p-adic integration and the theory of groups

Alexander Frolkin

March 17, 2005

Acknowledgements

Firstly, I would like to thank my supervisor, Con Griffin, for his input, advice, and feedback. Thanks are also due to Nadia Sidorova for the calculation on page 27. I am grateful to Ivan Yudin, Kelly Morley, and Konstantin Ardakov for spotting numerous typos and errors.

Contents

0 Introduction

0.1 *p*-adic numbers

There are several ways to motivate the construction of the p-adic numbers — a number-theoretic way, a topological way, and an algebraic way.

Let p be a prime and suppose we have an infinite set of congruences,

$$f(X) \equiv 0 \pmod{p^i}, \quad \text{for each } i \in \mathbb{N},$$

where $f(X)$ is a polynomial with integer coefficients. In order to make this more manageable, we would like to somehow reduce this to at least a finite set of equations. We can express the value of $f(X)$ in base p as

$$f(X) = \sum_{s=0}^{\infty} a_s(X) p^s,$$

with $0 \leqslant a_s(X) < p$ for all s, and where almost all[1] of the $a_s(X)$ are zero. Then the set of congruences says that we need an X such that all the $a_s(X)$ are zero. However, a finite subset of the congruences can only tell us about the first n of the $a_s(X)$, for some n. If instead of ordinary integers we work with expressions of the form

$$\sum_{s=0}^{\infty} a_s p^s, \tag{1}$$

with $0 \leqslant a_s < p$ for each s, we can write the infinite set of congruences as the single equation

$$f(X) = 0.$$

However, it is clear that an expression such as (1) will not converge in the usual norm (unless almost all of the a_s are zero). In spite of this, though, we shall see that the p-adic numbers allow us to do precisely what we want.

Another question which leads to the p-adic numbers concerns completions of $\mathbb{Q}$ with respect to different topologies. Recall how $\mathbb{R}$ is obtained from $\mathbb{Q}$. We equip $\mathbb{Q}$ with the usual Euclidean topology, induced by the norm $\|x\| = |x|$, and then $\mathbb{R}$ is simply the completion of $\mathbb{Q}$ with respect to this norm. A natural question arises at this point: what other norms can we put on $\mathbb{Q}$ and what completions would they give rise to? Let us fix a prime p and note that by the unique factorisation property of $\mathbb{Z}$, we can write any non-zero $q \in \mathbb{Q}$ uniquely as $p^k a/b$ for some $k, a \in \mathbb{Z}$, $b \in \mathbb{N}$, where p does not divide a or b and a and b are coprime. We can then define a new norm on $\mathbb{Q}$ by $\|q\| = p^{-k}$. Completing $\mathbb{Q}$ with respect to this new norm we get a new field, not isomorphic to $\mathbb{R}$, denoted by $\mathbb{Q}_p$. These are the p-adic numbers. It can be shown that any non-trivial norm on $\mathbb{Q}$ is equivalent to either the usual (Euclidean) norm or one of the p-adic norms. (This is known as Ostrowski's theorem; see [BF93].) We will see later that $\mathbb{Q}_p$ has a subring $\mathbb{Z}_p$ with exactly the properties we would like in the above example. Although we have not given a proof, it should be plausible that a series such as (1) would converge in the p-adic norm.

Another way the p-adic numbers arise is in algebra, in the study of so-called profinite and pro-p groups. Before we define what these terms mean, let us

[1] That is, all but finitely many.

look at a familiar concept from real analysis. Suppose that we have a sequence $\{a_i \mid i \in \mathbb{N}\}$ of real numbers which is bounded above and

$$a_0 \leqslant a_1 \leqslant a_2 \leqslant a_3 \leqslant \cdots .$$

Then we can find an element $a \in \mathbb{R}$, the supremum of the a_i, which satisfies $a_i \leqslant a$ for all i. Another crucial property of this element is that if we have another real number b with the property that $a_i \leqslant b$ for all i, then we must have that $a_i \leqslant a \leqslant b$ for all i. So in quite a clear sense, the a is the smallest element at the top of the "tower" of the a_i. It turns out that we can define a similar (in spirit) construction in algebra. Suppose that we have a family of groups $\{G_i \mid i \in \mathbb{N}\}$, with maps[2] $\phi_i : G_i \to G_{i-1}$ for each $i > 0$,

$$G_0 \xleftarrow{\phi_1} G_1 \xleftarrow{\phi_2} G_2 \xleftarrow{\phi_3} G_3 \xleftarrow{\phi_4} \cdots .$$

Then we can find a group G and maps $G_i \xleftarrow{\psi_i} G$ for each i (with the ψ_i respecting the ϕ_i in the sense that[3] $\psi_i\phi_i = \psi_{i-1}$ for each $i > 0$), with the crucial property that if X is another group with maps $G_i \xleftarrow{\alpha_i} X$ (with the α_i respecting the ϕ_i), then there is a map $G \xleftarrow{\beta} X$ such that the composition $G_i \xleftarrow{\psi_i} G \xleftarrow{\beta} X$ gives the same map as α_i. So, each α_i factors through G, via the same map β. We can therefore, as in the example with the real numbers, think of G (in a somewhat less clear sense) as being the "smallest" object at the top of the "tower" of the G_i. It can be shown that such a G always exists and is unique (up to isomorphism). This G is known as the **inverse limit** of the G_i, written

$$G = \varprojlim_i G_i.$$

Now, if $\mathcal{C}$ is a suitable class of groups, then a group is said to be **pro-$\mathcal{C}$** if it is an inverse limit of groups from class $\mathcal{C}$. A lot of current research in group theory concentrates on profinite and pro-p groups (the latter being inverse limits of p-groups). When looking at a class of groups, it is often instructive to look at the simplest examples, so we ask: what is the simplest non-trivial example of a pro-p group? This will naturally be given by an inverse limit of some simple family of p-groups. One such family is the family of cyclic groups of p-power order, namely $\{\mathbb{Z}/p^n\mathbb{Z} \mid n > 0\}$, which has maps $\phi_i : \mathbb{Z}/p^i\mathbb{Z} \to \mathbb{Z}/p^{i-1}\mathbb{Z}$ given by reduction modulo p^{i-1},

$$\mathbb{Z}/p\mathbb{Z} \xleftarrow{\phi_2} \mathbb{Z}/p^2\mathbb{Z} \xleftarrow{\phi_3} \mathbb{Z}/p^3\mathbb{Z} \xleftarrow{\phi_4} \cdots .$$

The inverse limit of this is precisely the group of p-adic integers $\mathbb{Z}_p$, with the maps $\psi_i : \mathbb{Z}_p \to \mathbb{Z}/p^i\mathbb{Z}$ being given by reduction modulo p^{i-1}, i.e.,

$$\mathbb{Z}_p = \varprojlim_i \mathbb{Z}/p^i\mathbb{Z}.$$

This approach helps to see what the p-adic integers look like. If we take an element $x + p^n\mathbb{Z} \in \mathbb{Z}/p^n\mathbb{Z}$, we can write it uniquely as

$$x + p^n\mathbb{Z} = \sum_{i=0}^{n-1} a_i p^i + p^n\mathbb{Z},$$

[2]We shall use the word map to mean a homomorphism.
[3]We are writing maps on the right.

where $0 \leqslant a_i < p$ for each i. If we think of $\mathbb{Z}_p$ as being the "smallest" object at the top of the "tower", then it should be intuitively clear that this forces $\mathbb{Z}_p$ to be the set of all expressions of the form

$$\sum_{i=0}^{\infty} a_i p^i,$$

with $0 \leqslant a_i < p$ for each i.

We shall give the precise definition of inverse limits later on, and rigorously define $\mathbb{Z}_p$ in this way. We shall define a topology on $\mathbb{Z}_p$, making it a compact topological group, and exploit this to construct a measure, leading to a theory of p-adic integration.

0.2 Zeta functions of groups

For an excellent introduction to this, see [dS03].

While the theory of groups has been a major topic of research during the past two centuries or so, there are still many elementary questions, the answers to which are not well understood. For example, let $f(p, n)$ denote the number of isomorphism classes of groups of order p^n (where p is a prime and n a positive integer). While it is possible to compute explicit formulae for $f(p, n)$ for a given n (at the time of writing, this has been done for $1 \leqslant n \leqslant 7$ — see [NOVL04] and [OVL05]), there is no general formula for arbitrary n. However, the formulae for $f(p, n)$ for $n = 5$, 6 and 7 start to reveal a pattern, namely that $f(p, n)$ is given by a polynomial which depends on the residue class of the prime p modulo some integer N. However, work on this question has still not yielded a proof of this result, known as Higman's PORC (Polynomial On Residue Class) conjecture.

Conjecture 0.2.1 (PORC). *Let n be a positive integer. Then there is a positive integer N and polynomials $P_{n,0}(X), \ldots, P_{n,N-1}(X) \in \mathbb{Q}[X]$ such that if $p \equiv i \pmod{N}$, then*

$$f(p, n) = P_{n,i}(p).$$

To illustrate this, the following was proved in [NOVL04].

$$f(p, 6) = 3p^2 + 39p + 344 + 24\gcd(p-1, 3) + 11\gcd(p-1, 4) + 2\gcd(p-1, 5),$$

for $p \geqslant 5$. In this case, $N = 3 \cdot 4 \cdot 5 = 60$. The formula for $f(p, 7)$ in [OVL05] is rather longer, but takes on a similar form, with $N = 2520$.

Over the centuries, analytic functions of the form

$$\sum_{n=1}^{\infty} \frac{a_n}{n^s}$$

have proven to be a very powerful tool initially in number theory, and later in other areas of mathematics. Functions of this form are referred to as **Dirichlet series**. The most famous example of such a function is the Riemann zeta function,

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}, \quad \Re(s) > 1,$$

which lies at the heart of analytic number theory by virtue of the fact that it encodes a surprising amount of information about the distribution of prime numbers. One of the most important of the Millennium Prize Problems concerns the zeros of this function — the solution of this problem would have far-reaching consequences in number theory.

In 1988, in a paper studying finite index subgroups of nilpotent groups, [GSS88], Grunewald, Segal, and Smith introduced the notion of a zeta function of a group. They defined this for torsion-free finitely generated nilpotent groups, what they called the class of $\boldsymbol{\mathcal{T}}$**-groups**, as follows. If G is a $\mathcal{T}$-group, then

$$\zeta_G^{\leqslant}(s) = \sum_{H \leqslant_f G} |G:H|^{-s} = \sum_{n=1}^{\infty} a_n^{\leqslant}(G) n^{-s},$$

where

$$a_n^{\leqslant}(G) = |\{H \leqslant G \mid |G:H| = n\}|.$$

There is a corresponding definition of $\zeta_G^{\trianglelefteq}$ for normal subgroups. Although zeta functions can be defined for a larger class of groups, attention is usually restricted to nilpotent groups. The reason is that nilpotent groups decompose into direct products of their Sylow subgroups, giving a corresponding decomposition of the zeta function into so-called **local factors** whenever the zeta function converges in a right half-plane in $\mathbb{C}$,

$$\zeta_G^*(s) = \prod_{p \text{ prime}} \zeta_{G,p}^*(s),$$

where

$$\zeta_{G,p}^*(s) = \sum_{n=0}^{\infty} a_{p^n}^*(G) p^{-ns},$$

for $* \in \{\leqslant, \trianglelefteq\}$. We can see that the local factors count subgroups of p-power index, and this suggests a way to approach PORC. We know that any group is a quotient of a free group, and that a finite p-group is nilpotent. Hence, a finite p-group is a quotient of a free nilpotent group and, moreover, of a finitely generated free nilpotent group. Now, a free group is torsion-free, so finitely generated free nilpotent groups fall into the class of $\mathcal{T}$-groups. The finite p-groups arise from finite p-power index normal subgroups of the free nilpotent group, so the hope is that we may be able to count these groups using the local factors of the zeta function counting normal subgroups of the free group.

Later on, we shall see how this approach develops, what other results have been deduced from studying zeta functions, and how the theory of p-adic numbers and p-adic integration applies.

1 A few prerequisites

In this section, we shall give a brief introduction to the necessary mathematics and prove some preliminary results. At this stage, the mathematics we are considering may well seem somewhat random and unrelated. Later on, though, we shall show how to tie all these diverse concepts together to produce many interesting results.

1.1 Inverse systems and their limits

We shall start where we left off in 0.1, where we gave a feeling for what an inverse limit should be. Here, we shall define inverse systems and inverse limits of rings (not groups, as we did in the introduction — this will save us some work later on), indexed by the natural numbers. In fact, it is possible to define inverse systems in much greater generality, where the objects come from arbitrary categories and are indexed by arbitrary partially ordered sets (and in fact, the proofs that we give require little modification in order to make them work in this more general setting). However, we shall not need this level of generality and refer the interested reader to [Rot79, Chapter 2]. From now on, whenever we say "inverse system", we shall mean an inverse system of rings, indexed by the natural numbers. We shall also use the word "map" to mean a ring homomorphism.

Let us launch straight into the definition.

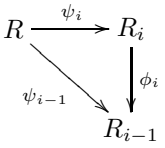
Definition 1.1.1. *Let $\{R_i \mid i \in \mathbb{N}\}$ be a family of rings, and suppose that for each $i > 0$ we have a map $\phi_i : R_i \to R_{i-1}$,*

$$R_0 \xleftarrow{\phi_1} R_1 \xleftarrow{\phi_2} R_2 \xleftarrow{\phi_3} R_3 \xleftarrow{\phi_4} \cdots .$$

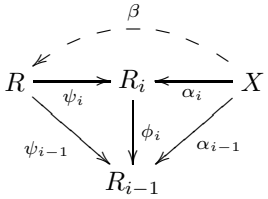
*Then the family of rings together with the family of maps is known as an **inverse system**, denoted by $\{R_i, \phi_i\}$.*

Intuitively, then, we have a sequence of rings, and we would like to define its "least upper bound", in some sense. We make the following definition.

Definition 1.1.2. *Let $\{R_i, \phi_i\}$ be an inverse system. Then a ring R is said to be an **inverse limit** of the system if there are maps $\psi_i : R \to R_i$, known as **connecting maps**, making the diagram*



commute for each $i > 0$ and satisfying the following universal mapping property.

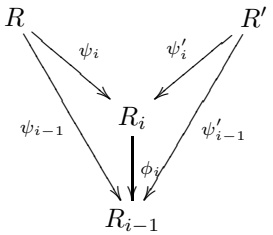


For every ring X and maps $\alpha_i : X \to R_i$ making the diagram commute (for each $i > 0$), there is a unique map $\beta : X \to R$ making the diagram commute (for each $i > 0$).

