The Cohen-Lenstra Heuristic for Finite Abelian Groups

Dissertation
zur Erlangung des Grades
des Doktors der Naturwissenschaften

vorgelegt
der Mathematisch-Naturwissenschaftlichen Fakultät I
der Universität des Saarlandes

von
Johannes Lengler

Saarbrücken,
Juli 2009

Tag des Kolloquiums:	16. Oktober 2009
Dekan:	Prof. Dr. Joachim Weickert

Prüfungsausschuss:

Vorsitzender:	Prof. Dr. Jörg Eschmeier
Gutachter:	Prof. Dr. Ernst-Ulrich Gekeler
	Prof. Dr. Rainer Schulze-Pillot-Ziemen
akademischer Mitarbeiter:	Dr. Dominik Faas

weiterer Gutachter:	Prof. Dr. Karim Belabas

Abstract

In the last decades, a method has gained ever-increasing influence which treats deterministic objects as if they were random objects and studies them with probability theoretic means. A major breakthrough for this method came in 1984, when Henri Cohen and Hendrik W. Lenstra noticed that the sequence of class groups of quadratic number fields behaves essentially like a random sequence with respect to a certain probability distribution on the space of all finite abelian groups.
Later on, it turned out that this distribution occurs also in many other contexts and plays the role of a "natural" distribution, regulating the structure of finite abelian groups in all situations where no obvious structural obstacles for a random-like behaviour exist.
This thesis is devoted to studying this "Cohen-Lenstra heuristic".

I will

- explain and motivate its fundamental assumption, namely that the larger the automorphism group of a group is, the less likely it should appear.
- review the basic formulas that were worked out by the pioneers of this topic.
- point out and study a deep connection between the Cohen-Lenstra probability measure and partitions.
- list methods for studying the distribution that were obtained by various research groups; also list formulas about interesting group theoretic quantities from a probabilistic point of view.
- show the difference between the local (i.e., for p-groups) and the global Cohen-Lenstra measure. The global case is considerably more difficult, due to convergence reasons. I give a solution for the fundamental problem underlying the global case.
- give applications of the Cohen-Lenstra measure to various fields of mathematics. My main focus will be on number theory, but I also mention applications from other mathematical areas.

Contents

Chapter 0

Introduction

Motivation: Randomness in real numbers

It is easy to show that there is no equidistributed probability measure on the real numbers or on the integers. Yet we have plenty of examples where such numbers do occur in nature:

- Data about our environment: The lengths of rivers, the weights of planets, ..., are reals. Other data like the number of inhabitants of cities are integers.

- Results from experiments in physics are typically reals (such as mass, speed, ...). However, since some physical parameters like electric charge are quantized, we also have integers occurring in physical context.

Whenever we can measure a parameter and if we can do this without any knowledge about the result, the parameter must be distributed with respect to some probability measure. As said above, this measure cannot be an equidistribution because such a distribution does not exist. But what is the distribution then?
It turns out that it is impossible to assign a non-zero probability to a real number or an integer in a way that fits the data. However, there are some subsets of the reals and the integers that appear to have a well-defined measure. The most prominent example was discovered by Simon Newcomb [New81] in 1881 and rediscovered by Frank Benford ([Ben38]) in 1938. They came to the conclusion that the set of all integers (or reals) with first digit 1 has measure $\log_{10}(2)$ (and not, as one might expect, $\frac{1}{9}$), and similarly for other digits. This is known as "Benford's law".

The exact value is already determined if we assume the distribution to be scaling invariant, i.e., if we require the distribution to stay the same when changing our units, e.g., from meter to inches.
The exciting feature of these prognoses is that they fit the real-world data stunningly well. All the examples listed above have been investigated. In all cases, the first digits are in perfect accordance with Benford's law. It seems that we have found a universal property of nature. The concurrence seems to hold for all physical parameters that come along with a non-trivial unit (e.g. meter, gram, newton), and for many others.
Note however that the law fails if the domain is not all of the integers but only a finite interval. If you investigate a telephone book, you will find a different pattern, since telephone numbers are bounded (e.g. from 100000 to 999999).

We see that although there is no total probability measure on $\mathbb{R}$ or $\mathbb{Z}$, it is possible to assign probabilities to certain subsets. These probabilities obviously reflect a deep principle of nature, so we might speak of a "natural" distribution. This thesis is dedicated to the study of a similar "natural" distribution (the Cohen-Lenstra distribution) — not of numbers, but rather of finite abelian groups.
The basic idea of this distribution is that each group must occur with a probability that is inverse proportional to the size of its automorphism group. I will introduce and motivate this distribution and point out the obstacles that arise. (We will see that in fact it is not a probability distribution on the whole power set of the set of all finite abelian groups — but as with Benford's law, we will find that we can make some sensible predictions if we restrict ourselves to certain subsets.) Also, I will prove some properties of the Cohen-Lenstra distribution.
In a seminar in our Zahlentheorie AG (number theory group), Prof. Ernst-Ulrich Gekeler, Bernd Mehnert, and myself discovered a mysterious connection between the distribution and partitions of natural numbers. We conjectured that there might be a link and we determined the nature of this link. This was the starting point for the work on this thesis.

Structure of the thesis
The structure of the thesis is as follows: In chapter 1, I will fix some notation and revise some common statements about finite abelian groups, and about partitions. Also, I will prove some preparatory, group-theoretical lemmas. In chapter 2, I will introduce the local Cohen-Lenstra heuristic and give several reasons why the heuristic is natural.

Chapter 3 and 4 are devoted to presenting tools in order to work with the Cohen-Lenstra heuristic. These tools should help us

(i) to compute interesting values about the Cohen-Lenstra heuristic, and

(ii) to prove that some interesting sequences of groups really behave like random groups with respect to this heuristic.

The method in chapter 3 relies on the connection between the Cohen-Lenstra heuristic and partitions mentioned above, and is new. The methods in chapter 4 are not my work, but most of them are unknown to number theorists. This strange fact is due to the incident that the very Cohen-Lenstra heuristic has been studied in a completely different context by group theorists without anyone (neither group theorist nor number theorist) recognizing the connection.
In chapter 5, I will develop a global Cohen-Lenstra theory. A naive approach must fail because it will lead to non-converging power-series. So the global setting is considerably harder than the local theories, and the problem of giving a sound global probabilistic interpretation to the Cohen-Lenstra heuristic was unsolved so far. I will give a satisfactory answer by restricting to a (still rich) class of measurable sets.
Finally, in chapter 6, I will give applications of the Cohen-Lenstra heuristic. In particular, I will explain in detail the (mostly conjectural) behaviour of class groups of number fields and function fields, which basically behave like random sequences with respect to the Cohen-Lenstra measure, and I will explain what can be proven. I will also give a "real-world" example where these heuristics play a crucial role for the success or failure of a cryptographic protocol. This chapter does not contain new work but rather tries to give a thorough but tight overview over the state of the art (at least in number theory).

My contribution
Here is a list of what I think are the four main contributions of this thesis:

(i) Connecting the Cohen-Lenstra heuristic with partitions. This includes the notion of CL-maps (def. 3.2.1), in particular order-preserving CL-maps (def. 3.2.7), the existence and explicit construction of such a map (alg. 3.3.2 and 3.3.6, cor. 3.3.26), and the demonstration of its usage for obtaining information about the Cohen-Lenstra probabilities (sect. 3.4).

(ii) Establishing a global Cohen-Lenstra theory. In particular, the definition of uniform properties (def. 5.2.2), and the proof that there exists

a global probability space compatible with the local Cohen-Lenstra probabilities (thm. 5.2.6), giving rise to a criterion when a sequence behaves randomly with respect to the Cohen-Lenstra heuristic (def. 5.4.1); also the discussion of other approaches, both of other researchers (sect. 5.1.1) and of my own (sect. 5.1.2).

(iii) Discovering the overlap between the number theorist and the group theorist community (thm. 4.2.1) and transferring the group theorists' results to the number theory setting. Although this is only a small thing to do, its importance lies in the big impact it may have on both communities since it is possible to transfer a whole arsenal of methods from one side to the other (all of sections 4.3–4.5 and part of 4.6).

(iv) Giving a sound overview about the state of the art in the probability-theoretical treatment of class groups of number fields and related objects (especially sect. 6.1 and 6.3).

Acknowledgements
I want to express my gratitude to my thesis advisor Prof. Dr. Ernst-Ulrich Gekeler for his guidance, not only through the process of writing this thesis, but also through a considerable part of my mathematical education. I take pride in discovering that both my mathematical thinking and my mathematical style bear the imprints of his education, although I still have a long way to go before reaching his mastery.
I thank my colleagues from the math department for creating a pleasant atmosphere which made it even more fun to work on mathematical problems there. I owe extra thank to Alice Keller, who helped me with LaTeX and all other practical questions, before she tragically died in an accident in 2007; to Ute Gebhardt, who joined me in exploring many interesting mathematical areas, and who carefully proof-read this thesis; and to Bernd Mehnert, whose enthusiasm and joy about mathematics was inevitably positively infectious.
Finally, I want to thank my friends and my family for supporting my work and for their endless moral support.

Chapter 1

Basic Theory

In this chapter, I will introduce notation and provide some important facts about finite abelian groups and about partitions. Since they are common knowledge I will not prove them except for some less common formulas concerning the size of the automorphism group and the number of homomorphisms with certain properties.

1.1 Notation

Throughout the thesis, I will use the following standard notation
$\mathbb{N} := \{0, 1, 2, \ldots\}$ the set of natural numbers, including 0.
$\mathbb{N}^+ := \{1, 2, \ldots\}$ the set of natural numbers without 0.
$\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$ the set of integers, rationals, reals, and complex numbers, respectively.
$\mathbb{P}$ the set of all (positive) primes.
$\mathbb{Z}/n := \mathbb{Z}/n\mathbb{Z} := \{\overline{0}, \ldots, \overline{n-1}\}$ the set of residue classes modulo n.
$\mathbb{F}_q$ the finite field with q elements, q a power of some prime p.
$\mathbb{Z}_p$ the ring of p-adic integers.
$\mathbb{Q}_p$ the field of p-adic rationals.
$\mathcal{G}$ the set of all (isomorphism classes of) finite abelian groups.
$\mathcal{G}_p$ the set of all (isomorphism classes of) p-primary finite abelian groups, for a prime p (def. 1.2.2).
$\mathcal{P}$ the set of all integer partitions (definition 1.3.1).
$\#M, |M|$ the cardinality of the set M.

1.1.1 Definition. *Let $\mathfrak{T}$ be the set of all tuples of non-negative integers of arbitrary size. Then $\mathfrak{T}$ carries a natural partial ordering: Let $t = (t_j)_{1\leq j\leq k}$, $u = (u_j)_{1\leq j\leq l} \in \mathfrak{T}$.*
We say that t dominates *u ($t \geq u$) if for all $i \geq 1$ we have*

$$\sum_{j=1}^{i} t_j \geq \sum_{j=1}^{i} u_j,$$

where $t_j, u_j := 0$ for $j > k, l$, respectively.
We say that t strictly dominates *u if t dominates u and $t \neq u$.*

1.2 Finite abelian groups and p-groups

I start with repeating some basic facts and fixing some notation about finite abelian groups and p-groups. I expect that the reader is familiar with the theorems in this section. If not so, they can be found in any introductory book about groups, e.g. [Lan65]. However, please check our definition of the p-adic exponent (definition 1.2.3).
Throughout the whole thesis, we will work only with finite abelian groups. So whenever I talk about groups, I automatically mean finite abelian groups. Moreover, we will consider groups only up to isomorphism. So when I use phrases or formulas like "sum over all groups", then I mean that the sum runs over all isomorphism classes of finite abelian groups. Also, when I use the formula "$G_1 = G_2$" for groups G_1 and G_2, I only mean that the groups are isomorphic.

1.2.1 Definition. *Let G be a finite abelian group. Then the* order $\mathrm{ord}(G)$ *of G is the number of elements in G. The* rank $\mathrm{rk}(G)$ *of G is the minimal number of elements generating G.*
We denote the set of all (isomorphism classes of) finite abelian groups by $\mathcal{G}$.

1.2.2 Definition. *A group is called a p-*group *or a (p-)*primary *group (for $p \in \mathbb{P}$), if the order of each element is a power of p. Unless otherwise stated, we assume all p-groups to be finite and abelian. We denote the set of all finite abelian p-groups by $\mathcal{G}_p$.*

1.2.3 Definition. *Let G be a finite abelian p-group. We define its p-*adic exponent $\exp(G) = \exp_p(G)$ *to be the smallest $n \in \mathbb{N}^+$ such that p^n annihilates every element of G.*
This definition deviates from the standard definition of the exponent. The usual definition would define p^n instead of n as the exponent. In particular, this definition applies only to p-groups.

The next theorems are versions or corollaries of the *Elementary Divisor Theorem*:

1.2.4 Theorem. *Every finite abelian group is a direct product of cyclic groups.*

1.2.5 Theorem. *Every finite abelian group G is the product of p-primary groups: $G = \prod_{p\in\mathbb{P}} G_p$. The groups G_p are uniquely determined by G and they are called* p-parts of G *or* p-primary parts of G.

1.2.6 Example. $\mathbb{Z}/12 \times \mathbb{Z}/4 \times \mathbb{Z}/18 = \underbrace{\mathbb{Z}/4 \times \mathbb{Z}/4 \times \mathbb{Z}/2}_{2-\text{primary part}} \times \underbrace{\mathbb{Z}/9 \times \mathbb{Z}/3}_{3-\text{primary part}}$

Recall that "=" means "isomorphic".

1.2.7 Theorem. *A finite abelian p-group G can be (up to isomorphism) uniquely written in the form*

$$\prod_{i=1}^{k} (\mathbb{Z}/p^{e_i})^{r_i},$$

where $k \in \mathbb{N}$, $e_i, r_i \in \mathbb{N}^+$ for all i, and where $e_1 > e_2 > \ldots > e_k$.
We call this the standard form *of a finite abelian p-group. .*

1.2.8 Definition. *Fix a prime p. Recall that $\mathcal{G}_p$ denotes the set of all finite abelian p-groups.*
We define $\mathcal{G}_\mathcal{P}$ to be the set of all ordered tuples $\underline{e} = (e_1, \ldots, e_n)$, $n \in \mathbb{N}$, $e_1 \geq \ldots \geq e_n > 0$ of positive integers, of arbitrary size n. By the preceding theorem, we may identify $\mathcal{G}_p$ and $\mathcal{G}_\mathcal{P}$ and whenever we are in a local situation (i.e., with one fixed prime p), we will regularly do so without any further comment.
In particular, by $\underline{0} := () \in \mathcal{G}_\mathcal{P} = \mathcal{G}_p$, we denote the trivial p-group.
We endow $\mathcal{G}_\mathcal{P}$ (and hence, $\mathcal{G}_p$) with the partial ordering of domination *(cf. 1.1.1). Note that if $\underline{n}$ dominates $\underline{m}$ then* $\mathrm{ord}(\underline{n}) \geq \mathrm{ord}(\underline{m})$.

1.2.9 Remark.

- The ordering of domination is not the most natural ordering one could impose on the set of all p-groups. One could, for example, define G_1 to be smaller than G_2 if G_1 is isomorphic to a subgroup of G_2. But the ordering we have chosen is finer (in the sense that more elements are comparable – so if G_1 is isomorphic to a subgroup of G_2 then G_2 also dominates G_1, but not vice versa), and it will turn out to be more helpful in some proofs, especially in chapter 3.

However, the ordering does *not* play any role in the motivational part of chapter 3. Particularly, it does not contribute to the definition of an order-preserving CL-map, cf. 3.2.7. To put it bluntly, if you think that the ordering is unnatural, that's just fine. It is only needed for proofs, and a reader not interested in the proofs may completely ignore it.

- Ordered tuples as in $\mathcal{G}_{\mathcal{P}}$ are known as partitions, and that is why the index $\mathcal{P}$ is chosen. The set of all partitions will get a symbol of its own, $\mathcal{P}$. So $\mathcal{P}$ and $\mathcal{G}_{\mathcal{P}}$ denote the same set. However, partitions will occur in a very different role and to a wide extent in chapter 3. Since there is a considerable danger of confusing the different roles, I have introduced a distinct notation for $\mathcal{G}_{\mathcal{P}}$.

  In the local theory (chapters 2 – 4), we deal only with a single prime p and we need not distinguish between $\mathcal{G}_p$ and $\mathcal{G}_{\mathcal{P}}$.

1.2.1 The size of the automorphism group

The Cohen-Lenstra heuristic makes extensive use of the size of the automorphism group $\mathrm{Aut}(G)$ of G. I give an elementary computation of this size. First note that if we have a p_1-group G_1 and a p_2-group G_2, where p_1 and p_2 are two distinct primes, then there are no non-trivial homomorphisms $G_1 \to G_2$. Indeed, every element of G_1 must be mapped onto an element of G_2 of order a power of p_1, and the only such element in G_2 is the neutral element. Consequently, we have a natural bijection

$$\mathrm{Aut}(G_1 \times G_2) \cong \mathrm{Aut}(G_1) \times \mathrm{Aut}(G_2) \qquad (G_1 \text{ a } p_1\text{-group}, G_2 \text{ a } p_2\text{-group}).$$

Thus we must only compute $|\mathrm{Aut}(G)|$ for G a p-group, which is done with the following theorem.

1.2.10 Theorem. *Let $G = \prod_{i=1}^{k} (\mathbb{Z}/p^{e_i})^{r_i}$ be a finite abelian p-group in standard form, i.e., $k \geq 0$, $e_1 > \ldots > e_k > 0$, $r_i > 0$. The size of the automorphism group of G is*

$$|\mathrm{Aut}(G)| = \left(\prod_{i=1}^{k} \left(\prod_{s=1}^{r_i} (1 - p^{-s}) \right) \right) \left(\prod_{1 \leq i,j \leq k} p^{\min(e_i,e_j) r_i r_j} \right).$$

The rest of this section is devoted to proving this theorem. We will first translate automorphisms into certain matrices. Afterwards, we will be able to count the automorphisms quite easily.

Let $G = \prod_{i=1}^{k} (\mathbb{Z}/p^{e_i})^{r_i}$, with notation as above. In this decomposition, we may choose generating elements $(a_{i,j})_{i=1,\ldots,k, j=1,\ldots,r_i}$, where $a_{i,j}$ has order p^{e_i}.

In order to determine an automorphism ϕ of G, we may as well specify a matrix M of the following form:

$$\begin{array}{c} r_1 \\ r_2 \\ r_3 \\ \vdots \\ r_k \end{array} \left\{ \begin{array}{c} \\ \\ \\ \\ \\ \end{array} \right. \left(\begin{array}{ccccc} \boxed{A_1} & B_{1,2} & B_{1,3} & \cdots & B_{1,k} \\ C_{2,1} & \boxed{A_2} & B_{2,3} & \cdots & B_{2,k} \\ C_{3,1} & C_{3,2} & \boxed{A_3} & \cdots & B_{3,k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ C_{k,1} & C_{k,2} & C_{k,3} & \cdots & \boxed{A_k} \end{array} \right),$$

where A_i is an $r_i \times r_i$-square matrix, $B_{i,j}$ is an $r_i \times r_j$-matrix and $C_{i,j}$ is an $r_i \times r_j$-matrix.
What entries do the submatrices have? First consider A_i. It must describe a map from $(\mathbb{Z}/p^{e_i})^{r_i}$ into itself. So we need it to be a $r_i \times r_i$-matrix with entries in $\mathbb{Z}/p^{e_i}$.
Now turn to $B_{i,j}$, $i < j$. This submatrix must map $\mathbb{Z}/p^{e_j}$ to $\mathbb{Z}/p^{e_i}$. Since $e_j < e_i$, we must take the entries of $B_{i,j}$ to be in $p^{e_i - e_j}\mathbb{Z}/p^{e_i}$, which is (as a group) isomorphic to $\mathbb{Z}/p^{e_j}$.
Finally, $C_{i,j}$, $i > j$, must map $\mathbb{Z}/p^{e_j}$ to $\mathbb{Z}/p^{e_i}$, where $e_i < e_j$. So we take the entries of $C_{i,j}$ to be elements of $\mathbb{Z}/p^{e_i}$.
So far, we have described all *endomorphisms* of G. However, we want ϕ to be an *automorphism*, i.e., to be bijective. The following lemma shows that this is the case if and only if the diagonal blocks are invertible.

1.2.11 Lemma. *Let M be a matrix as above, representing the homomorphism $\phi : G \to G$. Then the following statements are equivalent:*

(i) ϕ is bijective.

(ii) A_i is invertible for all $i = 1, \ldots, k$.

(iii) The reduction $\overline{A}_i$ of A_i mod p is invertible for all $i = 1, \ldots, k$.

Proof. "*(ii)* $\Leftrightarrow$ *(iii)*": This is clear: A matrix is invertible if and only if its determinant is invertible, which means in both cases that the determinant is not divisible by p.
"*(i)* $\Leftrightarrow$ *(iii)*": M is a module over the local ring $\mathbb{Z}_p$ with maximal ideal $p\mathbb{Z}_p$ and residue field $\mathbb{F}_p$. Then the endomorphism $\overline{\phi}$ of $\overline{G} := G/pG$ is represented by a matrix $\overline{M}$ with entries in $\mathbb{F}_p$ of the form

$$\overline{M} = \begin{pmatrix} \boxed{\overline{A}_1} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ * & \boxed{\overline{A}_2} & \mathbf{0} & \cdots & \mathbf{0} \\ * & * & \boxed{\overline{A}_3} & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ * & * & * & \cdots & \boxed{\overline{A}_k} \end{pmatrix}.$$

We see that $\overline{\phi}$ is bijective if and only if all $\overline{A}_i$ are invertible. On the other hand, by Nakayama's lemma, $\overline{\phi}$ is bijective if and only if ϕ is bijective, which proves the lemma.

□

The preceding lemma specifies the matrices we need to count. In order to arrange the calculation neatly, we first treat a helpful special case.

1.2.12 Lemma. *Let* $H := \mathrm{Aut}\,((\mathbb{Z}/p^e)^r)$. *Then*

$$|H| = p^{r^2 e} \prod_{s=1}^{r} (1 - p^{-s}).$$

Proof. Let us first do the case $e = 1$. So we count invertible $r \times r$-matrices with entries in $\mathbb{Z}/p$. The first column may contain any non-zero vector, which gives $p^r - 1$ possibilities.
The second column may contain any vector v which does not lie in the span of the first column vector. This rules out p vectors, leaving $p^r - p$ possibilities for the second column.
In the same manner, the s-th column may contain any vector which is not in the span of the first $s-1$ column vectors, giving $p^r - p^{s-1}$ possibilities for the s-th column. Altogether, we get

$$\begin{aligned} \prod_{s=1}^{r} (p^r - p^{s-1}) &= p^{r^2} \prod_{s=1}^{r} (1 - p^{s-1-r}) \\ &= p^{r^2} \prod_{s=1}^{r} (1 - p^{-s}) \end{aligned}$$

matrices.

For the case $e > 1$, we may use the exact sequence

$$0 \to H_1 \hookrightarrow H \to \mathrm{Aut}((\mathbb{Z}/p)^r) \to 1,$$

where $H_1 := \{M \in H \mid M \equiv \mathrm{Id} \mod p\}$ and where the second map is reduction mod p.
Since the sequence is exact, we have

$$\begin{aligned} \#H &= \#H_1 \cdot \#\mathrm{Aut}((\mathbb{Z}/p)^r) \\ &= p^{r^2(e-1)} \cdot p^{r^2} \prod_{s=1}^{r} (1 - p^{-s}) \\ &= p^{r^2 e} \prod_{s=1}^{r} (1 - p^{-s}). \end{aligned}$$

□

Now we are ready to compute the size of the automorphism group of a general finite abelian group, thus to complete the proof of theorem 1.2.10:

Proof of Theorem 1.2.10. We must count all matrices M of the form

$$\begin{array}{l} r_1 \left\{ \right. \\ r_2 \left\{ \right. \\ r_3 \left\{ \right. \\ \vdots \\ r_k \left\{ \right. \end{array} \left(\begin{array}{ccccc} \boxed{A_1} & B_{1,2} & B_{1,3} & \cdots & B_{1,k} \\ C_{2,1} & \boxed{A_2} & B_{2,3} & \cdots & B_{2,k} \\ C_{3,1} & C_{3,2} & \boxed{A_3} & \cdots & B_{3,k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ C_{k,1} & C_{k,2} & C_{k,3} & \cdots & \boxed{A_k} \end{array} \right),$$

where A_i is an invertible $r_i \times r_i$-square matrix with entries in $\mathbb{Z}/p^{e_i}$, $B_{i,j}$ is a $r_i \times r_j$-matrix with entries in $\mathbb{Z}/p^{e_j}$, and $C_{i,j}$ is a $r_i \times r_j$-matrix with entries in $\mathbb{Z}/p^{e_i}$.
By the previous lemma, there are $p^{(r_i)^2 e_i} \prod_{s=1}^{r_i}(1-p^{-s})$ possibilities to choose A_i.
For $B_{i,j}$ we have $p^{e_j r_i r_j}$ possibilities, for $C_{i,j}$ we have $p^{e_i r_i r_j}$ possibilities. Note that in the first case we always have $i < j$ ($\Leftrightarrow e_i > e_j$), and in the latter case we always have $i > j$ ($\Leftrightarrow e_i < e_j$). We may unify both formulas by saying that we have $p^{\min(e_i,e_j) r_i r_j}$ possibilities for each pair (i, j), $i \neq j$.

Altogether, we get

$$\begin{aligned}|\mathrm{Aut}(G)| &= \left(\prod_{i=1}^{k}\left(p^{(r_i)^2e_i}\prod_{s=1}^{r_i}(1-p^{-s})\right)\right)\left(\prod_{1\le i,j\le k, i\neq j} p^{\min(e_i,e_j)r_ir_j}\right)\\ &= \left(\prod_{i=1}^{k}\left(\prod_{s=1}^{r_i}(1-p^{-s})\right)\right)\left(\prod_{1\le i,j\le k} p^{\min(e_i,e_j)r_ir_j}\right),\end{aligned}$$

as required.

□

1.2.2 Counting homomorphisms

In this section I collect some facts about group homomorphisms that we will need in later chapters.

1.2.13 Lemma. *Let $e' > e \geq 0$. The number of $r \times r$-matrices A over $R := \mathbb{Z}/p^{e'}\mathbb{Z}$ which satisfy $(p^eR)^r \not\subseteq \mathrm{im(A)}$ is less than or equal to*

$$p^{r^2e'+r^2-re}\prod_{s=1}^{r}(1-p^{-s}).$$

Proof. Given such a matrix A, we may compose the homomorphism defined by A with an automorphism φ of $\mathbb{Z}/p^e$ that is the identity when restricted to $\mathbb{Z}/p^{e-1}$ in order to achieve that $(0,0,\ldots,0,p^e) \notin \mathrm{im}(\varphi \circ \mathrm{A})$. (More precisely, take such an automorphism $\varphi \bmod p^e$ and extend it in an arbitrary way to an automorphism mod $p^{e'}$.) Then

$$(\mathrm{im}(\varphi \circ \mathrm{A})) \subseteq \underbrace{\mathbb{Z}/p^{e'} \times \mathbb{Z}/p^{e'} \times \ldots \times \mathbb{Z}/p^{e'}}_{\mathrm{r-1}} \times p^{e+1}\mathbb{Z}/p^{e'}.$$

We have $p^{r(re'-e)}$ possibilities for $\varphi \circ A$.
How many possibilities do we have for φ? Obviously, choosing φ is the same as choosing an automorphism of $(\mathbb{Z}/p\mathbb{Z})^r$, so by lemma 1.2.12 we get

$$p^{r^2}\prod_{s=1}^{r}(1-p^{-s})$$

possible choices for φ.
The choice of φ may not be unique, but in any case we have at most

$$p^{r(re'-e)}p^{r^2}\prod_{s=1}^{r}(1-p^{-s}) = p^{r^2e'+r^2-re}\prod_{s=1}^{r}(1-p^{-s})$$

matrices A with $(p^eR)^r \not\subseteq \mathrm{im(A)}$.

□

1.2.14 Lemma. *Let $G \in \mathcal{G}_p$. Let $N := \mathrm{ord}(G)$ and $r := \mathrm{rk}(G)$. Let $n \geq r$. Then*

$$\#\{\Gamma \subseteq \mathbb{Z}_p^n \mid \mathbb{Z}_p^n/\Gamma \cong G\} = \frac{N^n}{|\mathrm{Aut}(G)|}\left(\prod_{i=n-r+1}^{n}(1-p^{-i})\right).$$

Here, Γ runs through all submodules of $\mathbb{Z}_p^n$.

Proof. The lemma is a consequence of [CL84, 3.1.], and I present a simplified version of the proof given there.
Let s_n be the number of surjective homomorphisms (of $\mathbb{Z}_p$-modules) from $\mathbb{Z}_p^n$ to G and let $\overline{s}_n$ be the number of surjective homomorphisms (of $\mathbb{Z}/p\mathbb{Z}$-modules) from $(\mathbb{Z}/p\mathbb{Z})^n$ to G/pG.
By the lemma of Nakayama, a homomorphism $\varphi \in \mathrm{Hom}_{\mathbb{Z}_p}(\mathbb{Z}_p^n, G)$ is surjective if and only if its reduction $\overline{\varphi} = \varphi \bmod p \in \mathrm{Hom}_{\mathbb{Z}/p\mathbb{Z}}((\mathbb{Z}/p\mathbb{Z})^n, G/pG)$ is surjective. Hence,

$$s_n = \overline{s}_n \cdot \#\{\varphi \in \mathrm{Hom}_{\mathbb{Z}_p}(\mathbb{Z}_p^n, G/pG) \mid \overline{\varphi} = \overline{0}\}.$$

For the latter factor, $\overline{\varphi} = \overline{0}$ if and only if all basis elements of $\mathbb{Z}_p^n$ are mapped into pG, hence this factor equals

$$\#\{\varphi \in \mathrm{Hom}_{\mathbb{Z}_p}(\mathbb{Z}_p^n, G/pG) \mid \overline{\varphi} = \overline{0}\} = \left(\frac{N}{p^r}\right)^n.$$

Now let us compute $\overline{s}_n$. This number equals the number of $n \times r$-matrices of rank r over $\mathbb{Z}/p\mathbb{Z}$. In a similar way as in the proofs before, we see that the number of possibilities for the i-th column is successively $p^n - p^{i-1}$. Therefore,

$$\overline{s}_n = \prod_{i=1}^{r}(p^n - p^{i-1}).$$

Together, we get

$$s_n = \left(\prod_{i=1}^{r}(p^n - p^{i-1})\right)\left(\frac{N}{p^r}\right)^n = N^n \prod_{i=n-r+1}^{n}(1-p^{-i}).$$

Finally, each surjective homomorphism φ from $\mathbb{Z}_p^n$ to G defines a submodule $\Gamma := \ker(\varphi)$ such that $\mathbb{Z}_p^n/\Gamma \cong G$. Two homomorphisms φ_1 and φ_2 define the same Γ if and only if there exists an automorphism σ of G such that $\varphi_1 = \sigma \circ \varphi_2$. Hence,

$$\#\{\Gamma \subseteq \mathbb{Z}_p^n \mid \mathbb{Z}_p^n/\Gamma \cong G\} = \frac{s_n}{|\mathrm{Aut}(G)|},$$

which proves the assertion.

□

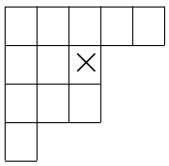
1.3 Partitions

1.3.1 Notation. A *partition* $\underline{n}$ is a tuple of non-increasing positive integers $\underline{n}_1 \geq \underline{n}_2 \geq \ldots \geq \underline{n}_k$. We call $n := \sum_{i=1}^{k} \underline{n}_i$ the *size* of $\underline{n}$ and say that $\underline{n}$ is a partition of n. We call k the *rank* of $\underline{n}$.
We always mark both partitions and their entries with underscores. The reason for this rather unusual convention is that we need to distinguish them from derivation, which will appear in chapter 3.
Every partition of n corresponds to a way of writing n as a sum of positive integers.
We call $\mathcal{P}$ the set of all partitions.
Recall that we denote by $\mathcal{G}_{\mathcal{P}}$ the set of all finite abelian p-groups. This notation is not incidental – you may have noticed that $\mathcal{G}_{\mathcal{P}}$ is also the set of all partitions. However, this is a rather unfortunate coincidence. We will talk extensively about partitions (especially in chapter 3), and they have a *completely different meaning* than the partitions in $\mathcal{G}_{\mathcal{P}}$. In particular, they come along with different partial orderings (definition 1.3.6 and 1.2.8). I will try to help you avoid confusion about the different roles of partitions: Whenever I refer to partitions $\underline{n} \in \mathcal{G}_{\mathcal{P}}$, then I will use these partitions *exclusively* to label groups. They are not subject to any algorithm whatsoever, nor are they derived, transformed, converted, processed, kneaded or anything else. The only exception is that two such partitions may be compared with respect to domination (definition 1.2.8).

1.3.2 Definition. *Let* $\underline{n} = (\underline{n}_1, \underline{n}_2, \ldots, \underline{n}_k) \in \mathcal{P}$. *Its* Young diagram *is a finite collection of boxes, arranged in left-justified rows, where the* i*-th row (counted from the top) contains* n_i *boxes.*
We label the boxes by pairs (i, j)*, where* i *is the row (counted from the top) and* j *is the column (counted from the left) of the box.*

1.3.3 Remark. Young diagrams are a very convenient tool for visualising partitions. My convention is called "English notation". Some authors will use another convention by counting the rows from the bottom (called "French notation", cf. [Mac79, p.2]).
The Young diagram is sometimes also referred to as "Ferrer's diagram".

1.3.4 Example. The diagram below shows the Young diagram of the partition $(5, 3, 3, 1)$, with its $(2, 3)$-box marked:



1.3.5 Definition. *For any $\underline{n} \in \mathcal{P}$, we call its* conjugate partition *the partition whose Young diagram is obtained by reflecting the Young diagram of $\underline{n}$ along its main diagonal.*
Equivalently, you can read off from the original diagram the columns instead of the rows.
For example, the partition $(5, 3, 3, 1)$ from the above example has conjugate $(4, 3, 3, 1, 1)$. Obviously, conjugation is an involution.

1.3.6 Definition. *We endow $\mathcal{P}$ with a partial ordering. We write $\underline{n} \leq \underline{m}$ if the Young diagram of $\underline{n}$ is contained in the Young diagram of $\underline{m}$. Equivalently, we have $\underline{n} \leq \underline{m}$ if and only if $\underline{n}_i \leq \underline{m}_i$ for all $i \geq 1$, where we declare undefined entries to be 0.*
Please do not confuse this ordering with the ordering on $\mathcal{G}_{\mathcal{P}}$ *(cf. definition 1.2.8)!*

1.3.7 Lemma. *The number $p(n)$ of partitions of n satisfies $p(n) \leq F_{n+1}$, where F_k denotes the k-th Fibonacci number, given by $F_1 := F_2 := 1$ and $F_{n+1} := F_n + F_{n-1}$.*
In particular, $p(n) \in O(\phi^n)$, where $\phi = \frac{1+\sqrt{5}}{2}$ is the golden ratio.

Proof. The proof is elementary and follows immediately from the facts $p(1) = 1$, $p(2) = 2$, and from the formula $p(n+1) < p(n) + p(n-1)$, which is a fun exercise.
For details, see [AE04, 3.3]. □

1.3.8 Remark. In fact, much more precise statements are known. Asymptotically, $p(n)$ grows as $\frac{1}{4n\sqrt{3}} \exp(\pi\sqrt{\frac{2}{3}n})$ ([And76, Thm. 6.3]). However, the weaker form given above is sufficient for us.

1.3.1 q-series identities

There are plenty of interesting power series identities for partitions. We will need some of them. For a very nice and down-to-earth introduction you may consult the book of Andrews and Eriksson [AE04]. A much more thorough treatment may be found in [And76]. The book [Sta97] lies somewhere in the middle, but contains less material than the other two books.
We want to study the function $p(n) := \#\{\text{partitions of } n\}$ and variants thereof. The basic tool for this (and many other purposes) is to consider the Fourier transform of this function, which in the combinatorial context is called generating function:

$$F(q) = \sum_{n=0}^{\infty} p(n) q^n.$$

The Fourier transform turns convolutions into multiplications, and derivations (discrete in this case, i.e., expressions like $a(n) - a(n-1)$) to mere multiplications with powers of q, thereby simplifying many relations. The series considered in this section have all a positive radius of convergence, so the generating functions are all well-defined as holomorphic functions.
We only need the most basic formulas: Investigating $F(q)$, we first get a product formula

$$F(q) = \prod_{i=1}^{\infty} (1 - q^i)^{-1},$$

which can immediately be checked by using the geometric series expansion $(1 - q^i)^{-1} = 1 + q^i + q^{2i} + \dots$ and multiplying out.
Analogously, we get a generating function for the number $p_a(n)$ of partitions of n into at most a integers:

$$\sum_{n=0}^{\infty} p_a(n) q^n = \prod_{i=1}^{a} (1 - q^i)^{-1}.$$

By conjugation, we see that $p_a(n)$ also equals the number of partitions into integers that are all $\leq a$.
We will need one more formula, which is not quite so obvious:

1.3.9 Proposition. *Let $p_{a,b}(n)$ be the number of partitions of n into at most a integers of size at most b. Then the generating function $\psi_{a,b}(q)$ satisfies:*

$$\psi_{a,b}(q) = \sum_{n=0}^{\infty} p_{a,b}(n) q^n = \frac{\prod_{i=1}^{a+b} (1 - q^i)}{(\prod_{i=1}^{a} (1 - q^i))(\prod_{i=1}^{b} (1 - q^i))}.$$

Proof. [AE04, 7.2] □

The preceding formula for the generating function of $p_a(n)$ may be considered as formula for $\psi_{a,\infty}(q) = \psi_{\infty,a}(q)$.

In the same manner, a lot of other formulas can be proven, and that is only the top of the iceberg. If the reader is not familiar with the topic, I can only encourage you to have a look at the treatment in the books mentioned above — not because it is needed in this thesis, but rather because it is such a beautiful field of mathematics.

Chapter 2

The Local Cohen-Lenstra Heuristic

In this chapter, I introduce the Cohen-Lenstra heuristic for finite abelian p-groups. We will call this the "local" Cohen-Lenstra heuristic.

2.1 The Cohen-Lenstra heuristic for finite abelian p-groups

Let me first give a probabilistic formulation of the Cohen-Lenstra heuristic:
Let p be a prime. Assume we have a "natural", unbiased stochastic process producing finite abelian p-groups. If we fix a finite abelian p-group G then the probability that an output of the process is isomorphic to G is inversely proportional to the size of its automorphism group $\mathrm{Aut}(G)$.
In this formulation, the heuristic is not a theorem but rather a meta-principle. It first became popular by the famous paper [CL84] of Henri Cohen and Hendrik W. Lenstra. In honor to this paper I call the principle "Cohen-Lenstra heuristic" or "Cohen-Lenstra principle". In their paper they claimed (without proof, but with some evidence) that the sequence of p-parts of class groups of imaginary quadratic number fields (which is a deterministic sequence!) behaves essentially like a random sequence in the above sense, for $p \neq 2$. We will see an exact formulation in section 6.1.1.
In the definition above, "unbiased" is not a precise term but rather means that we do not allow obvious obstacles. For example, there might well be stochastic processes that produce only cyclic groups. Or some that produce only groups of rank at most 2 (as is the case for the point group of elliptic curves over various finite fields). Such processes may well be modelled via

a probabilistic approach (done so for the elliptic curves in [Gek06]), but the probability distribution is clearly biased.
You might ask why we restrict ourselves to p-groups and do not allow arbitrary finite abelian groups. Indeed, in some way the Cohen-Lenstra heuristic seems to apply to the general setting as well. However, if one tries to make the heuristic precise, one runs into serious trouble. In fact, *for general finite abelian groups, there is no probability distribution that would allow us to perform a stochastic process generating random sequences of groups that is compatible with the Cohen-Lenstra heuristic*, as we can do when restricting to p-groups. So we can not compare a sequence of groups with a random sequence because there is no adequate stochastic process that could generate such a random sequence. We will examine this problem and analyze possible workarounds in much detail in chapter 6.
Returning to p-groups, let us make a more precise definition:

2.1.1 Definition. *The* Cohen-Lenstra weight w *is the measure on the set* $\mathcal{G}_p$ *of all finite abelian* p*-groups that is defined via*

$$w(\{G\}) = \frac{1}{|\mathrm{Aut}(G)|} \qquad \text{for all one-element sets } \{G\} \subset \mathcal{G}_p.$$

The (local) Cohen-Lenstra probability measure P *is the probability measure on* $\mathcal{G}_p$ *that is obtained by scaling* w*:*

$$P(M) := \frac{w(M)}{w(\mathcal{G}_p)} \qquad \text{for } M \subseteq \mathcal{G}_p$$

In slight abuse of notation I will write $w(G)$ *and* $P(G)$ *instead of* $w(\{G\})$ *and* $P(\{G\})$ *when we measure one-element sets* $\{G\} \subset \mathcal{G}_p$*.*
Note that w *and* P *depend on the prime* p*. If we need to distinguish several primes, I write* w_p *and* P_p *instead of* w *and* P*, respectively.*

We must check that P is well-defined, i.e., that the measure w is finite, $w(\mathcal{G}_p) < \infty$. More precisely, we prove the following theorem:

2.1.2 Theorem. *The Cohen-Lenstra weight of the set of all finite abelian* p*-groups is*

$$w(\mathcal{G}_p) = \prod_{i=1}^{\infty} (1 - p^{-i})^{-1}.$$

Proof. I give a nice combinatorial proof due to Hall [Hal38]. Cohen and Lenstra use a completely different method for their proof (cf. [CL84]). Their

approach is more complicated but has the advantage that it generalizes naturally to finite modules over the rings of integers of number fields.
Let $G \cong \underline{\lambda} = (\underline{\lambda}_1, \ldots, \underline{\lambda}_l) \in \mathcal{G}_p$ ($\cong \mathcal{G}_{\mathcal{P}}$), and let $\underline{\mu} = (\underline{\mu}_1, \ldots, \underline{\mu}_m)$ be its conjugate partition. Then the factor $\mathbb{Z}/p^i\mathbb{Z}$ occurs exactly $\underline{\mu}_i - \underline{\mu}_{i+1}$ times in G.
By theorem 1.2.10, we have

$$\begin{aligned} |\mathrm{Aut}(G)| &= \left(\prod_{i=1}^{m}\left(\prod_{s=1}^{\underline{\mu}_i-\underline{\mu}_{i+1}}(1-p^{-s})\right)\right)\left(\prod_{1\le i,j\le m} p^{\min(i,j)(\underline{\mu}_i-\underline{\mu}_{i+1})(\underline{\mu}_j-\underline{\mu}_{j+1})}\right) \\ &= \left(\prod_{i=1}^{m}\left(\prod_{s=1}^{\underline{\mu}_i-\underline{\mu}_{i+1}}(1-p^{-s})\right)\right)\left(\prod_{1\le i\le m} p^{\underline{\mu}_i^2}\right), \end{aligned}$$

where the latter equality is obtained by counting for each pair (i,j) how often the term $\underline{\mu}_i\underline{\mu}_j$ appears. It is easily seen that the number of occurrences sums up to 0 if $i \neq j$, and to 1 otherwise.
When $\underline{\lambda}$ runs through $\mathcal{G}_p$ then $\underline{\mu}$ runs through $\mathcal{G}_p$ as well. So we need to show that

$$\sum_{n=0}^{\infty} a_n p^{-n} = \sum_{n=0}^{\infty} \sum_{\substack{\underline{\mu}\in\mathcal{G}_p \\ \mathrm{size}(\underline{\mu})=n}} \mathrm{Aut}(G_{\underline{\mu}})^{-1},$$

where a_n is the number of partitions of size n and $G_{\underline{\mu}}$ is the group associated to $\underline{\mu}$. For the left hand side, we use the identity of power series

$$\sum_{n=0}^{\infty} a_n q^n = \sum_{n=0}^{\infty} q^n \prod_{i=1}^{n}(1-q^i)^{-1}.$$

Evidently we only need to show for any $n \geq 0$ the q-series identity

$$q^n \prod_{i=1}^{n}(1-q^i)^{-1} = \sum_{\substack{\underline{\mu}\in\mathcal{G}_p \\ \mathrm{size}(\underline{\mu})=n}} \left(\prod_{i=1}^{m}\left(\prod_{s=1}^{\underline{\mu}_i-\underline{\mu}_{i+1}}(1-q^s)^{-1}\right)\right)\left(\prod_{1\le i\le m} q^{\underline{\mu}_i^2}\right), \quad (2.1)$$

and the result will follow by plugging $q := p^{-1}$ and summing over all n.
We prove (2.1) by interpreting both sides as the generating functions of partitions of a certain kind. For the left hand side, the coefficient of q^{N+n}

equals the number of partitions of N with greatest part at most n. To such a partition $\underline{\nu}$, we define a partition $\underline{\mu}$ from the right hand side as follows:
Consider the Young diagram D of $\underline{\nu}$. Let $\underline{\mu}_1$ be the largest integer such that the point $(\underline{\mu}_1, \underline{\mu}_1)$ belongs to D. Think of it as the lower right corner point of the largest square fitting into D. Now we recursively define $\underline{\mu}_i$ to be the largest integer such that $(\underline{\mu}_1 + \underline{\mu}_2 + \ldots + \underline{\mu}_i, \underline{\mu}_i)$ belongs to D. Thus $\underline{\mu}_i$ is the size of the largest square that fits below the preceeding squares within D.
Let $M := N - \underline{\mu}_1^2 - \underline{\mu}_2^2 - \ldots$. Then there are exactly M blocks of D outside the mentioned squares. We divide those blocks up as follows: Define M_i to be the number of blocks of D at the right of the i-th square, i.e. number of blocks $(x, y) \in D$ such that

$$\underline{\mu}_1 + \underline{\mu}_2 + \ldots + \underline{\mu}_{i-1} < x \leq \underline{\mu}_1 + \underline{\mu}_2 + \ldots + \underline{\mu}_i$$

and

$$\underline{\mu}_i < y.$$

Then clearly $M = M_1 + \ldots + M_m$. Furthermore, the blocks forming M_i form a partition of height at most $\underline{\mu}_i$ and width at most $\underline{\mu}_{i-1} - \underline{\mu}_i$ (if the width were larger then $\underline{\mu}_{i-1}$ could have been chosen larger).
On the other hand it is clear that we can reverse our construction: Given any term q^N on the right hand side, specified by the choice of $\underline{\mu}$, numbers M_i such that $M_1 + \ldots + M_m = N - \underline{\mu}_1^2 - \ldots - \underline{\mu}_m^2$ and partitions of M_i of height at most $\underline{\mu}_i$ and width at most $\underline{\mu}_{i-1} - \underline{\mu}_i$, then we can reconstruct the Young diagram D and thus the partition $\underline{\nu}$.
Denoting by $\psi_{a,b}(q)$ the generating function for the partitions of n with height at most a and width at most b, we have proven that

$$q^n \prod_{i=1}^{n} (1 - q^i)^{-1} = \sum_{\substack{\underline{\mu} \in \mathcal{G}_{\mathcal{P}} \\ \mathrm{size}(\underline{\mu}) = n}} \left(\prod_{i=1}^{m} \psi_{\underline{\mu}_{i+1}, \underline{\mu}_i - \underline{\mu}_{i+1}}(q) \right) \left(\prod_{1 \leq i \leq m} q^{\underline{\mu}_i^2} \right),$$

where we put $\underline{\mu}_0 := \infty$ and $\underline{\mu}_{m+1} := 0$.
But we have already seen in proposition 1.3.9 that for finite a and b

$$\psi_{a,b} = \frac{\prod_{i=1}^{a+b}(1 - q^i)}{\left(\prod_{i=1}^{a}(1 - q^i)\right)\left(\prod_{i=1}^{b}(1 - q^i)\right)}$$

and

$$\psi_{\infty,b} = \frac{1}{\prod_{i=1}^{b}(1 - q^i)}.$$

Plugging in those formulas, we obtain equation (2.1). This finishes our proof.
□

2.1.3 Remark. In 3.2.1 I will define the notion of a CL-map, which exploits the same idea; it gives a bijection between the set of all partitions and the set of all terms in the sum over all $w(G)$. However, note that the above proof does *not* give a CL-map. It makes use of the cancellation of the $\psi_{a,b}$-terms, and therefore breaks up into two steps: There is one proper bijection establishing equation (2.1), and then there are formulas for the $\psi_{a,b}$ used to establish the link to equation (2.1).
The formulas for $\psi_{a,b}$ can be obtained in many different ways, but I know of no proof by bijection. But even if there should exist one, that does not mean that there is an obvious way to combine those two bijections. If you would apply the bijections for $\psi_{a,b}$ within the other bijection, then you would most likely drop out of the restriction that your local partition has height and width limited to a and b, respectively, and therefore you would collide with the definition of the $\underline{\mu}_i$.

