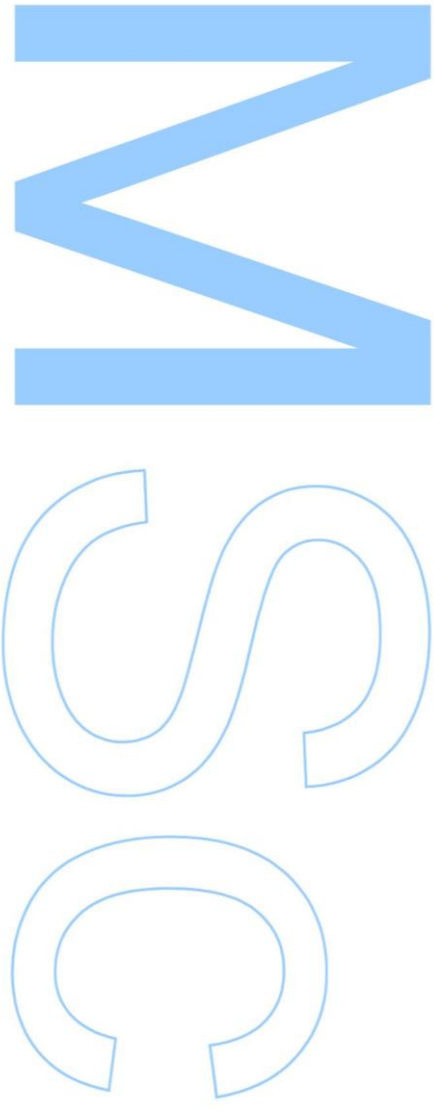
The Collatz function and some of its variants

Diogo Botelho
Mestrado em Matemática
Departamento de Matemática
2022

Orientador
António Machiavelo, Professor Auxiliar, Faculdade de Ciências



U.PORTO

FACULDADE DE CIÊNCIAS
UNIVERSIDADE DO PORTO

Todas as correções determinadas pelo júri, e só essas, foram efetuadas.

O Presidente do Júri,

Porto, _____/_____/_____

UNIVERSIDADE DO PORTO

MASTERS THESIS

The Collatz function and some of its variants

Author:
Diogo BOTELHO

Supervisor:
António MACHIAVELO

A thesis submitted in fulfilment of the requirements
for the degree of MSc. Mathematics

at the

Faculdade de Ciências da Universidade do Porto
Departamento de Matemática

December 19, 2022

Declaração de Honra

Eu, Diogo Gonçalo Maia Botelho, inscrito no Mestrado em Matemática da Faculdade de Ciências da Universidade do Porto declaro, nos termos do disposto na alínea a) do artigo 14.º do Código Ético de Conduta Académica da U.Porto, que o conteúdo da presente dissertação reflete as perspetivas, o trabalho de investigação e as minhas interpretações no momento da sua entrega.

Ao entregar esta dissertação declaro, ainda, que a mesma é resultado do meu próprio trabalho de investigação e contém contributos que não foram utilizados previamente noutros trabalhos apresentados a esta ou outra instituição.

Mais declaro que todas as referências a outros autores respeitam escrupulosamente as regras da atribuição, encontrando-se devidamente citadas no corpo do texto e identificadas na secção de referências bibliográficas. Não são divulgados na presente dissertação quaisquer conteúdos cuja reprodução esteja vedada por direitos de autor.

Tenho consciência de que a prática de plágio e auto-plágio constitui um ilícito académico.

Diogo Gonçalo Maia Botelho
Porto, 10 de outubro de 2022

Acknowledgements

First and foremost, I would like to thank my supervisor António Machiavelo for all the patience and encouragement towards me in these couple years. Never had given me classes before, but still kindly accepted to be my supervisor. Moreover, I am grateful for never giving up on me despite my slow and unsteady progress.

Secondly, it is paramount to thank my family for all the support given to me. My parents deserve a special place here, always asking me how the writing was going , encouraging me to keep moving forward and understanding my decision to extend the deadline for another year when I was diagnosed with a mild depression. To them, I am eternally grateful. I must also refer my paternal grandfather, which always wanted to know more about the phase in which I was regarding the thesis, and always motivated me to finish it. Concerning the rest of my family, I thank them for all the support given to me in this process.

Last but not least, I have to thank my close friends, specially Afonso, who always wanted to know how I was doing and motivated me to keep moving forward.

UNIVERSIDADE DO PORTO

Abstract

Faculdade de Ciências da Universidade do Porto
Departamento de Matemática

MSc. Mathematics

The Collatz function and some of its variants

by Diogo BOTELHO

Since the first half of the 20th century, the Collatz conjecture has established its place in the mathematical universe and since then it has been surprising a lot because it has not yet been solved, nor has there been progress in this direction, despite that even a child could understand what it states.

In this work, an approach to the conjecture is done with the help of the p-adic numbers, whose origin, construction and most important properties are presented, according to the introductory book to these numbers written by Fernando Gouvêa [11] .

An analysis of the conjecture itself from the 2-adic point of view is done based on the articles by Daniel Bernstein [1] and Ethan Akin [2]. In the ring of the 2-adic integers, the conjecture is equivalent to the restriction of a nowhere differentiable permutation to the thirds of the integers containing the natural numbers.

Keywords: Collatz conjecture, p-adic numbers

UNIVERSIDADE DO PORTO

Resumo

Faculdade de Ciências da Universidade do Porto
Departamento de Matemática

Mestrado em Matemática

A função de Collatz e algumas das suas variações

por Diogo BOTELHO

Desde a primeira metade do século XX que a conjetura de Collatz paira no universo matemático e desde então muito tem surpreendido por ainda não ter sido resolvida, nem haverem progressos nesse sentido, apesar de até uma criança conseguir entender o que ela afirma.

Neste trabalho, é feita uma abordagem à conjetura com a ajuda dos números p-ádicos, cuja origem, construção e propriedades mais importantes são apresentadas, de acordo com o livro introdutório a estes números de Fernando Gouvêa [11].

A análise à conjetura em si do ponto de vista 2-ádico é feita com base nos artigos de Daniel Bernstein [1] e Ethan Akin [2]. No anel dos 2-ádicos, tal conjetura é equivalente à restrição de uma permutação não diferenciável aos terços dos números inteiros conter os números naturais.

Palavras-chave: Conjetura de Collatz, números p-ádicos

Contents

Declaração de Honra iii

Acknowledgements v

Abstract vii

Resumo ix

Contents xi

1 Introduction 1

2 Problem's history 5
2.1 Origin and denomination . . . 5
2.1.1 When was the problem introduced and by whom? . . . 5
2.1.1.1 Designations of the problem . . . 9
2.1.2 Easy to understand versus hard to crack . . . 11
2.1.3 Properties inherited by the already computed Collatz orbits . . . 12
2.1.4 How can the conjecture be proved? . . . 13
2.2 Better partial known results . . . 14

3 p-adic numbers 17
3.1 Absolute values on a field . . . 22
3.1.1 Topology . . . 28
3.1.2 Algebra . . . 35
3.1.3 The case of $\mathbb{Q}$ . . . 37
3.1.3.1 Absolute values . . . 37
3.1.3.2 Completions . . . 42
3.2 Exploring $\mathbb{Q}_p$ . . . 52
3.2.1 Elementary analysis in $\mathbb{Q}_p$ . . . 54
3.2.1.1 Sequences and Series . . . 54
3.2.1.2 Functions: continuity and derivatives . . . 55

4 p-adic formulation of the Collatz Conjecture 57
4.1 The Two-adic Integers . . . 57
4.1.1 Construction . . . 57
4.1.2 Congruences . . . 61

4.1.3 The Two-adic Shift Map 65
4.1.4 The Two-adic $3x + 1$ Map 67
4.2 Applications 72

Bibliography **75**

Chapter 1

Introduction

The main goal of this thesis is to describe the easy to understand and infamous Collatz Conjecture with help of the p-adic numbers. The focus will be in two articles by Daniel Bernstein [1], and Ethan Atkin [2], which cover this problem using a 2-adic point of view.

Firstly, let us state the conjecture in its usual form. Given any positive integer n, if n is even, then we divide it by 2, and if n is odd, we triple it and add one. To the resulting number, we apply the same rules and so on repeatedly. The conjecture states that, no matter the initial choice of n, one will always end up in the loop $1 \to 4 \to 2 \to 1$, where $a \to b$ means that b is the image by a following the previous rules.

Let us compute the sequence of integers following the previous rules for small starting numbers of n. There is no reason to choose $n = 1, 2, 4$, as one can see above. For the starting number of $n = 3$, one gets:

$$3 \to 10 \to 5 \to 16 \to 8 \to 4 \to 2 \to 1 \to 4 \to 2 \to 1 \to \cdots$$

Now, 5 already appeared in the previous sequence so one might choose $n = 7$, getting:

$$7 \to 22 \to 11 \to 34 \to 17 \to 52 \to 26 \to 13 \to 40 \to 20 \to 10 \to 5 \to \cdots$$

and the rest is already known from the above computations. Now, notice that $\frac{3\times 9+1}{4} = 7$ which already appeared. The same is immediately true for 11 and 13. For $n = 15$, one gets:

$$15 \to 46 \to 23 \to 70 \to 35 \to 106 \to 53 \to 160 \to 80 \to 40 \to \cdots$$

It is clear that as n gets bigger, bigger integers appear in the sequence, but for now all sequences seem to end up in the referred loop. One should not try to compute more

examples by hand in order to disprove the conjecture, because it is pointless (a glimpse of what has already been proven regarding the conjecture will be seen later).

Back to the above rules, in other words, define a function f on the positive integers by:

$$f(n) = \begin{cases} \dfrac{n}{2} & \text{if n is even,} \\ 3n+1 & \text{if n is odd.} \end{cases}$$

This is the so-called *Collatz function*. Notice that by writing n in its binary expansion, say, $n = n_0 \times 2^0 + n_1 \times 2^1 + \cdots + n_k \times 2^k$, with $n_i \in \{0,1\}$ for all $i \in \{0, \ldots, k\}$, its image under f will be the first branch if $n_0 = 0$ and the second one otherwise.

In fact, this way of approaching the problem will play a major role in this thesis, since we will be focused in infinite extensions of binary sequences in order to state an equivalent formulation of the conjecture (we will see them later).

Returning to the problem, given any $n \in \mathbb{N}$, one can form the sequence

$$n, f(n), f^2(n), \ldots$$

where for all $i \in \mathbb{N}$, $f^i(n)$ is the i-th iterate of n by f, so that the conjecture being true is equivalent to

$$\forall\, n \in \mathbb{N}\ \exists\, m \in \mathbb{N} : f^m(n) = 1. \tag{1.1}$$

That is, every sequence of the iterates of a natural number by the Collatz function ends up in the loop $1 \to 4 \to 2 \to 1$. This sequence will be called the *Collatz orbit* of n.

Noticing that $3n+1$ is even for every odd n, the Collatz function can be simplified as

$$f(n) = \begin{cases} \dfrac{n}{2} & \text{if n is even,} \\ \dfrac{3n+1}{2} & \text{if n is odd.} \end{cases}$$

In fact, one could push this much further, by dividing at each step n or $3n+1$ (depending on the case) by the highest power of 2 in the factorization of that number, so that the Collatz orbit would only consist of odd numbers, eventually excepting the first one. This approach will be mentioned later, but for now we will maintain our previous form.

As mentioned before, given any sequence of positive integers $(a_0, a_1, \ldots, a_m)$, we say that it is a *cycle* or a *loop* with respect to the Collatz function if $f(a_0) = a_1, \ldots, f(a_{m-1}) = a_m, f(a_m) = a_0$, so that considering the previous shortcut of the Collatz function, the conjecture being true implies that $(1,2)$ is the only known cycle with respect

to the function, which is called the *trivial cycle*. Notice, however, that this consequence, if it were true, is not enough to prove the conjecture, as there could be some sequence increasing without bound towards infinity.

It remains to briefly explain how we are going to handle this problem in terms of p-adic numbers. At this point, they can only be introduced in a informal way, nevertheless with enough relevant information to give a glimpse of what one can expect in the future.

One can think of a p-adic number, associated to the prime number p, as a finite or infinite sequence of digits taking values from 0 to $p-1$. The sets of this kind with some well-defined operations will form a field, whose proof will be somewhat analogous to the construction of $\mathbb{R}$ from $\mathbb{Q}$.

In the particular case of $p = 2$, a conjecture regarding an interesting function defined in a restricted ring of the whole field will be surprisingly equivalent to the Collatz conjecture. The Collatz Conjecture's complexity will become patent in this point of view.

Chapter 2

Problem's history

2.1 Origin and denomination

2.1.1 When was the problem introduced and by whom?

"The Simple Math Problem We Still Can't Solve", as stated in the Quanta Magazine [3], has an uncertain origin. In fact, in mathematics, sometimes the first person to think about a problem is not the one who gives his name to it. One of the reasons for this to happen is that the one who publicizes the problem the most is thought to be its original creator. Despite of that, there is some evidence that the Collatz Conjecture was first introduced by the German mathematician Lothar Collatz.

Lothar Collatz (1910 – 1990) was a German mathematician and the Collatz Conjecture is named after him [6]. His interest in graphical representations of iterations of functions emerged in 1928 [4]. Indeed, according to himself, he took lecture courses in the following two years, where the latter included some graph theory. At that point in his career, Collatz had the idea of representing number-theoretic functions as directed graphs in the usual way: there is an edge from a vertex x to a vertex y, which are natural numbers, if the second one is the image of the first one under the respective function. By doing this, there is an immediate connection between the structure of such graphs and the behavior of the iterates of the underlying function. What is known after that is that in Collatz's notebook, see [5], dated July 1, 1932, Collatz was studying the function defined in the

natural numbers by

$$g(n) = \begin{cases} \dfrac{2n}{3} & \text{if } n \equiv 0 \pmod 3; \\ \dfrac{4n-1}{3} & \text{if } n \equiv 1 \pmod 3, \\ \dfrac{4n+1}{3} & \text{if } n \equiv 2 \pmod 3. \end{cases}$$

It is easy to see that this function is well-defined, that is, g sends natural numbers into natural numbers. At first glance the images under g seem to be a little bit complicated. In order to simplify those, one can look at a natural number m which is either even, or his congruence class modulo 4 is either 1 or 3. Then, m can be written as $2n$, $4n+1$ or $4n+3$ respectively, for some $n \in \mathbb{N}_0$. Now consider the function h on $\mathbb{N}$ defined by

$$h(m) = \begin{cases} 3n & \text{if } m = 2n; \\ 3n+1 & \text{if } m = 4n+1; \\ 3n+2 & \text{if } m = 4n+3, \end{cases}$$

which is sending congruence classes modulo 4 to congruence classes modulo 3, such that the two congruence classes modulo 4 corresponding to even numbers have the same type of image. The idea of writing m as $2n$ allows us to differentiate $2n$ from $2n+2 = 2(n+1)$. Notice also that since $\mathbb{N} = \{1 = 3 \times 0 + 1,\ 2 = 3 \times 0 + 2,\ 3 = 3 \times 1, \ldots\}$, it is clear that h is surjective. This function's injectivity follows immediately from looking at how it was defined. This means that h is invertible. Now, since

- $2n = \frac{6n}{3} = \frac{2(3n)}{3} = g(3n)$,
- $4n+1 = \frac{12n+3}{3} = \frac{4(3n+1)-1}{3} = g(3n+1)$ and
- $4n+3 = \frac{12n+9}{3} = \frac{4(3n+2)+1}{3} = g(3n+2)$,

it is clear that g is the inverse function of h. In particular, g is a permutation of the natural numbers. One can look at the orbit of 8, that is, the sequence formed by its iterates under the function. The first terms are:

$$8 \to 11 \to 15 \to 10 \to 13 \to 17 \to 23 \to 31 \to 41 \to 55 \to \cdots$$

At this point it is unclear what will happen to the next iterates. Until today, it is not known if the iterates will become predictable at some point [5]. Notice that this is not the case for other initial values, for instance:

- $g(1) = 1$, that is, 1 is a fixed point of g. In fact, 1 is the only fixed point of g;
- $g(2) = 3$ and $g^2(2) = 2$, so that 2 and 3 form a 2-cycle;
- The orbit of 4 consists of $4 \to 5 \to 7 \to 9 \to 6 \to 4$, and then it repeats. Notice that 9 is already bigger than 8, so that even if the orbit of 8 diverges, it is not true that the orbits of all the initial numbers bigger than 8 also diverge.

Since g is a bijection, what we do know is that 8 is the smallest number in its orbit, even if we consider its pre-images, because as we can see above, all other smaller numbers' images and pre-images have already been computed. Thus, if this orbit is not infinite, due to the fact that g is a permutation, it must contain a finite cycle which includes 8 and all its iterates written above. On the other hand, if it is infinite, does it contain all natural numbers except for the eight numbers mentioned in the items above? This is one of the unanswered questions concerning the g function.

Now, when does the Collatz function appear? The most reasonable answer is that Collatz wanted to simplify the previous function, since it has many branches and leads with natural numbers modulo 3, without losing its complexity of dynamics, and so the Collatz function may have appeared in the 1930's, probably in 1937 [7].

Although g is not the Collatz function, Lagarias calls the study of the iterates of g the "original Collatz problem" [5]. Similarly to the Collatz Conjecture, which is strongly believed to be true, the orbit of 8 under g is believed to be infinite.

Even though not clear, the Collatz function seems to provide at first glance an easier way of looking at the complexity of the dynamics of the g function. But due to g being the inverse function of h and h being similar to the Collatz function, there are several things one should notice in order to understand better the mentioned simplification.

- Concerning g, it is believed that the orbit of 8 diverges, on the contrary to the Collatz function where it is believed that any orbit converges to the trivial cycle.
- Neither of those beliefs being true implies the other being true: there is no immediate relation between the two functions and the behavior of their iterates.
- g is a permutation of the natural numbers. On the contrary, the Collatz function is not injective: every even number who is congruent to 1 modulo 3 has two pre-images.

There are some things, though, we can say about if they are false, for instance:

- In the case of g, if the orbit containing 8 is bounded, then its maximum element has to be a multiple of 3 (just look at how g is defined).
- In the case of the Collatz function, the smallest counterexample has to be an odd number congruent to 3 modulo 4, that is, the smallest natural number whose orbit does not reach 1.

Now, it seems to be consensual that no more information about the publicizing of this conjecture exists until the early 1950's. On the other hand, this decade embraced the rapid growth of this problem in the mathematical world.

According to Lagarias, during his talk at the International Congress of Mathematicians in Cambridge, Massachusetts in 1950, Collatz may have circulated the problem there [5]. With that being said, it is very unlikely that Bryan Thwaites discovered the problem in 1952, as he claimed [9] (one can look at who Bryan Thwaites is in the next section).

In a book translated to chinese by an unknown person, Collatz claims that he invented the problem, namely, he said that "Because I couldn't solve it I never published anything. In 1952 when I came to Hamburg I told it to my colleague Prof. Dr. Helmut Hasse. He was very interested in it. He circulated the problem in seminars and in other countries." [13].

According to John H. Conway and Richard Guy, the $3x + 1$ problem was already circulating in the university of Cambridge in the late 1950's [16].

Moreover, there is a letter dated 17 September, 1980, from Lothar Collatz to a mathematician named Michael Mays in which Collatz claims that he was the first to think about the $3x + 1$ problem. The letter was written in German, see [25] and [26], with a not so readable handwriting. The only problem about its traduction, see [27], is that the example $n = 80$ does not make any sense, since it trivially leads to the known cycle. The most reasonable alternative would be $n = 30$, which climbs up until 160 and then 80 until descending to the trivial cycle. Nevertheless, its orbit does not need difficult computations by hand, and if that were the case, $n = 15$ would make more sense, since it is odd and the next iterate of the Collatz function.

Resuming, it seems very likely that Collatz was the first one to consider the problem, probably in the 1930's, however since he did not make any progress towards the conjecture's goal, he failed to publicize it as soon as one could anticipate, only talking about it

years later to some of his colleagues and in an informal way. What is for sure certain is that the problem was already circulating in the mathematical community in the year 1952.

2.1.1.1 Designations of the problem

As one can anticipate, the diversity of names given to the conjecture is well explained. It was popularized by so many people in so many places, namely countries and universities.

The first publication concerning the problem only appeared in the 1970's. There are two main reasons for this, the first being part of the second. The arduous task of getting strong results concerning the problem did not make it very appealing to publicize, because who wants to talk about a problem which solution one knows little to nothing about? Moreover, during the 1960's, the mathematical community was under a strong influence of the Bourbaki-style mathematics, according to Lagarias [4]. The Collatz problem at that time did not meet any criteria according to that style to deserve being publicized, namely, the problem appeared to have no connection with other areas of mathematics, making it simply an onerous obstacle to surpass.

In fact, the Collatz Conjecture is also known as:

- $3n + 1$ problem, $3x + 1$ problem, $3n + 1$ conjecture or $3x + 1$ conjecture, due to the rule of the odd inputs;
- Ulam's conjecture or Ulam's problem. Stanisław Marcin Ulam (Lviv, April 13, 1909 – Santa Fe, May 13, 1984) was a Polish-American scientist working in the fields of mathematics and nuclear physics. He was one of the invited speakers at the International Congress of Mathematicians in Cambridge, Massachussets in 1950, which might have heard the problem at that time, and circulated it at Los Alamos and elsewhere, so the name Ulam's problem was popularized [5];
- Kakutani's problem. Shizuo Kakutani (Osaka, August 28, 1911 – New Haven, August 17, 2004) was a Japanese-American mathematician, best known for his eponymous fixed-point theorem. Just as Stanisław Ulam, he was one of the invited speakers there and was reportedly described the problem by Collatz himself [4]. If that did not happen to be the case, it is known that Kakutani already encountered the problem around 1960, because according to Lagarias [5], he said "For about a month everybody at Yale worked on it, with no result. A similar phenomenon happened when I mentioned it at the University of Chicago. A joke was made that this problem

was part of a conspiracy to slow down mathematical research in the U.S.", which is a pretty strong affirmation implying the hardness of the problem.

- Thwaites's conjecture or Thwaites's problem. Sir Bryan Thwaites, born on 6 of December of 1923, is an English applied mathematician, educator and administrator. He claimed to think about the problem for the first time in a hot summer day in 1952 where he was giving classes and his students were bored, so he had to entertain them for a while [12]. Thwaites even wrote a book in the 1990's where he called the problem "My Conjecture" [9].

- Hasse's algorithm. Helmut Hasse (Kassel, 25 August, 1898 – Ahrensburg, 26 December, 1979) was a German mathematician who worked in algebraic number theory, known for fundamental contributions to class field theory, the application of p-adic numbers to local class field theory and diophantine geometry (Hasse principle), and to local zeta functions. Recalling the previous alegated affirmation of Collatz, he explained the problem to Helmut Hasse in 1952, who become very interested in it and circulated the problem in other countries. The name Hasse's algorithm came up because Hasse discussed generalizations of the problem with many people [15].

- Syracuse problem. This name emerged from a visit by Hasse to the Syracuse university, where he publicized the problem and proposed this name to it [5].