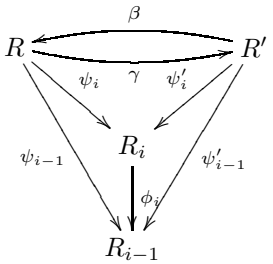
Before we're done, we need to show that inverse limits exist and are unique.

Theorem 1.1.3. *Let $\{R_i, \phi_i\}$ be an inverse system. Then an inverse limit, if it exists, is unique, up to isomorphism.*

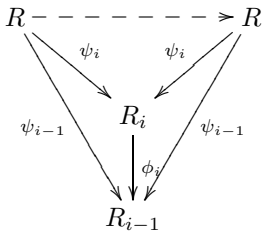
Proof. Suppose R and R' are two inverse limits of $\{R_i, \phi_i\}$. Then there are maps ψ_i, ψ'_i for each i, making the following diagram commute whenever $i > 0$.



Then by definition, there are maps $\beta : R' \to R$ and $\gamma : R \to R'$ making the following diagram commute whenever $i > 0$.



From the diagram, $\psi_i = \gamma\beta\psi_i$ and $\psi'_i = \beta\gamma\psi'_i$. Now consider the following diagram.



Since we have $\psi_i = \gamma\beta\psi_i$, for the dashed arrow, we can choose either $\gamma\beta$ or the identity. By uniqueness, $\gamma\beta = 1$. By considering a similar diagram with R', we deduce that $\beta\gamma = 1$, so that β and γ are isomorphisms, as required. □

Proposition 1.1.4. *Let $\{R_i, \phi_i\}$ be an inverse system. Then the system has an inverse limit.*

Proof. Define

$$R = \left\{ (x_i)_{i\in\mathbb{N}} \in \prod_{j\in\mathbb{N}} R_j \;\middle|\; x_{i-1} = x_i\phi_i,\ i > 0 \right\}.$$

Since the ϕ_i are ring homomorphisms, it easily follows that R is a ring (which is commutative if each R_i is). Define $\psi_i : R \to R_i$ to be the restriction to R of the projection $\pi_i : \prod_{j\in\mathbb{N}} R_j \to R_i$, that is, $\psi_i = \pi_i|_R$. Then if $(x_j)_{j\in\mathbb{N}} \in R$, $((x_j)_{j\in\mathbb{N}})\psi_i\phi_i = x_i\phi_i = x_{i-1} = ((x_j)_{j\in\mathbb{N}})\psi_{i-1}$, hence $\psi_i\phi_i = \psi_{i-1}$.

Now let X be a ring with maps $\alpha_i : X \to R_i$ such that $\alpha_i\phi_i = \alpha_{i-1}$. Then for $x \in X$, $x\alpha_{i-1} = x(\alpha_i\phi_i) = (x\alpha_i)\phi_i$, so $(x\alpha_j)_{j\in\mathbb{N}} \in \prod_{j\in\mathbb{N}} R_j$ defines an element of R. Hence, we can define a map $\beta : X \to R : x \mapsto (x\alpha_j)_{j\in\mathbb{N}}$. Then for $x \in X$, $x\beta\psi_i = ((x\alpha_j)_{j\in\mathbb{N}})\psi_i = x\alpha_i$, so $\beta\psi_i = \alpha_i$, as required. □

Now that we have established the existence and uniqueness of the inverse limit, let us introduce a notation for it. If $\{R_i, \phi_i\}$ is an inverse system, we shall denote the inverse limit, R, by

$$R = \varprojlim_i R_i.$$

1.2 The topology on an inverse limit

I feel I must apologise for the somewhat odd notation that follows, but this is what happens when an algebraist who enjoys writing maps on the right attempts to do point-set topology.

1.2.1 Products of topological spaces

We have seen (in the proof of 1.1.4) that an inverse limit can be construed as a subring of the product of the individual rings R_i. If we equip each ring R_i with a topology, we can define a natural topology on the inverse limit as the subspace topology induced by the topology on the product.

We shall show how we can define, in a natural way, the topology on a product of topological spaces. However, since our focus here is not on topology, we shall not go into too much detail and refer the reader to [Bou89], [Wil70] and [Mor04].

We shall make use of the following concept.

Definition 1.2.1. *Let $(Y, \mathcal{T})$ be a topological space. Then a subset $\mathcal{S}$ of $\mathcal{T}$ is said to be a **subbase** for the topology $\mathcal{T}$ on Y if any element of $\mathcal{T}$ can be expressed as a union of finite intersections of elements of $\mathcal{S}$.*

This generalises the concept of a base. In fact, it is easy to see that an equivalent definition of a subbase is that the set of all finite intersections forms a base. We can now state the definition of the product topology in terms of a subbase. The motivation for this definition is that we would like the projection maps to be continuous.

Definition 1.2.2. *Let $\{X_i \mid i \in I\}$ be a family of topological spaces, let $X = \prod_{i\in I} X_i$, and let $\pi_i : X \to X_i$ be the projection maps. Then the **Tychonoff topology** on X is defined to be the topology with subbase consisting of all sets of the form $U_i\pi_i^{-1}$ where U_i is open in X_i.*

This is in fact the coarsest topology with respect to which the projections maps are continuous.

We could also have given this definition in terms of a base, but this would have appeared less natural. Let us see what this says in terms of a base. Let U_{j_k} be open in X_{j_k} for $k = 1, \ldots, n$. Then $U_{j_k}\pi_{j_k}^{-1} = \prod_{i\in I} A_i^{j_k}$, where $A_{j_k}^{j_k} = U_{j_k}$

and $A_i^{j_k} = X_i$ for $i \neq j_k$. Now, $\bigcap_{j=1}^n U_{j_k}\pi_{j_k}^{-1} = \bigcap_{j=1}^n \prod_{i\in I} A_i^{j_k} = \prod_{i\in I} A_i$ where $A_i = X_i$ except for $i = j_1, \ldots, j_n$. Hence, a base for the Tychonoff topology is given by the sets $\prod_{i\in I} A_i$ where A_i is open in X_i and $A_i = X_i$ for almost all $i \in I$.

We can now state one of the major theorems of analytic topology, which we shall briefly, but crucially, make use of later.

Theorem 1.2.3 (Tychonoff). *Let $\{X_i \mid i \in I\}$ be a family of topological spaces and let $X = \prod_{i\in I} X_i$. Then X is compact in the Tychonoff topology if and only if each X_i is compact.*

The proof is somewhat non-trivial, and we shall omit it, referring the reader to [Bou89, Chapter 1, §9, 5] and [Wil70, 17.8].

1.2.2 The consequences for inverse limits

With the definitions of 1.2.1, we can now see, given an inverse system $\{R_i, \phi_i\}$, how we can naturally define a topology on $\varprojlim_i R_i$. Namely, we equip each R_i with the discrete topology and we equip $\prod_{i\in\mathbb{N}} R_i$ with the Tychonoff topology. Then the topology on $\varprojlim_i R_i$ is the subspace topology induced by the Tychonoff topology on the product.

What we would like to show is that an inverse limit of finite rings is compact. Clearly, if we equip each finite R_i with the discrete topology, it becomes compact, and then Tychonoff's theorem tells us that the product is compact in the Tychonoff topology. It hence remains to show that the inverse limit is a closed subset of the product. First, a lemma.

Lemma 1.2.4. *Let X, Y and Z be topological spaces and $\alpha : X \to Y$, $\beta : X \to Z$ be continuous. Then the map $\alpha \times \beta : X \to Y \times Z : x \mapsto (x\alpha, x\beta)$ is also continuous.*

Proof. Let U be open in X and V be open in Y and consider $(U \times V)(\alpha \times \beta)^{-1}$. Now,

$$\begin{aligned} x \in (U \times V)(\alpha \times \beta)^{-1} &\iff x\alpha \in U,\, x\beta \in V \\ &\iff x \in U\alpha^{-1} \cap V\beta^{-1}. \end{aligned}$$

By continuity, $U\alpha^{-1}$ and $V\beta^{-1}$ are open in X, so $(U \times V)(\alpha \times \beta)^{-1}$ is open in X, as required. (Note that it suffices to check the conditions just for the base.) □

Proposition 1.2.5. *Let $\{R_i, \phi_i\}$ be an inverse system of finite rings, with each ring equipped with the discrete topology. Then,*

$$\varprojlim_i R_i = \left\{ (x_i)_{i\in\mathbb{N}} \in \prod_{j\in\mathbb{N}} R_j \;\middle|\; x_{i-1} = x_i\phi_i,\, i > 0 \right\}$$

is a closed subspace of $\prod_{i\in\mathbb{N}} R_i$.

Proof. Let us forget about the ring structure for a moment and just consider the R_i and $R = \varprojlim_i R_i$ as additive abelian groups. For each $i \in \mathbb{N}$, define

$\lambda_i : \prod_{j\in\mathbb{N}} R_j \to R_i : (x_j)_{j\in\mathbb{N}} \mapsto x_i - x_{i+1}\phi_{i+1}$. It is easy to see that each λ_i is a homomorphism of abelian groups. Now note that

$$\begin{aligned}(x_j)_{j\in\mathbb{N}} \in R &\iff \forall i \in \mathbb{N}, x_i - x_{i+1}\phi_{i+1} = 0\\ &\iff \forall i \in \mathbb{N}, ((x_j)_{j\in\mathbb{N}})\lambda_i = 0\\ &\iff (x_j)_{j\in\mathbb{N}} \in \bigcap_{i\in\mathbb{N}} \mathrm{Ker}(\lambda_i),\end{aligned}$$

that is,

$$R = \bigcap_{i\in\mathbb{N}} \mathrm{Ker}(\lambda_i).$$

Now, if each λ_i is continuous, then $\mathrm{Ker}(\lambda_i) = \{0\}\lambda_i^{-1}$ is a closed subspace of $\prod_{j\in\mathbb{N}} R_j$ as $\{0\}$ is closed in R_i (since R_i has the discreet topology). It then follows that R, being the intersection of closed subspaces, is a closed subspace of $\prod_{j\in\mathbb{N}} R_j$.

Define $\sigma_i : R_i \times R_i \to R_i : (r, s) \mapsto r - s$ and let $\pi_i : \prod_{j\in\mathbb{N}} R_j \to R_i$ be the projection maps. Now note that $\lambda_i = (\pi_i \times \pi_{i+1}\phi_{i+1})\sigma_i$ and that σ_i and ϕ_{i+1} are continuous as they are maps between discreet spaces and the projections are continuous by definition of the Tychonoff topology. We thus conclude, with the help of 1.2.4, that each λ_i is continuous, as required. □

Corollary 1.2.6. *An inverse limit of finite rings is compact.*

Proof. This follows from 1.2.5, the fact that a finite set with the discrete topology is compact and the fact that a closed subspace of a compact space is compact. □

We shall refer to this topology on the inverse limit as the **profinite topology**. Another important result we shall need concerns the continuity of addition and subtraction in an inverse limit of finite rings.

Proposition 1.2.7. *Let R be an inverse limit of finite rings. Then the maps $\alpha : R \times R \to R : (x, y) \mapsto x + y$ and $\iota : R \to R : x \mapsto -x$ are continuous.*

Proof. Let $R = \varprojlim_i R_i$, and let

$$X = R \cap \prod_{i\in\mathbb{N}} A_i \subseteq R,$$

where $A_i = R_i$ for almost all i. Recall that this is an element of a base for the topology on R. Now,

$$\begin{aligned}x = (x_i)_{i\in\mathbb{N}} \in X\iota^{-1} &\iff -x \in R, \text{ and for all } i \in \mathbb{N}, -x_i \in A_i\\ &\iff x \in R, \text{ and for all } i \in \mathbb{N}, x_i \in -A_i\\ &\iff x \in R \cap \prod_{i\in\mathbb{N}} (-A_i).\end{aligned}$$

But for almost all i, $-A_i = -R_i = R_i$, so that

$$X\iota^{-1} = R \cap \prod_{i\in\mathbb{N}} (-A_i)$$

is an element of the base, and hence open.

Now consider addition. Suppose that $A_{i_k} \neq R_{i_k}$ for $k = 1, \ldots, n$ and that $A_i = R_i$ for $i \neq i_k$, $k = 1, \ldots, n$.

$$\begin{aligned} X\alpha^{-1} &= \{(x,y) \in R \times R \mid x_i + y_i \in A_i,\, i \in \mathbb{N}\} \\ &= \{(x,y) \in R \times R \mid x_{i_k} + y_{i_k} \in A_{i_k},\, k = 1, \ldots, n\}. \end{aligned}$$

since $x_i, y_i \in R_i$ so certainly $x_i + y_i \in R_i$ (for $i \neq i_k$, $k = 1, \ldots, n$). Note that

$$X\alpha^{-1} = \bigcap_{j=1}^{k} \{(x,y) \in R \times R \mid x_{i_j} + y_{i_j} \in A_{i_j}\}.$$

Now let α_{R_i} be the addition map on R_i, and let $\pi_i : R \to R_i$ be the projection maps. Then,

$$\begin{aligned} & \{(x,y) \in R \times R \mid x_{i_j} + y_{i_j} \in A_{i_j}\} \\ = \; & \{(x,y) \in R \times R \mid (x,y)(\pi_{i_j} \times \pi_{i_j})\alpha_{R_{i_j}} \in A_{i_j}\} \\ = \; & \{(x,y) \in R \times R \mid (x,y) \in A_{i_j}\alpha_{R_{i_j}}^{-1}(\pi_{i_j} \times \pi_{i_j})^{-1}\} \\ = \; & A_{i_j}\alpha_{R_{i_j}}^{-1}(\pi_{i_j} \times \pi_{i_j})^{-1}. \end{aligned}$$

This is open as R_{i_j} has the discrete topology (so $\alpha_{R_{i_j}}$ is continuous), the projection maps are continuous and by 1.2.4. □

1.2.3 Profinite completions

We shall take this opportunity to mention two useful constructions from profinite group theory. If we want to apply methods from profinite group theory to an arbitrary group, we need to find a profinite group which contains our original group. We define the following concept. (Note that this requires us to talk about inverse limits of groups, and in a more general setting than that in which we have analysed them above.)

