represents the integers [0,5,10,15,20,25, …].

At deeper levels, the diagram should be read from the outside in: the largest circles distinguish integers by their remainder modulo 5, while each subsequent layer of smaller circles distinguishes them more precisely by their remainder modulo 5^2, 5^3, …

At the second level, each of these five circles representing \(\mathbb{Z} / 5 \mathbb{Z} \) splits into five smaller circles. “Level 2” in figure 4 thus depicts the 25 residue classes;

labelled by 0,1,…, 24; of the integers modulo 5^2, also known as $\mathbb{Z} / 25 \mathbb{Z}$. At the second level, the numbers [0,5,10,15, 20], appear within the larger circle which was labelled 0 modulo 5 at the first level. This congruence relation becomes visible through the nesting structure: Their common "parent" circle at level 1 shows that they are congruent modulo 5, while their subdivision into five "child" circles at level 2 shows that they belong to distinct residue classes modulo 25. At the third level, describing the integers modulo 125 $\mathbb{Z} /125 \mathbb{Z}$, there are already 125 residue classes. This type of subdivision indefinitely continues at every subsequent power of 5, where progressively deeper levels are represented by darker shades of green in figure 4.

Analogous to how the 2-adic integers are represented by infinite paths through the Bruhat-Tits tree, a 5-adic integer thus corresponds to an infinite path through the green nested circles. In other words, any 5-adic integer can be uniquely defined by its residue class modulo 5, a contained class modulo 5^2=25, a contained class modulo 5^3=125, a contained class modulo 5^4=625, and so on.

The nested circles are thus really just a visual representation of the 5-adic expansion:

$$a_0 + a_1 5 + a_2 5^2 + a_3 5^3 + \cdots, \qquad a_i\in \{ 0,1,2,3,4 \}.$$

These nested circles also capture the behaviour of 5-adic distance: Two 5-adic integers are close when they remain in the same circles through many levels, implying that their difference is divisible by a large power of 5. In the previous article of this series, we looked at the examples of the 5-adic distances between 5, 25, and 50 to show that they satisfy the ultrametric triangle inequality. In that example, we found that $d_5(5,25)=d_5(5,50)=\frac{1}{5}$, whereas $d_5(25,50)=\frac{1}{25}.$ All three numbers are congruent modulo 5, so they occupy the same first-level circle, labelled by "0" in level 1 of figure 4. At the second level, 25 and 50 remain in the same remainder-circle; labelled by "0" in level 2 of figure 4; because both are congruent to 0 modulo 25. However, at the second level, the integer 5 belongs to a different congruence class, labelled 5 in level 2 of figure 4. Hence, as we already concluded in the previous article, 25 and 50 are 5-adically closer to one another than either is to 5. In general, the deeper (i.e. darker) the last circle shared by two numbers, the smaller their 5-adic distance. For another example, the integers 7 and 132 both lie in the same residue classes modulo 5, 25, and 125. This can be seen from the fact that $|132- 7| = 125 = 5^3.$ These two integers therefore are separated by a 5-adic distance of $|132- 7|_5 = d_5(132-7) = 5^{-3} = 1/125$ and follow the same path through the first three levels. By the same logic, 7 and 132 are 5-adically closer than 25 and 50, which only share a path up to the second level depicted in figure 4.

