The p-adic numbers

Holly Green[*]

24th September 2018

Abstract

This paper is the final product from my place on Imperial College's UROP programme. It provides an introduction to the p-adic numbers and their applications, based loosely on Gouvêa's *"p-adic numbers"*, [1]. We begin by discussing how they arise in mathematics, in particular emphasizing that they are both analytic and algebraic in nature. Proceeding, we explore some of the peculiar results accompanying these numbers before stating and proving Hensel's lemma, which concerns the solvability of p-adic polynomials. Specifically, we draw upon how Hensel's lemma allows us to determine the existence of rational solutions to a homogenous polynomial of degree 2 in 3 variables. We conclude with a thorough analysis of the p-adic aspects of the Bernoulli numbers.

Contents

[*] Department of Mathematics, Imperial College London, London, SW7 6AZ, United Kingdom. *E-mail address:* holly.green15@imperial.ac.uk

1 Introduction

The p-adic numbers are most simply a field extension of $\mathbb{Q}$, the rational numbers, which can be formulated in two ways, using either analytic or algebraic methods. The material presented in this chapter will be largely based upon that of Gouvêa [1] and Baker [2].

1.1 An analytic approach

We begin by recalling the notion of a metric space:

Definition 1.1. A metric space is a pair (X, d), where X is a set and

$$d\ : X \times X \longrightarrow \mathbb{R}_{\geq 0}$$

is a metric, i.e. it is such that the following hold:

(Da) $d(x, y) = 0 \Leftrightarrow x = y$, *the identity of indiscernibles,*

(Db) $\forall x, y \in R$, $d(x, y) = d(y, x)$, *symmetry,*

(Dc) $\forall x, y, z \in R$, $d(x, z) \leq d(x, y) + d(y, z)$, *the triangle inequality.*

Definition 1.2. A complete metric space (X, d) is one in which every Cauchy sequence, $(a_n)_{n\geq 0}$ for $a_n \in X$, converges to some limit, $a \in X$, with respect to the metric d.

Remark. *In a complete metric space, a sequence is convergent if and only if it is a Cauchy sequence.*

It is well known that $\mathbb{Q}$ equipped with the metric induced by the Euclidean norm is not a complete metric space. For example, the sequence

$$3,\ 3.1,\ 3.14,\ 3.141,\ 3.1415,\ 3.14159,\ ...$$

is Cauchy and clearly converges to π, which is not in $\mathbb{Q}$. This idea motivates the definition of a completion - a larger metric space which is complete. More precisely,

Definition 1.3. Given a metric space (X, d), the completion is $(\hat{X}, d)$, defined to be the quotient space

$$\hat{X} = \frac{CS(X)}{\sim},$$

where $CS(X)$, the sequence space of all Cauchy sequences in X, is quotiented by the equivalence relation:

$$(a_n) \sim (b_n) \Longleftrightarrow \lim_{n\to\infty} d(a_n, b_n) = 0. \tag{1}$$

Remark. *A metric space is a dense subset of its completion.*

Now recall the following,

Definition 1.4. A norm $||\cdot||$ on a ring R is defined to be a function

$$||\cdot|| : R \longrightarrow \mathbb{R}_{\geq 0}$$

such that the following hold:

(Na) $||x|| = 0 \Leftrightarrow x = 0$, *positive definiteness,*

(Nb) $\forall x, y \in R$, $||xy|| = ||x||||y||$, *multiplicativity,*

(Nc) $\forall x, y \in R$, $||x+y|| \leq ||x|| + ||y||$, *the triangle inequality.*

It can be shown that given a norm $||\cdot||$,

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

satisfies the criteria of a metric. We call this *'the metric induced by* $||\cdot||$*'*. We will be considering this case exclusively.

It is easy to spot here that

$$\forall x, y, z \in X, \quad d(x-z, y-z) = d(x,y),$$

which we refer to *'translation invariance, with respect to addition'* of the metric. Using this, we may now rephrase (1) as $(a_n - b_n) \to 0$ as $n \to \infty$, or $(a_n - b_n) \in Null(X)$, the sequence space of all sequences in X converging to 0. Hence, the completion $(\hat{X}, d)$, can be equivalently defined as

$$\hat{X} = \frac{CS(X)}{Null(X)}.$$

This alternative formulation allows us to observe that $\hat{X}$ is in fact a ring, something that we will show explicitly later on.

Using real analysis, the completion of $\mathbb{Q}$ with respect to the Euclidean norm is shown to be $\mathbb{R}$, the real numbers. This poses the question: what happens if we play around with different norms?

We aim to construct an alternative norm on $\mathbb{Q}$, which will arise from the following:

Definition 1.5. For $x \in \mathbb{Z}$, let

$$\mathrm{ord}_p(x) = \max\{r : p^r | x\} \geq 0,$$

and for $x = \frac{a}{b} \in \mathbb{Q}$, let

$$\mathrm{ord}_p\left(\frac{a}{b}\right) = \mathrm{ord}_p(a) - \mathrm{ord}_p(b).$$

The function $\mathrm{ord}_p : \mathbb{Q} \longrightarrow \mathbb{Z} \cup \{\infty\}$ is called the p-adic valuation.

Proposition 1.6. *For all* x, $y \in \mathbb{Q}$, *we have:*

(i) $\operatorname{ord}_p(x) = \infty \Leftrightarrow x = 0$,

(ii) $\operatorname{ord}_p(xy) = \operatorname{ord}_p(x) + \operatorname{ord}_p(y)$,

(iii) $\operatorname{ord}_p(x+y) \geq \min\{\operatorname{ord}_p(x), \operatorname{ord}_p(y)\}$, *with equality* $\Leftrightarrow \operatorname{ord}_p(x) \neq \operatorname{ord}_p(y)$.

Proof. (i) This is obvious,

$$\max\{r : p^r|x\} = \infty \Leftrightarrow p^\infty|x \Leftrightarrow x = 0.$$

Now, let

$$x = p^n \frac{a}{b} \text{ and } y = p^m \frac{c}{d}$$

for a, b, c, $d, \in \mathbb{Z}$, such that $p \nmid a$, b, c, d, i.e. $\operatorname{ord}_p(x) = n$ and $\operatorname{ord}_p(y) = m$.

(ii) Upon multiplying, we have

$$xy = p^{n+m} \frac{ac}{bd}$$

where $p \nmid ac$, bd, so $\operatorname{ord}_p(xy) = n + m$.

(iii) Without loss of generality, assume that $n \leq m$, so that

$$x + y = p^n \left(\frac{a}{b} + p^{m-n}\frac{c}{d}\right).$$

If $m = n$, then clearly $\operatorname{ord}_p(x+y) \geq n$. If $m \neq n$, then

$$x + y = p^n \left(\frac{ad + p^{m-n}bc}{bd}\right).$$

As $p \nmid ad$ and $m - n > 0$, we have that $\operatorname{ord}_p(x+y) = n$. □

These properties suggest a good candidate for our new norm.

Definition 1.7. The p-adic norm on $\mathbb{Q}$ is

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

Definition 1.8. A norm that satisfies the stronger property:

(Nc') $\forall x, y \in R$, $||x+y|| \leq \max\{||x||, ||y||\}$, with equality $\Leftrightarrow ||x|| \neq ||y||$,

is called non-Archimedean.

Corollary 1.9. $|\cdot|_p$ *does in fact satisfy the criteria for being a norm on* $\mathbb{Q}$. *Moreover, it is non-Archimedean.*

Proof. We omit the details, remarking that there are clear relationships between the properties of the p-adic valuation and (Na), (Nb), (Nc). □

Example 1.10. In a non-Archimedean metric space, every triangle is isosceles.

Let x, y, z denote the corners of a triangle which is not equilateral. Without loss of generality, we may assume that $d(x,y) \neq d(y,z)$, i.e. $||x-y|| \neq ||y-z||$. Then,

$$d(x,z) = ||x-z|| = \max\{||x-y||,\, ||y-z||\} = \max\{d(x,y),\, d(y,z)\}.$$

Thus two sides of the triangle must be of the same length. □

The p-adic norm is much more natural than you may think.

Definition 1.11. We say that two norms $||\cdot||_1$ and $||\cdot||_2$ are equivalent if there exists some $\lambda > 0$ such that

$$||\cdot||_1 = ||\cdot||_2^\lambda.$$

Theorem 1.12 (Ostrowski's theorem)**.** *Every non-trivial norm on $\mathbb{Q}$ is equivalent to either $|\cdot|_p$ for some prime p or $|\cdot|_\infty = |\cdot|$, the usual Euclidean absolute value.*

Proof. See [1, Theorem 3.1.3]. □

When working in the Euclidean norm, we have concrete notions of size and distance. The analogue of this in the p-adic norm may not appear so apparent at first.

Example 1.13. Considering only powers of p, we have that large negative powers are regarded as being "large" and large positive powers are "small". For example,

$$|1024|_2 = \frac{1}{1024},$$

thus we would say that 1024 is 2-adically small, despite being fairly large in the Euclidean sense. Conversely,

$$\frac{5}{294} \approx 0.017$$

is fairly small in the Euclidean sense, but

$$\left|\frac{5}{294}\right|_5 = \frac{1}{5} \text{ and } \left|\frac{5}{294}\right|_7 = 49.$$ □

Example 1.14. We also highlight that as p varies, so does the distance between any two fixed numbers. Considering the Euclidean norm, we have that

$$\left|\frac{1}{5} - \frac{5}{27}\right| = \frac{2}{135},$$

so we'd say that these numbers are very close. Now looking at the p-adic distances:

$$\left|\frac{1}{5} - \frac{5}{27}\right|_2 = \frac{1}{2}, \quad \left|\frac{1}{5} - \frac{5}{27}\right|_3 = 27, \quad \left|\frac{1}{5} - \frac{5}{27}\right|_5 = 5,$$

so 2-adically we may again say that the numbers are fairly close together, whereas 3-adically they are not.

Additionally, with respect to larger primes we have:

$$\left|\frac{1}{5} - \frac{5}{27}\right|_p = \left|\frac{2}{135}\right|_p = 1, \text{ for } p > 5.$$ □

In fact for any fixed pair $x, y \in \mathbb{Q}$, there always exists some p' such that for all $p > p'$, $|x-y|_p = 1$.

Other interesting peculiarities also arise working in $|\cdot|_p$, such as the convergence of some sequences.

Example 1.15. First, consider the sequence given by $a_n = p^n$, for $n \in \mathbb{N}$. I claim that,

$$\lim_{n\to\infty} p^n = \begin{cases} 0 & \text{with respect to } |\cdot|_p, \\ \text{does not exist} & \text{with respect to } |\cdot|_q \text{ for } q \neq p, \\ \infty & \text{with respect to } |\cdot|_\infty. \end{cases}$$

We'll discuss the cases in the order listed above. The first case is trivial upon noticing that,

$$|p^n|_p = \frac{1}{p^n} \to 0 \text{ as } n \to \infty.$$

In the second case, considering consecutive terms,

$$|p^{n+1} - p^n|_q = |p^n(p-1)|_q = |p-1|_q,$$

a non-zero constant. Hence, the sequence cannot be Cauchy and therefore it cannot be convergent. Finally, the third case follows from a routine real analysis argument. Fix $R > 0$. Notice that for some $q \geq 1$ we have that $p = 1 + q$ and so, $p^n = (1+q)^n > 1 + nq$. Thus,

$$\forall n > \frac{R}{q},\ p^n > 1 + nq > R. \qquad \square$$

Before moving onto the next example, it is necessary to introduce some notation concerning the base p expansion of integers. It is well known, by a direct application of the Euclidean algorithm, that for any fixed prime p and $n \in \mathbb{Z}$, there is a unique expression:

$$n = n_0 + n_1 p + n_2 p^2 + \cdots + n_k p^k,$$

for some $k \in \mathbb{N}$ and integers $0 \leq n_0, \ldots, n_k \leq p-1$. We denote,

$$s_p(n) = n_0 + n_1 + \cdots + n_k.$$

Lemma 1.16. *In the same notation as above,*

$$\operatorname{ord}_p(n!) = \frac{n - s_p(n)}{p-1}.$$

Proof. Writing $n!$ explicitly as $n! = 1 \cdot 2 \cdot 3 \cdots n$, we notice that in the product there are

$$\left\lfloor \frac{n}{p} \right\rfloor \text{ multiples of } p,\ \left\lfloor \frac{n}{p^2} \right\rfloor \text{ multiples of } p^2, \text{ and more generally } \left\lfloor \frac{n}{p^i} \right\rfloor \text{ multiples of } p^i.$$

Thus, there are in total

$$\left\lfloor \frac{n}{p} \right\rfloor + \left\lfloor \frac{n}{p^2} \right\rfloor + \cdots + \left\lfloor \frac{n}{p^k} \right\rfloor$$

factors of p in $n!$. Now, using the identity

$$\left\lfloor \frac{n}{p^i} \right\rfloor = n_k p^{k-i} + \cdots + n_{i+1} p + n_i,$$

and some simple algebraic manipulation, the desired expression is obtained. □

Remark. *In practice it is often more useful to use the estimate*

$$\mathrm{ord}_p(n!) < \frac{n}{p-1}, \quad \textit{or} \quad |n!|_p > \frac{1}{p^{\frac{n}{p-1}}}.$$

Example 1.17. Consequently, we have our next example of a convergent sequence with respect to $|\cdot|_p$. Let $a_n = n!$ for $n \geq 0$. By the above result, we have that for any p

$$|n!|_p = p^{\frac{p-1}{n-s_p(n)}} \to 0 \text{ as } n \to \infty,$$

hence $(a_n)_{n\geq 0}$ is a null sequence. □

Naturally, we question whether $\mathbb{Q}$ is complete with respect to $|\cdot|_p$. The answer is no.

Example 1.18. Let p be odd and for some integer $1 < b < p-1$, consider the sequence given by $a_n = b^{p^n}$ for $n \geq 0$. We see that this is Cauchy with respect to $|\cdot|_p$, by estimating

$$|a_{n+1} - a_n|_p = |b^{p^{n+1}} - b^{p^n}|_p = |b^{p^n}|_p |b^{p^n(p-1)} - 1|_p \leq \frac{1}{p^n},$$

having used Fermat's little theorem in the latter step, which says that $b^{\phi(p^n)} \equiv b \pmod{p^n}$, where ϕ denotes Euler's totient function, so clearly:

$$|a_{n+1} - a_n|_p \to 0 \text{ as } n \to \infty.$$

Now, suppose that the sequence converges to some limit $l \in \mathbb{Q}$ (we will obtain a contradiction). First, notice that

$$l = \lim_{n\to\infty} a_n = \lim_{n\to\infty} a_{n+1} = \lim_{n\to\infty} (a_n)^p = \left(\lim_{n\to\infty} a_n\right)^p = l^p,$$

so l is a $(p-1)$th root of unity $\Rightarrow l = \pm 1$. Estimating using the non-Archimedean property, it becomes immediate that

$$|b - \pm 1|_p < 1, \text{ i.e. } p|b-1 \text{ or } p|b,$$

but this cannot be as $1 < b < p-1$. □

Now we question, is the completion with respect to $|\cdot|_p$ still $\mathbb{R}$? Again, the answer is no. In fact, in this way we have constructed a totally new metric space which we call $\mathbb{Q}_p$, the p-adic numbers.

Theorem 1.19. *The space $\mathbb{Q}_p$ forms a field.*

Proof. Referring back to the definitions on page 3, we first claim that $Null(\mathbb{Q}) \triangleleft CS(\mathbb{Q})$, thus $\mathbb{Q}_p$ forms a ring. Taking the operation $+$ to be the natural one, it is clear that $(Null(\mathbb{Q}), +)$ is a subgroup of $(CS(\mathbb{Q}), +)$. For the claim to be true, it is now only necessary to show that if $(a_n) \in Null(\mathbb{Q})$ and $(x_n) \in CS(\mathbb{Q})$, then we have that $(a_n)(x_n) = (a_n x_n) \in Null(\mathbb{Q})$. As the sequence (x_n) is Cauchy, $\exists N > 0$ such that for all $n \geq 0$,

$$|x_n|_p < M = \max\{|x_1|_p, \ldots, |x_{N-1}|_p, |x_N|_p + 1\}.$$

Hence,

$$|a_n x_n|_p = |a_n|_p |x_n|_p < |a_n|_p M \to 0 \text{ as } n \to \infty.$$

Showing further that in fact $Null(\mathbb{Q})$ is a maximal ideal of $CS(\mathbb{Q})$, we obtain that $\mathbb{Q}_p$ is actually a field. For this, first notice that the that ideal is proper. Now, fix some sequence $(a_n) \in CS(\mathbb{Q}) - Null(\mathbb{Q})$, then $\exists N > 0$ such that $\forall n > N$, $a_n \neq 0$. Clearly, defining

$$\lambda_n = \begin{cases} 1 & 0 \leq n \leq N \\ 0 & N < n, \end{cases}$$

we have that the sequence $(\lambda_n) \in Null(\mathbb{Q})$ and considering now the ideal given by $(Null(\mathbb{Q}), (a_n))$, we have that $(b_n) \in (Null(\mathbb{Q}), (a_n))$, where $b_n = a_n - a_n \lambda_n + \lambda_n$. We have too that $(b_n) \notin Null(\mathbb{Q})$ and it has a multiplicative inverse (b_n^{-1}). For n, $m > N$

$$|b_n^{-1} - b_m^{-1}|_p = \frac{|a_m - a_n|_p}{|a_n|_p |a_m|_p} \to 0 \text{ as } n \to \infty,$$

using that (a_n) is both Cauchy and bounded away from 0. So we in fact have that $(b_n^{-1}) \in CS(\mathbb{Q})$ and therefore $(1) \in (Null(\mathbb{Q}), (a_n)) \Rightarrow (Null(\mathbb{Q}), (a_n)) = CS(\mathbb{Q})$. □

Analogously, we can define the space of p-adic integers, $\mathbb{Z}_p$, as the completion of $\mathbb{Z}$ with respect to $|\cdot|_p$.

