SJÄLVSTÄNDIGA ARBETEN I MATEMATIK

MATEMATISKA INSTITUTIONEN, STOCKHOLMS UNIVERSITET

Reimagining the Euler Characteristic: from Polyhedra to Lattices to Categories

av

Alex Nash

2025 - No M2

MATEMATISKA INSTITUTIONEN, STOCKHOLMS UNIVERSITET, 106 91 STOCKHOLM

Reimagining the Euler Characteristic: from Polyhedra to Lattices to Categories

Alex Nash

Självständigt arbete i matematik 30 högskolepoäng, avancerad nivå

Handledare: Gregory Arone

2025

Abstract

We will define the Euler characteristic as a sort of "zero dimensional measure" on the lattice of polyconvex subsets of $\mathbb{R}^d$. This enables us to view the Euler characteristic as an invariant first and foremost, and then later derive the classical alternating sum formula as a computational tool. This change in perspective facilitates the process of generalizing the Euler characteristic to other mathematical settings. We then show how a similar process may be carried out directly on the lattice of finite abstract simplicial complexes with a fixed vertex set. From that, we can draw on the connection between simplicial complexes and posets to present a definition of the Euler characteristic of a finite poset in terms of its Möbius function. Finally, we show how Euler characteristic can be generalised to a certain class of categories, which includes some categories whose classifying space is not equivalent to a finite complex. In particular, we will see examples where the Euler characteristic is not an integer.

Abstract

Vi definierar Euler-karakteristiken som ett slags "nolldimensionellt mått" på gittret av polykonvexa delmängder av $\mathbb{R}^d$. Detta gör det möjligt för oss att först och främst betrakta Euler-karakteristiken som en invariant, och därefter härleda den klassiska alternerande summaformeln som ett beräkningsverktyg. Detta perspektivskifte underlättar processen att generalisera Euler-karakteristiken till andra matematiska sammanhang. Vi visar sedan hur en liknande process kan utföras direkt på gittret av ändliga abstrakta simpliciella komplex med en fixerad hörnmängd. Därefter kan vi, genom kopplingen mellan simpliciella komplex och partialordnade mängder, presentera en definition av Euler-karakteristiken för en ändlig partialordnad mängd uttryckt genom dess Möbius-funktion. Slutligen visar vi hur Euler-karakteristiken kan generaliseras till en viss klass av kategorier, vilket inkluderar några kategorier vars klassificerande rum inte är ekvivalent med ett ändligt komplex. I synnerhet kommer vi att se exempel där Euler-karakteristiken inte är ett heltal.

Contents

1 Introduction

The Euler characteristic originated from Euler's investigation of three-dimensional polyhedra in the 18th century. Euler knew that the number of faces alone was not sufficient to classify an arbitrary three-dimensional polyhedron, which motivated him to consider counting edges and vertices as well [Ric19, Chapter 7]. This led him to define the Euler characteristic of the boundary of a three-dimensional polyhedron P as

$$\chi(\partial P) := V - E + F,$$

where V, E, and F denotes the number of vertices, edges, and faces, respectively. He soon realized that $\chi(\partial P) = 2$ for all known three-dimensional polyhedra, although he was unable to prove that this held in general for all convex three-dimensional polyhedra. Since the 18th century, the Euler characteristic has been an invaluable tool in mathematics. Perhaps the most well-known result is that orientable closed surfaces are entirely characterized by their Euler characteristic, and similarly for the non-orientable closed surfaces. However, it is considerably more difficult to define the Euler characteristic of a surface than it is to define the Euler characteristic of a polyhedron. Even in the case of polyhedra, it is not a priori clear that the Euler characteristic should be an invariant, or what sort of information it might be encoding.

Nevertheless, the Euler characteristic's powerful role as an invariant has motivated others to attempt to define it in more general settings. One such generalization is its definition in terms of homology

$$\chi(C) = \sum_{n=0}^{\infty} (-1)^n \operatorname{rank}(H_n(C)), \tag{1.1}$$

where C is any chain complex for which the direct sum of all homology groups has finite rank. This definition has the advantage of making it abundantly clear that the Euler characteristic is in fact an invariant, but it is still not quite clear what underlying properties are captured by it.

We will take a different approach, instead defining the Euler characteristic as a sort of rigid motion invariant 0-dimensional "measure" of polyconvex sets. Then, from this definition we will derive the standard alternating sum formula for the Euler characteristic. This approach thus flips the Euler characteristic on its head, viewing it first as foremost as a measurement of an intrinsic property of a space, and then afterwards providing a formula to facilitate its computation. Afterwards, we will show how this formula may be generalized to abstract simplicial complexes, then posets, and finally to certain kinds of finite categories. At each step of the process, we will make an effort to motivate these generalizations and hopefully make them seem more natural than they may at first appear.

This thesis will require a fair amount of lattice theory, so we will begin by briefly developing some necessary background material in the first chapter. This will include the basic definitions and properties of lattices. We will also define valuations in this chapter. The primary sources used in this chapter are [Grä11, Mun09, KR97].

In chapter 3, we will introduce the Euler characteristic as a valuation on polyconvex sets. We will follow the general approach set forth in [KR97, Chapters 2, 4,5] by first investigating the simpler case of parallelotopes, before moving to polytopes and finally polyconvex sets. This will have the advantage of illustrating how the Euler characteristic fits into a broader class of invariant valuations on sufficiently nice subsets of $\mathbb{R}^d$. In addition to realizing the Euler characteristic as a valuation, this approach has the added benefit of defining the Euler characteristic as a basis vector for the space of invariant valuations on the lattice of parallelotopes in $\mathbb{R}^d$. A similar result is also true for the lattice of polyconvex subsets of $\mathbb{R}^d$, but we will not cover this in detail, instead referring the interested reader to [Sch13, SW08, Sch86] and [KR97, Chapters 5-9]. We will also prove that this definition of the Euler characteristic satisfies the alternating sum formula for all polytopes, and thus coincides with the standard definition of the Euler characteristic.