Definition 1.2.8. *Let G be a group. Then the **profinite completion** of G is the profinite group*

$$\hat{G} = \varprojlim_{N \trianglelefteqslant_{\mathrm{f}} G} G/N,$$

*where $N \trianglelefteqslant_{\mathrm{f}} G$ denotes that N is a normal subgroup of G of finite index. The **pro-p completion** $\hat{G}_p$ of G is defined similarly, but the limit is taken over all normal subgroups of p-power index.*

We can look at the inverse limit in a similar way to how we looked at it in the proof of 1.1.4. Given a group G, we have a homomorphism $\phi : G \to \hat{G} : g \mapsto (Ng)_{N \trianglelefteqslant_{\mathrm{f}} G}$. The kernel of this is clearly

$$\mathrm{R}(G) = \bigcap_{N \trianglelefteqslant_{\mathrm{f}} G} N,$$

so we see that G injects into $\hat{G}$ if $\mathrm{R}(G) = 1$. In fact, G injects into $\hat{G}$ if and only if $\mathrm{R}(G) = 1$. If $\mathrm{R}(G) = 1$, we say that G is **residually finite**. We therefore see

that we can use the profinite completion as a tool for studying residually finite groups. In a similar way, we can see that G injects into $\hat{G}_p$ if and only if G is **residually p**, that is,

$$\mathrm{R}_p(G) = \bigcap_{N \trianglelefteqslant_p G} N = 1,$$

where $N \trianglelefteqslant_p G$ denotes that N is a normal subgroup of p-power index in G.

We shall not make significant use of these concepts, but we will see them arise in a few situations.

1.3 Lie algebras

In the study of zeta functions, we shall see that it is often helpful to linearise questions about groups. The tool that we shall require for this is the theory of Lie algebras.

Definition 1.3.1. *Let R be a commutative ring. Then an R-module $\mathfrak{g}$ together with a bilinear map $\mathfrak{g} \times \mathfrak{g} \to \mathfrak{g} : (x, y) \mapsto [x, y]$ (known as the **Lie bracket**) satisfying*

- *$[x, x] = 0$ for all $x \in \mathfrak{g}$,*
- *$[x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0$ for all $x, y, z \in \mathfrak{g}$ (the **Jacobi identity**),*

*is said to be a **Lie algebra**.*

Let $\mathfrak{g}$ be a Lie algebra over a ring R and let $x, y \in \mathfrak{g}$. Then,

$$0 = [x - y, x - y] = [x, x] - [y, x] - [x, y] + [y, y] = -[y, x] - [x, y].$$

Hence, $[x, y] = -[y, x]$. In fact, if we restrict R to being a field of zero or odd characteristic, it is easy to see that this becomes equivalent to the condition $[x, x] = 0$.

An example of a Lie algebra (over a field $\Bbbk$) is given by the general linear algebra $\mathfrak{gl}(n, \Bbbk)$, consisting of all $n \times n$ matrices over $\Bbbk$ with the Lie bracket defined as the matrix commutator, namely $[A, B] = AB - BA$. A (Lie) subalgebra of this is the special linear algebra $\mathfrak{sl}(n, \Bbbk) = \{A \in \mathfrak{gl}(n, \Bbbk) \mid \mathrm{tr} A = 0\}$. Another example is $\mathbb{R}^3$ with the usual vector cross product.

Lie algebras originally arose in the study of Lie groups (these are groups which are also differentiable manifolds in which the group operations are differentiable), as tangent spaces at the identity. For example, $\mathfrak{sl}(n, \mathbb{C})$ is the tangent space of the Lie group $\mathrm{SL}(n, \mathbb{C})$ at the identity. However, the study of Lie algebras has proved to be an interesting and useful area in its own right.

As usual with algebraic structures, we can define substructures. A Lie subalgebra is defined in the obvious way, and ideals are defined as follows.

Definition 1.3.2. *Let $\mathfrak{g}$ be a Lie algebra over a ring R, and let $A, B \subseteq \mathfrak{g}$. Then, we define*

$$[A, B] = \langle [a, b] \mid a \in A, b \in B \rangle_R.$$

*(That is, the R-span of the Lie brackets of elements of A with elements of B.) A submodule $\mathfrak{i}$ of $\mathfrak{g}$ is said to be an **ideal** in $\mathfrak{g}$ if $[\mathfrak{g}, \mathfrak{i}] \leqslant \mathfrak{i}$. (That is, $[x, y] \in \mathfrak{i}$ for all $x \in \mathfrak{g}$ and $y \in \mathfrak{i}$.) The **index** $|\mathfrak{g} : \mathfrak{i}|$ of $\mathfrak{i}$ in $\mathfrak{g}$ is the index of $\mathfrak{i}$ in $\mathfrak{g}$ as an additive subgroup.*

Note the resemblance of this definition to the definition of an ideal in a ring. As always, we can define quotients of Lie algebras by ideals and we have versions of the isomorphism theorems.

1.4 Measures and integration

The first step to a theory of integration is a measure. Intuitively, a measure is a function from some collection of sets, which gives each set a "size". For example, given a set X, $\mu(A) = |A|$, where $A \subseteq X$, defines a measure as does $\mu([a,b]) = b - a$, where $[a,b] \subseteq \mathbb{R}$ is an interval.

Here, we would like to define a p-adic integral, so the first thing we need is a p-adic measure.

Now, in order to have a good theory, we must restrict the collection of sets which we can "measure". This is done as follows.

Definition 1.4.1. *Let X be a set, and let $\Sigma \subseteq \mathcal{P}(X)$. Then Σ is said to be a **σ-algebra** on X if it contains X and is closed under taking complements and countable unions. That is,*

- $X \in \Sigma$,
- *if $A \in \Sigma$, then $X \backslash A \in \Sigma$,*
- *if $A_n \in \Sigma$ for each $n \in \mathbb{N}$, then $\bigcup_{n\in\mathbb{N}} A_n \in \Sigma$.*

Note that as a consequence, we have that any σ-algebra contains the empty set and is closed under countable intersections. We can now define precisely what we mean by a measure.

Definition 1.4.2. *Let X be a set and let Σ be a σ-algebra on X. Then a function $\mu : \Sigma \to [0,\infty]$ is said to be a **measure** if*

- $\mu(\emptyset) = 0$,
- *if $A_n \in \Sigma$ for each $n \in \mathbb{N}$, and $A_i \cap A_j = \emptyset$ for $i \neq j$, then*

$$\mu\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \sum_{n\in\mathbb{N}} \mu(A_n).$$

*The triple (X, Σ, μ) is known as a **measure space**.*

Note that we have included ∞ in the range of a measure μ, so that we allow $\mu(A) = \infty$.

Definition 1.4.3. *Let (X, Σ, μ) be a measure space. Then μ is said to be **finite** if $\mu(X)$ is finite.*

As noted above, we can obtain a measure space (X, Σ, μ) by taking any set X, with the σ-algebra $\Sigma = \mathcal{P}(X)$, and defining the measure by $\mu(A) = |A|$. This measure is known as the **counting measure** on X and is finite if and only if X is a finite set. A more sophisticated example of a measure is the Lebesgue measure on $[0,1]$, given by the Lebesgue integral

$$\mu(A) = \int_A 1 \, \mathrm{d}x,$$

where A belongs to the σ-algebra $\mathrm{B}[0,1]$ of Borel sets on $[0,1]$. For an interval $[a,b] \subseteq [0,1]$, this gives $\mu([a,b]) = b - a$.

Now, given a measure space (X, Σ, μ), a subset $A \subseteq X$, and a function $f : X \to C$ where $C \subseteq \mathbb{C}$ is countable, we can define an integral on X by

$$\int_A f \,\mathrm{d}\mu = \sum_{c \in C} f(c)\mu(A_f(c)),$$

where $A_f(c) = \{x \in A \mid f(x) = c\}$, whenever each $A_f(c) \in \Sigma$ and the series converges. Intuitively, what we are doing is splitting the integral into a sum over sets on which f is constant,

$$\int_A f \,\mathrm{d}\mu = \sum_{c \in C} \int_{A_f(c)} f \,\mathrm{d}\mu,$$

giving

$$\int_A f \,\mathrm{d}\mu = \sum_{c \in C} f(c) \int_{A_f(c)} 1 \,\mathrm{d}\mu = \sum_{c \in C} f(c)\mu(A_f(c)).$$

There is a (much more elaborate) way to define integrals of arbitrary functions $f : X \to \mathbb{C}$, but the above is all we shall need for our purposes, and we refer the reader to [Hal59, Chapter V] for more details.

1.4.1 Topological groups and Haar measure

Here, we shall be dealing with a particular measure, known as the Haar measure. The construction of this measure is quite sophisticated, so once again, we shall not go into the precise details, referring the reader to [Hal59, Chapter XI].

First, we shall need the following concepts.

Definition 1.4.4. *Let X be a Hausdorff topological space. Then X is said to be* ***locally compact*** *if every point has a compact neighbourhood.*

Clearly, a compact Hausdorff space is locally compact.

Definition 1.4.5. *Let G be a group. Then G is said to be a* ***topological group*** *if it is a topological space and the maps $G \to G : x \mapsto x^{-1}$ and $G \times G \to G : (x, y) \mapsto xy$ are continuous.*

As an aside, recall from the introduction that a profinite group is defined to be a group which is an inverse limit of some inverse system of finite groups. Such a group comes equipped with a natural topology, defined similarly to the topology on an inverse limit of finite rings, and it turns out that we can define profinite groups in purely topological terms. We can define a profinite group to be a compact Hausdorff topological group whose open subgroups form a base for the neighbourhoods of the identity.

Now let G be a locally compact topological group, and let Σ be the σ-algebra generated by all open subsets of G. (That is, the smallest σ-algebra on G containing all the open subsets of G.) This is known as the **Borel σ-algebra**, $\mathcal{B}(G)$. Note that it follows from the definitions, that since $\mathcal{B}(G)$ contains all open subsets of G, it must also contain all closed subsets of G.

Now, the fundamental result which we shall need is that G possesses a measure μ on the Borel σ-algebra, known as the **Haar measure**. This measure is

unique up to scalar multiplication and finite on all compact subsets of G. In particular, if G is compact, there is a canonical way to scale the measure, namely such that $\mu(G) = 1$. Crucially, the Haar measure is translation invariant[4], so that for an element $g \in G$ and $X \in \mathcal{B}(G)$, $\mu(Xg) = \mu(X)$.

When it comes to integrating, we observe that if we have a continuous function $f : G \to \mathbb{C}$, and $A \in \mathcal{B}(G)$, then for any $c \in Gf$, $A_f(c) = \{c\}f^{-1} \cap A \in \mathcal{B}(G)$ since $\{c\}f^{-1}$ is a closed subset of G since $\{c\}$ is closed because $\mathbb{C}$ is Hausdorff, and because $\mathcal{B}(G)$ is closed under taking intersections. We can therefore define the integral

$$\int_A f \, \mathrm{d}\mu$$

as described before, having made sure that the series defining it converges.

[4]We are actually talking about the *right* Haar measure. There is also a left Haar measure which is left-translation invariant, but we shall only need the case when G is abelian, so that these two measures coincide.

2 The p-adic numbers

In this section, we shall formally define the p-adic numbers, establish some results about them and their topology, and show how this leads to a theory of p-adic integration.

2.1 The formal definitions

Fix a prime p. We define the ring $\mathbb{Z}_p$ in terms of an inverse limit, as we did in the introduction. Note that for any $i \in \mathbb{N}$, $p^i\mathbb{Z}$ is an ideal in the ring $\mathbb{Z}$, and that for $i > 0$, the maps $\phi_i : \mathbb{Z}/p^i\mathbb{Z} \to \mathbb{Z}/p^{i-1}\mathbb{Z} : x + p^i\mathbb{Z} \mapsto x + p^{i-1}\mathbb{Z}$ are well-defined. These give an inverse system

$$\mathbb{Z}/p\mathbb{Z} \xleftarrow{\phi_2} \mathbb{Z}/p^2\mathbb{Z} \xleftarrow{\phi_3} \mathbb{Z}/p^3\mathbb{Z} \xleftarrow{\phi_4} \mathbb{Z}/p^4\mathbb{Z} \xleftarrow{\phi_5} \cdots .$$

We make the following definition.

Definition 2.1.1. *The ring $\mathbb{Z}_p$ of **p-adic integers** is defined by*

$$\mathbb{Z}_p = \varprojlim_i \mathbb{Z}/p^i\mathbb{Z}.$$

In a nutshell, the p-adic integers are the pro-p completion of the usual integers — $\mathbb{Z}_p = \hat{\mathbb{Z}}_p$.

To get some kind of an intuitive feel for $\mathbb{Z}_p$, we would like to know what a typical element of $\mathbb{Z}_p$ looks like. Let $x \in \mathbb{Z}_p$ and consider its images under the connecting maps $\psi_i : \mathbb{Z}_p \to \mathbb{Z}/p^i\mathbb{Z}$. For any n, we can uniquely write

$$x\psi_n = \sum_{i=0}^{n-1} a_{ni}p^i + p^n\mathbb{Z},$$

with $a_{ni} \in \mathbb{Z}$, $0 \leqslant a_{ni} < p$ for each i. Now consider what happens when $n = 1$.

$$x\psi_1 = a_{11} + p\mathbb{Z}.$$

When $n = 2$,

$$x\psi_2 = a_{21} + a_{22}p + p^2\mathbb{Z}.$$

Now recall the definition of the maps $\phi_i : \mathbb{Z}/p^i\mathbb{Z} \to \mathbb{Z}/p^{i-1}\mathbb{Z}$, namely reduction modulo p^{i-1}, that is, $x + p^i\mathbb{Z} \mapsto x + p^{i-1}\mathbb{Z}$. Now note that

$$x\psi_2\phi_2 = a_{21} + p\mathbb{Z},$$

and recall that the connecting maps satisfy $\psi_i\phi_i = \psi_{i-1}$, yielding

$$x\psi_1 = a_{21} + p\mathbb{Z} = a_{11} + p\mathbb{Z}.$$

Hence, $a_{21} = a_{11}$. Now suppose inductively, that $a_{ni} = a_{mi}$ for all $m < n$ and all $i \leqslant m$. Consider

$$x\psi_{n+1} = \sum_{i=0}^{n} a_{(n+1)i}p^i + p^{n+1}\mathbb{Z}.$$

Now,

$$x\psi_n = \sum_{i=0}^{n-1} a_{ni}p^i + p^n\mathbb{Z} = x\psi_{n+1}\phi_{n+1},$$

so $a_{ni} = a_{(n+1)i}$ for $i < n$. Hence, we can drop the first subscripts and write, for each n,

$$x\psi_n = \sum_{i=0}^{n-1} a_i p^i + p^n\mathbb{Z}.$$

We can therefore identify $\mathbb{Z}_p$ with the set of expressions of the form

$$\sum_{i=0}^{\infty} a_i p^i$$

with $a_i \in \mathbb{Z}$, and $0 \leqslant a_i < p$ for each i. Using a similar method, we can check that addition and multiplication of these expressions is what we expect, namely

$$\sum_{i=0}^{\infty} a_i p^i + \sum_{j=0}^{\infty} b_j p^j = \sum_{i=0}^{\infty}(a_i + b_i)p^i, \quad \sum_{i=0}^{\infty} a_i p^i \sum_{j=0}^{\infty} b_j p^j = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} a_i b_j p^{i+j}.$$

We can rewrite the resulting expressions in the form $\sum_{i=0}^{\infty} c_i p^i$ with $0 \leqslant c_i < p$ for each i by writing each $a_i + b_i$ (respectively, $a_i b_j$) in base p as $\sum_{i=0}^{n} d_i p^i$ for some n, with $0 \leqslant d_i < p$ for each i.