Theorem 1.20. *$\mathbb{Z}_p$ is in fact the unit disc in $\mathbb{Q}_p$ about 0, i.e.*

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

Proof. $(\subseteq)$ Let $a \in \mathbb{Z}_p$, then

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

with respect to $|\cdot|_p$, for some Cauchy sequence $(a_n)_{n \geq 0}$ in $\mathbb{Z}$. As for all $n \geq 0$, $a_n \in \mathbb{Z}$, we have that $|a_n|_p \leq 1$. Thus, taking the limit as $n \to \infty$, it is true that $|a|_p \leq 1$, i.e. $a \in \{x \in \mathbb{Q}_p : |x|_p \leq 1\}$.

$(\supseteq)$ First let $x \in \mathbb{Q}$ be such that $|x|_p \leq 1$, i.e. $x = \frac{a}{b}$ for a, $b \in \mathbb{Z}$ and without loss of generality, $p \nmid b$. Clearly, we have that

$$x_n = \frac{a + p^n}{b} \to \frac{a}{b}, \text{ as } n \to \infty,$$

so $x \in \mathbb{Z}_p$. Now, if $x \in \mathbb{Q}_p$ is such that $|x|_p \leq 1$, then x is the limit of some Cauchy sequence $(x_n)_{n\geq 0}$ in $\mathbb{Q}$. So $\exists N > 0$ such that $\forall n \geq N$,

$$|x_n - x|_p < 1 \Rightarrow |x_n|_p \leq \max\{|x_n - x|_p,\ |x|_p\} \leq 1$$

and the sequence

$$y_n = \begin{cases} 0 & 0 \leq n < N \\ x_n & N \leq n, \end{cases}$$

converges to x, so $x \in \mathbb{Z}_p$. □

Proposition 1.21. *Every $a \in \mathbb{Z}_p$, is the limit of a sequence of non-negative integers.*

Proof. Recalling an earlier remark, a space is dense in its completion, so given $a \in \mathbb{Z}_p$ and $n \in \mathbb{N}$ there exists some integer $0 \leq a_n \leq p^n - 1$ such that

$$|x - a_n|_p \leq \frac{1}{p^n} \to 0 \text{ as } n \to \infty.$$ □

1.2 An algebraic approach

Although the analytic formulation is generally the easiest to work with, we introduce an algebraic definition which will be particularly useful later on when we consider Hensel's lemma, a result that allows us to find solutions to polynomials modulo powers of primes.

From this algebraic perspective, we first work towards defining the p-adic integers, and will do so as a sequence space denoted $\widetilde{\mathbb{Z}}_p$. Consider the sequence of maps between the following groups:

$$\ldots \xrightarrow{\phi_4} \mathbb{Z}/p\mathbb{Z}^4 \xrightarrow{\phi_3} \mathbb{Z}/p^3\mathbb{Z} \xrightarrow{\phi_2} \mathbb{Z}/p^2\mathbb{Z} \xrightarrow{\phi_1} \mathbb{Z}/p\mathbb{Z},$$

where

$$\phi_i(a) = a \ (\text{mod } p^i).$$

This allows us to write:

$$\widetilde{\mathbb{Z}}_p = \varprojlim_n \mathbb{Z}/p^n\mathbb{Z},$$

called the *'projective limit'* of the system. More explicitly, it is a sequence space such that $(a_n)_{n\geq 1} \in \widetilde{\mathbb{Z}}_p$ if for all $n \geq 1$,

(i) $a_n \in \mathbb{Z}/p^n\mathbb{Z}$, and

(ii) $a_n \equiv a_n \pmod{p^n}$.

Example 1.22. Let $p = 2$. We can start to define such a sequence $(a_n)_{n\geq 1} \in \widetilde{\mathbb{Z}}_2$ as follows,

$$a_1 \in \{0,\ 1\},\ \text{take } a_1 = 1; \qquad a_2 \in \{1,\ 3\},\ \text{take } a_2 = 3; \qquad a_3 \in \{3,\ 7\},\ \text{take } a_3 = 3$$
$$a_4 \in \{3,\ 11\},\ \text{take } a_4 = 11; \qquad a_5 \in \{11,\ 27\},\ \text{take } a_5 = 27.$$

Thus $1, 3, 3, 11, 27, \ldots \in \widetilde{\mathbb{Z}}_2$, or alternatively we may view this as:

$$(1, 3, 3, 11, 27, \ldots) \in \widetilde{\mathbb{Z}}_2 \subseteq \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/4\mathbb{Z} \times \mathbb{Z}/8\mathbb{Z} \times \mathbb{Z}/16\mathbb{Z} \times \mathbb{Z}/32\mathbb{Z} \times \cdots.$$ □

Equipping $\widetilde{\mathbb{Z}}_p$ with the addition, multiplication, zero and unit defined by:

$$(a_n) + (b_n) = (a_n + b_n), \quad (a_n) \cdot (b_n) = (a_n \cdot b_n), \quad \mathbf{0} = (0, 0, 0, \ldots), \quad \mathbf{1} = (1, 1, 1, \ldots),$$

an easy check shows that we have given the space a ring structure. Noting too that $\widetilde{\mathbb{Z}}_p$ is a domain, and given a domain X there exists $\mathrm{Frac}(X)$, it's fraction field, we obtain a corresponding algebraic definition for $\mathbb{Q}_p$:

Definition 1.23. The p-adic numbers, from this point of view, are henceforth defined to be

$$\widetilde{\mathbb{Q}}_p = \mathrm{Frac}(\widetilde{\mathbb{Z}}_p).$$

To establish a norm on $\widetilde{\mathbb{Q}}_p$, we shall define the p-adic valuation in the context of sequence spaces.

Definition 1.24. For $a = (a_n)_{n \geq 0} \in \widetilde{\mathbb{Z}}_p$, let

$$\widetilde{\mathrm{ord}}_p(a) = \max\{n : a_{n-1} \equiv 0 \pmod{p^n}\} \geq 0,$$

and for $a = \frac{x}{y} \in \widetilde{\mathbb{Q}}_p$, as before let

$$\widetilde{\mathrm{ord}}_p\left(\frac{x}{y}\right) = \widetilde{\mathrm{ord}}_p(x) - \widetilde{\mathrm{ord}}_p(y).$$

From this, we again define the norm $\widetilde{|\cdot|}_p$ to be

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

We will soon see that the two definitions of the valuation, and consequently the corresponding norms, are equivalent.

1.3 The correspondence

We have introduced the p-adic numbers and integers in two different ways to be able to discuss results surrounding them more succinctly. Now we'll show that in fact these definitions are equivalent, but before doing so it's necessary to introduce the p-adic expansion of p-adic numbers - the analogue of writing integers in base p, [2].

1.3.1 The p-adic expansion

First, let $a \in \mathbb{Z}_p$, then by Proposition 1.21, a is the limit of a sequence of non-negative integers. This enables a choice of some integer $0 \leq a_0 \leq p - 1$, such that,

$$|a - a_0|_p < \frac{1}{p} \Rightarrow \frac{a - a_0}{p} \in \mathbb{Z}_p.$$

The same argument allows us to find an integer $0 \leq a_1 \leq p-1$, such that

$$|a - (a_0 + a_1 p)|_p < \frac{1}{p} \Rightarrow \frac{a - (a_0 + a_1 p)}{p} \in \mathbb{Z}_p.$$

Eventually, we obtain that the sequence $(b_n)_{n \geq 1}$ given by,

$$b_n = a_0 + a_1 p + a_2 p^2 + \cdots + a_{n-1} p^{n-1}$$

converges to a, allowing us to write

$$a = a_0 + a_1 p + a_2 p^2 + \cdots ,$$

an infinite sum which we call the p-adic expansion of a p-adic integer. More generally, let $a \in \mathbb{Q}_p$ be such that $|a|_p > 1$, i.e. $|a|_p = p^k$ for some $k \in \mathbb{N}$. Thus, considering $b = p^k a \in \mathbb{Z}_p$, if

$$b = b_0 + b_1 p + b_2 p^2 + \cdots ,$$

dividing by p^k we obtain

$$a = \frac{b_0}{p^k} + \frac{b_1}{p^{k-1}} + \frac{b_2}{p^{k-2}} + \cdots + b_k + b_{k+1} p + \cdots + b_{k+r} p^r + \cdots .$$

To summarise, for $a \in \mathbb{Q}_p$,

- if $|a|_p \leq 1$, then $a = \sum_{k=0}^{\infty} a_k p^k$ for some integers $0 \leq a_k \leq p-1$,
- if $|a|_p > 1$, then for some $r \in \mathbb{N}$, $a = \sum_{k=-r}^{\infty} a_k p^k$ for some integers $0 \leq a_k \leq p-1$.

This motivates the idea that we may think of $\mathbb{Z}_p$ as being analogous to $\mathbb{K}[[x]]$ and $\mathbb{Q}_p$ as being analogous to $\mathbb{K}((x))$, the rings of power series and Laurent series over some field $\mathbb{K}$. More precisely, what we've constructed here is the analogy

$$\mathbb{Z}_p \longleftrightarrow \mathbb{K}[[x]] \qquad \mathbb{Q}_p \longleftrightarrow \mathbb{K}((x)),$$

which we will return to look at more later on.

Theorem 1.25. *Unlike the more familiar decimal expansion, the p-adic expansion is unique.*

Proof. We need only show this for the p-adic integers, the case of the p-adic numbers then follows. For $a \in \mathbb{Z}_p$, suppose that

$$\begin{aligned} a &= a_0 + a_1 p + a_2 p^2 + \cdots \\ &= a_0' + a_1' p + a_2' p^2 + \cdots . \end{aligned}$$

Let $d \in \mathbb{N}$ be the smallest such that $a_d \neq a_d'$. Without loss of generality, suppose that $a_d > a_d'$, then $1 \leq a_d - a_d' \leq p-1$. Write

$$\begin{aligned} b_n &= a_0 + a_1 p + a_2 p^2 + \cdots + a_n p^n \\ b_n' &= a_0' + a_1' p + a_2' p^2 + \cdots + a_n' p^n, \end{aligned}$$

then we have that

$$b_d - b'_d = (a_d - a'_d)p^d \Rightarrow |b_d - b'_d|_p = \frac{1}{p^d}.$$

However, by the non-Archimedean property,

$$|b_d - b'_d|_p \leq \max\{|b_d - a|_p,\, |b'_d - a|_p\} < \frac{1}{p^d}.$$

Having obtained a contradiction, we conclude that the expansion is unique. □

We have formally introduced the p-adic expansion in the analytic context for the purposes of the next result, however an expansion also exists from the algebraic point of view and this will be established soon.

Theorem 1.26. $\mathbb{Z}_p \cong \widetilde{\mathbb{Z}}_p$.

Proof. (Based on [3, Section 1.1.2]). We first consider the map

$$\begin{aligned} \Phi:\ \mathbb{Z}_p &\longrightarrow \widetilde{\mathbb{Z}}_p, \\ a &\longmapsto (\phi_1(a),\, \phi_2(a),\, \ldots), \end{aligned}$$

where for $i \geq 1$ and $a = \sum_{k=0}^{\infty} a_k p^k$ (it's p-adic expansion), we have

$$\begin{aligned} \phi_i:\ \mathbb{Z}_p &\longrightarrow \mathbb{Z}/p^i\mathbb{Z}, \\ a &\longmapsto \sum_{k=0}^{i-1} a_k p^k \pmod{p^i}. \end{aligned}$$

Verifying that Φ defines a homomorphism is equivalent to doing so for each ϕ_i. Clearly, under ϕ_i, 1 always maps to 1 and the additive property also holds. Multiplicativity follows upon observing that

$$\begin{aligned} \sum_{k=0}^{2i}\sum_{l=0}^{k} a_l b_{k-l} p^k &= \sum_{l=0}^{2i}\sum_{k=l}^{2i} a_l b_{k-l} p^k \\ &= \sum_{l=0}^{2i}\sum_{k=0}^{2i-l} a_l b_k p^{l+k} \\ &= \sum_{l=0}^{i} a_l p^l \left(\sum_{k=0}^{i} b_k p^k + \sum_{k=i+1}^{2i-l} b_k p^k\right) + \sum_{l=i+1}^{2i}\sum_{k=0}^{2i-l} a_l b_k p^{l+k} \\ &= \left(\sum_{l=0}^{i} a_l p^l\right)\left(\sum_{k=0}^{i} b_k p^k\right) + \sum_{l=0}^{i-1} a_l p^l \sum_{k=i+1}^{2i-l} b_k p^k + \sum_{k=0}^{i-1} b_k p^k \sum_{l=i+1}^{2i-k} a_l p^l \end{aligned}$$

to establish the identity:

$$\left(\sum_{k=0}^{i} a_k p^k\right)\left(\sum_{k=0}^{i} b_k p^k\right) = \sum_{k=0}^{2i}\sum_{l=0}^{k} a_l b_{k-l} p^k - \sum_{k=0}^{i-1}\left(a_k \sum_{l=i+1}^{2i-k} b_l p^l + b_k \sum_{l=i+1}^{2i-k} a_l p^l\right) p^k,$$

i.e. for $i \geq 0$, $\phi_{i+1}(a)\phi_{i+1}(b) \equiv \phi_{i+1}(ab) \pmod{p^{i+1}}$. Next, choose $(a_1,\, a_2,\, \ldots) \in \widetilde{\mathbb{Z}}_p$, and for $n \geq 1$, let $0 \leq x_n \leq p^n - 1$ be the unique integer such that $x_n \equiv a_n \pmod{p^n}$. In base p, write

$$x_n = \lambda_{n,0} + \lambda_{n,1}p + \cdots + \lambda_{n,n-1}p^{n-1}.$$

For $0 \leq i \leq n-1$, we have that $\lambda_{n+1,i} = \lambda_{n,i} \Rightarrow \lambda_{n,i} \to \lambda_i$ for some λ_i, as $n \to \infty$. Using this idea, let

$$\begin{aligned} \Psi :\ \widetilde{\mathbb{Z}}_p &\longrightarrow \mathbb{Z}_p, \\ (a_1,\, a_2,\, \ldots) &\longmapsto \sum_{k=0}^{\infty} \lambda_k p^k. \end{aligned}$$

We'll show that this defines yet another ring homomorphism. Again, it is clear that 1 maps to 1. In addition to the notation used in constructing Ψ, let $(b_1,\, b_2,\, \ldots) \in \widetilde{\mathbb{Z}}_p$, $0 \leq y_n \leq p^n - 1$ be the unique integer such that $y_n \equiv b_n \pmod{p^n}$ and

$$y_n = \mu_{n,0} + \mu_{n,1}p + \cdots + \mu_{n,n-1}p^{n-1}$$

be the base p expansion. Considering first $(a_1 + b_1,\, a_2 + b_2,\, \ldots) \in \widetilde{\mathbb{Z}}_p$, we have that

$$x_n + y_n \equiv a_n + b_n \pmod{p^n}$$

and

$$x_n + y_n = (\lambda_{n,0} + \mu_{n,0}) + (\lambda_{n,1} + \mu_{n,1})p + \cdots + (\lambda_{n,n-1} + \mu_{n,n-1})p^{n-1}.$$

As $\lambda_{n,i} + \mu_{n,i} \to \lambda_i + \mu_i$ as $n \to \infty$, the additive result follows. Now consider $(a_1b_1,\, a_2b_2,\, \ldots) \in \widetilde{\mathbb{Z}}_p$, we have here that

$$x_n y_n \equiv a_n b_n \pmod{p^n}$$

so

$$x_n y_n = \theta_{n,0} + \theta_{n,1}p + \cdots + \theta_{n,n-1}p^{n-1}, \text{ for } \theta_{n,i} = \sum_{j=0}^{n-1} \lambda_{n,j}\mu_{n,i-j}.$$

Using this, clearly

$$\theta_{n,i} \to \sum_{j=0}^{n-1} \lambda_j \mu_{i-j} \text{ as } n \to \infty,$$

giving the desired expression for multiplicativity. Moreover, both $\Phi \circ \Psi$ and $\Psi \circ \Phi$ provide identity maps on $\widetilde{\mathbb{Z}}_p$ and $\mathbb{Z}_p$ respectively. □

Remark. *The function Ψ provides the p-adic expansion from the algebraic perspective.*

Having established this isomorphism, abusing notation slightly, we use it to show that

$$\operatorname{ord}_p(a) = \widetilde{\operatorname{ord}}_p(a), \text{ for all } a.$$

Proof. Let $a = \sum_{k=0}^{\infty} a_k p^k \in \mathbb{Z}_p$, for some integers $0 \leq a_k \leq p-1$, and let $\text{ord}_p(a) = N$. Then,

$$\frac{a}{p^N} = \sum_{k=0}^{\infty} a_k p^{k-N} = \frac{a_0}{p^N} + \frac{a_1}{p^{N-1}} + \cdots + a_N + a_{N+1}p + \cdots \in \mathbb{Z}_p.$$

Clearly $a_N + a_{N+1}p + \cdots$ is a p-adic integer and so we obtain that

$$\frac{a_0}{p^N} + \frac{a_1}{p^{N-1}} + \cdots \in \mathbb{Z}_p \Rightarrow a_0 + a_1 p + \cdots a_{N-1}p^{N-1} \equiv 0 \pmod{p^N}.$$

We now have that $\widetilde{\text{ord}}_p(\Phi(a)) \geq N$. Suppose that for some $k \geq 1$,

$$a_0 + a_1 p + a_2 p^2 + \cdots + a_{N+k-1}p^{N+k-1} \equiv 0 \pmod{p^{N+k}},$$

then $p^{N+k}|a$. This is clearly a contradiction, so $\widetilde{\text{ord}}_p(\Phi(a)) = N$. □

From now on we refer to both $\mathbb{Z}_p$ and $\widetilde{\mathbb{Z}}_p$ by the former. Similarly for $\mathbb{Q}_p$ and $\widetilde{\mathbb{Q}}_p$.