2.2 Motivation

So far, I have not given any justification for the Cohen-Lenstra heuristic. Unfortunately, for the most important applications there are no proofs known that the sequences in focus really behave as predicted by the Cohen-Lenstra heuristic. However, there are at least three reasons that support the Cohen-Lenstra principle:

(i) For probabilistic approaches, weighting the elements by the number of their automorphisms has turned out to work very well — not only in our setting with finite abelian (p-)groups, but also in many other cases like lattices, quadratic forms, elliptic curves, ...

(ii) Extensive tests have been carried out for some sequences. Especially the sequence of class groups of quadratic number fields (cf. section 6.1.2) has been listed for some 100,000,000s groups. The numerical data gives overwhelming support for the Cohen-Lenstra principle.

(iii) In [FW89], Friedman and Washington pointed out that the Cohen-Lenstra heuristic arises naturally (and provably!) if one models finite abelian p-groups in the following way: Take n generators and randomly assign n relations on them. This can be done because the set of all such relations is a matrix group over $\mathbb{Z}_p$ on which we have a Haar measure. If we do this, then in the limit ($n \to \infty$), we obtain the Cohen-Lenstra heuristic.

In the following subsections, I will describe these three reasons in more detail.

2.2.1 Automorphisms as weights

It seems to be a general principle in mathematics that counting should most of the time be done in a weighted way, where the weights are inversely proportional to the number of automorphisms. Examples are

- Quadratic forms: Instead of presenting a lot of formulas, I rather let a greater mind speak for me. Here is what Ferdinand Eisenstein [Eis47] says about classes K, K' of ternary quadratic forms with sizes of the automorphism groups δ, δ', respectively:
  "*Obgleich (...) jede Classe in der That unendlich viele Formen enthält, so kann man doch, wenn δ und δ' verschieden sind, nicht mit Recht sagen, daß die Classe K ebenso viele Formen enthielte, als die Classe K'; im Gegentheil kann man behaupten, daß die Formen-Anzahlen dieser beiden Classen im reciproken Verhältniß der beiden Zahlen δ und δ' stehen und daß der reciproke Werth von δ das wahre* Maaß *für die Totalität der Formen einer* Classe*, gewissermaßen für die* Dichtigkeit *der Classe sei. Man thut daher Unrecht, wenn man bei der Vergleichung oder Zusammenstellung mehrerer Classen, jede Classe als eine Einheit zählt, weil, um mich so auszudrücken, nicht jede Classe gleiche Berechtigung hat, man muß vielmehr jede Classe nach ihrem* Maaße $\frac{1}{\delta}$ *zählen.*
  Durch die Einführung dieses neuen Begriffs des Maaßes *der Classen, wird die ganze Theorie der ternären und die aller übrigen quadratischen Formen außerordentlich vereinfacht und, wie ich zu glauben wage, verschönert, während ohne denselben kaum irgendwie vorwärts zu kommen wäre.* "
  "Although every class contains indeed infinitely many forms, one can, when δ and δ' are different, not rightly say that the class K would contain as many forms as the class K'; contrariwise one can claim that the numbers of forms in either of those classes are in reciprocal proportion to the two numbers δ and δ' and that the reciprocal value of δ is the true *measure* for the totality of the forms of a *class*, quasi for the class's *density*. Thence, one does mischief if in comparing or composing several classes one counts every class as unit, for, to express it this way, not every class has the same qualification, rather one must count each class by its *measure* $\frac{1}{\delta}$.
  By introducing this new notion of the classes' *measures*, the whole theory of ternary and all other quadratic forms is extraordinarily simplified and, as I dare believe, beautified, while advancing without that would hardly be possible."

- Elliptic curves: If we count the number of elliptic curves over a finite base field $\mathbb{F}_q$ up to isomorphism naively, we do not get a closed formula but rather need to distinguish the residues of q modulo 12. (E.g., the number is $2q + 6$ if $q \equiv 1 \bmod 12$.)

  However, if we count the same thing in a weighted way, then we obtain simply q such curves, with no case distinction necessary. Similarly, many formulas (the number of Weierstraß normal forms per equivalence class; the number of elliptic curves with prescribed torsion group; ...) become "smoother", and for approximation formulas, higher rates of convergence are obtained. However, since almost all elliptic curves over $\mathbb{F}_p$ have an automorphism class of the same size (namely, 2), the difference is rather cosmetical.

  For an abundance of examples, [Pau08] (as well as many other sources) may be consulted.

  If we do not consider a fixed finite field, but rather consider all elliptic curves defined over any algebraic extension of $\mathbb{F}_p$, then the differences are bigger, and the weights do make a non-negligible contribution. A famous example is

$$\sum_{E \text{ s.s.}} \frac{1}{\#\mathrm{Aut}(E)} = \frac{p-1}{24}.$$

  It is originally due to Eichler and Deuring [Deu41], and can also be found in the very nice article of Tate [Tat74]. The sum needs explanation: We fix a prime p and then E runs over all isomorphism classes of super-singular elliptic curves defined over any algebraic extension of $\mathbb{F}_p$, up to isomorphism over the algebraic closure of $\mathbb{F}_p$.

  No unweighted analogue is known for this formula.

- Other applications include, for example, vector bundles over schemes $X/\mathbb{F}_q$, or lattices over Dedekind domains. These examples (as well as some of the examples above) are linked on a higher level by the theory of Tamagawa numbers, which integrates the weights into a canonical (i.e., with a canonical choice of scaling) Haar measure on $G(A)/G(K)$, where G is an algebraic group over the number field (or function field) K, and A is the adele ring over K. For details, you may consult [CF86] or [Vos98].

2.2.2 Numerical support

Extensive tests have been carried out that strongly support the Cohen-Lenstra principle. In particular, the sequence of class groups of quadratic number fields (cf. section 6.1.2) is extremely well-studied. Analyses of 100,000,000s of class groups give evidence that class groups distribute in perfect accordance with the Cohen-Lenstra distribution (e.g., see [Jac98] for real quadratic fields; [tRW03] for imaginary quadratic number fields; [Mal08] for more general number fields; [Fri00] for function fields).

2.2.3 Modelling p-groups by generators and relations

Assume that we would like to generate a random finite abelian p-group. How could we possibly do this? To make life easier, let us say that we want a group of rank $\leq r$. One way, of course, would be to make use of the structure theorems that classify finite abelian p-groups. So we could specify an increasing sequence of r non-negative integers e_i, obtaining the group $\prod \mathbb{Z}/p^{e_i}$. However, apart from practical problems (how to choose an increasing sequence of integers, or even a single integer!?), this approach is in some sense "unnatural": It makes crucial use of our knowledge about all possible structures of such groups. In a truly random process, we would rather expect that it is not necessary to have a structure theory in order to imitate this process.

Even more important: The process above is not very symmetric. If we think about the elements in the group, we would need to first generate the elements in the group with high exponent, and then generate other elements in strictly decreasing exponent order. Such a strict order is rather uncharacteristic for random processes.

Therefore, we might change our approach: Instead of using a structural description of the group, we could instead work directly with the elements. Since we want the group to have rank at most r, we might choose r generators of the group, and afterwards impose relations onto the generators. In this way, we would not make use of any structural knowledge and would preserve symmetry. Fortunately, it turns out that it is possible to choose random relations in a "natural" way.

This is the approach that Friedman and Washington proposed in [FW89]. I hope that I have convinced you that it is indeed the most natural way to generate a random finite abelian p-group. Not surprisingly, this process yields the Cohen-Lenstra distribution (when $r \to \infty$), therefore legitimating our claim that Cohen-Lenstra is a "natural" distribution.

Choosing relations

There is one question yet to be answered: How do we choose relations on the generators?
A relation in an abelian group is an equation of the form $e_1g_1 + \ldots + e_rg_r = 0$, where g_i are the generators and e_i are non-negative integers. Since we want r relations (in order to get a finite group), we need an $r \times r$-matrix A of integers. The output group G is then $\mathbb{Z}^r/\mathrm{im(A)}$. We need one more change in order for the group to be a p-group: We choose A to be an $r \times r$-matrix of *p-adic* integers and define $G := \mathbb{Z}_p^r/\mathrm{im(A)}$.
So a finite abelian p-group is given by a matrix in $\mathbb{Z}_p^{r\times r}$. How to choose this matrix? Note that $\mathbb{Z}_p$ is a compact group, so it comes along with a Haar measure (which we normalize to have total volume 1 in order to obtain a probability measure). Consequently, $\mathbb{Z}_p^{r\times r}$ also inherits a Haar (probability) measure, and we can choose A with respect to this Haar measure.
There is still one drawback: It might happen that A does not have full rank. In this case, the group G is infinite. But we will see that the probability for this is zero. So with probability 1 we get a finite abelian p-group.

Calculating probabilities

We need to show the following facts.

2.2.1 Theorem (Friedman-Washington)**.**
For a randomly (with respect to the Haar measure) chosen matrix $A \in \mathbb{Z}_p^{n\times n}$:

(i) $Pr(A \text{ has full rank}) = 1$ *for all* $n > 0$.

(ii) For any finite abelian p-group G,

$$Pr(\mathrm{coker}(A) \cong G) \to P(G) \qquad \text{for } n \to \infty,$$

where P is the Cohen-Lenstra probability.

Note that the probability on the left hand side implicitly depends on n.

Proof. (i) A matrix A has full rank if and only if there exists an e such that $(p^e\mathbb{Z}_p)^n \subseteq \mathrm{im(A)}$. For any $e' > e$, this is equivalent to saying that the reduction of A modulo $p^{e'}$ satisfies $(p^e\mathbb{Z}/p^{e'})^n \subseteq \mathrm{im(A \bmod p^{e'})}$. In other words, for all $e \geq 0$ and all $e' > e$ we have

$$\begin{aligned} Pr(A \text{ has full rank over } \mathbb{Z}_p) &\geq Pr\left((p^e\mathbb{Z}_p)^n \subseteq \mathrm{im(A)}\right) \\ &= Pr\left((p^e\mathbb{Z}/p^{e'})^n \subseteq \mathrm{im(A \bmod p^{e'})}\right). \end{aligned}$$

Since reduction mod p^e is compatible with the Haar measures, we may compute the latter probability simply by counting matrices. By lemma 1.2.13, we know that the number of matrices mod $p^{e'}$ which do *not* satisfy $(p^e\mathbb{Z}/p^{e'})^n \subseteq \mathrm{im(A)}$ is at most

$$p^{n^2+n^2e'-ne}\prod_{s=1}^{n}(1-p^{-s}).$$

By dividing by the total number $p^{n^2e'}$ of matrices in $(\mathbb{Z}/p^{e'}\mathbb{Z})^{n\times n}$ we may estimate the probability that a matrix in $(\mathbb{Z}/p^{e'}\mathbb{Z})^{n\times n}$ has this property:

$$Pr\Big(A \in (\mathbb{Z}/p^{e'})^{n\times n} \text{ satisfies } (p^e\mathbb{Z}/p^{e'})^n \not\subseteq \mathrm{im(A)}\Big) \leq p^{n^2-ne}\prod_{s=1}^{n}(1-p^{-s}).$$

Altogether, we have for all $e' > e \geq 0$:

$$\begin{aligned} Pr(A \in (\mathbb{Z}_p)^{n\times n} \text{ has full rank}) &\geq Pr\begin{pmatrix} A \in (\mathbb{Z}/p^{e'})^{n\times n} \text{ satis-} \\ \text{fies } (p^e\mathbb{Z}/p^{e'})^n \subseteq \mathrm{im(A)} \end{pmatrix} \\ &\geq 1 - p^{n^2-ne}\prod_{s=1}^{n}(1-p^{-s}) \\ &\stackrel{e\to\infty}{\longrightarrow} 1. \end{aligned}$$

(ii) We show that for $r := \mathrm{rk}(G) \leq n$,

$$Pr(\mathrm{coker}(A) \cong G) = \frac{1}{\#\mathrm{Aut}(G)}\left(\prod_{i=1}^{n}(1-p^{-i})\right)\left(\prod_{i=n-r+1}^{n}(1-p^{-i})\right). \quad (2.2)$$

The theorem will then follow from theorem 2.1.2 by taking $n \to \infty$.

Let Γ be a submodule of $\mathbb{Z}_p^n$ such that $\mathbb{Z}_p^n/\Gamma \cong G$. We compute the probability for the event $\mathrm{im(A)} = \Gamma$.

Let A_0 be a matrix with $\mathrm{im(A_0)} = \Gamma$. Then we have the identification $\{A \mid \mathrm{im(A)} = \Gamma\} = \mathrm{A_0 \cdot GL_n(\mathbb{Z}_p)}$. By the properties of the Haar measure (cf. [KS99]),

$$\begin{aligned} Pr(\mathrm{im(A)} = \Gamma) &= |\det(A_0)|^{-n} \cdot Pr(A \in \mathrm{GL}_n(\mathbb{Z}_p)) \\ &= \frac{Pr(A \text{ is invertible})}{(\mathrm{ord}(G))^n}. \end{aligned}$$

Since a matrix is invertible if and only if its reduction mod p is invertible, we have

$$\begin{aligned} Pr(A \text{ is invertible}) &= \frac{\#\mathrm{GL}_n(\mathbb{Z}/p\mathbb{Z})}{\#(\mathbb{Z}/p\mathbb{Z})^{n\times n}} \\ &\overset{1.2.12}{=} \frac{p^{n^2}\prod_{i=1}^{n}(1-p^{-i})}{p^{n^2}} \\ &= \prod_{i=1}^{n}(1-p^{-i}). \end{aligned}$$

In particular, this probability is independent of Γ. Therefore,

$$\begin{aligned} Pr(\mathrm{coker}(A)\cong G) &= \#\{\Gamma\subseteq\mathbb{Z}^n \mid \mathbb{Z}_p/\Gamma = G\}\cdot Pr(\mathrm{im(A)}=\Gamma) \\ &\overset{1.2.14}{=} \frac{(\mathrm{ord}(G))^n}{\#\mathrm{Aut}(G)}\left(\prod_{i=n-r+1}^{n}(1-p^{-i})\right)\frac{\prod_{i=1}^{n}(1-p^{-i})}{(\mathrm{ord}(G))^n} \\ &= \frac{1}{\#\mathrm{Aut}(G)}\left(\prod_{i=1}^{n}(1-p^{-i})\right)\left(\prod_{i=n-r+1}^{n}(1-p^{-i})\right). \end{aligned}$$

□

2.3 Computing special values

The explicit formulas (theorems 1.2.10 and 2.1.2) enable us to compute some values rather easily. For example, given a group $G \in \mathcal{G}_p$, we are given an explicit formula for $P(G)$. As a special case, let me give the probability that a p-group is the trivial group 0. Since $w(0) = 1$, we obtain

$$P(0) = \prod_{i=1}^{\infty}(1-p^{-i}).$$

Using q-series identities, we may compute some other probabilities. For example, the probability that a random group is cyclic (i.e., has rank ≤ 1), is (with $q = \frac{1}{p}$, as usual)

$$\begin{aligned} P(G \text{ cyclic}) &= \frac{1}{w(\mathcal{G}_p)} \sum_{G \text{ cyclic}} w(G) \\ &= \left(\prod_{i=1}^{\infty} (1-q^i) \right) \sum_{e=0}^{\infty} \frac{q^e}{1-q} \\ &= \left(\prod_{i=1}^{\infty} (1-q^i) \right) \frac{1}{(1-q)^2} \\ &= \frac{1}{1-q} \prod_{i=2}^{\infty} (1-q^i) \\ &= \frac{p}{p-1} \prod_{i=2}^{\infty} (1-p^{-i}). \end{aligned}$$

The calculation was pleasantly simple. Now let us compare this to what happens if we try to treat the slightly more complicated question of how likely it is for a random group to have rank 2. Within the computation we distinguish two different cases, corresponding to the possible group structures $G = (\mathbb{Z}/p^e)^2$, and $G = \mathbb{Z}/p^{e_1} \times \mathbb{Z}/p^{e_2}$, $e_1 > e_2$:

$$\begin{aligned} P(\mathrm{rk}(G)=2) &= \frac{1}{w(\mathcal{G}_p)} \sum_{\mathrm{rk}(G)=2} w(G) \\ &= \left(\prod_{i=1}^{\infty} (1-q^i) \right) \left(\sum_{e=1}^{\infty} \frac{q^{4e}}{(1-q)(1-q^2)} + \sum_{e_2=1}^{\infty} \sum_{e_1=e_2+1}^{\infty} \frac{q^{e_1+3e_2}}{(1-q)^2} \right) \\ &= \left(\prod_{i=1}^{\infty} (1-q^i) \right) \Bigg(\frac{q^4}{(1-q)(1-q^2)(1-q^4)} + \\ &\qquad + \frac{1}{(1-q)^2} \sum_{e_2=1}^{\infty} q^{3e_2} q^{e_2+1} \frac{1}{(1-q)} \Bigg) \\ &= \left(\prod_{i=1}^{\infty} (1-q^i) \right) \left(\frac{q^4}{(1-q)(1-q^2)(1-q^4)} + \frac{q^5}{(1-q)^3(1-q^4)} \right) \\ &= \left(\prod_{i=1}^{\infty} (1-q^i) \right) \frac{q^4 - q^5 + q^5 + q^6}{(1-q)^2(1-q^2)(1-q^4)} \\ &= \left(\prod_{i=1}^{\infty} (1-q^i) \right) \frac{q^4}{(1-q)^2(1-q^2)^2}. \end{aligned}$$

Recalling that this was still one of the "easier" cases, we see that this approach soon becomes quite cumbersome. It *is* possible to get general results about order and rank of a random group in this way (Bernd Mehnert will present some of these calculations in his PhD-thesis [Meh]), but this requires a highly skillful handling of q-series identities, which we do not want to expect from the user. Furthermore, while the formulas for rank and orders still behave rather "tame", the exponent gives even nastier expressions, and I know of no direct calculation with q-series which gives general results for the exponent of a group in more than the simplest case $\exp(G) \leq 1$. (We will obtain such results by other techniques, cf. sections 3.4 and 4.6.4.)
So we need other tools to enhance our ability to compute interesting values. The next two chapters will provide such tools.

Chapter 3

The Cohen-Lenstra Heuristic and Partitions

This chapter presents a deep connection between the Cohen-Lenstra measure and partitions. The basic relationship (equation (3.1) on page 38) was already discovered by Cohen and Lenstra, but they didn't further investigate it. It was studied in more detail in 2006 by Prof. Ernst-Ulrich Gekeler, Bernd Mehnert, and myself, culminating in the notion of a *CL-map* (def. 3.2.1). The existence of such a map is trivial, at least with the knowledge from [CL84]. However, in our research seminar we conjectured that there should be some "natural" such map, which should give us additional insight into the Cohen-Lensta probability. In definition 3.2.7, I will replace the vague term of a "natural" CL-map by the precise notion of an *order-preserving CL-map*. Then I show the existence and explicit construction of such a map (sect. 3.3), demonstrate its usage by drawing some conclusions about the exponent of a random group (sect. 3.4), and finally I discuss its uniqueness (sect. 3.5), although this last point still contains open questions.

Throughout the chapter, p denotes a fixed prime and we use the identification $\mathcal{G}_p \cong \mathcal{G}_{\mathcal{P}}$, so groups are given by certain tuples of integers (def. 1.2.8). Since there is no need to distinguish between the two sets in this chapter, I will write (under slight abuse of notation) $G \in \mathcal{G}_{\mathcal{P}}$ for finite abelian p-groups G. For this chapter, I expect that the reader is familiar with the basic properties of partitions, as presented in section 1.3. In particular, recall that $\mathcal{P}$ denotes the set of all partitions.
It turns out that for this chapter it is convenient to complement the notion of a partition by the equivalent notion of a *derivation* (which is a non-standard notion). This concept is not used until section 3.3.2, but it is so extensively used in this and the subsequent sections that it is worth a definition at a prominent place.

3.1 Derivations

3.1.1 Definition. *Let* $\underline{n} = (\underline{n}_1, \underline{n}_2, \ldots, \underline{n}_k) \in \mathcal{P}$. *Its* derivation *is the tuple* $\overline{n} = (\overline{n}_0, \overline{n}_1, \ldots, \overline{n}_k)$ *defined as follows:*

$$\overline{n}_i := \begin{cases} \underline{n}_1 - \underline{n}_2 & , \text{ for } i = 0, \\ \underline{n}_i - \underline{n}_{i+2} & , \text{ for } 1 \leq i \leq k-2, \\ \underline{n}_i & , \text{ for } i = k-1, k. \end{cases}$$

We define $\mathcal{D}$ *to be the set of all derivations, so*

$$\mathcal{D} = \{\overline{n} \mid \exists \underline{n} \in \mathcal{P} \text{ s.t. } \overline{n} \text{ is the derivation of } \underline{n}\}.$$

Note that the derivation of a partition is indeed similar to a discrete gradient vector, for step size 2. That is why I have chosen the word "derivation" for these objects. Only the $\overline{n}_0$-entry deviates a bit, but we could either neglect this entry (cf. lemma below), or we extend the partition by setting $\underline{n}_0 := \underline{n}_1$. Furthermore, if we fill up $\underline{n}$ with 0's on the right (i.e., $\underline{n}_i := 0$ for $i > k$), then we may define its derivation by the single formula $\overline{n}_i := \underline{n}_i - \underline{n}_{i+2}$.

The next lemma shows that partitions and derivation are in one-to-one-correspondence.

3.1.2 Lemma. *Let* $\underline{n} = (\underline{n}_1, \underline{n}_2, \ldots, \underline{n}_k) \in \mathcal{P}$ *and let* $\overline{n} = (\overline{n}_0, \overline{n}_1, \ldots, \overline{n}_k) \in \mathcal{D}$ *be its derivation.*
Then it is possible to recover $\underline{n}$ *from* $\overline{n}$ *by the following formula:*

$$\underline{n}_i = \sum_{\substack{i \leq j \leq k \\ i \equiv j \bmod 2}} \overline{n}_j, \qquad 1 \leq i \leq k.$$

Proof. For $i = k$ and $i = k-1$ the formula is obvious. For $i \leq k-2$ we have

$$\sum_{\substack{i \leq j \leq k \\ i \equiv j \bmod 2}} \overline{n}_j = \left(\sum_{\substack{i \leq j \leq k-2 \\ i \equiv j \bmod 2}} \underline{n}_j - \underline{n}_{j+2} \right) + \begin{cases} \underline{n}_k & , \text{ if } i \equiv k \bmod 2 \\ \underline{n}_{k-1} & , \text{ otherwise} \end{cases} = \underline{n}_i.$$

□

Note that we did not need $\overline{n}_0$ to recover $\underline{n}$.

3.1.3 Example. The tuple

$$\overline{n} = (2, 2, 0, 1, 4, 5, 2, 0, 1, 1, 1, 1)$$

is a derivation, and its associated partition is

$$\underline{n} = (10, 8, 8, 8, 7, 4, 2, 2, 2, 1, 1).$$

In order to work with them, we need to classify all derivations. This is done by the following lemma:

3.1.4 Lemma. *A sequence $\overline{n} = (\overline{n}_0, \overline{n}_1, \ldots, \overline{n}_k)$ is a derivation if and only if the following holds:*

(i) $\overline{n}_{k-1} \geq \overline{n}_k > 0$,

(ii) $\overline{n}_i \geq 0$ *for all* $i \geq 1$,

(iii) $\sum_{j=i_1}^{i_2} (-1)^{j-i_1} \overline{n}_j \geq 0$ *for all* $1 \leq i_1 < i_2 \leq k$, $i_1 \equiv i_2 \bmod 2$. *If* $i_2 = k$, *we drop the condition* $i_1 \equiv i_2 \bmod 2$.

(iv) $\sum_{j=0}^{k} (-1)^j \overline{n}_j = 0$.

Proof. Let $\overline{n}$ be the derivation of some $\underline{n} \in \mathcal{P}$. (i) and (ii) are obvious. Furthermore, we have by lemma 3.1.2

$$\overline{n}_0 = \underline{n}_1 - \underline{n}_2 = \left(\sum_{\substack{1 \leq j \leq k \\ i \equiv 1 \bmod 2}} \overline{n}_j \right) - \left(\sum_{\substack{2 \leq j \leq k \\ i \equiv 0 \bmod 2}} \overline{n}_j \right),$$

which implies (iv).
For (iii), let us for simplicity assume that $i_2 < k$. The case $i_2 = k$ is analogous. In the following calculation, if a term $\underline{n}_j$ does not exist because $j > k$, then it is to be replaced by 0:

$$\begin{aligned} \sum_{j=i_1}^{i_2} (-1)^{j-i_1} \overline{n}_j &= \left(\sum_{\substack{i_1 \leq j \leq i_2 \\ j \equiv i_1 \bmod 2}} \overline{n}_j \right) - \left(\sum_{\substack{i_1 \leq j \leq i_2 \\ j \not\equiv i_1 \bmod 2}} \overline{n}_j \right) \\ &= \underline{n}_{i_1} - \underline{n}_{i_2+2} - (\underline{n}_{i_1+1} - \underline{n}_{i_2+1}) \\ &= \underbrace{\underline{n}_{i_1} - \underline{n}_{i_1+1}}_{\geq 0} + \underbrace{\underline{n}_{i_2+1} - \underline{n}_{i_2+2}}_{\geq 0} \\ &\geq 0. \end{aligned}$$

Now we show that any tuple with properties (i)–(iv) is a derivation.
We define $\underline{n}$ as in lemma 3.1.2. Then we have

$$\overline{n}_0 \overset{(iv)}{=} \left(\sum_{\substack{1 \leq j \leq k \\ j \equiv 1 \bmod 2}} \overline{n}_j \right) - \left(\sum_{\substack{2 \leq j \leq k \\ j \equiv 0 \bmod 2}} \overline{n}_j \right) = \underline{n}_1 - \underline{n}_2.$$

Hence, if $\underline{n}$ is a partition, then it is clear by lemma 3.1.2 that $\overline{n}$ is its derivation.
By condition (i) and (ii), $\underline{n}_i \geq 0$ for $1 \leq i \leq k$. So we only need to show that $\underline{n}_i \geq \underline{n}_{i+1}$ for $1 \leq i \leq k-1$. For such i, we have

$$\begin{aligned} \underline{n}_i - \underline{n}_{i+1} &= \left(\sum_{\substack{i \leq j \leq k \\ j \equiv i \bmod 2}} \overline{n}_j \right) - \left(\sum_{\substack{i+1 \leq j \leq k \\ j \not\equiv i \bmod 2}} \overline{n}_j \right) \\ &= \sum_{j=i}^{k} (-1)^{j+i} \overline{n}_j \\ &\overset{(iii)}{\geq} 0. \end{aligned}$$

□

3.1.5 Remark.

- The equivalence remains true if we restrict (iii) to the case $i_2 = k$, i.e., if we replace (iii) by the condition $\sum_{j=i}^{k} (-1)^{j-i} \overline{n}_j \geq 0$ for all $1 \leq i \leq k$.
- We will especially make use of (iii) with $i_2 = i_1 + 2$. In this case the statement becomes
$$\overline{n}_{i_1} + \overline{n}_{i_1+2} \geq \overline{n}_{i_1+1}.$$

We have seen that the sets $\mathcal{P}$ and $\mathcal{D}$ are in bijection with each other, so I will use partitions and derivations interchangeably. In particular, by abuse of notation, I will apply functions that live on $\mathcal{P}$ to derivations and vice versa. Also, the partial ordering on $\mathcal{P}$ transfers to $\mathcal{D}$. To clarify things, I will throughout this thesis denote partitions by a lower bar and derivations by an upper bar ($\underline{n}$ and $\overline{n}$), and the same for their entries (e.g. $\underline{n}_1$ and $\overline{n}_1$). In this way, it will be easy to see with which object we are working at any point.

3.2 Why partitions? — CL-maps

We have seen that the Cohen-Lenstra measure of a finite abelian p-group

$$G = \prod_{i=1}^{k} \left(\mathbb{Z}/p^{e_i}\mathbb{Z}\right)^{r_i}, \text{ where } e_1 > e_2 > \ldots > e_k > 0$$

is

$$w(G) = \frac{1}{\mathrm{Aut}(G)} = \left(\prod_{i=1}^{k}\left(\prod_{s=1}^{r_i}(1-p^{-s})^{-1}\right)\right)\left(\prod_{1\le i,j\le k} p^{-\min(e_i,e_j)r_ir_j}\right).$$

Now we change the perspective. We introduce a new variable $q := \frac{1}{p}$. Of course, q is strictly speaking just a rational number, but from now on, I want to treat q rather as a formal variable. Then the above formula becomes

$$w(G) = \frac{1}{\mathrm{Aut}(G)} = \left(\prod_{i=1}^{k}\left(\prod_{s=1}^{r_i}(1-q^{s})^{-1}\right)\right)\left(\prod_{1\le i,j\le k} q^{\min(e_i,e_j)r_ir_j}\right),$$

and this is a *formal power series* in q.

Let me first explain why we do not lose any information by restricting ourselves to the *formal* variable q. Of course, we cannot create additional information: Whenever we derive a formula for q, then we just plug in $q = \frac{1}{p}$ and obtain a formula for p. On the other hand, suppose we have a formula (for example, formula (3.1) below) that is valid for all p. Then in terms of power series, we have an equality at the values $q = \frac{1}{p}$ for various p. Since all our power series have a positive radius of convergence, they give rise to holomorphic functions. But the Identity Theorem for holomorphic functions tells us that two holomorphic functions that coincide on the set $\{\frac{1}{p} \mid p \in \mathbb{P}\}$ must coincide everywhere, since the set has an accumulation point. Hence, the representing power series are identical. Altogether, we have shown that any identity valid for all p must necessarily also be valid for the formal variable q.

Hopefully, I could convince the reader that we do not lose anything by switching from p to q. But what do we gain? Let us consider the weight formula more closely. If we expand $(1-q^s)^{-1}$ as $\sum_{i=0}^{\infty}(q^s)^i$, we see that the weight is not only a power series in q, but it is a power series with *non-negative integers* as coefficients. I will call such a power series a *combinatorial power series* for the moment.

Keeping this in mind, we reinvestigate theorem 2.1.2, which states:

$$\sum_{\substack{G \text{ finite}\\ \text{abelian } p\text{-group}}} w(G) \;=\; \prod_{i=1}^{\infty}(1-q^i)^{-1}.$$

As I argued above, this equality holds not only for the particular values $q = \frac{1}{p}$, but is in fact an identity of power series. We continue:

$$\sum_{\substack{G \text{ finite} \\ \text{abelian } p\text{-group}}} w(G) \;=\; \prod_{i=1}^{\infty}(1-q^i)^{-1} \overset{\text{sect. } 1.3.1}{=} \sum_{n\in\mathbb{N}} \sum_{\substack{\underline{n} \text{ is a par-} \\ \text{tition of } n}} q^n. \tag{3.1}$$

Now the equation gets a new flavour. On the right hand side we have a sum over terms q^n. (I will always use the word "term" for a single monomial with coefficient 1 in this context.) On the left hand side, we have a sum of combinatorial power series in q, i.e., also a sum over terms q^n. The equality tells us that for each n we have the same number of terms on both sides.
This observation suggests that there might be some deeper connection. In fact, an identity of combinatorial power series often (but not always!) reflects some underlying bijection of combinatorial objects: If there are as many terms on the left hand side as there are on the right hand side, then one might expect that there exists some underlying *natural bijection* between the terms, whatever "natural" means in this context. This is the conjecture that came up in the research seminar by Gekeler, Mehnert, and myself.
How can we model such a bijection? We fix an n. Of course, we do not want to distinguish between the terms that belong to the same group G. After all, the weight is given by a power series, where we have no more information than *how many* terms q^n belong to a group. However, on the partition side, we do distinguish between different terms q^n. So for each partition, there should be a corresponding group on the left hand side. In other words, we need a map from the set of all partitions into the set of all finite abelian p-groups, telling us which power series the q^n-term of this partition belongs to. This leads to the following definition:

3.2.1 Definition. *A map $\Lambda : \mathcal{P} \to \mathcal{G}_{\mathcal{P}}$ is a* Cohen-Lenstra-map *(*CL-map*) if for any finite abelian p-group G,*

$$w(G) = \sum_{n\geq 0} a_G(n) q^n,$$

where

$$a_G(n) = \#\left(\left\{\Lambda^{-1}(G)\right\} \cap \{\underline{n} \in \mathcal{P} |\ \text{size}(\underline{n}) = n\}\right) \tag{3.2}$$

is the number of partitions of n that are mapped to G.

The existence of such a map is trivial. In fact, we know how many terms we need for each G, so we can just arbitrarily grab as many partitions as we need. Equation (3.1) guarantees that we have exactly as many partitions as we require in total for the various groups. But this process is so arbitrary

that we have little hope of finding a map that is natural enough to occur in applications. We need a way of saying when a map is natural.
Recall that the set of all partitions comes together with a natural partial ordering (def. 1.3.6). I will show that we also have a partial ordering on the set of all terms (= monomials) that occur in $w(G)$, for a fixed group G. We will then replace the vague concept of being "natural" by the rigorous requirement of being order-preserving.
To define the partial ordering on the set of all terms occurring in $w(G)$, recall that the weight of $G = \prod_{i=1}^{k} (\mathbb{Z}/p^{e_i}\mathbb{Z})^{r_i}$ is

$$\begin{aligned} w(G) &= \left(\prod_{i=1}^{k}\left(\prod_{s=1}^{r_i}(1-q^s)^{-1}\right)\right)\left(\prod_{1\leq i,j\leq k} q^{\min(e_i,e_j)r_ir_j}\right) \\ &= \left(\prod_{i=1}^{k}\left(\prod_{s=1}^{r_i}\sum_{t=0}^{\infty} q^{ts}\right)\right)\left(\prod_{1\leq i,j\leq k} q^{\min(e_i,e_j)r_ir_j}\right). \end{aligned}$$

If we multiply out, we get one monomial for each tuple $(t_{i,s})_{1\leq i\leq k, 1\leq s\leq r_i}$, where each entry is ≥ 0. So instead of listing all monomials, we can as well list all such tuples. We call the set of all such tuples I_G:

3.2.2 Definition. *Let $G = \prod_{i=1}^{k} (\mathbb{Z}/p^{e_i}\mathbb{Z})^{r_i}$ be a finite abelian p-group. We define the* index set *I_G of G as*

$$I_G := \{(t_{i,s})_{1\leq i\leq k, 1\leq s\leq r_i} \mid t_{i,s} \in \mathbb{N}\}.$$

For each vector $t \in I_G$, we define the exponent

$$\operatorname{expon}(t) := \sum_{1\leq i,j\leq k} \min(e_i,e_j)r_ir_j + \left(\sum_{i=1}^{k}\sum_{s=1}^{r_i} st_{i,s}\right).$$

3.2.3 Remark. By construction of the index set and the exponent, we may write the weight of G as

$$w(G) = \sum_{t\in I_G} q^{\operatorname{expon}(t)}.$$

3.2.4 Example.

- Let $G = \mathbb{Z}/p^e\mathbb{Z}$ be a cyclic group. Then $I_G = \{(t_{1,1})\} \cong \mathbb{N}$. The exponent of $t = (t_{1,1}) \in I_G$ is

$$\operatorname{expon}(t) = e + t_{1,1}.$$

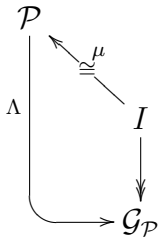
- Let $G = \mathbb{Z}/p\mathbb{Z} \times (\mathbb{Z}/p^4\mathbb{Z})^3$. Then $I_G = \{(t_{1,1}, t_{2,1}, t_{2,2}, t_{2,3})\} \cong \mathbb{N}^4$. The exponent of an element $t \in I_G$ is

$$\operatorname{expon}(t) = 43 + t_{1,1} + t_{2,1} + 2t_{2,2} + 3t_{2,3}.$$

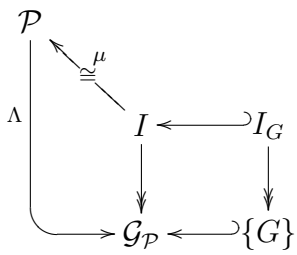
3.2.5 Definition.
The set I_G comes with a natural partial ordering: $(t_{i,s})_{1\le i\le k, 1\le s\le r_i} \in I_G$ is smaller or equal *than $(\tilde{t}_{i,s})_{1\le i\le k, 1\le s\le r_i} \in I_G$ if for all $1 \le i \le k$ the sequence $(t_{i,r_i}, t_{i,r_i-1}, \ldots, t_{i,1})$ is dominated by $(\tilde{t}_{i,r_i}, \tilde{t}_{i,r_i-1}, \ldots, \tilde{t}_{i,1})$. (Note the order of the sequences!) We denote this partial ordering by $\le$ (or $<$, if we do not allow equality), and we also say that $\tilde{t}$* lies above *t and t* lies below *$\tilde{t}$. If $t < \tilde{t}$ and there is no $u \in I_G$ such that $t < u < \tilde{t}$, then we say $\tilde{t}$* lies immediately above *t or t* lies immediately below *$\tilde{t}$.*

3.2.6 Remark. Let $t, \tilde{t} \in I_G$ such that $t < \tilde{t}$. Then $\operatorname{expon}(t) < \operatorname{expon}(\tilde{t})$. The proof is easy and is left to the reader.

Now we are ready to sort out things. Consider the formal disjoint union $I := \dot{\bigcup}_{G\in\mathcal{G}_\mathcal{P}} I_G$. By definition of the sets I_G, we may rephrase the CL-property as follows: A map $\Lambda : \mathcal{P} \to \mathcal{G}_\mathcal{P}$ is a CL-map if and only if there exists a bijective map $\mu : I \to \mathcal{P}$ such that $\operatorname{size}(\mu(t)) = \operatorname{expon}(t)$, and such that the diagram

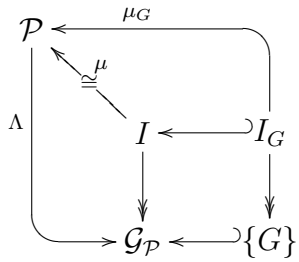


commutes, where the map $I \to \mathcal{G}_\mathcal{P}$ is simply projection on the index, so $I_G \ni t \mapsto G \in \mathcal{G}$. This projection may be characterized to be the unique map such that the extended diagrams



commute for all G. We see that maps $I_G \to \mathcal{P}$ are contained in this diagram, and these are maps between *partially ordered* sets. Now it becomes clear how the definition of order-preserving should look like:

3.2.7 Definition.
A CL-map $\Lambda : \mathcal{P} \to \mathcal{G}_\mathcal{P}$ is called order-preserving *if there exists a bijective map $\mu : I \to \mathcal{P}$ such that* $\mathrm{size}(\mu(t)) = \mathrm{expon}(t)$*, such that for each $G \in \mathcal{G}_\mathcal{P}$ the diagram*



commutes, and such that μ_G is monotone for each $G \in \mathcal{G}_\mathcal{P}$. In the diagram, $I := \dot{\bigcup}_{G \in \mathcal{G}_\mathcal{P}} I_G$.
Rephrasing the condition of being monotone for μ_G, we require that $t < t'$ implies $\mu_G(t) < \mu_G(t')$ for all groups G and all $t, t' \in I_G$.
Summarizing, the map

$$\begin{aligned} \mu : \dot{\bigcup_G} I_G &\to \mathcal{P} \\ I_G \ni t &\mapsto \mu_G(t) \end{aligned}$$

is a bijection which preserves order on each fiber.

3.2.8 Definition. *For any order-preserving CL-map Λ, let $\iota : \mathcal{G}_\mathcal{P} \to \mathcal{P}$ be the section of Λ such that $\iota(G)$ is the (unique) smallest element in $\Lambda^{-1}(G)$. We call ι the* canonical section *of Λ, and by $\mathcal{P}_{base} := \mathcal{P}_{base}(\Lambda) := \iota(\mathcal{G}_\mathcal{P})$, we denote its image.*

3.2.9 Remark.

- It is possible to define the canonical section also for CL-maps that are not necessarily order-preserving: For all G, the smallest non-zero coefficient in $w(G)$ is 1. Hence, for any CL-map, there is a unique element *of minimal size* in the fiber of G. (If Λ is order-preserving, then this element is even *minimal*.) So we can define the canonical section by assigning to each group the element in its fiber of minimal size.

- When we talk about CL-maps being natural, there is another condition one could impose: We could require that the canonical section is order-preserving as well, i.e., $G_1 \subseteq G_2$ implies $\iota(G_1) \leq \iota(G_2)$. Think of this as a "horizontal" property (a property for a section through all fibers), whereas definition 3.2.7 is a "vertical" property (we consider each fiber individually).

  I did not include this "horizontal" order-preserving property in my definitions for two reasons: Firstly, to keep things as simple as possible. Secondly, it does not seem to make a big difference. In particular, the CL-map Λ that we will define below (section 3.3.1 and 3.3.2) is order-preserving in this stronger sense, and so are all the following examples, including the examples which show that Λ is not unique (example 3.5.4).

  However, I do not want to claim that my definition is superior to the other variant. Rather, from my knowledge I can not decide which of the two possible definitions is the better one, and so I have picked one due to my personal taste.

  Let me give one warning: A natural idea would be to require some "horizontal" property that takes not only the canonical section – the "lowest level" – into account, but also partitions lying above that. However, there are two severe problems: Firstly, it is clearly true that, for $G_1 \hookrightarrow G_2$, $\mathrm{Aut}(G_2)$ is larger than $\mathrm{Aut}(G_1)$ because we have a canonical injection and surjection $\mathrm{Aut}(G_2) \hookleftarrow \{\varphi \in \mathrm{Aut}(G_2) \mid \mathrm{im}(\varphi) \subseteq \mathrm{G}_1\} \twoheadrightarrow \mathrm{Aut}(\mathrm{G}_1)$, respectively. However, it is *not* true in general that in this case $w(G_1)$ is bigger than $w(G_2)$ as power series (i.e., coefficient-wise). It already fails for $(\mathbb{Z}/p\mathbb{Z})^3 \hookrightarrow (\mathbb{Z}/p^2\mathbb{Z}) \times (\mathbb{Z}/p\mathbb{Z})^2$. So it is not clear how the more general horizontal property should look like.

  Secondly, does it make sense at all to talk about embeddings $G_1 \hookrightarrow G_2$ when G_1 and G_2 are only defined up to isomorphism? For example, there are essentially different embeddings $\mathbb{Z}/p\mathbb{Z} \hookrightarrow \mathbb{Z}/p\mathbb{Z} \times \mathbb{Z}/p^2\mathbb{Z}$. Do we want to distinguish those embeddings? Those are the problems one would need to address.

  A solution might be a combination of the notion of CL-maps with the Markov chain approach in section 4.3, but at present it is not clear how such a combination can be achieved.

The next question to address is whether natural CL-maps exist. In the next section, we will explicitly construct an order-preserving CL-map Λ. This will be done in two ways: via Young diagrams and numerically.

3.3 The existence of an order-preserving CL-map Λ

3.3.1 Definition of Λ (via Young diagrams)

As announced, this subsection contains the first of two possible definitions of the order-preserving CL-map Λ. Basically, this method stays on partition level, whereas the other definition works with derivations.
This Young diagram definition is probably harder to read than the one in the next section. However, since the notion of derivations is a non-standard invention, I find it convenient also to give a version working directly with partitions. Moreover, it has one further advantage: As example 3.3.5 indicates, the Young diagram algorithm is closely linked to the canonical section $\mathcal{P}_{base}$ of Λ.
First I introduce a new notation (which we use only for this algorithm). Recall that a box in the Young diagram is described by a pair $(i,j) \in \mathbb{N}^+ \times \mathbb{N}^+$.

3.3.1 Notation. Let $(i,j) \in \mathbb{Z} \times \mathbb{Z}$, and let $\lambda \in \mathbb{Z}$. The *λ-successor* $s_\lambda(i,j)$ of (i,j) is the point $(i+2, j-\lambda) \in \mathbb{Z} \times \mathbb{Z}$. For any $M \subset \mathbb{Z} \times \mathbb{Z}$, let $s_\lambda(M)$ be the image of M under s_λ.

Now Λ can be defined by the following algorithm:

3.3.2 Algorithm. Let $\underline{n} \in \mathcal{P}$.

1. Let $M_1 \subset \mathbb{N}^+ \times \mathbb{N}^+$ be the Young diagram of $\underline{n}$. Put $k := 1$.
2. Let $Q_k := \{(i,j) \in \mathbb{Z} \times \mathbb{Z} |\ j \geq 1,\ i \geq 2k-1\}$.
   Find $\lambda_k \in \mathbb{Z}$ minimal such that $s_{\lambda_k}(M_k) \cap Q_k \subset M_k$.
3. Find the maximal $i_k \in \mathbb{Z}$ such that there is a $j \in \mathbb{Z}$ with:
   - $(i_k, j) \in M_k$ and
   - $s_{\lambda_k - 1}(i_k, j) \in Q_k \setminus M_k$.
4. Let $C_k := \{(i,j) \in \mathbb{N}^+ \times \mathbb{N}^+ |\ i \leq i_k\} \setminus M_k$.
   Put $M_{k+1} := (M_k \setminus s_{\lambda_k}(C_k)) \cap Q_{k+1}$.
   Increase k by 1.
5. Repeat step 2–4 until $M_k \cap Q_k$ is empty.

If the algorithm terminates after k loops, it returns integers $\lambda_1, \ldots, \lambda_k$.
Put $\Lambda(\underline{n}) := (\lambda_1, \lambda_2, \ldots, \lambda_k) \in \mathcal{G}_{\mathcal{P}}$.

3.3.3 Remark.

- The algorithm always terminates, so Λ is well-defined.
- The λ_i are sorted: $\lambda_1 \geq \ldots \geq \lambda_k > 0$. So the algorithm outputs a partition.

If one wants to analyze the algorithm directly, the following facts are helpful. Otherwise, these facts may be ignored. We do not use them, since we will prove the correctness of the algorithm in an indirect way in section 3.3.2, after having shown the equivalence of the Young diagram algorithm and the numerical algorithm.