In chapter 4, we define the Euler characteristic as a valuation on lattices of abstract simplicial complexes, drawing inspiration from the definition given for parallelotopes in the previous chapter. This follows the general idea set forth in [KR97, Chapter 3] and has the advantage of viewing the Euler characteristic as fitting into a collection of invariant measures, analogous to the case of parallelotopes. However, we do deviate a bit from the presentation in [KR97], taking inspiration from [Mun09] by further developing the theory of finite distributive lattices and their connection to posets. In our presentation, we further alter a few definitions and proofs in order to make the analogy to parallelotopes more apparent. After proving that this definition of a simplicial complex coincides with the Euler characteristic of its geometric realization, we shift our attention to posets. We then define the Euler characteristic of a poset in terms of that poset's Möbius function, as discussed in [Rot64, BBR86].

Finally, in chapter 5, we follow in the footsteps of Tom Leinster [Lei06] and define the Euler characteristic of a finite category. We also present a slight generalization of [Lei06, Theorem 1.4] and further elaborate on the connection between the Euler characteristic and Möbius function of finite posets and finite categories. As Leinster mentioned in [Lei06], this approach also defines the Euler characteristic of categories whose classifying space is not equivalent to a finite complex and thus to categories for which (1.1) does not apply. This, in particular, includes categories whose Euler characteristic is not an integer.

2 Lattice Theory Background

We will spend this short chapter developing the lattice-theoretic concepts that will be necessary for our treatment of the Euler characteristic in chapters 3 and 4. We will primarily be interested in the following three lattices.

(i) The lattice $\mathrm{Par}(d)$ generated by parallelotopes in $\mathbb{R}^d$ with sides parallel to a fixed coordinate system.

(ii) The lattice $\mathrm{Polycon}(d)$ generated by convex sets in $\mathbb{R}^d$.

(iii) The lattice $\mathcal{J}\Delta^d$ of subcomplexes of the standard d-dimensional simplex.

A **lattice** Λ is a partially ordered set for which

$$\sup\{a,b\} \qquad \text{and} \qquad \inf\{a,b\}$$

exist for all $a, b \in \Lambda$. We will follow the lattice theoretic tradition of referring to $\sup\{a,b\}$ as the **join** of a and b, and referring to $\inf\{a,b\}$ as the **meet** of a and b. We will also adopt the standard notation

$$a \vee b := \sup\{a,b\} \qquad \text{and} \qquad a \wedge b := \inf\{a,b\}.$$

A lattice Λ is **distributive** if the operations $\wedge$ and $\vee$ distribute over each other, in the sense that for all $a, b, c \in \Lambda$

$$a \wedge (b \vee c) = (a \wedge b) \vee (a \wedge c) \qquad \text{and} \qquad a \vee (b \wedge c) = (a \vee b) \wedge (a \vee c).$$

Notice that for any lattice Λ, we have

$$a \wedge (b \vee c) \geq (a \wedge b) \vee (a \wedge c) \qquad \text{and} \qquad a \vee (b \wedge c) \leq (a \vee b) \wedge (a \vee c).$$

The first inequality follows from noting that a and $b \vee c$ are each greater than or equal to both $a \wedge b$ and $a \wedge c$. The second inequality may proved in the same manner. A more detailed treatment of this result may be found in [Grä11, I.4.3]. Thus we could have defined distributivity in terms of inequalities rather than equalities. However, in this paper, we will only be interested in distributive lattices.

Notice that for any $a, b \in \Lambda$,

$$a \wedge b = a \quad \leftrightsquigarrow \quad a \leq b \quad \leftrightsquigarrow \quad a \vee b = b. \tag{2.1}$$

Furthermore, if Λ is finite, then Λ has a minimum, denoted by 0, and a maximum, denoted by 1.

Example 2.1. The standard example of a lattice is given by the power set $\mathcal{P}(S)$ of a set S, with

$$A \wedge B := A \cap B \qquad \text{and} \qquad A \vee B := A \cup B.$$

In fact, it is because of this example that the notation $\wedge$ and $\vee$ is used. Compare (2.1) with the well known facts about subsets:

$$A \cap B = A \quad \leftrightsquigarrow \quad A \subseteq B \quad \leftrightsquigarrow \quad A \cup B = B.$$

In this thesis, all of the lattices we examine will be lattices of sets.

A **$\wedge$-semilattice**, or **lower-semilattice**, is a set Γ for which $a \wedge b \in \Gamma$ for all $a, b \in \Gamma$. Clearly every lattice is a $\wedge$-semilattice. There is an analogous notion of $\vee$-semilattices, but we will never discuss them in this thesis. For our purposes, "semilattice" will always mean $\wedge$-semilattice. Suppose that Λ is a lattice and $\Gamma \subset \Lambda$ is a (lower)-semilattice. Then Γ **generates** Λ if every element of Λ can be written as a finite join of elements of Γ.

We will conclude this chapter by introducing the notion of a valuation, which may be thought of as a lattice-theoretic analogue to measures. A **valuation** on a lattice Λ is a real-valued map $\mu : \Lambda \to \mathbb{R}$ such that

(i) $\mu(0) = 0$ (if Λ has a 0).

(ii) $\mu(x \vee y) = \mu(x) + \mu(y) - \mu(x \wedge y)$ for all $x, y \in \Lambda$.