Another application of the p-adic expansion allows us to establish the interesting nice result:

Corollary 1.27. *$\mathbb{Z}_p$ and $\mathbb{Q}_p$ have the cardinality of $\mathbb{R}$, the cardinality of 'continuum'.*

Proof. A consequence to the uniqueness of the p-adic expansion is that the map

$$\begin{aligned} f:\ \mathbb{Z}_p &\longrightarrow (\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}, \\ \sum_{k=0}^{\infty} a_k p^k &\longmapsto (a_0,\, a_1,\, a_2, \ldots), \end{aligned}$$

defines a bijection. Thus by [4, Page 2],

$$|\mathbb{Z}_p| = p^{|\mathbb{N}|} = 2^{|\mathbb{N}|},$$

which is the cardinality of $\mathbb{R}$. Now, as $\mathbb{Q}_p = \text{Frac}(\mathbb{Z}_p)$, it is a subset of $\mathbb{Z}_p \times \mathbb{Z}_p$, which also has cardinality of continuum. Injecting $\mathbb{Q}_p$ into $\mathbb{Z}_p \times \mathbb{Z}_p$, we see it has both at most and at least cardinality of continuum. □

2 Discrete valuation rings and fields

A generalisation of the p-adic norm is a valuation, ν. From here on we use $\mathbb{K}$ to denote an arbitrary field.

Definition 2.1. A valuation on $\mathbb{K}$ is a map

$$\nu : \mathbb{K} \longrightarrow \mathbb{R} \cup \{\infty\}$$

such that the following hold:

(Va) $\forall x, y \in \mathbb{K}$, $\nu(xy) = \nu(x) + \nu(y)$,

(Vb) $\forall x, y \in \mathbb{K}$, $\nu(x + y) \geq \min\{\nu(x),\, \nu(y)\}$,

(Vc) $\nu(x) = \infty \Leftrightarrow x = 0$.

We say that a valuation is discrete if the image of $\mathbb{K} - \{0\}$ under ν is isomorphic to $r\mathbb{Z}$ for some $r > 0$.

Definition 2.2. A field equipped with a non-trivial (discrete) valuation is called a (discrete) valuation field, a (D)VF.

Example 2.3. The p-adic valuation, unsurprisingly, provides a valuation on both $\mathbb{Q}$ and $\mathbb{Q}_p$, making them discrete valuation fields. □

Example 2.4. For an arbitrary element $f(x) \in \mathbb{K}[x]$, there exists some $f_0(x) \in \mathbb{K}[x]$, with a non-zero constant term, such that $f(x) = x^n f_0(x)$ where $n \in \mathbb{N}$. We define $\nu(f) = n$, and extend to the field $\mathbb{K}(x)$ by letting

$$\nu\left(\frac{f}{g}\right) = \nu(f) - \nu(g).$$

Checking that (Va), (Vb) and (Vc) hold is a straight forward task, thus ν is a discrete valuation and $\mathbb{K}(x)$ is a discrete valuation field. □

Example 2.5. For each non-zero Laurent series in $\mathbb{K}((x))$, there is an expression of the form

$$f(x) = \sum_{n \geq n_0} a_n x^n,$$

for some $a_{n_0} \neq 0$. Define $\nu(f) = n_0$. This time we will verify that this is a discrete valuation, thus $\mathbb{K}((x))$ equipped with it is a discrete valuation field.

First, notice that the image of this map is clearly $\mathbb{Z}$, so it is discrete. Now, let

$$f(x) = \sum_{n \geq n_0} a_n x^n, \quad g(x) = \sum_{n \geq m_0} b_n x^n,$$

where a_{n_0}, $b_{m_0} \neq 0$ and without loss of generality, assume that $n_0 \leq m_0$. We have that

$$f(x)g(x) = \sum_{n \geq n_0 + m_0} \left(\sum_{i+j=n} a_i b_j \right) x^n,$$

where the coefficient of $x^{n_0+m_0}$ is non-zero, verifying (Va). In the same notation,

$$f(x) + g(x) = \sum_{m_0 > n \geq n_0} a_n x^n + \sum_{n \geq m_0} (a_n + b_n) x^n,$$

hence $\nu(f+g) \geq n_0 = \min\{\nu(f),\, \nu(g)\}$. Finally, it is clear that $\nu(f) = \infty \Leftrightarrow f = 0$, so (Vc) follows. □

Definition 2.6. Given a (discrete) valuation ν on a (discrete) valuation field $\mathbb{K}$,

$$R = \{x \in \mathbb{K} : \nu(x) \geq 0\}$$

is the (discrete) valuation ring, the (D)VR, of $\mathbb{K}$.

Clearly, we have that $\nu(1) = 0$ thus both $0, 1 \in R$. Now let $x, y \in R$, then

$$\nu(x+y) \geq \min\{\nu(x),\, \nu(y)\} \geq 0 \text{ and } \nu(xy) = \nu(x) + \nu(y) \geq 0,$$

so $x + y$, $xy \in R$. Hence, R does in fact define a subring of $\mathbb{K}$. It can be further shown that $\mathbb{K}$ is the fraction field of R.

Example 2.7. $\mathbb{Z}_p$ is a discrete valuation ring, that of the valuation in Example 2.3 defined on $\mathbb{Q}_p$. □

Example 2.8. $\mathbb{K}[[x]]$ is a discrete valuation ring, that of the valuation in Example 2.5 defined on $\mathbb{K}((x))$. □

But what about the discrete valuation rings given rise to by the discrete valuation fields $\mathbb{Q}$ and $\mathbb{K}(x)$?

2.1 Localizations

Definition 2.9. Given R, a commutative ring with unit 1_R, and a subset $S \subseteq R$ which is closed under multiplication and such that $0 \notin S$, $1_R \in S$ we define the localization of R at S to be

$$R_S = \frac{\{\frac{a}{b} : a \in R,\, b \in S\}}{\sim},$$

where the equivalence relation is given by:

$$\frac{a}{b} \sim \frac{a'}{b'} \iff \exists u \in S \text{ such that } u(ab' - a'b) = 0.$$

In the context of the p-adics, we'll be considering a specific type of localization.

Proposition 2.10. *Let I be a prime ideal of R. I claim that $R - I \subset R$ provides such a candidate.*

Proof. To start, by the definition of being an ideal, $I \subset R$ is an additive subgroup, thus $0 \in I \Rightarrow 0 \notin R - I$. Additionally, as the ideal is proper, $1_R \notin I \Rightarrow 1_R \in R - I$. Finally, if $a, b \in R - I$ then, by definition of I being prime, $ab \in I \Rightarrow a \in I$ or $b \in I$, a contradiction. Thus $ab \in R - I$ and we have that $R - I$ is closed under multiplication. □

Contrary to the usual notation, we write R_I instead of R_{R-I} and refer to this as the *'localization of R at the prime ideal I'*.

Example 2.11. A simple example is given by $R = \mathbb{Z}$ and $I = (2)$, the principal ideal generated by 2. Explicitly,

$$(2) = \{2x : x \in \mathbb{Z}\} = \{\text{even integers}\},$$

which is clearly prime. Thus,

$$\mathbb{Z}_{(2)} = \left\{\frac{a}{b} \in \mathbb{Q} : b \text{ is odd}\right\} = \left\{\frac{a}{b} \in \mathbb{Q} : 2 \nmid b\right\}. \qquad \square$$

Example 2.12. More generally, as $(p) \triangleleft \mathbb{Z}$ is prime for any prime number p, we can write

$$\mathbb{Z}_{(p)} = \left\{\frac{a}{b} \in \mathbb{Q} : p \nmid b\right\}. \qquad \square$$

It turns out that the discrete valuation rings associated to $\mathbb{Q}$ as in Example 2.3 and $\mathbb{K}(x)$ as in Example 2.4, are in fact the localizations $\mathbb{Z}_{(p)}$ and $\mathbb{K}[x]_{(x)}$ respectively.

This discussion allows us to extend the analogy we noted earlier to the following:

$$\mathbb{Z} \longleftrightarrow \mathbb{K}[x] \qquad \mathbb{Z}_{(p)} \longleftrightarrow \mathbb{K}[x]_{(x)} \qquad \mathbb{Z}_p \longleftrightarrow \mathbb{K}[[x]]$$
$$\mathbb{Q} \longleftrightarrow \mathbb{K}(x) \qquad \mathbb{Q}_p \longleftrightarrow \mathbb{K}((x))$$

However, considering norms on these rings, we can conclude that this is not completely accurate. Theorem 1.12 told us that up to equivalence the non-Archimedean norms on $\mathbb{Q}$ are given by $|\cdot|_p$ and the only Archimedean norm is $|\cdot| = |\cdot|_\infty$. Noting that for $n, m \in \mathbb{N}$,

$$(n) \supseteq (m) \Leftrightarrow n|m,$$

we see that the maximal ideals of $\mathbb{Z}$ are exactly (p) for primes p, i.e. alternatively, we can view this as the non-Archimedean norms on $\mathbb{Q}$ being in correspondence with the maximal ideals of $\mathbb{Z}$.

> Venturing slightly beyond the scope of this paper, we can generalise this notion further.
>
> While $\mathbb{Z}$ and $\mathbb{K}[x]$ look different, they are actually prototypical examples of domains called *'Dedekind domains'*, for which the following holds:

Theorem 2.13. *Given a Dedekind domain X, there exists an injective map*

$$\{\textit{maximal ideals of } X\} \hookrightarrow \{\textit{non-Archimedean norms on } Frac(X), \ \textit{up to equivalence}\}.$$

Moreover, this map is an isomorphism if the class group of X is torsion, i.e. all elements have finite order.

This is the closest that we can get to an analogue of Theorem 1.12 for Dedekind domains in general. For additional details we refer the reader to [5, Theorem 4.2].

Working with $\mathbb{K} = \mathbb{C}$ (this field is nice to work in due to it being algebraically closed), the maximal ideals of $\mathbb{C}[x]$ are, by the fundamental theorem of algebra, given by

$$\{(x-a) : a \in \mathbb{C}\},$$

each of which gives rise to a non-Archimedean norm on $\mathbb{C}(x)$. However, in this case another non-Archimedean norm can be constructed from the valuation given by $\nu(f) = -\deg(f)$ for $f \in \mathbb{C}(x)$ and unlike when we considered $\mathbb{Q}$, an Archimedean norm doesn't exist.

Before reverting our attention to the discussion of DVFs and DVRs, we'll discuss some further results concerning localizations and the p-adics.

Lemma 2.14. *(i)* $\mathbb{Z} \subseteq \mathbb{Z}_{(p)} \subseteq \mathbb{Z}_p$

(ii) $\mathbb{Z}_{(p)} = \mathbb{Q} \cap \mathbb{Z}_p$

Proof. (i) The first inclusion is trivial, so we'll focus on the second. Suppose that $x = \frac{a}{b} \in \mathbb{Q}$ satisfies $p \nmid b$. Then $\text{ord}_p(b) = 0$ and so

$$\text{ord}_p(x) = \text{ord}_p(a) \geq 0 \Rightarrow |x|_p = \frac{1}{p^{\text{ord}_p(x)}} \leq 1.$$

(ii) $(\subseteq)$ The left inclusion is clear by the previous result.

$(\supseteq)$ For the right inclusion, let $x \in \mathbb{Q} \cap \mathbb{Z}_p$ so $x = \frac{a}{b} \in \mathbb{Q}$ and

$$\left|\frac{a}{b}\right|_p \leq 1 \Rightarrow \text{ord}_p(b) \leq \text{ord}_p(a). \tag{2}$$

Without loss of generality we may assume that not both a, b are multiples of p, if so,

$$\frac{a}{b} \sim \frac{\frac{a}{p}}{\frac{b}{p}}$$

with $u = 1_R \in R - (p)$. By (2), $\text{ord}_p(b) = 0$ so $x \in \mathbb{Z}_{(p)}$. □

A *'short exact sequence'* is a series of abelian group homomorphisms

$$0 \longrightarrow A \xrightarrow{f} B \xrightarrow{g} C \longrightarrow 0,$$

where f is injective, g is right-cancellative, $\{0\} = \ker(f)$, $\operatorname{im}(f) = \ker(g)$ and $\operatorname{im}(g) = \{0\}$. In this case, there exists the following isomorphism:

$$C \cong B/\operatorname{im}(f).$$

Further to this, a short exact sequence is *'split'* if there exists a homomorphism $h : C \longrightarrow B$ such that the composition $g \circ h$ is the identity on C. Thus, if the groups in question are all abelian,

$$B \cong A \times C.$$

This second quality is one which we will use later on. For now, we use this notion to derive:

Theorem 2.15. *The following isomorphism exists,*

$$\mathbb{Z}_p/p^n\mathbb{Z}_p \cong \mathbb{Z}/p^n\mathbb{Z} \cong \mathbb{Z}_{(p)}/p^n\mathbb{Z}_{(p)}.$$

Proof. Clearly the following defines a short exact sequence:

$$0 \longrightarrow \mathbb{Z}_p \xrightarrow{\cdot p^n} \mathbb{Z}_p \longrightarrow \mathbb{Z}/p^n\mathbb{Z} \longrightarrow 0,$$

and moreover $\operatorname{im}(\cdot p^n) = p^n\mathbb{Z}_p$. Thus, the first isomorphism holds. Replacing $\mathbb{Z}_p$ by $\mathbb{Z}_{(p)}$, we obtain the isomorphism between the first and last quotients. □

Now, upon noticing that $px - 1$ is irreducible over $\mathbb{Z}_p$ we introduce the notation

$$\mathbb{Z}_p\left[\frac{1}{p}\right] := \frac{\mathbb{Z}_p[x]}{(px-1)}.$$

Again, a similar argument this time involving the short exact sequence

$$0 \longrightarrow \mathbb{Z}_p[x] \xrightarrow{\cdot(px-1)} \mathbb{Z}_p[x] \longrightarrow \mathbb{Q}_p \longrightarrow 0,$$

provides the following isomorphisms:

Theorem 2.16. *(i)* $\mathbb{Z}_p[\frac{1}{p}] \cong \mathbb{Q}_p$,

(ii) $\mathbb{Z}_{(p)}[\frac{1}{p}] \cong \mathbb{Q}$.

Returning to the discussion of discrete valuation rings, we introduce a new concept.

Definition 2.17. A (commutative) ring is local if it has exactly one maximal ideal.

Example 2.18. $\mathbb{Z}$ is not local. □

Example 2.19. $\mathbb{K}$ is local. Notice that an ideal contains either no units and is the ideal (0), or it contains some unit and is the whole field $\mathbb{K}$. Thus (0) is the unique maximal ideal. □

Theorem 2.20. *A DVR is a local ring whose unique maximal ideal is*

$$\mathfrak{m} = \{x \in \mathbb{K} : \nu(x) > 0\}. \tag{3}$$

We will prove this by first stating some results that will be useful in later calculations.

Lemma 2.21. *(i) For $x, y \in R$, $\nu(x) \leq \nu(y) \Leftrightarrow x|y$.*

(ii) $R^{\times} = \{x \in \mathbb{K} : \nu(x) = 0\}$.

Proof. (i) ($\Rightarrow$) Suppose that $\nu(x) \leq \nu(y)$, and $x \neq 0$ (otherwise the statement is clear). Then

$$\nu\left(\frac{y}{x}\right) = \nu(y) - \nu(x) \geq 0, \text{ so } \frac{y}{x} \in R,$$

i.e. $x|y$.

($\Leftarrow$) If $x|y$ then for some $a \in R$, $y = ax$. So

$$\nu(y) = \nu(ax) = \nu(a) + \nu(x) \Rightarrow \nu(x) \leq \nu(y),$$

as each of these quantities is non-negative.

(ii) ($\subseteq$) Let $x \in R^{\times}$, then there exists some $y \in R^{\times}$ such that

$$0 = \nu(1) = \nu(xy) = \nu(x) + \nu(y) \Rightarrow \nu(x) = -\nu(y).$$

But both of these quantities are non-negative and therefore must be 0.

($\supseteq$) Let $x \in R$ be such that $\nu(x) = 0$, then

$$\nu\left(\frac{1}{x}\right) = -\nu(x) = 0 \Rightarrow \frac{1}{x} \in R \Rightarrow x \in R^{\times}. \qquad \square$$

Before proving the theorem, we may now utilise Lemma 2.21 to rewrite (3) as

$$\mathfrak{m} = R - R^{\times},$$

i.e. the ideal of non units in R.