It is also worth mentioning that the sequence of integers generated in the Collatz problem is sometimes referred to the hailstone numbers/sequence, due to the fact that values typically rise and fall, somewhat analogously to a hailstone inside a cloud. As one can recall, for a starting number of 7, the sequence is

$$7, 22, 11, 34, 17, 52, 26, 13, 40, 20, 10, 5, 16, 8, 4, 2, 1, 4, 2, 1, \ldots$$

While a hailstone eventually becomes so heavy that it falls to the ground, every starting positive integer ever tested has produced a hailstone sequence that eventually drops down to the number 1 and then consequently bounces into the small loop 4, 2, 1, ... [22].

Moreover, given any natural number n, it is called "wondrous" if the Collatz sequence starting at n eventually reaches 1. With that being said, the sequences associated to the Collatz problem are also known as "wondrous numbers". Thus, the Collatz Conjecture being true is equivalent to every positive integer being a "wondrous" number [21].

2.1.2 Easy to understand versus hard to crack

According to a presentation of Marc Chamberland in the 12th of April of 2013, at Grinnell College, this may be the most easily stated open problem! [12]. In fact, anyone who knows basic arithmetic operations between numbers can easily understand the problem. Moreover, given any positive integer, one can figure out what happens to its orbit under the Collatz function as long as one does not mess up in doing the computations (and sometimes be patient). Despite of that,"mathematics may not be ready for such problems", as said by mathematician Paul Erdös [5], that is, we may not be able to solve the problem using any resource known until this point: the hidden key to look for here is how to generalize the result to all positive integers, provided it has already been checked for a reasonable amount of positive integers as initial values. Erdös also offered a modest amount of five hundred dollars to whom could solve the conjecture. One reason for such a small prize in spite of his earlier commentary may be that even if the problem is solved, there has not been found an immediate correlation with other fields of mathematics, so that the problem is just an extremely arduous one, isolated, at least for now, from the rest of the world of mathematics.

As any other problem in mathematics that concerns something being true for all natural numbers, it has been checked for as big numbers as one can compute: as of 2020, all natural numbers up to 2^{68} have been verified to reach 1 after finitely many iterations of the Collatz function [17].

Concerning one specific number, apart from the trivial ones that we know for sure reach 1 after finitely many steps, that is, the powers of 2, and any power of 2 multiplied by a number which has already been checked to reach 1 after finitely many steps, the conjecture has already been proved for the number $2^{100000} - 1$: it needs 481603 times of $3x + 1$ computation, and 863323 times of $\frac{x}{2}$ computation [18]. This is the largest odd number for which the conjecture has been proved to hold and is a pretty much higher number that the previously mentioned ones, because many new algorithmic approaches to verify the conjecture for larger numbers are being optimized at every moment.

The fact that the conjecture has already been proved to be true for not so small numbers does not mean that the conjecture is true, as history has already told us (see Pólya's disproved Conjecture for instance [19], or Mertens disproved conjecture, whose counterexample is much higher [23]): the smallest counterexample may be a really big number, and we may have not been able to reach it yet. And even if there is a counterexample,

its Collatz orbit may diverge to infinity in a somewhat random way so that we may not be able to confirm that, that is, we may not find a pattern that justifies why it diverges. On the other hand, if a counterexample's Collatz orbit enters a non-trivial cycle, sooner or later we will find it. One has to be careful though, as the length of this cycle has to be greater than 17087915 as Elianouh proved in 1993 [20]. Indeed, he proved that the length q of that cycle has to be of the form

$$\begin{cases} q = 301994a + 17087915b + 85137581c, \\ b \geq 1,\ a, c \in \mathbb{Z} \text{ with } ac = 0. \end{cases}$$

As of 2020's computations, it is already known that q has to be greater than 114208327604, which can be written in the previous form with $b = 361$ and $c = 1269$, with the same source. There is no more information concerning the progress made so far in these computations [14].

2.1.3 Properties inherited by the already computed Collatz orbits

Let n be a positive integer and $f(n)$ be its image by the Collatz function. Let also

- $\mathrm{tst}(n) = \min\{k \in \mathbb{N} \mid f^k(n) = 1\}$, called the **total stopping time** of n;
- $\mathrm{st}(n) = \min\{k \in \mathbb{N} \mid f^k(n) < n\}$, called the **stopping time** of n;
- $\mathrm{st}_2(n) = \min\{k \in \mathbb{N}_0 \mid f^k(n) \text{ is a power of } 2\}$, called the **2-stopping time** of n;
- $\mathrm{p}_2(n) = f^{\mathrm{st}_2(n)}(n)$, that is, the first **power of** 2 appearing in the orbit of n,
- $f_{\max}(n) = \max\{f^k(n) \mid k \in \mathbb{N}_0\}$, that is, the maximum value on the collatz orbit of n.

Here we enumerate some relatively small examples for which the conjecture, as far as it is known, reaches its extreme values in terms of the previously mentioned values. For any $n, m \in \mathbb{N}$:

- $\mathrm{tst}(n) = m \Leftrightarrow \mathrm{st}_2(n) + \mathrm{tst}(\mathrm{p}_2(n)) = m$;
- 3 is the smallest number n for which $\mathrm{tst}(n) > n$. In fact, $\mathrm{tst}(3) = 7$;
- 7 is the smallest number n for which $\mathrm{st}_2(n) > n$. In fact, $\mathrm{st}_2(7) = 12$;

- $\text{tst}(n) > 4n \Leftrightarrow n = 27$: $\text{tst}(27) = 111 > 108 = 4 \times 27$ and the only other number m for which $\text{tst}(m) > 3m$ is $m = 31$, with $\text{tst}(31) = 106$;
- $\text{tst}(n) = n \Leftrightarrow n = 5$,
- 110 is the highest number n for which $\text{tst}(n) \geq n$ and 97 is the only number m smaller than 110 for which $\text{tst}(m)$ is bigger: $\text{tst}(97) = 118$.

So that from this point on, n is growing much faster compared to $\text{tst}(n)$. On the other hand, $\text{tst}(n)$ is growing much faster compared to $\text{tst}(2^m) = m$, where 2^m is the smallest power of 2 bigger than n, if n is not a power of 2.

2.1.4 How can the conjecture be proved?

In order to prove the Collatz Conjecture, there are several things one could do. Here we list the most reasonable ones:

1. Prove the non-existence of non-trivial cycles on any collatz orbit nor that it diverges for any starting natural number, which is the most natural approach, that is, the existence of these can only be the two reasons for why the conjecture cannot be true.
2. Find an upper bound on $\text{tst}(n)$ depending on n for every $n \in \mathbb{N}$. (this approach is reasonable due to the last property of the last section; for instance, proving $\text{tst}(n) < 5n$ for all $n \in \mathbb{N}$ is enough).
3. For all $n \in \mathbb{N}$, prove that $\text{st}(n)$ is finite. A small observation here: this is trivially true if $n \not\equiv 3 \pmod 4$: either n is even, or $n = 4m + 1 \implies f^2(n) = 3m + 1 < n$. The main result would follow easily by induction.
4. Letting G be the infinite digraph whose vertices are all the natural numbers and there is an edge from n to m if $f(n) = m$, then it is sufficient to prove that there is a path from any n to 1. This approach is strongly linked to Collatz's idea mentioned before.
5. Equivalently, for any $n \in \mathbb{N}$, it is enough to construct a path from 1 to n by either multiplying by 2 or subtracting 1 and dividing by 3.
6. Let us say a natural number m is *hard* if it is odd and $3m + 1$ is a power of 2. In order to prove the Collatz problem, it suffices to show that for every natural number n,

there is a hard number in is orbit, because the next iteration would be a power of 2, which we know reaches 1 as in many steps as the power itself. We can characterize these hard numbers: since $3m + 1 = 2^k$ for some k, then $m = \frac{2^k-1}{3}$, so that 2^k has to be congruent with 1 modulo 3, that is, k is even, and $m = \frac{4^\ell-1}{3}$ for some $\ell \in \mathbb{N}$.

7. Assuming by contradiction that the conjecture is false, by the well-order of $\mathbb{N}$, one may consider the smallest counterexample to the conjecture, which has to be odd, and try to reach some false statement. For instance, proving that there is a smaller number whose Collatz's orbit does not reach 1. One can also try to derive the properties of this number in terms of its factorization as a product of prime numbers.

2.2 Better partial known results

Concerning the conjecture, we now state, without proof, the more complete partial results known until this point. Whether if there is not a non-trivial cycle on any Collatz orbit or that if no orbit could increase without bound, there is nothing known.

Considering lower bounds, in a computer aided-proof, Krasikov and Lagarias showed that given any $x \in \mathbb{R}$, the amount of natural numbers in the interval $[1, x]$ whose orbits reach 1 is at least $x^{0.84}$ [24].

We shall also mention another form of progress thas has been made concerning the non-existence of non-trivial cycles. A k-cycle is a cycle that can be partitioned into $2k$ contiguous subsequences: k increasing sequences of odd numbers alternating with k decreasing sequences of even numbers. As of today, it is already known that there is no k-cycle for $k \leq 77$.

In 1976, mathematician Riho Terras proved that almost all natural numbers' orbits end up lower than them. Explicitely, it is known that the set of positive integers with finite stopping time has density one, that is,

$$F(k) \coloneqq \lim_{m\to\infty} \frac{1}{m} \mid \{n \leq m \colon \mathrm{tst}(n) \leq k\} \mid \text{exists for each } k \text{ and } \lim_{k\to\infty} F(k) = 1$$

This result has been improved just as recently as 2019 by Terence Tao, as stated by the Quanta Magazine: Tao "came away with one of the most significant results on the Collatz conjecture in decades." Indeed, Tao proved that for every not matter how slowly increasing sequence that goes off to infinity, the Collatz orbit of any natural number reaches a number smaller than the image of that number by any function of the mentioned type.

Tao said this result is "about as close as one can get to the Collatz conjecture without actually solving it.", so that it just takes any improvement on his result to prove the conjecture, even though that step may be unattainable. For detailed information, see [10].

Chapter 3

p-adic numbers

The book of Fernando Gouvêa, entitled "p-adic Numbers: An Introduction" , see [11], will be one's guide in the beginning of this theory. It talks about the universe of the p-adic numbers, which can be viewed as the halfway between algebra and analysis. Whatever concerns the p-adic numbers has been playing a very important role in modern number theory in the previous decades. For now, we can grossly say that the p-adic universe provides us with another language of talking about congruences between integers. Together with calculus and analysis' aid, one can deeply study the problems concerning the congruences between integers.

The p-adic numbers are introduced in the book as directly analogous to the real numbers: they are simply the field containing the rational numbers as a subfield that is complete with respect to some absolute value, where, if we take the standard absolute value, we simply get the field of real numbers, but we can consider many more absolute values, each one of them having a strong connection with a prime number.

Indeed, the origin of the p-adic numbers comes from German mathematician K.Hensel, who wanted to establish an analogy between the ring of integers together with its field of fractions and the ring of polynomials with complex coefficients together with its field of fractions as well. The latter contains the quotients of two polynomials with coefficients in the complex numbers, where the polynomial in the denominator cannot be the 0 polynomial, that is, the constant polynomial equal to 0, so that one can always compute the value of the rational function in almost any point, excluding the roots of the polynomial in the denominator. Similarly for the rational numbers, where the denominator cannot be 0.

One already knows from ring theory that both the ring of integers and the ring of such polynomials are unique factorization domains, so that the analogy mentioned above relates any prime number to any monic polynomial of first degree, that is, the linear polynomials. Notice that in the case of integers, one also has to consider the sign of the number, which is something one does not need to account for in the other ring.

There is so much about to be explored about this analogy. Indeed, now starting with the properties of the latter ring and field, one already knows from calculus that any polynomial with complex coefficients can be written in a form for each complex number which is called its Taylor's expansion. Using the previous starting point of the analogy, any positive integer can be written in base p for each prime p, which is the analogous to the Taylor expansion. Notice, however, that in the first case, the coefficients of the powers of the linear polynomial are allowed to be any complex numbers, although in the latter case, the coefficients of the prime powers are integers modulo p, so that there is a much stronger condition in writing an integer in base p.

These two ways of writing not only a polynomial but also a positive integer are indeed useful due to the fact that they give one local information regarding if the polynomial is divisible by the linear polynomial, that is, if the root corresponding to that linear polynomial is a root of the first polynomial, or if the positive integer is divisible by the prime, and in any case, to what order that happens.

Now, recalling complex analysis, one can in fact go much further to what concerns the Taylor's expansion. Indeed, given any rational function of that type, for every complex number, one can write it as a (usually infinite) sum of powers of the linear polynomial, finite-tailed to the left, so that one may take negative exponents. This case will only happen when, considering that the polynomials in the numerator and denominator are coprime in the usual way, if that complex number is a root of the denominator. Indeed, one only needs negative exponents for a finite number of complex numbers. If that is the case, one says that the rational function has a pole at that complex number.

Now, the first observation that is worth mentioning is that the Laurent expansion will be finite if and only if the polynomial in the denominator is a power of the chosen complex number.

Proposition 3.1. *Let* $f(X) = \frac{P(X)}{Q(X)} \in \mathbb{C}(X)$ *with* $P(X), 0 \neq Q(X) \in \mathbb{C}[X]$ *be in lowest terms, so that* $P(X)$ *and* $Q(X)$ *have no common zero,* $\alpha \in \mathbb{C}$ *and* $\sum_{n=n_0}^{\infty} a_n(X-\alpha)^n$ *be the Laurent*

expansion of $f(X)$ at α. Then

The expansion of $f(X)$ in powers of α is finite if and only if $Q(X) = (X-\alpha)^t$ for some $t \in \mathbb{N}_0$.

Proof. (only if) Suppose that the expansion of $f(X)$ in powers of α is finite, that is, there exists $m \in \mathbb{N}$ such that $a_{n_0+\ell} = 0$ for all $m < \ell \in \mathbb{N}$. We may assume that $n_0 < 0$, for, otherwise, $f(X)$ is a polynomial, thus equal to $P(X)$, so that $Q(X)$ is the constant polynomial equal to 1, which is the trivial power of $(X-\alpha)$. Then,

$$\begin{aligned} f(X) &= a_{n_0}(X-\alpha)^{n_0} + \cdots + a_{n_0+m}(X-\alpha)^{n_0+m} \\ &= a_{n_0}(X-\alpha)^{n_0} + \cdots + a_{n_0+m}(X-\alpha)^{m}(X-\alpha)^{n_0} \\ &= (X-\alpha)^{n_0}(a_{n_0} + \cdots + a_{n_0+m}(X-\alpha)^{m}) \\ &= \frac{a_{n_0} + \cdots + a_{n_0+m}(X-\alpha)^{m}}{(X-\alpha)^{-n_0}}, \end{aligned}$$

so that $Q(X)$ is written in the desired form, taking $t = n_0$.

(if) The exact same sequence of equalities, written from the bottom to the top, allows us to show the other implication, whence the proof is complete. □

The ideals generated by $(X-\alpha)$ are exactly the prime ideals of the ring $\mathbb{C}[X]$, so that in that sense one may view the linear polynomials as primes in that ring.

Now, from an algebraic point of view, one may view the set of such Laurent expansions at α, for fixed α, with the usual operations of sum and product as a field, which is easy to prove.

We want the analogy to go further so that we can accomplish something similar by writing any rational number in base p and agrouping all such expasions for every prime p.

The first thing to do is the obvious thing to do: just write the numerator and denominator of any rational number in base p. But now what one is supposed to do is not that immediate and is also tricky due to the more strict restrictions we have about the coefficients. Let us look at an example.

Example 3.1. *Take the rational number $\frac{24}{17} > 1$, and take $p = 3$. One has*

$$\begin{cases} 24 = 0 + 2x3 + 2x3^2 = 2p + 2p^2; \\ 17 = 2 + 2x3 + 1x3^2 = 2 + 2p + p^2, \end{cases}$$

so that $\frac{24}{17} = \frac{2p+2p^2}{2+2p+p^2}$.

At first, we may think that the way to go is to divide the numerator by the denominator in base p, but then we realize that bigger powers of p come into play, and our intuition tells us to stop there or we will be climbing too high with no clear end. Unexpectedly, this is really the way to go (the key here is that we will be able to predict the coefficients of the higher powers, so that there is no need to perform the long division indefinitely). Moreover, one has to accept this early the idea that the powers of a prime get smaller as the exponent grows, so that one can really cancel those powers to the right infinitely. With that being said, we shall return to our example. We are looking for a base p expansion that multiplied by the base p expansion of 17 gives us the base p expansion of 24 together with some other terms.

Now, it seems clear that the first parcel should be p, letting us with remainder p^3, which will be cancelled multiplying by p^3, so that the sum is $p^3 + 2p^3 = 3p^3 = p^4$ and so on....Indeed, we claim that

$$\frac{2p + 2p^2}{2 + 2p + p^2} = p + p^3 + 2p^5 + p^7 + p^8 + 2p^9 + \cdots$$

In order to check that this is correct, one may multiply the expansion of 17 by that infinite sum of powers of p, never forgetting that $p = 3$, so that one will be able to carry soon enough. Moreover, the result of the following products will be written in the ascending order of powers of p for convenient reasons. One gets $(2 + 2p + p^2)(p + p^3 + 2p^5 + p^7 +$

$$p^8 + 2p^9 + \cdots) = (*)$$

$$\begin{aligned}
(*) &= 2p + 2p^2 + \underbrace{p^3 + 2p^3}_{=p^4} + 2p^4 + p^5 + 4p^5 + 4p^6 + 2p^7 + 2p^7 + 2p^8 + 2p^8 + p^9 + 2p^9 + \cdots \\
&= 2p + 2p^2 + \underbrace{p^4 + 2p^4}_{=p^5} + p^5 + 4p^5 + 4p^6 + 2p^7 + 2p^7 + 2p^8 + 2p^8 + p^9 + 2p^9 + \cdots \\
&= 2p + 2p^2 + \underbrace{2p^5 + 4p^5}_{=2p^6} + 4p^6 + 2p^7 + 2p^7 + 2p^8 + 2p^8 + p^9 + 2p^9 + \cdots \\
&= 2p + 2p^2 + \underbrace{2p^6 + 4p^6}_{=2p^7} + 2p^7 + +2p^7 + 2p^8 + 2p^8 + p^9 + 2p^9 \cdots \\
&= 2p + 2p^2 + \underbrace{2p^7 + 2p^7 + 2p^7}_{=2p^8} + 2p^8 + 2p^8 + p^9 + 2p^9 + \cdots \\
&= 2p + 2p^2 + \underbrace{2p^8 + 2p^8 + 2p^8}_{=2p^9} + p^9 + 2p^9 + \cdots \\
&= \cdots \\
&= 2p + 2p^2.
\end{aligned}$$

Now, similarly to the rational functions case and reasonably assuming that our rational number is in lowest terms, as it always has been, the smallest exponent of the expansion will be negative if and only if the prime number divides the numerator, and will be 0 if both the numerator and denominator are coprime with that prime.

In fact, calling n_0 the smallest exponent of the expansion, we can write the rational number $\frac{a}{b}$ as

$$\frac{a}{b} = p^{n_0}\frac{a_1}{b_1} \text{ with } (p, a_1 b_1) = 1.$$

This works now for any positive rational number, so that it suffices to understand how it may work for negative rational numbers. For that, recalling the trick of cancelling bigger powers by adding what is left to attain the next power, we can clearly write -1 in base p: such expansion will be such that when we add 1 to it, we get 0. Hence, the most reasonable option should be

$$-1 = (p-1) + (p-1)p + (p-1)p^2 + \cdots.$$

Adding 1 to this will consecutively cancel all coefficients equal to $p-1$ by putting them equal to p, which gives the next power.

Analogously to the functions case, this expansion will be usually infinite (to the right,

of course), being finite only if the rational number is a non-negative integer, for obvious reasons.

Now, to return to the main focus, which is to establish a strong connection in the so referred analogy, we may as well consider the set of all "finite-tailed Laurent series in powers of p", which we will briefly call p-adic expansions. This will also be a field, although that is a more tricky result to prove because of carrying the coefficients when we add or multiply two of those expansions. It is not easy to write the general result, but we can do it step by step and comprehend what is going on when the exponent is large enough.

Indeed, we only wrote -1 in base p, so that it "only" suffices to write all other negative rational numbers in base p. Realizing the triviality of this process for any negative integer (we just need to find an expansion such that adding its symmetrical's expansions leads to 0), the same holds for non integers, a difference can only be in the first exponent of its symmetrical, that could be negative, and that the symmetrical's expansion is infinite. But just in the case of a non integer positive rational, the same argument still applies.

Now we may think of these expansions to be in some way periodic, relating them even further to the real numbers, and that is in fact true, and the proof can be similar to the real numbers case.

Again, analogously to the real numbers, one should find a similar way to write rational finite numbers. In the real numbers, the rational finite numbers can be written as a sequence of decimal numbers, separated by a comma in the position where it is necessary to establish to what power of 10 the number is being multiplied.

The main goal of this chapter is to establish solid foundations on absolute values in order to understand the construction of the p-adic numbers. Throughout this chapter, $\mathbb{K}$ will be a field, $\mathbb{R}_+ = \{x \in \mathbb{R} : x \geq 0\}$ will denote the set of all non-negative real numbers and p will be a prime number.

3.1 Absolute values on a field

The usual notion of absolute value can be extended to any field $\mathbb{K}$ in the following way.

Definition 3.2. An absolute value on $\mathbb{K}$ is a function $|\cdot| : \mathbb{K} \to \mathbb{R}_+$ satisfying the following conditions for all $x, y \in \mathbb{K}$:

i) $|x| = 0$ if and only if $x = 0$;

ii) $|xy| = |x||y|$;

iii) $|x + y| \leq |x| + |y|$.

Additionally, if it satisfies the following stronger condition

iv) $|x + y| \leq \max\{|x|, |y|\}$,

the absolute value is called *non-archimedean*, and *archimedean* otherwise.

Now, let us look at a couple of examples of absolute values.

Examples 1.

- Take $\mathbb{K} = \mathbb{Q}$ and put

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

  This is the usual absolute value on the rational numbers, that is, the restriction of the usual absolute value in the real numbers to the rational numbers. It is archimedean, since the fourth condition above fails for $x = y = 1$, for example. We will denote it by $|\cdot|_\infty$ and call it the *infinite absolute value*, the reason why will be revealed later.

- Put $|x| = 0$ if $x = 0$, and $|x| = 1$ otherwise. This is clearly a non-archimedean absolute value in any field: the first condition is satisfied, the second also, since a field has no zero divisors, and the fourth one due to the fact that $|x + y| = 1$ implies $\max\{|x|, |y|\} = 1$. This absolute value is known as the *trivial absolute value*, because it is the simplest one.

From the previous definition, it is easy to deduce that the trivial absolute value is the only one in any finite field $\mathbb{K}$. This can be seen as follows: if $\mathbb{K}$ is a finite field, then there exists an integer q such that $x^q = x$ for any $x \in \mathbb{K}$ (for instance, we can take q to be the number of elements of $\mathbb{K}$). One has:

- $|0| \overset{\text{i)}}{=} 0$;
- If $0 \neq x \in \mathbb{K}$, then $x^q = x \implies |x^q| = |x| \overset{\text{ii)}}{\implies} |x|^q = |x| \overset{\text{i)}}{\implies} |x| = 1$.