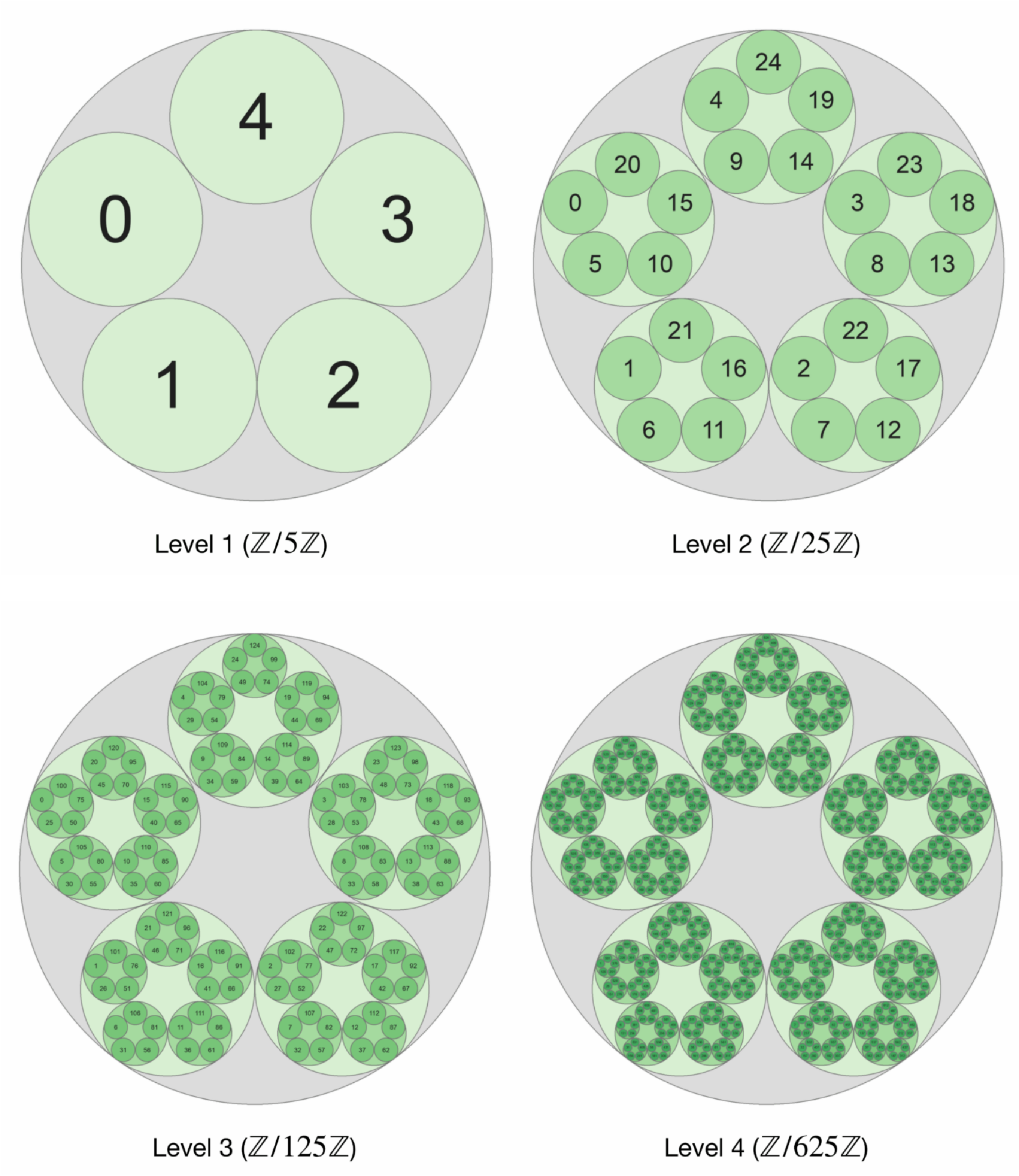

Figure 4. Nested-circle representation of the 5-adic integers. The four panels show the residue classes modulo 5 (level 1), 25 (level 2), 125 (level 3), and 625 (level 4). Each circle subdivides into five smaller circles at the next level, while its numerical label states the corresponding residue class. The nesting structure represents 5-adic closeness: the more circles two integers share before separating, the smaller their 5-adic distance. Progressively darker shades of green indicate increasingly deep levels of subdivision. Image inspired by Heiko Knospe's visualisations on [this website].

ii) Ultrametricity in the real world

One of the first scientific applications of ultrametricity outside mathematics emerged in biology, where it provides a natural way of describing evolutionary relationships between different species.[2] Rather than arranging organisms according to their physical locations, biologists classify them according to their **evolutionary history**. Evolution proceeds through a series of branching events, in which populations that share a common ancestor split and then continue to evolve independently and give rise to new species. In this way, evolutionary relationships naturally form a tree, known as a **phylogenetic tree** or **dendrogram**. Present-day species appear at the leaves of the tree, while branching points represent common ancestors in the past.