Proof. Some simple verification confirms that $\mathfrak{m}$ is non-empty, it is a subring of R and $\forall r \in R$, $\forall x \in \mathfrak{m}$ we have $xr \in \mathfrak{m}$, hence $\mathfrak{m}$ is an ideal. Let I be another ideal of R such that $\mathfrak{m} \subset I$. Then I contains some unit of R and is actually itself R, therefore $\mathfrak{m}$ is maximal. Finally we need to obtain uniqueness of $\mathfrak{m}$. Suppose that J is yet another ideal of R, and that it too is maximal. Assume that J contains no units of R so that $J \neq R$, but then $J \subseteq \mathfrak{m}$, contradicting maximality. □

Applying Theorem 2.20, we obtain these further examples:

Example 2.22. $\mathbb{Z}_p$ is local and it's maximal ideal is $p\mathbb{Z}_p = (p)$. □

Example 2.23. $\mathbb{Z}_{(p)}$ is local and it's maximal ideal is $\left\{\frac{a}{b} \in \mathbb{Q} : p \mid a,\, p \nmid b\right\}$. More generally, it is true that the localisation of a ring R at a prime ideal I of R is local. □

Moreover, we can work further to classify *all* the ideals of R.

Proposition 2.24. *R is a PID.*

Proof. Let $(0) \neq I \lhd R$, thus there exists some $x \in I$. If $\nu(x) = 0$, then $I = R = (1)$. Else, $\nu(x) > 0$ and there exists some $x_0 \in I$ such that

$$\nu(x_0) = n = \min\{\nu(x) : x \in I\} > 0.$$

Now, choose an arbitrary element $y \in I$ so,

$$\nu(y) \geq n \Rightarrow \nu\left(\frac{y}{x_0}\right) \geq 0 \Rightarrow \frac{y}{x_0} \in R \Rightarrow y \in (x_0).$$

As we chose y arbitrarily, it is the case that $I \subseteq (x_0)$. Thus I is principal. □

Definition 2.25. A generator of a maximal ideal is called a uniformizer.

Using that R is a PID, the maximal ideal $\mathfrak{m}$ has a uniformizer, call it m. Without loss of generality we may assume that $\nu(m) = 1$.

We now have enough information to be able to list all the ideals of R.

Theorem 2.26. *Let R be a DVR, then it's ideals are given by (0), $\mathfrak{m}^n$ for $n \in \mathbb{Z}_{\geq 1}$ and R.*

Proof. Clearly as R is a PID, its ideals are

$$\{(x) : x \in R\}.$$

By Lemma 2.21(i), $(x) = (y) \Leftrightarrow \nu(x) = \nu(y)$, so the distinct ideals are given by:

$$(x_0),\ (x_1),\ (x_2),\ (x_3),\ \ldots \text{ and } (0),$$

where $x_n \in R$ and $\nu(x_n) = n$ for all $n \in \mathbb{N}$. Noting that $\nu(m^n) = n$, we can take $x_n = m^n$ and the result follows. □

3 Solving polynomials in $\mathbb{Z}_p$

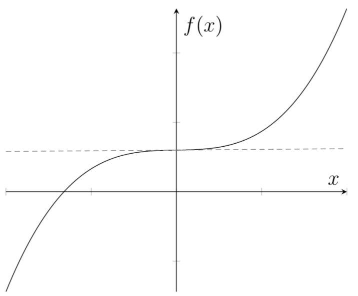
In this chapter we discuss Hensel's lemma, concerning the solvability of p-adic polynomials. Further, we apply this to be able to discuss the p-adic roots of unity and determine exactly when degree two homogenous polynomials in three variables have a non-trivial rational solution (for the latter, we refer to [1] for additional detail).

3.1 Hensel's lemma

Let's first restrict our attention to finding roots in $\mathbb{Z}_p$.

In calculus, when considering some function $f(x) \in \mathbb{R}[x]$, Newton's method is an effective way to find it's roots, i.e. given an initial guess it successively produces better approximations.

Geometrically, if the initial guess is x_0, the next approximation is the x-coordinate of the intersection of the x-axis with the tangent of the graph at $(x_0, f(x_0))$. For this method to be valid, it is crucial that such an intersection point exists or more precisely, that $f'(x_0) \neq 0$. In the figure, $f'(0) = 0$ and so the method would fail.



This argument for finding roots can be extended to $\mathbb{Z}_p$ and is summarised in the next result:

Theorem 3.1 (Hensel's lemma). *Let $f(x) \in \mathbb{Z}_p[x]$ and suppose that there exists some $\alpha_1 \in \mathbb{Z}_p$ such that the following are satisfied:*

(a) $f(\alpha_1) \equiv 0$ (mod p)

(b) $f'(\alpha_1) \not\equiv 0$ (mod p).

Then there exists some unique $\alpha \in \mathbb{Z}_p$ such that:

(i) $\alpha_1 \equiv \alpha$ (mod p)

(ii) $f(\alpha) = 0$.

Remark. *To make precise the definition of the derivative in this context; if*

$$f(x) = a_0 + a_1x + a_2x^2 + \ldots + a_nx^n \in \mathbb{Z}_p[x],$$

then

$$f'(x) = a_1 + 2a_2x + \ldots + na_nx^{n-1} \in \mathbb{Z}_p[x].$$

Before proving and discussing some applications of this theorem, we look further into the conditions it requires to be met.

Definition 3.2. A polynomial $f(x) \in \mathbb{K}[x]$ of degree n is said to be separable if it has n distinct roots in some extension $\mathbb{K} \subset \mathbb{L}$.

Proposition 3.3. *A polynomial $f(x) \in \mathbb{K}[x]$ is separable if and only if $f(x)$ and $f'(x)$ have no common roots in any extension of $\mathbb{K}$.*

Proof. ($\Rightarrow$) Suppose that $f(x) \in \mathbb{K}[x]$ is separable and in some $\mathbb{K} \subset \mathbb{L}$,

$$f(x) = \prod_{i=1}^{n} (x - \lambda_i),$$

so that

$$f'(x) = \sum_{i=1}^{n} (x - \lambda_1) \cdots (x - \lambda_{i-1})(x - \lambda_{i+1}) \cdots (x - \lambda_n).$$

If $f(x)$ and $f'(x)$ have a root in common in $\mathbb{L}$, then for some i, $(x - \lambda_i)|f'(x)$ which is impossible. Thus $f(x)$ and $f'(x)$ cannot have a root in common in $\mathbb{L} \Rightarrow f(x)$ and $f'(x)$ cannot have a root in common in $\mathbb{K}$.

($\Leftarrow$) Suppose that $f(x) \in \mathbb{K}[x]$ is inseparable, i.e. in all extensions $\mathbb{K} \subset \mathbb{L}$, there is some $\lambda \in \mathbb{L}$ and $g(x) \in \mathbb{L}[x]$ such that

$$f(x) = (x - \lambda)^2 g(x).$$

Hence,

$$f'(x) = (x - \lambda)\{2g(x) + (x - \lambda)g'(x)\}$$

and $(x - \lambda)$ is a common factor of $f(x)$ and $f'(x)$ in $\mathbb{L}[x]$. I claim that $f(x)$ and $f'(x)$ must therefore have a non-trivial common factor in $\mathbb{K}[x]$. Suppose not, then there exist $\phi(x)$, $\psi(x) \in \mathbb{K}[x]$ such that

$$\phi(x)f(x) + \psi(x)f'(x) = 1$$

in $\mathbb{K}[x]$, then this also holds in $\mathbb{L}[x]$ which is a contradiction. □

Returning to $\mathbb{Z}_p$ from this discussion of $\mathbb{K}[x]$, the notion of separability motivates us to rephrase the conditions stated in Hensel's lemma.

Proposition 3.4. *The conditions (a) and (b) stated in Hensel's lemma are equivalent to the following:*

$$\bar{\alpha}_1 \in \mathbb{F}_p \textit{ is a root of } \bar{f}(x) \in \mathbb{F}_p[x] \textit{ of multiplicity } 1, \tag{4}$$

where the overbar notation represents the reduction modulo p from $\mathbb{Z}_p$ to $\mathbb{F}_p$.

Proof. First assume that (a) and (b) hold. If $\bar{\alpha}_1 \in \mathbb{F}_p$ is a repeated root of $\bar{f}(x) \in \mathbb{F}_p[x]$ then $\bar{f}(x)$ is inseparable in $\mathbb{F}_p[x]$. By Proposition 3.3, $\bar{\alpha}_1$ is also a root of $\bar{f}'(x) \in \mathbb{F}_p[x]$, contradicting (b). Conversely, assuming that (4) holds, (a) follows immediately. Supposing that $f(\alpha_1) \equiv 0 \pmod{p}$, i.e. that $\bar{f}'(\bar{\alpha}_1) = 0$, then $\bar{f}(x)$ and $\bar{f}'(x)$ have the root $\bar{\alpha}_1$ in common in $\mathbb{F}_p$, a contradiction. So (b) follows too. □

Corollary 3.5. *Let $f(x) \in \mathbb{Z}_p[x]$. If $\bar{f}(x) \in \mathbb{F}_p[x]$ is separable, then the conditions of Hensel's lemma are satisfied by default.*

We shall now prove Hensel's lemma, a proof taken from [1, Section 3.4].

Proof. Using the notion of $\mathbb{Z}_p$ being the completion of $\mathbb{Z}$ with respect to the p-adic norm, we construct α as being the limit of some Cauchy sequence of integers $\alpha_1, \alpha_2, \alpha_3, \ldots$. We want to define our sequence according to the criteria: for $n \geq 1$,

- $f(\alpha_n) \equiv 0 \pmod{p^n}$,
- $\alpha_n \equiv \alpha_{n+1} \pmod{p^n}$.

We will show that such a choice of integers is possible by induction. Take $\alpha_2 = \alpha_1 + k_1 p$ for some $k_1 \in \mathbb{Z}_p$, then by means of a Taylor expansion:

$$f(\alpha_2) = f(\alpha_1) + f'(\alpha_1)k_1 p + \ldots \equiv f(\alpha_1) + f'(\alpha_1)k_1 p \pmod{p^2}.$$

By condition (a), there exists some $x_1 \in \mathbb{Z}_p$ such that $f(\alpha_1) = x_1 p$, so

$$\begin{aligned} f(\alpha_2) &\equiv 0 \pmod{p^2} \\ &\Leftrightarrow f(\alpha_1) + f'(\alpha_1)k_1 p \equiv 0 \pmod{p^2} \\ &\Leftrightarrow x_1 p + f'(\alpha_1)k_1 p \equiv 0 \pmod{p^2} \\ &\Leftrightarrow x_1 + f'(\alpha_1)k_1 \equiv 0 \pmod{p}. \end{aligned}$$

Now, using condition (b), that $f'(\alpha_1)$ is invertible, such a choice of k_1 is possible. In fact, adding the constraint that $0 \leq k_1 \leq p-1$ allows the choice of k_1 and consequently α_2 to be unique. A final verification is needed for the rest of the induction to work:

$$f'(\alpha_2) = f'(\alpha_1) + f''(\alpha_1)k_1 p + \ldots \equiv f'(\alpha_1) \pmod{p},$$

so $f'(\alpha_2) \not\equiv 0$. Now suppose that we have found $\alpha_1, \alpha_2, \alpha_3, \ldots, \alpha_n$ with $f(\alpha_n) = x_n p^n$ for some $x_n \in \mathbb{Z}_p$ and $f'(\alpha_n) \not\equiv 0 \pmod{p}$. Let $\alpha_{n+1} = \alpha_n + k_n p^n$ for some $k_n \in \mathbb{Z}_p$ that we want to find. Again using the Taylor expansion:

$$\begin{aligned} f(\alpha_{n+1}) &\equiv 0 \pmod{p^{n+1}} \\ &\Leftrightarrow f(\alpha_n) + f'(\alpha_n)k_n p^n \equiv 0 \pmod{p^{n+1}} \\ &\Leftrightarrow x_n p^n + f'(\alpha_n)k_n p^n \equiv 0 \pmod{p^{n+1}} \\ &\Leftrightarrow x_n + f'(\alpha_n)k_n \equiv 0 \pmod{p}, \end{aligned}$$

where a unique choice of $0 \leq k_n \leq p-1$ and α_{n+1} is possible. Furthermore, $f'(\alpha_{n+1}) \not\equiv 0$.

We now proceed to checking that the sequence we have constructed is actually a Cauchy sequence, so that defining α to be its limit makes sense. This is actually clear upon noticing that if $m > n$, then

$$\alpha_n \equiv \alpha_m \pmod{p^n}, \tag{5}$$

so

$$|\alpha_m - \alpha_n|_p \leq \frac{1}{p^n}.$$

Writing

$$\alpha = \lim_{n\to\infty} \alpha_n,$$

it remains to check that (i) and (ii) hold. Taking $n = 1$ and the limit as $m \to \infty$ in (5) confirms (i). More generally, using the continuity of $f(x)$,

$$\alpha \equiv \alpha_n \pmod{p^n} \Rightarrow f(\alpha) \equiv f(\alpha_n) \equiv 0 \pmod{p^n}.$$

Hence

$$\forall n,\ |f(\alpha)|_p \leq \frac{1}{p^n} \Rightarrow |f(\alpha)|_p = 0 \Leftrightarrow f(\alpha) = 0. \qquad \square$$

The content of this proof can be applied directly to construct a solution to any polynomial which has an approximate solution.

Example 3.6. Let $f(x) = x^2 - 11 \in \mathbb{Z}_7[x]$ and take $\alpha_1 = 2$. Clearly,

$$f(\alpha_1) = -7 \equiv 0 \pmod 7 \text{ and } f'(\alpha_1) = 4 \not\equiv 0 \pmod 7,$$

so the conditions of Hensel's lemma are satisfied.

Using the notation as in the proof of Hensel's lemma, we also have

$$x_1 = -1 \Rightarrow k_1 \equiv (4)^{-1} \pmod 7, \text{ i.e. } k_1 = 2 \Rightarrow \alpha_2 = 16 = 2 + 2\cdot 7.$$

We have now that

$$f(\alpha_2) = 245 \equiv 0 \pmod{49} \text{ and } f'(\alpha_2) = 4 \not\equiv 0 \pmod 7,$$

so

$$x_2 = 5 \Rightarrow k_2 \equiv -5\cdot(4)^{-1} \pmod 7, \text{ i.e. } k_2 = 4 \Rightarrow \alpha_3 = 212 = 2 + 2\cdot 7 + 4\cdot 7^2.$$

Repeating,

$$f(\alpha_3) = 44933 \equiv 0 \pmod{343} \text{ and } f'(\alpha_3) = 4 \not\equiv 0 \pmod 7,$$

so

$$x_3 = 131 \Rightarrow k_3 = 4 \Rightarrow \alpha_4 = 1584 = 2 + 2\cdot 7 + 4\cdot 7^2 + 4\cdot 7^3.$$

Performing one more iteration,

$$f(\alpha_4) = 2509045 \equiv 0 \pmod{2401} \text{ and } f'(\alpha_4) = 4 \not\equiv 0 \pmod 7,$$

so

$$x_4 = 1045 \Rightarrow k_4 = 3 \Rightarrow \alpha_5 = 8444 = 2 + 2\cdot 7 + 4\cdot 7^2 + 4\cdot 7^3 + 3\cdot 7^4.$$

So we are getting closer to approximating α, a solution to $f(x)$, where

$$\alpha \equiv \alpha_n \pmod{p^n}. \qquad \square$$

A less obvious consequence of Hensel's lemma is the following, which will prove useful later on in the paper.

Corollary 3.7. *(i) An element $a \in \mathbb{Z}_2^\times$ is a square in $\mathbb{Q}_2$ if and only if it's a quadratic residue modulo* 8.

(ii) For an odd prime p, an element $a \in \mathbb{Z}_p^\times$ is a square in in $\mathbb{Q}_p$ if and only if it's a quadratic residue modulo p.

Proof. (i) ($\Leftarrow$) The quadratic residues modulo 8 are 1 and 4. As $a \in \mathbb{Z}_2^\times$, clearly we must have

$$a \equiv 1 \pmod 8 \Rightarrow \exists k \in \mathbb{Z}_2 \text{ such that } a = 8k + 1.$$

Consider the equation

$$f(x) = x^2 + x - 2k \in \mathbb{Z}_2[x],$$

whose roots are the roots of

$$g(x) = (2x+1)^2 - (8k+1) \in \mathbb{Z}_2[x].$$

As $f(0) \equiv 0 \pmod 2$ and $f'(0) = 1 \not\equiv 0 \pmod 2$, we may apply Hensel's lemma to $f(x)$ to obtain a root of $g(x)$. We see that a is a square in $\mathbb{Z}_2^\times$.

(ii) ($\Leftarrow$) Let $a \in \mathbb{Z}_p^\times$ be a quadratic residue modulo p. Then,

$$\exists b \in \mathbb{Z}/p\mathbb{Z} \text{ such that } 1 \equiv a \equiv b^2 \pmod p.$$

Writing,

$$f(x) = x^2 - a \in \mathbb{Z}_p[x],$$

we have that $f(b) \equiv 0 \pmod p$ and $f'(b) \equiv 2b \not\equiv 0 \pmod p$, so Hensel's lemma applies and a root of $f(x)$ exists in $\mathbb{Z}_p$, i.e. a is a square in $\mathbb{Z}_p$ □

Example 3.8. 17 is a square in $\mathbb{Z}_2$. □

Example 3.9. As $6^2 \equiv 34 \equiv 2 \pmod{17}$, 2 is a quadratic residue modulo 17. Therefore, 2 is a square in $\mathbb{Z}_{17}$. □

3.1.1 Roots of unity

Another application of Hensel's lemma concerns roots of unity. We know that the only roots of unity lying in $\mathbb{Q}$, or even $\mathbb{R}$, are $+1$ and -1. Can we classify all roots of unity lying in $\mathbb{Q}_p$?