Fortunately, we will be focused in infinite fields, where much more can be said.

Now, in order to come up with other absolute values, namely in the field of rational numbers, it is more convenient for us to introduce an interesting infinite family of functions which will be strongly linked to the soon to come, p-adic absolute values.

Definition 3.3. The p-adic valuation on $\mathbb{Z}$ is the function

$$v_p \colon \mathbb{Z} \setminus \{0\} \to \mathbb{R}$$

defined by: for $0 \neq n \in \mathbb{Z}$, $v_p(n)$ is the unique positive integer satisfying

$$n = p^{v_p(n)} z \text{ with } p \nmid z \text{ and } z \in \mathbb{Z}.$$

Moreover, we extend this function to $\mathbb{Q}$ by putting, for $x = \frac{a}{b} \in \mathbb{Q} \setminus \{0\}$,

$$v_p(x) = v_p(a) - v_p(b),$$

and it is convenient to set $v_p(0) = +\infty$, and work with the usual conventions when adding and multiplying by $+\infty$.

Indeed, v_p just counts how many times some rational number can be divided by the prime number p, so that the result is positive if the numerator is divisible by p, negative if the denominator is divisible by p and 0 otherwise (we may assume that the rational number is written in lowest terms, because it is easy to check that the valuation does not depend on the representation of the rational number as a quotient of integers). The reasoning for $v_p(0)$ being $+\infty$ is that 0 can be divided by p as many times as one wants, always yielding 0.

In fact, we can say that, for any $x \in \mathbb{Q} \setminus \{0\}$,

$$x = p^{v_p(x)} \frac{a}{b} \text{ with } p \nmid ab \text{ and } \frac{a}{b} \in \mathbb{Q}. \tag{3.1}$$

Examples 2. For instance,

- $v_2(37) = 0$, because 37 is odd;
- $v_3(\frac{123}{48}) = v_3(\frac{41}{16}) = 0$, because 41 and 16 are not multiples of 3;
- $v_5(\frac{-180}{7}) = v_5(\frac{180}{7}) = 1$, because $5 \mid 180$, $5^2 = 25 \nmid 180$ and $5 \nmid 7$,
- $v_7(\frac{1}{49}) = -2$, due to $7 \nmid 1$ and $49 = 7^2$.

Let us explore the basic properties of the function v_p, which we will use often.

Lemma 3.4. *For all $x, y \in \mathbb{Q}$, the following hold:*

i. $v_p(xy) = v_p(x) + v_p(y)$,

ii. $v_p(x+y) \geq \min\{v_p(x), v_p(y)\}$.

Proof. According to equation (3.1), one can write

$$x = p^{v_p(x)}\frac{a}{b} \text{ with } p \nmid ab \text{ and } \frac{a}{b} \in \mathbb{Q};$$
$$y = p^{v_p(y)}\frac{c}{d} \text{ with } p \nmid cd \text{ and } \frac{c}{d} \in \mathbb{Q},$$

implying

$$xy = p^{v_p(x)+v_p(y)}\frac{ac}{bd} \text{ with } p \nmid abcd \text{ (because } p \text{ is prime) and } \frac{ac}{bd} \in \mathbb{Q},$$

so that the first property is true. For the second, it is sufficient to notice that common powers of p, if any, can be factored out from a sum. □

These two properties are very similar to conditions ii) and iv) of definition 3.2. Indeed, the product in condition ii) has been turned into a sum in property i. (as when taking a logarithm) and the inequality in iv) has been reversed in ii. By changing the sign of the latter inequality and turning the sum into a product by putting it into an exponent, one manages to go the other way around.

These observations already help state the following definition:

Definition 3.5. For any $x \in \mathbb{Q}$, the *p-adic absolute value* of x is given by

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

Remark 3.6. In the case $x = 0$, the definition matches with the expected value $p^{-\infty}$.

Proposition 3.7. *Definition 3.5 gives a non-archimedean absolute value on* $\mathbb{Q}$.

Proof. Definition 3.5 already covers condition i) of definition 3.2; Lemma 3.4 i. implies the second condition, and using Lemma 3.4 ii., one gets, for $x, y \in \mathbb{Q}$,

$$\begin{aligned} |x+y|_p &\overset{\text{def}}{=} p^{-v_p(x+y)} \\ &\overset{\text{ii.}}{\leq} p^{-\min\{v_p(x),v_p(y)\}} \\ &= \max\{p^{-v_p(x)}, p^{-v_p(y)}\} \\ &= \max\{|x|_p, |y|_p\}, \end{aligned}$$

proving the fourth condition of Definition 3.2, and therefore, the third as well. □

In constrast with the usual and trivial absolute values, the p-adic absolute value is measuring how divisible by p a certain rational is. If it is very divisible by p, its absolute value is going to be small. This is due to the fact that

$$\forall\, 0 \neq x \in \mathbb{Q}\ ,\ |xp^n|_p \to 0 \text{ as } n \to \infty.$$

Let us compute some p-adic absolute values.

Examples 3.

- $|\frac{56}{11}|_7 = \frac{1}{7}$ (because $7 \mid 56$, $49 = 7^2 \nmid 56$ and $7 \nmid 11$);
- $|\frac{2}{625}|_5 = 625$ (since 2 and 5 are coprime and 625 is a power of 5),
- $|\frac{10}{19}|_3 = 1$ (due to $\gcd(10,3) = \gcd(19,3) = 1$).

Before studying the p-adic absolute value's properties, one will look at the properties of any absolute value.

Lemma 3.8. *For any absolute value $|\cdot|$ in a field $\mathbb{K}$, the following hold:*

i) $|1| = 1$;

ii) *If $x \in \mathbb{K}$, $n \in \mathbb{N}$ are such that $|x^n| = 1$, then $|x| = 1$;*

iii) $|-1| = 1$,

iv) $\forall\, x \in \mathbb{K}, |-x| = |x|$.

Proof. i) Just note that $|1| = |1^2| = |1||1|$ and $|1| \neq 0$, because $1 \neq 0$; ii) follows from $|x^n| = |x|^n$, iii) follows from $|(-1)^2| = |1| = 1$ and ii), and finally iv) follows from $|-x| = |-1||x|$ and iii). □

Now we will state and prove a necessary and sufficient condition for an absolute value to be non-archimedean. But first, we recall from ring theory that for any field $\mathbb{K}$, there is a map $\mathbb{Z} \to \mathbb{K}$ defined by, for $n \in \mathbb{Z}$:

$$n \to \begin{cases} \underbrace{1 + \cdots + 1}_{n} & \text{if } n > 0 \\ 0 & \text{if } n = 0 \\ -\underbrace{(1 + \cdots + 1)}_{n} & \text{if } n < 0 \end{cases}$$

For instance, if $\mathbb{K} \supseteq \mathbb{Q}$, this map is the inclusion of $\mathbb{Z}$ into $\mathbb{K}$. And if $\mathbb{K}$ is finite, then the image of the map is a subfield of $\mathbb{K}$ with a prime number of elements, due to the characteristic of $\mathbb{K}$ being finite, and thus, a prime.

Theorem 3.9. *Let A be the image of $\mathbb{Z}$ in $\mathbb{K}$. An absolute value $|\cdot|$ in $\mathbb{K}$ is non-archimedean if and only if $|a| \leq 1$ for all $a \in A$. In particular, an absolute value $|\cdot|$ in $\mathbb{Q}$ is non-archimedean if and only if $|n| \leq 1$ for all $n \in \mathbb{Z}$.*

Proof. Suppose that the absolute value $|\cdot|$ in $\mathbb{K}$ is non-archimedean. Using the non-archimedean property, it follows that for all $a \in A$,

$$|a+1| \leq \max\{|a|,|1|\} \overset{3.8\ i)}{=} \max\{|a|,1\}. \tag{3.2}$$

There are three cases to consider:

- If $a = 0$, then there is nothing to show, because $|a| = 0 \leq 1$;
- If $a = \underbrace{1+\cdots+1}_{n}$ for some $n \in \mathbb{N}$, then one can proceed by induction on n. If $n = 1$, then $a = 1$ and so $|a| = 1 \leq 1$. Now, suppose $n > 1$ and that $|\underbrace{1+\cdots+1}_{n-1}| \leq 1$. Using inequality (3.2) and the induction hypothesis, it follows that

  $$|a| = |(a-1)+1| = \max\{|a-1|,|1|\} \leq \max\{1,1\} = 1,$$

  which completes the induction step;
- Finally, if $a = -(\underbrace{1+\cdots+1}_{n})$ for some $n \in \mathbb{N}$, then $|a| = |n| \leq 1$ by the previous item.

Conversely, suppose that $|a| \leq 1$ for all $a \in A$. It suffices to prove that

$$\forall\, x, y \in \mathbb{K}, |x+y| \leq \max\{|x|,|y|\}.$$

This is trivial if $y = 0$. Otherwise, it is equivalent to show that

$$\left|\frac{x}{y}+1\right| \leq \max\left\{\left|\frac{x}{y}\right|, 1\right\},$$

for which it is enough to prove that

$$|x+1| \leq \max\{|x|,1\}.$$

To show this, let $m \in \mathbb{N}$. Using the binomial theorem, it follows that:

$$\begin{aligned}
|x+1|^m &= \left|\sum_{k=0}^{m} \binom{m}{k} x^k\right| \\
&\leq \sum_{k=0}^{m} \left|\binom{m}{k}\right| |x^k| \\
&\leq \sum_{k=0}^{m} |x|^k \quad \left(\text{since } \binom{m}{k} \in \mathbb{N}_0 \subseteq \mathbb{Z}, \text{ our assumption yields } \left|\binom{m}{k}\right| \leq 1\right) \\
&\leq (m+1)\, \max\{1, |x|^m\} \text{ (since } \max\{x^k : 0 \leq k \leq m\} \text{is either 1 or } |x|^m)
\end{aligned}$$

This inequality implies that

$$\begin{aligned}
|x+1| &\leq \sqrt[m]{m+1}\, \max\{1, |x|\} \\
&\leq \lim_{m\to\infty} \sqrt[m]{m+1}\, \max\{1, |x|\} \\
&= \max\{1, |x|\},
\end{aligned}$$

as we wanted to prove. □

Thus, an absolute value $|\cdot|$ is archimedean if and only if $|n| > 1$ for some $n \in \mathbb{N}$, which is equivalent to

$$\forall\, 0 \neq x, y \in \mathbb{K}\,, \exists\, n \in \mathbb{N} : |nx| > |y|.$$

This is sometimes called the *Archimedean Property* of an absolute value.

From this observation and the previous theorem, the next result follows.

Corollary 3.10. *An absolute value $|\cdot|$ on a field $\mathbb{K}$ is non-archimedean if and only if* $\sup\{|n| : n \in \mathbb{Z}\} = 1$.

Proof. The direct implication is clear. For its converse, suppose that the supremum is $C > 1$. Then $|m| > 1$ for some $m \in \mathbb{Z}$ and so $\lim_{k\to\infty} |m^k| = +\infty$, thus at some point $|m^k| > C$, contradicting C being finite. It follows that $C \leq 1$ and since $|1| = 1$, $C = 1$. □

3.1.1 Topology

Now, the whole point of introducing absolute values is that we can use them to measure the distance between numbers, since they naturally yield a metric in the underlying field. Having that, it is therefore possible to investigate the underlying topology of the field.

Definition 3.11. Let $|\cdot|$ be an absolute value on $\mathbb{K}$. The distance between two elements x, $y \in \mathbb{K}$ is defined by

$$d(x,y) = |x-y|,$$

and the function $d\colon \mathbb{K} \times \mathbb{K} \to \mathbb{R}_+$ is called the metric induced by the absolute value $|\cdot|$.

The following result is then useful.

Proposition 3.12. *d is a metric in $\mathbb{K}$, meaning that it satisfies the following conditions, for any $x, y, z \in \mathbb{K}$:*

i) $d(x,y) \geq 0$ *and* $d(x,y) = 0 \Leftrightarrow x = y$;

ii) $d(x,y) = d(y,x)$;

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

Proof. Analogous to the proof concerning the metric on $\mathbb{R}$. □

Moreover, this function d satisfies some other properties that make $\mathbb{K}$ a topological field, that is, the operations of addition, multiplication and taking inverses are all continuous functions. Despite this being just an easy exercise of calculus, it will not be helpful in the remaining thesis, so there is no point in writing down the proof, but nevertheless it is worth mentioning.

Since now there is a metric on $\mathbb{K}$, Theorem 3.9 can be stated in terms of d. That is, the fact that the absolute value is non-archimedean can also be expressed in terms of the metric.

Lemma 3.13. *$|\cdot|$ is a non-archimedean absolute value on the field $\mathbb{K}$ if and only if*

$$\forall\, x,y,z \in \mathbb{K},\ d(x,z) \leq \max\{d(x,y), d(y,z)\}.$$

Proof. The non-archimedean condition applied to the equation

$$x - z = (x-y) + (y-z)$$

yields the direct implication. For the converse, our assumption implies

$$d(x,-z) \leq \max\{d(x,0), d(0,-z)\},$$

which is equivalent to

$$|x+z| \leq \max\{|x|, |z|\},$$

which is the non-archimedean condition. □

The inequality in the previous lemma is called the *ultrametric inequality* and any metric space satisfying it is, therefore, called an *ultrametric space*.

One question arises naturally from the previous result: when is the equality achieved? A sufficient condition for that to happen is given below.

Proposition 3.14. *If* $\mathbb{K}$ *is a field with a non-archimedean absolute value* $|\cdot|$ *and* $x, y \in \mathbb{K}$ *are such that* $|x| \neq |y|$*, then*

$$|x + y| = \max\{|x|, |y|\}.$$

Proof. Without loss of generality, one may suppose $|x| > |y|$. Lemma 3.13 then yields $|x + y| \leq |x|$. The non-archimedean condition gives $|x| = |x + y + (-y)| \leq \max\{|x + y|, |y|\}$. Together with $|x| > |y|$, this implies $\max\{|x + y|, |y|\} = |x + y|$ and, thus, $|x| \leq |x + y|$. Hence, $|x + y| = |x|$, as we wanted to show. □

The abnormal properties of the p-adic world become evident in the following result.

Corollary 3.15. *In an ultrametric space* $\mathbb{K}$*, all "triangles" are isosceles, meaning that for every triplet of points in the space, at least two of the three distances between two of them are equal.*

Proof. Given $x, y, z \in \mathbb{K}$, the respective "triangle" has sides with lengths

$$d(x, y) = |x - y|;$$
$$d(y, z) = |y - z|,$$
$$d(x, z) = |x - z|.$$

Now, either $d(x, y) = d(y, z)$, which proves that this triangle is isosceles, or last proposition applied to the equation

$$(x - y) + (y - z) = x - z$$

yields $|x - z| = \max\{|x - y|, |y - z|\}$, and hence, the triangle is isosceles. □

This is rather an unintuitive result and it will have an enormous impact on the topology on our field.

We will take a look at the case of the p-adic absolute value in order to understand what is going on behind the last result.

To simplify our approach, we will only look at the case where $x, y \in \mathbb{Z}$. Say that $v_p(x) = n$, $v_p(y) = m$, so that

$$x = p^n x', y = p^m y' \text{ and } p \nmid x'y'.$$

Hence, $|x| = p^{-n}$ and $|y| = p^{-m}$. We will have $|x| > |y|$ when $n < m$, say, $m = n + t$, with $t > 0$. Then

$$x + y = p^n x' + p^{n+t} y' = p^n (x' + p^t y').$$

Now, since $p \nmid x'$, we have $p \nmid (x' + p^t y')$ and thus $v_p(x + y) = n = v_p(x)$, that is, $|x + y| = |x|$.

On the other hand, we either have $|x| = |y|$ or analogously to the first case, $|x + y| = |y|$, so that in either case, two of the three absolute values $|x|$, $|y|$ and $|x + y|$ are equal, as expected.

Examples 4. For instance, giving $\mathbb{Q}$ the 5-adic topology and considering the triangle whose vertices are $x = \frac{2}{15}$, $y = \frac{1}{5}$ and $z = \frac{7}{15}$ yields

$$\begin{aligned} |x - y| &= \left| \frac{2}{15} - \frac{1}{5} \right| = \left| \frac{-1}{15} \right| = 5; \\ |y - z| &= \left| \frac{1}{5} - \frac{7}{15} \right| = \left| \frac{-4}{15} \right| = 5, \\ |x - z| &= \left| \frac{2}{15} - \frac{7}{15} \right| = \left| \frac{-5}{15} \right| = \left| \frac{-1}{3} \right| = 1. \end{aligned}$$

Now, let us introduce the usual notions we are so used to when we are dealing with a metric space. These will turn out to be pretty strange in the case of an ultrametric due to the previous corollary.

Definition 3.16. Let $\mathbb{K}$ be a field with an absolute value $|\cdot|$. Let also $a \in \mathbb{K}$ and $r \in \mathbb{R}_+$. The *open ball of radius r and center a* is the set

$$B(a, r) = \{x \in \mathbb{K} \colon d(x, a) < r\} = \{x \in \mathbb{K} \colon |x - a| < r\}.$$

The *closed ball of radius r and center a* is the set

$$\overline{B}(a, r) = \{x \in \mathbb{K} \colon d(x, a) \leq r\} = \{x \in \mathbb{K} \colon |x - a| \leq r\}.$$

One thing that is true for any absolute value is that the open ball is always an open set and the closed ball is always a closed set. The proof is the one we saw when we first studied metric spaces in an elementary calculus course.

For a non-archimedean absolute value, which is the case of any p-adic absolute value, surprising properties arise due to "all triangles being isosceles".

Proposition 3.17. *Let $\mathbb{K}$ be a field with a non-archimedean absolute value $|\cdot|$, $a, b \in \mathbb{K}$ and $r, s \geq 0$. We have the folowing properties.*

i) $b \in B(a,r) \Rightarrow B(a,r) = B(b,r)$, *i.e, any point inside an open ball is a center of that ball;*

ii) $b \in \overline{B}(a,r) \Rightarrow \overline{B}(a,r) = \overline{B}(b,r)$, *i.e., every point inside a closed ball is a center of that ball;*

iii) The set $B(a,r)$ is both open and closed;

iv) If $r \neq 0$, the set $\overline{B}(a,r)$ is both open and closed;

v) If $r, s > 0$, we have $B(a,r) \cap B(b,s) \neq \emptyset$ if and only if $B(a,r) \subseteq B(b,s)$ or $B(b,s) \subseteq B(a,r)$; that is, any two open balls are either disjoint or contained in one another,

vi) If $r, s > 0$, we have $\overline{B}(a,r) \cap \overline{B}(b,s) \neq \emptyset$ if and only if $\overline{B}(a,r) \subseteq \overline{B}(b,s)$ or $\overline{B}(b,s) \subseteq \overline{B}(a,r)$; that is, any two closed balls are either disjoint or contained in one another.

Proof.

i) By definition, $b \in B(a,r) \Leftrightarrow |b-a| < r$. If $x \in B(a,r)$, the non-archimedean property yields

$$|x-b| \leq \max\{|x-a|, |b-a|\} < r,$$

so that $x \in B(b,r)$, thus $B(a,r) \subseteq B(b,r)$. Switching a and b, we get the opposite inclusion.

ii) Analogous to the proof of i): just replace $<$ by $\leq$.

iii) The open ball is always an open set in any metric space. It suffices to show that it is closed in the non-archimedean case. So take x in the boundary of $B(a,r)$, that is, any open ball centered in x contains points of $B(a,r)$. We want to show that $x \in B(a,r)$. We may assume $r \neq 0$, because otherwise the empty set is trivially closed. So, choose $0 < s \leq r$. Since x is a boundary point of $B(a,r)$, then $B(a,r) \cap B(x,s) \neq \emptyset$. Let y be in this intersection, which implies that $|y-a| < r$ and $|y-x| < s$. The non-archimedean inequality yields

$$|x-a| \leq \max\{|x-y|, |y-a|\} < \max\{s,r\} < r,$$

so that $x \in B(a,r)$, hence any boundary point of $B(a,r)$ is in $B(a,r)$, implying that $B(a,r)$ is a closed set.

iv) This is a similar to iii).

v) Without loss of generality, we may assume $r \leq s$. If the intersection is not empty, then there exists $c \in B(a,r) \cap B(b,s)$. From i), it follows that $B(a,r) = B(c,r)$ and $B(b,s) = B(c,s)$. Hence,

$$B(a,r) = B(c,r) \subseteq B(c,s) = B(b,s),$$

as claimed.

vi) This is identical to v), but using ii) instead of i). □

Let us take a look at some examples and properties of open and closed balls and sets, where the absolute value is the p-adic one on the field $\mathbb{Q}$, if nothing is said.

Examples 5.

- $\overline{B}(0,1) = \left\{ \frac{a}{b} \in \mathbb{Q} \colon \left| \frac{a}{b} \right|_p \leq 1 \right\} = \left\{ \frac{a}{b} \in \mathbb{Q} \colon p \nmid b \text{ , } \gcd(a,b) = 1 \right\}$;
- $B(3,1) = \left\{ \frac{a}{b} \in \mathbb{Q} \colon \left| \frac{a}{b} - 3 \right|_p \leq 1 \right\} = \left\{ \frac{a}{b} \in \mathbb{Q} \colon \left| \frac{a-3b}{b} \right|_p \leq 1 \right\}$. If $\frac{a}{b}$ is in lowest terms, then $\frac{a-3b}{3}$ is also in lowest terms, so that $\frac{a}{b} \in B(3,1) \Leftrightarrow p \nmid b$ and $p \mid a - 3b$. The integers in this ball are the ones who are congruent with 3 modulo p (we see this by setting $b = 1$);
- $\overline{B}(0,1) = B(0,1) \dot{\cup} B(1,1) \dot{\cup} \cdots \dot{\cup} B(p-1,1)$: The fact that the union in the right hand side is disjoint follows from the numbers

  $$a,\ a-b,\ \cdots,\ a-(p-1)b$$

  being all different modulo p, where $\frac{a}{b} \in \mathbb{Q}$ with $p \nmid b$. Now, if $p \mid a - ib$, then $\frac{a}{b} \in B(i,1)$. This proves the equality. Notice that this gives another proof that the closed unit ball is an open set, because it is a union of open balls;
- $B(1,1) = B\left(1, \frac{1}{2}\right) = \cdots = B\left(1, \frac{1}{p-1}\right) = \overline{B}\left(1, \frac{1}{p}\right)$, due to the fact that the p-adic absolute value can only take values which are powers of p.
- Another difference concerning a non-archimedean absolute value on a field $\mathbb{K}$ is that the *sphere* $\{x \in \mathbb{K} \colon |x-a| = r\}$ for $a \in \mathbb{K}$ and $r > 0$ is a clopen set (that is, both open and closed). It is closed because it is the intersection of the underlying closed ball with the complement of the underlying open ball and both of them are closed. In fact, this is general for any absolute value. Now, if we take x in the sphere, so that $|x-a| = r$, choosing $0 < \epsilon < r$ gives: if $|x-y| < \epsilon$, then $|y-a| = r$ because

all triangles are isosceles. Hence, the open ball around x of radius ϵ is completely contained in the sphere, so that the sphere is open. In this case, the sphere is not the boundary of the underlying open ball, due to being itself open.