3.3.4 Remark. In the k-th loop, define

$$a_k := \#M_k - \#M_{k+1} - \#\{(i,j) \in M_k \mid i = 2k+1\} \\ -\#\{(i,j) \in M_{k+1} \mid i = 2k+3\}.$$

The a_k quantify the difference between M_k and M_{k+1}, where the two latter terms compensate (roughly speaking) for the two highest lines, which are cut off from M_{k+1}.
Define further $j_{k,\max} := \max\{j \mid \exists i \text{ s.t. } (i,j) \in M_k\}$. Then in each step after the first we have the invariant

$$n = |(M_{k+1})| + 2kj_{k,\max} + \left(\sum_{i=1}^{k} \lambda_i(2i-1)\right) + \sum_{i=1}^{k} a_i.$$

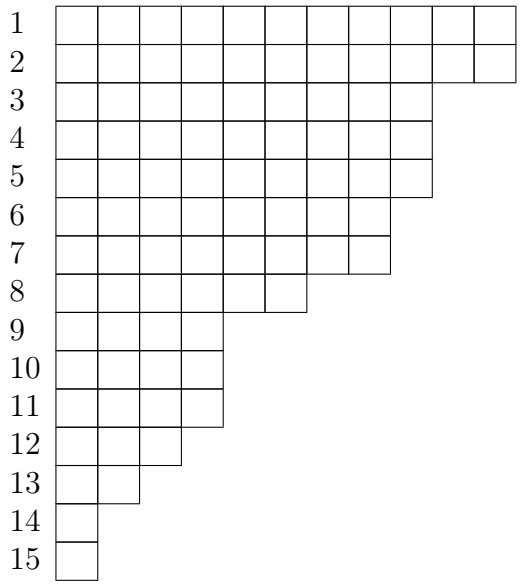
This assertion can easily be proven by induction.
In particular, after termination the first two terms will vanish, so we get

$$n = \left(\sum_{i=1}^{k} \lambda_i(2i-1)\right) + \sum_{i=1}^{k} a_i.$$

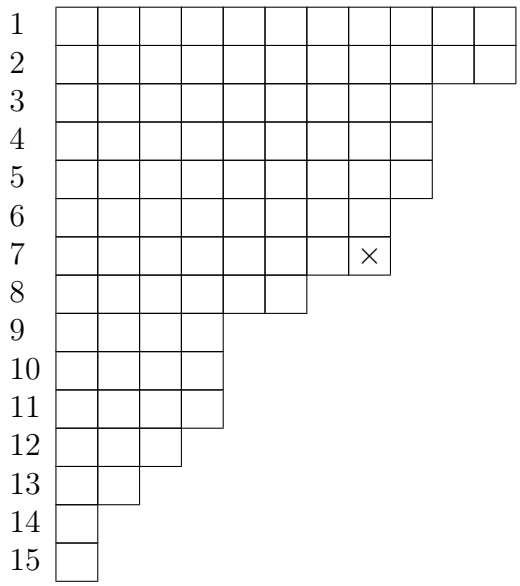
3.3.5 Example. Let us consider the partition

$$\underline{n} = (11, 11, 9, 9, 9, 8, 8, 6, 4, 4, 4, 3, 2, 1, 1).$$

Its Young diagram is

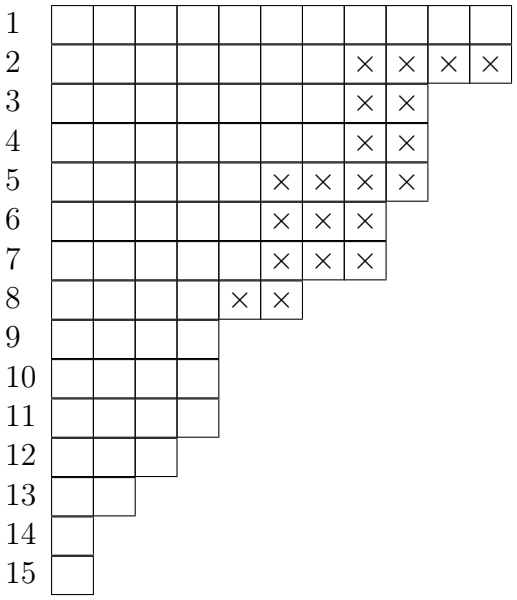


In each round, I will give a partition $\underline{n}^k$ that reflects M_k in the following sense: If you draw the Young diagram of $\underline{n}^k$ and intersect it with Q_k (i.e., you forget the $2k-2$ uppermost lines), then you get M_k. For the first round, we simply have $\underline{n}^1 = \underline{n}$.
Don't be confused by the fact that k is used as an upper index. As $\underline{n}^k$ is a partition, we will need lower indices to address its entries.
At the beginning, Q_k is the whole quadrant, so we consider the whole diagram. We find that $\lambda_1 = 4$ and $i_1 = 7$, because the box $(7, 8)$ is in M_1, but $s_{4-1} = s_3$ maps $(7, 8)$ to $(9, 5) \notin M_1$:



The boxes that are marked in the next diagram will be removed according

to step 4 of the algorithm. Note that also the two uppermost lines will be removed, so it is not really necessary to mark any box in line 2. However, in this way the number of marked boxes is exactly a_1 (cf. remark 3.3.4). (In general, in the i-th step the number of marked boxes will be a_i.)
The reader who is not interested in the proof may ignore these data.

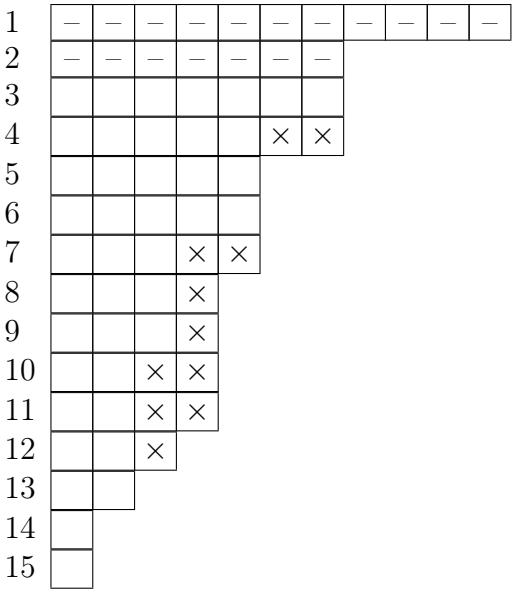


So we get the partition

$$\underline{n}^2 = (11, 7, 7, 7, 5, 5, 5, 4, 4, 4, 4, 3, 2, 1, 1).$$

Now we find that $\lambda_2 = 2$ and $i_2 = 12$, because the box $(12, 3)$ is not mapped into M_2 by $s_{\lambda_2 - 1}$. Remember that, in order to find λ_2 and i_2, we must ignore line 1 and 2, because they do not belong to Q_2. In the following, I will mark the lines not belonging to Q_k with "–".

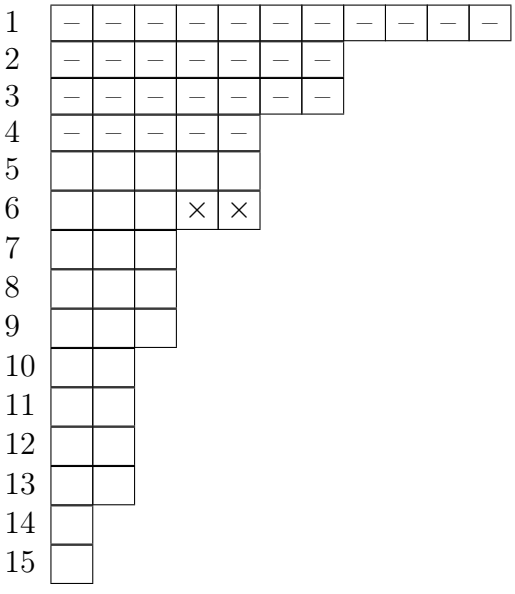
Again, we label the boxes which are going to be removed. And again, the reader who is not interested in the proof may ignore line 3 and 4:



We obtain

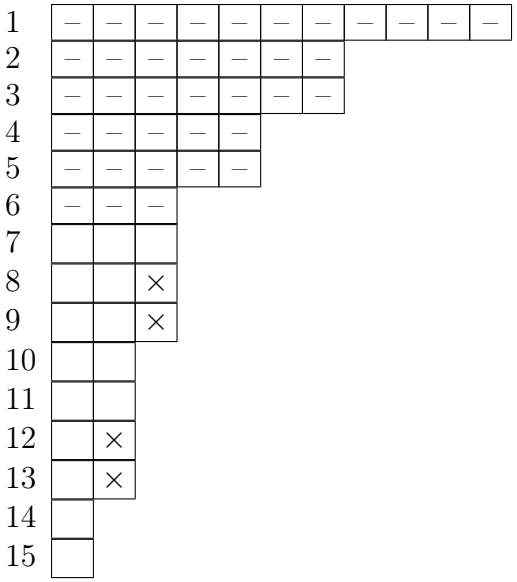
$$\underline{n}^3 = (11, 7, 7, 5, 5, 5, 3, 3, 3, 2, 2, 2, 2, 1, 1).$$

Now we look at M_3 and find $\lambda_3 = 2$ and $i_3 = 6$:



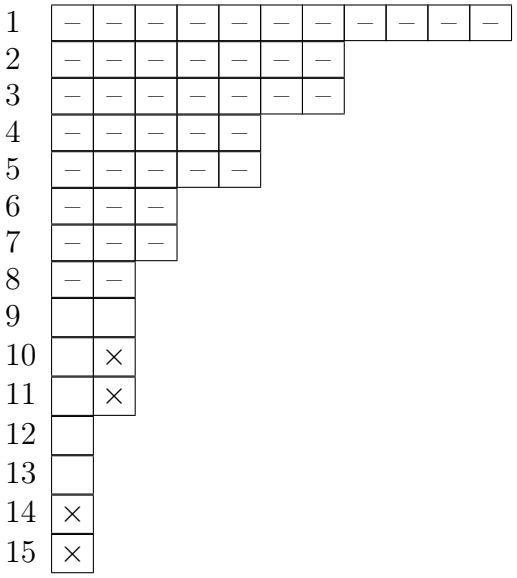
$$\underline{n}^4 = (11, 7, 7, 5, 5, 3, 3, 3, 3, 2, 2, 2, 2, 1, 1).$$

We get $\lambda_4 = 1$, $i_4 = 15$:



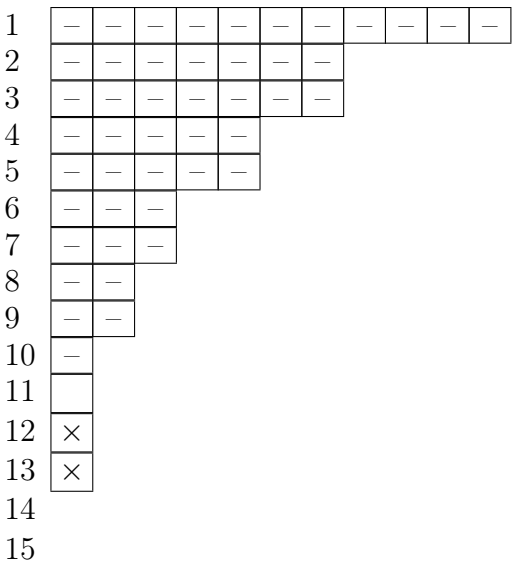
$$\underline{n}^5 = (11, 7, 7, 5, 5, 3, 3, 2, 2, 2, 2, 1, 1, 1, 1).$$

Next, $\lambda_5 = 1$, $i_5 = 15$:



$$\underline{n}^6 = (11, 7, 7, 5, 5, 3, 3, 2, 2, 1, 1, 1, 1, 0, 0).$$

Finally, $\lambda_6 = 1$, $i_6 = 13$:



$$\underline{n}^7 = (11, 7, 7, 5, 5, 3, 3, 2, 2, 1, 1, 0, 0, 0, 0).$$

M_7 is empty, so the algorithm has terminated and yields:

$$\Lambda(\underline{n}) = (\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5, \lambda_6) = (4, 2, 2, 1, 1, 1) \in \mathcal{G}_{\mathcal{P}}.$$

So the algorithm outputs the group

$$\mathbb{Z}/p^4\mathbb{Z} \times \mathbb{Z}/p^2\mathbb{Z} \times \mathbb{Z}/p^2\mathbb{Z} \times \mathbb{Z}/p\mathbb{Z} \times \mathbb{Z}/p\mathbb{Z} \times \mathbb{Z}/p\mathbb{Z}.$$

As mentioned above, I have given more information than is needed for running the algorithm. Apart from finding the quantities a_i as the number of crossed boxes in each step, there is another feature: Considering the final partition $\underline{n}^7$, we find that it is in the set $\mathcal{P}_{base}$, (cf. defintion def:canonical-section and the remark thereafter). More precisely, it equals $\iota((4, 2, 2, 1, 1, 1))$, i.e., it is the smallest partition that is mapped to $(4, 2, 2, 1, 1, 1)$ by Λ.
This is a general rule: By defining all entries of $\underline{n}^k$ in an appropriate way (not only those determined by the algorithm), we obtain in the end a partition that satisfies $\underline{n}^k = \iota(\Lambda(\underline{n}))$. So we may view the algorithm as a way of removing successively parts of the partition until we end up with an element of $\mathcal{P}_{base}$.
What is the appropriate way to defining the missing entries of $\underline{n}_i^k$? The entries are determined by the algorithm for $i \geq 2k - 1$. For $i \leq 2k - 3$, we don't change anything: $\underline{n}_i^k := \underline{n}_i^{k-1}$. For $i = 2k - 2$, we set $\underline{n}_{2k-2}^k := \underline{n}_{2k-1}^{k-1} - \lambda_{k-1} = \underline{n}_{2k-1}^k$. Since we will not make use of this idea, I will not discuss the details any further.

3.3.2 Definition of Λ (numerical)

Now to the numerical definition of Λ:

3.3.6 Algorithm (numerical)**.** Let $\underline{n} = (\underline{n}_1, \underline{n}_2, \ldots, \underline{n}_m) \in \mathcal{P}$. The algorithm works as follows:

1. Let $\overline{n}^1 := \overline{n}$ be the derivation of $\underline{n}$ (cf. 3.1). Let $k := 1$.

2. Let $\lambda_k := \max_l \{\overline{n}^k_l\}$, and let $i_k := \max\{l |\ \overline{n}^k_l = \lambda_k\}$.

3. Remove the entries with indices $i_k - 1$, i_k and $i_k + 1$ from $\overline{n}^k$ and replace them by the single new entry $\overline{n}^k_{i_k-1} + \overline{n}^k_{i_k+1} - \overline{n}^k_{i_k}$, thereby getting $\overline{n}^{k+1}$. For convenience, shift indices such that the index i of $\overline{n}^k$ runs from $2k-2$ to m.

   Increase k by 1.

   (We might need to use some $\overline{n}^k_l$ that is out of range at this point. In this case, we may add a 0 on the right. The invariants given below guarantee that this cannot happen on the left.)

4. Repeat steps 2 and 3 until $\overline{n}^k$ consists only of zeros.

The output of the algorithm is $(\lambda_1, \lambda_2, \ldots, \lambda_{k-1}, \lambda_k) \in \mathcal{G}_{\mathcal{P}}$.

3.3.7 Remark.

- For each k, $\overline{n}^k$ is a derivation (possibly up to some ending 0's, which may be ignored).

  In particular, we have $\overline{n}^k_{i-1} + \overline{n}^k_{i+1} \geq \overline{n}^k_i$ for all i, k (see 3.1.5).

- In loop k, all values in $\overline{n}^k$ are integers between 0 and λ_{k-1}. Thus the λ_k are monotonically decreasing, so the algorithm indeed outputs a partition.

  These statements will be proven after the illustrating example.

- One could choose other index conventions. For the algorithm, it would be no difference if in loop k, i runs e.g. from 0 to $m-2k+2$. However, our convention has advantages that will become obvious in the proofs of theorems 3.3.9 and 3.3.14.

- I have defined i_k to be the *maximal* index of a maximal element. In fact, one could allow to use *any* index of a maximal entry. The chosen

convention has two advantages: Firstly, it makes the algorithm deterministic, and secondly, in this way it coincides stepwise with the Young diagram algorithm, as we will see later.

If you would change this algorithm (e.g., by picking the *minimal* index of a maximal element), of course you could translate it into a Young diagram algorithm. However, be warned: *Small changes in the derivations tend to have big effects in the Young diagrams!* In particular, in the Young diagram algorithm it is *not* possible to simply pick another than the maximal i_k (cf. 3.3.2).

There is another reason why it is preferable to choose i_k to be maximal. As we will see later, this version is "more closely" related to the formula of the size of the automorphism groups. I will explain more precisely what I mean by this in remark 3.3.22.

- This form of the algorithm is much handier and should be used for computations rather than the Young diagram version.

3.3.8 Example. Let $\underline{n} = (11, 11, 9, 9, 9, 8, 8, 6, 4, 4, 4, 3, 2, 1, 1)$. This is the same partition as in example 3.3.5 for the Young diagram algorithm.
I mark the places where something will happen in the next step by bold type.
We compute

$$\overline{n}^1 = \overline{n} = (0, 2, 2, 0, 1, 1, \mathbf{2}, \mathbf{4}, \mathbf{2}, 0, 1, 2, 2, 1, 1, 1).$$

Obviously, $\lambda_1 = 4$ and $i_1 = 7$. (Recall that the first entry of $\overline{n}$ has index 0.)
We have to replace the part $2, 4, 2$ by the single entry $2 + 2 - 4 = 0$, getting:

$$\overline{n}^2 = (0, 2, 2, 0, 1, 1, 0, 0, 1, \mathbf{2}, \mathbf{2}, \mathbf{1}, 1, 1),$$

with indices running from 2 to 15.
We see that $\lambda_2 = 2$ and $i_2 = 12$. We replace $2, 2, 1$ by 1:

$$\overline{n}^3 = (0, \mathbf{2}, \mathbf{2}, \mathbf{0}, 1, 1, 0, 0, 1, 1, 1, 1),$$

with indices running from 4 to 15.
$\lambda_3 = 2$, $i_3 = 6$, so we must replace $2, 2, 0$ by 0:

$$\overline{n}^4 = (0, 0, 1, 1, 0, 0, 1, 1, \mathbf{1}, \mathbf{1}),$$

with indices running from 6 to 15.

Now $\lambda_4 = 1$ and $i_4 = 15$. We fill up one 0 at the right and replace $1, 1, 0$ by 0:

$$\overline{n}^5 = (0, 0, 1, 1, 0, 0, \mathbf{1}, \mathbf{1}, \mathbf{0}),$$

with indices running from 8 to 16.
$\lambda_5 = 1$, $i_5 = 15$ and we replace $1, 1, 0$ by 0:

$$\overline{n}^6 = (0, 0, \mathbf{1}, \mathbf{1}, \mathbf{0}, 0, 0),$$

with indices running from 10 to 16.
Finally, $\lambda_6 = 1$, $i_6 = 13$, and after replacing one last time, we get a sequence of zeros:

$$\overline{n}^7 = (0, 0, 0, 0, 0),$$

so we are done.
The result is $(\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5, \lambda_6) = (4, 2, 2, 1, 1, 1) \in \mathcal{G}_{\mathcal{P}}$, which corresponds to the group $\mathbb{Z}/p^4\mathbb{Z} \times \mathbb{Z}/p^2\mathbb{Z} \times \mathbb{Z}/p^2\mathbb{Z} \times \mathbb{Z}/p\mathbb{Z} \times \mathbb{Z}/p\mathbb{Z} \times \mathbb{Z}/p\mathbb{Z}$.

3.3.9 Theorem. *The output of the above algorithm coincides with the output of the Young diagram algorithm in section 3.3.1.*

Proof. One could say that both algorithms perform the same computation, only the Young diagram algorithm computes directly on the partition, whereas the numerical algorithm computes on the derivations. More precisely, we will prove:

$$\text{For all } k,\ \overline{n}^k \text{ is the derivation of } (\underline{n}^k)_{\geq 2k-1}. \tag{3.3}$$

Here, $\overline{n}^k$ is the intermediary result in loop k of the numerical algorithm, and $(\underline{n}^k)_{\geq 2k-1}$ is defined by the following property: If you draw the Young diagram of $\underline{n}^k$ and intersect it with Q_k (i.e., you forget the $2k-2$ uppermost lines), then you get M_k from the Young diagram algorithm (3.3.2). The starting partition $\underline{n}^1$ is defined to be $\underline{n}$. The lower index indicates that we cut off the first entries from $\underline{n}^k$. I have included $\underline{n}$ in example 3.3.5.
The above properties define $\underline{n}^k_i$ only for $i \geq 2k-1$. Also, $\overline{n}^k_i$ is only defined for $i \geq 2k-2$. In fact, this could be extended. In example 3.3.5, I have explained how to define $\underline{n}^k_i$ for all i, and in algorithm 3.3.18, we will find a way to define $\overline{n}^k_i$ for all i. With these definitions, property (3.3) holds in fact for the whole vectors. However, the cut-off version given above is sufficient for our purposes.
To prove (3.3), we use induction on k. The case $k = 1$ is clear from step 1 in the numerical algorithm. So let $k > 1$ and assume that $\overline{n}^k$ is the gradient vector of $(\underline{n}^k)_{g\geq 2k-1}$.

In the Young diagram algorithm, λ_k is chosen minimal with the property $s_{\lambda_k}(M_k) \cap Q_k \subset M_k$. By definition of s_{λ_k}, this means that for every box (i,j) in the Young diagram of $\underline{n}^k$, either $j \leq \lambda_k$ or the box $(i+2, j-\lambda_k)$ also belongs to the same Young diagram. But since j may take any value between 1 and $\underline{n}_i^k$, either $j \leq \lambda_k$, or $\underline{n}_{i+2}^k \leq \underline{n}_i^k - \lambda_k$. We may rewrite the second inequality as

$$\lambda_k \leq \underline{n}_i^k - \underline{n}_{i+2}^k = \overline{n}_i^k.$$

Therefore, λ_k coincides in both algorithms.

By the same argument, i_k is also the same in both algorithms. (It is the largest index for which the inequality is an equality.)

Now we turn to the definition of $\underline{n}^{k+1}$. Its Young diagram M_{k+1} is a subset of $Q_{k+1} := \{(i,j) \in \mathbb{Z} \times \mathbb{Z} |\ j \geq 1,\ i \geq 2k+1\}$, so the $2k$ top lines are cut off, corresponding to the fact that the first $2k$ entries of $\underline{n}^{k+1}$ are irrelevant.

Apart from these top lines, $M_{k+1} = (M_k \setminus s_{\lambda_k}(C_k))$, where $C_k := \{(i,j) |\ i \leq i_k\} \setminus M_k$. By definition of s_{λ_k}, this will not affect any line with index $i > i_k+2$. By definition of i_k, the line with index $i = i_k + 2$ is also unchanged. What happens to a line $i \leq i_k + 1$?

By definition of λ_k, we have $s_{\lambda_k}(M_k) \cap Q_k \subset M_k$. For the complement C_k of M_k this means that $M_{k+1} = M_k \setminus s_{\lambda_k}(C_k) = Q_k \setminus s_{\lambda_k}(C_k)$, at least for the rows with index $i \leq i_k + 2$ (since the complement is taken with respect to the rows with index $i \leq i_k$, and s_{λ_k} shifts this index by 2.)

Hence, M_{k+1} can be expressed in terms of $s_{\lambda_k}(Q_k \setminus M_k)$ (for rows $i \leq i_k + 2$), and so row i of M_{k+1} does only depend on row $i-2$ of M_k:

$$\underline{n}_i^{k+1} = \begin{cases} \underline{n}_i^k & \text{if } i \geq i_k + 2 \\ \underline{n}_{i-2}^k - \lambda_k & \text{if } i \leq i_k + 2 \end{cases}$$

The two cases coincide for $i = i_k + 2$.

Now let us compute the derivation of $\underline{n}^{k+1}$:

- For $i \geq i_k + 2$, we have

$$\begin{aligned} \overline{n}_i^{k+1} &= \underline{n}_i^{k+1} - \underline{n}_{i+2}^{k+1} \\ &= \underline{n}_i^k - \underline{n}_{i+2}^k \\ &= \overline{n}_i^k \end{aligned}$$

- For $3 \leq i \leq i_k$, we have

$$\begin{aligned} \overline{n}_i^{k+1} &= \underline{n}_i^{k+1} - \underline{n}_{i+2}^{k+1} \\ &= \left(\underline{n}_{i-2}^{k} - \lambda_k\right) - \left(\underline{n}_i^k - \lambda_k\right) \\ &= \underline{n}_{i-2}^{k} - \underline{n}_i^k \\ &= \overline{n}_{i-2}^{k} \end{aligned}$$

- For $i = i_k + 1$, we have

$$\begin{aligned} \overline{n}_i^{k+1} &= \underline{n}_{i_k+1}^{k+1} - \underline{n}_{i_k+3}^{k+1} \\ &= \left(\underline{n}_{i_k-1}^{k} - \lambda_k\right) - \underline{n}_{i_k+3}^{k} \\ &= \underline{n}_{i_k-1}^{k} - \underline{n}_{i_k+3}^{k} - \lambda_k \\ &= \left(\underline{n}_{i_k-1}^{k} - \underline{n}_{i_k+1}^{k}\right) + \left(\underline{n}_{i_k+1}^{k} - \underline{n}_{i_k+3}^{k}\right) - \lambda_k \\ &= \overline{n}_{i_k-1}^{k} + \overline{n}_{i_k+1}^{k} - \overline{n}_{i_k}^{k}. \end{aligned}$$

All three results fit the definition in the numerical algorithm. (Note that in the algorithm, three entries are replaced by a single new entry, and all entries with index $i \leq i_k - 1$ are shifted by 2.)
This completes our induction. We have successfully proven that the invariant (3.3) holds throughout the algorithm.
To conclude our proof, we only need to observe that the halting condition and the output coincide. For the output this is clear. For the halting condition, the two algorithms terminate when $\underline{n}_k$ or $\overline{n}_k$ become zero, respectively. But we have seen that $\overline{n}_k$ is the derivation of $\underline{n}_k$, and it is easily seen that a partition is zero if and only if it has zero derivation.

□

We have seen that both algorithms coincide. However, we have not yet shown that they terminate and that the output is a partition.
However, both facts are now rather easy to see from the numerical algorithm: Termination is clear, since in each step, the vector loses two entries. (Even in the special case where we add a "0" to the right, it is easy to see that the number of 0's increases by 2.) Hence, the algorithm runs at most $\lfloor \frac{m+1}{2} \rfloor$ steps, where m is the length of the input partition.
In order for the output to be a partition, we must show that the λ_i are sorted and positive: $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_k > 0$. Recall the definition $\lambda_k := \max_l \{\overline{n}_l^k\}$. We must only show that the maximum entry is not increased when going from $\overline{n}^k$ to $\overline{n}^{k+1}$, and that all entries are non-negative. The only new entry

in the latter vector is $\overline{n}^{k+1}_{i_k-1} = \overline{n}^k_{i_k-1} + \overline{n}^k_{i_k+1} - \overline{n}_{i_k}$, which is non-negative by 3.1.5. Since $\overline{n}_{i_k} = \lambda_k$ by definition of i_k, we obtain

$$\overline{n}^{k+1}_{i_k-1} = \overline{n}^k_{i_k-1} + \overline{n}^k_{i_k+1} - \lambda_k \leq 2\lambda_k - \lambda_k = \lambda_k.$$

Since the new entry is not bigger than λ_k, the maximum is not increased, and $\lambda_{k+1} \leq \lambda_k$.
So far we have seen two algorithms, terminating with a partition as output, both of which coincide. In the next section we will show that the function that is defined by these algorithms is a CL-map.

3.3.3 Proof of the CL-property

We will need one more definition before we tackle the proof: We will redefine the canonical section $\mathcal{P}_{base} \subset \mathcal{P}$ of Λ, as it is defined in 3.2.8. We need to redefine it because we do not yet know that Λ is a CL-map, so we cannot yet apply definition 3.2.8. So we need to define $\mathcal{P}_{base}$ directly, and we will do so by directly constructing a section $\iota : \mathcal{G}_\mathcal{P} \to \mathcal{P}$, i.e., $\Lambda \circ \iota = \mathrm{id}_{\mathcal{G}_\mathcal{P}}$. Then $\mathcal{P}_{base}$ will be the image under this map.

3.3.10 Definition. *Let $\underline{n} = (\underline{n}_1, \underline{n}_2, \ldots, \underline{n}_k) \in \mathcal{G}_\mathcal{P}$ be a group, $\underline{n}_1 \geq \underline{n}_2 \geq \ldots \geq \underline{n}_k > 0$. Then we define $\iota(\underline{n}) \in \mathcal{P}$ to be the partition*

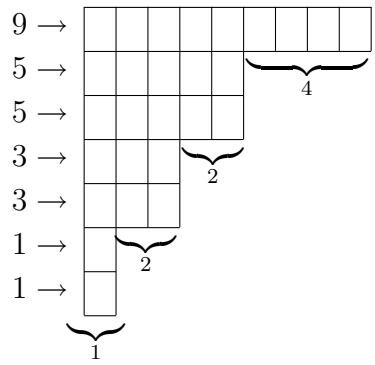
$$\iota(\underline{n}) := \left(\sum_{i=1}^{k} \underline{n}_i, \sum_{i=2}^{k} \underline{n}_i, \sum_{i=2}^{k} \underline{n}_i, \sum_{i=3}^{k} \underline{n}_i, \sum_{i=3}^{k} \underline{n}_i, \ldots \right.$$
$$\left. \ldots, \sum_{i=k-1}^{k} \underline{n}_i, \sum_{i=k-1}^{k} \underline{n}_i, \underline{n}_k, \underline{n}_k \right),$$

where each term appears twice, except for the first one, which appears only once.
Put $\mathcal{P}_{base} := \iota(\mathcal{G}_\mathcal{P})$. Similarly, let $\mathcal{D}_{base}$ be the set of all derivations whose partition is in $\mathcal{P}_{base}$.

Once we know that Λ is indeed an order-preserving CL-map, the definition above coincides with the general definition 3.2.8.

3.3.11 Example. The group $\mathbb{Z}/p^4\mathbb{Z} \times \mathbb{Z}/p^2\mathbb{Z} \times \mathbb{Z}/p^2\mathbb{Z} \times \mathbb{Z}/p\mathbb{Z}$ with partition $\underline{n} = (4, 2, 2, 1)$ corresponds to $\iota(\underline{n}) = (9, 5, 5, 3, 3, 1, 1)$.

The correspondence can be visualized in the Young diagram:



3.3.12 Lemma.

(i) A partition $\underline{m} = (\underline{m}_1, \underline{m}_2, \ldots, \underline{m}_k)$ *belongs to* $\mathcal{P}_{base}$ *if and only if it satisfies the following conditions:*

- *k is odd.*
- $\underline{m}_1 > \underline{m}_2 = \underline{m}_3 > \underline{m}_4 = \underline{m}_5 > \ldots > \underline{m}_{k-1} = \underline{m}_k$.
- $\underline{m}_1 - \underline{m}_3 \geq \underline{m}_3 - \underline{m}_5 \geq \underline{m}_5 - \underline{m}_7 \geq \ldots \geq \underline{m}_{k-2} - \underline{m}_k \geq \underline{m}_k$.

(ii) A derivation $\overline{m} = (\overline{m}_0, \overline{m}_1, \ldots, \overline{m}_k)$ *belongs to* $\mathcal{D}_{base}$ *if and only if it satisfies the following conditions:*

- *k is odd.*
- $\overline{m}_0 = \overline{m}_1 \geq \overline{m}_2 = \overline{m}_3 \geq \overline{m}_4 = \ldots = \overline{m}_{k-2} \geq \overline{m}_{k-1} = \overline{m}_k$.

Proof.

(i) I leave it to the reader to show that every element of $\mathcal{P}_{base}$ has the listed properties.

On the other hand, if $\underline{m} \in \mathcal{P}$ has these properties then it is the image of $(\underline{m}_1 - \underline{m}_3, \underline{m}_3 - \underline{m}_5, \ldots, \underline{m}_{k-2} - \underline{m}_k, \underline{m}_k) \in \mathcal{G}_{\mathcal{P}}$. It is immediate to check that the latter one is indeed in $\mathcal{G}_{\mathcal{P}}$.

(ii) This is completely analogous to the partition case. If $\overline{m} \in \mathcal{D}$ has the listed properties then it is the derivation of $\iota((\overline{m}_1, \overline{m}_3, \overline{m}_5, \ldots, \overline{m}_k))$.

Alternatively, it is possible to show that the characterization of $\underline{m}$ translates into the characterization of $\overline{m}$ and vice versa.

□

3.3.13 Corollary. *The map ι is a section of Λ, i.e., $\Lambda \circ \iota = \mathrm{id}_{\mathcal{G}_\mathcal{P}}$.*

Proof. This is immediate if we use lemma 3.3.12 and apply the numerical algorithm to $\overline{m}$. □

Now we can turn to the main theorem:

3.3.14 Theorem. *The map Λ is a CL-map.*

How can we prove such a statement? It is possible to prove it directly, preferably with the Young diagram algorithm. A sketch of the proof can be found in [Len08]. However, the proof is rather intransparent. Instead, we will follow an indirect approach. We have seen in 3.2.7 (and the preceding discussion) that Λ is a CL-map if and only if there exists a map μ with certain properties. We will directly define such a map. This approach has two advantages: Firstly, its structure is much clearer; secondly, we will get as spin-off that Λ is order-preserving.
Recall that we need to define maps $\mu_G : I_G \to \mathcal{P}$ such that for any $t \in I_G$ we have $\mathrm{expon}(t) = \mathrm{size}(\mu_G(t))$ and with some additional requirements (for details see 3.2.7).

3.3.15 Algorithm. Let

$$G = \prod_{i=1}^{k} (\mathbb{Z}/p^{e_i}\mathbb{Z})^{r_i}\text{, where } e_1 > e_2 > \ldots > e_k > 0,$$

be a p-group in standard form.
Let $t = (t_{i,s})_{1 \le i \le k, 1 \le s \le r_i} \in I_G$ (i.e., all $t_{i,s} \ge 0$).
We define $\mu_G(t)$ as the output of the following algorithm:

1. Start with the derivation $\overline{n}$ of $\iota(G)$.

2. FOR $i := k$ DOWNTO 1 DO
   FOR $s := r_i$ DOWNTO 1 DO
   REPEAT $t_{i,s}$ times:
   Let j be the maximal index such that $\overline{n}_{j-1} = e_i$ and $\overline{n}_j < e_i$.
   Decrease $\overline{n}_{j-2s}$ by 1.
   Increase $\overline{n}_j$ by 1.
   END REPEAT
   END FOR
   END FOR

3. Let $\underline{n}$ be the partition with derivation $\overline{n}$. Output $\underline{n}$.

If necessary, we may add an entry "0" at the right of $\overline{n}$ before choosing j. Note that the order of the loops is crucial: It is not possible to change the order of the inner nor of the outer loop.

3.3.16 Lemma. *The above algorithm has the following invariants: In any step, before decreasing and increasing entries,*

(i) $\overline{n}$ *is a derivation.*

(ii) $1 \leq \overline{n}_{j-2s} \leq e_i$.

(iii) $\overline{n}_{j-2s+1} = \overline{n}_{j-2s+2} = \ldots = \overline{n}_{j-1} = e_i$.

(iv) $\overline{n}_l < e_i$ *for all* $l \geq j$.

(v) $\overline{n}_{j-2s} + \overline{n}_j \geq e_i$.

(vi) For all $\tilde{i} < i$ *there is a* $\tilde{j} < j - 2s$ *such that* $\overline{n}_{\tilde{j}} = e_{\tilde{i}}$.

In particular, in each step (= the three lines within the REPEAT loop) there does exist a j *as required.*

Proof. We use induction on the number of steps. In the first step, i.e., before the first "decrease" operation, $\overline{n}$ is a derivation in $\mathcal{D}_{base}$, and the statements can be immediately verfied using lemma 3.3.12.
Now assume the invariants hold during all steps up to a certain step A of the algorithm. We want to show that they persist in the next step B. Let $\overline{n}$ be the state before step A, and let $\overline{n}'$ be the state before step B. Let e_i', s', j' be the data of step B. Then $e_i \leq e_i'$.
Let us first turn to (i). Lemma 3.1.4 gives us a characterization of derivations. Except for the third point, all the requirements there are trivial to see from our induction hypothesis. So we only need to show $\sum_{l=i_1}^{i_2}(-1)^{l-i_1}\overline{n}_l' \geq 0$ for all $1 \leq i_1 < i_2 \leq k$, $i_1 \equiv i_2 \bmod 2$ (if $i_2 = k$, without the condition $i_1 \equiv i_2 \bmod 2$). By remark 3.1.5 we may restrict ourselves to the case $i_2 = k$, so we need to show that

$$\sum_{l=i_1}^{k}(-1)^{l-i_1}\overline{n}_l' \geq 0 \qquad \text{for all } 1 \leq i_1 < k.$$

Of course, by induction hypothesis we know that the same formula is true for $\overline{n}$. Since $j \equiv j - 2s \bmod 2$, the only interesting case is when $j - 2s < i_1 \leq j$. If $i_1 \equiv j \bmod 2$, then

$$\sum_{l=i_1}^{k}(-1)^{l-i_1}\overline{n}_l' = \sum_{l=i_1}^{k}(-1)^{l-i_1}\overline{n}_l + 1 \geq 1 > 0.$$

Otherwise,

$$\begin{aligned}\sum_{l=i_1}^{k}(-1)^{l-i_1}\overline{n}'_l &\overset{(iii)}{\geq} \sum_{l=j-1}^{k}(-1)^{l-i_1}\overline{n}'_l \\ &= \underbrace{\overline{n}'_{j-1}-\overline{n}'_j}_{>e-e=0}+\sum_{l=j+1}^{k}(-1)^{l-i_1}\underbrace{\overline{n}'_l}_{=\overline{n}_l} \\ &\geq \sum_{l=j+1}^{k}(-1)^{l-i_1}\overline{n}_l \\ &\geq 0 \text{ by induction hypothesis.}\end{aligned}$$

Now we come to (ii)–(vi). First assume $e_i = e'_i$. Then $s \geq s'$. Since $\overline{n}_j < e_i$ we know $\overline{n}'_j \leq e_i$. If $\overline{n}'_j < e_i$ then $j = j'$ and the invariants (iii)–(vi) are obvious and (ii) follows from (v) and (vi) of the hypothesis. So assume $\overline{n}'_j = e_i$. Then $j' = j+1$ and (ii), (iii), (iv) and (vi) are obvious. (v) follows from the fact $\overline{n}_{j'-2s'} = e_i$.
Now assume $e_i < e'_i$. Then the invariants ensure that no entry $\geq e'_i$ has been modified in any preceeding step. Now recall that the start sequence was in $\mathcal{D}_{base}$, so when the algorithm started $\overline{n}$ was monotonously decreasing. Let l_0 be the index such that at the starting point of the algorithm $\overline{n}_l \geq e'_i$ for $l \leq l_0$ and $\overline{n}_l < e'_i$ for $l > l_0$. Then before step B, $\overline{n}_1, \ldots, \overline{n}_{l_0}$ are unchanged. Therefore, $\overline{n}_{l_0-2s+1} = \overline{n}_{l_0-2s+2} = \ldots = \overline{n}_{l_0} = e'_i$. On the other hand, for $l > l_0$, we still have $\overline{n}_l < e'_i$, so in step B we have $j = l+1$, and all invariants remain true. □

3.3.17 Lemma. *Let $G \in \mathcal{G}_{\mathcal{P}}$ be as in the algorithm above and let $t \in I_G$. Then* $\operatorname{expon}(t) = \operatorname{size}(\mu_G(t))$.

Proof. Recall that

$$\operatorname{expon}(t) = \sum_{1\leq i,j\leq k}\min(e_i,e_j)r_ir_j + \left(\sum_{i=1}^{k}\sum_{s=1}^{r_i}st_{i,s}\right).$$

Let $\overline{n}$ be the generic variable of algorithm 3.3.15. We know that $\overline{n} \in \mathcal{D}$, so let $\underline{n}$ be its partition. First we show that when the algorithm starts we have

$$\operatorname{size}(\underline{n}) = \sum_{1\leq i,j\leq k}\min(e_i,e_j)r_ir_j. \tag{3.4}$$

Recall now $\underline{n} = \iota(G)$. Let $G = (\underline{m}_1, \underline{m}_2, \ldots, \underline{m}_{k'}) \in \mathcal{G}_{\mathcal{P}}$. Recalling the definition of ι (definition 3.3.10) we see that

$$\operatorname{size}(\underline{n}) = \sum_{l=1}^{k'}\underline{m}_l + 2\sum_{j=2}^{k'}\sum_{l=j}^{k'}\underline{m}_l.$$

We know that exactly r_i of the $\underline{m}_i$ are equal to e_i, namely the entries $\underline{m}_{r_1+\ldots+r_{i-1}+1}, \ldots, \underline{m}_{r_1+\ldots+r_{i-1}+r_i}$.
How often does the term e_i occur in the sum? Counting yields

$$\begin{aligned} & r_i + 2\left(\sum_{j=2}^{r_1+r_2+\ldots+r_{i-1}} r_i + \sum_{j=r_1+r_2+\ldots+r_{i-1}+1}^{r_1+r_2+\ldots+r_{i-1}+r_i} (r_1+\ldots+r_{i-1}+r_i-j+1)\right) \\ = \;& r_i + 2\left(r_i(r_1+r_2+\ldots+r_{i-1}) - r_i + \frac{1}{2}r_i(r_i+1)\right) \\ = \;& r_i^2 + 2\sum_{j=1}^{i-1} r_i r_j, \end{aligned}$$

so the contribution of the e_i-terms is $e_i r_i^2 + 2\sum_{j=1}^{i-1} e_i r_i r_j$. Summing up over all i yields equation (3.4).
Now we show that one step in the algorithm (decreasing $\overline{n}_{j-2s}$ by 1 and increasing $\overline{n}_j$ by 1) increases size($\underline{n}$) by exactly s. Then the lemma follows from trivial induction.
So what does it mean for $\underline{n}$ if $\overline{n}_{j-2s}$ is decreased by 1 and $\overline{n}_j$ is increased by 1? It means that $\underline{n}_{j-2s}, \underline{n}_{j-2s+2}, \underline{n}_{j-2s+4}, \ldots, \underline{n}_{j-2}$ are increased by 1. The other entries of $\underline{n}$ remain unchanged. Hence, size($\underline{n}$) increases by s and we are done.

□

Next we want to show that $\Lambda \circ \mu_G$ is simply the projection onto the one-element set $\{G\}$. Unfortunately, the algorithms for Λ and μ_G are not step-by-step inverse. For this reason, I will define a variant of the Λ-algorithm which is equivalent to the algorithm 3.3.6 and inverts μ_G step by step.

3.3.18 Algorithm (numerical)**.** Let $\underline{n} = (\underline{n}_1, \underline{n}_2, \ldots, \underline{n}_m) \in \mathcal{P}$.

1. Let $\overline{n}^1 := \overline{n}$ be the derivation of $\underline{n}$. Let $k := 1$ and let $s_1 := 1$.

2. Let $j_k := \min\{j \mid \overline{n}^k_j < \overline{n}^k_{j+1}\}$.
   Let $\lambda_k := \max\{\overline{n}^k_l \mid j_k \leq l \leq m\}$, and let $i_k := \max\{l \mid \overline{n}^k_l = \lambda_k\}$.
   If $k > 1$ and $\lambda_k = \lambda_{k-1}$ then put $s_k := s_{k-1} + 1$; otherwise put $s_k := 1$.

3. If $\overline{n}^k_{i_k-2s_k+1} < \lambda_k$ then replace $\overline{n}^k_{i_k-2s_k+1}$ by λ_k and replace $\overline{n}^k_{i_k+1}$ by $\overline{n}^k_{i_k+1} + \overline{n}^k_{i_k-2s_k+1} - \lambda_k$. Otherwise, do nothing.

4. Shift the subsequence $\overline{n}^k_{i_k-2s_k+1}, \overline{n}^k_{i_k-2s_k+2}, \ldots, \overline{n}^k_{i_k}$ (all of which are equal to λ_k) to the left until its left neighbor entry is $\geq \lambda_k$, thereby getting

$\overline{n}^{k+1}$. (Here, "shifting" means only a reordering of the entries – no entry is destroyed. See also the example below.)

Increase k by 1.

5. Repeat step 2,3 and 4 until $\overline{n}^k$ is monotonously decreasing.

The output of the algorithm is $(\overline{n}_0, \overline{n}_2, \overline{n}_4, \ldots) \in \mathcal{G}_{\mathcal{P}}$, where possible 0's on the right are left out.

Before listing the crucial properties of this algorithm, I give an example:

3.3.19 Example. Let $\underline{n} = (11, 11, 9, 9, 9, 8, 8, 6, 4, 4, 4, 3, 2, 1, 1)$. This is the same partition as in examples 3.3.5 and 3.3.8.
I mark the area from $i_k - 2s_k + 1$ to i_k by bold type.
k=1: We compute

$$\overline{n}^1 = \overline{n} = (0, 2, 2, 0, 1, 1, \mathbf{2}, \mathbf{4}, 2, 0, 1, 2, 2, 1, 1, 1).$$

We have $s_1 = 1$, $j_1 = 0$, $\lambda_1 = 4$, $i_1 = 7$. (Recall that the first entry of $\overline{n}$ has index 0.)
Since $\overline{n}^1_6 < \lambda_1$, we replace $\overline{n}^1_6$ by 4 and $\overline{n}^1_8$ by $2 + 2 - 4$=0. We get

$$(0, 2, 2, 0, 1, 1, \mathbf{4}, \mathbf{4}, 0, 0, 1, 2, 2, 1, 1, 1).$$

Now we shift the subsequence $4, 4$ to the left. There is no other entry ≥ 4, so we shift it all the way to the left:

$$\overline{n}^2 = (4, 4, 0, 2, 2, 0, 1, 1, 0, 0, 1, \mathbf{2}, \mathbf{2}, 1, 1, 1).$$

k=2: We get $j_2 = 2$, $\lambda_2 = 2 \neq \lambda_1$, so $s_2 = 1$, $i_2 = 12$.
Since $\overline{n}^2_{11} = \lambda_2$, step 3 is void. We shift the sequence $2, 2$ to the left until we obtain $\overline{n}^2_4$ as left neighbor:

$$\overline{n}^3 = (4, 4, 0, \mathbf{2}, \mathbf{2}, \mathbf{2}, \mathbf{2}, 0, 1, 1, 0, 0, 1, 1, 1, 1).$$

k=3: We get $j_3 = 2$, $\lambda_3 = 2 = \lambda_2$, so $s_3 = 2$, $i_3 = 6$.
Since $\overline{n}^3_3 = \lambda_3$, step 3 is void. We shift the sequence $2, 2, 2, 2$ to the left until we obtain $\overline{n}^3_1$ as left neighbor:

$$\overline{n}^4 = (4, 4, 2, 2, 2, 2, 0, 0, 1, 1, 0, 0, 1, 1, \mathbf{1}, \mathbf{1}).$$

k=4: We get $j_4 = 7$, $\lambda_4 = 1 \neq \lambda_3$, so $s_4 = 1$, $i_4 = 15$.
Since $\overline{n}^4_{14} = \lambda_4$, step 3 is void. Since the left neighbor of the marked $1, 1$-sequence is already ≥ 1, we do not need to shift:

$$\overline{n}^5 = (4, 4, 2, 2, 2, 2, 0, 0, 1, 1, 0, 0, \mathbf{1}, \mathbf{1}, \mathbf{1}, \mathbf{1}).$$

k=5: We get $j_5 = 7$, $\lambda_5 = 1 = \lambda_4$, so $s_5 = 2$, $i_5 = 15$.
Since $\overline{n}^5_{12} = \lambda_5$, step 3 is void. We shift the sequence $1, 1, 1, 1$ to the left until we obtain $\overline{n}^5_9$ as left neighbor:

$$\overline{n}^6 = (4, 4, 2, 2, 2, 2, 0, 0, \mathbf{1}, \mathbf{1}, \mathbf{1}, \mathbf{1}, \mathbf{1}, \mathbf{1}, 0, 0).$$

k=6: We get $j_6 = 7$, $\lambda_6 = 1 = \lambda_5$, so $s_6 = 3$, $i_6 = 13$.
Since $\overline{n}^6_8 = \lambda_6$, step 3 is void. We shift the sequence $1, 1, 1, 1, 1, 1$ to the left until we obtain $\overline{n}^6_5$ as left neighbor:

$$\overline{n}^7 = (4, 4, 2, 2, 2, 2, 1, 1, 1, 1, 1, 1, 0, 0, 0, 0).$$

$\overline{n}^7$ is monotonously decreasing, so we are done.
The result is $(\overline{n}_0, \overline{n}_2, \overline{n}_4, \overline{n}_6, \overline{n}_8, \overline{n}_{10}) = (4, 2, 2, 1, 1, 1) \in \mathcal{G}_{\mathcal{P}}$.

If we compare this example with example 3.3.8 we can already see that this algorithm does essentially the same computation as the standard algorithm 3.3.6. Before we compare these two, I list some important invariances.

3.3.20 Lemma. *In the algorithm above, we have for any k:*

(i) $\overline{n}^k$ *is a derivation.*

(ii) $\overline{n}^k_{i_k-2s+1} = \ldots = \overline{n}^k_{i_k} = \lambda_k$.

(iii) If $\lambda_k \neq \lambda_{k-1}$ *then* $\overline{n}^k$ *starts with the sequence*

$$\lambda_1, \lambda_1, \lambda_2, \lambda_2, \lambda_3, \lambda_3, \ldots, \lambda_{k-1}, \lambda_{k-1}.$$

In this case, this part of the sequence remains unchanged during the rest of the algorithm.

(iv) When the algorithm terminates, $\overline{n}^k \in \mathcal{D}_{base}$.

Proof. The lemma is obtained by straightforward induction. The only non-trivial point is (i). Inspecting the characterization of derivations in lemma 3.1.4, we see that shifting an even number of entries maintains the derivation property, so step 4 of the algorithm is ok. Step 3 requires proof, but the proof is almost identical to the proof of 3.3.16, so I omit it here. □

3.3.21 Proposition. *The two algorithms 3.3.6 and 3.3.18 coincide.*

Proof. This is almost trivial by the invariants of the preceding lemma. Precisely speaking, we have the following: Let $\overline{n}^k$ be the intermediate result of our modified algorithm 3.3.18 and let $\overline{n}'^k$ be the intermediate result after the same number of step of the original algorithm 3.3.6. Then $\overline{n}'^k$ is obtained from $\overline{n}^k$ by removing all entries $\overline{n}^k_i$ with

- $0 \leq i \leq 2(k - s_k)$ (those are precisely the entries with $\overline{n}_i^k > \lambda_k$), *or*
- $i_k - 2s_k + 2 \leq i < i_k$.