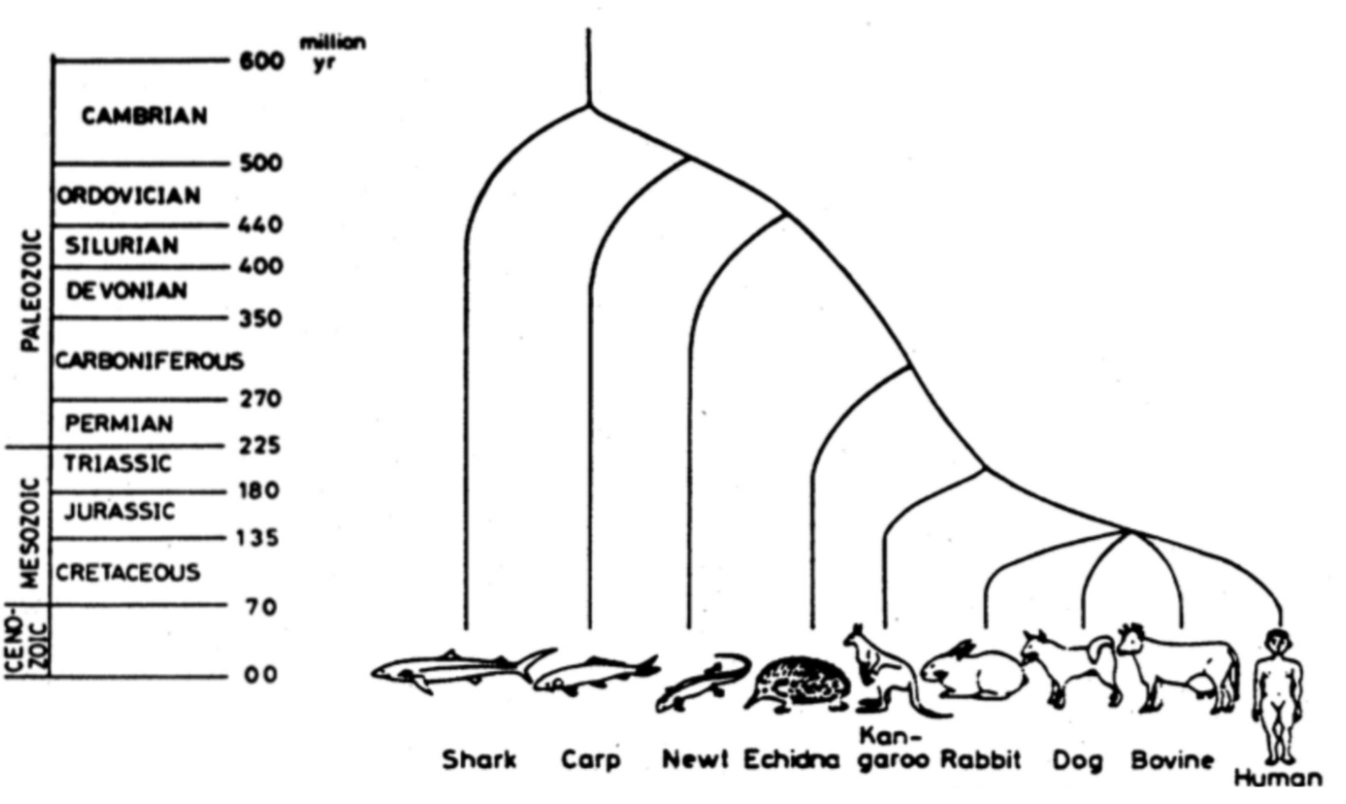

Figure 5. A phylogenetic tree. Phylogenetic tree showing the evolutionary relationships between several modern vertebrate species. The branching points indicate when different lineages diverged from a common ancestor, with the approximate timing indexed by the geological timescale on the left. Adapted from a figure in [2].
Two species are considered close if they share a recent common ancestor, whereas species whose lineages separated much earlier are regarded as further apart. In modern biology, such relationships are predominantly studied through molecular analyses. As evolution progresses, genes gradually change over time. Comparing differences in genetic information between species thus provides clues about how closely related they are and when their lineages diverged in the past.[2]

In this way, evolutionary relationships naturally organise themselves into nested groups: species into genera, genera into families, families into orders, and so on. If we measure the distance between two present-day species by how far back in time we must travel to reach their most recent common ancestor, this evolutionary distance approximately satisfies the ultrametric property! For any three modern-day species, their pairwise evolutionary distances correspond to the side lengths of a nearly equilateral or isosceles triangle, with the two longer sides equal.