As we have just seen, the rare clopen sets we were used when dealing with the usual calculus are much more common when we are dealing with non-archimedean absolute values. This will make the topology of these fields rather strange, but we could already guess that due to all triangles being isosceles. In order to understand this eccentricity deeply, we may recall the concept of a disconnected set. A set S is disconnected if one can find two open sets U_1 and U_2 such that

- $U_1 \cap U_2 = \emptyset$;
- $S = (S \cap U_1) \cup (S \cap U_2)$,
- $S \cap U_1, S \cap U_2 \neq \emptyset$.

Another big difference concerning a non-archimedean absolute value is that every closed ball with positive radius is disconnected. To see this, recall that it is clopen and any strictly smaller ball inside it is also clopen, so that the complement of the smaller ball inside the bigger ball is again clopen.

Recall now the meaning of a connected component of a set. Let $x \in \mathbb{K}$. The connected component of x is the union of all the connected sets containing x, that is, the largest connected set containing x. For example, the connected component of any point $x \in \mathbb{R}$ is $\mathbb{R}$ itself, simply because $\mathbb{R}$ is connected. This turns out to be way more different in the non-archimedean case.

Proposition 3.18. *In a field $\mathbb{K}$ with a non-archimedean absolute value, the connected component of any point $x \in \mathbb{K}$ is the set x consisting of only that point. That is, $\mathbb{K}$ is a totally disconnected topological space.*

Proof. Since every one-point set is trivially connected, it suffices to show that there are no more connected sets. So, suppose S contains both the points x and y and let $r = |x - y|$. Now, take $U_1 = B\left(x, \frac{r}{2}\right)$ and U_2 be its complement. Clearly U_1 is clopen, because it is an open ball, so that U_2 is also clopen, $x \in U_1$, $y \in U_2$, and these two sets are enough according to the definition of disconnected set to prove that S is disconnected, as we wanted. □

This result says that there are no interesting connected sets in $\mathbb{K}$: only the sets with one element are connected. On the other hand, if the absolute value on $\mathbb{K}$ is non-trivial, every one-point set is not open, because otherwise they all would have to be open, implying the underlying topology to be the discrete one, which only happens with the trivial absolute value.

To conclude this section, we return to our most studied case, where $\mathbb{K} = \mathbb{Q}$, and we are going to compare open and closed sets with respect to the usual absolute value and the p-adic absolute value.

Indeed, with respect to the usual absolute value, $\mathbb{Q}$ is totally disconnected. The key here is that between two rational numbers there is an irrational number and that the rational numbers are not connected as a set. For instance, the set of the rational numbers whose square is less than 2 is clopen. It is trivially open, and it is closed due to its complement being trivially open. Hence, if x, y are two different rationals, they belong to different connected components of $\mathbb{Q}$, which means that the connected components of $\mathbb{Q}$ are the singletons.

On the other hand, in the non-archimedean case, open balls are disconnected in a rather dramatic way. That is, every open ball is the disjoint union of open balls. The third and fourth item of the last examples says that the open unit ball around 0 is equal to the closed ball of radius $\frac{1}{p}$ around 0, which in turn is a disjoint union of open balls. By scaling and translating, we get the result for a general open ball.

3.1.2 Algebra

We are now going to take a brief look at the algebraic structure of the underlying fields provided the absolute values on them. The first one is a subring connected to any non-archimedean absolute value.

Proposition 3.19. *Let $|\cdot|$ be a non-archimedean absolute value on a field $\mathbb{K}$. We have the following properties:*

- *The closed unit ball $\Theta = \overline{B}(0,1) = \{x \in \mathbb{K} \colon |x| \leq 1\}$ is a subring of $\mathbb{K}$.*
- *The open unit ball $\beta = B(0,1) = \{x \in \mathbb{K} \colon |x| < 1\}$ is an ideal of Θ.*
- *Moreover, β is a maximal ideal in Θ and every element in $\Theta \backslash \beta$ is invertible in Θ.*

Proof. For the first property, notice that Θ contains 0, 1, is closed for multiplication due to $|xy| = |x||y|$, and is closed for subtraction due to the non-archimedean property, so that Θ is indeed a subring of $\mathbb{K}$.

Secondly, β contains 0, is closed for subtraction due to the non-archimedean property, and if $x \in \beta$, $y \in \Theta$, then $|x| < 1$, $|y| \leq 1$, so that $|xy| = |x||y| < 1$, i.e., $xy \in \beta$, hence β is an ideal of Θ.

Finally, if $x \in \Theta \backslash \beta$, then $|x| = 1$, so that $\left|\frac{1}{x}\right| = \frac{|1|}{|x|} = 1$, meaning that $\frac{1}{x} \in \Theta$ proving that every element in $\Theta \backslash \beta$ is invertible in Θ. Furthermore, any ideal of Θ strictly containing β would have to contain an invertible element, and thus, 1, so that it would be Θ, implying that β is a maximal ideal of Θ, which completes the proof. □

The important sets from the previous proposition are named in the following definition.

Definition 3.20. Let $\mathbb{K}$ be a field with a non-archimedean absolute value $|\cdot|$ and consider the last proposition's notation.

- The subring Θ is called the *valuation ring* of $|\cdot|$.
- The ideal β is called the *valuation ideal* of $|\cdot|$.
- The quotient $\kappa = \Theta/\beta$ is called the *residue field* of $|\cdot|$ (since the quotient of a ring by a maximal ideal is always a field).

Since our biggest concern is the case of the p-adic absolute values, we are going to find out what the previous sets are in this case.

Proposition 3.21. *Let $|\cdot| = |\cdot|_p$ be the p-adic absolute value on $\mathbb{K} = \mathbb{Q}$.*

- *The valuation ring of $|\cdot|$ is $\Theta = \{\frac{a}{b} \in \mathbb{Q} \colon p \nmid b\}$ and denoted $\mathbb{Z}_{(p)}$.*
- *The valuation ideal of $|\cdot|$ is $\beta = \{\frac{a}{b} \in \mathbb{Q} \colon p \nmid b \text{ and } p \mid a\}$ and denoted $p\mathbb{Z}_{(p)}$.*
- *The residue field of $|\cdot|$ is $\kappa = \mathbb{F}_p$, that is, the field with p elements.*

Proof. Letting $\frac{a}{b} \in \mathbb{Q}$, then there exist unique integers a_1, b_1 and v such that $\frac{a}{b} = p^v \frac{a_1}{b_1}$ with $p \nmid a_1 b_1$ and $\gcd(a_1, b_1) = 1$. So, we get:

- $\frac{a}{b} \in \Theta \Leftrightarrow |\frac{a}{b}| \leq 1 \Leftrightarrow v \geq 0$, thus if $\frac{a}{b}$ is in lowest terms, this means $p \nmid b$, as claimed.
- $\frac{a}{b} \in \beta \Leftrightarrow |\frac{a}{b}| < 1 \Leftrightarrow v > 0$, thus if $\frac{a}{b}$ is in lowest terms, this means $p \mid a$ and $p \nmid b$, as claimed.

- The inclusion $\mathbb{Z} \hookrightarrow \mathbb{Z}_{(p)}$ is an injective homomorphism which maps $p\mathbb{Z}$ into $p\mathbb{Z}_{(p)}$. Thus, we get a map $\Phi : \mathbb{Z}/p\mathbb{Z} \to \mathbb{Z}_{(p)}/p\mathbb{Z}_{(p)}$. In order to prove injectivity, just note that $\Phi(a) = 0 \Leftrightarrow \frac{a}{1} \in p\mathbb{Z}_{(p)} \Leftrightarrow p \mid a \Leftrightarrow a = 0$ (in $\mathbb{Z}/p\mathbb{Z}$). To check surjectivity, if $p \nmid b \in \mathbb{N}$, then there exists a positive integer c such that $bc \equiv 1 \pmod{p}$. Hence, for any $\frac{a}{b} \in \mathbb{Z}_{(p)}$, we get $\Phi(ac) = ac + p\mathbb{Z}_{(p)} = \frac{a}{b} + p\mathbb{Z}_{(p)}$, since $ac - \frac{a}{b} = \frac{abc-a}{b} = (bc-1)\frac{a}{b} \in p\mathbb{Z}_{(p)}$. □

3.1.3 The case of $\mathbb{Q}$

3.1.3.1 Absolute values

In order to construct the fields of p-adic numbers, it is important to understand first that we have already considered all possible absolute values in $\mathbb{Q}$, with a little hidden concept, which is, up to equivalence. The following definition will provide just that.

Definition 3.22. Two absolute values $|\cdot|_1$ and $|\cdot|_2$ on the same field $\mathbb{K}$ are said to be *equivalent* if they define the same topology on $\mathbb{K}$.

The following properties are equivalent to two absolute values being equivalent.

Lemma 3.23. *Let $|\cdot|_1, |\cdot|_2$ be two absolute values in a field $\mathbb{K}$. The next statements are equivalent:*

i) $|\cdot|_1$ and $|\cdot|_2$ are equivalent;

ii) For any $x \in \mathbb{K}$, $|x|_1 < 1 \Leftrightarrow |x|_2 < 1$,

iii) There exists $0 < \alpha \in \mathbb{R}$ such that for all $x \in \mathbb{K}$, $|x|_1 = |x|_2^{\alpha}$.

Proof. We will prove the implications in the usual order.

i) $\implies$ ii) : Supposing that $|\cdot|_1$ and $|\cdot|_2$ are equivalent, they define the same topology on $\mathbb{K}$, so that any sequence converges with respect to both of them or neither of them. Given any $x \in \mathbb{K}$, we get:

$$|x|_1 < 1 \Leftrightarrow \lim_{n\to\infty} |x|_1^n = 0 \Leftrightarrow \lim_{n\to\infty} x^n = 0 \Leftrightarrow \lim_{n\to\infty} |x|_2^n = 0 \Leftrightarrow |x|_2 < 1,$$

which gives ii).

ii) $\implies$ iii) : Let $0 \neq x_0 \in \mathbb{K}$ such that $|x_0|_1 < 1$ (if there is no such x_0, condition ii) implies that both the absolute values are the trivial ones, which are equivalent). Now, condition ii) implies $|x_0|_2 < 1$, so that there exists $0 < \alpha \in \mathbb{R}$ such that $|x_0|_1 = |x_0|_2^{\alpha}$ (just take $\alpha = \log_{|x_0|_1}(\log|x_0|_2)$).

Now, let $0 \neq x \in \mathbb{K}$. We want to show that $|x|_1 = |x|_2^\alpha$. Let us first consider the trivial cases. If $|x|_1 = |x_0|_1$, then $\left|\frac{x}{x_0}\right|_1 = \frac{|x|_1}{|x_0|_1} = \frac{|x_0|_1}{|x|_1} = \left|\frac{x_0}{x}\right|_1 = 1$, so that condition ii) implies that $\left|\frac{x}{x_0}\right|_2 = \left|\frac{x_0}{x}\right|_1 = 1$, and, thus, $|x|_2 = |x_0|_2$ and $|x|_1 = |x|_2^\alpha$ holds. The converse is true for $|x|_2 = |x_0|_2$. Also, if $|x|_1 = 1$, then $\left|\frac{1}{x}\right|_1 = 1$ and again condition ii) yields $|x|_2 = 1$ so that $|x|_1 = |x|_2^\alpha$ trivially holds. Hence, we may assume that

$$|x|_i \neq \begin{cases} 1 & \text{for } i = 1, 2, \\ |x_0|_i & \text{for } i = 1, 2. \end{cases}$$

As before, let $0 < \beta \in \mathbb{R}$ be such that $|x|_1 = |x|_2^\beta$. Our goal is to show that $\alpha = \beta$. We can also assume that $|x|_1 < 1$, otherwise we replace x by $\frac{1}{x}$. Condition ii) then gives $|x|_2 < 1$. Now, let m, n be two positive integers. Then we have:

$$|x|_1^n < |x_0|_1^m \Leftrightarrow \left|\frac{x^n}{x_0^m}\right|_1 < 1 \Leftrightarrow \left|\frac{x^n}{x_0^m}\right|_2 < 1 \Leftrightarrow |x|_2^n < |x_0|_2^m.$$

Taking logarithms in the first and last equations, we get:

$$\log|x|_1^n < \log|x_0|_1^m \Leftrightarrow \log|x|_2^n < \log|x_0|_2^m,$$

which is equivalent to

$$n\log|x|_1 < m\log|x_0|_1 \Leftrightarrow n\log|x|_2 < m\log|x_0|_2,$$

which we can write as

$$\frac{n}{m} < \frac{\log|x_0|_1}{\log|x|_1} \Leftrightarrow \frac{n}{m} < \frac{\log|x_0|_2}{\log|x|_2}.$$

Since $\mathbb{Q}$ is a dense subset of $\mathbb{R}$, this means that $\frac{\log|x_0|_1}{\log|x|_1} = \frac{\log|x_0|_2}{\log|x|_2}$, which is equivalent to $\frac{\log|x_0|_1}{\log|x_0|_2} = \frac{\log|x|_1}{\log|x|_2}$. Considering that $|x_0|_1 = |x_0|_2^\alpha$ and $|x|_1 = |x|_2^\beta$, we get that

$$\frac{\log|x_0|_2^\alpha}{\log|x_0|_2} = \frac{\log|x|_2^\beta}{\log|x|_2},$$

from where it follows that $\alpha = \beta$, as we wanted to show.

Since $0 \neq x \in \mathbb{K}$ was arbitrary, we get the equality $|x|_1 = |x|_2^\alpha$ for any $x \in \mathbb{K}$, which gives iii).

iii) $\Longrightarrow$ i): Assuming iii) we get for any $a \in \mathbb{K}$ and $r > 0$:

$$|x-a|_1 < r \Leftrightarrow |x-a|_2^\alpha < r \Leftrightarrow |x-a|_2 < r^{\frac{1}{\alpha}},$$

which shows that any open ball with respect to $|\cdot|_1$ is an open ball with respect to $|\cdot|_2$ and vice-versa, which means that the topologies induced by $|\cdot|_1$ and $|\cdot|_2$ are the same, so that the two absolute values are equivalent. This closes the circle of implications and completes the proof. □

Concerning $\mathbb{Q}$, are their some equivalent absolute values? and if so, to which extent? Firstly, the trivial one. If $|\cdot|$ is equivalent to the trivial absolute value, the previous lemma states that they differ by a raising of a positive power α. Since $1^\alpha = 1$ and $0^\alpha = 0$ for any such α, we conclude that $|\cdot|$ must also be trivial, thus, the same absolute value.

Analyzing the primes now, we can see that if p, q are two different primes, then the respective absolute values are not equivalent. This is due to $|p|_p < 1$, but $|p|_q = 1$, so that the lemma yields the rest. The same happens if $q = \infty$.

The following theorem shows how crucial the p-adic absolute values are.

Theorem 3.24 (Ostrowski)**.** *Up to equivalence, the non-trivial absolute values on* $\mathbb{Q}$ *are* $|\cdot|_p$*, where* p *is either prime or* $p = \infty$*.*

Proof. Let $|\cdot|$ be a non-trivial absolute value on $\mathbb{Q}$. This absolute value is either archimedean or non-archimedean, so let us separate the proof in these two possible cases.

a) If it is archimedean, we are going to show that it is equivalent to the ∞-adic absolute value. Corollary 3.10 implies that there exists a positive integer with absolute value bigger than 1, so that we may take the least one in these conditions, calling it n_0. Let α be the unique positive real number satisfying $|n_0| = n_0^\alpha$, that is, let $\alpha = \log_{|n_0|} n_0$. In order to prove the desired equivalence, this α has to be the one satisfying the condition in Lemma 3.23. Given the multiplicative property of an absolute value, together with $|1| = 1$, we know that the absolute value of an inverse is the inverse of the absolute value of the initial number, so that the absolute value of a quotient of integers is the quotient of the absolute values, thus we may only restrict ourselves to positive integers, that is, we just need to show that $|n| = n^\alpha$ for every positive integer n. So, take an arbitrary positive integer n and we are going to write it in base n_0. Say,

$$n = a_0 + a_1 n_0 + \cdots + a_k n_0^k,$$

where $0 \leq a_i \leq n_0 - 1$ and $a_k \neq 0$ are non-negative integers. Since k is such that $n_0^k \leq n < n_0^{k+1}$, we get that $k = \left\lfloor \frac{\log n}{\log n_0} \right\rfloor$. Taking absolute values, we get:

$$\begin{aligned}
|n| &= |a_0 + a_1 n_0 + \cdots + a_k n_0^k| \\
&\leq |a_0| + |a_1 n_0| + \cdots + |a_k n_0^k| \\
&= |a_0| + |a_1||n_0| + \cdots + |a_k||n_0^k| \\
&= |a_0| + |a_1| n_0^\alpha + \cdots + |a_k| n_0^{k\alpha} \\
&\leq 1 + n_0^\alpha + \cdots + n_0^{k\alpha} \\
&= n_0^{k\alpha}(1 + n_0^{-\alpha} + \cdots + n_0^{-k\alpha}) \\
&= n_0^{k\alpha} \sum_{i=0}^{k} n_0^{-i\alpha} \\
&\leq n_0^{k\alpha} \sum_{i=0}^{\infty} n_0^{-i\alpha} \\
&= n_0^{k\alpha} \frac{n_0^\alpha}{n_0^\alpha - 1} \\
&\leq n^\alpha \frac{n_0^\alpha}{n_0^\alpha - 1}.
\end{aligned}$$

Since n was arbitrary, we get the above inequality for every positive integer. Now we are going to apply it for any power of n. Given $N \in \mathbb{N}$, we get:

$$\left| n^N \right| = |n|^N \leq n^{N\alpha} \frac{n_0^\alpha}{n_0^\alpha - 1},$$

implying

$$|n| \leq \sqrt[N]{\frac{n_0^\alpha}{n_0^\alpha - 1}} n^\alpha.$$

Letting $N \to \infty$, we get

$$\lim_{N \to \infty} \sqrt[N]{\frac{n_0^\alpha}{n_0^\alpha - 1}} n^\alpha = n^\alpha.$$

This proves $|n| \leq n^\alpha$, which is one of the desired inequalities. In order to prove the reverse inequality, we see that

$$n_0^{(k+1)\alpha} = |n_0^{k+1}| = |n + n_0^{k+1} - n| \leq |n| + |n_0^{k+1} - n|,$$

implying

$$\begin{aligned}
|n| &\geq n_0^{(k+1)\alpha} - |n_0^{k+1} - n| \\
&\geq n_0^{(k+1)\alpha} - (|n_0^{k+1}| - |n|) \\
&\geq n_0^{(k+1)\alpha} - (n_0^{(k+1)\alpha} - n^\alpha) \\
&\geq n_0^{(k+1)\alpha} - (n_0^{k+1} - n)^\alpha \\
&\geq n_0^{(k+1)\alpha} - (n_0^{k+1} - n_0^k)^\alpha \\
&= n_0^{(k+1)\alpha} - (n_0^k(n_0 - 1))^\alpha \\
&= n_0^{(k+1)\alpha}\left(1 - \left(\frac{n_0 - 1}{n_0}\right)^\alpha\right) \\
&= n_0^{(k+1)\alpha}\left(1 - \left(\frac{1}{1 - n_0}\right)^\alpha\right) \\
&> \left(1 - \left(\frac{1}{1 - n_0}\right)^\alpha\right) n^\alpha.
\end{aligned}$$

The same trick as before gives the desired inequality. That is, we finished the archimedean case, as expected.

b) Now, supposing that $|\cdot|$ is non-archimedean, we have to find a prime p for which $|\cdot| = |\cdot|_p$. Corollary 3.10 implies $|n| \leq 1$ for every positive integer n. The non-triviality of $|\cdot|$ implies the previous inequality is not always an equality, so let n_1 be the smallest positive number with that property. n_1 must be a prime number, because otherwise, if $n_1 = ab$ with $a, b < n_1$, positive integers, the minimality of n_1 implies $|a| = |b| = 1$, so that $|n_1| = |ab| = |a||b| = 1$, a contradiction. Let us call p to n_1. This will naturally be the prime we were looking for. Letting n be a positive integer coprime with p, the division algorithm yiels $n = rp + s$ for some non-negative integer r and an integer $0 < s < p$. The minimality of p, again, implies $|s| = 1$. Also, $|rp| = |r||p| < 1$. Proposition 3.14 then gives $|n| = |s| = 1$. More generally, we can write any $m \in \mathbb{Z}$ as $m = p^v m'$ with $p \nmid m'$. We get:

$$|m| = |p^v m'| = |p^v||m'| = |p|^v|m'| = |p|^v = (|p|^{-1})^{-v} = (|p|^{-1})^{-v_p(m)}.$$

Since $|p| < 1$, then $|p|^{-1} > 1$, implying that $|\cdot|$ is equivalent to the p-adic absolute value, as claimed. □

To conclude this section, remains to be stated a result which has some unexpected applications.

Proposition 3.25 (Product Formula). *For any $0 \neq x \in \mathbb{Q}$, we have*

$$\prod_{p \leq \infty} |x|_p = 1,$$

where $p \leq \infty$ means that we take the product of all primes, including the one at infinity.

Proof. The general case will follow trivially if one proves the statement for any positive integer. So, let n be a positive integer with prime factorization $n = p_1^{\alpha_1} \cdots p_k^{\alpha_k}$. One gets

$$|x|_p = \begin{cases} 1 & \text{if } p \neq p_i \text{ for every } 1 \leq i \leq k; \\ p_i^{-\alpha_i} & \text{if } p = p_i \text{ for some } 1 \leq i \leq k, \\ p_1^{\alpha_1} \cdots p_k^{\alpha_k} & \text{if } p = \infty. \end{cases}$$

The product of all the absolute values clearly yields 1, as claimed. □

Although at this point we will not take full advantage of this result, one application of it includes determining the missing absolute value of some $x \in \mathbb{Q}$ if we know the rest of the absolute values.

3.1.3.2 Completions

The construction of the p-adic fields $\mathbb{Q}_p$ will be made in this section. As the construction of $\mathbb{R}$ from $\mathbb{Q}$, we will use Cauchy sequences, which we will briefly recall.

Definition 3.26. Let $|\cdot|$ be an absolute value on a field $\mathbb{K}$.

- A sequence of elements $x_n \in \mathbb{K}$ is called a *Cauchy sequence* if

$$\forall\, \epsilon > 0\ \exists\, M \in \mathbb{N} : m, n \geq M \implies |x_n - x_m| < \epsilon.$$

- $\mathbb{K}$ is called *complete* with respect to $|\cdot|$ if every Cauchy sequence in $\mathbb{K}$ has a limit in $\mathbb{K}$.

- A subset $S \subseteq \mathbb{K}$ is called *dense* in $\mathbb{K}$ if

$$\forall x \in \mathbb{K},\ \forall \epsilon > 0,\ B(x, \epsilon) \cap S \neq \emptyset.$$

Remark 3.27. Notice that the terms of every Cauchy sequence $(x_n)_n$ must be bounded, because, for instance, for $\epsilon = 1$, there exists $M \in \mathbb{N}$ such that $m, n \geq M \implies |x_n - x_m| < 1$, so that for every $n \in \mathbb{N}$, $|x_n| \leq \max\{|x_1|, \cdots, |x_{M-1}|, |x_M| + 1\}$.