Proposition 2.1.2. $\mathbb{Z}_p \cong \{\sum_{i=0}^{\infty} a_i p^i \mid a_i \in \mathbb{Z},\ 0 \leqslant a_i < p\}$.

In order to be able to define the field of p-adic numbers, we need to show the following.

Proposition 2.1.3. *$\mathbb{Z}_p$ is an integral domain.*

Proof. Let $a = \sum_{i=0}^{\infty} a_i p^i, b = \sum_{i=0}^{\infty} b_i p^i \in \mathbb{Z}_p$ be non-zero p-adic integers. Suppose that $a_j \neq 0$, but $a_r = 0$ for $r < j$ and that $b_k \neq 0$ but $b_r = 0$ for $r < k$. Then, $p^{j+1} \nmid a$ and $p^{k+1} \nmid b$. Now, the first term of ab is $a_j b_k p^{j+k}$, and so $p^{j+k+2} \nmid ab$ (noting that $0 \leqslant a_j b_k < p^2$). Hence ab is also non-zero. □

This now allows us to define the p-adic numbers as follows.

Definition 2.1.4. *The field of **p-adic numbers** is defined by $\mathbb{Q}_p = \mathrm{Q}(\mathbb{Z}_p)$ (that is, the field of fractions of $\mathbb{Z}_p$).*

2.2 The topology on $\mathbb{Z}_p$

Let us now look at the topological approach which leads to the p-adic numbers. First, we shall introduce a norm on $\mathbb{Q}$.

Definition 2.2.1. *For a non-zero integer $n = p^k m$, where $p \nmid m$, we define the **p-adic valuation** by*

$$\nu_p(n) = k.$$

For example, $\nu_5(25) = 2$, $\nu_7(13) = 0$, and $\nu_2(768) = 8$. Note that for integers m and n, $\nu_p(mn) = \nu_p(m) + \nu_p(n)$, so that the p-adic valuation behaves somewhat like a logarithm.

We can now extend the definition of the p-adic valuation to the whole of $\mathbb{Q}$ and define the associated norm.

Definition 2.2.2. *For $q = a/b \in \mathbb{Q}$, define the **p-adic valuation** by*

$$\nu_p(q) = \nu_p(a) - \nu_p(b).$$

*We define the **p-adic norm** on $\mathbb{Q}$ by*

$$\|q\|_p = \begin{cases} 0, & q = 0 \\ p^{-\nu_p(q)}, & q \neq 0 \end{cases}.$$

Our observation above ensures that our definition of the valuation on $\mathbb{Q}$ is sound — $\nu_p(ca/cb) = \nu_p(ca) - \nu_p(cb) = \nu_p(c) + \nu_p(a) - \nu_p(c) - \nu_p(b) = \nu_p(a) - \nu_p(b) = \nu_p(a/b)$.

From now on, we shall feel free to drop the subscripts when it is clear which prime p we have in mind.

We have the following simple but important result.

Lemma 2.2.3. *For $q, r \in \mathbb{Q}$, $\nu(qr) = \nu(q) + \nu(r)$.*

Proof. Let $q = a/b$, $r = c/d$. Then,

$$\begin{aligned} \nu(qr) = \nu(ac/bd) = \nu(ac) - \nu(bd) &= \nu(a) - \nu(b) + \nu(c) - \nu(d) \\ &= \nu(a/b) + \nu(c/d) = \nu(q) + \nu(r). \end{aligned}$$

□

The p-adic norm satisfies a stronger version of the triangle inequality, known as the **ultrametric inequality**.

Proposition 2.2.4. *For $x, y \in \mathbb{Q}$, $\|x + y\| \leqslant \max\{\|x\|, \|y\|\}$.*

Proof. First note that for integers m and n, with $\nu(n) \leqslant \nu(m)$, we have $p^{\nu(n)} \mid n$ and $p^{\nu(n)} \mid m$. Hence $p^{\nu(n)} \mid n+m$, and so $\nu(n+m) \geqslant \nu(n) = \min\{\nu(n), \nu(m)\}$.

Now let $x = a/b$ and $y = c/d$. Then,

$$\begin{aligned} \nu(x+y) = \nu\left(\frac{ad+bc}{bd}\right) &= \nu(ad+bc) - \nu(bd) \geqslant \min\{\nu(ad), \nu(bc)\} - \nu(bd) \\ &= \min\{\nu(ad) - \nu(bd), \nu(bc) - \nu(bd)\} = \min\{\nu(a) - \nu(b), \nu(c) - \nu(d)\} \\ &= \min\{\nu(x), \nu(y)\}. \end{aligned}$$

Hence,

$$p^{-\nu(x+y)} \leqslant p^{-\min\{\nu(x),\nu(y)\}} = \max\{p^{-\nu(x)}, p^{-\nu(y)}\},$$

as required. □

Proposition 2.2.5. *For $x, y \in \mathbb{Q}$,*

(i) $\|x\| \geqslant 0$ *and* $\|x\| = 0$ *if and only if* $x = 0$,

(ii) $\|xy\| = \|x\| \, \|y\|$,

(iii) $\|x + y\| \leqslant \|x\| + \|y\|$.

Proof. (i) Clearly $\|x\| \geqslant 0$. Also, $p^n \neq 0$ for $n \in \mathbb{Z}$, so $\|x\| = 0$ if and only if $x = 0$.

(ii) This is clear if $x = 0$ or $y = 0$. Otherwise, $\|xy\| = p^{-\nu(xy)} = p^{-\nu(x)-\nu(y)} = \|x\|\,\|y\|$.

(iii) This follows immediately from 2.2.4.

□

The p-adic norm thus really is a norm on $\mathbb{Q}$. What we can now do is take the completion of $\mathbb{Q}$ with respect to this norm to get a complete field. It can be shown (although we shall not need this) that this gives a field isomorphic to the field $\mathbb{Q}_p$ which we defined earlier. Note that we can write any $q \in \mathbb{Q}$ in the form

$$q = \sum_{i=-N}^{M} a_i p^i,$$

for some N and M, where $0 \leqslant a_i < p$ for each i. $\mathbb{Q}_p$ is then the set of expressions of the form

$$\sum_{i=-N}^{\infty} a_i p^i,$$

for some N with $0 \leqslant a_i < p$ for each i. This expression has norm p^N. We can therefore define the ring $\mathbb{Z}_p$ as a subring of $\mathbb{Q}_p$ by

$$\mathbb{Z}_p = \{x \in \mathbb{Q}_p \mid \|x\|_p \leqslant 1\}.$$

This brings us to the central result of this section. We have now defined two topologies on $\mathbb{Z}_p$, and we would like to know how they are related. In an ideal world, these two topologies would coincide. This is indeed so.

Theorem 2.2.6. *The profinite topology on $\mathbb{Z}_p$ and the topology induced by the p-adic norm coincide.*

Proof. First, let us look at a base for the norm topology, namely the collection $\{\mathrm{B}_\delta(x) \mid \delta > 0, x \in \mathbb{Z}_p\}$ of norm-open balls (where $\mathrm{B}_\delta(x) = \{y \in \mathbb{Z}_p \mid \|x - y\| < \delta\}$, as usual). Since the norm only takes values in $\{p^m \mid m \leqslant 0\}$, we can restrict to looking at values of δ in this set. Let $\delta = p^{-n+1}$ $(n \geqslant 1)$. Then $\mathrm{B}_\delta(x) = \{y \in \mathbb{Z}_p \mid \|x - y\| < p^{-n+1}\} = \{y \in \mathbb{Z}_p \mid \|x - y\| \leqslant p^{-n}\} = \{y \in \mathbb{Z}_p \mid \nu(x - y) \geqslant n\}$. Now note that

$$\begin{aligned} y \in \mathrm{B}_\delta(x) &\iff \nu(x - y) \geqslant n \\ &\iff p^n \mid x - y \\ &\iff x - y \in p^n\mathbb{Z}_p \\ &\iff y \in x + p^n\mathbb{Z}_p \end{aligned}$$

Hence, $\mathrm{B}_\delta(x) = x + p^n\mathbb{Z}_p$.

Now carry out a similar analysis of the profinite topology. First, consider the topology on the whole product $\prod_{i=1}^{\infty} \mathbb{Z}/p^i\mathbb{Z}$. Recall that a subbase for the Tychonoff topology on this is given by the collection $\{\mathrm{B}_i(A_i) \mid A_i \subseteq \mathbb{Z}/p^i\mathbb{Z},\ i > 0\}$, where $\mathrm{B}_i(A_i) = \prod_{j=1}^{i-1} \mathbb{Z}/p^j\mathbb{Z} \times A_i \times \prod_{j=i+1}^{\infty} \mathbb{Z}/p^j\mathbb{Z}$. (Remember that we gave each $\mathbb{Z}/p^i\mathbb{Z}$ the discrete topology, so any subset is open.) Recall how the inverse limit was defined in the proof of 1.1.4. This gives

$$\mathbb{Z}_p = \left\{ (x_i)_{i>0} \in \prod_{i=1}^{\infty} \mathbb{Z}/p^i\mathbb{Z} \;\middle|\; x_{i-1} = x_i\phi_i,\ i > 1 \right\},$$

where $\phi_i : \mathbb{Z}/p^i\mathbb{Z} \to \mathbb{Z}/p^{i-1}\mathbb{Z} : a + p^i\mathbb{Z} \mapsto a + p^{i-1}\mathbb{Z}$. Now take an element $B = \mathrm{B}_i(A_i)$ of the subbase for the topology on the product and note that

$$B = \bigcup_{y \in A_i} B_i^y,$$

where $B_i^y = \mathrm{B}_i(\{y\})$. Now for $y \in A_i$, consider $B_i^y \cap \mathbb{Z}_p$. Writing $y = \sum_{j=0}^{i-1} b_j p^j + p^i\mathbb{Z}$ $(0 \leqslant b_j < p)$, we have[5]

$$B_i^y \cap \mathbb{Z}_p = \left\{ \left(\sum_{j=0}^{k-1} b_j p^j + p^k\mathbb{Z} \right)_{k>1} \;\middle|\; 0 \leqslant b_r < p, \text{ for } r \geqslant i \right\}.$$

That is, the set of elements of the product where the first $i-1$ are constant (they are the reductions of y modulo p^k for $k = 1, \ldots, i$) and the rest are variable, but reduce to y modulo p^i. Now note that we can write this in the form

$$\left(b_0 + p\mathbb{Z}, b_0 + b_1 p + p^2\mathbb{Z}, \ldots, \sum_{j=0}^{i-1} b_j p^j + p^i\mathbb{Z}, \sum_{j=0}^{i-1} b_j p^j + p^{i+1}\mathbb{Z}, \ldots \right) + \left\{ \left(\sum_{j=0}^{k-1} c_j p^j + p^k\mathbb{Z} \right)_{k>1} \;\middle|\; c_r = 0 \text{ for } 0 \leqslant r < i,\ 0 \leqslant c_r < p \text{ for } r \geqslant i \right\},$$

and that we can identify this, in a natural way, with

$$y + p^i\mathbb{Z}_p.$$

We can now identify B with

$$\bigcup_{y \in A_i} y + p^i\mathbb{Z}_p,$$

which is open in the norm topology and hence B is open in the norm topology. Moreover, B_i^y is open (it's an element of a subbase), showing that any element of the base for the norm topology is open in the profinite topology.

Hence, the two topologies on $\mathbb{Z}_p$ coincide, as required. □

In the next section, we shall need another crucial result about the topology on $\mathbb{Z}_p$. This result ties together a lot of the work we've done so far, and follows very easily from the results that we now have at hand.

Theorem 2.2.7. *$\mathbb{Z}_p$ is a compact Hausdorff additive topological group.*

Proof. $\mathbb{Z}_p$ is compact by 1.2.6, Hausdorff as it is a metric space, and an additive topological group by 1.2.7. □

[5]This particular expression doesn't seem to lend itself well to being clearly expressed!

2.3 Integration over $\mathbb{Z}_p$

By 2.2.7, we know that $\mathbb{Z}_p$ is a locally compact additive topological group. This means that we have an additive Haar measure μ on it, which we will normalise, so that $\mu(\mathbb{Z}_p) = 1$ (we can do this since $\mathbb{Z}_p$ is compact, so its measure is finite).

We can find the measure of the subgroup $p^n\mathbb{Z}_p$ by using the fact that cosets partition the group. Let $m = |\mathbb{Z}_p : p^n\mathbb{Z}_p|$ and let $\{a_1, \ldots, a_m\}$ be a transversal (that is, a set of coset representatives) for $p^n\mathbb{Z}_p$ in $\mathbb{Z}_p$. Then

$$\begin{aligned} 1 = \mu(\mathbb{Z}_p) &= \mu\left(\bigcup_{i=1}^{m}(a_i + p^n\mathbb{Z}_p)\right) = \sum_{i=1}^{m}\mu(a_i + p^n\mathbb{Z}_p) \\ &= \sum_{i=1}^{m}\mu(p^n\mathbb{Z}_p) = m\mu(p^n\mathbb{Z}_p), \end{aligned}$$

using the translation invariance of μ. Now x and y represent the same coset if and only if $p^n \mid x - y$, so we see that a set of coset representatives is given by

$$\sum_{i=0}^{n-1} a_i p^i,$$

where $0 \leqslant a_i < p$, as usual. There are p^n of these numbers (p ways to choose each of the n numbers a_i), so that $m = p^n$ and hence

$$\mu(p^n\mathbb{Z}_p) = p^{-n}.$$

Recall from the proof of 2.2.6 that a base for the topology on $\mathbb{Z}_p$ is given by the cosets $a + p^n\mathbb{Z}_p$ for $a \in \mathbb{Z}_p$, and $n > 0$. Hence, we can now find the measure of any open subset of $\mathbb{Z}_p$ by expressing it in terms of this base, and any closed subset by finding the measure of its complement.

Now that we have a measure, we have an integral, as discussed before. Let us look at a couple of examples before proceeding further.