This claim can be checked immediately by comparing the steps of the two algorithms. It reflects the fact that the original algorithm removes twice the entry λ_k whereas the modified version shifts them to the left.
In particular, whenever $\lambda_k \neq \lambda_{k-1}$ the second condition is empty and $\overline{n}'^k$ is obtained from $\overline{n}^k$ by removing the first $2k - 2$ entries.
Now we investigate the algorithms and see that they choose λ_{k+1} as the maximum from the same set of numbers. Therefore, the algorithms give the same result. □

3.3.22 Remark. The preceeding proposition and example indicate that the modified version of the numerical algorithm differs only in details from the original one. This is true — as long as you let them work on *derivations*. However, you should keep in mind that small changes in derivations like shifting a few entries does have a non-trivial effect on the corresponding partitions. In particular, *a shift in derivations will change the size of the corresponding partitions!* Therefore, the two algorithms do have considerable differences when translated into algorithms that work on partitions.
The difference is exemplified in our proof that Λ is a CL-map. As we will shortly see, the modified algorithm inverts the μ-algorithm step by step. The original version does not so, and is not even close because it does not even have the correct partition sizes in the intermediate steps. Of course, as we have seen before, the two algorithms essentially coincide whenever $\lambda_{k-1} \neq \lambda_k$. With respect to μ that means that the original Λ-algorithm inverts every FOR-loop of the μ-algorithm as a whole, but not the single steps within a FOR-loop (i.e., not the REPEAT-loops).
Finally, note that in the modified algorithm it is crucial that we choose i_k to be maximal – not for coinciding with the other algorithm (for that we could have chosen any i with $\overline{n}_i^k = \lambda_k$) but for being inverse to the μ-algorithm. Also, it is not clear how one could modify the μ-algorithm in such a way that it would be inverse to a Λ-algorithm with non-maximal i_k. That is one reason why I chose i_k to be maximal already in the original Λ-algorithm (cf. 3.3.7).

3.3.23 Proposition. *Let $G \in \mathcal{G}_\mathcal{P}$ and let $t \in I_G$.*
Then $(\Lambda \circ \mu_G)(t) = G$.

Proof. We use the modified numeric algorithm for computing Λ (3.3.18).
When the μ_G-algorithm 3.3.15 runs on the input t, it modifies a derivation $\overline{n}$ in each step, starting with $\overline{n}_{start} := \iota(G)$ and terminating with the state

$\overline{n}_{end} := \mu_G(t)$. Furthermore, it uses a counter s. After termination we let the Λ-algorithm run on the input $\overline{n}_{end}$, producing states $\overline{n}^k$ and parameters s_k in each step.
We use an inductive argument. Let $\overline{n}$ be the state of the μ_G-algorithm at some point before going into a (non-trivial) REPEAT-loop, and let $\overline{n}'$ be the state after this loop. Let s and s' be the value of the counter corresponding to the two states, respectively. Comparing the invariants of the two algorithms (lemmas 3.3.16 and 3.3.20) we find:

If there is a k with $\overline{n}' = \overline{n}^k$ and $s' = s_k$ then $\overline{n} = \overline{n}^{k+1}$ and $s = s_{k+1}$.

In other words, one step of the Λ-algorithm inverts the REPEAT-loop of the μ_G-algorithm. This gives the induction step. For the base case of the induction, just note that the μ_G-algorithm terminates with $\overline{n}_{end}$ and $s = 1$ and the Λ-algorithm starts with $\overline{n}^1 = \overline{n}_{end}$ and $s_1 = 1 = s$.
Therefore, the Λ-algorithm terminates with $\overline{n}^k = \overline{n}_{start} := \iota(G)$. Its output is then G. This proves the proposition. □

3.3.24 Corollary. *For various G, the images of μ_G are mutually disjoint and their union is all of $\mathcal{P}$.*

Proof. By proposition 3.3.23, we have $\Lambda \circ \mu_G \equiv G \neq G' \equiv \Lambda \circ \mu_{G'}$ for $G \neq G'$, so the images of the μ_G are mutually disjoint.
For the second statement, fix an $n \in \mathbb{N}$ and let $\mathcal{P}_n := \{\underline{n} \mid \mathrm{size}(\underline{n}) = n\}$. Consider

$$M := \bigcup_G (\mu_G(I_G) \cap \mathcal{P}_n).$$

Clearly, $M \subseteq \mathcal{P}_n$. Consider the coefficient a_n of q^n in the power series $\sum_G w(G)$. We know already

$$w(G) \overset{3.2.3}{=} \sum_{t \in I_G} q^{\mathrm{expon}(t)} \overset{3.3.17}{=} \sum_{t \in I_G} q^{\mathrm{size}(\mu_G(t))}.$$

Since the images of all μ_G are mutually disjoint, we conclude $a_n = \#M$, where a_n is the coefficient of q^n. On the other hand, by equation (3.1) on page 38 we know $a_n = \#\mathcal{P}_n$. Therefore M cannot be a proper subset of $\mathcal{P}_n$ and we have equality. □

As we have seen earlier, the existence of such maps μ_G ensures that Λ is a CL-map. Hence *we have concluded our proof of the main theorem 3.3.14.*

We conclude this section by showing that Λ is order-preserving (in the sense of 3.2.7). Since we have already defined the maps μ_G, we only need to show that μ_G is order-preserving for each G.

3.3.25 Proposition. *Let $G \in \mathcal{G}_{\mathcal{P}}$. Then μ_G is order-preserving.*

Proof. We need to show that for $t, t' \in I_G, t < t'$ we have $\mu_G(t) < \mu_G(t')$. The first inequality means that t is dominated by t', the second one means that the Young diagram of $\mu_G(t)$ is contained in the Young diagram of $\mu_G(t')$.
Let $G = \prod_{i=1}^{k} (\mathbb{Z}/p^{e_i}\mathbb{Z})^{r_i}$ in standard form (with e_i strictly decreasing). Then $I_G = \{(t_{i,s})_{1\le i\le k, 1\le s\le r_i} \mid t_{i,s} \ge 0\}$. We may assume that there is no t'' such that $t < t'' < t'$. This is the case only if

- t and t' differ only for one i_0, *and*
- *Either A:* $t_{i_0,1} = t'_{i_0,1} - 1$ and $t_{i_0,s} = t_{i_0,s}$ for all $s > 1$.

  Or B: There is an $s_0 > 1$ such that $t_{i_0,s_0} = t'_{i_0,s_0} + 1$, $t_{i_0,s_0+1} = t'_{i_0,s_0+1} - 1$ and $t_{i_0,s} = t_{i_0,s}$ for all $s \neq s_0, s_0 + 1$.

I will argue for case B, which is a bit more complicated. Case A is completely analogous. Let $\overline{n}$, $\overline{n}'$ be the generic variable in the μ-algorithm when applied to t, t', respectively. Obviously, $\overline{n} = \overline{n}'$ until we come to the REPEAT loop indexed by $i = i_0$ and $s = s_0 + 1$. If we compare the t' case with the t case, in the former one the algorithm performs one decrease/increase action with $s_0 + 1$ instead of one with s_0. Hence after this operation $\overline{n}'$ differs from $\overline{n}$ only in two entries $\overline{n}_j$, $\overline{n}_{j+2}$ of distance 2: $\overline{n}_j = \overline{n}'_j - 1$ and $\overline{n}_{j+2} = \overline{n}'_{j+2} + 1$. (Exacly the same property is generated in case A). Inspecting the algorithm, we see that this property is maintained throughout the rest of the algorithm. Hence, after termination there is a j_0 such that $\overline{n}_{j_0} = \overline{n}'_{j_0} - 1$, $\overline{n}_{j_0+2} = \overline{n}'_{j_0+2} + 1$ and $\overline{n}_j = \overline{n}'_j$ for $j \neq j_0, j_0 + 2$. Therefore, returning to partitions we get $\underline{n}_{j_0+2} = \underline{n}'_{j_0+2} + 1$ and $\underline{n}_j = \underline{n}'_j$ for $j \neq j_0 + 2$. Evidently, $\underline{n}$ is dominated by $\underline{n}'$ and we have proven our assertion. □

3.3.26 Corollary. Λ *is an order-preserving CL-map.*

Proof. We know that Λ is a CL-map by 3.3.14. For being order-preserving, by definition 3.2.7 it is sufficient to check that each μ_G is order-preserving, which is true by corollary 3.3.25. □

3.4 Some consequences

Theorem 3.3.14 enables us to compute the probability of a group to have a certain exponent. Recall that I use the p-adic exponent, i.e., if a p-group has exponent e I mean that it is annihilated by p^e.
As always, $q = \frac{1}{p}$, but q may also be viewed as a formal variable.

3.4.1 Theorem. *Let $e \geq 0$ be fixed. Then we have*

$$\sum_{\substack{G\text{is a } p\text{-group} \\ \text{of exponent } \leq e}} w(G) = \prod_{\substack{j \not\equiv 0, \pm(e+1) \\ \bmod 2e+3}} (1-q^j)^{-1}.$$

(In the product, j runs through all positive integers staisfying the congruence conditions, not only through all residue classes mod $2e+3$*.)*

Proof. Recall that, since Λ is a CL-map,

$$w(G) = \sum_{n \geq 0} a_G(n) q^n,$$

where

$$a_G(n) = \left| \Lambda^{-1}(G) \cap \{ \underline{n} \in \mathcal{P} \mid \underline{n} \text{ is a partition of } n \in \mathbb{N} \} \right|.$$

Hence,

$$\sum_{\substack{G \text{ is a } p\text{-group} \\ \text{of exponent } \leq e}} w(G) = \sum_{n \geq 0} \left| \left\{ \underline{n} \in \mathcal{P} \;\middle|\; \begin{array}{l} \underline{n} \text{ is a partition of } n \text{ and} \\ \Lambda(\underline{n}) \text{ has exponent } \leq e \end{array} \right\} \right| q^n.$$

But if G is interpreted as a partition in $\mathcal{G}_{\mathcal{P}}$, then the exponent is simply the largest part. Given a partition $\underline{n} = (\underline{n}_1, \ldots, \underline{n}_m) \in \mathcal{P}$, the largest part of $\Lambda(\underline{n})$ will be λ_1, since the λ_i are sorted. On the other hand, it is easy to see that $\lambda_1 = \max_i (n_{i+2} - n_i)$, where we put $n_0 := n_{-1} := 0$. So we know that

$$\sum_{\substack{G \text{ is a } p\text{-group} \\ \text{of exponent } \leq e}} w(G) = \sum_{n \geq 0} \left| \left\{ \underline{n} \in \mathcal{P} \;\middle|\; \begin{array}{l} \underline{n} \text{ is a partition of n and} \\ \underline{n}_{i+2} - \underline{n}_i \leq e \text{ for all } i \end{array} \right\} \right| q^n.$$

But the right hand side is a well-known generating function, and its value is

$$\prod_{\substack{j \not\equiv 0, \pm(e+1) \\ \bmod 2e+3}} (1-q^j)^{-1}$$

(cf. [And76], Thm 7.5, $k := i := e+1$), which proves the theorem. □

3.4.2 Corollary. *The probability (in the Cohen-Lenstra heuristic) that a p-group has exponent $\leq e$ is*

$$\prod_{\substack{j \equiv 0, \pm(e+1) \\ \bmod 2e+3}} (1-p^{-j}).$$

Here, the product runs over all positive integers j that satisfy one of the congruences.

Proof. The heuristic tells us that the volume of the one-element set $\{G\}$ is $\frac{w(G)}{w(\mathcal{G}_p)}$ (here $w(G)$ is interpreted as an evaluated, not a formal series), so the probability of a p-group having exponent $\leq e$ is

$$\begin{aligned}\frac{1}{w(\mathcal{G}_p)}\left(\sum_{\substack{G \text{ is a } p\text{-group}\\ \text{of exponent } \leq e}} w(G)\right) &= \left(\prod_{j\geq 1}(1-p^{-j})\right)\left(\prod_{\substack{j\not\equiv 0,\pm(e+1)\\ \bmod 2e+3}}(1-p^{-j})^{-1}\right)\\ &= \prod_{\substack{j\equiv 0,\pm(e+1)\\ \bmod 2e+3}}(1-p^{-j}).\end{aligned}$$

□

To give a feeling for those probabilities, here is a table that lists the probability for a finite abelian p-group to have p-exponent e.

	$e=0$	$e=1$	$e=2$	$e=3$	$e\geq 4$
$p=2$	28.879%	33.965%	18.521%	9.361%	9.374%
$p=3$	56.013%	29.178%	9.871%	3.292%	1.646%
$p=5$	76.033%	19.167%	3.840%	0.768%	0.192%
$p=7$	83.680%	13.988%	1.999%	0.286%	0.048%
$p=11$	90.083%	9.015%	0.820%	0.075%	0.007%

3.4.3 Remark. This corollary is a generalization of [CL84, Example 5.3], where the case $e = 1$ is treated. Also, similar formulas for the rank of a p-group have long been known ([CL84, Thm. 6.1]). However, rank and exponent behave rather antipodal: It is pretty straightforward to derive results about the rank from the original Cohen-Lenstra approach, but the exponent gives very tough problems (except for $e = 1$).
On the other hand, with the given partition-theoretic interpretation (theorem 3.3.14), the exponent formula above is an almost trivial consequence, whereas it is not clear at all what it means for a partition to be mapped under Λ to a group of some given rank.
The same formula, although in a different context, was independently discovered by Fulman [Ful97]. See section 4.6 for a discussion.

3.5 Uniqueness of Λ

We have seen that our definition of Λ indeed establishes a powerful connection between the Cohen-Lenstra distribution and partitions and improves our understanding of the Cohen-Lenstra distribution. However, so far the

definition of Λ seems to be somewhat arbitrary. Recall that the key equation

$$w(G) = \sum_{n \geq 0} a_G(n) q^n$$

with

$$a_G(n) = \left| \Lambda^{-1}(G) \cap \{\underline{n} \in \mathcal{P} |\ \underline{n} \text{ is a partition of } n \in \mathbb{N}\} \right|,$$

is true for *any* order-preserving CL-map (cf. definition 3.2.7).

On the other hand, some corollaries (theorem 3.4.1 f.) rely heavily on the specific definition of Λ. So a question naturally arises: Is there anything special about our definition of Λ, or are there other order-preserving CL-maps, which might reveal other facts about the Cohen-Lenstra heuristic?

This section will answer the above question partially. There are two ways in which we can alter Λ: Firstly, we may concatenate Λ with any order-preserving automorphism of the set of all partitions. Fortunately, there is only one non-trivial such automorphism (see section 3.5.1 below).

Secondly, we may look for a completely different CL-map. This map then induces a canonical section $\iota : \mathcal{G}_{\mathcal{P}} \overset{1:1}{\to} \mathcal{P}_{base}$, and this canonical section may or may not coincide with the canonical section of Λ. Indeed, I will give an example that shows that the canonical section need not coincide, not even up to automorphism. However, we will show that if we require the canonical section to coincide with the canonical section of Λ (as described in definition 3.3.10) then the CL-map already equals Λ. In other words, Λ is unique over its canonical section. Note that this does not necessarily imply that for any other choice of the canonical section there is at most one order-preserving map over this section. Further work would be necessary to rule out this possibility. However, since this part of the thesis serves only as a motivation, I have excluded these further considerations.

For the same reason, I have not addressed the question which sections appear as canonical sections of CL-maps. The forthcoming proof indicates that such a section must satisfy very strict requirements, so I would be rather surprised if their was an "essentially different" canonical section (and hence, an "essentially different" order-preserving CL-map). But this is only a gut feeling, and I may well be mistaken.

Another drawback of this section is that even the proof of the conditional uniqueness is not yet fully satisfactory: It is messy and lengthy and does not give a "high-level reason" for this uniqueness. A better understanding of the proof would perhaps come along with a better understanding of Λ and hence might give us deeper insight into the Cohen-Lenstra heuristic.

3.5.1 Automorphisms of the set of partitions

3.5.1 Definition. *A map $\varphi : \mathcal{P} \to \mathcal{P}$ is an* automorphism *of $\mathcal{P}$ if it is bijective and both φ and φ^{-1} are compatible with the natural ordering on $\mathcal{P}$. In other words, we require the equivalence*

$$\underline{n} < \underline{m} \Leftrightarrow \varphi(\underline{n}) < \varphi(\underline{m}).$$

We have already encountered one non-trivial automorphism: The map which assigns to each partition its *conjugate* partition (i.e., the map which reflects the Young diagram of a partition along the diagonal of the quadrant) is clearly bijective and order-preserving, hence an automorphism. It will turn out that this is the only non-trivial automorphism of $\mathcal{P}$:

3.5.2 Theorem. *The set of automorphisms of $\mathcal{P}$ consists of only two elements: The identity map and the conjugation map.*

We postpone the proof until we have established the following helpful lemma:

3.5.3 Lemma. *The size of a partition is invariant under automorphisms of $\mathcal{P}$.*

Proof. We note that the size n of a partition $\underline{n}$ can be described as the length of the longest strictly increasing chain of partitions from $\underline{1} = (1)$ to $\underline{n}$.
As a formula:

$$n = \max\{k \mid \exists \text{ sequence } (1) = \underline{n_1} < \underline{n_2} < \underline{n_3} < \ldots < \underline{n_k} = \underline{n}\}.$$

Why is that formula true? Since $\underline{n}_i < \underline{n}_{i+1}$, we also have $\text{size}(\underline{n}_i) < \text{size}(\underline{n}_{i+1})$, therefore any sequence from $\underline{1}$ to $\underline{n}$ has length at most n. On the other hand, it is trivial to see that a sequence of length n exists (just fill up the rows in the Young diagram successively). Hence, we have equality.
Now notice that we have redefined the size in terms of the ordering only. Since the ordering is invariant under automorphism, everything that can be computed in terms of the ordering is invariant, too. In particular, the size is invariant, as required. □

Now we are ready to prove theorem 3.5.2:

Proof of Theorem 3.5.2. Let φ be an automorphism of $\mathcal{P}$. Since φ preserves the ordering, it must preserve the empty partition and the partition $\underline{1} = (1)$, because they are the only partitions with size 0 and 1, respectively.
We have two partitions of size 2, namely (2) and $(1, 1)$. Hence, the orbit of (2) under the automorphism group consists of at most 2 elements. If we can

show that the stabilizer of (2) consists only of the identity, we may conclude that the size of the automorphism group is at most $1 \cdot 2 = 2$. Since we have already shown that the identity and the conjugation map are automorphisms, this will finish our proof.
So let us turn to the stabilizer of (2), i.e., we assume $\varphi((2)) = (2)$. We need to show that φ is the identity map, i.e., it preserves every partition $\underline{n}$. For this, we use induction on the size n of $\underline{n}$. Since we make extensive use of the ordering, recall that we say $\underline{n}$ lies above $\underline{m}$ if $\underline{n} > \underline{m}$.

- $n \leq 2$*:* We have already shown this case.

- $n > 2$*:* We assume that φ preserves every partition of size $\leq n-1$.

  We consider two cases: First, let us assume that there are at least two different partitions $\underline{m_1}$ and $\underline{m_2}$ of size $n-1$ that lie under $\underline{n}$. Then the Young diagram of $\underline{n}$ contains both the Young diagrams of $\underline{m_1}$ and $\underline{m_2}$, and hence also the union of their Young diagrams. Since $\underline{m_1}$ and $\underline{m_2}$ are not equal (and not contained in each other due to their equal size), the union of their Young diagrams has size $\geq n$ and is thus equal to the Young diagram of $\underline{n}$. So $\underline{m_1}$ and $\underline{m_2}$ already determine $\underline{n}$, and since they are invariant under φ by our induction hypothesis, $\underline{n}$ is also invariant.

  Now let us consider the second case: There is only one partition $\underline{m}$ of size $n-1$ under $\underline{n}$. Call this property (*). Since automorphisms are order-preserving, this property is also preserved by φ. A short moment of thought shows that in this case all entries of $\underline{n}$ must be equal: $\underline{n} = (i, i, \ldots, i)$, and $\underline{m} = (i, i, \ldots, i, i-1)$. Now we must show that $\underline{n}$ can be reconstructed uniquely from $\underline{m}$. But this is trivial if $\underline{m}$ has at least two entries, or if it has one entry ≥ 2. Together, if the size of $\underline{m}$ is $n-1 \geq 2$, then there is only one $\underline{n}$ of size n with the property (*). Hence, $\underline{n}$ is invariant under φ.

We have concluded our induction and proven the theorem. □

3.5.2 Uniqueness modulo the canonical section

In the following example, I present an order-preserving CL-map the canonical section of which differs from the one of Λ. Furthermore, it is non-trivial, i.e., it is not a mere concatenation of Λ with some order-preserving automorphism of partitions.

3.5.4 Example. By the properties of Λ, we may write

$$\mathcal{P} = \dot{\bigcup_{G \in \mathcal{G}_\mathcal{P}}} \Lambda^{-1}(G) \cong \dot{\bigcup_{G \in \mathcal{G}_\mathcal{P}}} I_G,$$

where I_G is defined as in 3.2.2, and the union on the right hand side is a formal disjoint union.
We define $\Lambda' : \mathcal{P} \to \mathcal{G}_\mathcal{P}$ as follows

$$\Lambda'(\underline{n}) := \begin{cases} 0 & \text{if } \underline{n} = (), \\ \mathbb{Z}/p\mathbb{Z} & \text{if } \underline{n} \in \Lambda^{-1}(\mathbb{Z}/p\mathbb{Z}) \text{ (i.e. } \underline{n} = (1, \dots, 1)), \\ \mathbb{Z}/p^2\mathbb{Z} & \text{if } \underline{n} = (i), i \geq 2, \\ \mathbb{Z}/p^i\mathbb{Z} & \text{if } \underline{n} \in \Lambda^{-1}(\mathbb{Z}/p^{i-1}\mathbb{Z}) \setminus \{(i-1)\}, i \geq 3, \\ \Lambda(\underline{n}) & \text{otherwise.} \end{cases}$$

Before we investigate the example, let me for convenience list the values of Λ in these cases:

$$\Lambda(\underline{n}) = \begin{cases} 0 & \text{if } \underline{n} = (), \\ \mathbb{Z}/p\mathbb{Z} & \text{if } \underline{n} \in \Lambda^{-1}(\mathbb{Z}/p\mathbb{Z}) \text{ (i.e. } \underline{n} = (1, \dots, 1)), \\ \mathbb{Z}/p^i\mathbb{Z} & \text{if } \underline{n} = (i), i \geq 2, \\ \mathbb{Z}/p^{i-1}\mathbb{Z} & \text{if } \underline{n} \in \Lambda^{-1}(\mathbb{Z}/p^{i-1}\mathbb{Z}) \setminus \{(i-1)\}, i \geq 3, \\ \Lambda(\underline{n}) & \text{else.} \end{cases}$$

First, let us show that Λ' is a CL-map:

- For non-cyclic groups G, we have $\Lambda'^{-1}(G) = \Lambda^{-1}(G)$, so there is nothing to show.
- The same is true for $G = \{0\}$ and $G = \mathbb{Z}/p\mathbb{Z}$.
- For $G = \mathbb{Z}/p^2\mathbb{Z}$, we need exactly one partition of size k in the fiber $\Lambda'^{-1}(G)$, for any $k \geq 2$. By definition $\Lambda'^{-1}(G) = \{(i) \mid i \geq 2\}$, so for each $k \geq 2$, there is exactly one partition of k in $\Lambda'^{-1}(G)$, as required.
- For other cyclic groups $G = \mathbb{Z}/p^i\mathbb{Z}$, $i \geq 3$, we again need exactly one partition of size k in the fiber $\Lambda'^{-1}(G)$, for any $k \geq i$. We have $\Lambda'^{-1}(G) = \Lambda^{-1}(\mathbb{Z}/p^{i-1}\mathbb{Z}) \setminus \{(i-1)\}$. Hence, for any $k \geq i$, there is exactly one partition in the fiber, as required.

Now we show that Λ' is order-preserving. We need to show that it preserves order on each fiber:

- For non-cyclic groups G, we have $\Lambda'^{-1}(G) = \Lambda^{-1}(G)$. The same is true for $G = \{0\}$ or $G = \mathbb{Z}/p\mathbb{Z}$. In these cases, there is nothing to show.
- For $G = \mathbb{Z}/p^2\mathbb{Z}$, we need that the partition ordering induces a *total* ordering on $\Lambda'^{-1}(G)$, because we also have a total ordering on $I_G \cong \mathbb{N}$ in this case. This requirement is met, since $\Lambda'^{-1}(G) = \{(i) \mid i \geq 2\}$.
- For other cyclic groups $G = \mathbb{Z}/p^i\mathbb{Z}$, $i \geq 3$, we again need a total ordering on $\Lambda'^{-1}(G) = \Lambda^{-1}(\mathbb{Z}/p^{i-1}\mathbb{Z}) \setminus \{(i-1)\}$. Since the fiber $\Lambda^{-1}(\mathbb{Z}/p^{i-1}\mathbb{Z})$ is totally ordered, so is $\Lambda'^{-1}(G)$.

Note that Λ' is order-preserving even in a stronger sense: For groups $G_1 \subset G_2$, we have $\iota'(G_1) \leq \iota'(G_2)$, where ι' denotes the canonical section of Λ' (cf. definition 3.2.8 and remark 3.2.9).

So we have seen that Λ is not unique. But at least we will prove the following theorem:

3.5.5 Theorem. *Let Λ be the map defined in sections 3.3.1 and 3.3.2, and let $\iota : \mathcal{G}_\mathcal{P} \to \mathcal{P}_{base}$ be its canonical section. Then there is no other order-preserving CL-map with canonical section ι.*

3.5.6 Remark. Note that we do not only require $\mathcal{P}_{base}$ to be fixed, but also the section ι. A slight modification of example 3.5.4 produces a different order-preserving CL-map Λ' with identical base-set $\mathcal{P}_{base}$, but with different section $\iota : \mathcal{G}_\mathcal{P} \to \mathcal{P}_{base}$.

We split the proof of the theorem into several steps. Before we turn to the main arguments, let us first show the following lemmas:

3.5.7 Lemma. *Let $G = \prod_{i=1}^{k} (\mathbb{Z}/p^{e_i}\mathbb{Z})^{r_i}$ be a p-group in standard form, and let $t = (t_{i,s})_{1\leq i\leq k, 1\leq s\leq r_i} \in I_G$, such that $t_{i,s} > 0$ for (at least) two different index pairs (i_1, s_1) and (i_2, s_2).*
Then there exist two different tuples u and v in the fiber I_G that lie immediately below t. (I.e., $u < t$, and any tuple between u and t equals either u or t; same for v.)

Proof. Define u as follows:

$$u_{i,s} = \begin{cases} t_{i,s} & \text{if } i \neq i_1, \\ t_{i,s} & \text{if } i = i_1 \text{ and } s \neq s_1, s \neq s_1 - 1, \\ t_{i,s} - 1 & \text{if } i = i_1 \text{ and } s = s_1, \\ t_{i,s} + 1 & \text{if } i = i_1 \text{ and } s = s_1 - 1. \end{cases}$$

The last case is empty if $s_1 = 1$.
Then $(u_{i_1,s})_{1\le s\le r_{i_1}}$ is dominated by $(t_{i_1,s})_{1\le s\le r_{i_1}}$. Since all other entries coincide, $u < t$. Furthermore, $\operatorname{expon}(t) = \operatorname{expon}(u) + 1$ by a straightforward calculation. Since the exponent is integral and compatible with the ordering (remark 3.2.6), u lies *immediately* below t.
We define v analogously:

$$v_{i,s} = \begin{cases} t_{i,s} & \text{if } i \neq i_2, \\ t_{i,s} & \text{if } i = i_2 \text{ and } s \neq s_2, s \neq s_2 - 1, \\ t_{i,s} - 1 & \text{if } i = i_2 \text{ and } s = s_2, \\ t_{i,s} + 1 & \text{if } i = i_2 \text{ and } s = s_2 - 1. \end{cases}$$

By the same argument, v also lies immediately below t.
It only remains to show that $u \neq v$. If $i_1 \neq i_2$, this is clear. It is also clear if $i_1 = i_2$ and s_1, s_2 differ by at least 2. Finally, if $i_1 = i_2$ and $s_1 = s_2 + 1$, then we note that $u_{i_1,s_1} = t_{i_1,s_1} - 1 \neq t_{i_1,s_1} = v_{i_1,s_1}$. Hence, in all cases $u \neq v$.
□

The most difficult part of the proof is concerned with tuples $t \in I_G$ which are *not* of the form given above. We need a characterization of the images of all such t under μ_G for various G. The next lemma gives a necessary criterion for a partition to be such an image. In fact, the criterion is also sufficient, but since we do not need sufficiency, we will not prove it.

3.5.8 Lemma. *Let $G = \prod_{i=1}^{k} (\mathbb{Z}/p^{e_i}\mathbb{Z})^{r_i}$ be a p-group in standard form, and let $t = (t_{i,s})_{1\le i\le k, 1\le s\le r_i} \in I_G$ such that there is an index (i_0, s_0) with $t_{i_0,s_0} > 0$ and $t_{i,s} = 0$ for all $(i,s) \neq (i_0,s_0)$.*
Let $\underline{n} := \mu_G(t)$ and let $\overline{n} = (\overline{n}_0, \ldots, \overline{n}_m)$ be its derivation. Then there exists $j_0 \in \{0, \ldots, m-2\}$ and an even $k_0 > 0$ ($k_0 = 2s_0$) such that $j_0 + k_0 \le m$, the last index m is even if $j_0 + k_0 < m$, and one of the following two cases holds:

A)
- *j_0 is even.*
- *$\overline{n}_0 = \overline{n}_1 \ge \overline{n}_2 = \ldots \ge \overline{n}_{j_0-2} = \overline{n}_{j_0-1}$ (alternating "$\ge$" and "$=$").*
- *$\overline{n}_{j_0} < \overline{n}_{j_0+1} = \ldots = \overline{n}_{j_0+k_0-1} \ge \overline{n}_{j_0+k_0}$ (all "$=$").*
- *$\overline{n}_{j_0} + \overline{n}_{j_0+k_0} - \overline{n}_{j_0+1} = \overline{n}_{j_0+k_0+1}$ ($:= 0$ if $j_0 + k_0 + 1 > m$).*
- *$\overline{n}_{j_0+k_0} > \overline{n}_{j_0+k_0+1}$.*
- *$\overline{n}_{j_0+k_0+1} \ge \overline{n}_{j_0+k_0+2} = \overline{n}_{j_0+k_0+3} \ge \overline{n}_{j_0+k_0+4} = \ldots \ge \overline{n}_{m-1} = \overline{n}_m$.*
- *$\overline{n}_{j_0-1} \ge \overline{n}_{j_0+k_0+1}$.*

(No comparison between $\overline{n}_{j_0-1}$ and $\overline{n}_{j_0}$!)

B)
- j_0 *is odd.*
- $\overline{n}_0 = \overline{n}_1 \geq \overline{n}_2 = \overline{n}_3 \geq \ldots \geq \overline{n}_{j_0-1}$ *(alternating "≥" and "=").*
- $\overline{n}_{j_0-1} \leq \overline{n}_{j_0}$.
- $\overline{n}_{j_0} < \overline{n}_{j_0+1} = \ldots = \overline{n}_{j_0+k_0-1} \geq \overline{n}_{j_0+k_0}$ *(all "=").*
- $\overline{n}_{j_0} + \overline{n}_{j_0+k_0} - \overline{n}_{j_0+1} = \overline{n}_{j_0-1}$.
- $\overline{n}_{j_0+k_0} > \overline{n}_{j_0+k_0+1}$.
- $\overline{n}_{j_0+k_0+1} = \overline{n}_{j_0+k_0+2} \geq \overline{n}_{j_0+k_0+3} = \ldots \geq \overline{n}_{m-1} = \overline{n}_m$.
- $\overline{n}_{j_0-1} \geq \overline{n}_{j_0+k_0+1}$.

Furthermore, in both cases, there lies exactly one partition immediately below $\underline{n}$ *in* $\Lambda^{-1}(G)$*, and it has derivation*

$$(\overline{n}_1, \overline{n}_2, \ldots, \overline{n}_{j_0-1}, \overline{n}_{j_0} + 1, \overline{n}_{j_0+1}, \overline{n}_{j_0+2} - 1, \overline{n}_{j_0+3}, \overline{n}_{j_0+4}, \ldots, \overline{n}_m).$$

In case A, the multiset $\{\overline{n}_i \mid i \text{ odd}\}$ *is up to zeroes equal to the multiset* $\{e_i\}$*. (But the* $\overline{n}_i$ *are in a different order than the* e_i*!)*
In case B, the multiset $\{\overline{n}_i \mid i \text{ even}\}$ *is up to zeroes equal to the multiset* $\{e_i\}$*. (But the* $\overline{n}_i$ *are in a different order than the* e_i*!)*
In both cases, $\overline{n}_{j_0+1} = e_{i_0}$.
No derivation $\overline{n}$ *of a partition is of type A and B at the same time. (I.e., there are no* j_0*,* j_0' *such that* $\overline{n}$ *is of type A with respect to* j_0 *and of type B with respect to* j_0'*.)*

Proof. The key idea is to look at

$$(\underbrace{e_1, \ldots, e_1}_{2r_1 \text{ times}}, \underbrace{e_2, \ldots, e_2}_{2r_2 \text{ times}}, \ldots, \underbrace{e_k, \ldots e_k}_{2r_k \text{ times}}),$$

which is the derivation of $\iota(G)$ (and therefore, which corresponds to the tuple $t = (0, \ldots, 0) \in I_G$), and successively increase t_{i_0,s_0}.
Assume we have a derivation $\overline{n}$, belonging to some $t \in I_G$, for which all coefficients except possibly t_{i_0,s_0} are 0. Then increasing t_{i_0,s_0} by 1 can be done by the following simple procedure:

(i) Let j_0 be the largest index with the properties
- $\overline{n}_{j_0} > 0$.
- $\overline{n}_j < e_{i_0}$ for all $j \geq j_0 + 2s_0$.

(ii) Decrease $\overline{n}_{j_0}$ by 1.

(iii) Increase $\overline{n}_{j_0+2s_0}$ by 1.

From this algorithm, it is pretty easy to obtain the A-/B-assertions in the lemma. Simply define j_0 as the index that has been increased in the last step, and define $k_0 := 2s_0$. Since I fear that some readers might already be getting lost in too much notation, I will omit the formal proof and give instead a (hopefully) illuminating example:
Assume $G = \mathbb{Z}/p^6\mathbb{Z} \times \mathbb{Z}/p^5\mathbb{Z} \times \mathbb{Z}/p^5\mathbb{Z} \times \mathbb{Z}/p^5\mathbb{Z} \times \mathbb{Z}/p^4\mathbb{Z} \times \mathbb{Z}/p^1\mathbb{Z}$.
Then the derivation of $\iota(G)$ is simply $\overline{n} := (6, 6, 5, 5, 5, 5, 5, 5, 4, 4, 1, 1)$. This corresponds to $t = (0, \dots, 0)$.
What happens if we increase an entry of t? Assume the non-trivial entry is $t_{2,2}$, corresponding to the $e_2 = 5$-part. Then we obtain the sequence

$$\begin{aligned}
& (6, 6, 5, 5, 5, 5, 5, 5, 4, 4, 1, 1) \\
\to\ & (6, 6, 5, 5, 4, 5, 5, 5, 5, 4, 1, 1) \\
\to\ & (6, 6, 5, 5, 4, 4, 5, 5, 5, 5, 1, 1) \\
\to\ & (6, 6, 5, 5, 4, 4, 4, 5, 5, 5, 2, 1) \\
\to\ & (6, 6, 5, 5, 4, 4, 3, 5, 5, 5, 3, 1) \\
\to\ & (6, 6, 5, 5, 4, 4, 2, 5, 5, 5, 4, 1) \\
\to\ & (6, 6, 5, 5, 4, 4, 1, 5, 5, 5, 5, 1) \\
\to\ & (6, 6, 5, 5, 4, 4, 1, 4, 5, 5, 5, 2) \\
\to\ & (6, 6, 5, 5, 4, 4, 1, 3, 5, 5, 5, 3) \\
\to\ & \cdots
\end{aligned}$$

Note how the sequence of 5's moves through the derivation. If you want to prove the lemma, it is helpful to note the following: If you remove the wandering 5's in the manner of the numerical algorithm (replace a 5 and its two neighbors by their sum minus 10), then you always end up with the derivation $(6, 6, 5, 5, 4, 4, 1, 1)$, which is the original derivation with k_0 times the entry $5 = e_{i_0}$ removed. This is true in general and already implies most of the statements.
Let us turn to the statements listed below the case distinction. In I_G, it is clear that there is exactly one tuple that lies immediately below the tuple $t = (0, \dots, 0, t_{i_0,s_0}, 0 \dots, 0)$. Namely, it is $(0, \dots, 0, \underbrace{1}_{i_0,s_0-1}, t_{i_0,s_0} - 1, 0 \dots, 0)$ if $s_0 > 0$ and $(0, \dots, 0, t_{i_0,s_0} - 1, 0 \dots, 0)$ otherwise. Hence, there is also exactly one partition in $\Lambda^{-1}(G)$ immediately below $\underline{n}$, and since we know its t-tuple, we can compute it. (It is feasible to apply the above algorithm $t_{i_0,s_0} - 1$ times with s_0 and afterwards(!) once with $s_0 - 1$.)

The two statements concerning the multisets $\{\overline{n}_i\}$ are clear by what we have seen so far.
It only remains to show that no derivation is of type A and B at the same time. So let $\overline{n}$ be of type A with respect to j_0 and k_0. We need to show that there do not exist j_0' and k_0' such that $\overline{n}$ is of type B with respect to j_0' and k_0'. Assume they exist. Then j_0' is odd and $\overline{n}_{j_0'} < \overline{n}_{j_0'+1}$. By A, this implies $j_0' = j_0 - 1$. Now B implies that $\overline{n}_{j_0} = \overline{n}_{j_0+1}$ (if $k_0 > 2$) or $\overline{n}_{j_0} \geq \overline{n}_{j_0+1}$ (if $k_0 = 2$). In any case, this is a contradiction to A, which states $\overline{n}_{j_0} < \overline{n}_{j_0+1}$. □

Proof of Theorem 3.5.5 (Uniqueness modulo canonical section).
Let Λ' be an order-preserving CL-map with canonical section ι. We prove that for any $\underline{n} \in \mathcal{P}$, $\Lambda'(\underline{n}) = \Lambda(\underline{n})$, and both maps are induced by the same map $\mu : \bigcup_G I_G \to \mathcal{P}$ (cf. definition 3.2.7). We do this by induction on the size n of $\underline{n}$.
For $n = 0$, the statement is clear because () is the only partition of size 0 and is therefore mapped to the trivial group by both maps.
Let $n > 0$ and let $\underline{n}$ be a partition of n. Let $\Lambda(\underline{n}) =: G = \prod_{i=1}^{k} (\mathbb{Z}/p^{e_i}\mathbb{Z})^{r_i}$ in standard form. Then $\underline{n}$ corresponds to its preimage $\mu^{-1}(\underline{n}) =: t = (t_{i,s})_{1\leq i\leq k, 1\leq s\leq r_i} \in I_G$ (all $t_{i,s} \geq 0$). We distinguish several cases:

(i) $t_{i,s} = 0$ for all i and s, or equivalently $\underline{n} \in \mathcal{P}_{base}$. Then

$$\Lambda'(\underline{n}) \overset{\iota \text{ section}}{=} \Lambda(\iota(\Lambda'(\underline{n}))) \overset{\underline{n}\in\mathcal{P}_{base}}{=} \Lambda(\underline{n}).$$

(ii) $t_{i,s} > 0$ for (at least) two different index pairs (i_1, s_1) and (i_2, s_2). Then by lemma 3.5.7, there exist two different tuples $u, v \in I_G$ that lie immediately below t. Let $\underline{m_1} = \mu(u)$ and $\underline{m_2} = \mu(v)$. Then $\underline{m_1}$ and $\underline{m_2}$ lie immediately below $\underline{n}$. Therefore, in the fiber $\Lambda'^{-1}(G)$ there exists a partition that lies immediately above $\underline{m_1}$ and $\underline{m_2}$.

But $\underline{n}$ is the only such partition. (Its Young diagram is the union of the Young diagrams of $\underline{m_1}$ and $\underline{m_2}$.) Therefore, $\underline{n} \in \Lambda'^{-1}(G)$, or in other words, $\Lambda'(\underline{n}) = G$. Note that we have also proven that $\mu^{-1}(\underline{n})$ is uniquely determined.

(iii) There exists an index (i_0, s_0) such that $t_{i_0,s_0} > 0$ and $t_{i,s} = 0$ for all $(i, s) \neq (i_0, s_0)$.

Lemma 3.5.8 tells us that the derivation $\overline{n}$ of $\underline{n}$ is either of type A or of type B. Abusing terminology, I will say in this case that $\underline{n}$ is of type A or B, respectively.

Let $\overline{n} = (\overline{n}_0, \ldots, \overline{n}_m)$ denote the derivation of $\underline{n}$. Let j_0 and k_0 be as in lemma 3.5.8. Then the lemma tells us that the partition $\underline{m}$ with derivation $\overline{m} = (\overline{n}_0, \overline{n}_1, \ldots, \overline{n}_{j_0-1}, \overline{n}_{j_0}+1, \overline{n}_{j_0+1}, \overline{n}_{j_0+2}-1, \overline{n}_{j_0+3}, \overline{n}_{j_0+4}, \ldots, \overline{n}_m)$ is in $\Lambda^{-1}(G)$ and lies immediately below $\underline{n}$.

Since the size of $\underline{m}$ is smaller than the size of $\underline{n}$, we know by induction hypothesis that $\underline{m} \in \Lambda'^{-1}(G)$. Thus there must be a partition immediately above $\underline{m}$ in $\Lambda'^{-1}(G)$. Now the proof runs as follows: We split up the set of all partitions of type A and B into several cases and go through these cases one by one. Then we would like to show that $\underline{n}$ is the only partition immediately above $\underline{m}$ that has not yet been covered by prior cases. Since all the other partitions immediately above $\underline{m}$ are ruled out, we then conclude that $\underline{n} \in \Lambda'^{-1}(G)$, as required.

Unfortunately, it is not true in general that $\underline{n}$ is the only partition which is immediately above $\underline{m}$ and is not covered by former cases. We must relax the notion a bit: We show that the number of partitions immediately above $\underline{m}$ and not covered by prior cases equals the size of the set $\{\underline{n'} \in \Lambda'^{-1}(G) \mid \underline{n'} \text{ lies immediately above } \underline{m}\}$. Then we may again conclude that $\underline{n} \in \Lambda'^{-1}(G)$.

Before we turn to this part of the proof, let me finish the overall argument. So assume we have shown that for any partition $\underline{m}$ of size $n-1$ we have

$$\left\{ \underline{n} \in \Lambda'^{-1}(G) \,\middle|\, \begin{matrix} \underline{n} \text{ lies immedia-} \\ \text{tely above } \underline{m} \end{matrix} \right\} = \left\{ \underline{n} \in \Lambda^{-1}(G) \,\middle|\, \begin{matrix} \underline{n} \text{ lies immedia-} \\ \text{tely above } \underline{m} \end{matrix} \right\}.$$

On the first glance, it seems that we are already done. But in the inductive step, we have not only used that $\underline{m}$ lies in the fiber $\Lambda'^{-1}(G)$ and $\Lambda^{-1}(G)$ of the same group G, but also that $\underline{m}$ corresponds under μ and μ' to the same tuple $(t_{i,s})_{1 \le i \le k, 1 \le s \le r_i} \in I_G$. Formally speaking, this is not true!

However, it is rather easy to show that there is not much variation possible in μ'. More precisely, if $r_{i_1} = r_{i_2}$ for some indices $1 \le i_1, i_2 \le k$, then we may switch $(t_{i_1,s})_{1 \le s \le r_{i_1}}$ and $(t_{i_2,s})_{1 \le s \le r_{i_2}}$, and μ' can be obtained from μ by a sequence of such operations. This can be verified inductively, in parallel with the induction we are yet to complete. Although the proof of this is simple, I omit the details. The missing parts of the induction are complicated, so I want to keep them as "clean" as possible from other reasoning. However, once the reader has worked through the proof, it will be easy to recapitulate it and add the missing details about μ'.

If you believe me in the differences between μ and μ', please note that μ and μ' induce the same ordering on $\mathcal{P}$ — more precisely, for any partition $\underline{m}$, we have

$$\#\left\{\underline{n}\in\Lambda'^{-1}(G)\,\middle|\,\begin{matrix}\underline{n}\text{ lies immedia-}\\ \text{tely above }\underline{m}\end{matrix}\right\}=\#\left\{\underline{n}\in\Lambda^{-1}(G)\,\middle|\,\begin{matrix}\underline{n}\text{ lies immedia-}\\ \text{tely above }\underline{m}\end{matrix}\right\}.$$

This is the reason why our inductive argument works.

So let us return to the missing part of the proof: Recall that $\underline{n}$ is of type A or B, and $\underline{m}$ lies immediately below $\underline{n}$. We have described $\underline{m}$ explicitly in lemma 3.5.8. Let $\underline{n'}$ be any partition that is of type A or B lying immediately above $\underline{m}$, and let $\overline{n'}$ be its derivation. Then we can construct $\overline{n'}$ from $\overline{m}$ by picking some index $j_1\neq j_0$, decreasing $\overline{m}_{j_1}$ by 1 and increasing $\overline{m}_{j_1+2}$ by 1. (This is evidently true for any two derivations where one lies immediately above the other.)

Then extensive case distinction shows that either $\overline{n}=\overline{n'}$, or $\overline{n}$ and $\overline{n'}$ belong to one of the cases 1–17 listed below. Let me first describe how to read the case distinction:

In each case, the first sequence of inequalities refers to $\overline{n}$ and is the prerequisites (including those conditions that are given by underbrackets). There is one further prerequisite that I do not write down explicitly, namely that

$$\overline{n}_{j_0}+\overline{n}_{j_0+k_0}-\overline{n}_{j_0+1}=\begin{cases}\overline{n}_{j_0+k_0+1} & \text{for type A,}\\ \overline{n}_{j_0-1} & \text{for type B.}\end{cases}$$

If a derivation $\overline{n}$ meets the prerequisites, then it is of type A or B and there exists a feasible derivation $\overline{n'}$ as described in the second sequence that lies immediately above $\overline{m}$. This statement is true if you continue the sequence at the left and right in the obvious way, i.e., with alternating "="- and "$\geq$"-symbols. If a derivation lacks entries (e.g., if in case 1 we have $j_0=1$, so $\overline{n}_{j_0-2}$ and $\overline{n}_{j_0-3}$ do not exist), we may fill up the derivation on the left with "∞"'s and on the right with 0's and then check the inequalities. (Note that it is allowed to add something to an entry "0" and get a feasible derivation.)

The description of $\overline{n'}$ is straightforward: I have given a subsequence of $\overline{n'}$ with the entries separated (for convenience only) by relations $(<,=,\ldots)$ or by "," if two entries are not comparable. In the cases 8 and 15, $\overline{n'}$ is not of type A or B, but is in $\mathcal{P}_{base}$. Strictly speaking, it is

not necessary to include those cases, but I felt the case distinction to be more complete in this way.

The cases 16 and 17 include an index j_1, which is to be read as follows: If there exist indices j_0 and j_1 such that both the sequences with j_0 and j_1 occur in $\overline{n}$, then there exists an $\overline{n'}$ as specified. It does not matter which of the two sequences comes first. Note further that j_1 is automatically even. The reason why those two cases are qualitatively different is that $\overline{m} \in \mathcal{P}_{base}$ in these cases. And for such an $\overline{m}$ there may be lots of derivations that lie immediately above $\overline{m}$. In fact, the case 16 describes only derivations that are regularly (i.e., with respect to Λ) above $\overline{m}$.

1.

$\overline{n}$: Type B, $k_0 = 2$

$$\begin{aligned}\ldots \geq \overline{n}_{j_0-3} = \overline{n}_{j_0-2} \geq \underbrace{\overline{n}_{j_0-1}}_{>\overline{n}_{j_0+3}} &\leq \overline{n}_{j_0} \\ &< \overline{n}_{j_0+1} \geq \overline{n}_{j_0+2} > \overline{n}_{j_0+3} = \overline{n}_{j_0+4} \geq \ldots\end{aligned}$$

$\overline{n'}$: Type B, $k_0' = 2$, $j_0' = j_0$

$$\begin{aligned}\ldots \geq \overline{n}_{j_0-3} = \overline{n}_{j_0-2} > \overline{n}_{j_0-1} - 1 &< \overline{n}_{j_0} + 1 \\ &< \overline{n}_{j_0+1} + 1 > \overline{n}_{j_0+2} - 1 > \overline{n}_{j_0+3} = \overline{n}_{j_0+4} \geq \ldots\end{aligned}$$

2.