For instance, one already knows from a real analysis' course that $\mathbb{Q}$ is not complete with respect to the usual absolute value, since one can construct a Cauchy sequence in $\mathbb{Q}$ converging to $\sqrt{2} \notin \mathbb{Q}$.

But this problem can be easily handled if we extend the archimedean absolute value $|\cdot|_\infty$ to $\mathbb{R}$, that is,

- $\mathbb{R}$ is complete with respect to this absolute value,
- $\mathbb{Q}$ is dense in $\mathbb{R}$.

In other words, $\mathbb{R}$ is the completion of $\mathbb{Q}$ with respect to the absolute value $|\cdot|_\infty$, that is, $\mathbb{R}$ is the smallest field containing $\mathbb{Q}$ which is complete with respect to this absolute value.

It is now clear what we are going to do: construct the analogous completions of $\mathbb{Q}$ with respect to the p-adic absolute values, which will not be an easy task.

Remark 3.28. Note that there is no need for such construction concerning the trivial absolute value, since $\mathbb{Q}$ is already complete with respect to it.

So, throughout this section, $|\cdot|$ will be reserved for the p-adic absolute value on $\mathbb{Q}$. The next result characterizes a Cauchy sequence when the absolute value in non-archimedean, which is the case for the p-adic absolute value.

Lemma 3.29. *A sequence of rationals x_n is a Cauchy sequence with respect to $|\cdot|$ if and only if*

$$\lim_{n\to\infty} |x_{n+1} - x_n| = 0.$$

Proof. (only if) This is true for every Cauchy sequence.

(if) If $m, n \in \mathbb{N}$ with $m > n$, say, $m = n + r$, then we get:

$$\begin{aligned} |x_m - x_n| &= |x_{n+r} - x_{n+r-1} + x_{n+r-1} - \cdots - x_{n+1} + x_{n+1} - x_n| \\ &\leq \max\{|x_{n+r} - x_{n+r-1}|, \{x_{n+r-1} - x_{n+r-2}\}, \cdots, \{x_{n+1} - x_n\}\} \to_{n\to\infty} 0, \end{aligned}$$

so that (x_n) is a Cauchy sequence. □

Remark 3.30. Notice that the previous lemma is not true for archimedean absolute values. For instance, if we let $\mathbb{K} = \mathbb{R}$, the harmonic series $x_n = \sum_{k=1}^{n} \frac{1}{k}$ is a counterexample: its terms are getting closer (they differ by $\frac{1}{n}$), but the sequence cannot be bounded, that is, has no limit.

Now, in order to proceed, it remains to prove that $\mathbb{Q}$ is not complete with respect to the p-adic absolute value, so that the completion process is really going to be non-trivial.

Lemma 3.31. $\mathbb{Q}$ *is not complete with respect to any of its nontrivial absolute values.*

Proof. According to Ostrowski's Theorem 3.24, we have to prove this for $|\cdot|_p$ where $p \leq \infty$. The case $p = \infty$ was already mentioned and proved in a real analysis course. Let then, $|\cdot| = |\cdot|_p$ for some prime p. It is enough to construct a Cauchy sequence in $\mathbb{Q}$ that does not have a limit within the field. We will divide the proof in two cases: first, when $p \neq 2$, and then, when $p = 2$.

i) If $p \neq 2$, then let $a \in \mathbb{Z}$ be a non-trivial quadratic residue modulo p which is not a square in $\mathbb{Q}$ (recall from Number Theory that half of $\mathbb{Z}_p^*$ are quadratic residues modulo p), so that there must be at least a non-trivial one which is not a square in $\mathbb{Q}$. Now, we want to construct a Cauchy Sequence $(x_n)_n$ which will not converge in $\mathbb{Q}$.

- Let x_1 be a solution of $x^2 \equiv a \pmod{p}$ and let $b = \frac{x_1^2 - a}{p} > 0$.
- Let x_2 be such that $x_2 \equiv x_1 \pmod{p}$ and $x_2^2 \equiv a \pmod{p^2}$. In order to prove the existence of x_2, one has to find $y \in \mathbb{Z}$ such that

$$\begin{aligned}
(x_1 + yp)^2 &\equiv a \pmod{p^2} \\
x_1^2 + 2x_1yp + y^2p^2 &\equiv a \pmod{p^2} \\
x_1^2 + 2x_1yp &\equiv a \pmod{p^2} \\
x_1^2 - a + 2x_1yp &\equiv 0 \pmod{p^2} \\
bp + 2x_1yp &\equiv 0 \pmod{p^2} \\
b + 2x_1y &\equiv 0 \pmod{p} \\
2x_1y &\equiv -b \pmod{p}.
\end{aligned}$$

  Now, if $p \mid b$, then we can take $y = 0$, so that $x_2 = x_1$. Otherwise, since $p \nmid x_1$, then $2x_1$ is invertible modulo p, say, $2x_1z \equiv 1 \pmod{p}$, so that one can take $y = -bz$.

- Now, the general step is to choose x_n so that $x_n \equiv x_{n-1} \pmod{p^n}$ and $x_n^2 \equiv a \pmod{p^{n+1}}$, and its existence can be analogously proven as the existence of x_2.

  In order to prove that $(x_n)_n$ is a Cauchy sequence, notice that for every $n \in \mathbb{N}$, there is $y_n \in \mathbb{N}$ such that $x_{n+1} - x_n = y_n p^{n+1}$ and so:

$$|x_{n+1} - x_n| = |y_n p^{n+1}| \leq |p^{n+1}| = p^{-(n+1)} \to_{n\to\infty} 0$$

  so that Lemma 3.29 implies that $(x_n)_n$ is a Cauchy sequence.

On the other hand, for every $n \in \mathbb{N}$, there is $w_n \in \mathbb{N}$ such that $x_n^2 - a = z_n p^{n+1}$ and so:

$$|x_n^2 - a| = |z_n p^{n+1}| \leq |p^{n+1}| \to_{n\to\infty} 0$$

implying that the limit of the sequence $(x_n)_n$, if it existed, had to be the square root of a, which is not a rational by construction, meaning that $(x_n)_n$ does not converge in $\mathbb{Q}$ and hence $\mathbb{Q}$ is not complete with respect to $|\cdot|_p$.

ii) If $p = 2$, then since the equation $x^3 \equiv 3 \pmod 2$ has a solution ($x = 1$ for instance) and 3 is not a cube in $\mathbb{Q}$, we are going to construct a Cauchy sequence converging to the cubic root of 3 in the following way:

- Let x_1 be a solution of $x^3 \equiv 3 \pmod 2$ and let $b = \frac{x^3-3}{2} > 0$.
- Let x_2 be such that $x_2 \equiv x_1 \pmod 2$ and $x_2^3 \equiv 3 \pmod 4$. In order to prove the existence of x_2, we have to find $y \in \mathbb{Z}$ such that

$$\begin{aligned}
(x_1 + 2y)^3 &\equiv 3 \pmod 4 \\
x_1^3 + 6x_1^2 y + 12x_1 y^2 + 8y^3 &\equiv 3 \pmod 4 \\
x_1^3 + 2x_1^2 y &\equiv 3 \pmod 4 \\
x_1^3 - 3 + 2x_1^2 y &\equiv 0 \pmod 4 \\
2b + 2x_1^2 y &\equiv 0 \pmod 4 \\
b + x_1^2 y &\equiv 0 \pmod 2 \\
x_1^2 y &\equiv -b \pmod 2.
\end{aligned}$$

  Now, if $2 \mid b$, then one can take $y = 0$, so that $x_2 = x_1$. Otherwise, since $2 \nmid x_1$, then x_1 is invertible modulo 2, so that one can take $y = -b$.

- Now, the general step is to choose x_n so that $x_n \equiv x_{n-1} \pmod{2^n}$ and $x_n^3 \equiv 3 \pmod{2^{n+1}}$, and its existence can be analogously proven as the existence of x_2.

  In order to prove that $(x_n)_n$ is a Cauchy sequence, notice that for every $n \in \mathbb{N}$, there is $y_n \in \mathbb{N}$ such that $x_{n+1} - x_n = y_n 2^{n+1}$ and so:

$$|x_{n+1} - x_n| = |y_n 2^{n+1}| \leq |2^{n+1}| = 2^{-(n+1)} \to_{n\to\infty} 0,$$

  so that Lemma 3.29 implies that $(x_n)_n$ is a Cauchy sequence.

On the other hand, for every $n \in \mathbb{N}$, there is $w_n \in \mathbb{N}$ such that $x_n^3 - 3 = z_n 2^{n+1}$ and so:

$$|x_n^2 - 3| = |z_n 2^{n+1}| \leq |2^{n+1}| \to_{n\to\infty} 0,$$

so that the limit of the sequence $(x_n)_n$, if it existed, had to be the cubic root of 3, which is not a rational by construction, meaning that $(x_n)_n$ does not converge in $\mathbb{Q}$ and hence $\mathbb{Q}$ is not complete with respect to $|\cdot|_2$, which finishes the proof. □

Now the goal is to construct a completion of $\mathbb{Q}$ with respect to any non-archimedean absolute value. The most straightforward way would be to "add to $\mathbb{Q}$ the limits of all the Cauchy sequences", but the problem is that those limits do not exist, so we need to find a way of turning this around. In order to solve this, we may replace the limits that we do not know with the Cauchy sequences themselves. Let us now consider the set of all the Cauchy sequences.

Definition 3.32. Let $|\cdot| = |\cdot|_p$ be a non-archimedean absolute value on $\mathbb{Q}$. We will denote by $C_p(\mathbb{Q})$ the set of all Cauchy sequences of elements of $\mathbb{Q}$ with respect to the absolute value $|\cdot|$, that is,

$$C_p(\mathbb{Q}) = \{(x_n)_n \colon (x_n)_n \text{ is a Cauchy sequence in } \mathbb{Q} \text{ with respect to} |\cdot|_p\}.$$

In fact, $C_p(\mathbb{Q})$ has a natural ring structure if we define the sum and product in the natural way, that is, $(x_n) + (y_n) = (x_n + y_n)$, and $(x_n) \cdot (y_n) = (x_n y_n)$.

Proposition 3.33. *$C_p(\mathbb{Q})$ is a commutative ring with unity with the above operations.*

Proof. It is enough to prove that the sum and product of Cauchy sequences is still a Cauchy sequence, since the rest of the properties follow easily from $\mathbb{Q}$ being a field. Now, let $(x_n), (y_n) \in C_p(\mathbb{Q})$. For the sum,

$$\begin{aligned}
\lim_{n\to\infty} |(x_{n+1} + y_{n+1}) - (x_n + y_n)| &= \lim_{n\to\infty} |(x_{n+1} - x_n) + (y_{n+1} - y_n)| \\
&\leq \lim_{n\to\infty} |(x_{n+1} - x_n)| + |(y_{n+1} - y_n)| \\
&= \lim_{n\to\infty} |(x_{n+1} - x_n)| + \lim_{n\to\infty} |(y_{n+1} - y_n)| \\
&\overset{3.29}{=} 0 + 0 = 0,
\end{aligned}$$

so that Lemma 3.29 implies that $(x_n + y_n)$ is a Cauchy sequence. For the product,

$$\begin{aligned}
\lim_{n\to\infty}|(x_{n+1}y_{n+1}) - (x_ny_n)| &= \lim_{n\to\infty}|(x_{n+1}(y_{n+1}-y_n) + y_n(x_{n+1}-x_n)| \\
&\leq \lim_{n\to\infty}|(x_{n+1}(y_{n+1}-y_n)| + |(y_n(x_{n+1}-x_n)| \\
&= \lim_{n\to\infty}|x_{n+1}||(y_{n+1}-y_n)| + |y_n||(x_{n+1}-x_n)|| \\
&= \lim_{n\to\infty}|x_{n+1}||(y_{n+1}-y_n)| + \lim_{n\to\infty}|y_n||(x_{n+1}-x_n)| \\
&= \lim_{n\to\infty}|x_{n+1}|\lim_{n\to\infty}|(y_{n+1}-y_n)| + \lim_{n\to\infty}|y_n|\lim_{n\to\infty}|(x_{n+1}-x_n)| \\
&\overset{3.29}{=} \lim_{n\to\infty}|x_{n+1}|\cdot 0 + \lim_{n\to\infty}|y_n|\cdot 0 \\
&\overset{3.27}{=} 0+0=0,
\end{aligned}$$

and Lemma 3.29 implies that (x_ny_n) is a Cauchy sequence, completing the proof. □

Remark 3.34. Notice that the zero element of $C_p(\mathbb{Q})$ is the constant sequence equal to 0 and the unit element is the constant sequence equal to 1. Moreover, $C_p(\mathbb{Q})$ is clearly not a field: the Cauchy sequence $(x_n)_n$ defined by $x_1 = 1$ and $x_n = 0$ for $n > 1$ is non-zero and not invertible. In fact, letting $(y_n)_n$ be defined by $y_1 = 0$ and $y_n = 1$ for $n > 1$, we get that $(x_ny_n)_n$ is the zero sequence, so that these two sequences are zero divisors of $C_p(\mathbb{Q})$. In fact, if $(z_n)_n \in C_p(\mathbb{Q})$ is invertible, then every z_n must be non-zero. But this is not a sufficient condition to be invertible, since any Cauchy sequence converging to 0 is not invertible, because the only possibility for an inverse would be a sequence which is not Cauchy. Hence, the invertible elements are the Cauchy sequences not converging to 0 with every term non-zero, that is, $(w_n)_n \in C_p(\mathbb{Q})$ is invertible if and only if there exists $\epsilon > 0$ such that $|w_n| > \epsilon$ for every n.

Now, in order to construct the desired completion, we will include $\mathbb{Q}$ into $C_p(\mathbb{Q})$ via the map $x \to (x)$ where (x) is the constant Cauchy sequence equal to x.

Recalling that our desired completion must contain all the limits of Cauchy sequences, we still have not solved this problem since different Cauchy sequences that get closer and closer to each other should have the same limit. In order to solve this, we must pass to a quotient of $C_p(\mathbb{Q})$ by identifying those sequences. The next set provides the first step for doing that.

Lemma 3.35. *Letting $N_p(\mathbb{Q}) \subseteq C_p(\mathbb{Q})$ be defined as*

$$N_p(\mathbb{Q}) = \{(x_n)_n \in C_p(\mathbb{Q}) \colon \lim_{n\to\infty}|x_n|_p = 0\},$$

we get that $N_p(\mathbb{Q})$ is a maximal ideal of $C_p(\mathbb{Q})$.

Proof. Firstly, let us see that $N_p(\mathbb{Q})$ is an ideal. Clearly, $(0) \in N_p(\mathbb{Q})$. If $(x_n)_n$, $(y_n)_n \in N_p(\mathbb{Q})$, then their subtraction is a Cauchy sequence, and for every $n \in \mathbb{N}$,

$$|(x_{n+1} - y_{n+1}) - (x_n - y_n)| \leq |x_{n+1} - y_{n+1}| + |x_n - y_n|,$$

so that

$$\lim_{n\to\infty} |(x_{n+1} - y_{n+1})| = \lim_{n\to\infty} |x_n - y_n| \ = 0,$$

implying that

$$\lim_{n\to\infty} |(x_{n+1} - y_{n+1}) - (x_n - y_n)| = 0,$$

implying that $(x_n - y_n)_n \in N_p(\mathbb{Q})$.

Finally, if $(z_n)_n \in C_p(\mathbb{Q})$, remark 3.27 implies $\lim_{n\to\infty}(x_n z_n)_n = 0$, so that $(x_n z_n)_n \in N_p(\mathbb{Q})$, which shows that $N_p(\mathbb{Q})$ is an ideal of $C_p(\mathbb{Q})$.

It remains to show that the ideal $N_p\mathbb{Q}$ of $C_p(\mathbb{Q})$ is maximal. So, let $(z_n) \in C_p(\mathbb{Q}) \setminus N_p(\mathbb{Q})$, that is, a Cauchy sequence that does not tend to 0. We want to show that the ideal generated by $N_p(\mathbb{Q})$ and $(z_n)_n$ is the whole ring $C_p(\mathbb{Q})$, for which it is enough to show that it contains the unit of the ring, the constant sequence equal to 1.

Well, $(z_n)_n$ not tending to 0 means that there exists $\epsilon > 0$ such that for all $M \in \mathbb{N}$, there exists $m_0 \geq M$ such that $|z_{m_0}| > \epsilon$. On the other hand, $(z_n)_n$ being Cauchy implies that there exists $M_0 > 0$ such that for every n, $m \geq M_0$, $|z_n - z_m| < \frac{\epsilon}{2}$. In particular, for every $m \geq M_0$, we have $|z_m - z_{m_0}| < \frac{\epsilon}{2}$. But since $|z_{m_0}| < \epsilon$, this implies $|z_m| > \frac{\epsilon}{2}$ for every $m > M_0$. Let $\delta = \frac{\epsilon}{2}$.

In particular, $z_m \neq 0$ for every such m, so that we can define a sequence $(w_n)_n$ by $w_n = \frac{1}{z_n}$ if $n \geq M_0$ and $w_n = 0$ otherwise. The first thing to check is that $(w_n)_n$ is Cauchy, since for $n \geq M_0$,

$$|w_{n+1} - w_n| = \left|\frac{1}{z_{n+1}} - \frac{1}{z_n}\right| = \left|\frac{z_n - z_{n+1}}{z_n z_{n+1}}\right| \leq \left|\frac{z_{n+1} - z_n}{\delta^2}\right| \overset{3.29}{\to} 0 \text{ as } n \to \infty,$$

which shows that $(w_n)_n \in C_p(\mathbb{Q})$ due to Lemma 3.29.

Noticing that $z_n w_n = 0$ if $n \leq M_0$ and $z_n w_n = 1$ otherwise, we get $(1) - z_n w_n = 1$ if $n \leq M_0$ and $(1) - z_n w_n = 0$ otherwise, meaning that $(1) - z_n w_n \in N_p(\mathbb{Q})$. This means that (1) can be written as a sum of an element of $N_p(\mathbb{Q})$ with $z_n w_n$, and hence (1) belongs to the ideal generated by $N_p(\mathbb{Q})$ and $(z_n)_n$, as we wanted to show. □

Next is the most crucial definition in this current chapter, the field of p-adic numbers.

Definition 3.36. We define the field of p-adic numbers, denoted $\mathbb{Q}_p$, to be the quotient of the ring $C_p(\mathbb{Q})$ by its maximal ideal $N_p(\mathbb{Q})$:

$$\mathbb{Q}_p = C_p(\mathbb{Q})/N_p(\mathbb{Q})$$

Since two different constant sequences differ by another non-zero constant sequence, we still have an inclusion $\mathbb{Q} \subseteq \mathbb{Q}_p$ by sending $x \in \mathbb{Q}$ to the equivalent class of the constant sequence (x), which simply contains the constant sequence itself and all the other Cauchy sequences whose limit is x.

Since for every prime p we have an inclusion of $\mathbb{Q}$ into a field $\mathbb{Q}_p$, it remains to check that this field has the desired properties to be a completion of $\mathbb{Q}$ with respect to the absolute value $|\cdot|_p$.

First, let us see that the absolute value $|\cdot|_p$ extends to $\mathbb{Q}_p$.

Lemma 3.37. *Let $(x_n)_n \in C_p(\mathbb{Q})/N_p(\mathbb{Q})$. The sequence of real numbers $|x_n|_p$ is eventually stationary, meaning that there exists $N \in \mathbb{N}$ such that $n, m \geq N \implies |x_n|_p = |x_m|_p$.*

Proof. Since $(x_n)_n$ is a Cauchy sequence which does not tend to 0, then we can find $\epsilon > 0$ and $N_1 > 0$ such that

$$n > N_1 \implies |x_n| \geq \epsilon.$$

On the other hand, since $(x_n)_n$ is Cauchy, there exists an integer N_2 such that

$$n, m > N_2 \implies |x_n - x_m| < \epsilon.$$

Setting $N = \max\{N_1, N_2\}$, we get

$$n, m > N \implies |x_n - x_m| < \max\{|x_n|, |x_m|\},$$

so that the non-archimedean property gives $|x_n| = |x_m|$, as we wanted to show. □

From this lemma, the following definition makes sense:

Definition 3.38. If $\lambda \in \mathbb{Q}_p$, and $(x_n)_n \in C_p(\mathbb{Q})$ represents λ, we define the p-adic absolute value of λ as

$$|\lambda|_p = \lim_{n\to\infty} |x_n|_p.$$

There are several things to notice here.

Observation 1. Firstly, let us prove that the previous limit always exists. If $\lambda = 0$, meaning that $(x_n)_n \in N_p(\mathbb{Q})$, then $\lim_{n\to\infty} |x_n|_p = 0$, and so $|\lambda|_p = 0$. Otherwise, Lemma 3.37

implies that $(x_n)_n$ is constant for sufficient large n, so that it has a limit. Secondly, we must check that the previous limit does not depend on the choice of the Cauchy sequence representing λ. Say $(y_n)_n$ is another Cauchy sequence representing λ, then by definition $\lim_{n\to\infty} x_n - y_n = 0$, so that $\lim_{n\to\infty}|x_n - y_n|_p = 0$ and so the triangular inequality gives $\lim_{n\to\infty}|x_n|_p - |y_n|_p = 0$, thus $\lim_{n\to\infty}|x_n|_p = \lim_{n\to\infty}|y_n|_p$.

Now, we have to show that the previous definition yields an absolute value on $\mathbb{Q}_p$.

Proposition 3.39. *$|\cdot|_p$ defined above in $\mathbb{Q}_p$ is a non-archimedean absolute value in $\mathbb{Q}_p$.*

Proof. Let $\lambda_1, \lambda_2 \in \mathbb{Q}_p$ and $(x_n)_n, (y_n)_n \in C_p(\mathbb{Q})$ be two Cauchy sequences representing λ_1 and λ_2 respectively.

- We get $|\lambda_1|_p = 0 \Leftrightarrow \lim|x_n|_p = 0 \Leftrightarrow (x_n)_n \in N_p(\mathbb{Q}) \Leftrightarrow \lambda_1 = 0$;
- $|\lambda_1|_p|\lambda_2|_p = \lim_{n\to\infty}|x_n|_p \lim_{n\to\infty}|y_n|_p = \lim_{n\to\infty}|x_n y_n|_p = |\lambda_1\lambda_2|_p$,
- The non-archimedean property follows from

$$\begin{aligned}|\lambda_1 + \lambda_2|_p = \lim_{n\to\infty}|x_n + y_n|_p &\leq \lim_{n\to\infty}\max\{|x_n|_p, |y_n|_p\} \\ &= \max\{\lim_{n\to\infty}|x_n|_p, \lim n \to \infty |y_n|_p\} \\ &= \max\{|\lambda_1|_p, |\lambda_2|_p\},\end{aligned}$$

so that the proof is complete. □

Finaly, it remains to check that if $x \in \mathbb{Q}$, then $|(x)|_p = |x|_p$, which is obvious, hence we showed that the p-adic absolute value extends to $\mathbb{Q}_p$.