For instance, one might naively imagine that a dog is, in some sense, evolutionarily intermediate between a rabbit and a human. Yet the three lineages diverged within a relatively short evolutionary interval, making their pairwise distances approximately equal, thereby forming a nearly equilateral triangle. This is moreover an example of the absence of intermediates in ultrametric spaces. In ordinary geometry, we are used to imagining one point as lying somewhere between two others. However, in a hierarchical evolutionary tree, this intuition fails: distances are determined by shared branching points in the tree, rather than by positions along a continuous line, leaving no natural place for an object to sit “halfway” between two others.

Similarly, one might imagine a cow as occupying an intermediate position between a kangaroo and a human. However, the evolutionary tree tells a different story: Kangaroos are **marsupials**, whereas cows and humans belong to the **placental mammals**, whose common lineage separated from marsupials much earlier. The kangaroo therefore shares the same earlier ancestral branching point with both the cow and the human. In terms of genetic distance, it is equally far away from both, while the cow and human are closer to one another. The three species thus form an isosceles triangle with the two longer sides equal, precisely one of the patterns required by ultrametric geometry.

Finally, I would like to mention that ideas related to ultrametricity have also been impactful in several areas of physics.[1,2] One of the most famous examples appears in the theory of **spin glasses**: disordered magnetic materials with many competing interactions, preventing them from settling easily into a single low-energy equilibrium state.[2] In the 1980s, the physicist Giorgio Parisi made the fascinating discovery that, in spin glasses, the enormous number of possible equilibrium states is not arranged randomly. Instead, these states form a nested, tree-like hierarchy with an ultrametric structure. Importantly, this does not mean that physical space itself is ultrametric at subatomic scales. Rather, ultrametricity applies to an abstract space of possible spin configurations, in which two states are considered close if their microscopic arrangements are similar. This hidden organisation became a central insight in the theory of disordered and complex systems, for which Parisi was awarded one half of the Physics Nobel Prize in 2021. (Click this link for a Nobel Prize overview of Parisi’s work.)

Another application of p-adic numbers can be found in **string theory,** which posits that the basic ingredients of nature are tiny vibrating strings, rather than point particles. The strings could either be closed loops, or just open strands with two distinct endpoints. As a string moves through spacetime, it sweeps out a two-dimensional surface called a **worldsheet.** Similar to the worldline of a particle, the worldsheet represents the string’s complete history. It is, however, important to note that the worldsheet is not really a physical sheet floating in space. Rather, it is a mathematical representation used to calculate how a given string moves and vibrates in spacetime or how different strings interact, split, and join. For closed strings, the worldsheet surface resembles a tube or cylinder, whereas open strings produce a sheet with edges traced by its endpoints, as shown in the image below.