$\overline{n}$: Type B, $k_0 = 2$

$$\begin{aligned}\ldots \geq \overline{n}_{j_0-3} = \overline{n}_{j_0-2} = \overline{n}_{j_0-1} = \underbrace{\overline{n}_{j_0}}_{=\overline{n}_{j_0+1}-1} &\\ < \overline{n}_{j_0+1} = \overline{n}_{j_0+2} > \underbrace{\overline{n}_{j_0+3}}_{<\overline{n}_{j_0}} &= \overline{n}_{j_0+4} \geq \ldots\end{aligned}$$

$\overline{n'}$: Type B, $k_0' = 4$, $j_0' = j_0 - 2$

$$\begin{aligned}\ldots > \overline{n}_{j_0-3} - 1 < \overline{n}_{j_0-2} < \overline{n}_{j_0-1} + 1 &= \overline{n}_{j_0} + 1 \\ = \overline{n}_{j_0+1} > \overline{n}_{j_0+2} - 1 &> \overline{n}_{j_0+3} = \overline{n}_{j_0+4} \geq \ldots\end{aligned}$$

3.

$\overline{n}$: Type B, $k_0 = 2$

$$\ldots \geq \overline{n}_{j_0-3} = \underbrace{\overline{n}_{j_0-2}}_{=\overline{n}_{j_0-1}+1} > \overline{n}_{j_0-1} = \underbrace{\overline{n}_{j_0}}_{=\overline{n}_{j_0+1}-1}$$
$$< \overline{n}_{j_0+1} = \overline{n}_{j_0+2} > \underbrace{\overline{n}_{j_0+3}}_{<\overline{n}_{j_0-2}} = \overline{n}_{j_0+4} \geq \ldots$$

$\overline{n'}$: Type A, $k_0' = 4$, $j_0' = j_0 - 3$

$$\ldots > \overline{n}_{j_0-3} - 1 < \overline{n}_{j_0-2} = \overline{n}_{j_0-1} + 1 = \overline{n}_{j_0} + 1$$
$$= \overline{n}_{j_0+1} > \overline{n}_{j_0+2} - 1 \geq \overline{n}_{j_0+3} = \overline{n}_{j_0+4} \geq \ldots$$

4.

$\overline{n}$: Type B, $k_0 = 2$

$$\ldots \geq \overline{n}_{j_0-3} = \overline{n}_{j_0-2} \geq \overline{n}_{j_0-1} = \underbrace{\overline{n}_{j_0}}_{=\overline{n}_{j_0+1}-2}$$
$$< \overline{n}_{j_0+1} = \overline{n}_{j_0+2} > \underbrace{\overline{n}_{j_0+3}}_{\leq\overline{n}_{j_0}} = \overline{n}_{j_0+4} \geq \ldots$$

$\overline{n'}$: Type A, $k_0' = 4$, $j_0' = j_0 - 1$

$$\ldots \geq \overline{n}_{j_0-3} = \overline{n}_{j_0-2} \geq \overline{n}_{j_0-1} < \overline{n}_{j_0} + 1$$
$$= \overline{n}_{j_0+1} - 1 = \overline{n}_{j_0+2} - 1 \geq \overline{n}_{j_0+3} + 1 > \overline{n}_{j_0+4} \geq \ldots$$

5.

$\overline{n}$: Type A, $k_0 \geq 6$

$$\ldots \geq \overline{n}_{j_0-2} = \overline{n}_{j_0-1} > \underbrace{\overline{n}_{j_0}}_{=\overline{n}_{j_0+1}-1} < \overline{n}_{j_0+1} = \overline{n}_{j_0+2} = \overline{n}_{j_0+3} = \ldots$$
$$\ldots = \overline{n}_{j_0+k_0-2} = \overline{n}_{j_0+k_0-1} > \underbrace{\overline{n}_{j_0+k_0}}_{=\overline{n}_{j_0+k_0+1}+1} > \overline{n}_{j_0+k_0+1}$$
$$\geq \overline{n}_{j_0+k_0+2} = \overline{n}_{j_0+k_0+3} \geq \ldots$$

$\overline{n'}$: Type A, $k_0' = k_0 - 4$, $j_0' = j_0 - 2$

$$\ldots \geq \overline{n}_{j_0-2} = \overline{n}_{j_0-1} \geq \overline{n}_{j_0} + 1 = \overline{n}_{j_0+1} > \overline{n}_{j_0+2} - 1 < \overline{n}_{j_0+3} = \ldots$$
$$\ldots = \overline{n}_{j_0+k_0-2} > \overline{n}_{j_0+k_0-1} - 1 \geq \overline{n}_{j_0+k_0} = \overline{n}_{j_0+k_0+1} + 1$$
$$> \overline{n}_{j_0+k_0+2} = \overline{n}_{j_0+k_0+3} \geq \ldots$$

6.

$\overline{n}$: Type A, $k_0 = 4$

$$\begin{aligned}\ldots \geq \overline{n}_{j_0-2} = \overline{n}_{j_0-1} > \underbrace{\overline{n}_{j_0}}_{=\overline{n}_{j_0+1}-1} &< \overline{n}_{j_0+1} = \overline{n}_{j_0+2} \\ = \overline{n}_{j_0+3} > \overline{n}_{j_0+4} > \underbrace{\overline{n}_{j_0+5}}_{=\overline{n}_{j_0+4}-1} &\geq \overline{n}_{j_0+6} = \overline{n}_{j_0+7} \geq \ldots\end{aligned}$$

$\overline{n'}$: Type A, $k_0' = 2$, $j_0' = j_0 + 2$

$$\begin{aligned}\ldots \geq \overline{n}_{j_0-2} = \overline{n}_{j_0-1} \geq \overline{n}_{j_0} + 1 = \overline{n}_{j_0+1} > \overline{n}_{j_0+2} - 2 \\ < \overline{n}_{j_0+3} \geq \overline{n}_{j_0+4} + 1 > \overline{n}_{j_0+5} \geq \overline{n}_{j_0+6} = \overline{n}_{j_0+7} \geq \ldots\end{aligned}$$

7.

$\overline{n}$: Type A, $k_0 = 4$

$$\begin{aligned}\ldots \geq \overline{n}_{j_0-2} = \overline{n}_{j_0-1} > \underbrace{\overline{n}_{j_0}}_{=\overline{n}_{j_0+1}-1} &< \overline{n}_{j_0+1} = \overline{n}_{j_0+2} \\ = \overline{n}_{j_0+3} = \overline{n}_{j_0+4} > \underbrace{\overline{n}_{j_0+5}}_{=\overline{n}_{j_0+4}-1} &\geq \overline{n}_{j_0+6} = \overline{n}_{j_0+7} \geq \ldots\end{aligned}$$

$\overline{n'}$: Type B, $k_0' = 2$, $j_0' = j_0 + 3$

$$\begin{aligned}\ldots \geq \overline{n}_{j_0-2} = \overline{n}_{j_0-1} \geq \overline{n}_{j_0} + 1 = \overline{n}_{j_0+1} > \overline{n}_{j_0+2} - 2 \\ < \overline{n}_{j_0+3} < \overline{n}_{j_0+4} + 1 > \overline{n}_{j_0+5} \geq \overline{n}_{j_0+6} = \overline{n}_{j_0+7} \geq \ldots\end{aligned}$$

8.

$\overline{n}$: Type A, $k_0 = 4$

$$\begin{aligned}\ldots \geq \overline{n}_{j_0-2} = \overline{n}_{j_0-1} > \underbrace{\overline{n}_{j_0}}_{=\overline{n}_{j_0+1}-1} &< \overline{n}_{j_0+1} = \overline{n}_{j_0+2} \\ = \overline{n}_{j_0+3} > \overline{n}_{j_0+4} > \underbrace{\overline{n}_{j_0+5}}_{=\overline{n}_{j_0+4}-1} &\geq \overline{n}_{j_0+6} = \overline{n}_{j_0+7} \geq \ldots\end{aligned}$$

$\overline{n'} \in \mathcal{P}_{base}$

$$\begin{aligned}\ldots \geq \overline{n}_{j_0-2} = \overline{n}_{j_0-1} \geq \overline{n}_{j_0} + 1 = \overline{n}_{j_0+1} > \overline{n}_{j_0+2} - 1 \\ = \overline{n}_{j_0+3} - 1 \geq \overline{n}_{j_0+4} = \overline{n}_{j_0+5} + 1 > \overline{n}_{j_0+6} = \overline{n}_{j_0+7} \geq \ldots\end{aligned}$$

9.

$\overline{n}$: Type A, $k_0 = 4$

$$\begin{aligned}\ldots \geq \overline{n}_{j_0-2} = \overline{n}_{j_0-1} = \underbrace{\overline{n}_{j_0}}_{=\overline{n}_{j_0+1}-1} &< \overline{n}_{j_0+1} = \overline{n}_{j_0+2} \\ = \overline{n}_{j_0+3} > \overline{n}_{j_0+4} > \underbrace{\overline{n}_{j_0+5}}_{=\overline{n}_{j_0+4}-1} &\geq \overline{n}_{j_0+6} = \overline{n}_{j_0+7} \geq \ldots\end{aligned}$$

$\overline{n'}$: Type B, $k_0' = 2$, $j_0' = j_0 - 1$

$$\begin{aligned}\ldots \geq \overline{n}_{j_0-2} = \overline{n}_{j_0-1} &< \overline{n}_{j_0} + 1 = \overline{n}_{j_0+1} > \overline{n}_{j_0+2} - 1 \\ = \overline{n}_{j_0+3} - 1 \geq \overline{n}_{j_0+4} &= \overline{n}_{j_0+5} + 1 > \overline{n}_{j_0+6} = \overline{n}_{j_0+7} \geq \ldots\end{aligned}$$

10.

$\overline{n}$: Type A, $k_0 = 4$

$$\begin{aligned}\ldots \geq \overline{n}_{j_0-2} = \overline{n}_{j_0-1} > \underbrace{\overline{n}_{j_0}}_{=\overline{n}_{j_0+1}-1} &< \overline{n}_{j_0+1} = \overline{n}_{j_0+2} \\ = \overline{n}_{j_0+3} > \overline{n}_{j_0+4} > \underbrace{\overline{n}_{j_0+5}}_{=\overline{n}_{j_0+4}-1} &\geq \overline{n}_{j_0+6} = \overline{n}_{j_0+7} \geq \ldots\end{aligned}$$

$\overline{n'}$: Type B, $k_0' = 2$, $j_0' = j_0 + 3$

$$\begin{aligned}\ldots \geq \overline{n}_{j_0-2} = \overline{n}_{j_0-1} &\geq \overline{n}_{j_0} + 1 = \overline{n}_{j_0+1} > \overline{n}_{j_0+2} - 1 \\ = \overline{n}_{j_0+3} - 1 < \overline{n}_{j_0+4} &= \overline{n}_{j_0+5} + 1 > \overline{n}_{j_0+6} = \overline{n}_{j_0+7} \geq \ldots\end{aligned}$$

11.

$\overline{n}$: Type A, $k_0 = 2$

$$\begin{aligned}\ldots \geq \overline{n}_{j_0-2} = \underbrace{\overline{n}_{j_0-1}}_{=\overline{n}_{j_0+1}} > \underbrace{\overline{n}_{j_0}}_{=\overline{n}_{j_0+1}-2} &< \overline{n}_{j_0+1} \\ \geq \overline{n}_{j_0+2} > \underbrace{\overline{n}_{j_0+3}}_{=\overline{n}_{j_0+2}-2} &\geq \overline{n}_{j_0+4} = \overline{n}_{j_0+5} \geq \ldots\end{aligned}$$

$\overline{n'}$: Type A, $k_0' = 4$, $j_0' = j_0 - 2$

$$\begin{aligned}\ldots \geq \overline{n}_{j_0-2} - 1 < \overline{n}_{j_0-1} &= \overline{n}_{j_0} + 2 = \overline{n}_{j_0+1} \\ > \overline{n}_{j_0+2} - 1 > \overline{n}_{j_0+3} &\geq \overline{n}_{j_0+4} = \overline{n}_{j_0+5} \geq \ldots\end{aligned}$$

12.

$\overline{n}$: Type A, $k_0 = 2$

$$\begin{aligned}\ldots \geq \overline{n}_{j_0-2} = \underbrace{\overline{n}_{j_0-1}}_{\geq \overline{n}_{j_0+3}+1} , \underbrace{\overline{n}_{j_0}}_{<\overline{n}_{j_0+1}-2} &< \overline{n}_{j_0+1} \\ \geq \overline{n}_{j_0+2} > \underbrace{\overline{n}_{j_0+3}}_{<\overline{n}_{j_0+2}-2} &\geq \overline{n}_{j_0+4} = \overline{n}_{j_0+5} \geq \ldots\end{aligned}$$

$\overline{n'}$: Type A, $k_0' = 2$, $j_0' = j_0$

$$\begin{aligned}\ldots \geq \overline{n}_{j_0-2} = \overline{n}_{j_0-1} , \ \overline{n}_{j_0} + 1 &< \overline{n}_{j_0+1} - 1 \\ \geq \overline{n}_{j_0+2} - 1 > \overline{n}_{j_0+3} + 1 &> \overline{n}_{j_0+4} = \overline{n}_{j_0+5} \geq \ldots\end{aligned}$$

13.

$\overline{n}$: Type A, $k_0 = 2$

$$\begin{aligned}\ldots \geq \overline{n}_{j_0-2} = \overline{n}_{j_0-1} = \underbrace{\overline{n}_{j_0}}_{=\overline{n}_{j_0+1}-2} &< \overline{n}_{j_0+1} \\ = \overline{n}_{j_0+2} > \underbrace{\overline{n}_{j_0+3}}_{=\overline{n}_{j_0+2}-2} &\geq \overline{n}_{j_0+4} = \overline{n}_{j_0+5} \geq \ldots\end{aligned}$$

$\overline{n'}$: Type B, $k_0' = 4$, $j_0' = j_0 - 1$

$$\begin{aligned}\ldots \geq \overline{n}_{j_0-2} = \overline{n}_{j_0-1} < \overline{n}_{j_0} + 1 &= \overline{n}_{j_0+1} - 1 \\ = \overline{n}_{j_0+2} - 1 = \overline{n}_{j_0+3} + 1 &> \overline{n}_{j_0+4} = \overline{n}_{j_0+5} \geq \ldots\end{aligned}$$

14.

$\overline{n}$: Type A, $k_0 = 2$

$$\begin{aligned}\ldots \geq \overline{n}_{j_0-2} = \overline{n}_{j_0-1} \leq \underbrace{\overline{n}_{j_0}}_{=\overline{n}_{j_0+1}-2} &< \overline{n}_{j_0+1} \\ > \overline{n}_{j_0+2} > \underbrace{\overline{n}_{j_0+3}}_{=\overline{n}_{j_0+2}-2} &\geq \overline{n}_{j_0+4} = \overline{n}_{j_0+5} \geq \ldots\end{aligned}$$

$\overline{n'}$: Type B, $k_0' = 2$, $j_0' = j_0 - 1$

$$\begin{aligned}\ldots \geq \overline{n}_{j_0-2} = \overline{n}_{j_0-1} < \overline{n}_{j_0} + 1 &= \overline{n}_{j_0+1} - 1 \\ > \overline{n}_{j_0+2} - 1 = \overline{n}_{j_0+3} + 1 &> \overline{n}_{j_0+4} = \overline{n}_{j_0+5} \geq \ldots\end{aligned}$$

15.

$\overline{n}$: Type A, $k_0 = 2$

$$\begin{aligned}\ldots \geq \overline{n}_{j_0-2} = \overline{n}_{j_0-1} > \underbrace{\overline{n}_{j_0}}_{=\overline{n}_{j_0+1}-2} < \overline{n}_{j_0+1} \\ \geq \overline{n}_{j_0+2} > \underbrace{\overline{n}_{j_0+3}}_{=\overline{n}_{j_0+2}-2} \geq \overline{n}_{j_0+4} = \overline{n}_{j_0+5} \geq \ldots\end{aligned}$$

$\overline{n'} \in \mathcal{P}_{base}$

$$\begin{aligned}\ldots \geq \overline{n}_{j_0-2} = \overline{n}_{j_0-1} \geq \overline{n}_{j_0} + 1 = \overline{n}_{j_0+1} - 1 \\ \geq \overline{n}_{j_0+2} - 1 = \overline{n}_{j_0+3} + 1 > \overline{n}_{j_0+4} = \overline{n}_{j_0+5} \geq \ldots\end{aligned}$$

16.

$\overline{n}$: Type A, $k_0 = 2$

$$\begin{aligned}\ldots \geq \overline{n}_{j_0-2} = \overline{n}_{j_0-1} > \underbrace{\overline{n}_{j_0}}_{=\overline{n}_{j_0+1}-1} < \overline{n}_{j_0+1} \\ \geq \overline{n}_{j_0+2} > \underbrace{\overline{n}_{j_0+3}}_{=\overline{n}_{j_0+2}-1} \geq \overline{n}_{j_0+4} = \overline{n}_{j_0+5} \geq \ldots \\ \ldots = \overline{n}_{j_1-1} \geq \overline{n}_{j_1} = \overline{n}_{j_1+1} > \overline{n}_{j_1+2} = \overline{n}_{j_1+3} \geq \ldots\end{aligned}$$

$\overline{n'}$: Type A, $k'_0 = 2$, $j'_0 = j_1$

$$\begin{aligned}\ldots \geq \overline{n}_{j_0-2} = \overline{n}_{j_0-1} \geq \overline{n}_{j_0} + 1 = \overline{n}_{j_0+1} \\ > \overline{n}_{j_0+2} - 1 = \overline{n}_{j_0+3} \geq \overline{n}_{j_0+4} = \overline{n}_{j_0+5} \geq \ldots \\ \ldots = \overline{n}_{j_1-1} > \overline{n}_{j_1} - 1 < \overline{n}_{j_1+1} \geq \overline{n}_{j_1+2} + 1 > \overline{n}_{j_1+3} \geq \ldots\end{aligned}$$

17.

$\overline{n}$: Type A, $k_0 = 2$

$$\begin{aligned}\ldots \geq \overline{n}_{j_0-2} = \overline{n}_{j_0-1} > \underbrace{\overline{n}_{j_0}}_{=\overline{n}_{j_0+1}-1} < \overline{n}_{j_0+1} \\ \geq \overline{n}_{j_0+2} > \underbrace{\overline{n}_{j_0+3}}_{=\overline{n}_{j_0+2}-1} \geq \overline{n}_{j_0+4} = \overline{n}_{j_0+5} \geq \ldots \\ \ldots \geq \overline{n}_{j_1} = \overline{n}_{j_1+1} = \overline{n}_{j_1+2} = \overline{n}_{j_1+3} > \overline{n}_{j_1+4} = \overline{n}_{j_1+5} \geq \ldots\end{aligned}$$

$\overline{n'}$: Type B, $k_0' = 2$, $j_0' = j_1 + 1$

$$\begin{aligned} \ldots \geq \overline{n}_{j_0-2} = \overline{n}_{j_0-1} \geq \overline{n}_{j_0} + 1 = \overline{n}_{j_0+1} \\ > \overline{n}_{j_0+2} - 1 = \overline{n}_{j_0+3} \geq \overline{n}_{j_0+4} = \overline{n}_{j_0+5} \geq \ldots \\ \ldots > \overline{n}_{j_1} - 1 < \overline{n}_{j_1+1} < \overline{n}_{j_1+2} + 1 > \overline{n}_{j_1+3} > \overline{n}_{j_1+4} = \overline{n}_{j_1+5} \geq \ldots \end{aligned}$$

Be aware that the case distinction is not exclusive: It might be that $\overline{n}$ belongs to several cases. This is unavoidable, because it reflects the fact that several different derivations may lie immediately above $\overline{m}$.

I will not prove that the case distinction is exhaustive. The proof is straightforward, but rather lengthy. If you want to check it, then it is very helpful to use the following observations: The entries of any derivation of type A or B increase at most twice, and at succeeding positions: from j_0-1 to j_0 and from j_0 to j_0+1. Furthermore, any even j such that $\overline{n}_j \neq \overline{n}_{j+1}$ satisfies $j_0-1 \leq j \leq j_0$ or $j_0+k_0-1 \leq j \leq j_0+k_0$. In particular, there are at most two even indices j such that $\overline{n}_j \neq \overline{n}_{j+1}$. Since both $\overline{n}$ and $\overline{n'}$ are of type A or B, these two properties will rule out most possibilities. For a complete check, just start with any derivation $\overline{n}$ of type A or B and go through all possible values of k_0 ($k_0 = 2, k_0 = 4, k_0 = 6, \ldots$). Now pass from $\overline{n}$ to $\overline{n'}$ as described on page 78 – on partition level this means removing one block from the Young diagram and adding one other block – and check whether the resulting $\overline{n'}$ is of type A or B. You will not encounter any difficulties except for possibly losing your patience.

So let us proceed by working through all the cases. Since the case distinction is exhaustive, we know that for any derivation $\overline{n}$ of type A or B which does not belong to one of the cases 1–17, $\underline{n}$ is the only partition of type A or B immediately above $\underline{m}$. Since there must exist a partition immediately above $\underline{m}$ in $\Lambda'^{-1}(G)$ of type A or B, we conclude $\underline{n} \in \Lambda'^{-1}(G)$, or $\Lambda'(\underline{n}) = G = \Lambda(\underline{n})$, as required.

So we only need to consider the cases 1–17. Let the notation be as in the case distinction, so $\underline{n}$ and $\underline{n'}$ lie immediately above $\underline{m}$. We need to show that in all the cases $\Lambda'(\underline{n'}) \neq \Lambda(\underline{n})$, unless $\Lambda(\underline{n'}) = \Lambda(\underline{n})$. In most cases, we will use the argument that $\underline{n'}$ was already treated in a prior case, and therefore $\Lambda'(\underline{n'}) = \Lambda(\underline{n'}) \neq \Lambda(\underline{n})$, which proves the assertion.

Let me rephrase this argument in all detail, because it is at the heart of our proof: We want to show that Λ' and Λ coincide, specifically we want to show $\Lambda'(\underline{n}) = \Lambda(\underline{n})$. Our case distinction tells us that $\Lambda'(\underline{n})$ could take not only the value $\Lambda(\underline{n})$, but also possibly the value $\Lambda(\underline{n'})$. If

both values are equal, i.e., $\Lambda(\underline{n'}) = \Lambda(\underline{n})$, then we are done. If they are not equal we need to rule out the second possibility, so we need to show $\Lambda'(\underline{n}) \neq \Lambda(\underline{n'})$. Typically we will find that we have already treated $\underline{n'}$ in prior cases, so we already know $\Lambda'(\underline{n'}) = \Lambda(\underline{n'})$. But Λ and Λ' map equally many partitions to $\Lambda(\underline{n'})$. (Strictly speaking, in order to apply induction hypothesis we count not the total number of partitions, but only the number of partitions lying immediately above $\underline{m}$ – but this is also the same for Λ and Λ'). This contradicts the possibility that $\Lambda'(\underline{n}) = \Lambda(\underline{n'})$, because then Λ' would map *more* partitions to $\Lambda(\underline{n'})$ than $\Lambda(\underline{n})$, namely $\underline{n}$, $\underline{n'}$ and all others that are mapped to $\Lambda(\underline{n'})$ by Λ. Therefore, the possibility $\Lambda'(\underline{n}) = \Lambda(\underline{n'})$ is ruled out. When we have treated all cases in which $\underline{n}$ occurs (which may be more than one), then we have ruled out all other possibilities – the only remaining one is $\Lambda'(\underline{n}) = \Lambda(\underline{n})$, which we wanted to show.

We will treat the cases in the order 2, 1, 8, 7, 10, 3, 9, 13, 15, 4, 6, 11, 14, 17, 16, 12, 5.

Case 2: $\overline{n'}$ is of type B with $k'_0 = 4$. This already implies that $\overline{n'}$ does not belong to case 1–17, so it was covered by prior cases. (I.e., either there exist several partitions in $\Lambda^{-1}(\overline{n'})$ immediately below $\overline{n'}$ or there is no non-trivial $\overline{n''}$ such that $\Lambda'(\overline{n'}) = \Lambda(\overline{n''})$ is possible. In both cases, we may conclude $\Lambda'(\overline{n'}) = \Lambda(\overline{n'})$.) Hence, our standard argument works, and we conclude $\Lambda'(\underline{n'}) \neq \Lambda(\underline{n})$.

Case 1: $\overline{n'}$ is of type B. Therefore, it could only occur in the cases 1–4. Case 2 is already ruled out, and looking at $\overline{n'}$ we notice that $\overline{n}'_{j'_0+1}$ is strictly larger than its two neighbors, which rules out cases 3 and 4. Hence $\overline{n'}$ could only again be in case 1.

Now we do a kind of induction within case 1. Recall that the set $\mathcal{G}_{\mathcal{P}}$ of groups is partially ordered by domination (definition 1.2.8). Therefore, Λ induces a partial ordering on the set

$$S := \{\underline{n} \in \mathcal{P} \mid \underline{n} \text{ has size } n \text{ and belongs to case 1}\}$$

as follows: For $\underline{n}^1, \underline{n}^2 \in S$, we say $\underline{n}^1$ strictly dominates $\underline{n}^2$ if and only if $\Lambda(\underline{n}^1)$ strictly dominates $\Lambda(\underline{n}^2)$. (Note that the word "strictly" is crucial in order to obtain a partial ordering.) We may apply induction with respect to this ordering because S is finite. Our induction hypothesis will be that $\Lambda'(\underline{n}^1) = \Lambda(\underline{n}^1)$ for all $\underline{n}^1 \in S$ strictly dominating $\underline{n}$.

Now we only have to look at $\underline{n'}$. Its image $\Lambda(\underline{n'})$ is almost identical with $\Lambda(\underline{n})$, only the two entries $\overline{n}_{j_0+2}$ and $\overline{n}_{j_0}$ are replaced by $\overline{n}_{j_0+2} + 1$ and

$\overline{n}_{j_0} - 1$. (Recall that you can read off the image under Λ by taking the entries of the derivation with odd or even indices if $\underline{n}$ is of type A or B, respectively.) Since $\overline{n}_{j_0+2} > \overline{n}_{j_0}$, $\underline{n'}$ strictly dominates $\underline{n}$. So we can apply the induction hypothesis and get $\Lambda'(\underline{n'}) = \Lambda(\underline{n'}) \neq \Lambda(\underline{n})$. Hence, by our standard argument $\Lambda'(\underline{n}) = \Lambda(\underline{n})$.

Case 8: This case is trivial because $\overline{n'} \in \mathcal{P}_{base}$, and therefore $\Lambda'(\overline{n'}) = \Lambda(\iota(\Lambda'(\overline{n'}))) = \Lambda(\overline{n'})$.

Case 7+10: If $\overline{n}$ belongs to one of those cases, then it is of type A with $k_0 = 4$ and has 4 equal coefficients $\overline{n}_{j_0+1} = \ldots = \overline{n}_{j_0+4}$. Therefore it is not compatible with any other case. Let us first investigate case 7: $\overline{n'}$ is of type B, so it could only belong to case 1–4. Cases 1 and 2 are already treated, and looking closer at $\overline{n'}$, we see that $\overline{n}'_{j'_0+1}$ is strictly larger than its two neighbors, which contradicts cases 3 and 4.

In case 10, we have $\Lambda(\overline{n}) = \Lambda(\overline{n'})$: The values $\overline{n}_{j_0}, \overline{n}_{j_0+2}, \overline{n}_{j_0+4}$ are replaced by $\overline{n}_{j_0+1}, \overline{n}_{j_0+3} - 1, \overline{n}_{j_0+5} + 1$, which are the same three integers. Together we see that any derivation in cases 7 and 10 must satisfy $\Lambda'(\overline{n}) = \Lambda(\overline{n})$.

Case 3: $\overline{n'}$ is of type A with $k'_0 = 4$. Furthermore, it has four equal values $\overline{n}'_{j'_0+1} = \ldots = \overline{n}'_{j'_0+4}$. Therefore, it could only belong to Case 7 or 10, which are already done.

Case 9: $\overline{n'}$ is of type B. Since 1,2,3 are done, this leaves only possibly case 4. But $\overline{n}'_{j'_0+1} = \overline{n}'_{j'_0} + 1$, which is a contradiction to the condition $\overline{n}_{j_0+1} = \overline{n}_{j_0} + 2$ of case 4.

Case 13: $\overline{n'}$ is of type B with $k'_0 = 4$. No case fits into this pattern.

Case 15: This case is trivial, since $\overline{n'} \in \mathcal{P}_{base}$.

Case 4,6,11,14: This is the hardest part. First look at case 4. Then $\overline{n'}$ is of type A with $k_0 = 4$. There is only one case left which fits into that pattern, namely case 6. I claim that $\Lambda(\overline{n'})$ is strictly dominated by $\Lambda(\overline{n})$ (in the same sense as in case 1.) In fact, it emerges from the latter one by replacing $\overline{n}_{j_0-1}$, $\overline{n}_{j_0+1}$ and $\overline{n}_{j_0+3}$ by $\overline{n}_{j_0} + 1$, $\overline{n}_{j_0+2} - 1$ and $\overline{n}_{j_0+4}$. Since we have the equalities $\overline{n}_{j_0-1} = \overline{n}_{j_0} = \overline{n}_{j_0+1} - 2 = \overline{n}_{j_0+2} - 2$ and $\overline{n}_{j_0+3} = \overline{n}_{j_0+4}$, the claim becomes obvious.

Now inspect case 6. Obviously, $\Lambda(\overline{n'}) = \Lambda(\overline{n})$ because the odd entries remain unchanged. Further, $\overline{n'}$ is of type A with $k_0 = 2$ and with $\overline{n}'_{j'_0} = \overline{n}'_{j'_0+1} - 2$. Therefore, $\overline{n'}$ could fit the conditions of cases 11, 13, 14 and 15, of which 13 and 15 are already ruled out. So the cases 11 and 14 are left.

Next we turn to case 11. Again the odd enries are unchanged so we have $\Lambda(\overline{n'}) = \Lambda(\overline{n})$. Further, $\overline{n'}$ is of type A with $k_0 = 4$. The only remaining case with these parameters is case 6.

Finally we look at case 14. Here, $\overline{n'}$ has type B and could therefore be contained in case 4. We see that $\Lambda(\overline{n'})$ is obtained from $\Lambda(\overline{n})$ by replacing $\overline{n}_{j_0+1}$ and $\overline{n}_{j_0+3}$ by $\overline{n}_{j_0} + 1$ and $\overline{n}_{j_0+2} - 1$, or in other terms by $\overline{n}_{j_0+1} - 1$ and $\overline{n}_{j_0+3} + 1$. Since $\overline{n}_{j_0+1} > \overline{n}_{j_0+3} + 2$, this implies that $\Lambda(\overline{n})$ strictly dominates $\Lambda(\overline{n'})$.

Now we have all the ingredients together to start an induction with respect to domination. We define the set

$$S_4 := \{\overline{n} \in \mathcal{P} \mid \overline{n} \text{ has size } n \text{ and belongs to case 4.}\}$$

and in the same manner sets S_6, S_{11} and S_{14}. Let $\overline{n}$ be in any of these four sets. Our induction hypothesis is that for any derivation $\overline{n}^1$ in one of the four sets such that $\overline{n}$ dominates $\overline{n}^1$, we already know that $\Lambda(\overline{n}^1) = \Lambda'(\overline{n}^1)$.

Assume that $\overline{n} \in S_4$. Then we know that $\overline{n'} \in S_6$ is dominated by $\overline{n}$, so we may apply the induction hypothesis to $\overline{n'}$, and therefore $\Lambda'(\overline{n'}) = \Lambda(\overline{n}) \neq \Lambda'(\overline{n})$. Thus we conclude by our standard argument that $\Lambda'(\overline{n}) = \Lambda(\overline{n})$.

For $\overline{n} \in S_{14}$ the reasoning is completely analogous. So let us assume $\overline{n'} \in S_6$ or $\overline{n'} \in S_{11}$. Let $T_1 := (S_6 \cup S_{11} \cup S_{14}) \cap \Lambda^{-1}(G)$, and let $T_2 := (S_6 \cup S_{11} \cup S_{14}) \cap \Lambda'^{-1}(G)$. Then $T_2 \subseteq T_1$ by our above reasoning. On the other hand, by the properties of Λ' we know that T_1 and T_2 have equally many elements. Hence $T_1 = T_2$. In particular, $\overline{n'} \in T_1 = T_2$, which implies $\Lambda'(\overline{n'}) = G = \Lambda(\overline{n'})$.

Case 17: $\overline{n'}$ is of type B and there are no cases left with $\overline{n}$ of type B.

Case 16: $\overline{n'}$ is of type A and $k_0' = 2$. The only fitting case that remains is again 16. But a brief look shows that the entries with odd index are unchanged, so $\Lambda(\overline{n'}) = \Lambda(\overline{n})$. Furthermore, $\overline{n}$ is not compatible with any other open case. Thus $\Lambda'(\overline{n'}) = \Lambda(\overline{n'})$.

Case 5: $\overline{n'}$ is of type A and $k_0' = k_0 - 4$. Since no other cases are left, we may use a trivial induction by k_0.

This completes the case distinction and the proof.

□

Chapter 4

Computing Interesting Values

In this chapter, I will present other methods for computing concrete values concerning the Cohen-Lenstra probability measure, such as expected values and higher moments. These fall into two categories: Firstly, the theory of zeta functions invented by Cohen and Lenstra, and secondly, methods invented for studying conjugacy classes of the general linear group $\mathrm{GL}(n, p)$. This distinction is purely historical. We will see that the theory of conjugacy classes provides us with a plentitude of tools.

Although this theory is fully developed (e.g., cf. [Ger61], [Kun81], [RS88], [Sto93], [Ful97], [Ful99], [Ful00b]), the connection to the Cohen-Lenstra heuristic seems to have slipped general attention in both direction: Neither were the group theorists aware of the Cohen-Lenstra heuristic [Ful08], nor did the number theorists recognize the full connection to conjugacy classes (although Washington was aware of corollary 4.6.4 about fixed spaces [Was86], which is a special case of the general relationship).

These circumstances give me the golden opportunity to reap the fruits of other people's hard work. I want to emphasize that all the results in this chapter are not my own work. My humble contribution is only to re-interpret established results in the notion of the Cohen-Lenstra heuristic. However, since this connection was generally unnoticed until now I gather the most important results from conjugacy class theory that transfer to statements about the Cohen-Lenstra heuristic. Particularly of interest is the work of Jason Fulman [Ful97], [Ful99], [Ful00b] who examined precisely the conjugacy theory analogue of the Cohen-Lenstra probability.

But first I review Cohen and Lenstra's zeta function approach.

4.1 Zeta functions

Cohen and Lenstra embed what I call the Cohen-Lenstra weight w into a larger family of measures w_k as follows. For a finite abelian p-group G, let $s_k(G)$ be the number of surjective homomorphisms $\mathbb{Z}^k \to G$ (or, equivalently, $\mathbb{Z}_p^k \to G$). Then they define

$$w_k(G) := \frac{s_k(G)}{|G|^k} w(G).$$

Note that the denominator equals the number of *all* (not necessarily surjective) homomorphisms $\mathbb{Z}^k \to G$.
Then we may compute $w_k(G)$ as

$$w_k(G) = \begin{cases} w(G) \prod_{i=k-r+1}^{k} (1-q^i) & \text{if } k \geq r, \\ 0 & \text{otherwise.} \end{cases} \tag{4.1}$$

This follows from our proof of theorem 2.2.1.
In particular, we may recover $w(G)$ as

$$w(G) = \lim_{k \to \infty} w_k(G).$$

Now we define the k-ζ-function over $\mathcal{G}_p$ as

$$\zeta_k^{(p)}(s) := \sum_{G \in \mathcal{G}_p} \frac{w_k(G)}{|G|^s}.$$

Then $\zeta_k^{(p)}$ converges for $\Re(s) > -1$ and may be computed explicitly by

$$\zeta_k^{(p)}(s) = \prod_{i=1}^{k} \frac{1}{(1-p^{-s-i})}$$

([CL84, Cor. 3.7]).
In particular, this implies the formula $\zeta_{k_1+k_2}^{(p)}(s) = \zeta_{k_1}^{(p)}(s+k_2)\zeta_{k_2}^{(p)}(s)$.
We need one last definition: Let $f : \mathcal{G}_p \to \mathbb{C}$ be a integrable function. We define

$$\zeta_k^{(p)}(f;s) := \sum_{G \in \mathcal{G}_p} \frac{w_k(G) f(G)}{|G|^s}.$$

Then the expected value $E(f)$ of f may be computed as

$$E(f) = \lim_{k\to\infty} \frac{\zeta_k^{(p)}(f;0)}{\zeta_k^{(p)}(0)}.$$

(This is an analogue of [CL84, Cor. 5.5], only for local groups.)

Often, it is easier to compute the ζ-function of f than to compute the expected value of f directly. In this way, Cohen and Lenstra compute explicit formulas for the rank and the order of groups, and for some other functions (cf. the discussion in 4.6).
Their approach has two more advantages. Firstly, we get almost for free a treatment of the twisted probability measure P_u discussed in section 4.6.5, which is of special interest for number field extensions that are not imaginary quadratic (see 6.1.2 for details).
More precisely, we may compute the expected value $E_u(f)$ of f with respect to the twisted probability measure P_u as

$$E_u(f) = \lim_{k\to\infty} \frac{\zeta_k^{(p)}(f;u)}{\zeta_k^{(p)}(u)}$$

([CL84, Cor. 5.5]).
The second advantage is that the approach gives a way to obtain some statements about the global setting. We may analogously define a ζ-function over the global set $\mathcal{G}$, it only has a smaller domain of convergence. More precisely, it converges for $\Re(s) > 0$ and has a simple pole in 0. Therefore, under some technical conditions the expected value of certain *global* functions $f : \mathcal{G} \to \mathbb{C}$ may be computed as

$$E(f) = \lim_{s\to 0} \lim_{k\to\infty} \frac{\zeta_k(f;s)}{\zeta_k(s)}$$

([CL84, Thm. 5.5]), and we only need to compute the residues of the global ζ-functions. However, note that we cannot use this approach to define a probability measure on $\mathcal{G}$. Taking the sets for which the above limit exists only yields a content (definition 5.1.1), and does not avoid the problems we address in chapter 5.

4.2 The Cohen-Lenstra heuristic: Interpretation via conjugacy classes

For the rest of the chapter, we fix a prime p. All following definitions implicitly depend on p.

Consider the general linear group $\mathrm{GL}(n,p)$ of invertible $n \times n$-matrices over $\mathbb{F}_p$. Then each conjugacy class can be represented by a matrix in *Jordan-Chevalley normal form*.
Before I describe this form, let me define the *companion matrix* $C(\varphi)$ of a normalized polynomial $\varphi = X^m + a_{m-1}X^{m-1} + \ldots + a_1 X + a_0$. We set $C(\varphi)$ to be the $m \times m$-matrix

$$C(\varphi) := \begin{pmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & 1 \\ -a_0 & -a_1 & -a_2 & \ldots & -a_{m-1} \end{pmatrix}.$$

Now back to the normal form. It looks as follows: For every monic irreducible polynomial ϕ of degree m over $\mathbb{F}_p$ and every positive integer s we may have an arbitrary number (possibly 0) of (ϕ, s)-Jordan blocks. Each Jordan block is a square of size sm and is the companion matrix of the polynomial ϕ^s. The normal form then has the form

$$\begin{pmatrix} J_1 & 0 & 0 & \ldots & 0 \\ 0 & J_2 & 0 & \ldots & 0 \\ 0 & 0 & J_3 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & J_r \end{pmatrix},$$

where J_k runs through all the Jordan blocks. We only require that the sizes of the Jordan blocks add up to n.
The normal form works over every field. In section 4.5, we will also work over the field $\mathbb{F}_{p^i}$, but for the basic theorems it suffices to consider $\mathbb{F}_p$. Note that over an algebraically closed field (such as $\mathbb{C}$) all irreducible polynomials are linear and the Jordan-Chevalley normal form reduces to a slight variation of the ordinary Jordan normal form.
In order to specify a normal form we must specify for every monic irreducible polynomial ϕ and any $s > 0$ how many (ϕ, s)-Jordan blocks occur. In other words, for each ϕ we must specify a partition. We call this partition $\underline{\lambda}_\phi$. For example, if we have 2 blocks of size $3m$ and 3 blocks of size m then this corresponds to the partition $(3, 3, 1, 1, 1)$. In order for the matrix to be invertible we must require that $\underline{\lambda}_X = ()$.
On the other hand, every collection of partitions $(\underline{\lambda}_\phi)_\phi$ with the properties

- $\underline{\lambda}_X = \underline{0}$ and

- $\sum_{\phi,s}(\deg\phi)\lambda_{\phi,s} = n$

defines a (unique) conjugacy class in $\mathrm{GL}(n,p)$.

From now on, we fix a monic polynomial $\phi \neq X$ over $\mathbb{F}_p$ of degree 1.
Let $\underline{\lambda}$ be a partition. Pick a random matrix in $\mathrm{GL}(n,p)$ uniformly at random. Then we get a certain probability for the event $\underline{\lambda}_\phi = \underline{\lambda}$.
Fulman proved the following theorem.

4.2.1 Theorem. *Let ϕ be any monic polynomial over $\mathbb{F}_p$ of degree 1 and let $\underline{\lambda}$ be a partition. As $n \to \infty$, the probability (in the sense above) that $\underline{\lambda}_\phi = \underline{\lambda}$ for a random matrix in $\mathrm{GL}(n,p)$ (chosen uniformly at random) converges to the CL-probability $P(\underline{\lambda})$.*

Proof. [Ful97, Sect. 3.3, Cor. 5 and Sect. 2.7, Lemma 6 and Thm. 5 with $u = 1$ and $N \to \infty$]. □

4.2.2 Remark.

- Fulman uses in his thesis a slightly different way of taking the $n \to \infty$ limit. Rather, he chooses a parameter $0 < u < 1$, then picks the integer n with probability $(1-u)u^n$ and chooses a random matrix from $\mathrm{GL}(n,p)$ (cf. [Ful99, p.557f.]). Then he proceeds as above. However, it is easy to see that letting $u \to 1$ in this setting yields the same limit as letting $n \to \infty$ in the theorem above. We only need to interchange two limits, but this is no problem since all statements concern formal power series identities with positive convergence radius.

  The reason why Fulman chose the parameter u instead of n will become clear in section 4.5 about the cycle index.

- Fulman studies also the probability distribution for monic polynomials ϕ of higher degree. This yields similar distributions with similar formulas, only it does not give exactly the Cohen-Lenstra probability. We will encounter these other distributions in the context of the Kung-Stong cycle index in section 4.5.

The theorem allows us to transfer a multitude of methods and results from a whole community of researchers to the Cohen-Lenstra heuristic. I start with reviewing a very interesting interpretation of the Cohen-Lenstra heuristic in terms of Markov chains due to Fulman.

4.3 Interpretation via Markov chains

In his PhD thesis, Fulman gave two interpretations of the Cohen-Lenstra probability. One as the outcome of a probabilistic algorithm, one as the weight in the Young lattice with certain transition probabilities. I review both interpretations in the setting that is relevant to us.
First I present what Fulman calls the "Young Tableau Algorithm" ([Ful99]). Recall that p is a fixed prime.

4.3.1 Algorithm.

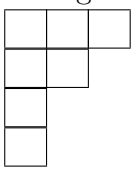
0. Start with $\underline{\lambda}$ the empty partition. Also start with $N = 1$ and with a collection of coins indexed by the natural numbers, such that coin i has probability $\frac{1}{p^i}$ of heads and $1 - \frac{1}{p^i}$ of tails.

1. Flip coin N. If the outcome is tails then set $N := N + 1$ and redo step 1, otherwise go to step 2.

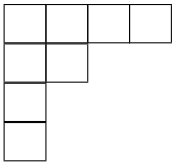
2. Choose an integer $S > 0$ according to the following rule. Set $S := 1$ with probability $\frac{p^{N-\underline{\lambda}_1}-1}{p^N-1}$. For $s > 1$, set $S := s$ with probability $\frac{p^{N-\underline{\lambda}_s}-p^{N-\underline{\lambda}_{s-1}}}{p^N-1}$. Then increase $\underline{\lambda}_S$ by 1 and go to step 1.

In step 2, we use the convention that all undefined entries of $\underline{\lambda}$ are 0. In particular, if we increase some $\underline{\lambda}_s$ that is not defined then after increasing the entry is 1.
The algorithm does not halt, but $\underline{\lambda}$ converges against some limit partition $\underline{\lambda}_\infty$ (cf. theorem 4.3.4 below). The ouput of the algorithm is the *conjugate* partition $\underline{\lambda}'_\infty$ of $\underline{\lambda}_\infty$.

4.3.2 Example. Assume that we are at step 1 with $\underline{\lambda} = (3, 2, 1, 1)$, so the Young diagram of $\underline{\lambda}$ is



Assume further that $N = 4$ and that coin 4 comes up heads, so we go to step 2. We add to $\underline{\lambda}_1$ with probability $\frac{p-1}{p^4-1}$, to $\underline{\lambda}_2$ with probability $\frac{p^2-p}{p^4-1}$, to $\underline{\lambda}_3$ with probability $\frac{p^3-p^2}{p^4-1}$, to $\underline{\lambda}_4$ with probability 0, and to $\underline{\lambda}_5$ with probability $\frac{p^4-p^3}{p^4-1}$.
Assume that we choose $S = 1$ and increase $\underline{\lambda}_1$, thus getting $\underline{\lambda} = (4, 2, 1, 1)$ with Young diagram



We return to step 1 and still have $N = 4$. Assume that again coin 4 comes up heads and we go to step 2. Now we add to $\underline{\lambda}_1$ with probability 0, to $\underline{\lambda}_2$ with probability $\frac{p^2-1}{p^4-1}$, to $\underline{\lambda}_3$ with probability $\frac{p^3-p^2}{p^4-1}$, to $\underline{\lambda}_4$ with probability 0, and to $\underline{\lambda}_5$ with probability $\frac{p^4-p^3}{p^4-1}$. Then we return to step 1.

4.3.3 Remark. The name "Young Tableau Algorithm" refers to the concepts of Young tableaux. A Young tableau is a Young diagram where the boxes are labelled with $1, \ldots, n$ (n the size of the Young diagram). The labels must be given in a way that for any $1 \leq i \leq n$ the boxes $1, \ldots, i$ form again a Young diagram. You may think of a Young tableau as a Young diagram together with an ordering which tells you how to build up the diagram from scratch. Since the algorithm does exactly this (building up Young diagrams block by block), the name is appropriate.

4.3.4 Theorem. *With probability* 1*, the algorithm outputs a* finite *partition. For any given partition* $\underline{\lambda}$*, the probability that the algorithm outputs* $\underline{\lambda}$ *equals the Cohen-Lenstra probability* $P(\underline{\lambda})$*.*

Since the concept of such an algorithm may be unfamiliar to the reader, let me rephrase the finiteness statement of the theorem. Let us say the algorithm has been running for some (finite) time and is in some state $\underline{\lambda}$. Then there is a *positive* probability that the algorithm will not add any more blocks to $\underline{\lambda}$ in all the (infinitely many) forthcoming steps of the algorithm. Thus, there is a positive probability that the algorithm outputs $\underline{\lambda}$. On the other hand, the probability that the algorithm adds infinitely many blocks to $\underline{\lambda}$ in the (infinite) sequel of the algorithm is 0. Hence, with probability 1 the algorithm outputs a finite partition.

Proof of theorem 4.3.4.
[Ful99, Thm. 1] with $u = 1$ and $q = p$. The author states termination of the algorithm only for the case $u < 1$, but his proof implies termination for $u = 1$ as well.
It may be of interest to state one intermediate result in Fulman's proof. Namely, the probability $P^N_{alg}(\underline{\lambda})$ that the generic partition of the algorithm equals $\underline{\lambda}$ at the time when coin N comes up tails is

$$P^N_{alg}(\underline{\lambda}) = \begin{cases} \left(\prod_{i=N-\underline{\lambda}'_1+1}^{N}(1-p^{-i})\right)\left(\prod_{i=1}^{N}(1-p^{-i})\right) w(\underline{\lambda}') & \text{if } \underline{\lambda}'_1 \leq N, \\ 0 & \text{if } \underline{\lambda}'_1 > N, \end{cases} \tag{4.2}$$

where $w(\underline{\lambda}')$ is the Cohen-Lenstra weight of the conjugate $\underline{\lambda}'$ of $\underline{\lambda}$.
Evidently, this converges to $P(\underline{\lambda}')$ as $N \to \infty$. Since the algorithm conjugates the output at the end, it will eventually output $\underline{\lambda}'$. □

4.3.5 Remark. Formula (4.2) is of particular interest because it is identical with formula (2.2) in the proof of theorem 2.2.1.(ii) (up to conjugation of $\underline{\lambda}$). This means that the probability that $\underline{\lambda}$ is the intermediary result in Fulman's algorithm when coin N comes up tails equals the probability that a random matrix $A \in \mathbb{Z}_p^{n\times n}$ has cokernel $\underline{\lambda}' \in \mathcal{G}_{\mathcal{P}} = \mathcal{G}_p$.
So the algorithm is compatible with the graded (by n) structure of the process of choosing generators and relations described in section 2.2.3.