Now, we are going to show that the image of $\mathbb{Q}$ under the inclusion $\mathbb{Q} \subseteq \mathbb{Q}_p$ is a dense subset of $\mathbb{Q}_p$.

Proposition 3.40. *The image of $\mathbb{Q}$ under the inclusion $\mathbb{Q} \subseteq \mathbb{Q}_p$ is a dense subset of $\mathbb{Q}_p$.*

Proof. Let $\lambda \in \mathbb{Q}_p$ and $\epsilon > 0$. It suffices to show that $B(\lambda, \epsilon)$ contains a constant sequence. So, let $(x_n)_n$ be a Cauchy sequence representing λ. Thus, in particular, there exists $N > 0$ such that for all n, $m > N$, $|x_n - x_m| < \frac{\epsilon}{2}$. We claim that $(x_N) \in B(\lambda, \epsilon)$. Clearly, $\lambda - (x_N)$ is represented by the sequence $(x_n - x_N)_n$ and by definition we have $|(x_n - x_N)| = \lim_{n\to\infty}|x_n - x_N|$. But for any $n \geq N$ we have $|x_n - x_N| < \frac{\epsilon}{2}$, hence in the limit we get

$$\lim_{n\to\infty}|x_n - x_N| \leq \frac{\epsilon}{2} < \epsilon,$$

implying that $(x_N) \in B(\lambda, \epsilon)$, which completes the proof. □

Before the big theorem is stated, we still need to show that $\mathbb{Q}_p$ is complete with respect to this absolute value. The only problem here is that we are dealing with Cauchy sequences whose elements are themselves Cauchy sequences up to equivalence, but it will all be overcome soon.

Proposition 3.41. *For every prime p, $\mathbb{Q}_p$ is complete with respect to the absolute value $|\cdot|_p$.*

Proof. Let $(\lambda_n)_n$ be a Cauchy sequence of elements of $\mathbb{Q}_p$, meaning that each λ_n is an equivalent class of Cauchy sequences of rational elements. For each $n \in \mathbb{N}$, let also z_n be a representative of λ_n. We want to show that this Cauchy sequence converges in $\mathbb{Q}_p$ to some equivalent class of Cauchy sequences in $\mathbb{Q}_p$. In order to do that, we are going to find a representative λ of that equivalent class.

From the previous proposition, the image of $\mathbb{Q}$ in $\mathbb{Q}_p$ being dense in $\mathbb{Q}_p$, one can already choose convenient radius, so that it is possible to find rational numbers y_1, y_2, …, y_n, …, such that for every $n \in \mathbb{N}$, one gets

$$(y_n) \in B\left(\lambda_n, \frac{1}{n}\right),$$

from where it follows that

$$\lim_{n\to\infty} |\lambda_n - (y_n)| = 0.$$

We will now prove that the sequence $(y_n)_n$ is Cauchy in $\mathbb{Q}$ with respect to $|\cdot|_p$. By Lemma 3.29, it suffices to show that

$$\lim_{n\to\infty} |y_{n+1} - y_n|_p = 0.$$

Noticing that

$$\begin{aligned} |y_{n+1} - y_n| &= |\lambda_{n+1} - \lambda_{n+1} + \lambda_n - \lambda_n + y_{n+1} - y_n| \\ &= |y_{n+1} - \lambda_{n+1} + \lambda_{n+1} - \lambda_n + \lambda_n - y_n| \\ &\leq |y_{n+1} - \lambda_{n+1}| + |\lambda_{n+1} - \lambda_n| + |\lambda_n - y_n|, \end{aligned}$$

it follows that

$$\begin{aligned} \lim_{n\to\infty} |y_{n+1} - y_n|_p &\leq \lim_{n\to\infty} |y_{n+1} - \lambda_{n+1}| + |\lambda_{n+1} - \lambda_n| + |\lambda_n - y_n| \\ &= \lim_{n\to\infty} |y_{n+1} - \lambda_{n+1}| + \lim_{n\to\infty} |\lambda_{n+1} - \lambda_n| + \lim_{n\to\infty} |\lambda_n - y_n| \\ &= 0 + 0 + 0 = 0, \end{aligned}$$

implying that $(y_n)_n$ is Cauchy in $\mathbb{Q}$ with respect to $|\cdot|_p$, as we wanted to show. Letting λ denote the element of $\mathbb{Q}_p$ corresponding to this sequence, it follows that

$$\lim_{n\to\infty} |\lambda_n - \lambda|_p = 0,$$

implying that $\lim_{n\to\infty} \lambda_n = \lambda$, so that the Cauchy sequence $(\lambda_n)_n$ in $\mathbb{Q}_p$ converges in $\mathbb{Q}_p$, which completes the proof. □

To conclude this section, remains to be stated the most powerful theorem so far, whose proof is immediate because we have already showed everything.

Theorem 3.42. *For each prime $p \in \mathbb{N}$, there exists a field $\mathbb{Q}_p$ with a non-archimedean absolute value $|\cdot|_p$, such that:*

- *There exists an inclusion $\mathbb{Q} \hookrightarrow \mathbb{Q}_p$ via $x \to (x)$ and the absolute value induced by $|\cdot|_p$ in $\mathbb{Q}_p$ is the p-adic absolute value;*
- *The image of $\mathbb{Q}$ under this inclusion is a dense subset of $\mathbb{Q}_p$ with respect to the absolute value $|\cdot|_p$,*
- *$\mathbb{Q}_p$ is complete with respect to the absolute value $|\cdot|_p$.*

Proof. The 3 items stated were shown in the 3 previous propositions, respectively. □

3.2 Exploring $\mathbb{Q}_p$

The last theorem from the previous section completely characterizes the fields of p-adic numbers. In this section we will explore their inherent properties, so that the construction process is no longer in our radar.

From now on, we will identify $\mathbb{Q}$ with its image under $\mathbb{Q}_p$, in other words, we will think at the first as a subfield of the latter. Recalling the p-adic valuations, it is clear that they extend to $\mathbb{Q}_p$ in the expected way.

Lemma 3.43. *For each $0 \neq x \in \mathbb{Q}_p$, there exists an integer $v_p(x)$ such that $|x|_p = p^{-v_p(x)}$.*

As before, we put $v_p(0) = +\infty$ so that the extension is complete.

Recalling definition 3.20, and since now $\mathbb{Q}_p$ is a field with a non-archimedean absolute value, we might as well look at its valuation ring, which turns out to be crucial in the upcoming pages.

Definition 3.44. The ring of p-adic integers is the valuation ring

$$\mathbf{Z}_p = \{x \in \mathbb{Q}_p \colon |x|_p \leq 1\}.$$

The next result will expose its good properties.

Proposition 3.45. *The ring $\mathbf{Z}_p$ of p-adic integers is a local ring whose maximal ideal is the principal ideal $p\mathbf{Z}_p = \{x \in \mathbb{Q}_p \colon |x| < 1\}$. Furthermore,*

i) $\mathbb{Q} \cap \mathbf{Z}_p = \mathbb{Z}_{(p)} = \{\frac{a}{b} \in \mathbb{Q} \colon p \nmid b\}$, *i.e., the valuation ideal of* $\mathbb{Q}$ *with respect to* $|\cdot|_p$.

ii) The inclusion $\mathbb{Z} \hookrightarrow \mathbf{Z}_p$ has dense image. In particular, given $x \in \mathbf{Z}_p$ and $n \geq 1$, there exists a unique integer $0 \leq \alpha \leq p^n - 1$ such that $|x - \alpha| \leq p^{-n}$.

iii) For any $x \in \mathbf{Z}_p$, there exists a unique Cauchy sequence α_n converging to x satisfying the following properties:

- *For every $n \in \mathbb{N}$, $\alpha_n \in \mathbb{Z}$ satisfies $0 \leq \alpha_n \leq p^n - 1$,*
- *For every $n \in \mathbb{N}$, we have $\alpha_n \equiv \alpha_{n-1} \pmod{p^{n-1}}$.*

Proof. First of all, $\mathbf{Z}_p$ is a valuation ring, so it only has one maximal ideal, thus it is a local ring (recall Proposition 3.20).

Now, to see that the valuation ideal $\beta = \{x \in \mathbb{Q}_P \colon |x| < 1\}$ is generated by p, Lemma 3.43 comes in hand.

$$|x| < 1 \overset{3.43}{\implies} |x| \leq \frac{1}{p} \implies \left|\frac{x}{p}\right| \leq 1 \implies p \mid x \implies x \in p\mathbf{Z}_p.$$

Since the valuation ideal is a maximal ideal and since $p\mathbf{Z}_p \neq \mathbf{Z}_p$, the latter inclusion is enough to prove equality, as we wanted to show.

It remains to prove the items stated. i) is clear because we already know that $\mathbb{Z}_{(p)}$ is the valuation ring in $\mathbb{Q}$ corresponding to the p-adic valuation.

In order to check ii), let $x \in \mathbf{Z}_p$ and let $n \in \mathbb{N}$. By definition, $\mathbb{Q}$ is dense in $\mathbb{Q}_p$ and since $x \in \mathbf{Z}_p \subseteq \mathbb{Q}_p$, it follows that there exists $\frac{a}{b} \in \mathbb{Q}$ such that

$$\left|x - \frac{a}{b}\right| \leq p^{-n} < 1.$$

We want to choose an integer in the place of $\frac{a}{b}$. First, notice that, by the non-archimedean property, we have

$$\left|\frac{a}{b}\right| \leq \max\left\{|x|, \left|x - \frac{a}{b}\right|\right\} \leq 1,$$

which implies that $\frac{a}{b} \in \mathbb{Z}_{(p)}$, that is, $p \nmid b$. From here it follows that there exists an integer b' such that $bb' \equiv 1 \pmod{p^n}$, that is, $p^n \mid (bb' - 1)$. Thus $|bb' - 1| \leq p^{-n}$ and hence

$$\left|\frac{a}{b} - bb'\right| = \left|\frac{a(1-bb')}{b}\right| = \left|\frac{a}{b}\right| |1 - bb'| \leq p^{-n}.$$

Now just choose α as the unique integer such that

$$0 \leq \alpha \leq p^n - 1 \text{ and } \alpha \equiv ab' \pmod{p^n}.$$

We claim that this α will do, for which it remains to check that $|x - \alpha| \leq p^{-n}$. First, notice that

$$|x - ab'| \leq \max\left\{\left|x - \frac{a}{b}\right|, \left|\frac{a}{b} - ab'\right|\right\} \leq p^{-n}.$$

Now, $\alpha \equiv ab' \pmod{p^n}$ implies that there exists $y \in \mathbb{Z}$ such that $ab' - \alpha = p^n y$, so that $x - \alpha = x - (ab' - p^n y) = x - ab' + p^n y$. Therefore,

$$|x - \alpha| = |x - ab' + p^n y| \leq \max\{|x - ab'|, |p^n y|\} \leq p^{-n},$$

as we wanted to show.

Finally, (iii) follows from (ii) for the sequence of positive integers. All properties are true by construction. □

3.2.1 Elementary analysis in $\mathbb{Q}_p$

3.2.1.1 Sequences and Series

Some elementary results which we will need will be stated in this section.

Lemma 3.46. *A sequence $(a_n)_n$ in $\mathbb{Q}_p$ is a Cauchy sequence, and so convergent, if and only if it satisfies*

$$\lim_{n\to\infty} |a_{n+1} - a_n| = 0.$$

Proof. Same proof as Lemma 3.29. □

Here are some examples where we will use the lemma.

Examples 6.

- The sequence $(a_n)_n$ given by $a_n = n!$ converges to 0 in $\mathbb{Q}_p$, because $(n+1)! - n! = n!(n+1-1) = nn!$ and $\lim_{m\to\infty} |p^m (p^m)!| = 0$;

- The sequence $(a_n)_n = \frac{1}{n}$ diverges, due to $\frac{1}{n+1} - \frac{1}{n} = \frac{n-(n+1)}{n(n+1)} = -\frac{1}{n(n+1)}$ whose absolute value has no limit as $n \to \infty$, since $\left|\frac{1}{n(n+1)}\right| = 1$ if $p \nmid n, n+1$, and $\left|\frac{1}{n(n+1)}\right| = p$ otherwise,
- The sequence $(a_n)_n = p^n$ clearly converges to 0.

Now we need a similar result for series, which in the case of the p-adic world, will be a very convenient one.

Corollary 3.47. *An infinite series $\sum_{n=0}^{\infty} a_n$ with $a_n \in \mathbb{Q}_p$ is convergent if and only if*

$$\lim_{n\to\infty} a_n = 0,$$

in which case we also have

$$\left|\sum_{n=0}^{\infty} a_n\right| \leq \max_n |a_n|.$$

Proof. It is known that a series converges when the sequence of partial sums does. Since we are dealing with $\mathbb{Q}_p$, then the sequence of partial sums converges if and only if it is Cauchy. Realizing that the n-th term a_n is the difference between the n-th and $(n-1)$-th partial sums, we see that the sequence of partial sums is a Cauchy sequence if and only if $\lim_{n\to\infty} a_n = 0$ according to the previous lemma.

For the inequality part, if $\sum_{n=0}^{\infty} = 0$, it is trivial. Otherwise, for any $N \in \mathbb{N}$, the non-archimedean property gives

$$\left|\sum_{n=0}^{N} a_n\right| \leq \max_n |a_n|. \tag{3.3}$$

Now, because the a_n tend to zero, for large enough N one has

$$\max_{0\leq n\leq N} |a_n| = \max_n |a_n|,$$

so that Lemma 3.37 gives

$$\left|\sum_{n=0}^{\infty} a_n\right| = \left|\sum_{n=0}^{N} a_n\right|.$$

Hence, the result follows by the inequality 3.3. □

3.2.1.2 Functions: continuity and derivatives

Finally we arrive at the last section before the actual goal of this thesis. The definitions here stated will be used to understand the applications that will be seen in the end. Firstly, the continuity.

Definition 3.48. Let $U \subseteq \mathbb{Q}_p$ be an open set. A function $f\colon U \to \mathbb{Q}_p$ is said to be *continuous* at $a \in U$ if for every $\epsilon > 0$ there exists $\delta > 0$ such that, for every $x \in U$,

$$|x - a| < \delta \implies |f(x) - f(a)| < \epsilon.$$

Next is the definition of derivative.

Definition 3.49. Let $U \subseteq \mathbb{Q}_p$ be an open set and let $f\colon U \to \mathbb{Q}_p$ be a function. We say that f is *differentiable* at $x \in U$ if the limit

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$$

exists. If that is the case for every $x \in U$, we say that f is *differentiable* in U and we write $f'\colon U \to \mathbb{Q}_p$ for the function $x \to f'(x)$.

Chapter 4

p-adic formulation of the Collatz Conjecture

4.1 The Two-adic Integers

First of all, it is worth mentioning a distinct perspective of looking at the 2-adics, which arises from generalizing from the set of natural numbers. This point of view can be seen in the Ethan Atkin's article.

4.1.1 Construction

So, why is the $3n + 1$ problem connected to the 2-adics? Notice that dividing an even number by 2 in base 2 is just removing the digit of its binary expansion corresponding to the exponent 0, which is a 0. Besides, the other operation in the Collatz function is not that difficult either in base 2, because multiplying by 3 is the same as multiplying by 2 (which corresponds to a shift in the binary expansion) and adding the number, and then we just have to add 1. We can give an explicit formula for doing this.

We will write any natural number a in base 2 as a finite sequence $a_0a_1a_2\ldots = a_0 2^0 + a_1 2^1 + a_2 2^2 + \cdots$, with each $a_i \in \mathbb{Z}_2$. Notice that the base 2 digits are being written in the reverse of the standard order, which is to write first the biggest power of 2 and all the way down to the lowest one, so that when performing the sum, we can carry to the right, which is more convenient to perform the desired extension and matches with the usual sinistrodextral script. Indeed, letting $a = a_0a_1a_2\ldots$, $b = b_0b_1b_2\ldots$ be two finite sequences of this form, which just means that a_i, b_i are 0, for all sufficiently large i, in order to sum a

with b one can write:

$$\begin{aligned} & a_0a_1a_2\ldots \\ +\ & \underline{b_0b_1b_2\ldots} \\ & c_0c_1c_2\ldots \end{aligned}$$

so that the answer, $c = c_0c_1c_2\ldots$, can be computed analogously to the addition algorithm in base 10, the only differences being in the carry occurring at 2 instead of 10 and the order of the digits that we are considering, which is reverse of the standard.

And in order to multiply a by b, one can sum the products of a by every digit in b, where the i-th row below is included in the sum if and only if $b_i = 1$.

$$\begin{aligned} & a_0a_1a_2\ldots \\ \times\ & \underline{b_0b_1b_2\ldots} \\ b_0 \cdot\ & a_0a_1a_2\ldots \\ b_1 \cdot\ & \quad a_0a_1\ldots \\ b_2 \cdot\ & \qquad a_0\ldots \\ \ldots \cdot\ & \qquad\quad \ldots \\ +\ & \_\_\_\_\_ \\ & d_0d_1d_2\ldots \end{aligned}$$

Notice that in order to compute the i-th place of the answer, at most $2i$ 1's are involved in the sum, which correspond to the first $i+1$ multiplications in the rows regarding the i-th digit and at most $i-1$ carries.

Notice that these operations will work just fine if any of the sequences is of infinite length, that is, $a_0a_1a_2\ldots$ being infinite means that for all $n \in \mathbb{N}$, there exists $m \geq n$ such that $a_m = 1$. Thus, they just allow one to compute as many bits as one wants when both sequences are infinite.

As one has seen in the previous chapter, the infinite sequences of this type, together with these two operations are called the *two-adic integers*, which are denoted by $\mathbf{Z}_2$. This set is uncountable, because it is in a clear bijection with $2^{\mathbb{N}}$, that is, it has the same cardinality as the set of parts of $\mathbb{N}$.

So, let us recap what we have just done. We constructed an uncountable set, closed for the sum and product, by generalizing from binary sequences of positive integers. At

first glance, since subtraction and division applied to two natural numbers do not usually yield a natural number, we could guess that may be the case with respect to the 2-adics as well. But it is not, at least concerning subtraction, and that is clearly good news.

In order to see this, we will first try to write the negative integers in base 2, because in order to prove our claim, all the negative integers must be possible to write in base 2. For instance, how can -1 be written in base 2?

Since $1 + (-1) = 0$, and noticing that

$$\begin{array}{r} 1111\ldots \\ +\ \underline{1000\ldots} \\ 0000\ldots \end{array}$$

one gets that $-1 = 1111\ldots$. From here, it is just a matter of adding -1 enough times to reach any negative integer and thus, its base 2 expansion, however this is a long and boring process. Fortunately, there is a straightforward way of doing so. Given $a = a_0a_1a_2\ldots \in \mathbf{Z}_2$, we define $\overline{a} \in \mathbf{Z}_2$ by $\overline{a}_i = 0$ if $a_i = 1$ and vice-versa. It is clear that $a + \overline{a} = 1111\ldots = -1$, so that $a + (\overline{a} + 1) = 0$.

Thus, $-a = \overline{a} + 1$, hence the binary expansion of any negative integer can be therefore computed through the binary expansion of its symmetric, which is at easy reach.

Noticing that the sequence corresponding to any natural number terminates in a string of 0's, it is clear that the sequence corresponding to any negative integer terminates in a string of 1's. Conversely, any binary sequence terminating in a string of 0's represents a natural number in base 2 and if it ends in a string of 1's, it represents a negative integer.

So, now, a question that arises is if one can extend these base 2 expansions to the rational numbers. Indeed, every rational number can be written as an infinite binary sequence except for the rational numbers whose denominator is even. This has to do with how one can define division in the 2-adic integers.

Given a 2-adic $a = a_0a_1a_2\ldots$, we will naturally call it even if $a_0 = 0$ and odd otherwise. Now, notice that multiplying any a by 2 yields $2a = 0a_0a_1a_2\ldots$, that is, a 2-adic is even if and only if it is divisible by 2, meaning that dividing it by 2 yields a 2-adic. On the other hand, one cannot divide any odd 2-adic by 2, and hence, by any even 2-adic. Therefore, division in the 2-adics is only defined when the denominator is odd (which already implies that the 2-adics do not form a field, since in any field every nonzero element has an

inverse, that is, one can divide by any nonzero element (which is the same as multiplying by its inverse)).

We can of course apply the division algorithm in order to divide a 2-adic b by an odd 2-adic a:

$$\begin{array}{r} b_0b_1b_2\ldots \,|\underline{1a_1a_2\ldots} \\ -\underline{b_0c_1c_2\ldots}\; b_0\ldots \\ d_1d_2\ldots \end{array}$$

where

$$c_1c_2\ldots = b_0 \times a_1a_2\ldots$$

and then continue the division process for as long as we want.

However, there is another approach: to compute the binary expansion of the reciprocal of a, $\frac{1}{a}$, by reversing the multiplication algorithm, and then simply multiply b by $\frac{1}{a}$ using the multiplication algorithm above. Letting $\frac{1}{a} = a'_0a'_1a'_2\ldots$ and operating modulo 2, we can exemplify how to compute the first 3 bits of the reciprocal, the remaining bits being possible to compute in a similar way:

$$\begin{array}{llll} a'_0\ 1a_1a_2\ldots & 1\ \ 1a_1a_2\ldots & 1\ \ 1a_1a_2\ldots & 1\ \ 1a_1a_2\ldots \\ a'_1\quad 1a_1\ldots & a'_1\quad 1a_1\ldots & a_1\quad 1a_1\ldots & a_1\quad 1a_1\ldots \\ a'_2\qquad 1\ldots \overset{a'_0=1}{\Longrightarrow} & a'_2\qquad 1\ldots \overset{a_1+a'_1=0}{\Longrightarrow} & a'_2\qquad 1\ldots \overset{a_2+a_1+a'_2+\left\lfloor\frac{a_1+a'_1}{2}\right\rfloor=0}{\Longrightarrow} & a_2\qquad 1\ldots \\ \ldots\qquad\ldots & \ldots\qquad\ldots & \ldots\qquad\ldots & \ldots\qquad\ldots \\ +\ \underline{\qquad\qquad} & +\ \underline{\qquad\qquad} & +\ \underline{\qquad\qquad} & +\ \underline{\qquad\qquad} \\ 1000\ldots & 1000\ldots & 1000\ldots & 1000\ldots \end{array}$$

Notice that the reciprocal of a always coincides with a in the first 3 bits, but it may not coincide from the fourth one on, and that is the reason why we stopped computing them above, since

- $a'_0a_3 + a'_1a_2 + a'_2a_1 + a'_3 + \left\lfloor\frac{a'_0a_2+a'_1a_1+a'_2+a_1}{2}\right\rfloor = 0 \implies a_3 + 2a_1a_2 + a'_3 + a_1 + a_2 = 0 \implies a'_3 = a_1 + a_2 + a_3$.