By iterating the above identity, we obtain

$$\mu(x_1 \vee \cdots \vee x_n) = \sum_{1 \le i \le n} \mu(x_i) - \sum_{1 \le i < j \le n} \mu(x_i \wedge x_j) + \cdots + (-1)^n \mu(x_1 \wedge \cdots \wedge x_n).$$

We will often abbreviate the above expansion as

$$\mu(x_1 \vee \cdots \vee x_n) = \sum_{k=1}^{n} (-1)^{k+1} \sum_{1 \le i_1 < \cdots < i_k \le n} \mu\left(x_{i_1} \wedge \cdots \wedge x_{i_k}\right). \tag{2.2}$$

Condition (ii) is sometimes referred to as the **inclusion-exclusion principle**. If the map μ satisfies (2.2) for all $x_1, \ldots, x_n$ then we will say that μ is **n-additive**.

3 Parallelotopes and Polyconvex Sets

In this chapter, we introduce the Euler characteristic as a valuation on the lattice consisting of finite unions of compact convex subsets of $\mathbb{R}^d$. Following the general approach laid out in [KR97, Chapters 4-5], we will begin by considering the much simpler sublattice generated by parallelotopes with sides parallel to a fixed coordinate system. This has the advantage of more clearly illustrating how the Euler characteristic fits into the broader picture of all sufficiently nice valuations on that lattice.

In order to facilitate these endeavours, we will first establish some general results. Recall that if A is a subset of some set X, then the **indicator function** of A is the function $\mathbb{I}_A : X \to \{0, 1\}$ defined by

$$\mathbb{I}_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A. \end{cases}$$

The following lemma follows immediately from the definition.

Lemma 3.1. *Let Λ be a lattice consisting of subsets of a set X, and let $A, B \in \Lambda$.*

(i) $\mathbb{I}_{A\cap B} = \mathbb{I}_A \mathbb{I}_B$.

(ii) $\mathbb{I}_{A\cup B} = \mathbb{I}_A + \mathbb{I}_B - \mathbb{I}_{A\cap B}$.

Note that (ii) implies that for every $x \in X$, the map $A \mapsto \mathbb{I}_A(x)$ is a valuation on Λ.

Given a collection $\mathcal{C}$ of subsets of some set X, a $\mathcal{C}$-**simple function** is a function of the form

$$\sum_{i=1}^{n} \alpha_i \mathbb{I}_{A_i},$$

where each $A_i \in \mathcal{C}$. The **integral** of a Λ-simple function with respect to a valuation μ on Λ is defined by

$$\int \sum_{i=1}^{n} \alpha_i \mathbb{I}_{A_i} \mathrm{d}\mu = \sum_{i=1}^{n} \alpha_i \mu(A_i).$$

Notice that the collection of simple functions over a lattice is a vector space. This space is spanned by the indicator functions. But, as is the case in measure theory, the indicator functions do not in general form a basis of this vector space. (In particular, Lemma 3.1(ii) may be used to find linearly dependent indicator functions.) As such, we will need to show that this definition of integration actually makes sense. This will be done in a slightly more general manner in Theorem 3.2.

We will now introduce a lattice-theoretic analogue of pre-measures. Let Γ be a lower-semilattice of sets. A **pre-valuation** on Γ is a map $\nu : \Gamma \to \mathbb{R}$ such that for all $x_1, \ldots, x_n \in \Gamma$, if $x_1 \vee \cdots \vee x_n \in \Gamma$, then

$$\nu(x_1 \vee \cdots \vee x_n) = \sum_{k=1}^{n} (-1)^{k+1} \sum_{1 \leq i_1 < \cdots < i_k \leq n} \nu\left(x_{i_1} \wedge \cdots \wedge x_{i_k}\right).$$

If Γ is a lattice, then every pre-valuation is a valuation and vice-versa. Notice that the definition of a pre-valuation is required to satisfy the inclusion-exclusion identities for all n, while the definition of a valuation only requires that the inclusion-exclusion principle is satisfied for $n = 2$. This is because it is not a priori clear for lower-semilattices that if the inclusion-exclusion principle holds for $n = 2$ then it holds for all values of n. Indeed, it is possible to have $x_1, \ldots, x_n \in \Gamma$ with $x_1 \vee \cdots \vee x_n \in \Gamma$ but $x_i \vee x_j \notin \Gamma$ for all $i \neq j$. This is explored further in Example 3.5.

We may define integration with respect to pre-valuations in much the same way it is defined for valuations. Namely, the **integral** of a Γ-simple function with respect to the pre-valuation ν is defined by

$$\int \sum_{i=1}^{n} \alpha_i \mathbb{I}_{x_i} \mathrm{d}\nu = \sum_{i=1}^{n} \alpha_i \nu(x_i).$$

Notice that if Γ generates Λ, then we may use Lemma 3.1 to write any indicator function on Λ as a simple function on Γ. Consequently, the pre-valuation ν also determines an integral of Λ-simple functions.

The content of the following theorem and the subsequent corollary are known as Groemer's integral theorem. The underlying argument used in the proof of the theorem is similar to the one in [KR97, Section 2.2], [SW08, Section 14.4] and [Gro78], but we have altered the presentation.