Finding m-th roots of unity is precisely the task of finding roots of

$$f(x) = x^m - 1.$$

A first observation is that p-adic roots of unity are necessarily units in the p-adic integers. Suppose that $\zeta \in \mathbb{Q}_p$ is an m-th root of unity, there exists some $k \in \mathbb{Z}$ such that

$$|\zeta|_p = \frac{1}{p^k},$$

thus we can conclude:

$$|1|_p = |\zeta^m|_p = |\zeta|_p^m = \frac{1}{p^{km}} \Rightarrow k = 0 \Rightarrow \zeta \in \mathbb{Z}_p^{\times}.$$

Corollary 3.10. *For odd p, $\mathbb{Z}_p$ contains exactly $p-1$ distinct $(p-1)$th roots of unity. Furthermore, they are distinct modulo p.*

Proof. The equation

$$f(x) = x^{p-1} - 1 \in \mathbb{Q}_p[x]$$

has at most $p-1$ roots in $\mathbb{Q}_p$, as it is a field. Let $1 \leq a \leq p-1$ be an integer, by Fermat's little theorem, it's true that

$$f(a) \equiv 0 \pmod{p} \text{ and } f'(a) = (p-1)a^{p-2} \not\equiv 0 \pmod{p},$$

so Hensel's lemma tells us that there's a root of $f(x)$ in $\mathbb{Z}_p$ which is congruent to a modulo p. We therefore obtain distinct roots of $f(x)$ for each a, i.e. there are exactly $p-1$ roots. □

Remark. *We will see the corresponding result for $p = 2$ later on.*

Recall that,

Definition 3.11. A primitive m-th root of unity is ζ such that m is the smallest non-zero natural number for which

$$\zeta^m = 1.$$

The harder problem of finding these primitive m-th roots of unity is closely related to what is called the m-th cyclotomic polynomial, $\Phi_m(x)$.

Definition 3.12. The mth cyclotomic polynomial, $\Phi_m(x)$, is the unique, irreducible, monic polynomial whose roots are exactly the m-th primitive roots of unity.

Notice that this gives rise to the following key relation:

$$x^n - 1 = \prod_{m|n} \Phi_m(x). \tag{6}$$

Proposition 3.13. *For any prime p and $m \in \mathbb{Z}_{\geq 0}$ not divisible by p, there exists a primitive m-th root of unity in $\mathbb{Q}_p$ if and only if m divides $p-1$.*

Proof. ($\Rightarrow$) Suppose that $m \nmid p-1$ and that a primitive m-th root of unity in $\mathbb{Q}_p$ exists, i.e. $\exists \alpha \in \mathbb{Q}_p$ such that $\Phi_m(x) \in \mathbb{Z}_p[x]$ is satisfied (in fact, by our earlier discussion $\alpha \in \mathbb{Z}_p^\times$). Then $\bar{\alpha} \in \mathbb{F}_p$ is a root of $\bar{\Phi}_m(x) \in \mathbb{F}_p[x]$, so $\bar{\alpha}$ is a primitive m-th root of unity in $\mathbb{F}_p$. This is equivalent to $\bar{\alpha}$ having order m in $\mathbb{F}_p^\times$, but $m \nmid p-1$ so this is impossible. Thus there cannot be a primitive m-th root of unity in $\mathbb{Q}_p$.
($\Leftarrow$) It is a well-known fact that $\mathbb{F}_p^\times$ is a cyclic group of order $p-1$, thus there exists some $a \in \mathbb{F}_p^\times$, of order $p-1$ such that

$$\langle a \rangle = \mathbb{F}_p^\times.$$

As $m|p-1$, we have that $\bar{\alpha} = a^{\frac{p-1}{m}} \in \mathbb{F}_p^\times$ is of order m, thus is a root of $\bar{\Phi}_m(x) \in \mathbb{F}_p[x]$. Using (6), we may write

$$x^m - 1 = \Phi_m(x)Q(x), \text{ where } Q(x) := \prod_{\substack{m|n \\ m\neq n}} \Phi_m(x).$$

By reducing modulo p, differentiating and evaluating at $\bar{\alpha}$, we obtain

$$m\bar{\alpha}^{m-1} = \bar{\Phi}'_m(\bar{\alpha})\bar{Q}(\bar{\alpha}) + \bar{\Phi}_m(\bar{\alpha})\bar{Q}'(\bar{\alpha}).$$

Additionally, using that $\bar{\Phi}_m(\bar{\alpha}) = 0$ and $Q(\bar{\alpha}) \neq 0$, it is clear that $\bar{\alpha}$ is not a root of $\bar{\Phi}'_m(x) \in \mathbb{F}_p[x]$. The conditions of Hensel's lemma are therefore met, so we obtain a root of $\Phi_m(x)$ in $\mathbb{Z}_p$, i.e. a primitive m-th root of unity in $\mathbb{Q}_p$. □

3.2 Local and global

The problem of finding the roots of a given polynomial in $\mathbb{Q}$ is a notoriously hard one. Notice that if a polynomial has a root in $\mathbb{Q}$, i.e. a *'global'* solution, then for all $p \leq \infty$, it must have a root in $\mathbb{Q}_p$, i.e. a *'local'* solution. Equivalently, if there exists a $p \leq \infty$ such that a given polynomial has no roots in $\mathbb{Q}_p$, then the polynomial cannot have roots in $\mathbb{Q}$.

Example 3.14. Consider the polynomial

$$3x^2 + 2y^2 - z^2 = 0.$$

By the following argument, this has no non-trivial solution in $\mathbb{Q}_3$. On the contrary, suppose that (x_0, y_0, z_0) is a non-trivial solution in $\mathbb{Q}_3$, without loss of generality we may assume that $y_0, z_0 \in \mathbb{Z}_3^\times$. Thus,

$$3{x_0}^2 + 2{y_0}^2 - {z_0}^2 \equiv 0 \pmod 3 \Rightarrow 2{y_0}^2 \equiv {z_0}^2 \pmod 3 \Rightarrow 2 \equiv \left(\frac{z_0}{y_0}\right)^2 \pmod 3,$$

using that ${y_0}^2$ has invertible image in $\mathbb{F}_3$. However, 2 is not a quadratic residue module 3.

From this, we may conclude that there are no non-trivial rational solutions to the polynomial. □

A situation in which the converse holds, and local implies global, is the following:

Proposition 3.15. *$a \in \mathbb{Q}$ is a square if and only if $a \in \mathbb{Q}_p$ is a square for all $p \leq \infty$.*

Proof. ($\Rightarrow$) This is trivial as for all $p \leq \infty$, $\mathbb{Q} \subseteq \mathbb{Q}_p$.
($\Leftarrow$) Suppose that $a \in \mathbb{Q}_p$ is a square for all $p \leq \infty$. When $p = \infty$ this is precisely the statement that a is non-negative. Notice that for $p < \infty$ and any non-negative $a \in \mathbb{Q}_p$, we may write

$$a = \prod_{p<\infty} p^{\mathrm{ord}_p(a)}.$$

So such an a being a square in $\mathbb{Q}_p$ for all $p < \infty$ is the statement that $\mathrm{ord}_p(a)$ is even for all $p < \infty$. Clearly, $a \in \mathbb{Q}$ is a square. □

Ideally we'd like to say that this converse holds all the time, but this is not the case. The following provides a counterexample.

Example 3.16. The equation

$$(x^2 - 2)(x^2 - 17)(x^2 - 34) = 0$$

clearly has no roots in $\mathbb{Q}$. We will show that is does however have roots in $\mathbb{Q}_p$ for all $p \leq \infty$.

The statement is clear for $p = \infty$ so we'll restrict to the case where $p < \infty$.

By Corollary 3.7, Example 3.8 and Example 3.9, the equation has roots in both $\mathbb{Q}_2$ and $\mathbb{Q}_{17}$.

When $p \neq 2$, 17, all of 2, 17 and 34 are p-adic units and moreover, one of them is always a quadratic residue modulo p. So again we obtain a root of the equation in $\mathbb{Q}_p$. □

What we can say with more certainty is that the existence or non-existence of global solutions can be detected by studying, for each $p \leq \infty$, the local solutions. A hugely successful application of this principle is the following:

Theorem 3.17 (Hasse–Minkowski). *Let $f(x_1, \dots, x_n) \in \mathbb{Q}[x_1, \dots, x_n]$ be a homogenous polynomial of degree 2. The equation*

$$f(x_1, \dots, x_n) = 0$$

has non-trivial solutions in $\mathbb{Q}$ if and only if it has non-trivial solutions in $\mathbb{Q}_p$ for all $p \leq \infty$.

We shall not provide a proof of this theorem here, instead see [6], but we shall discuss a direct application.

3.2.1 Finding the rational roots of $ax^2 + by^2 + cz^2$

To be precise, we are considering

$$ax^2 + by^2 + cz^2 \in \mathbb{Q}[x, y, z].$$

However, without loss of generality there are some additional assumptions we can make regarding the coefficients a, b, c.

- ◦ Clearly, we can take the coefficients to be non-zero, if any one of the coefficients is 0, there are countably many non-trivial rational solutions $(x_0,\, y_0,\, z_0)$.
- ◦ By 'clearing denominators' of the coefficients, we may also assume that $a,\, b,\, c \in \mathbb{Z}$.
- ◦ We may also assume that each coefficient is *'square-free'*, i.e. a product of distinct prime factors. In the case that a coefficient is not square free, it's square factors can be absorbed into it's corresponding unknown variable.
- ◦ Finally, we may assume that there is no common factor of all coefficients. Further, we may assume that the coefficients are pairwise coprime: suppose not, i.e that the highest common factor of a pair is λ, multiplying the whole equation by λ we may then absorb this factor into the unknowns corresponding to the coefficients, by repeating this the result follows. Note that this is the same as saying that the product abc is also square-free.

Example 3.18. A trivial case to consider is when $a = b = c = 1$, i.e. solving the equation

$$x^2 + y^2 + z^2 = 0.$$

Clearly the only solution to this in $\mathbb{R}$ is $x = y = z = 0$, hence there are no non-trivial solutions in $\mathbb{Q}$. □

Under some circumstances we may be able to apply a number theoretic argument to determine if non-trivial solutions exist.

Example 3.19. Consider the polynomial $x^2 - 5y^2 - 3z^2 = 0$. Reducing modulo 3,

$$x^2 - 5y^2 \equiv 0 \pmod 3 \Rightarrow x^2 \equiv 2y^2 \pmod 3,$$

thus $x \equiv y \equiv 0 \pmod 3$. Further,

$$x^2 \equiv y^2 \equiv 0 \pmod 9 \Rightarrow z \equiv 0 \pmod 3 \Rightarrow x^2 - 5y^2 \equiv 0 \pmod{27}.$$

Repeating, we obtain that x and y are both divisible by arbitrarily large powers of 3 and must therefore be 0. This structure of argument is called *'Fermat's method of infinite descent'*. □

The Hasse-Minkowski theorem provides a more consistent method for determining whether $ax^2 + by^2 + cz^2$ has non-trivial solutions in $\mathbb{Q}$. This gives rise to the next major theorem.

Theorem 3.20. *Let a, b, c be pairwise coprime, square-free, integers. The equation*

$$ax^2 + by^2 + cz^2 = 0$$

has non-trivial solutions in $\mathbb{Q}$, if and only if the following conditions are satisfied:

(i) a, b, c are not all positive or negative.

(ii) for each odd prime p dividing a, $-\frac{b}{c}$ is a quadratic residue modulo p and similarly for the odd primes dividing b and c.

(iii) if a, b, c are all odd, then there are two of them whose sum is divisible by 4.

(iv) if a is even, then either $b+c$ or $a+b+c$ is divisible by 8 and similarly if b or c is even.

We will work towards proving that the conditions of Theorem 3.20 being met are necessary to guarantee the existence of a non-trivial rational solution. For the proof in the opposite direction, see [1, Section 3.5] for some justification.

The conditions appearing at each prime p come from the following, to which we apply Hasse–Minkowski and obtain Theorem 3.20:

Theorem 3.21. *Let a, b, c be pairwise coprime, square-free, integers. The equation*

$$ax^2 + by^2 + cz^2 = 0$$

has non-trivial solutions in $\mathbb{Q}_p$ when:

(i) $p = \infty$, if and only if a, b, c are not all positive or negative.

(ii) p is odd and p does not divide a, b, c.

(iii) p is odd and p dividing a, if and only if $-\frac{b}{c}$ is a quadratic residue modulo p, similarly for p dividing b and c.

(iv) p is even and a, b, c are all odd, if and only if either $a+b$, $b+c$ or $a+c$ is divisible by 4.

(v) p is even and a is even, if and only if either $b+c$ or $a+b+c$ is divisible by 8, similarly if b or c is even.

A crucial ingredient to proving this is a slightly different formulation of Hensel's lemma, which can be applied to polynomials in multiple variables.

Lemma 3.22 (Hensel's lemma in three variables, [7, Page 451])**.** *Let $f(x_1, x_2, x_3) \in \mathbb{Z}_p[x_1, x_2, x_3]$ and suppose that there exists some $(\alpha_1, \alpha_2, \alpha_3) \in \mathbb{Z}_p^3$ such that there exists $n \geq 0$ satisfying the following:*

(a) $f(\alpha_1, \alpha_2, \alpha_3) \equiv 0 \pmod{p^{2n+1}}$,

(b) for some i, $\frac{\partial f}{\partial x_i}(\alpha_1, \alpha_2, \alpha_3) \equiv 0 \pmod{p^n}$ and $\frac{\partial f}{\partial x_i}(\alpha_1, \alpha_2, \alpha_3) \not\equiv 0 \pmod{p^{n+1}}$.

Then there exists some unique $(\beta_1, \beta_2, \beta_3) \in \mathbb{Z}_p^3$ such that:

(i) for all i, $\alpha_i \equiv \beta_i \pmod{p^{n+1}}$,

(ii) $f(\beta_1, \beta_2, \beta_3) = 0$.

Returning now to Theorem 3.21:

Proof. ($\Leftarrow$) (i) This is immediate.

(ii) Suppose that p is odd, dividing none of a, b, c. We claim that there is always a non-trivial integer solution $(x_0,\, y_0,\, z_0)$ modulo p, moreover, we can assume that $0 \leq x_0,\, y_0,\, z_0 \leq p-1$ and that not all are divisible by p. Our goal is to count the number, N, of the p^3 triples $(x_0,\, y_0,\, z_0)$, which do not provide a solution modulo p. Notice that by Fermat's little theorem:

$$(ax_0^2 + by_0^2 + cz_0^2)^{p-1} \equiv \begin{cases} 1 \pmod{p}, & \text{if } (x_0,\, y_0,\, z_0) \text{ is not a solution modulo } p \\ 0 \pmod{p}, & \text{if } (x_0,\, y_0,\, z_0) \text{ is a solution modulo } p, \end{cases}$$

thus we may write

$$\begin{aligned} N &\equiv \sum_{x_0,\, y_0,\, z_0 = 1}^{p-1} (ax_0^2 + by_0^2 + cz_0^2)^{p-1} \pmod{p} \\ &\equiv \sum_{i+j+k=p-1} \mu_{i,\,j,\,k} \sum_{x_0=0}^{p-1} x_0^{2i} \sum_{y_0=0}^{p-1} y_0^{2j} \sum_{z_0=0}^{p-1} z_0^{2k} \pmod{p} \end{aligned}$$

where

$$\mu_{i,\,j,\,k} = \frac{(p-1)!}{i!j!k!} a^i b^j c^k.$$

Without loss of generality, we may assume that $2k < p-1$ and henceforth claim that

$$\sum_{z_0=0}^{p-1} z_0^{2k} \equiv 0 \pmod{p}.$$

To justify this claim, we suppose that $2k > 0$, else the statement is trivial. Let $g \in \mathbb{F}_p$ be a primitive root, so that $g^{2k} \not\equiv 1 \pmod{p}$. As primitive roots are invertible,

$$\{0,\, 1,\, 2,\, \ldots,\, p-1\} = \{0,\, g,\, 2g,\, \ldots,\, (p-1)g\},$$

from which we conclude

$$\sum_{z_0=0}^{p-1} z_0^{2k} \equiv g^{2i} \sum_{z_0=0}^{p-1} z_0^{2k} \pmod{p} \Rightarrow \sum_{z_0=0}^{p-1} z_0^{2k} \equiv 0 \pmod{p}.$$

It now clearly follows that the number of non-solutions to the polynomial modulo p is divisible by p, thus so is the number of solutions. Finally, the existence of the trivial solution then guarantees the existence of a non-trivial solution. Given such a solution, there clearly exists some variable with respect to which the partial derivative evaluated at $(x_0,\, y_0,\, z_0)$ is non-zero modulo p, so Lemma 3.22 holds with $n = 0$ and a solution of f in $\mathbb{Q}_p$ exists.