2.3.1 Some examples

Consider

$$I = \int_{\mathbb{Z}_p} \|x\|^s \, \mathrm{d}\mu.$$

We can work this out quite easily by splitting it into a sum of integrals over subsets on which the integrand is constant.

$$\int_{\mathbb{Z}_p} \|x\|^s \, \mathrm{d}\mu = \sum_{i=0}^{\infty} \int_{W_i} \|x\|^s \, \mathrm{d}\mu,$$

where

$$W_i = \{x \in \mathbb{Z}_p \mid \nu(x) = i\}.$$

Hence,

$$I = \sum_{i=0}^{\infty} \int_{W_i} p^{-is} \, \mathrm{d}\mu = \sum_{i=0}^{\infty} p^{-is}\mu(W_i).$$

It now remains to compute $\mu(W_i)$. For this we note that

$$W_i = \left\{ x \in \mathbb{Z}_p \;\middle|\; x = \sum_{j=i}^{\infty} a_j p^j,\ a_i \neq 0 \right\} = p^i\mathbb{Z}_p \backslash p^{i+1}\mathbb{Z}_p.$$

Hence, as $p^{i+1}\mathbb{Z}_p \subseteq p^i\mathbb{Z}_p$,

$$\mu(W_i) = \mu(p^i\mathbb{Z}_p) - \mu(p^{i+1}\mathbb{Z}_p) = p^{-i} - p^{-i-1} = p^{-i-1}(p-1).$$

Finally then, the integral is given by

$$I = \sum_{i=0}^{\infty} p^{-is}p^{-i-1}(p-1) = (p-1)p^{-1}\sum_{i=0}^{\infty}(p^{-s-1})^i = \frac{p-1}{p}\frac{1}{1-p^{-s-1}},$$

for $\Re(s) > -1$. Hence,

$$\int_{\mathbb{Z}_p} \|x\|^s \, \mathrm{d}\mu = \frac{p-1}{p-p^{-s}}, \quad \Re(s) > -1.$$

We can also find integrals of functions of more than one variable — we have a measure on $\mathbb{Z}_p^2$ given by the product measure $\mu \times \mu$. We define $(\mu \times \mu)(A \times B) = \mu(A)\mu(B)$ and extend to the whole of $\mathcal{B}(\mathbb{Z}_p^2)$ using the additivity property of the measure. By uniqueness (and since $(\mu \times \mu)(\mathbb{Z}_p^2) = 1$), we see that this is the additive Haar measure on $\mathbb{Z}_p^2$. Consider the following example.

$$J = \int_{\mathbb{Z}_p^2} \|xy\|^s \, \mathrm{d}\mu,$$

where μ is now the additive Haar measure on $\mathbb{Z}_p^2$. We proceed as above.

$$J = \sum_{i=0}^{\infty} p^{-is}\mu(V_i),$$

where

$$V_i = \{(x,y) \in \mathbb{Z}_p^2 \mid \nu(xy) = i\}.$$

We now use the fact that $\nu(xy) = \nu(x) + \nu(y)$.

$$\begin{aligned}
\mu(V_i) &= \mu(\{(x,y) \in \mathbb{Z}_p^2 \mid \nu(y) = i - \nu(x)\}) \\
&= \sum_{k=0}^{i} \mu(\{(x,y) \in \mathbb{Z}_p^2 \mid \nu(x) = k,\ \nu(y) = i-k\} \\
&= \sum_{k=0}^{i} \mu(p^k\mathbb{Z}_p \backslash p^{k+1}\mathbb{Z}_p \times p^{i-k}\mathbb{Z}_p \backslash p^{i-k+1}\mathbb{Z}_p) \\
&= \sum_{k=0}^{i} p^{-k-1}(p-1)p^{-i+k-1}(p-1) \\
&= (i+1)(p-1)^2p^{-i-2}.
\end{aligned}$$

Substituting back,

$$\begin{aligned} J &= \left(\frac{p-1}{p}\right)^2 \sum_{i=0}^{\infty} (i+1)p^{-is-i} \\ &= \left(\frac{p-1}{p}\right)^2 \sum_{i=0}^{\infty} (i+1)p^{-i(s+1)} \\ &= \left(\frac{p-1}{p}\right)^2 \frac{1}{(1-p^{-s-1})^2}, \quad \Re(s) > -1 \\ &= \left(\frac{p-1}{p-p^{-s}}\right)^2, \quad \Re(s) > -1. \end{aligned}$$

Note that a simpler way to calculate this would have been to use Fubini's theorem (although we have not proved it) and note (once again denoting by μ the additive Haar measure on $\mathbb{Z}_p$) that

$$\int_{\mathbb{Z}_p^2} \|xy\|^s \, \mathrm{d}(\mu \times \mu) = \int_{\mathbb{Z}_p} \int_{\mathbb{Z}_p} \|x\|^s \, \|y\|^s \, \mathrm{d}\mu \, \mathrm{d}\mu = \int_{\mathbb{Z}_p} \|x\|^s \, \mathrm{d}\mu \int_{\mathbb{Z}_p} \|y\|^s \, \mathrm{d}\mu = I^2.$$

The important thing to notice about these two examples is that in both cases, the integrands involve only a monomial in the variables. This is precisely what makes these integrals easy to compute, since we know that $\nu(xy) = \nu(x) + \nu(y)$. p-adic integration gets much more complicated when we consider expressions involving, for example, $x + y$, since there is no easy way to relate $\nu(x + y)$ to $\nu(x)$ and $\nu(y)$. To illustrate this, we shall show how to compute

$$K = \int_{\mathbb{Z}_p^2} \|xy(x+y)\|^s \, \mathrm{d}\mu$$

in a relatively elementary way. We first proceed as before.

$$\begin{aligned} K &= \sum_{i=0}^{\infty} p^{-is} \mu(\{(x,y) \in \mathbb{Z}_p^2 \mid \nu(x) + \nu(y) + \nu(x+y) = i\}) \\ &= \sum_{i=0}^{\infty} \sum_{j=0}^{i} p^{-is} \mu(\{(x,y) \in \mathbb{Z}_p^2 \mid \nu(y) + \nu(x+y) = i-j, \, \nu(x) = j\}) \\ &= \sum_{i=0}^{\infty} \sum_{j=0}^{i} \sum_{k=0}^{i-j} p^{-is} \mu(U_{j,k,i-j-k}), \end{aligned}$$

where

$$U_{r,s,t} = \{(x,y) \in \mathbb{Z}_p \mid \nu(x) = r, \, \nu(y) = s, \, \nu(x+y) = t\},$$

so we see that the calculation boils down to finding the measure of this set. We first note that the problem is symmetric in x and y, so $\mu(U_{r,s,t}) = \mu(U_{s,r,t})$, hence it suffices to consider the case $r \leqslant s$. First consider the case $r < s$, and let $(x,y) \in U_{r,s,t}$. Write

$$\begin{aligned} x &= a_r p^r + a_{r+1} p^{r+1} + \cdots, \\ y &= b_s p^s + b_{s+1} p^{s+1} + \cdots, \end{aligned}$$

with $a_r, b_s \neq 0$ (and $0 \leqslant a_i, b_i < p$). Now

$$x + y = a_r p^r + \cdots + a_{s-1} p^{s-1} + (a_s + b_s) p^s + (a_{s+1} + b_{s+1}) p^{s+1} + \cdots ,$$

so that $\nu(x + y) = r$.

Now consider what happens if $r = s$. We now have

$$x + y = (a_r + b_r) p^r + (a_{r+1} + b_{r+1}) p^{r+1} + \cdots ,$$

so there are two possibilities, namely $p \mid a_r + b_r$ and $p \nmid a_r + b_r$. In the latter case, we have $\nu(x+y) = r$. In the former case, $a_r + b_r = p$ since $0 < a_r + b_r < 2p$. We can then write

$$x + y = (1 + a_{r+1} + b_{r+1}) p^{r+1} + (a_{r+2} + b_{r+2}) p^{r+2} + \cdots ,$$

and we now have to consider whether $1 + a_{r+1} + b_{r+1}$ is divisible by p. If not, $\nu(x + y) = r + 1$, otherwise we have to consider $1 + a_{r+2} + b_{r+2}$ and so on (noting that $0 < 1 + a_i + b_i < 2p$). Motivated by this, we define $n(x, y)$ to be r if $p \nmid a_r + b_r$ and otherwise we define it to be the smallest $j > r$ such that $p \nmid 1 + a_j + b_j$. We claim that $\nu(x + y) = n(x, y)$. If $p \nmid a_r + b_r$ (i.e., $n(x, y) = r$), then we've already seen that $\nu(x + y) = r$. Suppose $p \mid a_r + b_r$, and let $n = n(x, y)$. Then, $a_r + b_r = p$ and $a_i + b_i = p - 1$ for $r < i < n$. Hence, we can write

$$x + y = p p^r + (p - 1) p^{r+1} + \cdots + (p - 1) p^{n-1} + (a_n + b_n) p^n + \cdots ,$$

and we see (by induction) that this simplifies to

$$x + y = (1 + a_n + b_n) p^n + (a_{n+1} + b_{n+1}) p^{n+1} + \cdots .$$

Since we know that $p \nmid 1 + a_n + b_n$, $\nu(x + y) = n$, as required.

Let us summarise our findings.

$$U_{r,s,t} = \begin{cases} \{(x, y) \in \mathbb{Z}_p^2 \mid \nu(x) = r,\, \nu(y) = s\}, & r < s,\, t = r \\ & \text{or } r > s,\, t = s \\ \{(x, y) \in \mathbb{Z}_p^2 \mid \nu(x) = \nu(y) = r,\, n(x, y) = t\}, & r = s \\ \emptyset, & \text{otherwise} \end{cases} .$$

We already know the measure in the first and last cases, so it remains to compute the measure of

$$T_{r,t} = \{(x, y) \in \mathbb{Z}_p^2 \mid \nu(x) = \nu(y) = r,\, n(x, y) = t\}.$$

We shall do the case $r = t$ separately. Write

$$T_{r,r} = \{(x, y) \in \mathbb{Z}_p^2 \mid \nu(x) = \nu(y) = r,\, b_r \neq p - a_r\},$$

and rewrite this as

$$\begin{aligned} T_{r,r} &= \bigcup_{a=1}^{p-1} \{(x, y) \in \mathbb{Z}_p^2 \mid \nu(x) = \nu(y) = r,\, a_r = a,\, b_r \neq p - a\} \\ &= \bigcup_{a=1}^{p-1} \left\{ (x, y) \in \mathbb{Z}_p^2 \;\middle|\; \begin{array}{c} x \in a p^r + p^{r+1} \mathbb{Z}_p, \\ y \in p^r \mathbb{Z}_p \backslash (((p - a) p^r + p^{r+1} \mathbb{Z}_p) \cup p^{r+1} \mathbb{Z}_p) \end{array} \right\}. \end{aligned}$$

Hence,

$$\mu(T_{r,r}) = (p-1)p^{-r-1}(p^{-r} - 2p^{-r-1}) = p^{-2r-2}(p-1)(p-2).$$

Suppose now that $r < t$. Define

$$S^{r,t}_{c_r,\ldots,c_t} = \left\{ (x,y) \in \mathbb{Z}_p^2 \;\middle|\; \begin{array}{c} \nu(x) = \nu(y) = r,\ a_j = c_j \text{ for } r \leqslant j \leqslant t, \\ b_r = p - c_r,\ b_t \neq p - c_t - 1, \\ b_j = p - c_j - 1 \text{ for } r < j < t \end{array} \right\}.$$

Then we can express $T_{r,t}$ in terms of this, as

$$T_{r,t} = \bigcup_{c_r=1}^{p-1} \bigcup_{c_{r+1}=0}^{p-1} \cdots \bigcup_{c_t=0}^{p-1} S^{r,t}_{c_r,\ldots,c_t}.$$

Now,

$$S^{r,t}_{c_r,\ldots,c_t} = \left\{ (x,y) \in \mathbb{Z}_p^2 \;\middle|\; \begin{array}{c} x \in \sum_{j=r}^{t} c_j p^j + p^{t+1}\mathbb{Z}_p, \\ y \in (z + p^t\mathbb{Z}_p) \backslash (z + (p - c_t - 1)p^t + p^{t+1}\mathbb{Z}_p), \\ z = (p - c_r)p^r + \sum_{j=r+1}^{t-1} (p - c_j - 1)p^j \end{array} \right\}.$$

Piecing all this together,

$$\mu(S^{r,t}_{c_r,\ldots,c_t}) = p^{-t-1}(p^{-t} - p^{-t-1}) = p^{-2t-2}(p-1),$$

giving

$$\mu(T_{r,t}) = (p-1)p^{t-r}p^{-2t-2}(p-1) = p^{-t-r-2}(p-1)^2.$$

Thus, we have

$$\mu(U_{r,s,t}) = \begin{cases} p^{-r-s-2}(p-1)^2, & r < s,\ t = r \\ & \text{or } r > s,\ t = s \\ p^{-2r-2}(p-1)(p-2), & r = s = t \\ p^{-t-r-2}(p-1)^2, & r = s,\ t > r \\ 0, & \text{otherwise} \end{cases}.$$

The final step is to substitute this back into the sum. The calculation that follows is a result of [Sid05]. First, rewrite $\mu(U_{r,s,t})$ in terms of i, j, and k.

$$\mu(U_{j,k,i-j-k}) = \begin{cases} p^{-j-k-2}(p-1)^2, & j < k = i - 2j \\ & \text{or } k < j = i - 2k \\ p^{-2j-2}(p-1)(p-2), & j = k = i/3 \\ p^{-i+k-2}(p-1)^2, & j = k < i/3 \\ 0, & \text{otherwise} \end{cases}.$$

In the first case ($j < k = i - 2j$), we have $j < i/3$, and the first two cases ($j < k = i - 2j$ and $k < j = i - 2k$) are symmetric in j and k, so their contribution to the sum will be the same. Writing

$$\chi(i) = \begin{cases} 1, & 3 \mid i \\ 0, & 3 \nmid i \end{cases}$$

and

$$\eta(i) = \max\{n \in \mathbb{Z} \mid n < i/3\},$$

we have the following.

$$\begin{aligned}
&\sum_{j=0}^{i}\sum_{k=0}^{i-j}\mu(U_{j,k,i-j-k})\\
=\;& 2\sum_{j=0}^{\eta(i)}p^{-i+j-2}(p-1)^2+\chi(i)p^{-2i/3-2}(p-1)(p-2)+\sum_{j=0}^{\eta(i)}p^{-i+j-2}(p-1)^2\\
=\;& 3(p-1)^2p^{-i-2}\sum_{j=0}^{\eta(i)}p^j+\chi(i)p^{-2i/3-2}(p-1)(p-2)\\
=\;& 3(p-1)p^{-i-2}(p^{\eta(i)+1}-1)+\chi(i)p^{-2i/3-2}(p-1)(p-2).
\end{aligned}$$

Now write the sum over i as three sums over the congruence classes modulo 3.