Theorem 3.2 (Groemer's Integral Theorem)**.** *Suppose that Λ is a lattice of sets and Γ generates Λ. Then integration with respect to pre-valuations on Γ is well-defined.*

Proof. Let ν be a pre-valuation on Γ and let Λ be the lattice generated by Γ. Recall that the indicator functions span the vector space of Γ-simple functions. Since integration is defined to be the linear extension of the map acting on indicator functions by $\mathbb{I}_A \mapsto \nu(A)$, it suffices to show that

$$\sum_{i=1}^{n} \alpha_i \mathbb{I}_{A_i} = 0 \qquad \rightsquigarrow \qquad \sum_{i=1}^{n} \alpha_i \nu(A_i) = 0. \tag{3.1}$$

Our approach will to be prove by induction on n that (3.1) holds whenever the A_i are closed under intersections and listed in a non-decreasing order (i.e. $i < j$ implies $A_i \not\supset A_j$). To see that this will be sufficient, let $\sum_{j=1}^{m} \beta_j \mathbb{I}_{B_j}$ be an arbitrary simple function. For $1 \leq i \leq m$, define $A_i := B_i$. Let $A_{m+1}, \ldots, A_n$ run over all the intersections of the B_j that are not included in $\{B_1, \ldots, B_m\}$. Then define coefficients α_i by

$$\alpha_i = \begin{cases} \beta_i & \text{if } 0 \leq i \leq m \\ 0 & \text{if } m < i \leq n \end{cases}.$$

Notice that

$$\sum_{i=1}^{n} \alpha_i \mathbb{I}_{A_i} = \sum_{j=1}^{m} \beta_j \mathbb{I}_{B_j} \qquad \text{and} \qquad \sum_{i=1}^{n} \alpha_i \nu(A_i) = \sum_{j=1}^{m} \beta_j \nu(B_j).$$

We can then reindex the A_i (and α_i) so that the A_i are listed in non-decreasing order while still maintaining the above two equalities. Hence, if (3.1) holds for $\sum_{i=1}^{n} \alpha_i \mathbb{I}_{A_i}$, then it also holds for $\sum_{j=1}^{m} \beta_j \mathbb{I}_{B_j}$.

We will now proceed with the proof by induction. For $n = 1$, we see that if $\alpha \mathbb{I}_A = 0$ then either $\alpha = 0$ or $A = \varnothing$. In either case, $\alpha \nu(A) = 0$.

Now, suppose $n > 1$ and that

$$\sum_{j=1}^{m} \beta_j \mathbb{I}_{B_j} = 0 \qquad \rightsquigarrow \qquad \int \sum_{j=1}^{m} \beta_j \mathbb{I}_{B_j} \mathrm{d}\nu = 0$$

whenever $m < n$ and the B_j are closed under intersections and non-decreasing. Let $\sum_{i=1}^{n} \alpha_i \mathbb{I}_{A_i} = 0$ where the A_i are closed under intersections and non-decreasing.

First consider the case when $\alpha_n = 0$. Since $A_i \not\supset A_n$ for all $1 \leq i < n$, it follows that $\{A_1, \ldots, A_{n-1}\}$ is closed under intersections and non-decreasing; thus, the result follows by induction in this case.

Now suppose $\alpha_n \neq 0$. Then for all $x \notin A_1 \cup \cdots \cup A_{n-1}$,

$$\mathbb{I}_{A_n}(x) = \frac{-1}{\alpha_n} \sum_{i=1}^{n-1} \alpha_i \mathbb{I}_{A_i}(x) = 0,$$

hence $x \notin A_n$ so that $A_n \subseteq A_1 \cup \cdots \cup A_{n-1}$. Consequently,

$$\bigcup_{i=1}^{n-1} (A_i \cap A_n) = (A_1 \cup \cdots \cup A_{n-1}) \cap A_n = A_n \in \Gamma.$$

Since ν is a pre-valuation,

$$\nu(A_n) = \sum_{k=1}^{n-1} (-1)^{k+1} \sum_{1 \le i_1 < \cdots < i_k \le n-1} \nu\left(A_{i_1} \cap \cdots \cap A_{i_k} \cap A_n\right). \tag{3.2}$$

As the A_i are closed under intersections and non-decreasing,

$$A_{i_1} \cap \cdots \cap A_{i_k} \cap A_n \in \{A_1, \ldots, A_{n-1}\} \qquad (1 \le i_1 < \cdots < i_k \le n-1).$$

Thus by collecting terms in (3.2), we see that there exist coefficients β_i, depending only on the A_i and not on ν, such that

$$\nu(A_n) = \sum_{i=1}^{n-1} \beta_i \nu\left(A_i\right).$$

Similarly, Lemma 3.1 shows that for these same coefficients β_k,

$$\mathbb{I}_{A_n} = \sum_{k=1}^{n-1} (-1)^{k+1} \sum_{1 \le i_1 < \cdots < i_k \le n-1} \mathbb{I}_{A_{i_1} \cap \cdots \cap A_{i_k} \cap A_n} = \sum_{i=1}^{n-1} \beta_i \mathbb{I}_{A_i}.$$

Since $\{A_1, \ldots, A_{n-1}\}$ is closed under meets and non-decreasing, and since

$$\sum_{i=1}^{n-1} (\alpha_i + \alpha_n \beta_i) \mathbb{I}_{A_i} = \sum_{i=1}^{n} \alpha_i \mathbb{I}_{A_i} = 0,$$

it follows by induction that

$$0 = \sum_{i=1}^{n-1} (\alpha_i + \alpha_n \beta_i) \nu(A_i) = \sum_{i=1}^{n} \alpha_i \nu(A_i).$$

Therefore, integration with respect to ν is well-defined. □

Recall that a **relative boolean algebra** of sets is a lattice of sets $\mathcal{B}$ such that $A \smallsetminus B \in \mathcal{B}$ for all $A, B \in \mathcal{B}$. Notice that if μ is a valuation on $\mathcal{B}$, then for all