(iii) Suppose that p is odd and does divide one of a, b, c, without loss of generality, assume this to be a. It is key here to notice that

$$-\frac{b}{c} \text{ is a quadratic residue modulo } p \Leftrightarrow \exists r \in \mathbb{Z} \text{ such that } b + cr^2 \equiv 0 \pmod{p}.$$

This suggests a candidate for a solution of f modulo p: let $(\alpha_1,\, \alpha_2,\, \alpha_3) = (1,\, 1,\, r)$ and $n = 0$. Clearly,

$$f(1,\, 1,\, r) = a + b + cr^2 \equiv 0 \pmod{p} \text{ and } \frac{\partial f}{\partial y}(1,\, 1,\, r) = 2b \not\equiv 0 \pmod{p}.$$

So by Lemma 3.22 we obtain a solution of f in $\mathbb{Q}_p$.

(iv) Suppose that $p = 2$ and all of a, b, c are odd, without loss of generality, suppose too that $b + c$ is divisible by 4. Here there are two separate cases to consider. First, let $b + c \equiv 0 \pmod{8}$, $(\alpha_1,\, \alpha_2,\, \alpha_3) = (4,\, 1,\, 1)$ and $n = 1$. Clearly,

$$\begin{aligned} f(4,\, 1,\, 1) = 16a + b + c \equiv 0 \pmod{8} \text{ and } \frac{\partial f}{\partial y}(4,\, 1,\, 1) = 2b &\equiv 0 \pmod{2} \\ &\not\equiv 0 \pmod{4} \end{aligned}$$

(similarly for the partial derivative with respect to z), so a solution of f in $\mathbb{Q}_2$ exists by Lemma 3.22. Next, let $b + c \equiv 4 \pmod{8}$, $(\alpha_1,\, \alpha_2,\, \alpha_3) = (2,\, 1,\, 1)$ and $n = 1$. In this case, we have that,

$$\begin{aligned} f(2,\, 1,\, 1) = 4a + b + c \equiv 0 \pmod{8} \text{ and } \frac{\partial f}{\partial y}(2,\, 1,\, 1) = 2b &\equiv 0 \pmod{2} \\ &\not\equiv 0 \pmod{4} \end{aligned}$$

(similarly for the partial derivative with respect to z), so again a solution of f in $\mathbb{Q}_2$ exists by Lemma 3.22.

(v) Suppose that $p = 2$ and one of a, b, c is even, without loss of generality, assume this to be a. Again there are two separate cases: when $b + c \equiv 0 \pmod{8}$ take $(\alpha_1,\, \alpha_2,\, \alpha_3) = (2,\, 1,\, 1)$ and $n = 1$; when $a + b + c \equiv 0 \pmod{8}$, take $(\alpha_1,\, \alpha_2,\, \alpha_3) = (1,\, 1,\, 1)$ and $n = 1$. The argument follows identically to that of (iv), providing a solution in $\mathbb{Q}_2$. □

To summarise, what we have just proved explains the conditions of Theorem 3.21, and proves that there are solutions in $\mathbb{Q}$ only if the conditions of Theorem 3.20 are satisfied. A proof of the conditions of Theorem 3.20 being sufficient to produce a non-trivial rational solution has been omitted, but [1, Section 3.5] provides some justification.

We conclude this section with one final example.

Example 3.23. Consider the polynomial

$$3x^2 + 5y^2 - 7z^2 = 0.$$

We will check whether the conditions of Theorem 3.21 are met to determine exactly for which p a non-trivial solution in $\mathbb{Q}_p$ exists. We will then be able to determine if a non-trivial rational solution exists. For clarity, we write $a = 3$, $b = 5$ and $c = -7$.

It is immediate that (i) holds, i.e. there is a non-trivial solution in $\mathbb{Q}_\infty = \mathbb{R}$.

Condition (ii) tells us that for odd $p \nmid 3 \cdot 5 \cdot 7 = 105$, a non-trivial solution in $\mathbb{Q}_p$ can always be found.

Looking next at condition (iii), the only odd primes dividing a, b, c are 3, 5, 7 respectively. As there is no solution in $\mathbb{F}_3$ to

$$5 - 7r^2 \equiv 0 \pmod 3,$$

we see that $\frac{5}{7}$ is not a quadratic residue modulo 3, thus there is no non-trivial solution in $\mathbb{Q}_3$. By the same argument, neither is $-\frac{3}{5}$ modulo 7 and so there is no non-trivial solution in $\mathbb{Q}_7$. However, as

$$-7 + 3 \cdot 2^2 \equiv 0 \pmod 5,$$

we do have that $\frac{7}{3}$ is a quadratic residue modulo 5, i.e. a non-trivial solution in $\mathbb{Q}_5$ does exist, but this alone is not sufficient for (ii) of Theorem 3.20 to be satisfied. We now know enough to conclude that the polynomial has no non-trivial rational solutions.

Despite this, as a, b, c are all odd, we proceed to look at condition (iv) of Theorem 3.21. Upon noticing that $a + b = 8$ is divisible by 4, this final condition is met and so there is a non-trivial solution in $\mathbb{Q}_2$. □

4 Analysis in $\mathbb{Q}_p$ and its applications

We conclude this paper with some analytic arguments. Specifically, we look at a group isomorphism arising from p-adic power series, referring once again to Gouvea [1].

4.1 Series

We begin this section with the discussion of series: recall, an infinite series over some metric space $(X,\, d)$ is an expression of the form

$$\sum_{n=0}^{\infty} a_n, \text{ for } a_n \in X.$$

Such a series is said to be convergent in X if and only if for some $l \in X$, the partial sums

$$S_k = \sum_{n=0}^{k} a_n \to l, \text{ as } k \to \infty.$$

Proposition 4.1. *In a non-Archimedean, complete normed space* $(X, ||\cdot||)$,

$$\sum_{n=0}^{\infty} a_n \textit{ converges } \Leftrightarrow \lim_{n\to\infty} ||a_n|| = 0.$$

Proof. ($\Rightarrow$) Suppose that $\sum_{n=0}^{\infty} a_n$ converges, i.e. that $S_k \to l$ as $k \to \infty$ for some $l \in X$. Clearly it is also true that $S_{k-1} \to l$ as $k \to \infty$ and by the triangle inequality

$$||a_k|| = ||S_k - S_{k-1}|| \leq ||S_k - l|| + ||S_{k+1} - l|| \to 0 \text{ as } k \to \infty.$$

($\Leftarrow$) Suppose that $||a_n|| \to 0$ as $n \to \infty$ and fix $k,\, m \in \mathbb{N}$ such that $k > m$.

$$||S_k - S_m|| = ||\sum_{n=m+1}^{k} a_n|| \leq \max\{||a_n||\}_{n=m+1}^{k}$$

using the non-Archimedean property, and taking the limit of the right hand side as both $k,\, m \to \infty$ we obtain that S_n defines a Cauchy sequence. As the space we are working in is complete, S_n converges to some limit. □

Remark. *The right implication holds in a general Archimedean metric space too.*

A special type of series is a power series, i.e. one of the form

$$\sum_{n=0}^{\infty} a_n x^n.$$

The values of x for which a power series converges is of great interest in analysis, and we summarise such values by defining the radius of convergence:

Definition 4.2. The radius of convergence is defined to be some $r \in \mathbb{R}_{\geq 0} \cup \{\infty\}$ such that the power series

$$\sum_{n=0}^{\infty} a_n x^n$$

converges for $||x|| < r$ and diverges for $||x|| > r$.

Analogously to the Archimedean case, the radius of convergence can be defined explicitly by the formula

$$r = \frac{1}{\lim_{n\to\infty} \sup \sqrt[n]{||a_n||}}.$$

We will now revert our attention to working over $\mathbb{Q}_p$, where we observe that the radius of convergence will be a power of p.

The natural logarithm is defined as

$$\log(1+x) = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} x^n$$

and can be considered as being defined on $\mathbb{R}$ or $\mathbb{Q}_p$. On the former, with respect to the Euclidean norm, the radius of convergence is well known as being 1. For clarity, we'll write $\log_p$ for the power series on $\mathbb{Q}_p$, and shall proceed to determine its radius of convergence, r_p.

As r_p is a power of p, an equivalent problem is finding

$$\begin{aligned} \mathrm{ord}_p(r_p) &= -\lim_{n\to\infty} \inf \frac{\mathrm{ord}_p(a_n)}{n} \\ &= \lim_{n\to\infty} \inf \frac{\mathrm{ord}_p(n)}{n} \\ &= 0, \end{aligned}$$

having used that

$$\mathrm{ord}_p(a_n) = \mathrm{ord}_p\left(\frac{(-1)^{n-1}}{n}\right) = -\mathrm{ord}_p(n).$$

So the radius of convergence is

$$r_p = p^{\mathrm{ord}_p(r_p)} = 1,$$

i.e. convergence occurs on the *open* unit disc about 1, with respect to $|\cdot|_p$. This is same as saying that $\log_p$ converges on $1 + p\mathbb{Z}_p$.

Over $\mathbb{R}$, a close companion to the natural logarithm is its inverse, the exponential, defined as

$$\exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!},$$

whose radius of convergence is ∞. Considering this over $\mathbb{Q}_p$, we denote it as $\exp_p$ whose radius of convergence is a lot less obvious. We argue as above:

$$\begin{aligned}\mathrm{ord}_p(r_p) &= -\lim_{n\to\infty}\inf\frac{\mathrm{ord}_p(a_n)}{n}\\ &= -\lim_{n\to\infty}\inf\frac{n-s_p(n)}{n(p-1)}\\ &= -\frac{1}{p-1},\end{aligned}$$

having used Lemma 1.16 to obtain that

$$\mathrm{ord}_p(a_n)=\mathrm{ord}_p\left(\frac{1}{n!}\right)=\frac{n-s_p(n)}{p-1}.$$

So

$$r_p=p^{\mathrm{ord}_p(r_p)}=p^{-\frac{1}{p-1}}.$$

As this value is dependent on p we will further consider this case-by-case.

- When $p=2$, we have that $r_2=\frac{1}{2}$, i.e. $\exp_2(x)$ converges when $|x|_2<\frac{1}{2}$ or $|x|_2\leq\frac{1}{4}$. This is the same as saying that $\exp_2$ converges on $4\mathbb{Z}_2$.

- When p is odd, it holds that

$$\frac{1}{p-1}<1\Rightarrow\frac{1}{p}<\frac{1}{p^{\frac{1}{p-1}}}.$$

  This tells us that $\exp_p(x)$ converges when $|x|_2<\frac{1}{p}$, or $\exp_p$ converges on $p\mathbb{Z}_p$. This is different to that of above.

We can summarise the above discussion more concisely upon introducing the parameter

$$q=\begin{cases}4, & p=2,\\ p, & p \text{ odd},\end{cases}$$

i.e. we have shown that $\exp_p$ converges on $q\mathbb{Z}_p$. Although we know that $\log_p$ converges on $1+p\mathbb{Z}_p$, from here on we restrict the domain of this function to $1+q\mathbb{Z}_p$ to obtain the desired results.

The next step in our analysis is to determine the image of these convergent functions.

We'll consider first the p-adic logarithm, and let $1+x$ lie in it's domain, i.e. $|x|_p\leq\frac{1}{q}$ or $\mathrm{ord}_p(x)>\frac{1}{p-1}$. For $n=1,\,2,\,\ldots$ the estimate

$$\begin{aligned}\mathrm{ord}_p\left(\frac{(-1)^{n-1}}{n}x^n\right)-\mathrm{ord}_p(x) &= (n-1)\mathrm{ord}_p(x)-\mathrm{ord}_p(n)\\ &> (n-1)\left(\frac{1}{p-1}-\frac{\mathrm{ord}_p(n)}{n-1}\right)\\ &\geq 0,\end{aligned}$$

gives rise to the corresponding result regarding the norm, i.e. for $n = 1, 2, \dots$

$$\left|\frac{(-1)^{n-1}}{n}x^n\right|_p < |x|_p.$$

Hence, using the non-Archimedean property of the norm

$$|\log_p(1+x)|_p = \left|\sum_{n=1}^{\infty}\frac{(-1)^{n-1}}{n}x^n\right|_p \leq \max_{n=1,\,2,\,\dots}\left|\frac{(-1)^{n-1}}{n}x^n\right|_p < |x|_p,$$

so $\log_p(1+x) \in q\mathbb{Z}_p$.

Now, considering the exponential: let x lie in its domain so that $|x|_p \leq \frac{1}{q}$. Checking case-by-case, i.e. for even and odd p, it holds that

$$|\exp_p(x) - 1|_p = \left|\sum_{n=1}^{\infty}\frac{x^n}{n!}\right|_p \leq \max_{n=1,\,2,\,\dots}\left|\frac{x^n}{n!}\right|_p < 1.$$

So, $\exp_p(x) \in 1 + q\mathbb{Z}_p$.

Additionally, closure of each of the domains follows noting that

$$|(1+x)(1+y) - 1|_p = |x + y + xy|_p \leq \max\{|x|_p,\, |y|_p,\, |x|_p|y|_p\},$$

so if $1+x$, $1+y$ are in the domain of the logarithm, then so is $(1+x)(1+y)$ and

$$|xy|_p = |x|_p|y|_p,$$

so if x, y are in the domain of the exponential, then so is xy. Thus by the usual power series arguments,

$$\log_p((1+x)(1+y)) = \log_p(1+x) + \log_p(1+y), \text{ and}$$
$$\exp_p(x+y) = \exp_p(x)\exp_p(y).$$

To summarise, we have established two homomorphisms given by

$$\log_p : 1 + q\mathbb{Z}_p \longrightarrow q\mathbb{Z}_p, \qquad \exp_p : q\mathbb{Z}_p \longrightarrow 1 + q\mathbb{Z}_p.$$

A natural question is whether these functions still provide inverses to one another in $\mathbb{Q}_p$. Once again, the usual power series analysis confirms that on the domains stated above, the two compositions provide identity maps.

We have shown that:

Theorem 4.3. *For q as defined above,*

$$1 + q\mathbb{Z}_p \cong q\mathbb{Z}_p,$$

i.e. the p-adic logarithm and exponential provide a a group isomorphism between the multiplicative group $1 + q\mathbb{Z}_p$ and the additive group $q\mathbb{Z}_p$.

Remark. *Further to this, there is a clear isomorphism between the additive groups $\mathbb{Z}_p$ and $p^k\mathbb{Z}_p$ for $k \geq 0$. Thus we may note too that*

$$\mathbb{Z}_p \cong q\mathbb{Z}_p \cong 1 + q\mathbb{Z}_p.$$

Upon having made these observations, we're now able to formulate the following result concerning the structure of the unit group of the p-adic integers.

Theorem 4.4.

$$\mathbb{Z}_p^\times \cong \begin{cases} \mathbb{Z}_2 \times \mathbb{Z}/2\mathbb{Z}, & p = 2, \\ \mathbb{Z}_p \times \mathbb{Z}/(p-1)\mathbb{Z}, & p \text{ odd.} \end{cases}$$

Proof. This argument will rely on the short exact sequence

$$0 \longrightarrow 1 + q\mathbb{Z}_p \xrightarrow{f} \mathbb{Z}_p^\times \xrightarrow{g} (\mathbb{Z}/q\mathbb{Z})^\times \longrightarrow 0,$$

where f denotes the inclusion map and g the reduction modulo q. If we can construct a homomorphism $h : (\mathbb{Z}/q\mathbb{Z})^\times \longrightarrow \mathbb{Z}_p^\times$ such that $g \circ h = id_{(\mathbb{Z}/q\mathbb{Z})^\times}$, then we have split the short exact sequence and obtain

$$\begin{aligned} \mathbb{Z}_p^\times &\cong 1 + q\mathbb{Z}_p \times (\mathbb{Z}/q\mathbb{Z})^\times \\ &\cong \mathbb{Z}_p \times (\mathbb{Z}/q\mathbb{Z})^\times. \end{aligned}$$

As q varies with the parity of the prime p, we proceed with a case-by-case approach. First, let $p = 2$ so $q = 4$. Observe that $(\mathbb{Z}/4\mathbb{Z})^\times = \{1, 3\}$ and $\pm 1 \in \mathbb{Z}_2^\times$, so $\mathbb{Z}_2^\times$ contains both 2nd roots of unity. This motivates us to define $\omega : (\mathbb{Z}/4\mathbb{Z})^\times \longrightarrow \mathbb{Z}_2^\times$ so that $\omega(a) \in \mathbb{Z}_2^\times$ is the unique 2nd root of unity such that $\omega(a) \equiv a \pmod 4$. It can be easily shown that $g \circ \omega = id_{(\mathbb{Z}/4\mathbb{Z})^\times}$ and so it follows that

$$\mathbb{Z}_2^\times \cong \mathbb{Z}_2 \times (\mathbb{Z}/4\mathbb{Z})^\times \cong \mathbb{Z}_2 \times \mathbb{Z}/2\mathbb{Z}.$$

Now let p be odd, so that $q = p$. We find $h : (\mathbb{Z}/p\mathbb{Z})^\times \longrightarrow \mathbb{Z}_p^\times$ based on the same principle as before. Utilising Corollary 3.10 of Hensel's lemma, it's possible for us to map $a \in (\mathbb{Z}/p\mathbb{Z})^\times$ to the unique $(p-1)$th root of unity, $\omega(a) \in \mathbb{Z}_p$ such that $\omega(a) \equiv a \pmod p$. This construction again makes it clear that $g \circ \omega$ provides the identity on $(\mathbb{Z}/p\mathbb{Z})^\times$, thus

$$\mathbb{Z}_p^\times \cong \mathbb{Z}_p \times (\mathbb{Z}/p\mathbb{Z})^\times \cong \mathbb{Z}_p \times \mathbb{Z}/(p-1)\mathbb{Z}. \qquad \square$$

Remark. *The homomorphism of multiplicative groups, $\omega : (\mathbb{Z}/q\mathbb{Z})^\times \longrightarrow \mathbb{Z}_p^\times$, as defined above is called the Teichmüller character.*

Corollary 4.5. *The only roots of unity in $\mathbb{Q}_2$ are the second roots of unity, ± 1.*

Proof. As $\mathbb{Z}_2$ is torsion-free, the only roots of unity in $\mathbb{Z}_2 \times \mathbb{Z}/2\mathbb{Z}$ are $(\pm 1, 0)$. □

4.1.1 Application: quadratic extensions of $\mathbb{Q}_p$

A well-known fact is that the quadratic field extensions of $\mathbb{Q}$ are given by $\mathbb{Q}(\sqrt{d})$ for square-free $d \in \mathbb{Z}$. More generally, it is true that the quadratic extensions of a field $\mathbb{K}$ are in bijection with the quotient

$$\mathbb{K}^\times/(\mathbb{K}^\times)^2.$$

Thus, classifying the quadratic extensions of $\mathbb{Q}_p$ is equivalent to understanding the quotient

$$\mathbb{Q}_p^\times/(\mathbb{Q}_p^\times)^2.$$

A first observation is that

$$\mathbb{Q}_p^\times \cong \mathbb{Z}_p^\times \times \langle p \rangle,$$

and so quotienting by the subgroup of squares gives

$$\mathbb{Q}_p^\times/(\mathbb{Q}_p^\times)^2 \cong \mathbb{Z}_p^\times/(\mathbb{Z}_p^\times)^2 \times \langle p \rangle/\langle p^2 \rangle,$$

i.e. we have reduced the problem to studying $\mathbb{Z}_p^\times/(\mathbb{Z}_p^\times)^2$. As we saw before, the structure of $\mathbb{Z}_p^\times$ varies with p, so it is necessary to take a case by case approach.