4.4 Interpretation in the Young lattice

Fulman's second interpretation is perhaps even more interesting from our point of view, since it connects more directly to the *CL-weight* rather than to the *CL-probability*.
This approach makes use of the *Young lattice*. The Young lattice is a directed graph with vertex set $\mathcal{G}_{\mathcal{P}}$ ($= \mathcal{G}_p$, but independent of p!). There is a directed edge from $\underline{\lambda}$ to $\underline{\mu}$ if and only if the Young diagram of $\underline{\lambda}$ is contained in the Young diagram of $\underline{\mu}$ and $\text{size}(\underline{\lambda}) = \text{size}(\underline{\mu}) - 1$.
For the algorithm we will index the vertices by the conjugate $\underline{\lambda}'$ of $\underline{\lambda}$. This does not affect the edge set. Note that there is a directed edge from $\underline{\lambda}$ to $\underline{\mu}$ if and only if there is an index i_0 such that $\underline{\mu}'_{i_0} = \underline{\lambda}'_{i_0} + 1$ and $\underline{\mu}'_i = \underline{\lambda}'_i$ for all $i \neq i_0$.

4.4.1 Theorem. *Put weights $m_{\underline{\lambda}',\underline{\mu}'}$ on the edges in the Young lattice as follows:*

(i)

$$m_{\underline{\lambda}',\underline{\mu}'} = \frac{1}{p^{\underline{\lambda}'_1}(p^{\underline{\lambda}'_1+1}-1)} \qquad \text{if } \underline{\mu}'_1 = \underline{\lambda}'_1 + 1.$$

(ii)

$$m_{\underline{\lambda}',\underline{\mu}'} = \frac{p^{-\underline{\lambda}'_s} - p^{-\underline{\lambda}'_{s-1}}}{p^{\underline{\lambda}'_1} - 1} \qquad \text{if } \underline{\mu}'_s = \underline{\lambda}'_s + 1 \text{ for } s > 1.$$

Then the following formula holds for the Cohen-Lenstra weight w and for any $\underline{\lambda} \in \mathcal{G}_{\mathcal{P}}$ of size λ:

$$w(\underline{\lambda}) = \sum_{\gamma'} \prod_{i=0}^{\lambda-1} m_{\gamma'_i,\gamma'_{i+1}},$$

where $\gamma' = (\gamma'_1, \ldots, \gamma'_\lambda)$ runs over all directed paths from the empty partition to $\underline{\lambda}'$ in the Young lattice.

Proof. [Ful99, Thm. 2] □

4.4.2 Remark. A brief calculation shows that for any partition $\underline{\lambda} \in \mathcal{P}$ the sum of the weights of edges out of $\underline{\lambda} \neq ()$ is $\frac{p}{p^{\underline{\lambda}'_1+1}-1} < 1$. (For $\lambda = ()$, it is $\frac{1}{p-1} < 1$.) Therefore, the edge weights can also be viewed as transition probabilities, provided that we allow for halting.

4.5 The Kung-Stong cycle index

This is a powerful tool for investigating conjugacy classes of groups, developed by Kung, Stong and Fulman. The techniques apply also to more general algebraic groups, but for us only the group $\mathrm{GL}(n,p)$ is of interest. Recall (section 4.2) that a conjugacy class of a matrix $M \in \mathrm{GL}(n,p)$ is described by assigning a partition $\underline{\lambda}_\phi(M)$ to each monic irreducible polynomial $\phi \neq X$ such that $\sum_{\phi,s}(\deg\phi)\underline{\lambda}_{\phi,s}(M) = n$.

4.5.1 Definition. *For all $\phi \neq X$ and all partitions $\underline{\lambda}$, let $x_{\phi,\underline{\lambda}}$ be a variable. Then the* cycle index $Z_{\mathrm{GL}(n,p)}$ *is defined as follows:*

$$Z_{\mathrm{GL}(n,p)} := \frac{1}{|\mathrm{GL}(n,p)|} \sum_{M \in \mathrm{GL}(n,p)} \prod_{\phi \neq X} x_{\phi,\underline{\lambda}_\phi(M)}.$$

This cycle index is connected with the Cohen-Lenstra probability. In order to formulate the connection, we embed the CL-probability in a larger class of probability measures on $\mathcal{G}_p$. For any power p^i of p and real number $0 < u < 1$, we define a probability distribution P_{u,p^i} on $\mathcal{G}_p$ as follows. Fix a monic polynomial $\phi \neq X$ over $\mathbb{F}_{p^i}$ of degree 1. Choose an integer n randomly according to the probability distribution $k \mapsto (1-u)u^k$. Now pick a matrix

$M \in \mathrm{GL}(n, p^i)$ uniformly at random. Then the pair (M, ϕ) defines a partition $\underline{\lambda}_\phi(M)$. We define $P_{u,p^i}(\underline{\lambda})$ to be the probability that $\underline{\lambda}_\phi(M) = \underline{\lambda}$. (This is easily seen to be independent of the choice of ϕ.)
Recall that the CL-probability is obtained from P_{u,p^i} by setting $i := 1$ and letting $u \to 1$.
Explicit formulas for P_{u,p^i} are given in [Ful99, sect. 2]. (The author writes $M_{(u,q)}$ instead of P_{u,p^i}.)
Now we can state the following theorem due to Kung [Kun81] and Stong [Sto88]:

4.5.2 Theorem.

$$(1-u)\left(1+\sum_{n=1}^{\infty} Z_{\mathrm{GL}(n,p)} u^n\right) = \prod_{\phi \neq X} \sum_{\underline{\lambda}} x_{\phi,\underline{\lambda}} P_{u,p^{\deg(\phi)}}(\underline{\lambda}).$$

Proof. [Ful97, Thm. 10] □

We will not go into too detail about the techniques that extract interesting consequences from this formula, but the essential point is – possibly after some formula manipulation – comparing the coefficients of u^n on both sides. I refer to [Ful97], [Ful99] and [Ful00b] for tons of examples.

4.6 A collection of results

In this section I cite results that were obtained by the number theory community and the group theoretic community. Some of them were found by both communities, some not.
Recall that (as everywhere else in this thesis except for chapter 5) a "randomly chosen group" really means a randomly chosen finite abelian p-group with respect to the Cohen-Lenstra probability with $q = \frac{1}{p}$ regarded as a formal variable. It was explained in section 3.2 why we may consider q as a formal variable.

4.6.1 Order

4.6.1 Theorem. *The probability that a randomly chosen group has order p^n is*

$$P(\mathrm{ord}(G) = p^n) = q^n \prod_{i=n+1}^{\infty} (1-q^i).$$

Proof. [CL84, Cor. 3.8] □

Higher moments of the order

Recall that the k-th moment of a random variable X is the expected value of X^k.
The higher moments of the order of a random group do not exist if $k \geq 1$. (I.e., their values are ∞.) However, for the p-logarithm of the order (which we will define as *local order* $\mathrm{ord}_p(G)$ in definition 5.2.1) we obtain something meaningful. In his PhD thesis [Meh], yet to appear, Bernd Mehnert gives a stunning description in terms of Eisenstein series:
For $k \geq 1$ let

$$E_k(q) := \sum_{n=1}^{\infty} \sigma_{k-1}(n) q^n$$

be the k-th *Eisenstein series* deprived of its constant term, where $\sigma_i(n) = \sum_{1 \leq d | n} d^i$ is the i-th divisor sum. Note that we have defined the Eisenstein series both for odd and even k.
For a group $G = \prod_{i=1}^{l} (\mathbb{Z}/p^{e_i})^{r_i}$ in standard form (in particular, all e_i are mutually distinct) of order p^k, let

$$f_G(X_1, \ldots, X_k) := k! \prod_{i=1}^{l} \frac{X_{e_i}^{r_i}}{r_i!(e_i!)^{r_i}}$$

and

$$f_k(X_1, \ldots, X_k) := \sum_{G \text{ group of order } p^k} f_G(X_1, \ldots, X_k).$$

4.6.2 Theorem. *With the above notation, the k-th moment M_k of the local order of a random p-group is*

$$\sum_{n \geq 0} n^k \cdot \Pr(\mathrm{ord}_p(G) = n) = f_k(E_1, E_2, \ldots, E_k).$$

Proof. [Meh]. □

For example, $M_1 = E_1$, $M_2 = E_1^2 + E_2$, $M_3 = E_1^3 + 3E_1E_2 + E_3$, $M_4 = E_1^4 + 6E_1^2E_2 + 3E_2^2 + 4E_1E_3 + E_4$, and so on. Remarkably, we see that the local order of a random group has expected value E_1 and variance $M_2 - M_1^2 = E_2$.
Since this is the first time the result is published, let me list some computations. As formal power series, we get expected value

$$M_1 = E_1 = q + 2q^2 + 2q^3 + 3q^4 + 2q^5 + 4q^6 + \ldots,$$

variance

$$V = E_2 = q + 3q^2 + 4q^3 + 7q^4 + 6q^5 + 12q^6 + \dots,$$

and higher moments

$$\begin{aligned} M_2 &= q + 4q^2 + 8q^3 + 15q^4 + 20q^5 + 32q^6 + \dots \\ M_3 &= q + 8q^2 + 26q^3 + 63q^4 + 116q^5 + 208q^6 + \dots, \\ M_4 &= q + 16q^2 + 80q^3 + 255q^4 + 608q^5 + 1280q^6 + \dots, \end{aligned}$$

and so on.
Finally, I give a table giving (approximatively) expected value M_1, variance V, and higher moments M_2, M_3 and M_4 of the local order for various primes p. Recall that all values are simply obtained from the power series by plugging in $q = \frac{1}{p}$:

	$p=2$	$p=3$	$p=5$	$p=7$	$p=11$	$p=13$	$p=17$
M_1	1.6067	0.6822	0.3017	0.1909	0.1091	0.0898	0.0662
V	2.7440	0.9494	0.3660	0.2191	0.1192	0.0968	0.0701
M_2	5.3255	1.4148	0.4571	0.2556	0.1311	0.1048	0.0745
M_3	24.4734	3.9984	0.8848	0.4173	0.1817	0.1387	0.0926
M_4	145.5087	14.7677	2.2088	0.8596	0.3053	0.2189	0.1340

Recall that the local order is the p-logarithm of the usual order, so the trivial group has local order 0. This is why moments of less than 1 are possible.

4.6.2 Rank

4.6.3 Theorem. *The probability that a randomly chosen group has rank r is*

$$P(\mathrm{rk}(G) = r) = \left(\prod_{i=1}^{\infty}(1-q^i)\right) \frac{q^{r^2}}{\left(\prod_{i=1}^{r}(1-q^i)\right)^2}.$$

This formula was already contained in Cohen and Lenstra's original paper [CL84, Thm. 6.3], but was independently proven by Rudvalis and Shinoda [RS88]. Later on, a new proof by means of the cycle index was given by Fulman [Ful97, Thm. 15].
In fact, the theorems of Rudvalis and Shinoda look very different from the version given above. They make statements about the probability that a random matrix from $\mathrm{GL}(n,p)$ has a fixed space of dimension r. But it is easy to see (cf. [Ful97, Lemma 11]) that the dimension of the fixed space

of a matrix $M \in \mathrm{GL}(n,p)$ equals the rank of $\underline{\lambda}'_{X-1}$, i.e., the number of parts of the partition corresponding to the polynomial $X-1$ in the Jordan-Chevalley normal form. Since for $n \to \infty$ the distribution of this partition is given by the Cohen-Lenstra probability, the above theorem is equivalent to the following corollary, and this is the form in which Rudvalis/Shinoda and Fulton have given their theorems:

4.6.4 Corollary. *The probability that a randomly chosen matrix in* $\mathrm{GL}(n,p)$ *has a fixed space of dimension* r *approaches, as* $n \to \infty$*,*

$$\left(\prod_{i=1}^{\infty}(1-p^{-i})\right)\frac{p^{-r^2}}{\left(\prod_{i=1}^{r}(1-p^{-i})\right)^2}.$$

Washington, who is clearly in the number theory fraction, published this as a remarkable observation [Was86], but he did not deduce the general theorem 4.2.1. Also, no immediate reason for this coincidence is known (or for the general agreement between the Cohen-Lenstra probability and the probability of partitions appearing in the Jordan-Chevalley normal form), although this might be simply due to lack of research.

Higher moments of the rank

A closed formula for the higher moments of the rank of a random group is not known. However, if we consider the quantity $p^{\mathrm{rk}(G)}$ instead of $\mathrm{rk}(G)$, then more can be said. Cohen and Martinet [CM87, (1.1)(d)] give the following formula for its higher moments:

4.6.5 Theorem. *The* k*-th moment of* $p^{\mathrm{rk}(G)}$ *is (with* $q = \frac{1}{p}$*)*

$$\sum_{r\geq 0} p^{kr}\cdot P(\mathrm{rk}(G)=r) = \sum_{i=0}^{k}\left(q^{-i(k-i)}\frac{\prod_{j=1}^{k}(1-q^j)}{\left(\prod_{j=1}^{i}(1-q^j)\right)\left(\prod_{j=1}^{k-i}(1-q^j)\right)}\right).$$

The same formula was independently proven by Fulman [Ful97, Thm. 18,19]. He also pointed out that the summands may be interpreted as the q-analogue $S_q(k,i)$ of the Stirling numbers of second kind (cf. [BDS94]).

4.6.3 Rank and order combined

4.6.6 Theorem. *The probability that a finite abelian* p*-group has order* p^n *and rank* r *is*

$$P\begin{pmatrix}\operatorname{ord}(G)=n,\\ \operatorname{rk}(G)=r\end{pmatrix}=\left(\prod_{i=1}^{\infty}(1-q^i)\right)\frac{q^{n-r}\prod_{i=1}^{n-1}(1-q^i)}{|\mathrm{GL}(r,p)|\left(\prod_{i=1}^{r-1}(1-q^i)\right)\left(\prod_{i=1}^{n-r}(1-q^i)\right)}.$$

This theorem seems to be missing in the number theory community. It was proven by Fulman [Ful97, Thm. 16] using the cycle index.

4.6.4 Exponent

4.6.7 Theorem. *The probability that a random group has (p-adic) exponent at most e is*

$$P(\exp G\leq e)=\prod_{\substack{i=1\\ i\equiv 0,\pm(e+1) \bmod (2e+3)}}^{\infty}(1-q^i),$$

where the index runs through all positive integers that satisfy one of the congruences.
This theorem was first proven by Cohen [Coh85] and was independently rediscovered by Fulman [Ful97, Thm. 21] via his Young Tableau Algorithm. A different and very simple proof is given in [Len08] by means of CL-maps (cf. corollary 3.4.2 in this thesis).
All proof methods involve the generalized Ramanujan-Rogers identities [And76, Thm. 7.5]. The case $e=1$ occurred already in [CL84] and involves the original Ramanujan-Rogers identity.

4.6.5 u-probabilities

4.6.8 Definition. *Let u be a positive integer and G a finite abelian p-group. The* u-probability of G, *denoted by $P_u(G)$, is the probability that G is obtained by the following random process:*

(i) Choose randomly a p-group H with respect to the Cohen-Lenstra probability.

(ii) Choose u elements $g_1,\dots,g_u$ uniformly at random.

(iii) Output $H/\langle g_1,\dots,g_u\rangle$.

Here, $\langle g_1, \ldots, g_u \rangle$ denotes the subgroup generated by $g_1, \ldots, g_u$.

The u-probabilities are important for studying class groups of number fields (cf. section 6.1.2). They have extensively been studied by Cohen and Lenstra [CL84] and others. By means of ζ-functions, Cohen and Lenstra derived the following explicit formula:

4.6.9 Theorem. *Let $u > 0$ be an integer, and let G be a finite abelian p-group of order n. Then*

$$\begin{aligned} P_u(G) &= \frac{1}{n^u \prod_{i=1}^{u}(1-p^{-i})} P(G) \\ &= n^{-u} \frac{1}{\#\mathrm{Aut}(G)} \prod_{i=u+1}^{\infty} (1-p^{-i}). \end{aligned}$$

Proof. [CL84, Example 5.9] □

In the same paper, you can find explicit formulas for the u-probability that a p-group is of a certain order or certain rank, is cyclic, is elementary, and formulas for the expected values of the size of a group and the number of elements with given annihilator [CL84, examples 5.8–5.13, theorem 6.3].

Chapter 5

Global Theory

We have seen how the Cohen-Lenstra principle leads to a probability distribution on the set of (isomorphism classes of) all finite abelian p-groups, for arbitrary $p \in \mathbb{P}$. However, being a p-group is a restriction we would like to remove. Often we deal with non-primary groups, e.g., the class group of a number field (section 6.1) or the Jacobian of a hyperelliptic curve (section 6.3).
But when we try to transfer the techniques for p-groups to non-primary groups, we face a severe problem. Recall that we introduced the Cohen-Lenstra distribution by defining the weight of an atomic event $\{G\}$ to be proportional to $|\mathrm{Aut}(G)|^{-1}$. For this approach it is crucial that the measure is finite: $\sum_G |\mathrm{Aut}(G)|^{-1} < \infty$. We have seen in theorem 2.1.2 that this is the case if G runs over all p-groups for some $p \in \mathbb{P}$. Now what happens if G runs over all finite abelian groups? This clearly includes all groups of the form $\mathbb{Z}/p\mathbb{Z}$, where p runs over all primes. Hence,

$$\sum_{G \in \mathcal{G}} \frac{1}{|\mathrm{Aut}(G)|} \geq \sum_{p \in \mathbb{P}} \frac{1}{|\mathrm{Aut}(\mathbb{Z}/p\mathbb{Z})|} = \sum_{p \in \mathbb{P}} \frac{1}{p-1} = \infty.$$

So we see that we cannot transfer the approach for p-groups to arbitrary finite abelian groups. On the other hand, by product formulas (multiplying up all local probabilities) it is pretty clear what "measure" many sets should have. But this will not lead to a probability measure.
In this chapter we will discuss two different ways – their benefits and drawbacks – still to assign probabilities to certain events.
The first one invents a notion of restricted countable additivity, where we require that all the sets involved are measurable. However, we will see that this approach must necessarily fail to measure the most important global quantities. Still it is better than the approach that is undertaken in most current research papers (cf. section 5.1.1).

The second approach imitates the definition of the Lebesgue measure on $\mathbb{R}^n$. In real analysis, constructions like the Banach-Tarski-paradox (originally in [BT24]; for a more recent treatment see [Wag93]) show that there is no equivariant measure on the power set of $\mathbb{R}^n$. The solution is to designate only some σ-algebra of sets as *measurable* and to define the measure only on those. We copy this approach by defining *uniform properties* and designating these as a basis for the σ-algebra. In my eyes, this is a quite satisfactory solution.

Both approaches may be combined.

Actually, there is a third way, using densities. This is perhaps the most pragmatical way, and it is almost the unique way found in current research papers. However, this approach has severe theoretical and practical drawbacks, which will be discussed in section 5.1.1.

Let me finish the introduction with some remarks about terminology: Cohen and Lenstra speak of probabilities, although they are only talking about contents (cf. def. 5.1.1 below), and they are well aware of this terminological slackness. I will not use the term "probability" in a context where we do not have a probability measure – therefore, my terminology is different from the one of Cohen and Lenstra. When I talk about their concept, I use the words "content" or "density". Further, I use the word "heuristic" to refer to any one of the above concepts, so a "heuristic" is not a precise mathematical concept.

It would be very convenient to write down a definition of "the" Cohen-Lenstra content. Unfortunately, such a definition does not exist. (This is one of the circumstances that necessitate this chapter!) Rather, the precise definitions in the literature (which still include unprecise terms like "reasonable functions") work with the concept of densities (section 5.1.1) and always depend on the specific application. For different applications, one gets different densities: they differ in the set of "measurable" sets, but even if one set is assigned a content in several settings, these contents need not agree. These problems are discussed in more detail in section 5.1.1. I will define a global content in definition 5.1.3 as my personal proposal of a theoretic sound content, but you should be aware that in the literature there is no agreement on what a "Cohen-Lenstra content" should be (at least if you want it to be independent of the specific application).

Opposed to that, when I talk about "the global Cohen-Lenstra measure" or about "the global Cohen-Lenstra probability", I mean the probability measure that I define in 5.2.5. Its existence is the central insight of this chapter, and section 5.3 is devoted to studying this measure.

5.1 Global contents

Before we start, let me repeat some basic notions from measure theory which we will use throughout the chapter.

5.1.1 Definition. *An* algebra *of sets over some set X is a set $\mathcal{A}$ of subsets of X that is closed under complements, finite unions and finite intersections, and with $\emptyset \in \mathcal{A}$.*
A σ-algebra *is an algebra that is also closed under countable unions and intersections. We usually denote σ-algebras by Σ.*
A content *on an algebra $\mathcal{A}$ is a map $\mu : \mathcal{A} \to \mathbb{R} \cup \{\infty\}$ such that*

- $\mu(\emptyset) = 0$.
- $\mu(A) \geq 0$ *for all* $A \in \mathcal{A}$.
- $\mu(A_1 \cup A_2) = \mu(A_1) + \mu(A_2)$ *for all disjoint* $A_1, A_2 \in \mathcal{A}$.

We will usually further assume that $\mu(X) = 1$.
A content that is defined on a σ-algebra is called a measure *if it is furthermore countably additive. If $\mu(X) = 1$, it is called a* probability measure.

5.1.2 Remark. In the literature, contents are more often referred to as *finite additive measures*. I have not adopted this notion because it suggests that finite additive measure are measures, which is not true in general.

Before coming to the different methods of defining contents, let me first illustrate the problems we face when we try to define a global probability measure. So assume we are more ambitious and want to construct a measure instead of a content.
What properties should a global probability measure have? Note that for any p, there is a natural projection $\mathcal{G} \xrightarrow{\pi_p} \mathcal{G}_p$. We would like our probability to be compatible with these maps, i.e., for any $M \subseteq \mathcal{G}_p$ we would like to have $P(\pi_p^{-1}(M)) = P_p(M)$. ($P_p$ is the local Cohen-Lenstra probability on $\mathcal{G}_p$.)
Moreover, the p-parts of each group should be independent (as the automorphism group of a group decomposes into a direct product of the automorphism groups of its p-parts, see section 3.5.1), i.e., for finitely many mutually distinct primes $p_1, \ldots, p_k$ and sets $M_i \subseteq \mathcal{G}_{p_i}$, $1 \leq i \leq k$ we require $\bigcap_i P(\pi_{p_i}^{-1}(M_i) = \prod_i P(M_i)$.
So the first attempt would be to define Σ as the coarsest σ-algebra that contains all $\pi_p^{-1}(M)$ for all primes p and $M \subseteq \mathcal{G}_p$, and to define the probabilities via the product formula.
Unfortunately, this does not lead to a measure: Obviously we can describe every group $G \in \mathcal{G}$ be specifying each of its p-parts. Since a measure is

defined on a σ-algebra, the set $\{G\}$ would be measurable as a countable intersection of measurable sets, and by an easy calculation it would have measure 0. But since $\mathcal{G}$ is countable, we would get the contradiction

$$1 = P(\mathcal{G}) = P\left(\bigcup_{G\in\mathcal{G}} \{G\}\right) = \sum_{G\in\mathcal{G}} P(\{G\}) = 0.$$

Note that this argument shows that Σ is the whole power set of $\mathcal{G}$.
We see that it is difficult to find a measure which is compatible with the local Cohen-Lenstra measures, although we will finally succeed in section 5.3. Before I come to this measure, let us discuss the alternatives. In the following sections, I illustrate several ways of defining contents instead of measures. However, we will also find that all these methods have severe drawbacks.

5.1.1 Densities

Cohen and Lenstra tried to avoid the problems illustrated above in the following way: They were interested in a very concrete sequence of finite abelian groups (the sequence of odd parts of class groups of imaginary quadratic number fields, see section 6.1.1). For us, the concrete sequence is of no importance, so let (G_n) be a sequence of finite abelian groups. Let D be the set of all subsets $S \subseteq \mathcal{G}$ which have a density in (G_n), i.e., all S for which the limit

$$\lim_{n\to\infty} \frac{\#\{k \leq n \mid G_k \in S\}}{n}$$

exists. Then D is an algebra of sets, and the limits define a content on D. This approach is copied by almost all currently active researchers. It has the philosophical drawback that we cannot speak of probabilities, and the practical drawback that we usually do not know D. Furthermore, it is at least annoying that we do not have countable additivity. But there are also much more severe obstacles.
Of course, we want to decide whether a sequence is compatible with the (local) Cohen-Lenstra distributions. But how do we decide this? In principle we would like D to be "reasonably" rich, and that the densities of sets $S \in D$ are compatible with the Cohen-Lenstra heuristic.
But what does "compatible" really mean? Often, researchers are only concerned with very special sets S, in particular sets that are direct products $\prod_{p\in\mathbb{P}} S_p$, for sets $S_p \in \mathcal{G}_p$. Then they declare the Cohen-Lenstra probability

to be $\prod_{p\in\mathbb{P}} P_p(S_p)$. This sounds quite reasonable, but in this way there is no hope whatsoever to gain countable additivity, as is proven in section 5.1.3.
If we are given such a direct product set S, are there other ways to define a "Cohen-Lenstra probability" for S? The answer is yes! We have two different limit processes going on: One in the definition of the local Cohen-Lenstra probability, where we average over all p-groups. And another one when we multiply the probabilities for various primes. Assigning the probabilities $\prod_{p\in\mathbb{P}} P_p(S_p)$ to a set S as above imposes an order on the limit process. Moreover, by what we have already shown, the limits do not commute! So we might with equal legitimation compute the double limit in a different way, and obtain a different "Cohen-Lenstra probability" for the same set S. This is highly unsatisfactory.
Another point is that for every sequence (G_n) we get a different content. Even if we would accept the order of the limit process for special sets $S = \prod_{p\in\mathbb{P}} S_p$, then it is not clear at all how to extend this to the whole power set of $\mathcal{G}$. For a set S which does not happen to be a direct product, there are many ways that lead to different contents for S, and we do not have a canonical way of choosing the "right" one. Thus for each sequence of groups, we would have to figure out the sets with densities and make up a new content on these sets. For different sequences of groups, the contents would in general not be compatible.
A related approach, which appears to be a bit less critical, is to define a content $P(S)$ for any set $S \subseteq \mathcal{G}$ for which the following limit exists:

$$P(S) := \lim_{x\to\infty} \frac{\sum_{G\in S,|G|<x} w(G)}{\sum_{|G|<x} w(G)}.$$

This yields a content. Basically, the approach imposes an ordering onto $\mathcal{G}$, namely by their size, and then sums up over all groups up to a certain threshold. This sounds very natural, but still it is a specific ordering. It corresponds to taking the density with respect to the sequence where the group of order 1 appears an appropriate number of times, then the group of order 2 appears, and so on. This analogy is not perfect, because it is only possible to construct the sequence for every finite start sequence $\{G \in \mathcal{G} \mid \mathrm{ord}(G) \leq x\}$ of the ordering. (For extending the sequence, we need to adjust the number of order-1-groups, order-2-groups, ... in order to get an integral number of appearances.) Nevertheless, in my eyes the analogy catches the essential point: There is no real reason to impose this specific ordering on $\mathcal{G}$, and it is not clear why a truly random sequence should respect this specific ordering.
Furthermore, is the ordering above really the most natural ordering? Or

would it perhaps be more natural to order the groups by their weight? This would give a different content, and so it would be a matter of taste which content one prefers. We see that this situation is quite unsatisfactory.
Finally, by this approach, we do not have any hope to get a measure. Clearly, every one-element set $S = \{G\}$ is measurable with measure 0, which already rules out countable additivity. We cannot even hope for restricted countable additivity (cf. section 5.1.2 below), since we still have the same problem $\mathcal{G} = \dot{\cup}_{G\in\mathcal{G}}\{G\}$, but $P(\mathcal{G}) \neq \sum_{G\in\mathcal{G}} P(\{G\})$.
Summarizing, the illustrated approaches only postpone the problems – the reason why they have worked so far is that only a very limited type of sets S has been investigated, and that often the researcher concentrates on only one specific sequence of groups and does not care about other sequences. Cohen and Lenstra were well aware of the problem (that is why they did not specify what a "reasonable function" [CL84, 8.1] should be), but apparently they saw no way to avoid it.

5.1.2 Restricted countability

We have seen that we need a general notion for sequences of groups to be "compatible" with the Cohen-Lenstra heuristic. Let us first try to define a content that does not depend on the specific approach. In order to be compatible with the local Cohen-Lenstra measures, we want the algebra of sets to contain all sets of the form $\pi_p^{-1}(M)$, where $p \in \mathbb{P}$ and $M \subseteq \mathcal{G}_p$. This leads to the following definition:

5.1.3 Definition. *Let $\mathcal{A}$ be the algebra of all subsets S of $\mathcal{G}$ for which there exists a finite index set $I \subset \mathbb{P}$ and a set $S_I \subseteq \prod_{p\in I} \mathcal{G}_p$ such that*

$$S = S_I \times \bigoplus_{p\in\mathbb{P}\setminus I} \mathcal{G}_p. \tag{5.1}$$

Informally speaking, S is only specified at finitely many local places.
We define the (global) Cohen-Lenstra content *P on $\mathcal{A}$ via*

$$P(S_I \times \bigoplus_{p\in\mathbb{P}\setminus I} \mathcal{G}_p) := \sum_{G\in S_I} \prod_{p\in I} P_p(G_p),$$

where G_p denotes the p-part of G.
I usually omit the attribute "global" if no confusion is possible and talk only of the Cohen-Lenstra content *on $\mathcal{G}$.*

5.1.4 Remark. In the definition above, the symbol "$\bigoplus$" denotes the outer direct sum, by which I simply mean for any index set $I \subseteq \mathbb{P}$:

$$\bigoplus_{p\in I} \mathcal{G}_p := \{(G_p)_{p\in I} \in \prod_{p\in I} \mathcal{G}_p \mid \text{ almost all } G_p \text{ are } 0\}.$$

So in particular,

$$\bigoplus_{p\in \mathbb{P}} \mathcal{G}_p \xrightarrow{\cong} \mathcal{G}.$$

5.1.5 Theorem. *The global Cohen-Lenstra content is a well-defined content.*

Proof. It is clear that $\mathcal{A}$ is an algebra of sets.
By measure theory we know that for any finite I we can endow $\prod_{p\in I} \mathcal{G}_p$ with a probability measure by defining $P(\{G\}) := \prod_{p\in I} P_p(G_p)$. (Note that this does not work for infinite I because $\mathcal{G}$ is not the product space but rather the direct sum of the $\mathcal{G}_p$ – only for finite I do $\prod_{p\in I} \mathcal{G}_p$ and $\bigoplus_{p\in I} \mathcal{G}_p$ agree.) Since any complement and any finite union or finite intersection of sets in $\mathcal{A}$ is only specified on a finite set S, we can restrict ourselves to a probability space of this kind. So we may restrict ourselves to the power set of $\prod_{i\in F} \mathcal{G}_p$, where F is some *finite* set of primes. But the finite product of probability spaces is again a probability space, so all formulas then become evident. □

The algebra $\mathcal{A}$ has a remarkable property: Whenever a countable disjoint union of sets $A_i \in \mathcal{A}$ is again an element of $\mathcal{A}$, then everything takes place only on finitely many primes, and therefore we have countable additivity:

$$P\left(\dot{\bigcup_i} A_i\right) = \sum_i P(A_i).$$

So in other words, we have countable additivity *provided that the union is measurable*. A similar statement holds for intersection. We will say that such a content has *restricted countable additivity*. This is not quite a probability measure, but it might be satisfactory. However, the algebra $\mathcal{A}$ is still too coarse to measure interesting quantities. So one approach would be to refine $\mathcal{A}$ and still keep the restricted countable additivity. Unfortunately, we will see in the next section that this approach is necessarily of limited success.

5.1.3 Global quantities

What kind of statements would we like to make about groups? We have already seen that we cannot measure all sets of groups. But there are some minimal requirements – at least to my feeling we should be able to measure the three most important quantities of a finite abelian group: its order, rank

and exponent. So we definitely want the following sets to be measurable for any n:

- $\{G \in \mathcal{G} \mid \mathrm{ord}(G) = n\}$.
- $\{G \in \mathcal{G} \mid \mathrm{rk}(G) = n\}$.
- $\{G \in \mathcal{G} \mid \exp(G) = n\}$.

Unfortunately, this is impossible. We will prove:

5.1.6 Theorem. *There is no algebra $\mathcal{A}$ on $\mathcal{G}$ with a content P with restricted countable additivity (i.e., countable additivity on measurable sets) that is compatible with the Cohen-Lenstra heuristic induced by the projections $\mathcal{G} \to \mathcal{G}_p$ such that order or exponent are measurable.*

Note that we implicitly assume that distinct primes are independent of each other. This is an assumption which is usually made whenever people work with the Cohen-Lenstra philosophy.

Proof. We will only show the statement for the measurability of the order. The statement for the exponent can be proven analogously.
Assume such an algebra and content exist. Then for all $n \in \mathbb{N}$, we can measure the set $S_n := \{G \in \mathcal{G} \mid \mathrm{ord}(G) = n\}$. We fix an n and define $I_n := \{p \in \mathbb{P} \mid p \nmid n\}$ and $T_p := \{G \in \mathcal{G} \mid \pi_p(G) = 0\}$ for all $p \in I_n$. Then T_p is measurable with measure $P(T_p) = P_p(\{0\}) = \prod_{i=1}^{\infty}(1 - p^{-i}) \leq 1 - \frac{1}{p}$. Since $S_n \subseteq T_p$ for all $p \in I_n$, we have for any finite subset F of I_n

$$S_n \subseteq \bigcap_{p \in F} T_p.$$

Since F is finite, both sides are measurable and by independence of distinct primes we obtain

$$P(S_n) \leq \prod_{p \in F} P(T_p).$$

The above inequality is true for any finite set $F \subset I_n$, so we may replace the

right hand side by the infimum over all such F:

$$\begin{aligned} P(S_n) &\leq \inf_{F \subset I_n \text{ finite}} \prod_{p \in F} P(T_p) \\ &= \prod_{p \in I_n} P(T_p) \\ &\leq \prod_{p \in I_n} (1 - \frac{1}{p}) \\ &\leq \exp \underbrace{\left(\sum_{p \in I_n} (-\frac{1}{p}) \right)}_{=-\infty} \\ &= 0. \end{aligned}$$

Therefore, $P(S_n) = 0$ for all $n \in \mathbb{N}$. But $\mathcal{G} = \bigcup_{n \in \mathbb{N}} S_n$, which would imply $P(\mathcal{G}) = 0$, a contradiction.

□

You may wonder why the theorem above only refers to the order and the exponent, but not to the rank. Surprisingly, it turns out that it is even possible to endow $\mathcal{G}$ with a *probability measure* compatible with the rank. The reason why the rank behaves differently is that it is a uniform quantity in the following sense: If you require the rank of a group $G \in \mathcal{G}$ to be k, then the information that you can extract about the local ranks r_p of G_p is independent of p. This seems to be a rather complicated way of saying that essentially the only thing we know for a fixed p is $r_p \leq r$. However, going through the proof of the theorem, this was the crucial point that forbade countable additivity for the order (and the exponent): If we know the order of a group, then we can compute the order of G_p for any $p \in \mathbb{P}$, so we get *individual* information about local quantities.
This leads us to the definition of the *uniform order* and the *uniform exponent*, which turn out to be better suited for the situation. Afterwards, we will define the notion of *uniform properties* in general.

5.2 Uniform properties

Since we have noticed that the rank behaves better than order and exponent, we want to catch the local behaviour of the rank and transfer it to order and exponent as follows:

5.2.1 Definition. *For a prime p, we define the* local order *on $\mathcal{G}_p$ as*

$$\mathrm{ord}_p(G) := \log_p(\mathrm{ord}(G)).$$

We have already defined the local exponent *on $\mathcal{G}_p$ as*

$$\exp_p(G) := \exp(G) = \log_p(\min\{n \in \mathbb{N}^+ \mid n \textit{ annihilates } G\}).$$

Now we define the uniform order ord_{uni} *on $\mathcal{G}$ and the* uniform exponent $\exp_{uni}$ *on $\mathcal{G}$ as*

$$\begin{aligned} \mathrm{ord}_{uni}(G) &:= \max_{p\in\mathbb{P}} \mathrm{ord}_p(G_p) \\ \exp_{uni}(G) &:= \max_{p\in\mathbb{P}} \exp_p(G_p) \end{aligned}$$

Note that this definition is completely analogous to the formula

$$\mathrm{rk}(G) = \max_{p\in\mathbb{P}} \mathrm{rk}(G_p)$$

for the rank. Therefore the "uniform rank" coincides with the ordinary rank. As we will show later, it turns out that there is a probability measure on $\mathcal{G}$ which allows to measure the uniform order, rank and exponent. So at least we can obtain the minimal program formulated in section 5.1.3, if we work with uniform quantities. But in fact, we can show much more. For this we need a general notion of uniform quantities. For the moment, we restrict ourselves to properties, i.e., to functions $\mathcal{G} \to \{0,1\}$, telling whether a group has a certain property or not.

5.2.2 Definition. *A* property *(on $\mathcal{G}$) is a function $E : \mathcal{G} \to \{0,1\}$. For properties E_1, E_2 we define $E_1 \vee E_2$ and $E_1 \wedge E_2$ by*

$$\begin{aligned} (E_1 \vee E_2)(G) &= \begin{cases} 1 & \textit{if } E_1(G) = 1 \textit{ or } E_2(G) = 1, \\ 0 & \textit{otherwise,} \end{cases} \\ (E_1 \wedge E_2)(G) &= \begin{cases} 1 & \textit{if } E_1(G) = 1 \textit{ and } E_2(G) = 1, \\ 0 & \textit{otherwise,} \end{cases} \end{aligned}$$

respectively.
A property E is called uniform *if there is a function, which by abuse of notation we also call E, from $\mathcal{G}_{\mathcal{P}}$ to $\{0,1\}$ such that for all $G \in \mathcal{G}$*

$$E(G) = 1 \textit{ if and only if } E(\underline{n}_p) = 1 \textit{ for all } p \in \mathbb{P},$$

where $\underline{n}_p \in \mathcal{G}_\mathcal{P}$ is the partition that corresponds to the p-part G_p of G via the identification $\mathcal{G}_p \overset{\cong}{\longleftrightarrow} \mathcal{G}_\mathcal{P}$.
If we want to distinguish explicitly between the two functions E, then we write $E_\mathcal{G}$ and $E_{\mathcal{G}_\mathcal{P}}$, respectively.
Finally, for a uniform property E we define $O(E) := E_\mathcal{G}^{-1}(\{1\})$.

5.2.3 Remark.

- Note that for every p we have used the correspondence $\mathcal{G}_p \overset{\cong}{\longleftrightarrow} \mathcal{G}_\mathcal{P}$. This is our way to "compare" the groups $\mathcal{G}_p$ for various p, which is necessary for defining uniform properties.

  I feel quite confident that uniform properties are the "right" concept, for the following reason: It has proven very convenient not to work with the number p^{-1} but rather with the formal variable q. But this replacement is reflected by the identification between $\mathcal{G}_p$ and $\mathcal{G}_\mathcal{P}$ and indicates that we should rather work with the latter object. If the identification is appropriate in this case then it makes sense to identify the copies of $\mathcal{G}_\mathcal{P}$ that correspond to various p.

  Recall that this correspondence is almost canonical: The only non-trivial, order-preserving automorphism of $\mathcal{G}_\mathcal{P}$ is conjugation (theorem 3.5.2). Although this gives us – in principle – two ways of identifying $\mathcal{G}_p$ with $\mathcal{G}_\mathcal{P}$, the identification $\mathcal{G}_{p_1} \cong \mathcal{G}_{p_2}$ for $p_1, p_2 \in \mathbb{P}$ is canonical, due to the fact that the non-trivial automorphism of $\mathcal{P}$ is not compatible with the weight w.

- The question whether a group has rank r is a uniform property. Indeed a group has rank r if each p-part has rank $\leq r$ and if it is not true that each p-part has rank $\leq r-1$. Analogously, the uniform order and uniform exponent are uniform properties.

5.2.4 Remark.

- For all uniform properties E_1 and E_2, we have

$$O(E_1 \wedge E_2) = O(E_1) \cap O(E_2).$$

- In general, it is *not* true that $O(E_1 \vee E_2) = O(E_1) \cup O(E_2)$ for local properties E_1, E_2.

5.2.5 Definition. *Let $\Sigma_\mathcal{G}$ be the coarsest σ-algebra on $\mathcal{G}$ that contains the fibers $O(E)$ of all uniform properties E of $\mathcal{G}$. We define the* Cohen-Lenstra probability measure *$P_\mathcal{G}$ on $\Sigma_\mathcal{G}$ via:*

$$P_{\mathcal{G}}(E) := P_{\mathcal{G}}(E = 1) := P_{\mathcal{G}}(O(E)) := \prod_{p \in \mathbb{P}} P_p(E^{-1}_{\mathcal{G}_{\mathcal{P}}}(\{1\})), \qquad (5.2)$$

where P_p is the Cohen-Lenstra probability on $\mathcal{G}_{\mathcal{P}} \overset{\cong}{\longleftrightarrow} \mathcal{G}_p$. Be aware that for each p we have a different probability measure on $\mathcal{G}_{\mathcal{P}} = \mathcal{G}_p$. If no confusion with the local Cohen-Lenstra probability measures is possible then we omit the index and write P instead of $P_{\mathcal{G}}$.

The main result in this chapter is that $P_{\mathcal{G}}$ is indeed a probability measure on $\mathcal{G}$ that makes all uniform properties measurable. This justifies many calculations that researchers have carried out without specifying the probability space in which their calculations are supposed to happen. (Of course, the computations were usually carried out in terms of formal series, and the results are definitely true as identities of formal series. But in order to translate the results into probability statements, one needs to specify a probability space.) There are very few statements in the literature which are not uniform statements. There are only two wide-spread non-uniform examples I know of, both of them due to Cohen and Lenstra:
Firstly, they state that the "probability" of a one-element set $\{G_0\}$ is 0 for every $G_0 \in \mathcal{G}$ [CL84, §9,II]. However, it is obvious that this statement is not compatible with a probability measure, since that would mean that we have a countable probability space with probability 0 for each atomic event, which is impossible. (Cohen and Lenstra were well aware of the fact that this gives only a content instead of a measure.) Secondly, they state that the "probability" that a finite abelian group has p-part G_0, for a fixed p-group G_0, is $P_p(G_0)$. This is highly problematic. As we have seen before, there is no probability measure on $\mathcal{G}$ which is compatible with this statement, so we should at least avoid talking about probabilities in this context.
Now let us come to the main theorem:

5.2.6 Theorem. *The Cohen-Lenstra probability measure $P_{\mathcal{G}}$ is indeed a probability measure, and it makes all uniform properties measurable.*

The proof is complicated and the whole next section is devoted to it.
Before we come to the proof, let me first summarize our discussion about measurable functions: The theorem asserts that the rank, the uniform order, the uniform exponent and all other uniform properties are measurable, and so are all functions defined in these terms, for example, the expected value or higher moments of these functions.
Not measurable are the classical order and exponent, and the property that the p-part of a group is isomorphic to some fixed p-group G_0. But for any

single one of these properties we have shown (cf. theorem 5.1.6 and page 106, respectively) that there is no probability measure which would make these functions measurable, so we could not expect to be able to measure these functions. More generally, essentially no function is measurable that is defined via the p-part of the group, for some fixed p.

5.3 The existence of a global measure

In this section, we prove theorem 5.2.6. We proceed as follows: First, we construct an outer measure on the power set of $\mathcal{G}$ that coincides on certain key sets with our desired probability measure $P_{\mathcal{G}}$. Then we use the theorem of Carathéodory to deduce the existence of a σ-algebra of measurable sets such that the outer measure is a measure on these sets. Finally we show that uniform properties are measurable with respect to this σ-algebra.
Let me start with some general remarks. First of all, note that the product that defines $P_{\mathcal{G}}$ consists only of factors ≤ 1. Therefore, we either have absolute convergence or we have definite divergence to 0. In both cases, we may arbitrarily reorder the factors, and we may apply the formula

$$\prod_i a_i = \exp\left(\sum_i \log(a_i)\right).$$

Since this is a major tool for us, we will be concerned about estimating $\log(a_i)$. We will use the formula

$$-2h \leq \log(1-h) \leq -h,$$

which is true for any $0 \leq h \leq \frac{1}{2}$ (by Jensen's inequality) and in particular for $h = \frac{1}{p}$, for any prime p.

5.3.1 First properties of the global measure

This section contains essentially some technical lemmas about $P = P_{\mathcal{G}}$. However, lemma 5.3.3 is of intrinsic interest, independent of its use in the construction of the probability space.

So let us check a couple of properties of P. First of all, in definition 5.2.5 we have not excluded the case that $E(\underline{0}) = 0$, where $\underline{0}$ stands for the trivial partition. But in this case $O(E)$ is empty, since any group has trivial p-parts for almost all $p \in \mathbb{P}$. In other words, we have non-trivial ways to describe the empty set, so the formula in 5.2.5 had then better give $P(E) = P(\emptyset) = 0$, if it is supposed to make sense. Indeed this is the case:

5.3.1 Lemma. *If E is a uniform property with $E(\underline{0}) = 0$, then $P(E) = 0$.*

Proof. We have $E^{-1}_{\mathcal{G}_\mathcal{P}}(\{1\}) \subseteq \mathcal{G}_\mathcal{P} \setminus \{\underline{0}\}$, so we have

$$\begin{aligned} P_p(E^{-1}_{\mathcal{G}_\mathcal{P}}(\{1\})) &\leq P_p(\mathcal{G}_p \setminus \{\underline{0}\}) \\ &= 1 - P_p(\{\underline{0}\}) \\ &= 1 - \prod_{i=1}^{\infty}(1 - p^{-i}) \\ &\leq 1 - \left(1 - 2\sum_{i=1}^{\infty} p^{-i}\right) \\ &= 2\sum_{i=1}^{\infty} p^{-i} \\ &= \frac{2}{p-1}. \end{aligned}$$

Therefore,

$$\begin{aligned} P(E = 1) &= \prod_{p\in\mathbb{P}} P_p(E^{-1}_{\mathcal{G}_\mathcal{P}}(\{1\})) \\ &\leq \prod_{p\in\mathbb{P}} \frac{1}{p-1} \\ &= 0. \end{aligned}$$

□

So from now on we may assume that $E(\underline{0}) = 1$.
We continue with a lemma, which is of interest in its own right:

5.3.2 Lemma. *Let E be a uniform property with $E(\underline{1}) = 0$, where $\underline{1}$ is the unique partition of 1. Then $P(E) = 0$.*

Proof. We have $E^{-1}_{\mathcal{G}_\mathcal{P}}(\{1\}) \subseteq \mathcal{G}_\mathcal{P} \setminus \{\underline{1}\}$, so we get

$$\begin{aligned}
P_p(E^{-1}_{\mathcal{G}_\mathcal{P}}(\{1\})) &\leq P_p(\mathcal{G}_\mathcal{P} \setminus \{\underline{1}\}) \\
&= 1 - P_p(\{\underline{1}\}) \\
&= 1 - \frac{1}{p-1}\prod_{i=1}^{\infty}(1-p^{-i}) \\
&= 1 - p^{-1}\prod_{i=2}^{\infty}(1-p^{-i}) \\
&\leq 1 - p^{-1}\left(1 - 2\sum_{i=2}^{\infty}p^{-i}\right) \\
&= 1 - p^{-1} + \frac{2p^{-2}}{p-1} \\
&\overset{\text{for } p>2}{\leq} 1 - \frac{1}{2}p^{-1}.
\end{aligned}$$

Therefore,

$$\begin{aligned}
P(E=1) &= \prod_{p\in\mathbb{P}} P_p(E^{-1}_{\mathcal{G}_\mathcal{P}}(\{1\})) \\
&\leq \prod_{p\in\mathbb{P}\setminus\{2\}}\left(1-\frac{1}{2}p^{-1}\right) \\
&= \exp\left(\sum_{p\in\mathbb{P}\setminus\{2\}}\log\left(1-\frac{1}{2}p^{-1}\right)\right) \\
&\leq \exp\underbrace{\left(\sum_{p\in\mathbb{P}\setminus\{2\}}\left(-\frac{1}{2}p^{-1}\right)\right)}_{=-\infty} \\
&= 0.
\end{aligned}$$

□

In fact, we even have equivalence:

5.3.3 Lemma. *Let E be a uniform property. Then $P(E) > 0$ if and only if $E(\underline{0}) = E(\underline{1}) = 1$.*

Proof. We have already shown one direction, so now assume that the latter statement is true. Then we have $E^{-1}(\{1\}) \supseteq \{\underline{0}, \underline{1}\}$, so we get

$$\begin{aligned} P_p(E^{-1}(\{1\})) &\geq P_p(\{\underline{0}, \underline{1}\}) \\ &= \left(\sum_{i=0}^{\infty} p^{-i}\right) \prod_{i=1}^{\infty}(1-p^{-i}) \\ &= \frac{1}{1-p^{-1}} \prod_{i=1}^{\infty}(1-p^{-i}) \\ &= \prod_{i=2}^{\infty}(1-p^{-i}). \end{aligned}$$

Therefore,

$$\begin{aligned} P(E) &= \prod_{p\in\mathbb{P}} P_p(E^{-1}(\{1\})) \\ &\geq \prod_{p\in\mathbb{P}} \prod_{i=2}^{\infty}(1-p^{-i}) \\ &= \prod_{i=2}^{\infty} \prod_{p\in\mathbb{P}}(1-p^{-i}) \\ &= \prod_{i=2}^{\infty} \zeta^{-1}(i), \end{aligned}$$

where ζ denotes the Riemann ζ-function.
The latter product is well-known and converges against a positive constant $\approx 0.435757...$ (see e.g. [CL84, §7]).