$A, B \in \mathcal{B}$ with $A \subseteq B$, it holds that

$$\mu(B) = \mu(B \setminus A) + \mu(A) - \mu(\varnothing),$$

and thus

$$\mu(B \setminus A) = \mu(B) - \mu(A). \tag{3.3}$$

Suppose $\Lambda \subseteq \mathcal{P}(X)$ is a lattice of sets. Let $\{\mathcal{B}_\alpha\}_\alpha$ denote the collection of all relative boolean algebras such that $\Lambda \subseteq \mathcal{B}_\alpha \subseteq \mathcal{P}(X)$. Since $\mathcal{P}(X)$ is itself a relative boolean algebra, this collection is nonempty. Note that $\bigcap_\alpha \mathcal{B}_\alpha$ is also a relative boolean algebra that contains Λ, and it is necessarily the smallest such algebra. We will define the **relative boolean closure** of Λ to be the relative boolean algebra $\bigcap_\alpha \mathcal{B}_\alpha$. As mentioned in [KR97, Chapter 2.2], it is the smallest collection of subsets of X that contains Λ and is closed under finite unions, intersections, and relative complements.

The proof of the following corollary follows the proofs given in [KR97, Section 2.2], [SW08, Section 14.4] and [Gro78].

Corollary 3.3 (Groemer's Extension Theorem)**.** *Let Λ be a lattice of sets and let Γ generate Λ.*

(i) Every pre-valuation on Γ extends uniquely to a valuation on Λ.

(ii) Every valuation on Λ extends uniquely to a valuation on the relative boolean closure of Λ.

(iii) There is a one-to-one correspondence between valuations on Λ and linear functionals on the vector space of Λ-simple functions.

Proof. (i) Suppose we are given a pre-valuation ν. To prove uniqueness, suppose μ is any valuation extending ν. By viewing μ as a pre-valuation on Λ, it follows from Theorem 3.2 that

$$\int \mathbb{I}_A \mathrm{d}\mu = \mu(A) \qquad (A \in \Lambda).$$

Now, let $B \in \Lambda$ be arbitrary. Lemma 3.1 shows that we may write $\mathbb{I}_B$ as a Γ-simple function, say

$$\mathbb{I}_B = \sum_{i=1}^{n} \alpha_i \mathbb{I}_{A_i},$$

where each $A_i \in \Gamma$. By what we have just shown, and by the linearity of integration, we obtain

$$\mu(B) = \int \mathbb{I}_B \mathrm{d}\mu = \int \sum_{i=1}^{n} \alpha_i \mathbb{I}_{A_i} \mathrm{d}\mu = \sum_{i=1}^{n} \alpha_i \mu(A_i).$$

Since μ extends ν, we must have

$$\mu(B) = \sum_{i=1}^{n} \alpha_i \nu(A_i).$$

This proves uniqueness. To prove existence, we may simply define an extension by $\mu(B) = \int \mathbb{I}_B \mathrm{d}\nu$.

(ii) Let $\overline{\Lambda}$ denote the relative boolean closure of Λ. Then for all $A, B \in \overline{\Lambda}$,

$$\mathbb{I}_{A \setminus B} = \mathbb{I}_A - \mathbb{I}_{A \cap B}.$$

By using the above identity in conjunction with Lemma 3.1, it is thus possible to write any $\overline{\Lambda}$-simple function as a Λ-simple function. We may thus define an extension $\overline{\mu} : \overline{\Lambda} \to \mathbb{R}$ of a valuation $\mu : \Lambda \to \mathbb{R}$ via

$$\overline{\mu}(A) = \int \mathbb{I}_A \mathrm{d}\mu.$$

(iii) Given a valuation μ, we obtain a linear functional via $s \mapsto \int s \mathrm{d}\mu$. Conversely, suppose we are given a linear functional f. Since

$$f\left(\mathbb{I}_{A \cup B}\right) = f(\mathbb{I}_A) + f(\mathbb{I}_B) - f\left(\mathbb{I}_{A \cap B}\right),$$

it follows that the map $A \mapsto f(\mathbb{I}_A)$ is a valuation on Λ. □

We will also define a notion of distance for compact convex sets as follows. For $x \in \mathbb{R}^d$ and $Y \subset \mathbb{R}^d$, the distance from x to Y is given by

$$d(x, Y) := \inf_{y \in Y} \|x - y\|.$$

For compact sets $K, L \subset \mathbb{R}^d$, the **Hausdorff distance** $\rho(K, L)$ is defined by

$$\rho(K, L) := \max \left\{ \sup_{k \in K} d(k, L),\ \sup_{\ell \in L} d(K, \ell) \right\}.$$

Intuitively, the distance $\rho(K, L)$ measures "how closely K resembles L". It is not too difficult to see that d defines a metric on the collection of compact subsets of $\mathbb{R}^d$. The Hausdorff metric is discussed in more detail in [SW08, Section 12.3] and [KR97, Section 4.1].

3.1 Parallelotopes

Fix an orthogonal coordinate system for $\mathbb{R}^d$. Let $\mathcal{P}^d$ denote the set of all parallelotopes in $\mathbb{R}^d$ with sides parallel to the fixed coordinate system. Note that a parallelotope $P \in \mathcal{P}^d$ may be viewed as a product of intervals $P = I_1 \times \cdots \times I_d$, where each I_j is an interval in $\mathbb{R}$. Notice also that the parallelity asssumption means that $\mathcal{P}^d$ is closed under intersections and hence a lower-semilattice. Let $\text{Par}(d)$ denote the lattice generated by $\mathcal{P}^d$. It consists of all finite unions of parallelotopes in $\mathcal{P}^d$.