When $p = 2$, we saw that $\mathbb{Z}_2^\times \cong \mathbb{Z}_2 \times \mathbb{Z}/2\mathbb{Z}$ or more explicitly that

$$\begin{aligned}\mathbb{Z}_2^\times &\cong (1+4\mathbb{Z}_2)^\times \times (\mathbb{Z}/4\mathbb{Z})^\times \\ &\Rightarrow \mathbb{Z}_2^\times/(\mathbb{Z}_2^\times)^2 \cong (1+4\mathbb{Z}_2)^\times/((1+4\mathbb{Z}_2)^\times)^2 \times (\mathbb{Z}/4\mathbb{Z})^\times/((\mathbb{Z}/4\mathbb{Z})^\times)^2.\end{aligned}$$

Choosing representatives of these factors will allow use to formulate a complete list of the quadratic extensions of $\mathbb{Q}_2$. The second factor is easy: an obvious choice for a representative is -1. The first factor however is significantly harder. We utilise the isomorphism of $(1+4\mathbb{Z}_2)^\times$ with the additive group $4\mathbb{Z}_2^\times$, to spot that 4 is a representative of the latter quotiented by its subgroup of squares. So the image of 4 under $\exp_2$ is the representative we want. However, this is not a particularly nice number. If we reduce modulo 4, we get 1 which is clearly a square. Instead, notice that

$$\exp_2(4) = 1 + 4 + \frac{4^2}{2!} + \frac{4^3}{3!} + \ldots \equiv -3 \pmod 8,$$

we may take this to be our representative as -3 is not a square. Now, from the expression

$$\mathbb{Q}_2^\times/(\mathbb{Q}_2^\times)^2 \cong (1+4\mathbb{Z}_2)^\times/((1+4\mathbb{Z}_2)^\times)^2 \times (\mathbb{Z}/4\mathbb{Z})^\times/((\mathbb{Z}/4\mathbb{Z})^\times)^2 \times \langle 2 \rangle/\langle 4 \rangle$$

we see that the set of representatives will be of order 8, given by -3, -1, 2 and multiples thereof. Hence, there are 7 quadratic field extensions, being

$$\mathbb{Q}_2(\sqrt{-3}),\ \mathbb{Q}_2(\sqrt{-1}),\ \mathbb{Q}_2(\sqrt{2}),\ \mathbb{Q}_2(\sqrt{3}),\ \mathbb{Q}_2(\sqrt{-6}),\ \mathbb{Q}_2(\sqrt{-2}),\ \text{and } \mathbb{Q}_2(\sqrt{6}).$$

Now let p be an odd prime. Here we have the isomorphism

$$\mathbb{Z}_p^\times \cong \mathbb{Z}_p \times \mathbb{Z}/(p-1)\mathbb{Z} \cong \mathbb{Z}_p \times \mathbb{F}_p^\times.$$

Thus the quotient is question is merely,

$$\mathbb{Z}_p^\times/(\mathbb{Z}_p^\times)^2 \cong \mathbb{F}_p^\times/(\mathbb{F}_p^\times)^2 \Rightarrow \mathbb{Q}_p^\times/(\mathbb{Q}_p^\times)^2 \cong \mathbb{F}_p^\times/(\mathbb{F}_p^\times)^2 \times \langle p\rangle/\langle p^2\rangle.$$

From this, it is clear that the set of representatives will be of order 4 and given by $\{1,\ n,\ p,\ np\}$ where n is a non-quadratic residue modulo p. Hence, there are 3 quadratic field extensions:

$$\mathbb{Q}_p(\sqrt{n}),\ \mathbb{Q}_p(\sqrt{p}),\ \text{and}\ \mathbb{Q}_p(\sqrt{np}).$$

Example 4.6. Let $p \equiv 3 \pmod 4$. The first supplement to the law of quadratic reciprocity suggests an obvious choice for n, i.e. $n = -1$. Thus the quadratic field extensions are

$$\mathbb{Q}_p(\sqrt{-1}),\ \mathbb{Q}_p(\sqrt{p}),\ \text{and}\ \mathbb{Q}_p(\sqrt{-p}). \qquad \square$$

Example 4.7. Let $p \equiv 1 \pmod 4$. Additionally, when $p \equiv 5 \pmod 8$, the second supplement to the law of quadratic reciprocity suggests $n = 2$. Thus the quadratic field extensions are

$$\mathbb{Q}_p(\sqrt{2}),\ \mathbb{Q}_p(\sqrt{p}),\ \text{and}\ \mathbb{Q}_p(\sqrt{2p}).$$

But what about when $p \equiv 1 \pmod 8$? Determining that there is no such n which works for all p is a finite problem. The smallest p of this kind is 17, whose non-quadratic residues are $\{3,\ 5,\ 6,\ 7,\ 10,\ 11,\ 12,\ 14\}$. For each number in this list, a prime $q \equiv 1 \pmod 8$ can be found with respect to which the number is a quadratic residue, hence confirming the claim. For instance, $10 \equiv 16^2 \equiv 1 \pmod{41}$. □

4.2 The Bernoulli numbers and p-adic analyticity

Advancing from the discussion of series in $\mathbb{Q}_p$, we now discuss what it means for a function to be p-adically analytic. This subsection summarises some of Washington's work in [8].

First, we introduce the Bernoulli numbers which are objects of analysis that can be defined in various ways, two of which we will make use of:

Definition 4.8. The Bernoulli numbers B_n satisfy,

$$\frac{t}{e^t - 1} = \sum_{n=0}^{\infty} \frac{B_n t^n}{n!}$$

i.e. they appear in the exponential generating function.

Definition 4.9. Recursively, for $m \geq 0$

$$B_n = \delta_{n,0} - \sum_{k=0}^{n-1} \binom{n}{k} \frac{B_k}{n-k+1}.$$

In practice, the second definition is easier to work with and we use this to determine the first few Bernoulli numbers:

$$B_0 = 1,\ B_1 = -\frac{1}{2},\ B_2 = \frac{1}{6},\ B_3 = 0,\ B_4 = -\frac{1}{30},\ B_5 = 0,\ B_6 = \frac{1}{42},\ \ldots$$

Observe that:

Proposition 4.10. *For odd $n \geq 3$, $B_n = 0$.*

Proof. Notice that we may write

$$\sum_{n=0}^{\infty} \frac{B_n t^n}{n!} = \frac{t}{e^t - 1} = \frac{t}{2} \cdot \frac{(e^t + 1) - (e^t - 1)}{e^t - 1}.$$

Thus,

$$1 + \sum_{n=2}^{\infty} \frac{B_n t^n}{n!} = \frac{t}{2} \cdot \frac{(e^t + 1)}{e^t - 1},$$

where the latter is an even function. □

We also begin to see the following which shall be justified later on:

Proposition 4.11. *For $n \neq 0$, $B_n \leq 0 \Leftrightarrow 4|n$.*

But why do we call these numbers objects of analysis?

Example 4.12. The Bernoulli numbers appear in the Taylor series expansions of the tangent and hyperbolic tangent functions:

$$\tan(x) = \sum_{n=1}^{\infty} \frac{(-1)^{n-1} 2^{2n} (2^{2n} - 1) B_{2n} x^{2n-1}}{(2n)!},$$

$$\tanh(x) = \sum_{n=1}^{\infty} \frac{2^{2n} (2^{2n} - 1) B_{2n} x^{2n-1}}{(2n)!},$$

for $|x| < \frac{\pi}{2}$. □

Example 4.13. The Bernoulli numbers appear in the expression for the sum of powers of the first positive integers. Write

$$S_m(n) = \sum_{k=0}^{n-1} k^m.$$

Noting the power series expansion of the exponential,

$$e^{kx} = \sum_{m=0}^{\infty} k^m \frac{x^m}{m!}$$

and summing these functions for $k = 0, 1, \ldots n - 1$, we obtain

$$\sum_{m=0}^{\infty} S_m(n) \frac{x^m}{m!} = \sum_{k=0}^{n-1} e^{kx} = \frac{e^{nx} - 1}{e^x - 1} = \sum_{k=1}^{\infty} n^k \frac{x^{k-1}}{k!} \sum_{m=0}^{\infty} \frac{B_m t^m}{m!},$$

where the last equality uses the generating function definition of B_n. Comparing coefficients of the powers of x gives the desired result that

$$S_m(n) = \sum_{k=0}^{m} \binom{m}{k} B_{m-k} \frac{n^{k+1}}{k+1}.$$ □

Lemma 4.14. *For $n \geq 0$, $pB_n \in \mathbb{Z}_p$. Moreover, if $n \geq 2$ is even then*

$$pB_n \equiv S_n(p) \ (mod\ p).$$

Proof. We prove the first statement by induction. First notice, $pB_0 = p$ which is certainly a p-adic integer. The formula established in the above example gives that

$$pB_n = S_n(p) - \sum_{k=1}^{n} \binom{n}{k} pB_{n-k} \frac{p^k}{k+1}.$$

Let's assume that $pB_0,\ pB_1,\ \ldots,\ pB_{n-1} \in \mathbb{Z}_p$. As $S_n(p) \in \mathbb{Z}$ and for $k = 1,\ 2,\ \ldots n$, $\binom{n}{k} \in \mathbb{Z}$, it follows that $pB_n \in \mathbb{Z}_p$ upon noticing that $k+1 \leq p^k \Rightarrow \frac{p^k}{k+1} \in \mathbb{Z}_p$ for $k \geq 1$. The stronger statement that $k+1 < p^k$ for $k \geq 2$ implies that

$$pB_n \equiv S_n(p) - npB_{n-1}\frac{p}{2} \pmod{p}.$$

So when $n \geq 1$ is even we obtain the desired result. □

Lemma 4.15.

$$S_n(p) \equiv \begin{cases} 0 \ (mod\ p) & p-1 \nmid n, \\ -1 \ (mod\ p) & p-1 \mid n. \end{cases}$$

Proof. Let g be a primitive root modulo p. Then

$$\{1,\ 2,\ \ldots, p-1\} = \{1,\ g,\ \ldots, g^{p-2}\},$$

so

$$S_n(p) \equiv 0^n + 1^n + g^n + \ldots + g^{n(p-2)} \pmod{p}.$$

Thus we may determine that

$$\begin{aligned}(g^n - 1)S_n(p) &\equiv 0^n + g^n + g^{2n} + \ldots + g^{n(p-1)} - (0^n + 1^n + g^n + \ldots + g^{n(p-2)}) \\ &\equiv g^{n(p-1)} - 1 \equiv 0 \pmod{p}.\end{aligned}$$

Now, if $p-1 \nmid n$, then $g^n - 1 \not\equiv 0 \pmod{p}$, so $S_n(p) \equiv 0 \pmod{p}$. If instead $p-1 \mid n$, then $g^n \equiv 1 \pmod{p} \Rightarrow S_n(p) \equiv p-1 \pmod{p}$. □

These preliminary results (from [9]) allow us to prove our first major result concerning the Bernoulli numbers.

Theorem 4.16 (Von Staudt–Clausen)**.** *If $n \geq 2$ is even, then*

$$B_n + \sum_{(p-1)|n} \frac{1}{p} \in \mathbb{Z}.$$

Proof. Let $n \geq 2$ be even and q be a prime. If $q - 1 \nmid n$, then $qB_n \equiv 0 \pmod{q}$, so $B_n \in \mathbb{Z}_q$ and clearly it follows that

$$B_n + \sum_{(p-1)|n} \frac{1}{p} \in \mathbb{Z}_q.$$

If $q - 1 \mid n$, then $qB_n \equiv -1 \pmod{q}$. We find too that,

$$B_n + \sum_{(p-1)|n} \frac{1}{p} \in \mathbb{Z} = B_n + \frac{1}{q} + \sum_{\substack{(p-1)|n \\ p \neq q}} \frac{1}{p} = \frac{qB_n + 1}{q} + \sum_{\substack{(p-1)|n \\ p \neq q}} \frac{1}{p} \in \mathbb{Z}_q.$$

Therefore, the quantity in question lies in $\mathbb{Z}_q$ for all primes q, hence it must be in $\mathbb{Z}$. □

The next theorem we will work towards proving is one of Kummer:

Theorem 4.17 (Kummer's congruences)**.** *Let m, $n \geq 1$ be even and such that for some a, $m \equiv n$ (mod $(p-1)p^a$). If also $n \not\equiv 0$ (mod $p - 1$), then*

$$(1 - p^{m-1})\frac{B_m}{m} \equiv (1 - p^{n-1})\frac{B_n}{n} \ (mod\ p^{a+1}).$$

In order to do so, we must work in a much more general setting. The material presented from here on is based on that of Washington in [8, Sections 3 - 5].

Definition 4.18. A Dirichlet character χ of modulus F is a multiplicative homomorphism

$$\chi : (\mathbb{Z}/F\mathbb{Z})^\times \longrightarrow \mathbb{C}^\times.$$

The conductor of χ, denoted f or f_χ, is the minimal possible conductor, i.e. if $F|G$, then $\chi : (\mathbb{Z}/G\mathbb{Z})^\times \longrightarrow (\mathbb{Z}/F\mathbb{Z})^\times \longrightarrow \mathbb{C}^\times$ is an induced homomorphism.

Remark. *We may extend a Dirichlet character to the whole of $\mathbb{Z}$ by defining*

$$\chi(a) = 0 \ \textit{if}\ (a, f) \neq 1.$$

Example 4.19. The Teichmüller character, ω, defined previously is a Dirichlet character of conductor q. □

Definition 4.20. The Dirichlet L-function is

$$L(s, \chi) := \sum_{n \geq 1} \frac{\chi(n)}{n^s},$$

for $s \in \mathbb{C}$ with $Re(s) > 1$ and χ a Dirichlet character.

Although the notation here omits the modulus associated to the Dirichlet character, this value does in fact matter as is seen below.

Example 4.21. Consider the trivial character with modulus $F = 1$,

$$\begin{aligned}\chi :\ \mathbb{Z} &\longrightarrow \mathbb{C}^\times,\\ a &\longmapsto 1.\end{aligned}$$

Here we see that

$$L(s,\chi) = \zeta(s),$$

the usual Riemann Zeta function. □

Example 4.22. Now let $F = p$, thus

$$\begin{aligned}\chi :\ (\mathbb{Z}/p\mathbb{Z})^\times &\longrightarrow \mathbb{C}^\times,\\ a &\longmapsto 1.\end{aligned}$$

Using Euler's product formula for the Riemann Zeta function, we obtain that

$$L(s,\chi) = \sum_{\substack{n\geq 1\\ p\nmid n}} \frac{1}{n^s} = (1-p^{-s})\zeta(s).$$ □

The Riemann Zeta function is closely related to the Bernoulli numbers. In particular, for integers $n \geq 1$:

- $\zeta(1-n) = (-1)^{n+1}\frac{B_n}{n}$,
- $\zeta(2n) = (-1)^{n+1}\frac{(2\pi)^{2n}B_n}{2(2n)!}$.

Combining this second identity with the well-known fact that $\zeta(s) > 0$ on the positive real numbers, we obtain that the sign of B_{2n} alternates, proving Proposition 4.11.

The Kubota-Leopoldt p-adic zeta function, $\zeta_p(s)$, is a one of a p-adic variable which interpolates values of $\zeta(s)$ at negative integers in the sense that

$$\zeta_p(1-n) = (1-p^{n-1})\zeta(1-n), \text{ for } n \equiv 0 \pmod{p-1}.$$

The first identity listed above combined with Kummer's congruences tells us that $\zeta_p(s)$ is uniformly continuous when restricted to negative integers in a fixed residue class modulo $p-1$.