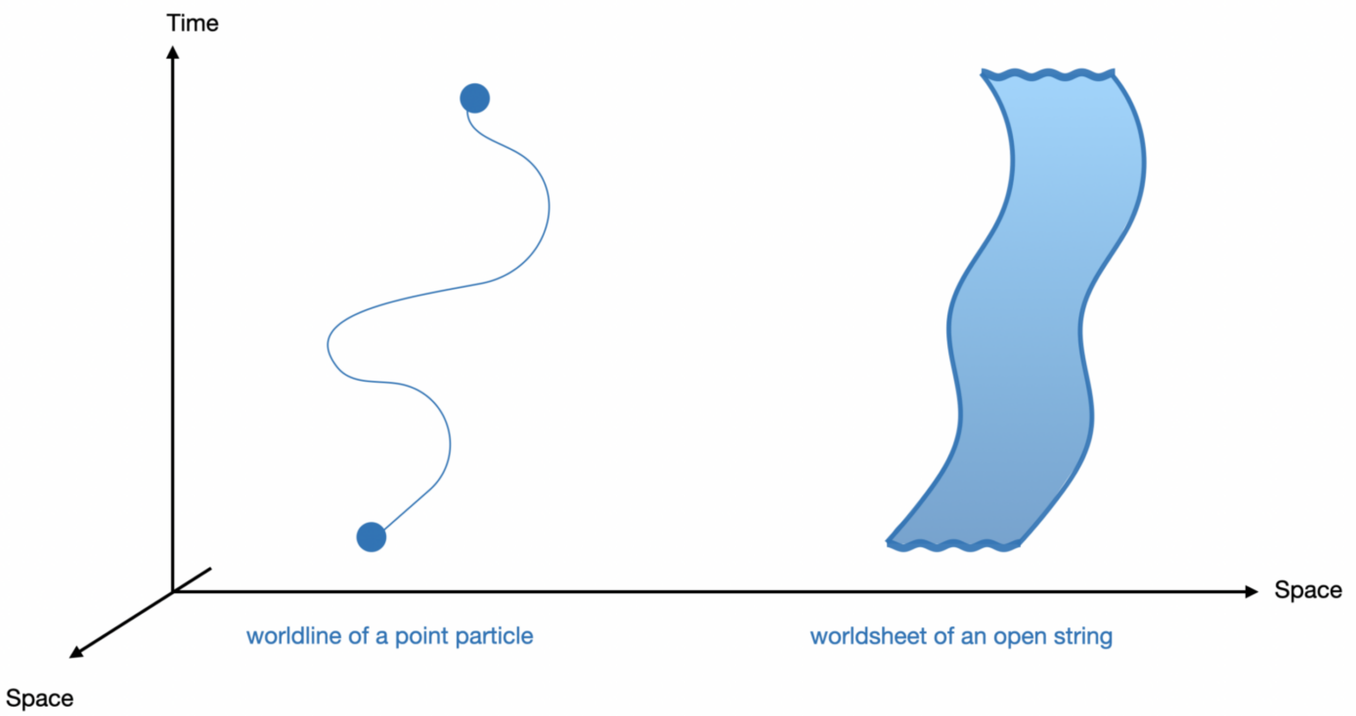

Figure 6. Modelling time evolution in physics. The entire trajectory, i.e. a history of positions, of a point particle moving through spacetime can be represented by its worldline. For a propagating string-like object, the history of all its positions sweeps out a smooth two-dimensional sheet. The points on the boundaries can be identified with the real number line.

For ordinary open strings, the boundary of this worldsheet can be represented by the real number line. Contrary to this, in *p*-adic string theory, the positions along the edges of the worldsheet get labelled by *p*-adic numbers. This gives the worldsheet a remarkably different, ultrametric structure: instead of a smooth, continuous surface, it is a discrete, endlessly branching network, like the Bruhat–Tits tree. In this sense, *p*-adic string theory provides us with discretised versions of the usual worldsheet picture.[1] This does, however, not mean that p-adic string theory has proven spacetime to be a discrete tree. Rather, one can view it as a toy model to understand some fundamental aspects of more realistic theories of quantum gravity. In particular, this ultrametric geometry is interesting because it allows us to formulate ideas about how smooth spacetime geometry might emerge from more elementary microscopic units. Moreover, Bruhat–Tits trees have also been used in recent simplified models of holography, i.e. the idea that there is a connection between gravity inside a space and a quantum theory on its boundary.[3] Among experts, this is known as the p-adic version of the AdS/CFT conjecture, which was previously described in this article.

These examples from biology and physics show why ultrametricity, and the p-adic numbers, are more than counterintuitive mathematical curiosities. While the familiar real numbers are well-suited to describing continuous quantities and distances, they are not necessarily the most natural language for every kind of structure. From evolutionary trees to disordered magnets: when relationships are branching, nested, or hierarchical, rather than along a simple line or within a smooth space, the p-adic notion of closeness may provide a more fitting description. A seemingly small change in how we define distance between numbers can therefore open up an entirely different way of modelling patterns in the world around us.

References

[1] Brekke, L., & Freund, P. G. (1993). p-Adic numbers in physics. *Physics Reports*, *233*(1), 1-66.

[2] Rammal, R., Toulouse, G., & Virasoro, M. A. (1986). Ultrametricity for physicists. *Reviews of Modern Physics*, *58*(3), 765.

[3] Gubser, S.S., Knaute, J., Parikh, S., Samberg, A. and Witaszczyk, P., 2017. p-adic AdS/CFT. *Communications in Mathematical Physics*, *352*(3), pp.1019-1059.