In this section, we will focus on studying the Euler characteristic as a valuation on $\text{Par}(d)$. We begin with some definitions. The **dimension of a parallelotope** P, denoted $\dim P$, is defined to be the smallest integer k for which there exists a k-dimensional hyperplane in $\mathbb{R}^d$ that contains P. Notice that the boundary of an n-dimensional paralellotope can be written as a finite union of $(n-1)$-dimensional parallelotopes. These parallelotopes are called the **facets** of P. Each of the facets of P may in turn be written as a finite union of $(n-2)$-dimensional parallelotopes. A **face** of P is a parallelotope obtained by iterating this boundary decomposition process. In other words, a face of P is a facet of P, or a facet of a facet of P, and so on.

Our first task is to prove that a stronger version of the Groemer extension theorem holds for $\text{Par}(d)$. Groemer originally proved the following for polytopes [Gro78], but we will follow the approach of Rota and Klain [KR97, Chapter 4] and instead prove it for parallelotopes. The reason for this is that it is considerably easier to define the intrinsic valuations for parallelotopes.

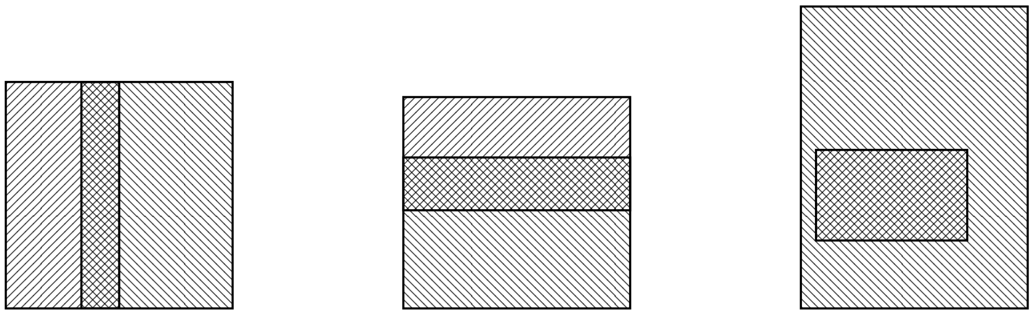
Theorem 3.4 (Groemer's Extension Theorem for *Par*(d))**.** *Suppose $\nu : \mathcal{P}^d \to \mathbb{R}$ is any function satisfying*

$$\nu(P_1 \cup P_2) = \nu(P_1) + \nu(P_2) - \nu(P_1 \cap P_2)$$

whenever $P_1 \cup P_2 \in \mathcal{P}^d$. Then ν extends to a valuation on $\text{Par}(d)$.

The above theorem is equivalent to the statement that if a map $\nu : \mathcal{P}^d \to \mathbb{R}$ satisfies the inclusion-exclusion principle (2.2) for $n = 2$, then ν satisfies the inclusion-exclusion principle for all n.

Example 3.5. Consider the statement for Par(2), the collection of rectangles in $\mathbb{R}^2$ with sides parallel to a fixed coordinate system. Let $P_1, P_2 \in \mathcal{P}^2$ be such that $P_1 \cup P_2 \in \mathcal{P}^2$. Then there are, without loss of generality, essentially three possibilities: P_1 is "left" of P_2, or P_1 is "above" P_2, or P_1 is contained in P_2. Notice that the case $P_1 \subseteq P_2$ also includes all scenarios in which $\dim P_1 < \dim P_2$. These possibilities are depicted in Figure 1, where P_1 is the rectangle filled with lines running north east, while P_2 is the rectangle filled with lines running north west; the intersection $P_1 \cap P_2$ is consequently depicted by the hatch pattern.



(a) P_1 is left of P_2 (b) P_1 is above P_2 (c) P_1 is contained in P_2

Figure 1: The three distinct ways in which $P_1 \cup P_2 \in \mathcal{P}^2$.

Notice that in Figure 1c we have $P_1 \cup P_2 = P_2$ and $P_1 \cap P_2 = P_1$ so that any function $\nu : \mathcal{P}^2 \to \mathbb{R}$ satisfies

$$\nu(P_1 \cup P_2) = \nu(P_2) = \nu(P_1) + \nu(P_2) - \nu(P_1 \cap P_2)$$

for these values of P_1 and P_2. Thus, it is sufficient to only consider cases (a) and (b) when proving that a map $\nu : \mathcal{P}^2 \to \mathbb{R}$ satisfies the hypothesis of Theorem 3.4.

However, there are considerably more ways in which a rectangle can be formed as the union of three rectangles. Figure 2 illustrates one possible method of decomposition, where P_1 is filled with lines running north east, P_2 is filled with lines running north west, and P_3 is filled with a solid grey. Notice that the union of any two of the three rectangles depicted in Figure 2 is not itself a rectangle. As such, it is not immediately obvious that a function ν satisfying the inclusion-exclusion principle in each of the three scenarios outlined in Figure 1 should satisfy the inclusion-exclusion principle for all values of n.

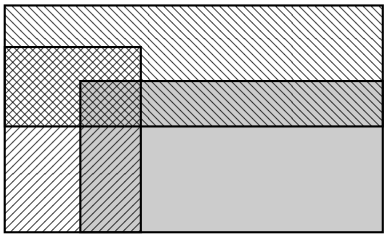


Figure 2: The decomposition of a rectangle in $\mathcal{P}^2$ as a union of three rectangles in $\mathcal{P}^2$.

The situation is analogous for $\mathcal{P}^d$, in which there are essentially $d+1$ ways of decomposing a parallelotope as a union of two parallelotopes. Of these $d+1$ possibilities, d of them are the same up to rotation.