> A slight digression motivates why there is so much interest surrounding this function.
>
> **Definition 4.23.** A prime p is called irregular if it divides the class number of $\mathbb{Q}(\xi_p)$, the pth cyclotomic field (obtained by adjoining a primitive pth root of unity ξ_p to the rational numbers), and regular if not.
>
> A first application of this notion is in proving Fermat's last theorem where the exponent is a regular prime. This proof was accessible in the 19th century, while the full proof was not established until the 1990s. Another is the following:

Theorem 4.24 (Kummer's criterion)**.** *A prime p is regular if and only if p does not divide the numerator of any B_n for $n = 2, 3, \dots, p-3$.*

Example 4.25. 37 is irregular: the numerator of B_{32} is $-7709321041217 = 37 \times 683 \times 305065927$. In fact, 37 can be shown to be the least irregular prime. □

Using this fact, it holds that the smallest p for which there exists $s \in \mathbb{Z}_p^\times$ with $\zeta_p(s) = 0$ is $p = 37$.

Definition 4.26. The generalized Bernoulli numbers are defined by

$$\sum_{a=1}^{f} \frac{\chi(a)te^{at}}{e^{ft}-1} = \sum_{n=0}^{\infty} \frac{B_{n,\chi}t^n}{n!},$$

where χ is a Dirichlet character with conductor f.

Returning to the standard Bernoulli numbers we also define:

Definition 4.27. The Bernoulli polynomials for $n \geq 0$ satisfy,

$$\frac{te^{xt}}{e^t-1} = \sum_{n=0}^{\infty} \frac{B_n(x)t^n}{n!},$$

or explicitly are defined as

$$B_n(x) = \sum_{k=0}^{n} \binom{n}{k} (B_k) x^{n-k}.$$

Remark. *Note that $B_n(0) = B_n$.*

From here on the notation $B_n(x)$ will be used for the n-th Bernoulli polynomial evaluated at x and $(B_n)(x)$ will denote the n-th Bernoulli number multiplied by x.

Proposition 4.28.

$$B_{n,\chi} = f^{n-1} \sum_{a=1}^{f} \chi(a) B_n\left(\frac{a}{f}\right)$$

Proof. Multiplying the right-hand side by $\frac{t^n}{n!}$ and summing over n we obtain

$$\begin{aligned}\sum_{n=0}^{\infty} f^{n-1} \sum_{a=1}^{f} \chi(a) B_n\left(\frac{a}{f}\right)\frac{t^n}{n!} &= \sum_{a=1}^{f} \chi(a)\frac{1}{f} \sum_{n=0}^{\infty} B_n\left(\frac{a}{f}\right)\frac{(ft)^n}{n!} \\ &= \sum_{a=1}^{f} \frac{\chi(a)te^{at}}{e^{ft}-1} \\ &= \sum_{n=0}^{\infty} \frac{B_{n,\chi}t^n}{n!},\end{aligned}$$

and so the result follows. □

Example 4.29. Let χ be the trivial Dirichlet character with conductor 1, then

$$B_{n,1} = B_n(1) = \sum_{k=0}^{n} \binom{n}{k} B_k.$$

Clearly, $B_{0,1} = 1$ and $B_{1,1} = \frac{1}{2}$. We claim that for $n > 1$, $B_{n,1} = B_n$ and proceed to prove this by induction.

When $n = 2$, clearly we have that $B_{2,1} = B_2 = \frac{1}{6}$. Now suppose the hypothesis is true, and consider the following:

$$\begin{aligned}\sum_{k=0}^{n+1} \binom{n+1}{k} B_k &= \sum_{k=0}^{n} \binom{n+1}{k} B_k + B_{n+1} \\ &= (n+1) \sum_{k=0}^{n} \binom{n}{k} \frac{B_k}{n+1-k} + B_{n+1} \\ &= (n+1) \sum_{k=0}^{n-1} \binom{n}{k} \frac{B_k}{n+1-k} + B_{n+1} + (n+1)B_n \\ &= -(n+1)B_n + B_{n+1} + (n+1)B_n = B_{n+1}.\end{aligned}$$

Thus we are done. □

We now take the opportunity to set up the appropriate field in which the remaining discussion will take place. As usual, let $\overline{\mathbb{Q}}_p$ denote the algebraic closure of $\mathbb{Q}_p$.

Theorem 4.30. *$\overline{\mathbb{Q}}_p$ is not complete.*

Studying analysis in a complete field is of great convenience to us, so we let $\mathbb{C}_p$ denote the completion of $\overline{\mathbb{Q}}_p$ with respect to $|\cdot|_p$.

Theorem 4.31. *$\mathbb{C}_p$ is algebraically closed.*

The proof of the Kummer congruences will revolve around the following function of a complex variable $s \in \mathbb{C}_p$ and a Dirichlet character χ of conductor f:

$$L_p(s,\chi) = \sum_{\substack{a=1 \\ p\nmid a}}^{F} \chi(a) H_p(s,a,F),$$

where F is a multiple of f and q and

$$H_p(s,a,F) = \frac{1}{s-1}\frac{1}{F}\{\omega(a)^{-1}a\}^{1-s} \sum_{j=0}^{\infty} \binom{1-s}{j} (B_j) \left(\frac{F}{a}\right)^j.$$

From here on we restrict our attention to the case where p is odd for simplicity of notation, i.e. $q = p$. When $p = 2$ and $q = 4$, the argument will follow very similarly.

Proposition 4.32. *$H_p(s, a, F)$ is analytic except for a simple pole at $s = 1$.*

To prove this we require a stronger result than that of Mahler which says "any continuous function

$$f : \mathbb{Z}_p \longrightarrow \mathbb{Q}_p$$

may be written uniquely as

$$f(x) = \sum_{n=0}^{\infty} a_n \binom{x}{n},$$

with $a_n \to 0$ as $n \to \infty$."

Theorem 4.33. *Let $r < p^{-\frac{1}{p-1}}$ and*

$$f(x) = \sum_{n=0}^{\infty} a_n \binom{x}{n}.$$

If there exists some $M > 0$ such that $|a_n| \leq Mr^n$ for all n, then f is analytic with radius of convergence at least $(rp^{\frac{1}{p-1}})^{-1}$.

Remark. *Notice that $(rp^{\frac{1}{p-1}})^{-1} > 1$ so f is analytic on $\mathbb{Z}_p$.*

We can now prove Proposition 4.32:

Proof. First consider the sum

$$\sum_{j=0}^{\infty} \binom{s}{j} (B_j) \left(\frac{F}{a}\right)^j, \text{ i.e. } a_j = (B_j) \left(\frac{F}{a}\right)^j$$

in the notation of Theorem 4.33. As we have taken F to be a multiple of p, and $0 < a < F$ to be such that $p \nmid a$, it holds that

$$|a_j|_p \leq |(B_j)|_p \cdot \frac{1}{p^j} \leq p \cdot \frac{1}{p^j},$$

where the last inequality follows from Lemma 4.14. We apply Theorem 4.33 with the parameters $r = \frac{1}{p}$ and $M = p$ to obtain that the series above is analytic on $D = \{s \in \mathbb{C}_p \mid |s| < p^{\frac{p-2}{p-1}}\}$. Notice that as $p^{\frac{p-2}{p-1}} > 1$, we have too that $D = \{s \in \mathbb{C}_p \mid |1 - s| < p^{\frac{p-2}{p-1}}\}$, hence analyticity of

$$\sum_{j=0}^{\infty} \binom{1-s}{j} (B_j) \left(\frac{F}{a}\right)^j.$$

The remaining term to consider is

$$\{\omega(a)^{-1}a\}^s = \sum_{j=0}^{\infty} \binom{s}{j} (\omega(a)^{-1}a - 1)^j, \text{ i.e. } a_j = (\omega(a)^{-1}a - 1)^j.$$

We have

$$\omega(a) \equiv a \pmod{p} \Rightarrow \omega(a)^{-1}a \equiv 1 \pmod{p} \Rightarrow |a_j|_p \leq \frac{1}{p^j},$$

so again applying Theorem 4.33 this time with $r = \frac{1}{p}$ and $M = 1$ we obtain that this series in analytic on D, and correspondingly that $\{\omega(a)^{-1}a\}^{1-s}$ is analytic on D. The definition of $H_p(s, a, F)$ as a product of these series makes it clear that it is analytic with a simple pole at $s = 1$. □

Corollary 4.34. *When χ is non-trivial, $L_p(s, \chi)$ is analytic.*

Proposition 4.35. *For $n \geq 1$,*

$$H_p(1-n, a, F) = -\frac{\omega^{-n}(a)F^{n-1}}{n} B_n\left(\frac{a}{F}\right).$$

Proof. A simple manipulation using the Bernoulli polynomial gives

$$\begin{aligned} H_p(1-n, a, F) &= -\frac{1}{nF}\{\omega(a)^{-1}a\}^n \sum_{j=0}^{n} \binom{n}{j} (B_j) \left(\frac{F}{a}\right)^j \\ &= -\frac{1}{nF}\{\omega(a)^{-1}a\}^n \frac{F^n}{a^n} B_n\left(\frac{a}{F}\right) \\ &= -\frac{\omega^{-n}(a)F^{n-1}}{n} B_n\left(\frac{a}{F}\right). \end{aligned}$$

□

From now on we will also focus exclusively on $\chi = \omega^n$ for some $n \in \mathbb{N}$, so that $f = q = p$ and we may let $F = p$.

Corollary 4.36. *For $n > 1$,*

$$L_p(1-n, \omega^n) = -(1-p^{n-1})\frac{B_n}{n}.$$

Proof. Beginning with the definition,

$$\begin{aligned} L_p(1-n, \omega^n) &= \sum_{a=1}^{p-1} \omega^n(a) H_p(1-n, a, p) \\ &= -\sum_{a=1}^{p-1} \omega^n(a) \frac{\omega^{-n}(a)p^{n-1}}{n} B_n\left(\frac{a}{p}\right), \\ &= -\sum_{a=1}^{p-1} \frac{p^{n-1}}{n} B_n\left(\frac{a}{p}\right), \quad \text{as } \omega^n(a)\omega^{-n}(a) = \omega^n\omega^{-n}(a) = 1 \\ &= -\frac{B_{n,1} - p^{n-1}B_n(1)}{n}, \\ &= -(1-p^{n-1})\frac{B_n}{n}, \quad \text{as } B_{n,1} = B_n(1) = B_n \text{ for } n > 1. \end{aligned}$$

□

A finally necessary result is:

Theorem 4.37. *There exists a power series expansion of the form*

$$L_p(s,\omega^n) = a_0 + a_1(s-1) + a_2(s-2)^2 + \ldots,$$

where $|a_0| \leq 1$ *and for all* $n \geq 1$, $p|a_n$.

Proof. Writing $L_p(s,\omega^n)$ as

$$\frac{1}{s-1}\sum_{a=1}^{p-1}\omega^n(a)\left(\sum_{j=0}^{\infty}\frac{1}{j!}(1-s)^j\log_p\{\omega(a)^{-1}a\}^j\right)\left((1-s)\cdots(1-s-(j-1))\frac{B_j}{j!}\frac{p^{j-1}}{a^j}\right),$$

it makes sense to look at the two bracketed power series in turn. The coefficient of $(1-s)^j$ in the former is clearly

$$\frac{1}{j!}\log_p\{\omega(a)^{-1}a\}^j.$$

When $j \leq 2$,

$$\left|\frac{1}{j!}\log_p\{\omega(a)^{-1}a\}^j\right|_p < p^{\frac{j}{p-1}}\cdot\frac{1}{p^j} = \frac{1}{p^{j\frac{p-2}{p-1}}} \leq \frac{1}{p}$$

using that $\log_p\{\omega(a)^{-1}a\} \in p\mathbb{Z}_p$ and the remark following Lemma 1.16. Thus, the contribution of this series left to consider is

$$1 + (1-s)\log_p\{\omega(a)^{-1}a\}.$$

Upon noticing that $(1-s)(1-s-1)\cdots(1-s-(j-1))$ is a polynomial in $(1-s)$ with integer coefficients of degree j, it is enough to just consider the coefficient

$$\frac{B_j}{j!}\frac{p^{j-1}}{a^j}.$$

Again, when $j \geq 2$,

$$\left|\frac{xfB_j}{j!}\frac{p^{j-1}}{a^j}\right|_p < p\cdot p^{\frac{j}{p-1}}\cdot\frac{1}{p^{j-1}} = \frac{1}{p^{j\frac{p-2}{p-1}}} \leq \frac{1}{p},$$

(this time having also made use of Lemma 4.14), so the contribution of this series left to consider is

$$\frac{1}{p} - \frac{1-s}{2a}.$$

Hence,

$$L_p(s,\omega^n) \equiv \frac{1}{s-1}\sum_{a=1}^{p-1}\omega^n(a)(1+(1-s)\log_p\{\omega(a)^{-1}a\})\left(\frac{1}{p} - \frac{1-s}{2a}\right) \pmod{p}.$$

Firstly,

$$a_0 \equiv -\sum_{a=1}^{p-1} \omega^n(a) \left(\frac{\log_p\{\omega(a)^{-1}a\}}{p} - \frac{1}{2a} \right) \pmod{p}$$

$$\Rightarrow |a_0|_p \leq \max \left\{ \omega^n(a) \left(\frac{\log_p\{\omega(a)^{-1}a\}}{p} - \frac{1}{2a} \right) : a = 1, \ldots, p-1, \frac{1}{p} \right\}.$$

Clearly $p \nmid a \Rightarrow \frac{1}{2a} \in \mathbb{Z}_p$ and along with $\omega^n(a)$, $\frac{\log_p\{\omega(a)^{-1}a\}}{p} \in \mathbb{Z}_p$ for each a, we obtain that $|a_0|_p \leq 1$. Finally,

$$a_1 \equiv \sum_{a=1}^{p-1} \omega^n(a) \frac{\log_p\{\omega(a)^{-1}a\}}{2a} \pmod{p}$$

$$\Rightarrow |a_1|_p \leq \max \left\{ \omega^n(a) \frac{\log_p\{\omega(a)^{-1}a\}}{2a} : a = 1, \ldots, p-1, \frac{1}{p} \right\}.$$

We estimate

$$\operatorname{ord}_p \left(\frac{\log_p\{\omega(a)^{-1}a\}}{2a} \right) = \operatorname{ord}_p(\log_p\{\omega(a)^{-1}a\}) - \operatorname{ord}_p(2a) = \operatorname{ord}_p(\log_p\{\omega(a)^{-1}a\}) \geq 1,$$

so $|a_1|_p \leq \frac{1}{p}$, i.e. $p \mid a_1$, and we're done. □

Having established these preliminaries, the proof of Kummer's theorem follows rather concisely.

Theorem 4.17 (Kummer's congruences)**.** *Let m, $n \geq 1$ be even and such that for some a, $m \equiv n$ (mod $(p-1)p^a$). If also $n \not\equiv 0$ (mod $p-1$), then*

$$(1 - p^{m-1}) \frac{B_m}{m} \equiv (1 - p^{n-1}) \frac{B_n}{n} \ \textit{(mod } p^{a+1}).$$

Proof. Using first that $m \equiv n$ (mod $(p-1)p^a$), the character ω^{m-n} is trivial, i.e. $\omega^m = \omega^n$. Thus, it holds that

$$L_p(s, \omega^m) = L_p(s, \omega^n).$$

By Theorem 4.37, we may write

$$\begin{aligned} L_p(1-m, \omega^m) &= a_0 - a_1 m + a_2 m^2 - \ldots \\ &\equiv a_0 - a_1 n + a_2 n^2 - \ldots \pmod{p^{a+1}} \\ &= L_p(1-n, \omega^n), \end{aligned}$$

where the congruence follows from $m \equiv n$ (mod $(p-1)p^a$) and $p|a_i$ for $i \geq 1$. Finally, by Corollary 4.36, we obtain that

$$(1 - p^{m-1}) \frac{B_m}{m} \equiv (1 - p^{n-1}) \frac{B_n}{n} \pmod{p^{a+1}}.$$ □

Acknowledgments

I would like thank the EPSRC for supporting this project. I would also like to thank my supervisor Carl Wang-Erickson for his patience and commitment.

References

[1] Fernando Q Gouvêa. *p-adic Numbers*. Springer, 1997.

[2] Alan J Baker. *An Introduction to p-adic Numbers and p-adic Analysis*. Available at `http://www.maths.gla.ac.uk/~ajb/dvi-ps/padicnotes.pdf`, 2011.

[3] Xavier Caruso. *Computations with p-adic numbers*. Available at `http://arxiv.org/abs/1701.06794`, 2017.

[4] Claudio Hüni, Kathrin Naef, and Daniel Schmitter. *p-adic Analysis Compared with Real: Lecture 3*. Available at `https://www2.math.ethz.ch/education/bachelor/seminars/hs2011/p-adic/report3.pdf`, 2011.

[5] Brian Conrad. *Some basics concerning absolute values*. Available at `http://math.stanford.edu/~conrad/676Page/handouts/ostrowski.pdf`.

[6] Jean-Pierre Serre. *Local fields*, volume 67. Springer Science & Business Media, 2013.

[7] Kiyosi Itō. *Encyclopedic dictionary of mathematics*, volume 1. MIT press, 1993.

[8] Lawrence C. Washington. *Introduction to cyclotomic fields*, volume 83 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1997.

[9] *Proof of congruence of Clausen and von Staudt*. Available at `http://planetmath.org/proofofcongruenceofclausenandvonstaudt`, 2013.