□

5.3.2 The global outer measure

In order to define an outer measure, we first need to specify a family $\mathcal{D}$ of subsets with non-negative values ("Method I" in [Mun53]).

5.3.4 Definition. *Let $E_1, \ldots, E_r$ be uniform properties. In accordance with the former definition of $O(E)$ we define*

$$O(E_1, \ldots, E_r) := \bigcup_{i=1}^{r} E_i^{-1}(1),$$

and we set

$$\mathcal{O} := \{O(E_1, \ldots, E_r) \mid r \geq 0, E_1, \ldots, E_r \text{ uniform properties}\}.$$

Let $E_1, \ldots, E_r, F_1, \ldots, F_s$ *be uniform properties. Then we define*

$$D(E_1, \ldots, E_r; F_1, \ldots, F_s) := O(E_1, \ldots, E_r) \setminus O(F_1, \ldots, F_s),$$

and we set

$$\begin{aligned} \mathcal{D} \;:=\; & \{D(E_1, \ldots, E_r; F_1, \ldots, F_s) \mid r, s \geq 0, \\ & E_1, \ldots, E_r, F_1, \ldots, F_s \text{ uniform properties}\}. \end{aligned}$$

By slight abuse of notation, I will sometimes write $O(\mathcal{E})$ *and* $D(\mathcal{E}; \mathcal{F})$ *instead of* $O(E_1, \ldots, E_r)$ *and* $D(E_1, \ldots, E_r; F_1, \ldots, F_s)$*, respectively, where* $\mathcal{E}$ *and* $\mathcal{F}$ *are the families* $\{E_1, \ldots, E_r\}$ *and* $\{F_1, \ldots, F_s\}$*.*
By even stronger abuse of notation, I will occasionally write $O(E_i)$ *and* $D(E_i; F_j)$ *in these cases.*

5.3.5 Remark.

- The notion $\mathcal{D}$ conflicts with our notion for derivations in chapter 3. Since no derivations are used in this section, no confusion can arise.
- $\mathcal{O}$ is embedded into $\mathcal{D}$ by setting $s := 0$.
- For all uniform properties $E_1, \ldots, E_r$ and $E'_1, \ldots, E'_s$:

$$\begin{aligned} O(E_1, \ldots, E_r) \cap O(E'_1, \ldots, E'_s) &= O(E_1 \wedge E'_1, E_1 \wedge E'_2, \ldots, E_r \wedge E'_s). \\ O(E_1, \ldots, E_r) \cup O(E'_1, \ldots, E'_s) &= O(E_1, \ldots, E_r, E'_1, \ldots, E'_s). \end{aligned}$$

- For all uniform properties $E_1, \ldots, E_r, F_1, \ldots, F_s$ and $E'_1, \ldots, E'_t$:

$$\begin{aligned} D(E_1, \ldots, E_r; F_1, \ldots, F_s) \cap O(E'_1, \ldots, E'_t) \\ = D(E_1 \wedge E'_1, \ldots, E_r \wedge E'_t; F_1, \ldots, F_s). \end{aligned}$$

  Caution: No similar formula for the union exists.

- $\mathcal{D}$ is closed under intersection. More precisely, we have

$$\begin{aligned} & D(E_1, \ldots, E_{r_1}; F_1, \ldots, F_{s_1}) \cap D(\tilde{E}_1, \ldots, \tilde{E}_{r_2}; \tilde{F}_1, \ldots, \tilde{F}_{s_2}) \\ & = D(E_1 \wedge \tilde{E}_1, E_1 \wedge \tilde{E}_2, \ldots, E_{r_1} \wedge \tilde{E}_{r_2}; F_1, \ldots, F_{s_1}, \tilde{F}_1, \ldots, \tilde{F}_{s_2}) \end{aligned}$$

- $\mathcal{D}$ is *not* closed under union!
- $\mathcal{D}$ is *not* closed under set difference!

We would like to exend the definition of the function P from single uniform properties to the whole set $\mathcal{D}$. In order to do so, we need one more remark:

5.3.6 Remark.

- We may always assume that the defining uniform properties $E_1, \ldots, E_r$, $F_1, \ldots, F_s$ of a set $D(E_i, F_j) \in \mathcal{D}$ satisfy the condition

$$O(F_1, \ldots, F_s) \subseteq O(E_1, \ldots, E_r).$$

  In fact, if E_i and F_j are uniform properties which do *not* satisfy this condition, then we may use the identity

$$D(E_1, \ldots, E_r; F_1, \ldots, F_s) = D(E_1, \ldots, E_r; F_1 \wedge E_1, F_1 \wedge E_2, \ldots, F_s \wedge E_r)$$

  to enforce the condition.

- It is easy to see that $O(F_1, \ldots, F_s) \subseteq O(E_1, \ldots, E_r)$ if and only if for any $i \in \{1, \ldots, s\}$ there is a $j \in \{1, \ldots, r\}$ such that $O(F_i) \subseteq O(E_j)$.

5.3.7 Definition/Proposition. *We extend P to $\mathcal{D}$ as follows: We have already defined $P(O(E))$ for a single uniform property E in definition 5.2.5. Because of the formula $O(E_1) \cap O(E_2) = O(E_1 \wedge E_2)$ the function P is also defined on intersections of sets in $\mathcal{O}$. Hence we may extend P to sets of the form $O(E_1, E_2)$ $(= O(E_1) \cup O(E_2))$ via*

$$P(O(E_1, E_2)) := P(O(E_1)) + P(O(E_1)) - P(O(E_1 \wedge E_2)).$$

Continuing inductively, we extend P on the set $\mathcal{O}$. Finally, for uniform properties $E_1, \ldots, E_r, F_1, \ldots, F_s$ with $O(F_1, \ldots, F_s) \subseteq O(E_1, \ldots, E_r)$ we set

$$P(D(E_1, \ldots, E_r; F_1, \ldots, F_s)) := P(O(E_1, \ldots, E_r)) - P(O(F_1, \ldots, F_s)).$$

This yields a well-defined map $P : \mathcal{D} \to [0, 1]$.

Proof. The procedure for computing $P(O(E_1, \ldots, E_r))$ yields the Inclusion-Exclusion formula, which is independent of the order of the E_i. So we only need to show that whenever

$$D(E_1, \ldots, E_{r_1}; F_1, \ldots, F_{s_1}) = D(E'_1, \ldots, E'_{r_2}; F'_1, \ldots, F'_{s_2}) \tag{5.3}$$

then the value of P coincides for both sets.
Let us first consider the case that $O(E_1, \ldots, E_{r_1}) = O(E'_1, \ldots, E'_{r_2})$. If for some i, j we have $O(E_i) \subseteq O(E_j)$, then the result of the Inclusion-Exclusion

formula does not change if we omit E_i. So we may assume that all E_i are maximal in the sense that $O(E_i)$ is not a proper subset of $O(E_j)$, for all $j \neq i$. We assume the same for the E'_i. Then I claim that E_1 occurs also on the right hand side. By symmetry, this will imply the statement for $\mathcal{O}$.

Because of the maximality of E_1, it suffices to show that $O(E_1) \subseteq O(E'_i)$ for some i. (Then by symmetry, $O(E'_i) \subseteq O(E_j)$ for some j, and by maximality of E_1 we conclude $j = 1$ and $O(E_1) = O(E'_i)$). Assume not. Then for all $1 \leq i \leq r_2$ there is a partition $\underline{n}_i$ such that $E_1(\underline{n}_i) = 1$ and $E'_i(\underline{n}_i) = 0$. Now take r_2 distinct primes $p_1, \ldots, p_{r_2}$ and consider a group with p_i-part equal to $\underline{n}_i$, for $i = 1, \ldots, r_2$. Then this group is contained in $O(E_1)$ but in none of the $O(E'_i)$, contradicting $O(E_1, \ldots, E_{r_1}) = O(E'_1, \ldots, E'_{r_2})$.

This finishes our proof for $\mathcal{O}$. For $\mathcal{D}$, first notice that by the preceding remark, P is indeed defined on the whole set $\mathcal{D}$. To show that it is well-defined we use essentially the same argument as for $\mathcal{O}$. But beforehand, we replace each property F_i by properties $F_{i,1} := F_i \cap E_1, \ldots, F_{i,r_1} := F_i \cap E_{r_1}$. Since this does not change $O(F_{\ldots})$, it does not affect P. Now we may further assume that no E_i equals an F_j. Otherwise, we replace the tuple

$$(E_1, \ldots, E_{r_1}; F_{1,1}, \ldots, F_{s_1,r_1})$$

by

$$(E_1, \ldots, \widehat{E_i}, \ldots, E_{r_1}; F_{1,1}, \ldots, \widehat{F_{i,1}}, \widehat{\ldots}, \widehat{F_{i,r_1}}, \ldots, F_{s_1,r_1}),$$

where a hat indicates that the entry is removed. (The change of the F is necessary to ensure that each $O(F)$ is still contained in some $O(E)$). You can easily check that this procedure does not change the value of P.

Furthermore, we may assume that all E_i, E'_i are maximal and all F_i, F'_i are maximal (in the sets $\{F_j\}, \{F'_j\}$, respectively). If not, then remove the superfluous sets.

Now we proceed as in the proof for $\mathcal{O}$. First we show that the E_i and the E'_i coincide. Assume E_1 does not appear in the right hand side. Choose mutually distinct primes $p_i, p_{i,j}$ for each E'_i and each $F_{i,j}$, respectively. Then construct a group such that its p_i-part corresponds to a partition in $E_1^{-1}(1) \backslash E_i'^{-1}(1)$ and its $p_{i,j}$-part corresponds to a partition in $E_1^{-1}(1) \setminus F_{i,j}^{-1}(1)$. The assumptions above ensure that the latter sets are all non-empty. Then the group is in E_1, but it is neither in any E'_i nor in any $F_{i,j}$. Therefore, it is contained in the left hand side, but not in the right hand side of (5.3). Contradiction! So the assumption was wrong, and the E_i and the E'_j coincide.

Now turn to the $F_{i,j}$ and $F'_{i,j}$. Since $O(E_i) = O(E'_i)$, $O(F_{i,j}) \subseteq O(E_i)$, $O(F'_{i,j}) \subseteq O(E'_i)$, and $O(E_i) \setminus O(F_{i,j}) = O(E'_i) \setminus O(F'_{i,j})$, we can deduce $O(F_{i,j}) = O(F'_{i,j})$. Now we may apply the first part of the proof (for $\mathcal{O}$) to

conclude that $P(O(F_{i,j})) = P(O(F'_{i,j}))$. Putting things together, we see that

$$P(D(E_1, \ldots, E_{r_1}; F_1, \ldots, F_{s_1})) = P(D(E'_1, \ldots, E'_{r_2}; F'_1, \ldots, F'_{s_2})),$$

as required. □

5.3.8 Remark.
In the following proofs (as well as in the proof above), be aware that the formula

$$P(D(E_1, \ldots, E_r; F_1, \ldots, F_s)) = P(O(E_1, \ldots, E_r)) - P(O(F_1, \ldots, F_s))$$

is *not true* if we omit the condition

$$O(F_1, \ldots, F_s) \subseteq O(E_1, \ldots, E_r).$$

We will use the function P to define an outer measure. But before that, we prove a technical lemma about P:

5.3.9 Lemma.

(i) Let $D_1, \ldots, D_n \in \mathcal{D}$ be mutually disjoint, and let $D_0 \in \mathcal{D}$ be such that

$$\bigcup_{i=1}^{n} D_i \subseteq D_0.$$

Then

$$\sum_{i=1}^{n} P(D_i) \leq P(D_0).$$

In particular, this implies that P is monotone, i.e., for $D_1 \subseteq D_0$ we have $P(D_1) \leq P(D_0)$.

(ii) Let $D_0, D_1, \ldots, D_n \in \mathcal{D}$ be such that

$$D_0 \subseteq \bigcup_{i=1}^{n} D_i.$$

Then

$$P(D_0) \leq \sum_{i=1}^{n} P(D_i).$$

Proof. We only prove the first statement, which is slightly more complicated. The proof of the second case is completely analogous, except that we do not have to worry about the D_i being disjoint.

Let $D_i = D(\mathcal{E}_i, \mathcal{F}_i)$ for all $i = 0, \dots, n$, where $\mathcal{E}_i, \mathcal{F}_i$ are collections of uniform properties. We may assume $O(\mathcal{F}_i) \subseteq O(\mathcal{E}_i)$ for all i. Then $P(D_i) = P(O(\mathcal{E}_i)) - P(O(\mathcal{F}_i))$ for all $i = 0, \dots, n$.
Therefore, we need to show that

$$\sum_{i=1}^{n} P(O(\mathcal{E}_i)) - \sum_{i=1}^{n} P(O(\mathcal{F}_i)) \leq P(O(\mathcal{E}_0)) - P(O(\mathcal{F}_0)), \tag{5.4}$$

or equivalently by expanding the $P(O(\mathcal{E}_i))$:

$$\begin{aligned} \sum_{i=1}^{n} \sum_{S \subseteq \mathcal{E}_i} (-1)^{\#S} P\Big(\bigwedge_{E \in S} E \Big) - \sum_{i=1}^{n} \sum_{S \subseteq \mathcal{F}_i} (-1)^{\#S} P\Big(\bigwedge_{F \in S} F \Big) \\ \leq \sum_{S \subseteq \mathcal{E}_0} (-1)^{\#S} P\Big(\bigwedge_{E \in S} E \Big) - \sum_{S \subseteq \mathcal{F}_0} (-1)^{\#S} P\Big(\bigwedge_{F \in S} F \Big). \end{aligned} \tag{5.5}$$

Let us first examine the prerequisites of the statement. We may assume that no $E \in \mathcal{E}_i$ is contained in any $F \in \mathcal{F}_i$, for $i = 0, \dots, n$. Then it is easy to see that the prerequisites are satisfied if and only if the following conditions are satisfied:

1. $O(\mathcal{E}_i) \subseteq O(\mathcal{E}_0)$ for $i = 1, \dots, n$.
2. $O(\mathcal{E}_i) \cap O(\mathcal{E}_j) \subseteq O(\mathcal{F}_i) \cup O(\mathcal{F}_j)$ for all $1 \leq i < j \leq n$.
3. $O(\mathcal{F}_0) \subseteq O(\mathcal{F}_i)$ for $i = 1, \dots, n$.

Now let $\mathbb{P}_{\leq x} := \{p \in \mathbb{P} \mid p \leq x\}$ and let

$$\mathcal{G}_{\leq x} := \prod_{p \in \mathbb{P}_{\leq x}} \mathcal{G}_p.$$

Then $\mathcal{G}_{\leq x}$ is the direct product of probability spaces and carries a unique product probability measure. The set $\mathcal{G}_{\leq x}$ embeds naturally into $\mathcal{G}$. So for each uniform property E, we may define $O_{\leq x}(E) := O(E) \cap \mathcal{G}_{\leq x}$. By definition of the product probability, we have for these sets the probabilities

$$P_{\leq x}(E) := P_{\mathcal{G}_{\leq x}}(O_{\leq x}(E)) = \prod_{p \in \mathbb{P}_{\leq x}} P_p(E).$$

Then it is evident that for any uniform property E,

$$P(E) = \lim_{x\to\infty} P_{\le x}(E).$$

Conditions 1.–3. are still satisfied if we intersect both sides with $\mathcal{G}_{\le x}$, so we also have

1'. $O_{\le x}(\mathcal{E}_i) \subseteq O_{\le x}(\mathcal{E}_0)$ for $i = 1, \dots, n$.

2'. $O_{\le x}(\mathcal{E}_i) \cap O_{\le x}(\mathcal{E}_j) \subseteq O_{\le x}(\mathcal{F}_i) \cup O_{\le x}(\mathcal{F}_j)$ for all $1 \le i < j \le n$.

3'. $O_{\le x}(\mathcal{F}_0) \subseteq O_{\le x}(\mathcal{F}_i)$ for $i = 1, \dots, n$.

Now for sufficiently large x (we need more primes than uniform properties involved), the conditions 1.'–3.' are equivalent to the statement

$$\begin{aligned} &D_1 \cap \mathcal{G}_{\le x}, \dots, D_n \cap \mathcal{G}_{\le x} \text{ are mutually disjoint, and} \\ &\bigcup_{i=1}^{n} D_i \cap \mathcal{G}_{\le x} \subseteq D_0 \cap \mathcal{G}_{\le x}. \end{aligned}$$

Since $\mathcal{G}_{\le x}$ is a probability space, we deduce

$$\sum_{i=1}^{n} P(D_i \cap \mathcal{G}_{\le x}) \le P(D_0 \cap \mathcal{G}_{\le x}),$$

or equivalently

$$\begin{aligned} \sum_{i=1}^{n} \sum_{S\subseteq\mathcal{E}_i} (-1)^{\#S} P_{\le x}\Big(\bigwedge_{E\in S} E\Big) - \sum_{i=1}^{n} \sum_{S\subseteq\mathcal{F}_i} (-1)^{\#S} P_{\le x}\Big(\bigwedge_{F\in S} F\Big) \\ \le \sum_{S\subseteq\mathcal{E}_0} (-1)^{\#S} P_{\le x}\Big(\bigwedge_{E\in S} E\Big) - \sum_{S\subseteq\mathcal{F}_0} (-1)^{\#S} P_{\le x}\Big(\bigwedge_{F\in S} F\Big). \end{aligned}$$

Since we have finite sums and differences on both sides, we obtain equation (5.5) by taking the limit $x \to \infty$. This proves the claim.

□

Now we come to the definition of the outer measure:

5.3.10 Definition. *For any $A \subset \mathcal{G}$, we define the* outer measure ν *as*

$$\nu(A) := \inf\left\{ \sum_{i=1}^{\infty} P(A_i) \,\middle|\, A_i \in \mathcal{D} \text{ and } A \subset \bigcup_{i=1}^{\infty} A_i \right\}.$$

5.3.11 Remark. The definition above always yields an outer measure, for any map $P : S \to [0, \infty]$, where S is any subset of the power set of $\mathcal{G}$ containing $\emptyset$ and $P(\emptyset) = 0$ [Mun53].
Recall that an outer measure is almost a measure, only we replace the Σ-additivity by Σ-subadditivity. More precisely, an outer measure on a space X is a function ν from the power set of X into the interval $[0, \infty]$ satisfying the three conditions:

- $\nu(\emptyset) = 0$.
- *Monotonicity*: $\nu(A) \leq \nu(B)$ for all $A \subseteq B \subseteq X$.
- Σ-*subadditivity*:
$$\nu(\bigcup_{i=1}^{\infty} A_i) \leq \sum_{i=1}^{\infty} \nu(A_i)$$
for all $A_i \subseteq X$.

5.3.3 The global measure

Next we check that ν and P coincide on $\mathcal{D}$. We divide up the proof into several steps. First, we prove a helpful lemma:

5.3.12 Lemma. *Let $D = D(E_1, \ldots, E_r; F_1, \ldots, F_s) \in \mathcal{D}$ such that for all $1 \leq i \leq r$ we have finite fibers $E_i^{-1}(1)$, and assume without loss of generality that $O(E_i) \not\subseteq O(F_{i'})$ for all i, i'. Let $\tilde{D}_j = D(\tilde{E}_{j,k}; \tilde{F}_{j,k'})$ be an arbitrary family in $\mathcal{D}$ such that*

$$D \subseteq \bigcup_j \tilde{D}_j.$$

Then for each i there exists a j such that $O(E_i) \subseteq O(\tilde{E}_{j,1}, \tilde{E}_{j,2}, \ldots)$ and such that $O(E_i) \not\subseteq O(\tilde{F}_{j,1}, \tilde{F}_{j,2}, \ldots)$.

Proof. We use a similar argument as in the proof of 5.3.7. Assume that the assertion is wrong for some i. Then for all j there exists an $\underline{n} \in E_i^{-1}(1)$ such that $\underline{n} \notin \tilde{D}_j$, and in particular $\underline{n} \notin O(\tilde{E}_{j,1}, \tilde{E}_{j,2}, \ldots)$.
Now choose mutually distinct primes $p_{\underline{n}}$ for each $\underline{n} \in E_i^{-1}(1)$. Consider a group G with $p_{\underline{n}}$-part $\underline{n}$ for all $\underline{n} \in E_i^{-1}(1)$. Then $G \in O(E_i)$, but $G \notin O(F_{i'})$ for all i', since otherwise $O(E_i) \subseteq O(F_{i'})$.
Furthermore, $G \notin D(\tilde{E}_{j,1}, \tilde{E}_{j,2}, \ldots; \tilde{F}_{j,1}, \tilde{F}_{j,2}, \ldots)$ for all j, contradicting the prerequisite $D \subset \bigcup_j \tilde{D}_j$. This proves the lemma. □

Next we prove that P and ν coincide on a certain subset of $\mathcal{D}$.

5.3.13 Lemma. *Let D be the finite disjoint union of sets*

$$D^{(k)} = D(E_1^{(k)}, \dots, E_{r_k}^{(k)}; F_1^{(k)}, \dots, F_{s_k}^{(k)}) \in \mathcal{D}$$

such that the fibers $(E_i^{(k)})^{-1}(1)$ are finite for all i and k. Then we have

$$\nu(D) = \sum_k P(D^{(k)}).$$

Proof. We set $D(E_1, \dots, E_r; F_1, \dots, F_s) < D(E_1', \dots, E_r'; F_1', \dots, F_s')$ if and only if $O(E_1, \dots, E_r) \subsetneq O(E_1', \dots, E_r')$; in this way, we impose a partial ordering on $\mathcal{D}$. It is a well-ordering on sets in $\mathcal{D}$ with finite fibers $E_i^{-1}(1)$, and the lexicographic ordering extends this to the set of all finite tuples $(D^{(k)})$ of elements of $\mathcal{D}$ with finite fibers $E_i^{-1}(1)$. The lexicographic ordering is still a well-ordering so we may use induction with respect to this ordering.
For the sake of clarity, I will restrict the proof to the case where we have only one set $D^{(k)}$ and simplify the notation to $D = D(E_1, \dots, E_r; F_1, \dots, F_s)$. The extension to a disjoint union is straightforward: Just apply the descending step to the largest (possibly several) $D^{(k)}$'s with respect to the ordering.
As usual we assume that the sets $O(E_i)$ are mutually not contained in each other.
If D is minimal then $D = \emptyset$ and we have $P(D) = 0 = \nu(D)$.
So assume $D \neq \emptyset$. Since $D \in \mathcal{D}$ we have $\nu(D) \leq P(D)$. So we only need to show that for any countable familiy A_j in $\mathcal{D}$ with $D \subset \bigcup_{j=1}^{\infty} A_j$ we have $P(D) \leq \sum_{j=1}^{\infty} P(A_j)$.
Let A_j be such a family. Consider E_1. If E_1 is contained in any of the sets $F_1, \dots, F_s$, then we simply omit it and we are done by induction hypothesis. So assume otherwise. Then by lemma 5.3.12 there exists an index j_0, $A_{j_0} = D(\tilde{E}_k; \tilde{F}_k) =: D(\tilde{\mathcal{E}}; \tilde{\mathcal{F}})$, such that

$$\begin{aligned} O(E_1) &\subseteq O(\tilde{\mathcal{E}}), \text{ and} \\ O(E_1) &\not\subseteq O(\tilde{\mathcal{F}}). \end{aligned}$$

Now consider the set $D_0 := D \setminus A_{j_0}$. We will see that we need to compute the measure of this set. Unfortunately, D_0 is not in $\mathcal{D}$ in general, but it is the disjoint union of two elements in $\mathcal{D}$. Basically we will use the decomposition

$$D_0 = D \setminus A_{j_0} = D \setminus (O(\tilde{\mathcal{E}}) \setminus O(\tilde{\mathcal{F}})) = (\underbrace{D \setminus O(\tilde{\mathcal{E}})}_{=:D_1}) \dot\cup (\underbrace{D \cap O(\tilde{\mathcal{F}})}_{=:D_2}).$$

We need to show that $D_1, D_2 \in \mathcal{D}$: We write $D_1 = D(\mathcal{E}_1; \mathcal{F}_1)$, where

$$\begin{aligned} \mathcal{E}_1 &:= \{E_2, \dots, E_r\} \cup \{E_1 \wedge \tilde{F}_1, E_1 \wedge \tilde{F}_2, \dots\}, \text{ and} \\ \mathcal{F}_1 &:= \{F_1, \dots, F_s\} \cup \{\tilde{E}_1, \tilde{E}_2, \dots\}, \end{aligned}$$

and $D_2 = D(\mathcal{E}_2; \mathcal{F}_2)$, where

$$\begin{aligned} \mathcal{E}_2 &:= \{E_i \wedge \tilde{F}_{i'} \mid i, i' = 1, 2, \ldots\}, \text{ and} \\ \mathcal{F}_2 &:= \{F_1, \ldots, F_s\}. \end{aligned}$$

Then D_1 and D_2 are disjoint with union $D \setminus A_{j_0}$, they have finite fibers and are strictly smaller (in the inductive sense) than D, so we may apply the induction hypothesis and conclude $\nu(D_1 \cup D_2) = P(D_1) + P(D_2)$.
Since $D_1 \cup D_2 = D \setminus A_{j_0}$, we have

$$D_1 \cup D_2 \subset \bigcup_{\substack{j=1 \\ j \neq j_0}}^{\infty} A_j,$$

and consequently

$$\nu(D_1 \cup D_2) \leq \sum_{\substack{j=1 \\ j \neq j_0}}^{\infty} P(A_j).$$

Now we can put everything together: Reusing the formula $D \subseteq D_1 \cup D_2 \cup A_{j_0}$, we see that $P(D) \leq P(D_1) + P(D_2) + P(A_{j_0})$ by lemma 5.3.9, and therefore

$$\begin{aligned} P(D) &\leq P(D_1) + P(D_2) + P(A_{j_0}) \\ &= \nu(D_1 \cup D_2) + P(A_{j_0}) \\ &\leq \left(\sum_{\substack{j=1 \\ j \neq j_0}}^{\infty} P(A_j) \right) + P(A_{j_0}) \\ &= \sum_{j=1}^{\infty} P(A_j). \end{aligned}$$

This proves $P(D) \leq \nu(D)$, as required.

□

Now we are ready to tackle the general case:

5.3.14 Proposition. *For any $D \in \mathcal{D}$, we have $\nu(D) = P(D)$.*

Proof. Let $D = D(E_1, \ldots, E_r; F_1, \ldots, F_s)$. Let $\mathcal{E}$ be the r-tuple $(E_1, \ldots, E_r)$ and let $\mathcal{F}$ be the s-tuple $(F_1, \ldots, F_s)$. In the following, $\mathcal{E}'$ will always denote an r-tuple of uniform properties that is finite in the sense that the fibers $E'^{-1}(1)$ are finite for all properties E' in $\mathcal{E}'$. We shall write $\mathcal{E}' \leq \mathcal{E}$ if for all $1 \leq i \leq r$ we have $O(E'_i) \subseteq O(E_i)$.

The crucial step in this proof is to show

$$P(D) = \sup_{\mathcal{E}' \leq \mathcal{E} \text{ finite}} P(D(\mathcal{E}', \mathcal{F})). \tag{5.6}$$

The inequality "$\geq$" is trivial. For the other direction, note that for any finite $\mathcal{E}' \leq \mathcal{E}$, we have

$$P(D(\mathcal{E}, \mathcal{F})) - P(D(\mathcal{E}', \mathcal{F})) \leq P(O(\mathcal{E})) - P(O(\mathcal{E}')).$$

Therefore, it suffices to show that $P(O(\mathcal{E})) = \sup_{\mathcal{E}'} P(O(\mathcal{E}'))$.

Furthermore, it suffices to consider the case $r = 1$ (i.e., $\mathcal{E}$ consists of only one uniform property), because by the Inclusion-Exclusion formula $P(O(\mathcal{E}))$ can be computed as a finite sum (with signs) from values $P(O(E))$, where E is a single uniform property.

Altogether, we need to show that for each uniform property E, we have

$$P(O(E)) = \sup_{E' \leq E \text{ finite}} P(O(E')),$$

where "$E' \leq E$ finite" means that $E'^{-1}(1) \subseteq E^{-1}(1)$ and $E'^{-1}(1)$ is finite.

We may assume that $P(O(E)) > 0$, otherwise the statement is trivial.

Let us look at the local situation: Let $p \in \mathbb{P}$ and let $n_0 \in \mathbb{N}$. For any $n \in \mathbb{N}^+$, it is possible to choose $E' \leq E$ finite such that $w_p(E) \leq w_p(E') + \sum_{i=n}^{\infty} a_i q^i$ (a_i = number of partitions of i) *as power series*, i.e., coefficient-wise.

By lemma 1.3.7 we know that $a_i \in O(\phi^i)$, where $\phi = 1.618\ldots$ is the golden ratio. There exists a constant $d < 1$ (e.g., $d := 0.7$) such that $2^d > \phi$ and such that $2^{4-d} > 2^3 + 1$. Then it is easy to see that for all primes p we have $p^{4-d} > p^3 + 1$. By choosing n large enough, we may further assume that $a_i \leq 2^{di - n_0 - 3}$ for all $i \geq n$. Then in particular $a_i \leq p^{di - n_0 - 3}$ for all primes p. Also by lemma 5.3.3, we may assume that $P(O(E')) \geq c$ for some $c > 0$, and therefore also $P_p(E') \geq P(O(E')) \geq c$ for all $p \in \mathbb{P}$.

Then we have

$$\begin{aligned} w_p(E) &\leq w_p(E') + \sum_{i=n}^{\infty} a_i q^i \\ &\leq w_p(E') + \sum_{i=n}^{\infty} p^{di-n_0-3} p^{-i} \\ &= w_p(E') + p^{-n_0-3} \sum_{i=n}^{\infty} p^{(d-1)i} \\ &= w_p(E') + p^{-n_0} p^{-3} \frac{p^{n(d-1)}}{1-p^{d-1}} \\ &= w_p(E') + p^{-n_0} \frac{p^{(n-1)(d-1)}}{p^{4-d}-p^3} \\ &\leq w_p(E') + p^{-n_0}, \end{aligned}$$

where in the last inequality we use that the fraction has numerator ≤ 1 and denominator ≥ 1.
For the probability, we must multiply with $\prod_{i=0}^{\infty}(1-p^{-i})$:

$$P_p(E) \leq P_p(E') + p^{-n_0} \prod_{i=0}^{\infty}(1-p^{-i}) \leq P_p(E') + p^{-n_0}$$

Since our choice of E' and of n was independent of p, the analysis works for all p. Putting this together, we get

$$\begin{aligned} P(E) &= \prod_{p\in\mathbb{P}} P_p(E) \\ &\leq \prod_{p\in\mathbb{P}} \left(P_p(E') + p^{-n_0}\right) \\ &= \left(\prod_{p\in\mathbb{P}} P_p(E')\right) \prod_{p\in\mathbb{P}} \left(1 + \frac{p^{-n_0}}{P_p(E')}\right) \\ &\leq \left(\prod_{p\in\mathbb{P}} P_p(E')\right) \prod_{p\in\mathbb{P}} \left(1 + \frac{1}{c} p^{-n_0}\right) \\ &\leq \left(\prod_{p\in\mathbb{P}} P_p(E')\right) \left(1 + \sum_{p\in\mathbb{P}} \left(\frac{1}{c} p^{-n_0}\right)\right) \end{aligned}$$

$$\begin{aligned} &= P(E') \left(1 + \frac{1}{c} \underbrace{\sum_{p \in \mathbb{P}} p^{-n_0}}_{\to 0 \text{ for } n_0 \to \infty} \right) \\ &\overset{n_0 \to \infty}{\to} P(E'). \end{aligned}$$

This proves equation (5.6).

Now let $A_j \in \mathcal{D}$ be a countable family with $D \subseteq \bigcup_{j=1}^{\infty} A_j$. We need to show that $P(D) \leq \sum_{j=1}^{\infty} P(A_j)$.
Recall that $D = D(\mathcal{E}, \mathcal{F})$. Let $\mathcal{E}' \leq \mathcal{E}$ be finite. Then $D(\mathcal{E}', \mathcal{F}) \subseteq D \subseteq \bigcup_{j=1}^{\infty} A_j$, so by lemma 5.3.13, we have

$$P(D(\mathcal{E}', \mathcal{F})) \leq \sum_{j=1}^{\infty} P(A_j).$$

Therefore,

$$P(D) \overset{5.6}{=} \sup_{\mathcal{E}' \leq \mathcal{E} \text{ finite}} P(D(\mathcal{E}', \mathcal{F})) \leq \sum_{j=1}^{\infty} P(A_j),$$

which finishes the proof. □

For the last step, we use

5.3.15 Theorem (Carathéodory)**.** *Let X be some space with outer measure ν. We call a set $A \subseteq X$* measurable*, if for all $B \subseteq X$ we have*

$$\nu(B) = \nu(B \setminus A) + \nu(B \cap A).$$

Then the set of all measurable sets is a σ-algebra, and ν is a measure when restricted to measurable sets.

Proof. [Hal50] □

So we only need to show that all uniform properties are measurable (in the sense of Carathéodory):

5.3.16 Proposition. *Let E be a uniform property. Then $O(E)$ is measurable.*

Proof. Let $A \subseteq \mathcal{G}$. We need to show that $\nu(A) = \nu(A \setminus O(E)) + \nu(A \cap O(E))$. Since ν is subadditive (as outer measure), we only need to show the direction

$$\nu(A) \geq \nu(A \setminus O(E)) + \nu(A \cap O(E)).$$

Let $A_i \in \mathcal{D}$ be a family such that $A \subseteq \bigcup_{i=1}^{\infty} A_i$. By definition of ν, it suffices to show that for any such family

$$\sum_{i=1}^{\infty} P(A_i) \geq \nu(\underbrace{A \setminus O(E)}_{=:B}) + \nu(\underbrace{A \cap O(E)}_{=:C}).$$

Since $A_i \in \mathcal{D}$, we also have $B_i := A_i \setminus O(E) \in \mathcal{D}$ and $C_i := A_i \cap O(E) \in \mathcal{D}$. Therefore, by proposition 5.3.14, we have $\nu(B_i) = P(B_i)$, $\nu(C_i) = P(C_i)$, and $P(A_i) = P(B_i) + P(C_i)$.
Clearly the B_i cover B, and the C_i cover C, so by definition of ν

$$\begin{aligned} \nu(B) &\leq \sum_{i=1}^{\infty} P(B_i), \text{ and} \\ \nu(C) &\leq \sum_{i=1}^{\infty} P(C_i). \end{aligned}$$

Putting things together, we obtain

$$\begin{aligned} \sum_{i=1}^{\infty} P(A_i) &= \sum_{i=1}^{\infty} (P(B_i) + P(C_i)) \\ &= \sum_{i=1}^{\infty} P(B_i) + \sum_{i=1}^{\infty} P(C_i) \\ &\geq \nu(B) + \nu(C), \end{aligned}$$

as required.

□

So we have successfully concluded the proof and shown that the Cohen-Lenstra probability measure (def. 5.2.5) is indeed a probability measure.

5.4 Modifications of the global measure

As we will see in chapter 6, there are some important cases where we need to exclude certain primes. E.g., for quadratic number fields we need to exclude $p = 2$. In this case, we proceed as follows: We consider the set $\mathcal{G}^{\neq 2}$ of all

finite abelian groups with trivial 2-part and modify our definition of uniform properties to these groups. It is clear that all our proofs work also for $\mathcal{G}^{\neq 2}$ instead of $\mathcal{G}$, so we get a probability measure on $\mathcal{G}^{\neq 2}$ that makes all (modified) uniform properties measurable.
Then either we stop at this point and do not make any statements about groups with non-trivial 2-part. In this case we often replace a random G by G/G_2, where G_2 is the 2-part of G. Or, if we are given a probability measure on the set $\mathcal{G}_2$ of all finite abelian 2-groups, then we take the product space of $\mathcal{G}^{\neq 2}$ and $\mathcal{G}_2$ and obtain automatically a probability measure on the product space. Candidates for such probability measures for "bad" primes are known for number fields (cf. the discussion in section 6.1.2).
Of course, all this applies also to other primes than $p = 2$, and also to a finite number of primes.
CAUTION: We get a different probability space for each finite set of primes, and *those probability spaces are not compatible*. As we have seen in section 5.1, there is no rich probability measure whose σ-algebra would make all projections $\mathcal{G} \to \mathcal{G}_p$ continuous.
So there are no objections against ruling out some bad primes in a number field situation (in the sense of section 6.1.2), since these primes are fixed. But if you fix one situation and make statements about the p-parts of the class groups for various p (as it is often done, e.g. in [CL84]), then you must be extremely careful, because our analysis above has shown that you will *inevitably lose countable additivity*. Therefore, the interpretation as probabilities is not valid in this context! Unfortunately, this point is usually ignored in the literature.
A more general way of extending uniform properties is to split up the primes into a finite number of subsets, e.g., into $\mathbb{P}_1 := \{p \in \mathbb{P} \mid p \equiv 1 \bmod 4\}$, $\mathbb{P}_2 := \{2\}$, and $\mathbb{P}_3 := \{p \in \mathbb{P} \mid p \equiv 3 \bmod 4\}$. Then we may define uniform properties for each of the sets $\mathcal{G}_{\mathbb{P}_1}$, $\mathcal{G}_{\mathbb{P}_2}$, and $\mathcal{G}_{\mathbb{P}_3}$ (in the obvious way), and by combining them we obtain a probability measure on $\mathcal{G}$ that is an extension of the probability measure we have defined in the preceding sections. In this way, we may formulate equidistribution statements for congruence classes of primes. However, we have the same restriction as we have when taking out finitely many primes: Each partition of the set $\mathbb{P}$ yields its own probability measure, and combining more than finitely many of them will eventually result in losing the countable additivity.

Another extension is obvious from measure theory: Of course, we are not restricted to measuring *properties*, but we may measure any measurable function, which includes measuring expected values, higher moments of random variables and many other things. This seems like a trivial remark, but so

far it has been an unsolved problem which quantities to consider in the Cohen-Lenstra context. Cohen and Lenstra declared that we should take "reasonable" functions without specifying what "reasonable" means, and this handwaving concept was adapted in basically all subsequent papers. By our preparatory work, we get the solution for this problem for free from measure theory.
For convenience, let me explicitly state what it means for a sequence of groups to be random (more precisely: equidistributed) with respect to the Cohen-Lenstra heuristic:

5.4.1 Definition. *Let $(G_i)_{i=1}^{\infty}$ be a sequence of finite abelian groups. Let Σ be the σ-algebra on $\mathcal{G}$ generated by uniform properties and let μ be the probability measure on Σ as defined in 5.2.5. We say that G_i* behaves as a random sequence *or* is equidistributed with respect to the Cohen-Lenstra measure *if for all measurable functions $f : \mathcal{G} \to \mathbb{C}$ we have*

$$\lim_{n\to\infty} \frac{\sum_{i=1}^{n} f(G_i)}{n} = \int_{\mathcal{G}} f d\mu.$$

5.5 Combination of both methods

The methods of restricted countability and uniform properties may be combined. Let $\mathcal{A}$ denote the algebra of subsets of $\mathcal{G}$ generated by sets $\pi_p^{-1}(M)$, for various $p \in \mathbb{P}$ and $M \subseteq \mathcal{G}_p$, and let Σ denote the σ-algebra over $\mathcal{G}$ that is generated by uniform properties. Then we may consider the algebra $\mathcal{A}'$ that is generated by $\mathcal{A} \cup \Sigma$, and we naturally have a content on $\mathcal{A}'$. It turns out that this is still a content with restricted countable additivity (in the sense of 5.1.2). I will only give a sketch of the proof, because in my eyes it is (at least from a theoretical point of view) only a minor extension of the σ-algebra Σ. Nevertheless, it justifies at least to some extent the convenient habit of researchers to switch between local and global (uniform) properties when studying a sequence of global groups.
Essentially, the reason for the two concepts being compatible is that uniform properties are "horizontal" (uniform over the set of all primes), whereas the algebra $\mathcal{A}$ is "vertical" (each set lives only on finitely many primes), and therefore they do not interfere with each other. More precisely, whenever a set $S \in \Sigma$ is a superset of a non-empty set $A \in \mathcal{A}$, then there are primes p such that for any $G_p \in \mathcal{G}_p$, there is a group $G \in \mathcal{A} \subseteq S$ with p-part G_p. Since S is uniform, this already implies $S = \mathcal{G}$. Hence, sets in Σ cannot contribute to a disjoint cover of an element of $\mathcal{A}$ in a non-trivial way, and

the restricted countability for $\mathcal{A}'$ reduces to the restricted countability of $\mathcal{A}$, which was already shown in section 5.1.2.

Chapter 6

Applications and Extensions

I already mentioned two applications of the Cohen-Lenstra heuristic: Firstly, it is the probability distribution that one obtains by randomly choosing generators and relations (cf. section 2.2.3), which is perhaps the most natural way to produce a finite abelian p-group.

Secondly, the size of conjugacy classes of $\mathrm{GL}(n,p)$ is essentially governed by the Cohen-Lenstra probability, as explained in detail in chapter 4. Consequently, one can deduce statements about random matrices from the Cohen-Lenstra measure. Examples are the probability that a random matrix has fixed spaces of given dimension (corollary 4.6.4), is regular, is semisimple ([Ful97, Thm. 27, Thm. 25], respectively), or satisfies a given polynomial equation ([Ger61], [Sto88], and [Ful97, Thm. 13]).

However, the main application for us are the class groups of number fields. This was also the motivation for Cohen and Lenstra to invent this distribution. We first look at the case of imaginary quadratic number fields, which is in some sense the "generic case". Afterwards, I will demonstrate how to extend the Cohen-Lenstra heuristic to arbitrary number fields. Note that this section is completely conjectural (unless otherwise stated). To make things more concrete, I will exemplarily show how this extension looks like for real quadratic number fields.

In section 6.2, I give a "real-world application" for which it is important to know the distribution of number fields for a certain cryptographic protocol. In the final sections, I give some other applications of the Cohen-Lenstra heuristics, as well as its transfer from number fields to function fields.

6.1 Number fields

6.1.1 Imaginary quadratic number fields

An imaginary quadratic number field is of the form $K = \mathbb{Q}(\sqrt{d})$, where $d < 0$ is a square-free integer. They are naturally ordered by their discriminants D_K (where $D_K = d$ if $d \equiv 1 \bmod 4$, and $D_K = 4d$ otherwise). We write H_K for the class group of K.
By $H_K^{\neq 2}$ we denote the group $H_K/H_{K,2}$ where $H_{K,2}$ is the 2-part of H_K. Why do we exclude the 2-part? By genus theory, we know that $H_{K,2}$ has rank $r-1$, where r is the number of distinct prime factors of D_K. Thus the 2-part is clearly "biased" in the sense mentioned in the introduction. This is why the following prediction can only be valid for the non-2-part.
Cohen and Lenstra have predicted in [CL84] that the sequence $(H_K^{\neq 2})_K$ of class groups of imaginary quadratic number fields, ordered by their discriminant, behaves like a random sequence of groups w.r.t. to the Cohen-Lenstra probability subject to the condition that the 2-part is trivial. Of course, they did not have a probability measure, nor the notion of measurable functions, so their formulation was a bit vague. I will say a more about their formulation in remark 6.1.3. With our knowledge about the global theory, we may state the conjecture as follows:

6.1.1 Conjecture. *If f is a measurable random variable on $\mathcal{G}^{\neq 2}$ with existing expected value $E(f)$ (w.r.t. the global Cohen-Lenstra probability measure, cf. chapter 5), then the limit*

$$\lim_{x\to\infty} \frac{\sum_{0>D_K\geq -x} f(H_K^{\neq 2})}{\sum_{0>D_K\geq -x} 1}$$

exists and is equal to $E(f)$, where the sums run over all imaginary quadratic number fields of discriminant $0 > D_K \geq -x$.

Let me remind you that characteristic functions (that output 1 if a group satisfies a certain property, and 0 otherwise) are of particular interest, because they enable us to compute the probability that a class group satisfies this property.
I want to emphasize that a proof of this conjecture seems far out of reach. Some partial results have been proven for the 3-part of the class group (cf. section 6.1.2), but these only confirm some implications of the Cohen-Lenstra conjecture, and they could so far not be generalized to other primes.
On the other hand, extensive numerical tests strongly support the prediction, and a huge majority of researchers has no doubt that the Cohen-Lenstra conjectures will prove true.

6.1.2 Arbitrary number fields

The probabilistic model

Following the notation of Malle [Mal06], we fix a *situation* $\Sigma = (K_0, G_0, \sigma)$ consisting of a number field K_0 (the base field), a transitive permutation group $G_0 \subset S_n$ for some $n \geq 2$, and a possible signature σ (which is a pair (r_1, r_2) of the number of real and pairs of complex embeddings, respectively) of a degree n galois extension K/K_0 with group G. By $\mathcal{K} := \mathcal{K}(\Sigma)$ we denote the set of all galois extensions K/K_0 (inside a fixed algebraic closure) with Galois group G_0 and signature σ.
As in the previous section, for a field $K \in \mathcal{K}$ we denote by D_K the (relative) discriminant of K/K_0 and by H_K the class group of K. For each discriminant there are at most finitely many fields in $\mathcal{K}$. If $B = \{p_1, p_2, \ldots\}$ ("B" like "bad") is a finite set of primes, then $H_K^{\notin B}$ is the group $H_K/(H_{K,p_1} \times H_{K,p_2} \times \ldots)$ where H_{K,p_i} is the p_i-part of the class group H_K.
Let further $u := u_\Sigma$ be the rank of the group of units in K. By Dirichlet's unit theorem, $u = r_1 + r_2 - 1$, where r_1 is the number of real embeddings and r_2 is the number of conjugate pairs of complex embeddings of K into the (fixed) algebraic closure.
Then the Cohen-Lenstra heuristic predicts that except for a finite set B of "bad primes", the class group of fields in $\mathcal{K}$ behaves like a random finite group G modulo the image of a random homomorphism $\phi : \mathbb{Z}^u \to G$. In formula:

6.1.2 Conjecture. *If f is a measurable random variable on $\mathcal{G}^{\notin B}$, then the limit*

$$\lim_{x\to\infty} \frac{\sum_{K\in\mathcal{K},|D_K|<x} f(H_K^{\notin B})}{\sum_{K\in\mathcal{K},|D_K|<x} 1}$$

exists and equals the limit

$$\lim_{x\to\infty} \left(\frac{\sum_{G\in\mathcal{G}^{\notin B},\mathrm{ord}(G)\leq x} \frac{w(G)}{\#\mathrm{Hom}(\mathbb{Z}^u,G)} \sum_{\phi\in\mathrm{Hom}(\mathbb{Z}^u,G)} f(G/\mathrm{im}(\phi))}{\sum_{G\in\mathcal{G}^{\notin B},\mathrm{ord}(G)\leq x} w(G)} \right). \tag{6.1}$$

6.1.3 Remark.

- The conjecture was formulated by Cohen and Lenstra [CL84] for (imaginary and real) quadratic number fields and was extended by Cohen and Martinet [CM87], [CM90] to arbitrary number fields. The generalizations of Cohen and Martinet seem to be wrong in details (at least they do not fit the numerical data). More precisely, Cohen and Martinet

have seemingly chosen the set B to be too small. But of course, the above conjecture is in the general spirit of their statement, and therefore it is only fair to call it "Cohen-Lenstra conjecture" for quadratic fields, and "Cohen-Lenstra-Martinet conjecture" when number fields of higher degrees are included.