The following proof uses essentially the same argument as the one given for [KR97, Theorem 4.1.3]

Proof. By Corollary 3.3, it suffices to show that ν induces an integral on the space of $\mathcal{P}^d$-simple functions. We will show by induction on d, that for all $\mathcal{P}^d$-simple functions, we have

$$\sum_{i=1}^{n} \alpha_i \mathbb{I}_{P_i} = 0 \qquad \rightsquigarrow \qquad \sum_{i=1}^{n} \alpha_i \nu(P_i) = 0. \tag{3.4}$$

If $d=0$, then $\mathcal{P}^d$ is a point and we are done.

So, suppose $d>0$ and the theorem holds on $\mathrm{Par}(d-1)$. We will show by induction on the number of P_i having dimension equal to d that (3.4) holds for all $\mathcal{P}^d$ simple functions. We will break the proof up into three cases.

Case 1: Suppose all P_i have dimension less than d. Let $H_1, \ldots, H_\ell$ be distinct hyperplanes such that

$$\bigcup_{i=1}^{n} P_i \subseteq \bigcup_{j=1}^{\ell} H_j.$$

We will prove this case by induction on ℓ. If $\ell = 1$, then then all P_i lie in a single hyperplane and we may view $\sum_i \alpha_i \mathbb{I}_{P_i}$ as a $\mathcal{P}^{d-1}$-simple function, from which the result follows by induction on d. Now suppose $\ell > 1$. Since $P_1 \cap H_1, \ldots, P_n \cap H_1$ all lie inside a single hyperlane, and since

$$\sum_{i=1}^{n} \alpha_i \mathbb{I}_{P_i \cap H_1} = \left(\sum_{i=1}^{n} \alpha_i \mathbb{I}_{P_i}\right) \mathbb{I}_{H_1} = 0,$$

it follows that
$$\sum_{i=1}^{n} \alpha_i \nu(P_i \cap H_1) = 0.$$
Consequently,

$$\sum_{i=1}^{n} \alpha_i \nu(P_i) = \sum_{i=1}^{n} \alpha_i (\nu(P_i) - \nu(P_i \cap H_1)).$$

Since the above sum runs over parallelotopes contained in $\ell - 1$ or fewer distinct hyperplanes, and since

$$\sum_{i=1}^{n} \alpha_i (\mathbb{I}_{P_i} - \mathbb{I}_{P_i \cap H_1}) = 0,$$

it follows by induction on ℓ that

$$\sum_{i=1}^{n} \alpha_i \nu(P_i) = 0.$$

Case 2: Suppose that exactly one P_i has dimension d, and all other P_j have dimension strictly less than d. Without loss of generality, $\dim P_1 = d$. Then we may pick $x \in P_1 \setminus \bigcup_{i=2}^{n} P_i$ to see that

$$\alpha_1 = \sum_{i=1}^{n} \alpha_i \mathbb{I}_{P_i}(x) = 0,$$

so that by Case 1,
$$\sum_{i=1}^{n} \alpha_i \nu(P_i) = \sum_{i=2}^{n} \alpha_i \nu(P_i) = 0.$$

Case 3: Suppose at least two of the P_i have dimension equal to d. Without loss of generality, $\dim P_1 = \dim P_2 = d$. Let $H_1, \ldots, H_\ell$ be hyperplanes such that

$$\partial P_1 = \bigcup_{i=1}^{\ell} (H_i \cap P_1).$$

For each i, let H_i^+ and H_i^- denote the two connected components of $\mathbb{R}^d \setminus H_i$, labeled so that the interior of P_1 is contained in H_i^+. Then

$$P_1 = \overline{H_1^+} \cap \cdots \cap \overline{H_\ell^+}.$$

Notice that for each i, j,

$$\nu(P_i) = \nu\left(P_i \cap \overline{H_j^+}\right) + \nu\left(P_i \cap \overline{H_j^-}\right) - \nu\left(P_i \cap H_j\right). \tag{3.5}$$

Since

$$\sum_{i=1}^{m} \alpha_i \mathbb{I}_{P_i \cap H_j} = \sum_{i=1}^{m} \alpha_i \mathbb{I}_{P_i} \mathbb{I}_{H_j} = 0,$$

and each $P_i \cap H_j$ has dimension strictly less than d, it follows from Case 1 that

$$\sum_{i=1}^{m} \alpha_i \nu\left(P_i \cap H_j\right) = 0.$$

Similarly, since

$$\sum_{i=1}^{m} \alpha_i \mathbb{I}_{P_i \cap \overline{H_j^-}} = \sum_{i=1}^{m} \alpha_i \mathbb{I}_{P_i} \mathbb{I}_{\overline{H_j^-}} = 0,$$

and $\dim(P_1 \cap \overline{H_j^-}) < d$, it follows by induction on the number of parallelotopes of dimension d that

$$\sum_{i=1}^{m} \alpha_i \nu\left(P_i \cap \overline{H_j^-}\right) = 0.$$

Thus, by (3.5),

$$\sum_{i=1}^{m} \alpha_i \nu\left(P_i\right) = \sum_{i=1}^{m} \alpha_i \nu\left(P_i \cap \overline{H_j^+}\right).$$

By repeating this argument, we see that

$$\sum_{i=1}^{n} \alpha_i \nu\left(P_i\right) = \sum_{i=1}^{n} \alpha_i \nu\left(P_i \cap \overline{H_1^+} \cap \cdots \cap \overline{H_m^+}\right) \tag{3.6}$$

for every integer $m \geq 1$. In particular, taking $m = \ell$, we have

$$\sum_{i=1}^{n} \alpha_i \nu\left(P_i\right) = \sum_{i=1}^{n} \alpha_i \nu\left(P_i \cap P_1\right).$$