$$\begin{aligned}
&\sum_{i=0}^{\infty}\sum_{j=0}^{i}\sum_{k=0}^{i-j}p^{-is}\mu(U_{j,k,i-j-k})\\
=\;& \sum_{n=0}^{\infty}3p^{-3ns}(p-1)p^{-3n-2}(p^n-1)+\sum_{n=0}^{\infty}p^{-3ns}p^{-2n-2}(p-1)(p-2)\\
&+\sum_{n=0}^{\infty}3p^{-(3n+1)s}(p-1)p^{-3n-3}(p^{n+1}-1)\\
&+\sum_{n=0}^{\infty}3p^{-(3n+2)s}(p-1)p^{-3n-4}(p^{n+1}-1)\\
=\;& 3(p-1)p^{-2}\sum_{n=0}^{\infty}p^{-(3s+3)n}(p^n-1)+(p-1)(p-2)p^{-2}\sum_{n=0}^{\infty}p^{-(3s+2)n}\\
&+3(p-1)p^{-s-3}\sum_{n=0}^{\infty}p^{-(3s+3)n}(p^{n+1}-1)\\
&+3(p-1)p^{-2s-4}\sum_{n=0}^{\infty}p^{-(3s+3)n}(p^{n+1}-1)\\
=\;& (p-1)p^{-2}\left(3+p-2+3p^{-s}+3p^{-(2s+1)}\right)\sum_{n=0}^{\infty}p^{-(3s+2)n}\\
&-3(p-1)p^{-2}\left(1+p^{-(s+1)}+p^{-(2s+2)}\right)\sum_{n=0}^{\infty}p^{-(3s+3)n}\\
=\;& \frac{p-1}{p}\left[\frac{1+p+3p^{-s}+3p^{-(2s+1)}}{p(1-p^{-(3s+2)})}-\frac{3+3p^{-(s+1)}+3p^{-(2s+2)}}{p(1-p^{-3(s+1)})}\right]\\
=\;& \frac{p-1}{p}\left[\frac{1+p^{-1}+3p^{-(s+1)}+3p^{-(2s+2)}}{1-p^{-(3s+2)}}-\frac{3p^{-1}}{1-p^{-(s+1)}}\right]\\
=\;& \frac{p-1}{p}\frac{1-2p^{-1}-p^{-(s+2)}+2p^{-(s+1)}}{(1-p^{-(3s+2)})(1-p^{-(s+1)})}.
\end{aligned}$$

Thus, we finally have that the integral is given by

$$\left(1-p^{-1}\right)\left(1-2p^{-1}+2p^{-(s+1)}-p^{-(s+2)}\right)\left(1-p^{-(s+1)}\right)^{-1}\left(1-p^{-(3s+2)}\right)^{-1},$$

which agrees with the result in [dS03]. (Note that the author uses ζ to denote the local factor of the Riemann zeta function at p, not the Riemann zeta function itself. That is, $\zeta(s) = (1 - p^{-s})^{-1}$.)

2.3.2 More advanced techniques

As we have just seen, while p-adic integrals may look innocent at first sight, they get very complicated as soon as we introduce addition into the integrand — any integrand that involves addition is likely to be at least as complicated as the above example.

Given a p-adic integral, we may be interested in either calculating it explicitly, or just obtaining information about it, such as whether or not it is a rational function. Thankfully, there are methods to achieve both the former and the latter. We shall attempt to describe them here, although the details are somewhat beyond our reach — the techniques involve mathematics ranging from algebraic geometry to model theory!

Algebraic geometry comes to the rescue when we need to compute a p-adic integral explicitly. The technique, known as resolution of singularities, breaks the region of integration into several regions over which the integrand becomes a monomial. An example of the use of this technique appears in [dS03], where the integral which we were evaluating above is computed.

If we want to decide whether or not a given p-adic integral is a rational function, we can use the model-theoretic technique from [DvdD88]. This technique is employed in [dS93] to establish the rationality of certain zeta functions of groups.

We will discuss these techniques and their applications to group theory later on.

3 Applications to the theory of groups

3.1 A few more prerequisites

Before embarking on our journey through the theory of zeta functions of groups, we shall need to review a few definitions. First, we define a class of groups which we shall be dealing with frequently, for reasons which will become apparent later.

Definition 3.1.1. *Let G be a group and let $x, y \in G$. Then we define the* ***commutator*** *of x and y by*

$$[x, y] = x^{-1}y^{-1}xy.$$

Given two subsets $A, B \subseteq G$, we define

$$[A, B] = \langle [a, b] \mid a \in A,\, b \in B \rangle,$$

that is, the group generated by commutators of elements of A with elements of B.

The idea of a commutator, and the reason for the name, is that $xy = yx[x, y]$. It's clear that x and y commute if and only if $[x, y] = 1$.

We now have the necessary background to define the class of groups that we need, as well as a related invariant.

Definition 3.1.2. *Let G be a group. We define the* ***lower central series*** *by*

$$\begin{aligned} \gamma_0(G) &= G, \\ \gamma_{i+1}(G) &= [G, \gamma_i(G)], \quad i \geqslant 0. \end{aligned}$$

If there exists an i such that $\gamma_i(G) = 1$, we say that G is ***nilpotent****. If G is nilpotent and c is such that $\gamma_{c+1}(G) = 1$, but $\gamma_c(G) \neq 1$, we call c the* ***nilpotency class*** *of G.*

If G is an abelian group, then for each $x, y \in G$, $[x, y] = 1$, so $\gamma_2(G) = [G, G] = 1$. Conversely, if $[G, G] = 1$, then $[x, y] = 1$ for any $x, y \in G$, so G is abelian. We therefore see that nilpotent groups of class 1 are precisely the non-trivial abelian groups and so the concept of nilpotency generalises the concept of an abelian group.

We shall also make use of the following finiteness condition.

Definition 3.1.3. *Let G be a group. Then G is said to be* ***finitely generated*** *if there exists a finite subset $X \subseteq G$ such that $G = \langle X \rangle$.*

Finally, we have the following.

Definition 3.1.4. *Let G be a group. An element x of G is said to be a* ***torsion*** *element if $x^n = 1$ for some $n > 0$; it is said to be* ***torsion-free*** *otherwise. G is said to be* ***torsion-free*** *if it contains no torsion elements different from the identity.*

3.1.1 Free groups and presentations

Definition 3.1.5. *Let F be a group, and let X be a subset of F. Then F is said to be* ***free*** *on X if for any group G, any map $\phi : X \to G$ extends uniquely to a homomorphism $\hat{\phi} : F \to G$.*

It is not hard to show that two free groups on the same set are isomorphic. (Suppose both F and F' are free on X. We can extend the inclusion $X \hookrightarrow F$ to a homomorphism $\phi : F' \to F$ and the inclusion $X \hookrightarrow F'$ to a homomorphism $\phi' : F \to F'$. Now $\phi\phi'$ fixes X pointwise so, by uniqueness, coincides with the extension of the inclusion $X \hookrightarrow F'$ to F', which coincides with the identity map. Hence $\phi\phi' = 1$ and by symmetry, $\phi'\phi = 1$.) Let us denote the free group on X by $\mathrm{F}(X)$. It is also not hard to show that $\mathrm{F}(X)$ and $\mathrm{F}(Y)$ are isomorphic if and only if $|X| = |Y|$, and we shall denote the free group on a set with d elements by F_d.

We can show that F_n is isomorphic to the group of equivalence classes of words from an alphabet consisting of n symbols and their inverses, where two words are equivalent if we can get from one to the other by inserting xx^{-1} and $x^{-1}x$ into (or deleting them from) the word, where x is a letter in the alphabet. For example, consider F_1. Take the alphabet $\{x\}$. Then it is easy to see that the representatives of the equivalence classes will be x^n for $n \in \mathbb{Z}$, so that $\mathrm{F}_1 \cong \mathbb{Z}$.

Now, given any group G, it is easy to show, using the universal property, that G is a factor of a free group. (Take a generating set X for G, and then extend the inclusion map $X \hookrightarrow G$ to an epimorphism $\mathrm{F}(X) \to G$. Now use the first isomorphism theorem.) Moreover, if G is finitely generated (or, in particular, finite), it is a factor of a finitely generated free group (that is, F_n, for some n). We can therefore describe any group by giving a generating set, taking the free group on that set and taking the quotient by a normal subgroup, where the latter is determined by the relations between the generators. We formalise this concept as follows.

Definition 3.1.6. *Let G be a group and let X be a subset of G. We define the* ***normal closure*** *of X in G to be the intersection of all normal subgroups of G containing X (that is, the smallest normal subgroup of G containing X); it is denoted by $\langle\langle X \rangle\rangle_G$.*

Definition 3.1.7. *Let G be a group. Then G has* ***presentation*** *$\langle X | R \rangle$ if*

$$G \cong \mathrm{F}(X)/\langle\langle R \rangle\rangle_{\mathrm{F}(X)}.$$

This is in fact a very natural concept to define. For example, the dihedral group, D_n, of order $2n$ has presentation

$$\langle a, b \mid a^n, b^2, (ab)^2 \rangle.$$

The words on the right are to be interpreted as the words in the generators which are the identity in the group. A simpler example is given by the cyclic group of order n, C_n, which has presentation

$$\langle c \mid c^n \rangle.$$

This says that C_n has one generator and that the nth power of this generator is the identity.

We can restrict the class (category) of groups in which we are working and work only with abelian groups, say. In this class, we can also find free groups. That is, we can find an abelian group F which is free on $X \subseteq F$, satisfying the universal property that for any *abelian* group G, any map $\phi : X \to G$ extends uniquely to a homomorphism $\hat{\phi} : F \to G$. The free abelian group on d generators is given by the presentation

$$\langle x_1, \ldots, x_d \mid [x_i, x_j],\ 1 \leqslant i, j \leqslant d \rangle.$$

(That is, we have d generators, and the relations say that they all commute.) It is not too hard to see that this is the direct product $\prod_{i=1}^{d} \langle x_i \rangle$ of d copies of the free abelian group on one generator. As we saw above, the latter is simply $\mathbb{Z}$. Therefore, the free abelian group of rank d (that is, on d generators) is $\mathbb{Z}^d$.

In the same way, we can construct free groups in the category of nilpotent groups. Let us denote the free class c, rank d nilpotent group by $\mathrm{F}_{c,d}$. Then, for example, $\mathrm{F}_{2,2}$ is given by

$$\langle x,\ y \mid [x,[x,y]],\ [y,[x,y]] \rangle,$$

or equivalently, by

$$\langle x,\ y,\ z \mid [x,y] = z,\ [x,z] = [y,z] = 1 \rangle$$

(where we interpret the relation $a = b$ as ab^{-1}). The relations here force commutators of three elements to be the identity. In other words, $\gamma_3(\mathrm{F}_{2,2}) = 1$. This group is also known as the **discrete Heisenberg group**, a concrete realisation of which is

$$\left\{ \begin{pmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{pmatrix} \;\middle|\; a,\ b,\ c \in \mathbb{Z} \right\}.$$

As above, we can prove that any nilpotent group of class at most c is a quotient of $\mathrm{F}_{c,d}$ for some d. Also note that $\mathrm{F}_{1,d} \cong \mathbb{Z}^d$.

3.2 $\mathcal{T}$-groups and zeta functions

The study of zeta functions of groups started when Grunewald, Segal, and Smith published their seminal paper [GSS88]. In that paper, they introduced the notion of a zeta function of a group. They restricted to looking at a certain class of groups. We shall the reasons for this restriction later on.

Definition 3.2.1. *Let G be a group. Then G is said to be a **$\mathcal{T}$-group** if it is finitely generated, torsion-free, and nilpotent.*

Lemma 3.2.2. *Let G be a finitely generated group and let n be a positive integer. Then G has only finitely many subgroups of index n.*

Proof. We shall follow the proof given in [Hal50]. Let H be a subgroup of G of index n and let G act on the coset space $(G : H)$ by right multiplication. This gives a homomorphism $\rho : G \to \mathrm{S}_n$. Now, $g \in H$ if and only if g fixes H in this action. By renumbering, we have $H = \{g \in G \mid 1(g\rho) = 1\} = \mathrm{Stab}_G(1)$. H is therefore completely determined by ρ. But if X is a finite generating set for G, then ρ is determined by $x\rho$ for $x \in X$, so we can choose ρ in at most $|S_n|^{|X|} = n!^{|X|}$ ways, so there are at most $n!^{|X|}$ subgroups of index n. □

Armed with this, we can safely make the following definition.

Definition 3.2.3. *Let G be a finitely generated group and let $\mathscr{X}$ be a family of finite index subgroups of G. Then we define*

$$a_n(\mathscr{X}) = |\{H \in \mathscr{X} \mid |G : H| = n\}|,$$

*and we define the **zeta function** of the family $\mathscr{X}$ to be the formal Dirichlet series*

$$\zeta_{\mathscr{X}}(s) = \sum_{H \in \mathscr{X}} |G : H|^{-s} = \sum_{n=1}^{\infty} a_n(\mathscr{X}) n^{-s}.$$

Grunewald, Segal, and Smith considered four different families of subgroups of a group G. We shall be considering the families $\mathscr{S}(G)$ and $\mathscr{N}(G)$ of all subgroups of finite index in G and of all normal subgroups of finite index in G, respectively, and we will write $\zeta_{\mathscr{S}(G)} = \zeta_G^{\leqslant}$, $\zeta_{\mathscr{N}(G)} = \zeta_G^{\trianglelefteqslant}$, $a_n(\mathscr{S}(G)) = a_n^{\leqslant}(G)$ and $a_n(\mathscr{N}(G)) = a_n^{\trianglelefteqslant}(G)$. Grunewald, Segal and Smith also considered the families $\mathscr{J}(G) = \{H \in \mathscr{S}(G) \mid H \cong G\}$ and $\mathscr{H}(G) = \{H \in \mathscr{S}(G) \mid \hat{H} \cong \hat{G}\}$, where $\hat{}$ denotes the profinite completion. The latter was considered for a technical reason, which we shall mention later.

3.2.1 Convergence of zeta functions

Clearly, we would like our Dirichlet series to define a holomorphic function in the complex plane. Given a finitely generated group G, and a family $\mathscr{X}$ of finite index subgroups of G, define

$$s_n(\mathscr{X}) = \sum_{i=1}^{n} a_i(\mathscr{X}).$$

Then we can show that

$$\alpha_{\mathscr{X}} = \limsup_{n \to \infty} \frac{\log s_n(\mathscr{X})}{\log n},$$

is the abscissa of convergence of the zeta function $\zeta_{\mathscr{X}}$ (see [HR15]). That is, $\zeta_{\mathscr{X}}$ defines a holomorphic function in the right half-plane $\{z \in \mathbb{C} \mid \Re(z) > \alpha_{\mathscr{X}}\}$. We can also show that

$$\alpha_{\mathscr{X}} = \inf\{\alpha \geqslant 0 \mid \exists c > 0 : s_n(\mathscr{X}) < cn^{\alpha} \ \forall n > 0\}.$$

If we write $\alpha_G^{\leqslant} = \alpha_{\mathscr{S}(G)}$, then $\alpha_G^{\leqslant}$ is the best possible upper bound for the degree of polynomial subgroup growth of G. If $\alpha_G^{\leqslant}$ is finite, that is, if G has polynomial subgroup growth, then the zeta function $\zeta_G^{\leqslant}$ converges in some right half-plane. Clearly, if $\mathscr{Y} \subseteq \mathscr{X}$, then $\alpha_{\mathscr{Y}} \leqslant \alpha_{\mathscr{X}}$, hence if $\zeta_G^{\leqslant}$ converges, so does $\zeta_G^{\trianglelefteqslant}$, possibly on a larger subset of $\mathbb{C}$.