We might as well consider a vital example, which will be used in the future:

Example 4.1. *Since* $3 = 11000\ldots$*, then we can find the binary expansion of* $\frac{1}{3}$*, say,* $b_0b_1b_2\ldots$*:*

$$
\begin{array}{rrl}
b_0 = 1 & 1100000\ldots & \textit{clearly} \\
b_1 = 1 & 110000\ldots & \textit{since } b_0 + b_1 = 0 \\
b_2 = 0 & 11000\ldots & \textit{since } b_1 + b_2 + 1 = 0 \\
b_3 = 1 & 1100\ldots & \textit{since } b_2 + b_3 + 1 = 0 \\
b_4 = 0 & 110\ldots & \textit{since } b_3 + b_4 + 1 = 0 \\
b_5 = 1 & 11\ldots & \textit{since } b_4 + b_5 + 1 = 0 \\
\cdots & \ldots & \ldots \\
+\_\_\_\_\_ & & \\
1000000\ldots & &
\end{array}
$$

That is, $\frac{1}{3} = 110101\ldots$. *Also, from* $3 \times 101010\ldots = 1111\ldots$, *we conclude that* $-\frac{1}{3} = 101010\ldots$, *and this expansion matches with the one we would get if we consider that* $-\frac{1}{3} = \overline{\frac{1}{3}} + 1$*:* $\overline{\frac{1}{3}} + 1 = 001010\ldots + 1000\ldots = 101010\ldots$.

4.1.2 Congruences

Congruences played a major role in the beginning of the previous chapter. We are going to apply to them again in order to view the set of 2-adics from an abstract algebraic perspective, which itself will be of extreme importance in understanding some of the following results.

It is easy to see that congruence preserves both addition and multiplication, so that for every positive integer k, both these operations can be defined in the modulo 2^k congruence classes, from where the ring of integers modulo 2^k arises, which we know by $\mathbb{Z}/2^k$.

Identifying every positive integer with its congruence class mod 2^k, two natural numbers are congruent modulo 2^k if and only if their binary expansions agree on the first k bits. Moreover, adding and multiplying in these congruence classes only requires to doing the computations until the referred bits.

An interesting function arises from all of this, which is the function $\rho_k \colon \mathbb{Z}/2^k \to \mathbb{Z}/2^{k-1}$, the obvious restriction mapping which can be defined by just erasing the last bit. This map is a ring homomorphism with respect to the two operations mentioned above, and this is easy to check as well.

We can go much further and define for every k the restriction mapping from $\mathbf{Z}_2$ to the ring of integers mod 2^k, that is, that only accounts for the first k bits. These maps also preserve the two operations.

For every 2-adic a, we will denote by $[a]_k$ its modulo 2^k congruence class, that is:

$$[a]_k \stackrel{\text{def}}{=} \{b \in \mathbf{Z}_2 \colon b \equiv a \bmod 2^k\}$$

Hence, for $a, b \in \mathbf{Z}_2$, $[a]_k = [b]_k \Leftrightarrow a_i = b_i$ for $i = 0, 1, \ldots, k-1$.

There is a correspondence between the restriction mapping and set inclusion in the following way.

$$\rho_k([a]_k) = [b]_{k-1} \Leftrightarrow a \equiv b \pmod{2^{k-1}} \Leftrightarrow [a]_k \subsetneq [b]_{k-1}$$

In fact, given a sequence $\{\alpha_k \in \mathbb{Z}/2^k \colon k \in \mathbb{N}\}$, we say that it is *coherent* if $p_k(\alpha_k) = \alpha_{k-1}$ for every $k > 1$. Clearly, $a \in \mathbf{Z}_2$ is uniquely described by the coherent sequence of congruence classes $\{[a]_k \colon k \in \mathbb{N}\}$. As we run along the coherent sequence, the new information provided by α_k, given α_{k-1}, is precisely the k-th bit in the base 2 expansion: whether it is a 0 or a 1.

Now, this whole business allows us to confirm that the usual arithmetic rules are true for $\mathbf{Z}_2$, since the commutative, associative and distributive laws are inherited from the rings $\mathbb{Z}/2^k$.

One still needs a few definitions and properties inherent to them in order to proceed.

Definition 4.1. A 2-adic x is called *rational* if it can be written as $x = \frac{a}{b}$ for some integers a and b with b odd.

Remark 4.2. Not all rational numbers are rational 2-adics. Having the denominator odd is a sufficient and necessary condition to a rational number be a rational 2-adic.

The following definition will be proved later to be equivalent to the latter, constituting another characterization of the former.

Definition 4.3. A 2-adic x is said to be *terminally periodic* if its binary expansion at some point becomes periodic, that is, if there exist $k, i \in \mathbb{N}$ such that

$$\forall\, j \geq i\,,\, x_{j+k} = x_j.$$

The smallest such k, if it exists, is called the *period* of x.

Remark 4.4. In the particular case that $x_{i+k} = x_i$ for all $i \in \mathbb{N}$, we get that x is periodic.

Remark 4.5. As we might recall from the previous section, the integers are precisely the 2-adics which are terminally periodic of period 1, and $0 = 000\ldots$, $-1 = 111\ldots$ are the only ones periodic of period 1.

We will begin by giving a more concrete perspective of a rational 2-adic.

Lemma 4.6. *If $x = x_0x_1\ldots \in \mathbf{Z}_2$ is periodic of period k, then $x = -\frac{a}{2^k-1}$ for some $a \in \mathbb{N}_0$.*

Proof. If $k = 1$, then either $x = 0$ and we can put $a = 0$, or $x = -1$ and we can put $a = 1$. So, we may assume $k > 1$, so that $x = x_0x_1\ldots x_{k-1}x_0x_1\ldots x_{k-1}\ldots$. We can also assume that x is odd, because if it is even, then $x = 2^l y$ for some $l \in \mathbb{N}$ and an odd 2-adic y, so that if we can write $y = -\frac{a}{2^k-1}$ with $a \in \mathbb{N}$, then $x = -\frac{2^l a}{2^k-1}$ does the job. Now, notice that

$$\begin{aligned} 2^k x &= \underset{k \text{ zeros}}{00\ldots0}1x_1\ldots x_{k-1}x_0x_1\ldots x_{k-1}\ldots \\ \implies \overline{2^k x} &= \underset{k \text{ ones}}{11\ldots1}0\overline{x_1}\ldots\overline{x_{k-1}x_0x_1}\ldots\overline{x_{k-1}}\ldots \\ \implies \overline{2^k x}+1 &= \underset{k \text{ zeros}}{00\ldots0}1\overline{x_1}\ldots\overline{x_{k-1}}. \end{aligned}$$

Hence, we can compute $x - 2^k x = x + (\overline{2^k x} + 1)$ as follows

$$\begin{aligned} & x_0x_1\ldots x_{k-1}1x_1\ldots x_{k-1}\ldots \\ +&\underline{00\ldots01\overline{x_1}\ldots\overline{x_{k-1}}} \\ & x_0x_1\ldots x_{k-1}000\ldots \end{aligned}$$

This means that $x - 2^k x = a$ for the positive integer $a = x_0x_1\ldots x_{k-1}000\ldots$; hence, $x = -\frac{a}{2^k-1}$, as we wanted to show. □

The next result explores the properties of terminally periodic numbers.

Lemma 4.7. *If $x \in \mathbf{Z}_2$ is terminally periodic, then $-x$ is terminally periodic.*

Proof. We may write x as $x = x_0x_1\ldots x_{k-1}x_k\ldots x_{k+\ell}x_k\ldots x_{k+\ell}\ldots$. Thus, the 2-adic expansion of $-x = \overline{x} + 1$ can be computed as

$$\begin{aligned} & \overline{x_0x_1}\ldots\overline{x_{k-1}x_k}\ldots\overline{x_{k+\ell}x_k}\ldots\overline{xk+\ell}\ldots \\ +\ & \underline{10\ldots00\ldots00\ldots0} \end{aligned}$$

so that the only way not to have $-x$ with the same period after the k-th term is to have a carry in all the sums before that one and to have a carry in all of the sums corresponding to the bits of the period. But for that to happen, we must have $\overline{x_k}\ldots\overline{x_{k+\ell}} = 11\ldots1$, so that $x \in \mathbb{N}_0$ and so $-x$ is an integer, which is terminally periodic. □

We still need another result in order to prove the equivalence of definitions 4.1 and 4.7.

Lemma 4.8. *If n is an odd natural number, then n divides $2^m - 1$ for some $m \in \mathbb{N}$.*

Proof. Let $\mathbb{N} \ni m = \Phi(n)$, that is, the number of natural numbers smaller than n which are prime with n. Since n is odd, then 2 is prime with n and so by Euler's theorem from an elementary Number Theory course, we get that

$$2^m \equiv 1 \pmod{n},$$

that is, $n \mid 2^m - 1$, as we wanted to show. □

We are now ready to state and prove the last result of this section.

Proposition 4.9. *A 2-adic x is rational if and only if it is terminally periodic.*

Proof.

(if) Suppose that x is terminally periodic. Then we can find n, $\ell \in \mathbb{N}_0$ and $y \in \mathbf{Z}_2$ periodic of period k such that

$$x = n + 2^\ell y.$$

By Lemma 4.6, $y = -\frac{a}{2^k - 1}$ for some $a \in \mathbb{N}_0$, so that we get

$$x = n + 2^\ell \left(-\frac{a}{2^k - 1}\right) = n + \left(-\frac{a2^\ell}{2^k - 1}\right) = \frac{-2^\ell a + n(2^k - 1)}{2^k - 1},$$

thus, x is a rational 2-adic.

(only if) Suppose now that x is a rational 2-adic, say, $x = \frac{a}{b}$ for some integers a, b with b positive and odd. We may assume that $b \neq 1$, otherwise x is an integer and so trivially terminally periodic. By Lemma 4.6, it follows that $b \mid 2^m - 1$ for some $m \in \mathbb{N}$, say, $2^m - 1 = bz$ with $z \in \mathbb{N}$. Thus,

$$x = \frac{a}{b} = \frac{a}{\frac{2^m - 1}{z}} = \frac{az}{2^m - 1} = az\frac{1}{2^m - 1} = -az\sum_{i=0}^{\infty} 2^{im}.$$

We may assume that $-az \in \mathbb{N}$, so that a is negative. Now, if $-az < m$, then by writing $-az$ in base 2 as $-az = a_0 + 2a_1 + 4a_2 + \ldots + 2^k a_k$, it is clear that

$$-az\sum_{i=0}^{\infty} 2^{im} = a_0 + 2a_1 + 4a_2 + \ldots + 2^k a_k + a_0 2^m + a_1 2^{m+1} + \ldots + a_k 2^{m+k} + \ldots$$

is terminally periodic, because it is periodic.

Otherwise, let m_0 be the smallest positive integer such that $-az < 2^{m_0}$. Hence, we get $-az = 2^{m_0 - 1} + b_0 + 2b_1 + \ldots 2^j b_j$ for some $j \in \mathbb{N}$ with $j < m$ and $b_i \in \{0, 1\}$. Now, it is

clear that

$$-az\sum_{i=0}^{\infty}2^{im} = b_0 + \ldots + b_j 2^j + b_0 2^m + \ldots + b_j 2^{m+j} + \ldots + 2^{m_0-1} + b_0 2^{m_0} + \ldots + b_j 2^{m_0+j} + \ldots + 2^{2m_0-1} + \ldots$$

is a terminally periodic number of period $\leq j(m_0 - 1) + 1$. Lemma 4.7 does the job for $az > 0$, therefore the proof is complete. □

4.1.3 The Two-adic Shift Map

Before dealing concretely with the Collatz function itself, we are going to consider a gross simplification of the previously mentioned function by omitting the multiplication by 3. By doing this, we will see that the easy to check properties of the first one are no longer inherent to the latter.

Starting with an odd natural number x, divide away the highest power of 2 in $x + 1$ to get the odd number $S(x)$. In this way, $S(1) = 1$ and $S(x) < x$ otherwise. It is clear that the sequence of iterates of S, $(S^n(x))_n$, decreases monotonically to 1.

If we divide only by 2 at a time, as in the Collatz function, the function can be defined as follows.

$$s(x) = \begin{cases} \dfrac{x}{2} & \text{if } x \text{ is even,} \\ \dfrac{x+1}{2} & \text{if } x \text{ is odd.} \end{cases}$$

More generally, suppose we are given a non-empty set X and a function q in X. For each $x \in X$, the q-orbit of x is the sequence in X: $\{x, q(x), q \circ q(x), \ldots\} = \{q^n(x) : n \in \mathbb{N}_0\}$.

In our case, starting with an odd natural number x, it is clear that the S-orbit of x is precisely the subsequence of odd numbers in the s-orbit of x.

We already know that even 2-adics can be uniquely divided by 2, so that we can define a function $s : \mathbf{Z}_2 \to \mathbf{Z}_2$ as before. In order to introduce the shift map on $\mathbf{Z}_2$, since adding 1 to an odd number will most likely interfere with its base 2 expansion after the first bit, we have to go the another way around, literally, by subtracting 1. In other words, we define the shift map σ on $\mathbf{Z}_2$ as $\sigma(x) = -s(-x)$, that is,

$$\sigma(x) = \begin{cases} \dfrac{x}{2} & \text{if } x \text{ is even,} \\ \dfrac{x-1}{2} & \text{if } x \text{ is odd.} \end{cases}$$

Writing a 2-adic x as $x = x_0x_1x_2\ldots$, we get that

$$\sigma(x)_i = x_{i+1} \text{ for } i = 0, 1, 2, \ldots \tag{4.1}$$

Concerning now the σ-orbit of x, it is clear that

$$x_i = \sigma^i(x)_0 \text{ for } i = 0, 1, 2, \ldots \tag{4.2}$$

More simplified,

$$x_i = \begin{cases} 0 & \text{if } \sigma^i(x) \text{ is even,} \\ 1 & \text{if } \sigma^i(x) \text{ is odd.} \end{cases}$$

Concerning the shift map, it is clear that it is onto (for any $y \in \mathbf{Z}_2$, $\sigma(2y) = y$) and it is also clear that is is not one-to-one, since for any $y \in \mathbf{Z}_2$, $y = \sigma(2y) = \sigma(2y+1)$.

We will now look at some special σ-orbits. Given $x \in \mathbf{Z}_2$, we have:

- x is a nonnegative integer if and only if $\sigma^i(x) = 000\ldots = 0$ for some $i \in \mathbb{N}_0$;
- x is a negative integer if and only if $\sigma^i(x) = 111\ldots = -1$ for some $i \in \mathbb{N}_0$;

In fact, 0 and -1 are the only fixed points of σ and the integers are the 2-adics whose orbits arrive at these fixed points at some point.

- σ has one cycle of period 2. Recalling that $-\frac{1}{3} = 101010\ldots$ and $-\frac{2}{3} = 010101\ldots$, it is clear that $\sigma\left(-\frac{1}{3}\right) = -\frac{2}{3}$ and $\sigma\left(-\frac{2}{3}\right) = \frac{1}{3}$. It is the only cycle of period 2, because for that to happen, the first two bits of the expansion must repeat over time and be different from one another (otherwise it would be a fixed point).
- From the previous item, it is clear that $x \in \mathbf{Z}_2$ is periodic if and only if x is a fixed point for some iterate of σ, that is, $\sigma^k(x) = x$ for some $k \in \mathbb{N}$. In other words, the base 2 expansion of x is periodic if and only if its σ-orbit is periodic.
- Using the last result of the previous section, we see that x is rational if and only if at some point its σ-orbit becomes periodic.

In conclusion, the shift map is a function very easily to study. On the contrary, concerning the $3x+1$ map, despite only adding multiplication by 3, and apart from the trivial cycle, there are no other clues about the existence or not of any more cycles.

4.1.4 The Two-adic $3x + 1$ Map

As one already know from the previous chapter, there is a bijection between the increasing sequences of non-negative integers, finite or infinite, and the 2-adic integers in the following way:

$$(d_0, d_1, \ldots) \longleftrightarrow 2^{d_0} + 2^{d_1} + \cdots$$

We are now interested in considering a particular type of functions defined in the 2-adic integers. Let u_0, u_1, ... be odd 2-adics, that is, $u_i \in 1 + 2\mathbf{Z}_2$, for all $i \in \mathbb{N}_0$. For any increasing sequence, finite or infinite, of non-negative integers $0 \leq d_0 < d_1 < \cdots$, the sum

$$\sum_{i=0}^{\infty} u_i 2^{d_i}$$

converges to some 2-adic integer x, since $|u_n 2^n|_2 = |u_n|_2 |2^n|_2 = |2^n|_2 \to 0$ as $n \to \infty$.

The following map will play a crucial role in the p-adic formulation of the Collatz conjecture.

Definition 4.10. Let Φ be the function on $\mathbf{Z}_2$ defined by:

$$\sum_{n=0}^{\infty} 2^{d_n} \xrightarrow{\Phi} -\sum_{n=0}^{\infty} \frac{1}{3^{n+1}} 2^{d_n}$$

For instance,

$$\begin{aligned}
\Phi\left(-\frac{1}{3}\right) &= \Phi(2^0 + 2^2 + 2^4 + \cdots) \\
&= \frac{-1}{3}2^0 + \frac{-1}{9}2^2 + \frac{-1}{27}2^4 + \cdots \\
&= -\frac{1}{3}\sum_{i=0}^{\infty}\left(\frac{4}{3}\right)^i \\
&\overset{(*)}{=} \frac{-\frac{1}{3}}{1 - \frac{4}{3}} \\
&= \frac{-\frac{1}{3}}{-\frac{1}{3}} \\
&= 1,
\end{aligned}$$

where (*) follows from applying the geometric series formula, which is valid in $\mathbb{Q}_p$ with an analogous proof to $\mathbb{R}$ (notice that the ratio $\frac{4}{3}$ is such that $\left|\frac{4}{3}\right|_2 = \frac{1}{4} < 1$, so that the formula is valid).

We start by showing that Φ is a permutation of the 2-adic integers, as a corollary of the following result:

Proposition 4.11. *Given $u_0, u_1, \ldots \in 1 + 2\mathbf{Z}_2$, the function $F : \mathbf{Z}_2 \to \mathbf{Z}_2$ defined by:*

$$\sum_{n=0}^{\infty} 2^{d_n} \xrightarrow{F} \sum_{n=0}^{\infty} u_n 2^{d_n}$$

is a permutation.

Proof. (injectivity) Suppose $0 \leq a_0 < a_1 < \cdots$ and $0 \leq b_0 < b_1 < \cdots$ are two distinct increasing sequences of integers, finite or infinite, and let $i \in \mathbb{N}_0$ be the smallest index such that $a_i \neq b_i$.

If it were $\sum_{n=0}^{\infty} u_n 2^{a_n} = \sum_{n=0}^{\infty} u_n 2^{b_n}$, then it would imply $\sum_{n=i}^{\infty} u_n 2^{a_n} = \sum_{n=i}^{\infty} u_n 2^{b_n}$. But since $|\sum_{n=i}^{\infty} u_n 2^{a_n}|_2 = \max\{|u_i 2^{a_i}|_2 + |\sum_{n>i}^{\infty} u_n 2^{a_n}|_2\} = |u_i 2^{a_i}|_2 = |2^{a_i}|_2 = a_i$, with an equivalent sequence of equalities for the other increasing sequence, we conclude that we must have $a_i = b_i$, which contradicts the definition of i. Thus, this implies injectivity of F.

(surjectivity) Given $x \in \mathbf{Z}_2$, we just need to find an increasing sequence $0 \leq d_0 < d_1 < \cdots$, which can be finite or infinite, such that $F(2^{d_0} + 2^{d_1} + \cdots) = x$, that is, such that $x = \sum_{i=0}^{\infty} u_i 2^{d_i}$. Since x is a 2-adic, we can write it in the form $x = 2^{x_0} + 2^{x_1} + \cdots$, for an increasing sequence of integers $0 \leq x_0 < x_1 < \cdots$. For easier future computations, we may write $u_i = 1 + y_i$ for every $i \in \mathbb{N}_0$. Now, just take:

- $d_0 = x_0$.

If $F(2^{d_0}) = x$, then we are done. Otherwise, there exist $x_{1_0}, x_{1_1}, \ldots$ such that $x - u_0 2^{d_0} = 2^{x_{1_0}} + 2^{x_{1_1}} + \cdots$. Notice that $x_{1_0} > d_0$, because $x - u_0 2^{d_0} = x - (1 + y_0)2^{x_0} = 2^{x_0} + 2^{x_1} + \cdots - (2^{x_0} + y2^{x_0}) = 2^{x_1} + 2^{x_2} + \cdots - y2^{x_0} \in 2^{d_0+1}\mathbf{Z}_2$.

- $d_1 = x_{1_0}$;

Again, if $F(2^{d_0} + 2^{d_1}) = x$, we are done. Otherwise, there exist $x_{2_0}, x_{2_1}, \ldots$ such that $x - u_0 2^{d_0} - u_1 2^{d_1} = 2^{x_{2_0}} + 2^{x_{2_1}} + \cdots$, $x_{2_0} > d_1$ and we can take:

- $d_2 = x_{2_0}$;

We can continue with this process as long as we want, stopping it if for some m we get $F(2^{d_0} + 2^{d_1} + \cdots + 2^{d_m}) = x$. The general term for d_i is, thus, $d_i = x_{i_0}$, with $x - u_0 2^{d_0} - u_1 2^{d_1} - \cdots u_{i-1} 2^{d_{i-1}} = 2^{x_{i_0}} + 2^{x_{i_1}} + \cdots$. It is clear that the $d_i{}'s$ form an increasing sequence of non-negative integers.

In the case of never ending this process, it remains to check that this sequence $0 \leq d_0 < d_1 < \cdots$ provides x under F.

Since we have the convergence

$$\sum_{i=0}^{\infty} u_i 2^{d_i} \to x,$$

due to the fact that for all $n \in \mathbb{N}_0$, $|x - \sum_{i=0}^{n} u_i 2^{d_i}|_2 \leq |2^{d_n+1}|_2 \to 0$ as $n \to \infty$, it follows that the 2-adic $2^{d_0} + 2^{d_1} + \cdots$ is $x's$ pre-image under F. But x was an arbitrary 2-adic, hence F is onto.

Therefore, F is a permutation of the 2-adic integers, as we wanted to show. □

We can now state a p-adic conjecture, that is shown below to be equivalent to the $3n + 1$ conjecture.

Conjecture 1. $\mathbb{N} \subseteq \Phi\left(\frac{1}{3}\mathbb{Z}\right)$.

In order to prove the equivalence, besides Φ, we will consider the two functions on the 2-adics $H, C\colon \mathbf{Z}_2 \to \mathbf{Z}_2$, defined as:

$$H(x) = \begin{cases} \dfrac{x}{2} & \text{if } x \text{ is even,} \\ x - 1 & \text{if } x \text{ is odd.} \end{cases}$$

$$C(x) = \begin{cases} \dfrac{x}{2} & \text{if } x \text{ is even,} \\ 3x + 1 & \text{if } x \text{ is odd.} \end{cases}$$

That is, C is the Collatz function extended to the 2-adics, and H is a simplification of the Collatz function. It is easy to see that $H^{-1}\left(\frac{1}{3}\mathbb{Z}\right) \subseteq \frac{1}{3}\mathbb{Z}$, and hence,

$$\forall\, k \in \mathbb{N}\,,\, H^{-k}\left(\frac{1}{3}\mathbb{Z}\right) \subseteq \frac{1}{3}\mathbb{Z}. \tag{4.3}$$

The following equality of functions will be used in the proof:

Theorem 4.12. *$C \circ \Phi = \Phi \circ H$, that is, Φ defines a conjugacy between H and C.*

Proof. Let $x = 2^{x_0} + 2^{x_1} + 2^{x_2} + \cdots \in \mathbf{Z}_2$.