- The formulation of Cohen and Lenstra was in fact quite different from the formulation above. As I mentioned, they did not have the notion of a measurable function, so they said that the conjecture should be true for all "reasonable" functions f. However, although they did not specify it they clearly considered many more functions to be "reasonable" than just our measurable ones. In particular, they included functions that live only on one fixed prime p, e.g., the function that decides whether a group has p-part G_0 for some fixed p-group G_0. In this sense, our formulation is only a special case of the more general conjectures of Cohen and Lenstra. However, as we have seen earlier, including these functions makes it at least highly problematic to speak about "probabilities". In section 5.1.1, we have seen further unsatisfactory effects that such a broad choice of functions entails.

How to interpret formula (6.1)? Perhaps the most intuitive way is the following: We divide out the image of all maps $\phi \in \mathrm{Hom}(\mathbb{Z}^u, G)$ and average over the results. Note that this is equivalent to choosing u elements in G arbitrarily (i.e., uniformly at random) and dividing those out. Hence, the heuristic can be reformulated as follows:

The sequence of class groups within a situation Σ is random with respect to the following stochastic process: Pick a finite abelian group G w.r.t. the Cohen-Lenstra probability, then choose u elements uniformly at random, and divide out those elements.

This idea is due to Cohen and Lenstra [CL84], at least in the case $u = 1$. A similar formulation is the one of Friesen [Fri99], formulated for quadratic function field extensions and involving elliptic curves (see below).
How could a geometric object look like that might play the role of the random group G? Two analogies are known:
Firstly, as Cohen and Lenstra pointed out, in the case of real quadratic number fields the set of all reduced binary quadratic forms having the right discriminant carries a "group-like" structure. This set breaks up into cycles, all of the same length (given by the regulator), and the number of cycles equals the class number. Viewing the principal cycle as a "subgroup", we get operations that are very similar to dividing out a cyclic subgroup.

This has been properly formalized by Lenstra, Schoof and others in terms of Arakelov theory. They defined the *Arakelov class group* Pic_K^0 of a number field K, and showed that there is a natural exact sequence

$$0 \to T^0 \to \mathrm{Pic}_K^0 \to H_K \to 0,$$

where T^0 is the cokernel of the natural homomorphism from the unit group $\mathcal{O}_K^*$ into the group $(\mathbb{R}^u)^0$ of degree 0-divisors defined at the infinite primes, i.e., T^0 is a real torus of dimension u (see [Sch08] for details).
Concurrently, a theory of binary forms corresponding to higher number fields is emerging for extensions of degree ≤ 5 — for a brief overview you may have a look into the introduction of [Bha05].
Secondly, Friesen [Fri99, Thm. 2.4] has established the following: If we consider quadratic function fields of the form $\mathbb{F}_q(T, \sqrt{f(T)})$, where f is an irreducible monic polynomial of degree 4 without cubic term, and $\mathrm{char}(\mathbb{F}_q) \neq 2, 3$, then there is a $1-1$ correspondence between such polynomials and pairs (E, P) of non-singular elliptic curves E with $\#E(\mathbb{F}_q)$ even and points P on E such that $\#(E(\mathbb{F}_q)/\mathrm{ord}(P))$ is odd. Under this correspondence the ideal class group is isomorphic to the quotient $E(\mathbb{F}_q)/\langle P\rangle$. The analogy is even clearer than in the former example, but it applies only to a very special case and it is unclear how this correspondence could be generalized.
Let me return to the probabilistic process given above. There is one delicate point in the formulation: Despite of my nomenclature, the Cohen-Lenstra probability measure is not defined on the whole power set of $\mathcal{G}$ (not even the content in the original sense of Cohen and Lenstra is). In particular, it is not defined on one-element sets, unless we define (as Cohen and Lenstra did) every one-element set to have "measure" 0. So the probability measure (or even the content) does not allow us to choose a single group randomly. However, for each $u \geq 1$ we get a u-probability on the set of all finite abelian p-groups, and switching to the global case we obtain a probability measure on the whole power set of $\mathcal{G}$. Hence, for $u \geq 1$ all the problems that we addressed in chapter 5 collapse. I will not prove this assertion formally but instead refer to [CL84]. I only give the key calculation:
By theorem 4.6.9, for a p-group G_p of size n_p, the u-probability that a random p-group is isomorphic to G_p equals

$$P_u(G_p) = n_p^{-u} \frac{1}{\#\mathrm{Aut}(G_p)} \prod_{i=u+1}^{\infty} (1 - p^{-i}).$$

Multiplying up over all primes, we compute the u-probability that a random group is isomorphic to a particular group $G = \prod_{p\in\mathbb{P}} G_p$ of size $n = \prod_{p\in\mathbb{P}} n_p$ as

$$\begin{aligned} P_u(G) &= \prod_{p\in\mathbb{P}} P_u(G_p) \\ &= \prod_{p\in\mathbb{P}} n_p^{-u} \frac{1}{\#\mathrm{Aut}(G_p)} \prod_{i=u+1}^{\infty} (1-p^{-i}) \\ &= n^{-u} \frac{1}{\#\mathrm{Aut}(G)} \prod_{p\in\mathbb{P}} \prod_{i=u+1}^{\infty} (1-p^{-i}) \\ &= n^{-u} \frac{1}{\#\mathrm{Aut}(G)} \prod_{i=u+1}^{\infty} \zeta(i) \\ &\overset{u\geq 1}{>} 0. \end{aligned}$$

So every group is obtained with a positive probability, and this computation already implies that we have absolute convergence and that we may interchange limits, so P_u is a probability measure on $\mathcal{G}$. It is not completely trivial that this probability measure coincides with formula (6.1), but it is true ([CL84, §5]).

Bad primes

Let me comment on the set B of bad primes. In their original work [CM87], [CM90], Cohen and Martinet excluded the primes that divided the degree n of the extension. It is a (proven!) fact that those primes are indeed bad primes. As in the case of imaginary quadratic number fields, this is a consequence of genus theory.
However, there seem to be other bad primes. Eventually Cohen and Martinet noticed [CM94] that growing numerical evidence seemed not to support their predictions, particularly for $p = 2$ in the case of cubic extensions. Instead, they proposed to take the larger set $B := \{p \in \mathbb{P} \mid \gcd(p, |G_0|) > 1\}$, where G_0 is the shared Galois group of the situation. Indeed, numerical data supports the assumption that those primes are bad primes. However, this is still not the end. Gerth noticed ([Ger89], see also [Ger90]) that the distributions of class groups might be influenced by the existence of roots of unity in the base field. Although his results did not directly contradict the Cohen-Martinet choice of B, they led Malle [Mal08] to extend the set B to be

$$B := \left\{ p \in \mathbb{P} \;\middle|\; \begin{array}{c} \gcd(p, |G_0|) > 1, \text{ or } K_0 \\ \text{contains the } p\text{-th roots of unity} \end{array} \right\}.$$

In the same paper, Malle proposed modified probability distributions for the new bad primes (adding to work of Gerth [Ger89] and Wittmann [Wit05]), and tested it extensively by computer calculations, which gave good support for his modified predictions.

As you have noticed, the debate about the set B is still vivid and not finished. However, I want to emphasize that it is undoubted by the experts in the field (and strongly supported by computer tests) that the Cohen-Lentra heuristic is still true in principle — only the finite set of exceptions needs to be identified.

As I already mentioned, for bad primes there exist modified conjectures which seem to fit the data and which are still equidistribution statements — only for (modestly) modified probability distributions. They are still conjectural, but for ramified primes there are some partial results proven. Recently the most general results are obtained by Fouvry and Klüners [FK07] in the case of quadratic extensions, and by Wittmann [Wit05] in the case of general cyclic extensions – the latter paper relying heavily on the extensive work of Gerth (especially [Ger86] and [Ger89], but also [Ger82], [Ger84], [Ger87], [Ger90], [Ger05], only to mention the most important of his papers). The case of bad primes is in some sense easier because genus theory allows to formulate a probability distribution for each rank individually by restricting on regulators with a fixed number of prime divisors, which yields probability distributions on finite sets.

Real quadratic number fields

I give this as an example case. As long as $u > 0$ the formulas can be translated in a straightforward way from section 4.6.5 by multiplying up the local probabilities. Since all series and products converge absolutely, we may change the order of summation and multiplication just as we like. The only open question is to find the set B of bad primes (cf. the discussion above). For $u = 0$, we get exactly the same probabilities as in the imaginary quadratic case, up to correction terms because of a different set B of bad primes.

So considering the situation $\Sigma = (\mathbb{Q}, \mathbb{Z}/2\mathbb{Z}, (2, 0))$, we specialize on real quadratic number fields. The rank of the unit group is $u = 1$, so the class group is predicted to behave like a finite abelian group with one random element divided out.

Clearly, 2 is a bad prime, and we suspect that it is the only bad prime, so $B = \{2\}$.

Since $u > 0$, we get a probability measure (!) on the whole power set of $\mathcal{G}^{\neq 2}$. The probability for a finite group G of odd order n to be obtained is

$$Pr(G) = \frac{1}{n}\left(\prod_{p\in\mathbb{P}\setminus\{2\}}\prod_{s=2}^{\infty}(1-p^{-s})\right) w(G), \tag{6.2}$$

where $w(G) = \frac{1}{|\mathrm{Aut}(G)|}$.

Proof. [CL84, Example 5.9]. For formula (6.2), it suffices to multiply up the local probabilities $P_u(G)$ (for $u = 1$) for all $p \neq 2$, which are given in theorem 4.6.9 of this thesis. □

In particular, for G the trivial group we get a probability of $Pr(0) \approx 75.446\%$ that the odd part of a class group in this situation is trivial. Further values are $Pr(\mathbb{Z}/3\mathbb{Z}) \approx 12.574\%$, $Pr(\mathbb{Z}/5\mathbb{Z}) \approx 3.772\%$, $Pr(\mathbb{Z}/7\mathbb{Z}) \approx 1.796\%$, $Pr(\mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z}) \approx 0.175\%$ and $Pr(\mathbb{Z}/9\mathbb{Z}) \approx 1.397\%$. This already shows that the class groups (at least the odd parts) have a very strong tendency to be of rank 1 and exponent 1. Keep in mind, however, that each other group still accounts for a small but positive fraction.
The gap between our conjectural "knowledge" and provable theorems is immense. In fact, it was a conjecture of Gauß in his famous Disquisitiones Arithmeticae [Gau89] in 1801 that there are infinitely many real quadratic number fields with class number 1, and this has remained unproven since then. It is not even known whether there are infintely many *arbitrary* number fields (over $\mathbb{Q}$) with class number 1. On the other hand, at least if we suppress the 2-part of the class group, then the Cohen-Lenstra conjectures tell us that *the majority* of all class groups of real quadratic number fields is trivial!

Theoretical evidence

The Cohen-Lenstra conjectures for number fields have turned out to be very hard to prove. However, there are some theoretical results which support the conjectures.

The Spiegelungssatz

One of the historically first results concerns Leopoldt's Spiegelungssatz ("reflection theorem"), which gives for any (for simplicity odd) prime p a pairing of certain number fields, such that the ranks of the p-parts of the class groups of paired fields are closely related. Since a precise formulation would require a lot of additional notation, I will not give the general statement and refer to [Lee02] for an overview without proofs, or to [Lon77] for a complete treatment; for a treatment of the $p = 2$ case, see also [DP70].

Instead I give as an example Scholz's theorem, which is the $p = 3$ case of the Spiegelungssatz. This case was studied by Dutarte [Dut84], only one year after Cohen-Lenstra had published their heuristic: Let $K = \mathbb{Q}(\sqrt{m})$, for a positive integer m. For simplicity, we assume m is a square-free positive integer and $3 \nmid m$. Then the reflection field of K is $K' := \mathbb{Q}(\sqrt{-3m})$, and the 3-ranks r_K and $r_{K'}$ of the class groups of K and K' satisfy

$$r_{K'} - r_K \in \{0, 1\}.$$

Assuming some very natural equidistribution property concerning the behaviour of fundamental units (which was removed later on by Klüners), Dutarte computed the conditional probability that r_K and $r_{K'}$ are equal, provided that r_K is known, as

$$Pr(r_{K'} = r_K \mid r_K = a) = \frac{1}{3^{a+1}}.$$

Now let us summarize: The Cohen-Lenstra heuristic gives us a probability distribution for r_K, where K runs through all real quadratic number fields (if we neglect the technical condition that $3 \nmid m$). By Dutarte's result, we may compute the probability distribution for $r_{K'}$, where K' runs through all imaginary quadratic number fields of the form $\mathbb{Q}(\sqrt{-3m})$. Combining both probabilities, we obtain a probability distribution for $r_{K'}$, and it turns out that this is exactly the probability that the Cohen-Lenstra heuristic predicts for imaginary quadratic number fields. Although a different result would not immediately contradict the Cohen-Lenstra heuristic, since K' does not run through *all* imaginary quadratic number fields, it still indicates that the Cohen-Lenstra heuristics for several number fields are compatible with each other.
In the sequel, several similar results for other special cases of the Spiegelungssatz were obtained (e.g., [Ger01] for $p = 2$) until finally Lee proved that the general Spiegelungssatz is compatible with the Cohen-Lenstra heuristic in [Lee02], still making a similar equidistribution assumption as Dutarte. For the $p = 2$ and $p = 3$ case, this assumption could be removed by Fouvry and Klüners [FK09], and Belabas [Bel99], [Bel04], respectively.

Special cases
In the case of cubic field extensions, we have an additional tool: There is a discriminant-preserving correspondence between isomorphism classes of cubic number fields and certain equivalence classes of integral binary cubic forms, established by Delone and Faddeev [DF64]. Apparently unaware of this work, Davenport and Heilbronn rediscovered this connection [DH69], [DH71], and

were able to deduce formulas for the average size of the 3-part of the class number of *quadratic* number fields (by a theorem of Hasse [Has30, Satz 8], the 3-parts of the class groups of quadratic number fields are linked with cubic number fields) by counting binary cubic forms. Actually, the counting part was already done by Davenport in [Dav51a] and [Dav51b].
Their methods were extended by Datskovsky and Wright [DW88] to arbitrary global base fields K_0 (including function fields). Instead of binary cubic forms, they counted cubic extensions (which is essentially the same, because a cubic extension of $\mathbb{Q}$ can be described by an equivalence class of binary cubic forms $F(x,y) = aX^3 + bX^2Y + cXY^2 + dY^3$ by adjoining to $\mathbb{Q}$ a root of the polynomial $F(x,1)$ – this gives a bijection which preserves discriminants). In order to count such cubic extensions, they used ζ-functions methods invented by Shintani [Shi75]. For the case $K_0 = \mathbb{Q}$, Belabas, Bhargava, and Pomerance [Bel99], [Bel04], [BBP09] proved some upper bounds for the rate of convergence.

Following ideas of Wright and Yukie [WY92], Bhargava could count quartic extensions of $\mathbb{Q}$, for which there exists a connection to the 2-parts of the class groups of cubic number fields. In this way, he could prove some formulas about (in particular, the average size of) the 2-part of class groups of cubic extension of $\mathbb{Q}$ [Bha05]. Explicit error bounds are given in [BBP09].

Summarizing, we see that for some very specific cases partial results could be proven, but unfortunately the methods all rely on specific tools that are only available in these settings and cannot be easily transferred to other cases.

The methods used in these special cases show that the Cohen-Lenstra conjectures for number fields are closely related to the number of number fields of a certain type over a fixed base field. Malle [Mal02], [Mal04] has given precise conjectures on the asymptotics, and in fact, Klüners could prove some of the conjectures modulo the Cohen-Lenstra heuristics [Klü06]. (He also disproved some of Malle's conjectures by counterexamples [Klü05].) This may be seen as another application of the Cohen-Lenstra heuristic for number fields.

6.2 A Fiat-Shamir protocol based on real quadratic number fields

There are also some algorithms which rely on the hypothesis that the class groups obey the Cohen-Lenstra predictions. In particular, there are some cryptographic protocols which work over real quadratic number fields with

large discriminants but small class groups. Since the Cohen-Lenstra heuristic predicts that asymptotically 75.4 . . . % of all real quadratic number fields have class groups with trivial odd part, it ensures that there exist such number fields (if we assume further that the 2-part of the class group does not interfere in an unexpected way). As I mentioned earlier, this existence is not proven yet. So if the Cohen-Lenstra heuristic should turn out to be wrong (which none of us expects), then the algorithms might work over the empty set.
As an example for such an algorithm, I will briefly sketch a variant of the Fiat-Shamir identification protocol (FS) that was proposed by Buchmann, Maurer and Möller in [BMM00].
Let me first describe a generalized version of the FS protocol. I will follow the description of Buchmann, Maurer and Möller, simplifying where possible. For the basic version of FS, see [FS87]; an improved and more detailed version can be found e.g. in [BMM00]. There you will also find a description on how one can derive a digital signature scheme from FS.
The goal of the FS protocol is that one party, called the *prover*, convinces the other party, called the *verifier*, of his knowledge of a private key without revealing any relevant information about this key.
In the setup phase of the protocol, the two parties agree on two abelian groups G and H, and on an isomorphism $\varphi : G \to H$ which must be a one-way-function, i.e., it is possible to compute φ in polynomial time, but no way is known to compute the inverse φ^{-1} in polynomial time. The prover selects a group element $g \in G$ as his private key, and publishes $h := \varphi(g)$ as his public key.
Now the FS identification protocol works as follows:

(i) (Commitment and Witness) The prover randomly chooses a *commitment* $g_0 \in G$ and computes the *witness* $h_0 = \varphi(g_0)$. He sends the witness h_0 to the verifier.

(ii) (Challenge) The verifier selects a *challenge bit* $b \in \{0, 1\}$ and sends it to the prover.

(iii) (Response) The prover computes the *response* $r = g^b g_0$ and sends it to the verifier.

(iv) (Verification) The verifier checks whether $h^b h_0 = \varphi(r)$.

If the prover does not know the private key then it can be shown that it is impossible for him to give the correct answer to both challenges unless he can invert φ (which we assume is impossible). Hence, the probability to convict him is 50%. The probability can be increased by repeating the protocol.

For our purpose, we choose a real quadratic number field F with large discriminant and small class group, and maximal order $O \subseteq F$. Let $G := F^*$ be the group of invertible elements of F, let

$$H := \{\alpha O \mid \alpha \in F^*\},$$

and let $\varphi : G \to H; \alpha \mapsto \alpha O$. (Note that φ is not bijective. This problem can be solved by choosing the generator α of αO for which the euclidean length a of the logarithmic embedding is minimal. For details, see [BMM00].) The result is output in *standard form*

$$\alpha O = q\left(\mathbb{Z} + \frac{b+\sqrt{D}}{2a}\right),$$

where $a, b \in \mathbb{Z}$, $q \in \mathbb{Q}$, $a, q > 0$, $c = \frac{b^2-D}{4a} \in \mathbb{Z}$ and $\gcd(a, b, c) = 1$ (cf. [BB94, 5.1]).
In order for the FS-protocol to be secure, inverting φ must be intractable. Inverting φ is known as *principal ideal problem*. The two most efficient methods known for solving this problem are the babystep-giantstep algorithm and the index calculus method ([BB94]). The former algorithm takes $\Omega(\min(\sqrt{a}, \sqrt{R}))$ time, where R is the regulator of O, and a is the length of the logarithmic embedding of the generator α, as above. The index calculus method has running time $\exp(\Omega(\sqrt{\log(D)\log\log(D)}))$, where D is the discriminant of F.
One can show that if R is big then a is also big for a high proportion of all orders of F ([BMM00]), so we only need to make the regulator R sufficiently large.
The analytic class number formula relates the regulator to the class group. For real quadratic number fields, it reads:

$$2hR = L(1, \chi)\sqrt{D},$$

where L is the Dirichlet L-series and χ is the Kronecker symbol. It is well-known that $L(1, \chi)$ grows asymptotically much slower than $\sqrt{D}$ (only polylogarithmical), hence we have $2hR \approx \sqrt{D}$. Since we have chosen fields with small class number h (independent of D), we see that R becomes arbitrarily large, and hence the FS protocol is resistant against the babystep-giantstep algorithm.

6.3 Function fields

The first steps in order to transfer the Cohen-Lenstra heuristics to function fields came from Friedman and Washington [FW89]. Their approach was

not quite correct. (Essentially they thought that they might take the group $\mathrm{GL}(2g, \mathbb{Z}_p)$ instead of $\mathrm{Sp}_{2g}(\mathbb{Z}_p)$, which turned out to be wrong – cf. the description below.) The first one to correct them was seemingly Yu [Yu96], although his work has never been published. Friesen [Fri99] was probably the first one to publish results citing Yu's work. For the sake of clarity, I will describe only the case of quadratic extensions in detail; generalizations are possible (e.g. [Ach06]), and considerable parts of the conjectures have been proven in the general case as well [Pac04], [Ach06]. Also, I will restrict myself to the local heuristic. I will follow closely the text of Friedman and Washington, correcting their approach in the way of Yu where necessary.

We want to transfer the (imaginary quadratic) number field heuristic to (imaginary quadratic) function fields. We replace $\mathbb{Q}$ by the projective line P^1 over a fixed finite field $\mathbb{F}$ of characteristic $l \in \mathbb{P}$. Instead of an imaginary quadratic field, take a double cover C of P^1 defined over $\mathbb{F}$ and ramified at ∞ (i.e., for $l \neq 2$, a hyperelliptic curve, given by an equation $y^2 = f(x)$ for some separable polynomial f over $\mathbb{F}$ of odd degree). Fix an odd prime $p \neq l$, and replace the class group of the number field by the 0-Picard group of divisor classes of degree 0 defined over $\mathbb{F}$. Denote the p-part of this group by C_p. We order the function fields by the genus g_C. In this case, the Cohen-Lenstra conjecture would predict that for any finite abelian p-group H,

$$\lim_{g\to\infty} \frac{\sum_{\substack{C, g_C \leq g \\ C_p \cong H}} 1}{\sum_{C, g_C \leq g} 1} \stackrel{??}{=} \frac{1}{\#\mathrm{Aut}(H)} \prod_{i=1}^{\infty} (1 - p^{-i}). \tag{6.3}$$

Here C ranges over all hyperelliptic curves which are defined over $\mathbb{F}$, which are ramified at ∞, and which have genus $g_C \leq g$, and we always assume C to be complete, nonsingular and absolutely irreducible.

The distinguishing feature of the function field case is that the Frobenius map is available. Let Fr be the Frobenius map with respect to $\mathbb{F}$ and let $T_p(C)$ be the p-adic Tate module of C. We have $T_p(C) \cong \mathbb{Z}_p^{2g}$ (non-canonically). Then $M_C := I - \mathrm{Fr}$ is an endomorphism of $T_p(C)$ whose cokernel is isomorphic to C_p. Thus the groups whose distribution we seek all appear as $\mathrm{coker}(M_C)$, where C is a variable hyperelliptic curve.

In which matrix space does M_C live? Since $\mathrm{Fr} = I - M_C$ is a symplectic similitude with respect to the Weil pairing, M_C may be considered (non-canonically) as an element of $\mathrm{Sp}_{2g}(\mathbb{Z}_p)$ (see [Ach06] for details). And how is Fr distributed in this space? The most naive possible answer turns out to be correct: It is equidistributed by a result of Katz [KS99, Thm. 9.7.13]. Hence a random class group is obtained by choosing a random matrix A in $\mathrm{Sp}_{2g}(\mathbb{Z}_p)$ (with respect to the Haar measure, and for $g \to \infty$), and taking the cokernel of $I - A$.

Note the analogy with the basic Cohen-Lenstra probability measure for quadratic imaginary number fields, where a random class group is obtained by taking a random matrix in $\mathbb{Z}_p^{n\times n}$ (with respect to the Haar measure, and for $n \to \infty$) and taking its cokernel (cf. section 2.2.3). However, it turns out that the results are only analogous, not identical. The probability distribution is modified, *so equation* (6.3) *is false* and its right hand side needs to be replaced by some twisted Cohen-Lenstra probabilities.
I should add that the fact that the Frobenius is equidistributed is also proven with explicit error bounds [Ach08]. Moreover, Achter shows in [Ach06] that the equidistribution property is even true for every "sufficiently general" sequence of quadratic function fields.
It remains to compute the twisted probabilities. They were computed by Achter [Ach06], and Gekeler gives in [Gek06] explicit error bounds for the special case of groups of rank 2 (which is the interesting case for elliptic curves). This is again a case where the unfortunate unawareness between the number theory group and the combinatoric group has produced double statements: The formulas are already contained in Fulman's thesis [Ful97] and in a summary paper on the case of symplectic groups [Ful00a, Thm. 1]. Fulman computes not only the probabilities, but also provides a cycle index description and a description by Markov chains, and deduces formulas for the average size and rank of such groups and for some higher moments of these values.

6.4 Modules over group rings

There are various ways to generalize the Cohen-Lenstra heuristic. Greither proposed to fix a finite abelian p-group G_0 and consider all finite $\mathbb{Z}_p[G_0]$-modules M, where $\mathbb{Z}_p[G_0]$ denotes the group ring of G_0. He proves [Gre98], [Gre00] that

$$\sum_{\substack{M \text{ finite} \\ \mathbb{Z}_p[G_0]\text{-module}}} \frac{1}{\#\mathrm{Aut}_{\mathbb{Z}_p[G_0]}(M)} = \sum_{\substack{G \text{ finite} \\ \text{abelian } p\text{-group}}} \frac{1}{\#\mathrm{Aut}(G)} = w(\mathcal{G}_p), \qquad (6.4)$$

where the right hand side is just the standard overall Cohen-Lenstra weight. In other words, considering finite $\mathbb{Z}_p[G_0]$-modules is a refinement of considering finite abelian p-groups. In the former paper, Greither gives some indication that for certain families of number fields the p-class group of an extension $K/\mathbb{Q}$ should be considered as a $\mathbb{Z}_p[G_0]$-module.
Let me add that, following unpublished ideas of Lenstra and de Smit, Greither proves in [Gre00] that equation (6.4) is also valid for non-abelian finite

p-groups G_0 if we restrict M to be finite and cohomologically trivial (i.e., in this case, of projective dimension ≤ 1).
These ideas enable me to conclude my thesis with a highly speculative and not yet sufficiently formalized conjecture, which might lead to a vast generalization of the Cohen-Lenstra heuristic for number fields:
As an example, consider the following sequence of number fields: Fix an odd $p \in \mathbb{P}$, and consider the following sequence of fields K_r, where r runs through all primes in $\mathbb{P}_1 := \{r \in \mathbb{P} \mid r \equiv 3 \bmod 4 \text{ and } r \equiv 1 \bmod p\}$. We let K_r be the unique abelian extension of $\mathbb{Q}$ of degree $2p$ and conductor r. (The field K_r is then always imaginary.) Then the p-part $H_p^-(K_r)$ of the minus class group $H^-(K_r)$ of K_r may be considered as a $\mathbb{Z}_p[G_r]$-module, where G_r is the subgroup of order p in $\mathrm{Gal}(K_r/\mathbb{Q})$. Then a way of extending the Cohen-Lenstra heuristic would be:

6.4.1 Conjecture. *Consider the following stochastic process: For each $r \in \mathbb{P}_1$ choose randomly a module M over $\mathbb{Z}_p[G_r]$ according to the probability measure*

$$P(M) = \frac{1}{w(\mathcal{G}_p)} \frac{1}{\#\mathrm{Aut}_{\mathbb{Z}_p[G_r]}(M)},$$

where the proportionality factor $\frac{1}{w(\mathcal{G})}$ does not depend on G_r (and hence, K_r). Now consider the sequence $(H_p^-(K_r))_{r\in\mathbb{P}_1}$. The group $H_p^-(K_r)$ may be viewed as a $\mathbb{Z}_p[G_r]$-module.
Then both sequences are stochastically indistinguishable.

6.4.2 Remark.

- Note that the conjecture is not precise. One would need to specify what "stochastically indistinguishable" means in this context. But it would probably at least include the following: Let us fix an $r_0 \in \mathbb{P}_1$. Whenever we have for all r a morphism from the set of all $\mathbb{Z}_p[G_r]$-modules to the set of all $\mathbb{Z}_p[G_{r_0}]$-modules, for various r, so that we can map all the modules to $\mathbb{Z}_p[G_{r_0}]$, then the images of both sequences are stochastically indistinguishable in $\mathbb{Z}_p[G_{r_0}]$. Further research would be necessary in order to find out whether this notion is useful, or whether we should replace it by something else.

- The conjecture can be transferred to many other sequences of field extensions over a fixed ground field. We would only need to specify what G_r is in the general setting. Opposed to the classical Cohen-Lenstra conjectures for number fields, we are not restricted to a fixed Galois group.

- Note that this conjecture would imply the classical (local) Cohen-Lenstra conjecture. This would be the case where all Galois groups are isomorphic.

Summarizing, the conjecture is remarkable, because M is a module over different rings for each number field K_l. On the one hand, it looks like a step backwards, because it is not clear what the meaning of probability should be in this context. On the other hand, if it could be rigidly formalized, it would allow us to formulate conjectures of Cohen-Lenstra types for much more general sequences of class groups. This new perspective on the Cohen-Lenstra heuristic might also be useful when attention turns to non-abelian field extension, for which only little numerical evidence is available so far.

Nomenclature

$\#M$	cardinality of M 7
$\lvert M\rvert$	cardinality of M 7
$\bigoplus$	outer direct sum 109
$\mathcal{A}$	subalgebra of $\mathcal{G}$ induced by the projections π_p 109
$\mathrm{Aut}(G)$	group of automorphisms of G 10
$\mathbb{C}$	complex numbers 7
$\mathcal{D}$	set of derivations, cf. 3.1, except for chapter 5 34
$D(E_1, \ldots, E_r; F_1, \ldots, F_s)$	cf. 5.3.4 119
D_K	discriminant of K 137
$\mathcal{D}$	in chapter 5: cf. 5.3.4 119
$\exp_{uni}(G)$	uniform exponent of the group $G \in \mathcal{G}$ 113
$\mathrm{expon}(t)$	exponent associated to the tuple t, cf. 3.2.2 39
$\exp_p(G)$	p-adic exponent of the p-group G 8
$\mathbb{F}_q$	finite field of q elements 7
$\mathcal{G}$	set of isomorphism classes of finite abelian groups 8
$\mathrm{GL}(n, p)$	general linear group over $\mathbb{F}_p$ 92
$\mathcal{G}_p$	set of all finite abelian p-groups 8
$\mathcal{G}_{\mathcal{P}}$	set of partitions, $\cong \mathcal{G}_p$, cf. 1.2.8, 1.2.9 9

Bibliography

[Ach06] Jeffrey D. Achter, *The distribution of class groups of function fields*, Journal of Pure and Applied Algebra **204** (2006), no. 2, 316–333.

[Ach08] Jeffrey D. Achter, *Results of Cohen-Lenstra type for quadratic function fields*, Computational arithmetic geometry, Contemp. Math., vol. 463, Amer. Math. Soc., Providence, RI, 2008, pp. 1–7. MR 2459984

[AE04] George E. Andrews and Kimmo Eriksson, *Integer partitions*, Cambridge University Press, Cambridge, 2004. MR 2122332 (2006b:11125)

[And76] George E. Andrews, *The theory of partitions*, Addison-Wesley Publishing Co., Reading, Mass.-London-Amsterdam, 1976, Encyclopedia of Mathematics and its Applications, Vol. 2. MR 0557013 (58 #27738)

[BB94] Ingrid Biehl and Johannes Buchmann, *Algorithms for quadratic orders*, Mathematics of Computation 1943–1993: a half-century of computational mathematics (Vancouver, BC, 1993), Proc. Sympos. Appl. Math., vol. 48, Amer. Math. Soc., Providence, RI, 1994, pp. 425–449. MR 1314882 (95m:11146)

[BBP09] Karim Belabas, Manjul Bhargava, and Carl Pomerance, *Error estimates for the Davenport-Heilbronn theorems*, Duke Math. Journal (to appear, 2009).

[BDS94] Curtis Bennett, Kathy J. Dempsey, and Bruce E. Sagan, *Partition lattice q-analogs related to q-Stirling numbers*, J. Algebraic Combin. **3** (1994), no. 3, 261–283. MR 1285496 (95h:05014)

[Bel99] Karim Belabas, *On the mean* 3*-rank of quadratic fields*, Compositio Math. **118** (1999), no. 1, 1–9. MR 1705974 (2000g:11102)

[Bel04] ———, *Corrigendum: "On the mean 3-rank of quadratic fields" [Compositio Math.* **118** *(1999), no. 1, 1–9; mr1705974]*, Compos. Math. **140** (2004), no. 5, 1221. MR 2081161 (2005b:11176)

[Ben38] Frank Benford, *The law of anomalous numbers*, Proceedings of the American Philosophical Society **78** (1938), no. 4, 551–572.

[Bha05] Manjul Bhargava, *The density of discriminants of quartic rings and fields*, Ann. of Math. (2) **162** (2005), no. 2, 1031–1063. MR 2183288 (2006m:11163)

[BMM00] Johannes Buchmann, Markus Maurer, and Bodo Möller, *Cryptography based on number fields with large regulator*, J. Théor. Nombres Bordeaux **12** (2000), no. 2, 293–307, Colloque International de Théorie des Nombres (Talence, 1999). MR 1823187 (2002e:11180)

[BT24] Stefan Banach and Alfred Tarski, *Sur la décomposition des ensembles de points en parties respectivement congruentes*, Fund. Math. (1924), no. 6, 244–277.

[CF86] J. W. S. Cassels and A. Fröhlich (eds.), *Algebraic number theory*, London, Academic Press Inc. [Harcourt Brace Jovanovich Publishers], 1986, Reprint of the 1967 original. MR 911121 (88h:11073)

[CL84] H. Cohen and H. W. Lenstra, Jr., *Heuristics on class groups of number fields*, Number theory, Noordwijkerhout 1983 (Noordwijkerhout, 1983), Lecture Notes in Math., vol. 1068, Springer, Berlin, 1984, pp. 33–62. MR 756082 (85j:11144)

[CM87] Henri Cohen and Jacques Martinet, *Class groups of number fields: numerical heuristics*, Math. Comp. **48** (1987), no. 177, 123–137. MR 866103 (88e:11112)

[CM90] ———, *Étude heuristique des groupes de classes des corps de nombres*, J. Reine Angew. Math. **404** (1990), 39–76. MR 1037430 (91k:11097)

[CM94] ———, *Heuristics on class groups: some good primes are not too good*, Math. Comp. **63** (1994), no. 207, 329–334. MR 1226813 (94i:11087)

[Coh85] Henri Cohen, *On the p^k-rank of finite abelian groups and Andrews' generalizations of the Rogers-Ramanujan identities*, Nederl. Akad.

Wetensch. Indag. Math. **47** (1985), no. 4, 377–383. MR MR820930 (87g:20091)

[Dav51a] H. Davenport, *On the class-number of binary cubic forms. I*, J. London Math. Soc. **26** (1951), 183–192. MR 0043822 (13,323e)

[Dav51b] ———, *On the class-number of binary cubic forms. II*, J. London Math. Soc. **26** (1951), 192–198. MR 0043823 (13,323f)

[Deu41] Max Deuring, *Die Typen der Multiplikatorenringe elliptischer Funktionenkörper*, Abh. Math. Sem. Hansischen Univ. **14** (1941), 197–272. MR 0005125 (3,104f)

[DF64] B. N. Delone and D. K. Faddeev, *The theory of irrationalities of the third degree*, Translations of Mathematical Monographs, Vol. 10, American Mathematical Society, Providence, R.I., 1964. MR 0160744 (28 #3955)

[DH69] H. Davenport and H. Heilbronn, *On the density of discriminants of cubic fields*, Bull. London Math. Soc. **1** (1969), 345–348. MR 0254010 (40 #7223)

[DH71] ———, *On the density of discriminants of cubic fields. II*, Proc. Roy. Soc. London Ser. A **322** (1971), no. 1551, 405–420. MR 0491593 (58 #10816)

[DP70] Pierre Damey and Jean-Jacques Payan, *Existence et construction des extensions galoisiennes et non-abéliennes de degré* 8 *d'un corps de caractéristique différente de* 2, J. Reine Angew. Math. **244** (1970), 37–54. MR 0280466 (43 #6186)

[Dut84] Philippe Dutarte, *Compatibilité avec le Spiegelungssatz de probabilités conjecturales sur le p-rang du groupe des classes*, Number theory (Besançon), 1983–1984, Publ. Math. Fac. Sci. Besançon, Univ. Franche-Comté, Besançon, 1984, pp. Exp. No. 4, 11. MR 803700 (86m:11103)

[DW88] Boris Datskovsky and David J. Wright, *Density of discriminants of cubic extensions*, J. Reine Angew. Math. **386** (1988), 116–138. MR 936994 (90b:11112)

[Eis47] Ferdinand Eisenstein, *Neue Theoreme der höheren Arithmetik*, J. Reine Angew. Math. **35** (1847), 117–136.

[FK07] Étienne Fouvry and Jürgen Klüners, *On the 4-rank of class groups of quadratic number fields*, Invent. Math. **167** (2007), no. 3, 455–513. MR 2276261 (2007k:11187)

[FK09] ———, *On the Spiegelungssatz for the 4-rank*, preprint, www.math.u-psud.fr/ fouvry/prepublications/Spiegelungssatz.pdf, 2009.

[Fri99] Christian Friesen, *A special case of Cohen-Lenstra heuristics in function fields*, Number theory (Ottawa, ON, 1996), CRM Proc. Lecture Notes, vol. 19, Amer. Math. Soc., Providence, RI, 1999, pp. 99–105. MR 1684595 (2000b:11129)

[Fri00] ———, *Class group frequencies of real quadratic function fields: the degree 4 case*, Math. Comp. **69** (2000), no. 231, 1213–1228. MR 1659859 (2000j:11174)

[FS87] Amos Fiat and Adi Shamir, *How to prove yourself: practical solutions to identification and signature problems*, Advances in cryptology—CRYPTO '86 (Santa Barbara, Calif., 1986), Lecture Notes in Comput. Sci., vol. 263, Springer, Berlin, 1987, pp. 186–194. MR 907087 (88m:94023)

[Ful97] Jason Fulman, *Probability in the classical groups over finite fields: symmetric functions, stochastic algorithms and cycle indices*, Ph.D. thesis, Harvard University, 1997.

[Ful99] ———, *A probabilistic approach toward conjugacy classes in the finite general linear and unitary groups*, J. Algebra **212** (1999), no. 2, 557–590. MR 1676854 (2000c:20072)

[Ful00a] ———, *A probabilistic approach to conjugacy classes in the finite symplectic and orthogonal groups*, J. Algebra **234** (2000), no. 1, 207–224. MR 1799484 (2002j:20094)

[Ful00b] ———, *The Rogers-Ramanujan identities, the finite general linear groups, and the Hall-Littlewood polynomials*, Proc. Amer. Math. Soc. **128** (2000), no. 1, 17–25. MR 1657747 (2000h:05229)

[Ful08] ———, Email contact, 2008.

[FW89] Eduardo Friedman and Lawrence C. Washington, *On the distribution of divisor class groups of curves over a finite field*, Théorie des nombres (Quebec, PQ, 1987), de Gruyter, Berlin, 1989, pp. 227–239. MR 1024565 (91e:11138)

[Gau89] Carl Friedrich Gauß, *Disquisitiones arithmeticae*, Leipzig, 1801, Springer, Berlin, 1889.

[Gek06] Ernst-Ulrich Gekeler, *The distribution of group structures on elliptic curves over finite prime fields*, Doc. Math. **11** (2006), 119–142 (electronic). MR 2226271 (2007b:11143)

[Ger61] Murray Gerstenhaber, *On the number of nilpotent matrices with coefficients in a finite field*, Illinois J. Math. **5** (1961), 330–333. MR 0130875 (24 #A729)

[Ger82] Frank Gerth, III, *Counting certain number fields with prescribed l-class numbers*, J. Reine Angew. Math. **337** (1982), 195–207. MR 676052 (84c:12002)

[Ger84] ———, *The 4-class ranks of quadratic fields*, Invent. Math. **77** (1984), no. 3, 489–515. MR 759260 (85j:11137)

[Ger86] ———, *Densities for certain l-ranks in cyclic fields of degree l^n*, Compositio Math. **60** (1986), no. 3, 295–322. MR 869105 (88f:11110)

[Ger87] ———, *Densities for ranks of certain parts of p-class groups*, Proc. Amer. Math. Soc. **99** (1987), no. 1, 1–8. MR 866419 (88b:11067)

[Ger89] ———, *The 4-class ranks of quadratic extensions of certain imaginary quadratic fields*, Illinois J. Math. **33** (1989), no. 1, 132–142. MR 974015 (90b:11114)

[Ger90] ———, *On p-class groups of cyclic extensions of prime degree p of certain cyclotomic fields*, Manuscripta Math. **70** (1990), no. 1, 39–50. MR 1080901 (92a:11127)

[Ger01] ———, *Comparison of 4-class ranks of certain quadratic fields*, Proc. Amer. Math. Soc. **129** (2001), no. 9, 2547–2552 (electronic). MR 1838376 (2002c:11149)

[Ger05] ———, *On 3-class groups of certain pure cubic fields*, Bull. Austral. Math. Soc. **72** (2005), no. 3, 471–476. MR 2199648 (2007g:11143)

[Gre98] Cornelius Greither, *The structure of some minus class groups, and Chinburg's third conjecture for abelian fields*, Math. Z. **229** (1998), no. 1, 107–136. MR 1649330 (99f:11143)

[Gre00] ———, *Galois-Cohen-Lenstra heuristics*, Acta Math. Inform. Univ. Ostraviensis **8** (2000), no. 1, 33–43. MR 1800220 (2001i:11127)

[Hal38] P. Hall, *A partition formula connected with abelian groups*, Comm. Math. Helvetici **11** (1938), no. 1, 126–129.

[Hal50] Paul R. Halmos, *Measure Theory*, D. Van Nostrand Company, Inc., New York, N. Y., 1950. MR 0033869 (11,504d)

[Has30] Helmut Hasse, *Arithmetische Theorie der kubischen Zahlkörper auf klassenkörpertheoretischer Grundlage*, Math. Z. **31** (1930), no. 1, 565–582. MR 1545136

[Jac98] Michael J. Jacobson, Jr., *Experimental results on class groups of real quadratic fields*, ANTS-III: Proceedings of the Third International Symposium on Algorithmic Number Theory (London, UK), Springer-Verlag, 1998, pp. 463–474.

[Klü05] Jürgen Klüners, *A counterexample to Malle's conjecture on the asymptotics of discriminants*, C. R. Math. Acad. Sci. Paris **340** (2005), no. 6, 411–414. MR 2135320 (2005m:11214)

[Klü06] ———, *Asymptotics of number fields and the Cohen-Lenstra heuristics*, J. Théor. Nombres Bordeaux **18** (2006), no. 3, 607–615. MR 2330430 (2008j:11162)

[KS99] Nicholas M. Katz and Peter Sarnak, *Random matrices, Frobenius eigenvalues, and monodromy*, American Mathematical Society Colloquium Publications, vol. 45, American Mathematical Society, Providence, RI, 1999. MR 1659828 (2000b:11070)

[Kun81] Joseph P. S. Kung, *The cycle structure of a linear transformation over a finite field*, Linear Algebra Appl. **36** (1981), 141–155. MR 604337 (82d:15012)

[Lan65] Serge Lang, *Algebra*, Addison-Wesley, 1965.

[Lee02] Yoonjin Lee, *Cohen-Lenstra heuristics and the Spiegelungssatz: number fields*, J. Number Theory **92** (2002), no. 1, 37–66. MR 1880583 (2002j:11130)

[Len08] Johannes Lengler, *A combinatorial interpretation of the probabilities of p-groups in the Cohen-Lenstra measure*, J. Number Theory **128** (2008), no. 7, 2070–2084. MR 2423750

[Lon77] Robert L. Long, *Algebraic number theory*, Marcel Dekker Inc., New York, 1977, Monographs and Textbooks in Pure and Applied Mathematics, Vol. 41. MR 0469888 (57 #9668)

[Mac79] I. G. Macdonald, *Symmetric functions and hall polynomials*, Clarendon Press ; Oxford University Press, 1979 (English).

[Mal02] Gunter Malle, *On the distribution of Galois groups*, J. Number Theory **92** (2002), no. 2, 315–329. MR 1884706 (2002k:12010)

[Mal04] ———, *On the distribution of Galois groups. II*, Experiment. Math. **13** (2004), no. 2, 129–135. MR 2068887 (2005g:11216)

[Mal06] ———, *The totally real primitive number fields of discriminant at most* 10^9, Algorithmic number theory, Lecture Notes in Comput. Sci., vol. 4076, Springer, Berlin, 2006, pp. 114–123. MR 2282919 (2007j:11179)

[Mal08] ———, *Cohen-Lenstra heuristic and roots of unity*, J. Number Theory **128** (2008), no. 10, 2823–2835. MR 2441080

[Meh] Bernd Mehnert, , Ph.D. thesis, Universität des Saarlandes, Germany, – to appear –.

[Mun53] M. E. Munroe, *Introduction to measure and integration*, Addison-Wesley Publishing Company, Inc., Cambridge, Mass., 1953. MR 0053186 (14,734a)

[New81] Simon Newcomb, *Note on the frequency of use of the different digits in natural numbers*, American Journal of Mathematics **4** (1881), no. 1, 39–40.

[Pac04] Allison M. Pacelli, *Abelian subgroups of any order in class groups of global function fields*, J. Number Theory **106** (2004), no. 1, 26–49. MR 2029780 (2004m:11193)

[Pau08] Thorsten Paul, *Statistische Untersuchungen elliptischer Kurven über endlichen Primkörpern*, Master's thesis, Universität des Saarlandes, 2008, http://www.math.uni-sb.de/ag/gekeler/PERSONEN/exDiplom.html.

[RS88] Arunas Rudvalis and Ken-ichi Shinoda, *An enumeration in finite classical groups*, Preprint, Department of Mathematics, U-Mass Amherst, 1988.

[Sch08] René Schoof, *Computing Arakelov class groups*, Algorithmic number theory: lattices, number fields, curves and cryptography, Math. Sci. Res. Inst. Publ., vol. 44, Cambridge Univ. Press, Cambridge, 2008, pp. 447–495. MR MR2467554 (2009k:11212)

[Shi75] Takuro Shintani, *On zeta-functions associated with the vector space of quadratic forms*, J. Fac. Sci. Univ. Tokyo Sect. I A Math. **22** (1975), 25–65. MR 0384717 (52 #5590)

[Sta97] Richard P. Stanley, *Enumerative combinatorics. Vol. 1*, Cambridge Studies in Advanced Mathematics, vol. 49, Cambridge University Press, Cambridge, 1997, With a foreword by Gian-Carlo Rota, Corrected reprint of the 1986 original. MR 1442260 (98a:05001)

[Sto88] Richard Stong, *Some asymptotic results on finite vector spaces*, Adv. in Appl. Math. **9** (1988), no. 2, 167–199. MR 937520 (89c:05007)

[Sto93] ———, *The average order of a matrix*, J. Combin. Theory Ser. A **64** (1993), no. 2, 337–343. MR 1245166 (94j:11094)

[Tat74] John T. Tate, *The arithmetic of elliptic curves*, Invent. Math. **23** (1974), 179–206. MR 0419359 (54 #7380)

[tRW03] Herman te Riele and Hugh Williams, *New computations concerning the Cohen-Lenstra heuristics*, Experiment. Math. **12** (2003), no. 1, 99–113. MR 2002677 (2005d:11183)

[Vos98] V. E. Voskresenskiĭ, *Algebraic groups and their birational invariants*, Translations of Mathematical Monographs, vol. 179, American Mathematical Society, Providence, RI, 1998, Translated from the Russian manuscript by Boris Kunyavski [Boris È. Kunyavskiĭ]. MR 1634406 (99g:20090)

[Wag93] Stan Wagon, *The Banach-Tarski paradox*, Cambridge University Press, Cambridge, 1993, With a foreword by Jan Mycielski, Corrected reprint of the 1985 original. MR 1251963 (94g:04005)

[Was86] Lawrence C. Washington, *Some remarks on Cohen-Lenstra heuristics*, Math. Comput. **47** (1986), no. 176, 741–747.

[Wit05] Christian Wittmann, *p-class groups of certain extensions of degree p*, Math. Comp. **74** (2005), no. 250, 937–947 (electronic). MR 2114656 (2005h:11256)

[WY92] David J. Wright and Akihiko Yukie, *Prehomogeneous vector spaces and field extensions*, Invent. Math. **110** (1992), no. 2, 283–314. MR 1185585 (93j:12004)

[Yu96] Jiu-Kang Yu, *Toward a proof of the Cohen-Lenstra conjecture in the function field case*, preprint, 1996.