We can prove that for a $\mathcal{T}$-group G, the number $\alpha_G^{\leqslant}$ is finite. For this, Grunewald, Segal and Smith use an invariant known as the **Hirsch number**. This is defined for a family of groups known as **virtually polycyclic** groups. These are groups which have a polycyclic normal subgroup of finite index, where a polycyclic group is one that has a subnormal series with cyclic factors. If G is virtually polycyclic, then the Hirsch number, denoted $h(G)$, is the maximum

number of infinite cyclic factors in any subnormal series for G. If G is a finitely generated nilpotent group, then we can show that G is virtually polycyclic, and hence has a well-defined Hirsch number. Write $\alpha_G^{\trianglelefteqslant} = \alpha_{\mathscr{N}(G)}$. The following result is part of [GSS88, Proposition 1].

Proposition 3.2.4. *Let G be a $\mathcal{T}$-group. Then $\alpha_G^{\trianglelefteqslant} \leqslant \alpha_G^{\leqslant} \leqslant h(G)$.*

We have therefore established that for a $\mathcal{T}$-group G, $\zeta_G^{\leqslant}$ defines a holomorphic function in the right half-plane $\{z \in \mathbb{C} \mid \Re(z) > \alpha_G^{\leqslant}\}$ and $\zeta_G^{\trianglelefteqslant}$ defines a holomorphic function in $\{z \in \mathbb{C} \mid \Re(z) > \alpha_G^{\trianglelefteqslant}\}$.

3.2.2 Euler products

Although we can see from the above that we can define zeta functions for arbitrary finitely generated groups, there is a good reason for restricting attention to nilpotent such groups. The reason is a simplification afforded by the following fact.

Proposition 3.2.5. *Let G be a finite group. Then G is nilpotent if and only if it is the direct product of its Sylow p-subgroups.*

From this, we can deduce the following.

Lemma 3.2.6. *Let G be a finitely generated nilpotent group, let n be a positive integer with prime factorisation $\prod_{i=1}^{k} p_i^{e_i}$, and let $* \in \{\leqslant, \trianglelefteqslant\}$. Then,*

$$a_n^*(G) = \prod_{i=1}^{k} a_{p_i^{e_i}}^*(G).$$

We shall only give a proof for $* = \trianglelefteqslant$. The proof for $* = \leqslant$ is not quite as simple, and involves looking at the profinite completion of the group. However, the proof has the advantage that it proves both cases at the same time. We refer the reader to [GSS88, Proposition 1.3].

Proof. Let $* = \trianglelefteqslant$. We shall follow the approach taken in [dS00]. Let N be a normal subgroup of G of finite index n. Then G/N is a finite nilpotent group, so by 3.2.5,

$$G/N \cong \prod_{i=1}^{k} P_i/N,$$

where each P_i/N is a Sylow p_i-subgroup of G/N, that is, $|P_i : N| = p_i^{e_i}$. If we let

$$Q_i/N = \prod_{\substack{j=1 \\ j \neq i}}^{k} P_j/N,$$

then

$$G/Q_i \cong \frac{G/N}{Q_i/N} \cong P_i/N,$$

hence $|G : Q_i| = p_i^{e_i}$. Conversely, given subgroups Q_i of G of index $p_i^{e_i}$, set $Q = \bigcap_{i=1}^{k} Q_i$ and we have

$$G/Q \cong \prod_{i=1}^{k} G/Q_i,$$

so $|G : Q| = n$. Now note that $G/N \cong G/M$ if and only if the corresponding Sylow subgroups are isomorphic. This gives us the bijection we were looking for, and hence we have the result for $* = \trianglelefteqslant$. □

This gives the simplification we were looking for ([GSS88, Proposition 4]).

Theorem 3.2.7. *Let G be a $\mathcal{T}$-group and let $* \in \{\leqslant, \trianglelefteqslant\}$. Then*

$$\zeta^*_G(s) = \prod_{p \text{ prime}} \zeta^*_{G,p}(s),$$

as a product of a formal Dirichlet series, convergent for $\Re(z) > \alpha^*_G$, where

$$\zeta^*_{G,p}(s) = \sum_{i=0}^{\infty} a^*_{p^i}(G) p^{-is}.$$

$\prod_{p \text{ prime}} \zeta^*_{G,p}$ is known as the **Euler product** of the zeta function and each $\zeta^*_{G,p}$ is called a **local factor**.

Recall the families of subgroups $\mathscr{J}(G)$ and $\mathscr{H}(G)$ considered by Grunewald, Segal, and Smith. It turns out that the zeta function $\zeta_{\mathscr{J}(G)}$ does not have an Euler product decomposition while $\zeta_{\mathscr{H}(G)}$ does. This is due to a version of the Chinese remainder theorem which says that for a nilpotent group G, $\hat{G} = \prod_{p \text{ prime}} \hat{G}_p$.

Let us now look at an example. (This example is considered both in [dS03] and in [GSS88], and the proof below is taken from the former.) In everything that follows we use ζ to denote the Riemann zeta function.

Proposition 3.2.8. *Let d be a positive integer. Then,*

$$\zeta^{\trianglelefteqslant}_{\mathbb{Z}^d}(s) = \zeta^{\leqslant}_{\mathbb{Z}^d}(s) = \prod_{i=0}^{d-1} \zeta(s-i).$$

Proof. Any finite index additive subgroup of $\mathbb{Z}^d$ is of the form $\mathbb{Z}^d\mathsf{A}$ for some $\mathsf{A} \in \mathrm{M}_d(\mathbb{Z})$. Moreover, $|\mathbb{Z}^d : \mathbb{Z}^d\mathsf{A}| = |\det \mathsf{A}|$. Two matrices A and B give rise to the same subgroup if and only if they lie in the same left coset of $\mathrm{GL}(d, \mathbb{Z})$. The set of upper triangular matrices (a_{ij}) with $0 \leqslant a_{ij} < a_{jj}$ for each i and j gives a set $\mathcal{A}$ of left coset representatives. Let $\mathcal{A}(a_{11}, \ldots, a_{dd})$ denote the set of matrices in $\mathcal{A}$ with fixed diagonal entries $a_{11}, \ldots, a_{dd}$. We now have

$$\begin{aligned} \zeta^{\leqslant}_{\mathbb{Z}^d}(s) &= \sum_{\mathsf{A} \in \mathcal{A}} |\det \mathsf{A}|^{-s} \\ &= \sum_{a_{11}=1}^{\infty} \cdots \sum_{a_{dd}=1}^{\infty} a_{11}^{-s} \cdots a_{dd}^{-s} |\mathcal{A}(a_{11}, \ldots, a_{dd})|. \end{aligned}$$

We use a simple counting argument to find the size of $\mathcal{A}(a_{11}, \ldots, a_{dd})$. There are a_{dd} choices for each entry in the last column, giving a_{dd}^{d-1} choices for this column. Similarly, there are $a_{(d-1)(d-1)}$ choices for each entry in the penultimate column, giving a total of $a_{(d-1)(d-1)}^{d-2}$ choices for that column. This argument gives $|\mathcal{A}(a_{11}, \ldots, a_{dd})| = a_{22} a_{33}^2 \cdots a_{dd}^{d-1}$. Substituting this back gives

$$\begin{aligned} \zeta^{\leqslant}_{\mathbb{Z}^d}(s) &= \sum_{a_{11}=1}^{\infty} \cdots \sum_{a_{dd}=1}^{\infty} a_{11}^{-s} a_{22}^{-s+1} \cdots a_{dd}^{-s+d-1} \\ &= \zeta(s)\zeta(s-1)\cdots\zeta(s-d+1), \end{aligned}$$

as required. Clearly, $\zeta^{\leqslant}_{\mathbb{Z}^d} = \zeta^{\trianglelefteqslant}_{\mathbb{Z}^d}$ since the group is abelian. □

Now, according to 3.2.7, the zeta function of $\mathbb{Z}^d$ should have an Euler product decomposition. Indeed, it does. Recall the "original" Euler product for the Riemann zeta function,

$$\zeta(s) = \prod_{p \text{ prime}} \frac{1}{1-p^{-s}}.$$

This immediately gives the local factors of the zeta function of $\mathbb{Z}^d$ as

$$\zeta^{\trianglelefteqslant}_{\mathbb{Z}^d,p}(s) = \zeta^{\leqslant}_{\mathbb{Z}^d,p}(s) = \prod_{i=0}^{d-1} \frac{1}{1-p^{-s+i}}.$$

3.2.3 Rationality and uniformity of zeta functions

Notice a rather special property of the above example. If we define

$$\Psi(X,Y) = \prod_{i=0}^{d-1} \frac{1}{1-X^iY},$$

then for each prime p,

$$\zeta^{\trianglelefteqslant}_{\mathbb{Z}^d,p}(s) = \zeta^{\leqslant}_{\mathbb{Z}^d,p}(s) = \Psi(p,p^{-s}).$$

In other words, we have a single *rational* function $\Psi(X,Y) \in \mathbb{Q}(X,Y)$ which, with the appropriate values of X and Y, gives every local factor.

Inspired by this example (and others in [GSS88]), we make the following definition.

Definition 3.2.9. *Let G be a $\mathcal{T}$-group and let $* \in \{\trianglelefteqslant, \leqslant\}$. We say that the zeta function ζ^*_G of G is **finitely uniform** if there exist a positive integer n and rational functions $\Psi_1(X,Y), \ldots, \Psi_n(X,Y) \in \mathbb{Q}(X,Y)$ such that for each prime p,*

$$\zeta^*_{G,p}(s) = \Psi_i(p,p^{-s})$$

*for some i. If $n = 1$, we say that the zeta function is **uniform**.*

In this terminology we can say that the zeta functions of $\mathbb{Z}^d$ are uniform.

This naturally leads us to ask which $\mathcal{T}$-groups have zeta functions with rational local factors, and which of those have finitely uniform or uniform zeta functions. It turns out that the answer to the first of these questions is provided by the theory of p-adic integration. More specifically, the answer is provided by a class of integrals introduced in [dSG00].

Definition 3.2.10. *Let $\mathbf{x} = (x_1, \ldots, x_m)$ and let $f_i(\mathbf{x})$, $g_i(\mathbf{x}) \in \mathbb{Q}[\mathbf{x}]$ for $1 \leqslant i \leqslant l$. Then we say that the formula*

$$\psi(\mathbf{x}) = \bigwedge_{i=1}^{l} (\nu(f_i(\mathbf{x})) \leqslant \nu(g_i(\mathbf{x})))$$

*is a **cone condition** over $\mathbb{Q}$. Given a cone condition $\psi(\mathbf{x})$ and $f_0(\mathbf{x})$, $g_0(\mathbf{x}) \in \mathbb{Q}[\mathbf{x}]$, we say that*

$$Z_{\mathcal{D}}(s,p) = \int_{V_p} \|f_0(\mathbf{x})\|^s \, \|g_0(\mathbf{x})\| \, \mathrm{d}\mu$$

is a ***cone integral*** *defined over* $\mathbb{Q}$*, where* $V_p = \{\mathbf{x} \in \mathbb{Z}_p^m \mid \psi(\mathbf{x})\}$*,* μ *is the normalised additive Haar measure on* $\mathbb{Z}_p^m$ *and* $\mathcal{D} = \{(f_i, g_i) \mid 1 \leqslant i \leqslant l\}$ *is the* ***cone integral data****.*

The fundamental result which was proved in [dSG00] about this class of integrals is the following.

Theorem 3.2.11. *Let* $Z_{\mathcal{D}}$ *be a cone integral. Then there exists a rational function* $\Psi(X, Y) \in \mathbb{Q}(X, Y)$ *such that for almost all primes* p*,*

$$Z_{\mathcal{D}}(s, p) = \Psi(p, p^{-s}).$$

In fact, the complete statement of this theorem ([dSG00, Theorem 1.2]) gives more precise information about the rational functions occurring, in terms of varieties (in fact, schemes). (A variety is, roughly speaking, a subset of some space defined by polynomial equations.) The proof of the theorem depends on using resolution of singularities to break the space over which we are integrating into several varieties, and does not rely on the model-theoretic results of [DvdD88] which were used to establish such facts about rationality before [dSG00]. In [DvdD88], the authors define a first-order language and prove that if a p-adic integral is definable in the language, it must be a rational function. The latter technique is employed in earlier papers, in particular [dS00] and [dS93].

Recall the proof of 3.2.8. This was relatively easy as we could count subgroups by counting matrices. In a more general $\mathcal{T}$-group, things are not so simple, and this is where the theory of Lie algebras comes in. There is a somewhat deep result known as the Mal'tsev correspondence which associates to any finitely generated torsion-free nilpotent group G (that is, a $\mathcal{T}$-group), nilpotent[6] Lie algebras $\mathfrak{g}_p$ over the ring $\mathbb{Z}_p$, one for each prime p. These Lie algebras have the property that there exists a one-to-one index-preserving correspondence between ideals in $\mathfrak{g}_p$ and normal subgroups of p-power index in G. We can define the zeta function counting ideals in a Lie algebra in the natural way, and we then have

$$\zeta_{G,p}^{\trianglelefteqslant}(s) = \zeta_{\mathfrak{g}_p}^{\trianglelefteqslant}(s).$$

For a general $\mathcal{T}$-group, there may be a finite number of primes where this correspondence breaks down.

Now, since Lie algebras are linear structures, we can count ideals in them by counting matrices, as we did for $\mathbb{Z}^d$. However, we still lack a powerful enough theory of normal forms of matrices to be able to do this easily. Instead, though, we can use p-adic integration to obtain some information about $\zeta_{\mathfrak{g}_p}$.