If x is even, then $x_0 > 0$ and we have:

$$\begin{aligned} C(\Phi(x)) &= C(\Phi(2^{x_0} + 2^{x_1} + 2^{x_2} + \cdots)) \\ &= C\left(\frac{-1}{3}2^{x_0} + \frac{-1}{9}2^{x_1} + \frac{-1}{27}2^{x_2} + \cdots\right) \\ &= \frac{-1}{3}2^{x_0-1} + \frac{-1}{9}2^{x_1-1} + \frac{-1}{27}2^{x_2-1} + \cdots \\ &= \Phi(2^{x_0-1} + 2^{x_1-1} + 2^{x_2-1} + \cdots) \\ &= \Phi\left(\frac{x}{2}\right) \\ &= \Phi(H(x)). \end{aligned}$$

If x is odd, then $x_0 = 0$ and we have:

$$\begin{aligned} C(\Phi(x)) &= C(\Phi(1 + 2^{x_1} + 2^{x_2} + \cdots)) \\ &= C\left(\frac{-1}{3} + \frac{-1}{9}2^{x_1} + \frac{-1}{27}2^{x_2} + \cdots\right) \\ &= 1 + 3\left(\frac{-1}{3} + \frac{-1}{9}2^{x_1} + \frac{-1}{27}2^{x_2} + \cdots\right) \\ &= \frac{-1}{3}2^{x_1} + \frac{-1}{9}2^{x_2} + \cdots \\ &= \Phi(2^{x_1} + 2^{x_2} + \cdots) \\ &= \Phi(x - 2^{x_0}) \\ &= \Phi(x - 1) \\ &= \Phi(H(x)). \end{aligned}$$

Hence, the proof is complete. □

From here, the next corollary follows by induction:

Corollary 4.13. *For any $k \in \mathbb{N}$, $C^k \circ \Phi = \Phi \circ H^k$.*

Proof. We will proceed by induction on $k \geq 1$. The case $k = 1$ is the previous Theorem. Now, suppose $k > 1$ and assume the induction hypothesis (I.H). We get:
$C^k \circ \Phi = (C \circ C^{k-1}) \circ \Phi = C \circ (C^{k-1} \circ \Phi) \overset{\text{I.H.}}{=} C \circ (\Phi \circ H^{k-1}) = (C \circ \Phi) \circ H^{k-1} \overset{4.3}{=} (\Phi \circ H) \circ H^{k-1} = \Phi \circ (H \circ H^{k-1}) = \Phi \circ H^k$, which completes the induction step. □

Now we can finally prove that the Conjecture 1 is equivalent to the Collatz conjecture:

Theorem 4.14. *The Collatz conjecture is true if and only if the Conjecture 1 is true.*

Proof. (only if) Let $n \in \mathbb{N}$ and $x = \Phi^{-1}(n)$. Assuming the Collatz conjecture to be true, there exists $k \in \mathbb{N}$ such that $C^k(n) = 1$, that is, $C^k\left(\Phi(x)\right) = 1$. By the previous corollary, $\Phi\left(H^k(x)\right) = 1$ and so $H^k(x) = \Phi^{-1}(1) = -\frac{1}{3} \in \frac{1}{3}\mathbb{Z}$. Now, 4.3 implies $x \in \frac{1}{3}\mathbb{Z}$ and hence, $n \in \Phi\left(\frac{1}{3}\mathbb{Z}\right)$, as desired.

(if) Conversely, assume that $\mathbb{N} \subseteq \Phi\left(\frac{1}{3}\mathbb{Z}\right)$. Let $n \in \mathbb{N}$, $x = \Phi^{-1}(n)$ and let $0 \leq x_0 < x_1 < \cdots$ be integers such that $x = 2^{x_0} + 2^{x_1} + \cdots$. Let also $X = (x_0, x_1, \ldots)$. From our hypothesis, it follows that $x \in \frac{1}{3}\mathbb{Z}$, which implies that $3x \in \mathbb{Z}$. Moreover, x cannot be an integer, because if it was, then either:

- $x = 0$, which implies $n = 0 \notin \mathbb{N}$, contradiction;
- $x > 0$, which implies that X is finite, and so, n is negative, another contradiction;
- $x < 0$, which implies that the base 2 expansion of s terminates in a string of 1's, that is, that $x_{i+1} = x_i + 1$ for all sufficiently large i. Letting j be the smallest such i, it follows that:

$$\begin{aligned}
n &= \frac{-1}{3}2^{x_0} + \cdots + \frac{-1}{3^j}2^{x_{j-1}} + \frac{-1}{3^{j+1-x_j}}\sum_{k=x_j}^{\infty}\left(\frac{2}{3}\right)^k \\
&= \frac{-1}{3}2^{x_0} + \cdots + \frac{-1}{3^j}2^{x_{j-1}} + \frac{-1}{3^{j+1-x_j}}\sum_{k=0}^{\infty}\left(\frac{2}{3}\right)^{k+x_j} \\
&= \frac{-1}{3}2^{x_0} + \cdots + \frac{-1}{3^{i_0}}2^{x_{j-1}} + \frac{-2^{x_j}}{3^{j+1}}\sum_{k=0}^{\infty}\left(\frac{2}{3}\right)^k \\
&= \frac{-1}{3}2^{x_0} + \cdots + \frac{-1}{3^j}2^{x_{j-1}} + \frac{-2^{x_j}}{3^{j+1}}\frac{1}{1-\frac{2}{3}} \\
&= \frac{-1}{3}2^{x_0} + \cdots + \frac{-1}{3^j}2^{x_{j-1}} + \frac{-2^{x_j}}{3^j} < 0, \text{ a contradiction.}
\end{aligned}$$

Hence, $x = y \pm \frac{1}{3}$ for some $y \in \mathbb{Z}$ and by our previous computation of $\frac{1}{3}$ in base 2, it follows that we have $x_{\ell+1} = x_\ell + 2$ for all sufficiently large ℓ. Let m be the smallest of such ℓ.

Now, the idea is to define a map D on the increasing sequences of the type of X corresponding to the action of C on n, that is:

$$D(x_0, x_1, x_2, \ldots) = \begin{cases} (x_0 - 1, x_1 - 1, x_2 - 1, \ldots) & \text{if } x_0 > 0, \\ (x_1, x_2, x_3, \ldots) & \text{if } x_0 = 0. \end{cases}$$

We claim that

$$D^{x_m+m}(x_0,\ldots,x_{m-1},x_m,x_m+2,x_m+4,\ldots)=(0,2,4,\ldots),$$

which implies that $C^{x_m+m}(n)=1$, thus proving that n is a wondrous number. In order to prove this, we will proceed by induction on x_m+m. We can already assume that $X\neq(0,2,4,\ldots)$, otherwise there is nothing to be shown.

The case $x_m+m=0$ is trivial. So, suppose $x_m+m>0$ and that we have already proved that $D^{y_z+z}(y_0,\ldots,y_{z-1},y_z,y_z+2,y_z+4,\ldots)=(0,2,4,\ldots)$ for every integer $0\leq z$ and every increasing sequence of integers $0\leq y_0<y_1<\cdots<y_z$ such that $y_z+z<x_m+m$, which is our induction hypothesis.

If $x_0=0$, which implies that $m>0$, then we get:

$$\begin{aligned}D^{x_m+m}(X)&=D^{x_m+m}(x_0,\ldots,x_{m-1},x_m,x_m+2,x_m+4,\ldots)\\&=D^{x_m+m}(0,\ldots,x_{m-1},x_m,x_m+2,x_m+4,\ldots)\\&=D^{x_m+(m-1)}(x_1,\ldots,x_m,x_m+2,x_m+4,\ldots)\\&=(0,2,4,\ldots)\text{ by the induction hypothesis.}\end{aligned}$$

Otherwise, if $x_0>0$, then we get:

$$\begin{aligned}D^{x_m+m}(X)&=D^{x_m+m}(x_0,\ldots,x_{m-1},x_m,x_m+2,x_m+4,\ldots)\\&=D^{x_m+m-1}(x_0-1,\ldots,x_{m-1}-1,x_m-1,x_m+2-1,x_m+4-1,\ldots)\\&=D^{(x_m-1)+m}(x_0-1,\ldots,x_{m-1}-1,x_m-1,(x_m-1)+2,(x_m-1)+4,\ldots)\\&=(0,2,4,\ldots)\text{ by the induction hypothesis}\end{aligned}$$

In any case, the Collatz sequence beginning at n reaches 1. Since n was an arbitrary natural number, this process yields for every natural number, therefore it implies the veracity of the Collatz conjecture, which completes the proof. □

The main goal of this thesis, which is to express the Collatz Conjecture from a p-adic point of view, is therefore complete. As well as the Collatz conjecture, conjecture 1 is a very simple statement, yet far from being proved.

4.2 Applications

We will be following Bernstein's article here.

The equivalence of the conjectures has several applications. The first one we will state is the following.

Corollary 4.15. $\Phi(\mathbb{Q} \cap \mathbf{Z}_2) \subseteq \mathbb{Q} \cap \mathbf{Z}_2$.

Proof. Let $x = 2^{x_0} + 2^{x_1} + \cdots \in \mathbb{Q} \cap \mathbf{Z}_2$ and $y = \Phi(x)$. Then either $x \in \mathbb{N}_0$ or x is a negative integer or x is a rational in $\mathbf{Z}_2$. Let us look at the three cases.

- In the first one, the binary expansion of x is finite, say, of length μ, so that

$$3^{=}3^{\mu}\left(-\frac{1}{3}2^{x_0} - \frac{1}{9}2^{x_1} - \cdots - \frac{1}{3^{\mu}}2^{x_{\mu-1}}\right) = -3^{\mu-1}2^{x_0} - 3^{\mu-2}2^{x_1} - \cdots - 2^{x_{\mu-1}} \in \mathbb{Z},$$

  implying that y is rational in $\mathbb{Z}_2$, hence $y \in \mathbb{Q} \cap \mathbf{Z}_2$;

- In the second case, x is terminally periodic, as well in the third case (due to Proposition 4.9), so, say, $x_{n+\lambda} = x_n + k$ for some fixed $\lambda \in \mathbb{N}$ with $n \geq \mu$ and $k \in \mathbb{N}$. Then

$$\begin{aligned}
-3^{\mu}(3^{\lambda} - 2^{k})y &= (3^{\lambda} - 2^{k})\left(-3^{\mu}\left(-\sum_{i=0}^{\infty}\frac{1}{3^{i+1}}2^{x_i}\right)\right) \\
&= (3^{\lambda} - 2^{k})\left(\sum_{i=0}^{\infty}3^{\mu-i-1}2^{x_i}\right) \\
&= (3^{\lambda} - 2^{k})\left(3^{\mu-1}2^{x_0} + \cdots + 2^{x_{\mu-1}} + \sum_{i=\mu}^{\infty}3^{\mu-i-1}2^{x_i}\right) \\
&= \underbrace{(3^{\lambda} - 2^{k})(3^{\mu-1}2^{x_0} + \cdots + 2^{x_{\mu-1}})}_{(*)} + (3^{\lambda} - 2^{k})\left(\sum_{i=\mu}^{\infty}3^{\mu-i-1}2^{x_i}\right) \\
&= (*) + \sum_{i=\mu}^{\infty}3^{\lambda+\mu-1-i}2^{x_i} - \sum_{i=\mu}^{\infty}3^{\mu-i-1}2^{x_i+k} \\
&= (*) + 3^{\lambda-1}2^{\mu} + \cdots + 2^{\mu+\lambda-1} + \sum_{i=\mu+\lambda}^{\infty}3^{\lambda+\mu-1-i}2^{x_i} - \sum_{i=\mu}^{\infty}3^{\mu-i-1}2^{x_i+k} \\
&= (*) + 3^{\lambda-1}2^{\mu} + \cdots + 2^{\mu+\lambda-1} + \sum_{i=\mu}^{\infty}3^{\mu-1-i}2^{x_{i+\lambda}} - \sum_{i=\mu}^{\infty}3^{\mu-i-1}2^{x_i+k} \\
&= (*) + 3^{\lambda-1}2^{\mu} + \cdots + 2^{\mu+\lambda-1} + \sum_{i=\mu}^{\infty}3^{\mu-1-i}2^{x_i+k} - \sum_{i=\mu}^{\infty}3^{\mu-i-1}2^{x_i+k} \\
&= (3^{\lambda} - 2^{k})(3^{\mu-1}2^{x_0} + \cdots + 2^{x_{\mu-1}}) + 3^{\lambda-1}2^{\mu} + \cdots + 2^{\mu+\lambda-1} \in \mathbb{Z},
\end{aligned}$$

  so that $y = \frac{(3^{\lambda}-2^{k})(3^{\mu-1}2^{x_0}+\cdots+2^{x_{\mu-1}})+3^{\lambda-1}2^{\mu}+\cdots+2^{\mu+\lambda-1}}{-3^{\mu}(3^{\lambda}-2^{k})}$ is a rational in $\mathbf{Z}_2$, hence, $y \in \mathbb{Q} \cap \mathbf{Z}_2$.

Therefore in either case $\Phi(x)$ is a rational in $\mathbf{Z}_2$, which completes the proof. □

The next and last property highlight the complexity of the Φ function.

Corollary 4.16. *Φ is nowhere differentiable.*

Proof. Let $x = 2^{x_0} + 2^{x_1} + \cdots \in \mathbf{Z}_2$. According to definition 3.49, one needs to show that the limit

$$\lim_{h \to 0} \frac{\Phi(x+h) - \Phi(x)}{h} \tag{4.4}$$

does not exist.

Again, there are two cases to consider: if the binary expansion of x is finite or infinite. We start with the latter.

Recall that $\lim_{k\to\infty} |-2^{x_k}|_2 = 0$, so that -2^{x_k} can be our h of the definition. If $\mathbb{N} \ni k \geq 0$, then,

$$\begin{aligned}
\frac{\Phi(x - 2^{x_k}) - \Phi(x)}{-2^{x_k}} &= \frac{-\sum_{k\neq i=0}^{\infty} \frac{1}{3^{i+1}} 2^{x_i} - \left(-\sum_{i=0}^{\infty} \frac{1}{3^{i+1}} 2^{x_i}\right)}{-2^{x_k}} \\
&= \frac{\frac{1}{3^{k+1}} 2^{x_k}}{-2^{x_k}} \\
&= -\frac{1}{3^{k+1}}.
\end{aligned}$$

Now, we claim that $-\frac{1}{3^{k+1}} \equiv (-1)^k \pmod 4$. In order to show this, we will proceed by induction on k.

- For $k = 0$, $-\frac{1}{3^{k+1}} = -\frac{1}{3} = 101010\ldots \equiv 1 = (-1)^0 \pmod 4$.
- Now, assume $k \in \mathbb{N}$ and that we already showed, as the induction hypothesis, that $-\frac{1}{3^k} \equiv (-1)^{k-1} \pmod 4$. Then, it follows that

  $$-\frac{1}{3^{k+1}} = -\left(-\frac{1}{3}\right)\left(-\frac{1}{3^k}\right) \equiv (-1)(-1)^{k-1} = (-1)^k \pmod 4,$$

  which completes the induction step. Hence, we showed, by induction, that $-\frac{1}{3^{k+1}} \equiv (-1)^k \pmod 4$, from where it follows that the limit 4.4 does not exist in this case.

On the other hand, if $x = 2^{x_0} + \cdots + 2^{x_j} \in \mathbb{N}_0$, then for $\mathbb{N} \ni \ell \geq x_j$, $k \in \mathbb{N}$,

$$\begin{aligned}
\frac{\Phi(x + 2^\ell + 2^{\ell+k}) - \Phi(x)}{2^\ell + 2^{\ell+k}} &= \frac{-\left(\frac{1}{3^{\ell+1}}2^\ell + \frac{1}{3^{\ell+k+1}}2^{\ell+k}\right) - \sum_{i=0}^{j}\frac{1}{3^{i+1}}2^{x_i} - \left(-\sum_{i=0}^{j}\frac{1}{3^{i+1}}2^{x_i}\right)}{2^\ell + 2^{\ell+k}} \\
&= \frac{-\left(\frac{1}{3^{\ell+1}}2^\ell + \frac{1}{3^{\ell+k+1}}2^{\ell+k}\right)}{2^\ell + 2^{\ell+k}} \\
&= -\frac{\frac{2^\ell}{3^{\ell+1}}\left(1 + \frac{2^k}{3^k}\right)}{2^\ell(1 + 2^k)} \\
&= -\frac{1}{3^{\ell+1}}\frac{\frac{3^k+2^k}{3^k}}{1 + 2^k} \\
&= -\frac{1}{3^{\ell+k+1}}\frac{3^k + 2^k}{1 + 2^k}.
\end{aligned}$$

Recalling that $\frac{1}{3} = 110101\ldots \equiv 5 \pmod 8$, it follows that, for $k = 1$,

$$-\frac{1}{3^{\ell+k+1}}\frac{3^k + 2^k}{1 + 2^k} = -\frac{1}{3^{\ell+2}}\frac{3+2}{1+2} = -\frac{1}{3^{\ell+2}}\frac{5}{3} = -\frac{5}{3^{\ell+3}} \equiv -\frac{-3}{3^{\ell+1}} = \frac{1}{3^\ell} = \left(\frac{1}{3}^\ell\right) \equiv 5^\ell \pmod 8.$$

But for $k = 2$,

$$-\frac{1}{3^{\ell+k+1}}\frac{3^k + 2^k}{1 + 2^k} = -\frac{1}{3^{\ell+3}}\frac{9+4}{1+4} = -\frac{1}{3^{\ell+3}}\frac{13}{5} \equiv -\frac{1}{3^{\ell+1}} = -(\frac{1}{3})^{\ell+1} \equiv -5^{\ell+1} \pmod 8.$$

Notice that $5^\ell \not\equiv -5^{\ell+1} \pmod 8$ because $5^\ell - (-5^{\ell+1}) = 5^\ell + 5^{\ell+1} = 5^\ell(1+5) = 5^\ell 6$ which is not divisible by 8.

Thus, since $\lim_{\ell\to\infty}|2^\ell + 2^{\ell+1}| = \lim_{\ell\to\infty}|2^\ell + 2^{\ell+2}| = 0$, it follows that the limit 4.4 does not exist also in this case.

Hence, both cases are covered and therefore Φ is nowhere differentiable, as we wanted to show. □

The complexity of the Φ function is, therefore, presented, explaining why the equivalent form of the Collatz conjecture that was stated in this thesis does not seem likely to be proven any time soon.

Bibliography

[1] Daniel J. Bernstein (1994). *A Non-Iterative 2-adic Statement of the 3x + 1 Conjecture*, Proc. Amer. Math. Soc., **121** (1994), 405–408. [Cited on pages vii, ix, and 1.]

[2] Ethan Akin. *Why is the 3x + 1 Problem Hard?*, Chapel Hill Ergodic Theory Workshops (I.Assani, Ed.), Contemp. Math. vol 356, Amer. Math. Soc. 2004, pp. 1–20. [Cited on pages vii, ix, and 1.]

[3] https://www.quantamagazine.org/why-mathematicians-still-cant-solve-the-collatz-conjecture-20200922/ [Cited on page 5.]

[4] Jeffrey C. Lagarias. *The 3x + 1 Problem: An Annotated Bibliography (1963–1999) (Sorted by Author)*. Department of Mathematics, University of Michigan, Ann Arbor, MI 48109–1109, lagarias@umich.edu, (January 1, 2011 version). [Cited on pages 5 and 9.]

[5] Jeffrey C. Lagarias. *THE 3x + 1 PROBLEM AND ITS GENERALIZATIONS*. *AT & T Bell Laboratories, Murray Hill, NJ 07974* [Cited on pages 5, 6, 7, 8, 9, 10, and 11.]

[6] https://en.wikipedia.org/wiki/Lothar Collatz (consulted on 29th September, 2022) [Cited on page 5.]

[7] https://mathshistory.st-andrews.ac.uk/Biographies/Collatz/ [Cited on page 7.]

[8] Lothar. Collatz, Letter to Michael E. Mays, dated 17 Sept. 1980.

[9] B. Thwaites. *My conjecture*. Bull. Inst. Math. Appl. 21 (1985), 35–41. [Cited on pages 8 and 10.]

[10] https://www.quantamagazine.org/mathematician-proves-huge-result-on-dangerous-problem-20191211/ [Cited on page 15.]

[11] Fernando Q. Gouvêa *p-adic Numbers, An Introduction* Second Edition, 1997, Corrected 3rd printing, 2003, Springer. [Cited on pages vii, ix, and 17.]

[12] https://www.youtube.com/watch?v=t1I9uHF9X5Yt=494s. Minute 13. [Cited on pages 10 and 11.]

[13] Lothar Collatz (1986). *On the Motivation and Origin of the (3n+ 1)-Problem, J. of Qufu Normal University, Natural Science Edition [Qufu shi fan da xue xue bao. Zi ran ke xue ban]*. 12 (1986) No. 3, 9–11 (Chinese, transcribed by Zhi-Ping Ren). [Cited on page 8.]

[14] https://en.wikipedia.org/wiki/Collatz conjecture (consulted on 29th September, 2022) [Cited on page 12.]

[15] Helmut Hasse. *Unsolved problems in elementary number theory*. Lectures at U.Maine (Orono), Spring 1975, dittoed notes, p. 23 - 33. [Cited on page 10.]

[16] R. K. Guy. *private communication*. 2009. [Cited on page 8.]

[17] http://pcbarina.fit.vutbr.cz/ [Cited on page 11.]

[18] https://ieeexplore.ieee.org/document/8560077 [Cited on page 11.]

[19] Haselgrove, C. B. (1958). *A disproof of a conjecture of Pólya*. Mathematika. 5 (02): 141–145. [Cited on page 11.]

[20] Eliahou, Shalom (1993). *The 3x + 1 problem: new lower bounds on nontrivial cycle lengths*. Discrete Mathematics. 118 (1): 45–56. [Cited on page 12.]

[21] http://www.people.vcu.edu/ elhaij/GEB/Notes/Unit4/wondrous-numbers.pdf [Cited on page 10.]

[22] https://mathworld.wolfram.com/HailstoneNumber.html [Cited on page 10.]

[23] https://en.wikipedia.org/wiki/Mertens conjecture (consulted on 29th September, 2022) [Cited on page 11.]

[24] Krasikov, Ilia; Lagarias, Jeffrey C. (2003). *Bounds for the 3x + 1 problem using difference inequalities*. Acta Arithmetica. 109 (3): 237–258. arXiv:math/0205002. Bibcode:2003AcAri.109..237K. doi:10.4064/aa109-3-4. MR 1980260. S2CID 18467460. [Cited on page 14.]

[25] http://www.cecm.sfu.ca/organics/papers/lagarias/paper/goodies/letter1.gif [Cited on page 8.]

[26] http://www.cecm.sfu.ca/organics/papers/lagarias/paper/goodies/letter2.gif [Cited on page 8.]

[27] http://www.cecm.sfu.ca/organics/papers/lagarias/paper/goodies/ubersetzung/html/ubersetzung.html [Cited on page 8.]