Fix a prime p and consider the example given in [dS03], namely $G = \mathrm{F}_{2,2}$. Let $\mathfrak{g}_p$ be the $\mathbb{Z}_p$-Lie algebra given by the Mal'tsev correspondence. It turns out that in this case, the correspondence works for all primes. Here, an ideal $\mathfrak{i}$ of $\mathfrak{g}_p$ is an additive subgroup of $\mathbb{Z}_p^3$ (as $\mathfrak{g}_p$ has three generators), so there is a matrix $\mathsf{A} \in \mathrm{M}_3(\mathbb{Z}_p)$ such that $\mathfrak{i} = \mathbb{Z}_p^3\mathsf{A}$. For this ideal, consider the set of upper triangular matrices which generate it, namely $\mathcal{M}(\mathfrak{i}) = \{\mathsf{A} \in \mathrm{Tr}_3(\mathbb{Z}_p) \mid \mathfrak{i} = \mathbb{Z}_p^3\mathsf{A}\}$. To find the index of the ideal $\mathfrak{i}$, we integrate the determinant over $\mathcal{M}(\mathfrak{i})$ and divide by the measure of $\mathcal{M}(\mathfrak{i})$.

$$\zeta_{\mathfrak{g}_p}^{\trianglelefteqslant}(s) = \sum_{\mathfrak{i} \trianglelefteqslant \mathfrak{g}_p} |\mathfrak{g}_p : \mathfrak{i}|^{-s} = \sum_{\mathfrak{i} \trianglelefteqslant \mathfrak{g}_p} \int_{\mathsf{A} \in \mathcal{M}(\mathfrak{i})} \|\det \mathsf{A}\|^{-s} \, \mu(\mathcal{M}(\mathfrak{i}))^{-1} \, \mathrm{d}\mu,$$

[6] The definition of a nilpotent Lie algebra is the same as that for a group, but with commutators replaced by Lie brackets.

where μ is the additive Haar measure on $\mathrm{Tr}_3(\mathbb{Z}_p) \cong \mathbb{Z}_p^6$ (as an additive group). We can show that the measure of $\mathcal{M}(\mathfrak{i})$ is given by

$$\mu(\mathcal{M}(\mathfrak{i})) = (1-p^{-1})^3 \left\|a_{11}\right\| \left\|a_{22}\right\|^2 \left\|a_{33}\right\|^3,$$

where $(a_{ij}) \in \mathcal{M}(\mathfrak{i})$ is a representative matrix for this subset, and that

$$\left\|\det \mathsf{A}\right\| = \left\|a_{11}\right\|^{-1} \left\|a_{22}\right\|^{-1} \left\|a_{33}\right\|^{-1}.$$

We therefore have that $\zeta^{\triangleleft}_{\mathfrak{g}_p}(s)$ is given by

$$\sum_{\mathfrak{i} \trianglelefteq \mathfrak{g}_p} \int_{\mathsf{A} \in \mathcal{M}(\mathfrak{i})} \left\|a_{11}\right\|^s \left\|a_{22}\right\|^s \left\|a_{33}\right\|^s (1-p^{-1})^{-3} \left\|a_{11}\right\|^{-1} \left\|a_{22}\right\|^{-2} \left\|a_{33}\right\|^{-3} \mathrm{d}\mu$$
$$= (1-p^{-1})^{-3} \int_{\mathsf{A} \in \mathcal{M}} \left\|a_{11}\right\|^{s-1} \left\|a_{22}\right\|^{s-2} \left\|a_{33}\right\|^{s-3} \mathrm{d}\mu,$$

where $\mathcal{M} = \bigcup_{\mathfrak{i} \trianglelefteq \mathfrak{g}_p} \mathcal{M}(\mathfrak{i})$.

We have thus expressed the zeta function of the Lie algebra $\mathfrak{g}_p$, and hence the local factor of the zeta function of $\mathrm{F}_{2,2}$ at p, as a p-adic integral.

In order to count finite groups, we need to refine this zeta function to take account of the fact that different normal subgroups of $\mathrm{F}_{2,2}$ may give rise to isomorphic quotients. In fact, M and N give rise to isomorphic quotients if and only if there is an automorphism ϕ of $\mathrm{F}_{2,2}$ such that $M\phi = N$. Thus, we need to count the sizes of orbits of this action of $\mathrm{Aut}(\mathrm{F}_{2,2})$ on the family (lattice) of subgroups of $\mathrm{F}_{2,2}$. This is a simple application of the orbit-stabiliser theorem and gives the following expression for the zeta function counting non-isomorphic quotients.

$$\tilde{\zeta}^{\triangleleft}_{\mathrm{F}_{2,2}}(s) = \sum_{N \trianglelefteqslant \mathrm{F}_{2,2}} |\mathrm{F}_{2,2} : N|^{-s} |\mathrm{Aut}(\mathrm{F}_{2,2}) : \mathrm{Stab}_{\mathrm{Aut}(\mathrm{F}_{2,2})}(N)|^{-1}.$$

We can perform a similar analysis on this new zeta function, expressing it in terms of counting ideals in $\mathfrak{g}_p$ and subsequently as a p-adic integral.

We can do the same for a general free nilpotent group $\mathrm{F}_{c,d}$, and we can show that the integrals arising are in fact cone integrals. We can therefore apply 3.2.11 to conclude that $\zeta^{\triangleleft}_{\mathrm{F}_{c,d},p}(s)$ and $\tilde{\zeta}^{\triangleleft}_{\mathrm{F}_{c,d},p}(s)$ are rational in p and p^{-s}. (See also [dS00, Theorem 3.18].)

Returning to the question of uniformity, the following conjecture was made in [GSS88].

Conjecture 3.2.12. *The zeta functions $\zeta^{\triangleleft}_{\mathrm{F}_{c,d}}$ are uniform.*

This conjecture was shown to be true for class 2 free groups in the same paper.

Theorem 3.2.13. *The zeta functions of $\mathrm{F}_{2,d}$ are uniform.*

As an illustration, consider the example of $\mathrm{F}_{2,2}$ above. The integral can be computed explicitly to give

$$\zeta_{\mathrm{F}_{2,2}}(s) = \frac{\zeta(s)\zeta(s-1)\zeta(2s-2)\zeta(2s-3)}{\zeta(3s-3)}.$$

Once again, using the "original" Euler product, we get an Euler product for this zeta function, and hence the zeta function is uniform, as guaranteed by 3.2.13.

They also asked whether the zeta functions of a more general group are finitely uniform. The answer to this questions turned out to be negative, as shown in [dS01].

Theorem 3.2.14. *Let G be the class 2, Hirsch number 9 group with presentation*

$$\left\langle \begin{array}{c} x_1, x_2, x_3, x_4, x_5, x_6, y_1, y_2, y_3 \mid [x_1, x_4] = y_3,\ [x_1, x_5] = y_1, \\ [x_1, x_6] = y_2,\ [x_2, x_4] = y_1,\ [x_2, x_5] = y_3,\ [x_3, x_4] = y_2,\ [x_3, x_6] = y_1 \end{array} \right\rangle,$$

where all other commutators are defined to be 1. Then the zeta functions $\zeta_G^{\trianglelefteq}$ and $\zeta_G^{\leqslant}$ are not finitely uniform.

More precisely, if we let E be the elliptic curve $y^2 = x - x^3$, then there are rational functions $\Phi_1(X, Y)$, $\Phi_2(X, Y) \in \mathbb{Q}(X, Y)$ such that for almost all primes p,

$$\zeta_{G,p}^{\trianglelefteq}(s) = \Phi_1(p, p^{-s}) + |E(\mathbb{F}_p)|\Phi_2(p, p^{-s}),$$

where $|E(\mathbb{F}_p)| = |\{(x, y) \in \mathbb{F}_p^2 \mid y^2 = x - x^3\}|$ (the number of points on E modulo p). To see how this elliptic curve arises from the presentation, consider the determinant of the 3×3 matrix $([x_i, x_{j+3}])$.

$$\begin{vmatrix} [x_1, x_4] & [x_2, x_4] & [x_3, x_4] \\ [x_1, x_5] & [x_2, x_5] & [x_3, x_5] \\ [x_1, x_6] & [x_2, x_6] & [x_3, x_6] \end{vmatrix} = \begin{vmatrix} y_3 & y_1 & y_2 \\ y_1 & y_3 & 0 \\ y_2 & 0 & y_1 \end{vmatrix} = -y_2^2 y_3 + y_1 y_3^2 - y_1^3,$$

which is the projective version of E — if we set $y_1 = x$, $y_2 = y$, and $y_3 = 1$, we get $-y^2 + x - x^3$.

This counterexample, therefore, relies on the fact that $|E(\mathbb{F}_p)|$ varies wildly with the prime p. In particular, the behaviour of $|E(\mathbb{F}_p)|$ is not in line with the behaviour of counting p-groups predicted by the PORC conjecture.

3.2.4 PORC via zeta functions

Recall the PORC conjecture, 0.2.1, from the introduction. The approach suggested there was to look at the local factors of the normal zeta functions of free nilpotent groups. From the above we can see that we actually need to look at the functions $\tilde{\zeta}_{\mathrm{F}_{c,d},p}^{\trianglelefteq}$ to avoid the overcounting problem.

In order to prove PORC, we need to understand what these functions look like. Let us make the following conjecture, as in [dS03].

Conjecture 3.2.15. *For each c and d, there exist a positive integer N, and rational functions $\Psi_1(X, Y), \ldots, \Psi_N(X, Y) \in \mathbb{Q}(X, Y)$ such that if $p \equiv i \pmod N$, then*

$$\tilde{\zeta}_{\mathrm{F}_{c,d},p}^{\trianglelefteq}(s) = \Psi_i(p, p^{-s}).$$

Moreover, these rational functions have the form

$$\Psi_i(X, Y) = \frac{P_i(X, Y)}{(1 - X^{a_{i1}} Y^{b_{i1}}) \cdots (1 - X^{a_{id_i}} Y^{b_{id_i}})},$$

for some polynomial $P_i(X, Y) \in \mathbb{Q}[X, Y]$.

Since a group of order p^n is nilpotent of class at most $n-1$ and can be generated by a set of size at most n, it is not too hard to see that the truth of this conjecture would imply PORC.

However, even though we have 3.2.13 and even if 3.2.12 is true, this is not yet enough to establish 3.2.15, since the uniformity of zeta functions of free nilpotent groups does not immediately tell us about the uniformity of the zeta functions counting quotients up to isomorphism.

However, we can once again use p-adic integrals to analyse these zeta functions. As we mentioned, the integrals arising can be shown to be cone integrals. We can therefore use the full version of 3.2.11, [dSG00, Theorem 1.2], to extract the following information about $f(p,n)$, the number of isomorphism classes of groups of order p^n.

Theorem 3.2.16. *For each integer n, there exist finitely many subvarieties $\{E_{i,n}\}_{i\in T(n)}$ of a variety Y_n over $\mathbb{Q}$, and for each $I \subseteq T(n)$, a polynomial $H_{n,I}(X) \in \mathbb{Q}[X]$ such that for almost all primes p,*

$$f(p,n) = \sum_{I\subseteq T(n)} e_{n,p,I} H_{n,I}(p),$$

where

$$e_{n,p,I} = |\{a \in \bar{Y}_n(\mathbb{F}_p) \mid a \in \bar{E}_{i,n}(\mathbb{F}_p) \iff i \in I\}|$$

(where $\bar{X}$ denotes the reduction of the variety X modulo p).

This is not quite as strong as PORC, but is still quite a spectacular result. It essentially says that we can count p-power index subgroups of free nilpotent groups giving rise to non-isomorphic quotients, and hence isomorphism classes of finite p-groups, by counting points modulo p on certain varieties. It provides a rather surprising link between group theory and algebraic geometry. (This is what du Sautoy refers to in [dS03] as NOPOV — number of points on varieties, pointing out that this is not quite as catchy as Higman's culinary shorthand.)

This also suggests another way of approaching PORC, namely by analysing the nature of the varieties that occur. However, the varieties arising from counting subgroups can be quite exotic, as seen in the example of 3.2.14.

At the time of writing, PORC has still not been established, even for class 2 groups. Although the approaches outlined above look promising, there is still a lot more work to be done.

For further reading on the subjects of PORC and zeta functions of groups, we refer the reader to the survey paper [dS03] and the more mathematical introduction [dS00].

References

[BF93] L. Brekke and P. G. O. Freund, *p-adic numbers in physics*, Physics Reports (Review Section of Physics Letters) **233** (1993), no. 1, 1–66.

[Bou89] N. Bourbaki, *General topology, chapters 1–4*, Springer-Verlag, 1989.

[dS93] M. P. F. du Sautoy, *Finitely generated groups, p-adic analytic groups and Poincaré series*, Ann. of Math. **137** (1993), no. 3, 639–670.

[dS00] ———, *Counting finite p-groups and nilpotent groups*, Inst. Hautes Études Sci. Publ. Math. (2000), no. 92, 63–112 (2001).

[dS01] ———, *A nilpotent group and its elliptic curve: non-uniformity of local zeta functions of groups*, Israel J. Math. **126** (2001), 269–288.

[dS03] ———, *Zeta functions of groups: the quest for order versus the flight from ennui*, Groups St. Andrews 2001 in Oxford. Vol. I, London Math. Soc. Lecture Note Ser., vol. 304, Cambridge Univ. Press, 2003, pp. 150–189.

[dSG00] M. P. F. du Sautoy and F. J. Grunewald, *Analytic properties of zeta functions and subgroup growth*, Ann. of Math. (2) **12** (2000), no. 3, 793–833.

[DvdD88] J. Denef and L. van den Dries, *p-adic and real subanalytic sets*, Ann. of Math. **128** (1988), no. 1, 79–138.

[GSS88] F. J. Grunewald, D. Segal, and G. C. Smith, *Subgroups of finite index in nilpotent groups*, Invent. math. **93** (1988), 185–223.

[Hal50] M. Hall, Jr., *A topology for free groups and related groups*, Ann. of Math. **52** (1950), no. 1, 127–139.

[Hal59] P. R. Halmos, *Measure theory*, D. van Nostrand Company, 1959.

[HR15] G. H. Hardy and M. Riesz, *The general theory of Dirichlet's series*, Cambridge Univ. Press, 1915.

[Mor04] K. Morley, *Topology: the mystery math*, `http://users.ox.ac.uk/~mert1113/topology.shtml`, 2004.

[NOVL04] M. F. Newman, E. A. O'Brien, and M. R. Vaughan-Lee, *Groups and nilpotent Lie rings whose order is the sixth power of a prime*, J. Algebra **278** (2004), no. 1, 383–401.

[OVL05] E. A. O'Brien and M. R. Vaughan-Lee, *The groups with order p^7 for odd prime p*, to appear in J. Algebra (2005), `http://www.math.auckland.ac.nz/~obrien/research/p7/paper-p7.pdf`.

[Rot79] J. J. Rotman, *An introduction to homological algebra*, Academic Press, 1979.

[Sid05] N. Sidorova, 2005, private communication.

[Wil70] S. Willard, *General topology*, Addison-Wesley, 1970.