Since P_2 also has dimension equal to d, we can use the same argument to show that

$$\sum_{i=1}^{n} \alpha_i \nu\left(P_i\right) = \sum_{i=1}^{n} \alpha_i \nu\left(P_i \cap P_1 \cap P_2\right).$$

Furthermore, since

$$\sum_{i=1}^{m} \alpha_i \mathbb{I}_{P_i \cap P_1 \cap P_2} = \sum_{i=1}^{m} \alpha_i \mathbb{I}_{P_i} \mathbb{I}_{P_1} \mathbb{I}_{P_2} = 0,$$

it then follows by induction on the number of parallelotopes of dimension d that

$$0 = (\alpha_1 + \alpha_2)\nu(P_1 \cap P_2) + \sum_{i=3}^{n} \alpha_i \nu(P_i \cap P_1 \cap P_2) = \sum_{i=1}^{n} \alpha_i \nu(P_i).$$

This completes the inductive step. □

We will say that a valuation μ on $\mathrm{Par}(d)$ is

(i) **continuous** on $\mathcal{P}^d$ if its restriction to $\mathcal{P}^d$ is continuous in the Hausdorff metric;

(ii) **translation invariant** if $\mu(x + P) = \mu(P)$ for all $P \in \mathrm{Par}(d), x \in \mathbb{R}^d$;

(iii) **invariant** if it is both translation invariant and invariant under permutations of coordinates.

We will now shift our attention to invariant valuations that are continuous on $\mathcal{P}^d$. Notice that a translation invariant valuation on $\mathrm{Par}(d)$ is uniquely determined by its values on parallelotopes of the form $[0, \ell_1] \times \cdots \times [0, \ell_d]$. Let $s_0, \ldots, s_d$ denote the elementary symmetric polynomials in d variables, under the convention that $s_0 = 1$. We will define the **intrinsic valuations** $\mu_0, \ldots, \mu_d$ on $\mathrm{Par}(d)$ by setting

$$\mu_k ([0, \ell_1] \times \cdots \times [0, \ell_d]) := s_k (\ell_1, \ldots, \ell_d).$$

Then for $k > 0$, we have

$$\mu_k ([0, \ell_1] \times \cdots \times [0, \ell_d]) = \sum_{1 \le i_1 < \cdots < i_k \le d} \ell_{i_1} \cdots \ell_{i_k}.$$

The μ_k are called the intrinsic valuations because their definition does not depend on the ambient dimension d. To see this, suppose $P \in \mathcal{P}^d$ has side lengths $\ell_1, \ldots, \ell_d$. Let $d' > d$ and embed $\mathbb{R}^d$ into $\mathbb{R}^{d'}$ by identifying $\mathbb{R}^d \approx \mathbb{R}^d \times \{0\}^{d'-d} \subset \mathbb{R}^{d'}$. Then in $\mathbb{R}^{d'}$, the parallelotope P has side lengths given by $\ell_1, \ldots, \ell_d, 0, \ldots, 0$. Let μ_k^d and $\mu_k^{d'}$ denote the intrinsic valuations on $\mathbb{R}^d$ and $\mathbb{R}^{d'}$ respectively. Then by definition,

$$\mu_k^d(P) = \sum_{1 \le i_1 < \cdots < i_k \le d} \ell_{i_1} \cdots \ell_{i_k} = \sum_{1 \le i_1 < \cdots < i_k \le d'} \ell_{i_1} \cdots \ell_{i_k} = \mu_k^{d'}(P),$$

as all terms involving a side length ℓ_i with $i > d$ vanish.

Lemma 3.6. *The intrinsic valuations are actually valuations.*

Proof. By Theorem 3.4, it suffices to verify that each μ_k is 2-additive. Suppose $P, P' \in \mathcal{P}^d$ with $P \cup P' \in \mathcal{P}^d$. For $1 \leq i \leq d$, let ℓ_i, ℓ_i' and ℓ_i'' denote the side lengths of P, P', and $P \cap P'$ respectively. Then $P \cup P'$ has side lengths $\ell_i + \ell_i' - \ell_i''$. In light of Example 3.5 and the invariance of μ_k under rotations, we see that it is sufficient to consider the case when P and P' are "side by side" (in the sense of Figure 1a). This means that, without loss of generality, $\ell_i = \ell_i'$ for $i \geq 2$. Then

$$\begin{aligned}\mu_k(P \cup P') &= \sum_{2 \leq i_2 < \cdots < i_k \leq d} (\ell_1 + \ell_1' - \ell_1'')\ell_{i_2} \cdots \ell_{i_k} + \sum_{2 \leq i_1 < \cdots < i_k \leq d} \ell_{i_1}\ell_{i_2} \cdots \ell_{i_k} \\ &= \mu_k(P) + \mu_k(P') - \mu_k(P \cap P'). \end{aligned}$$

□

It is worth noting that our proof of Lemma 3.6 differs substantially from the proof of the corresponding result in [KR97, Theorem 4.2.1], which defines μ_0 and μ_1 independently and then defines the μ_k as the coefficients (with respect to t) of the polynomial given by evaluating the product valuation

$$(\mu_0 + t\mu_1) \times \cdots \times (\mu_0 + t\mu_1)$$

on parallelotopes. One of the insights that is lost in our approach is the remarkable fact that the higher dimensional intrinsic valuations can be built up from only the valuations μ_0 and μ_1.

We will define the **Euler characteristic** to be the valuation $\chi := \mu_0$. It assigns the value 1 to every (nonempty) parallelotope in $\mathcal{P}^d$. Notice that each of the valuations μ_k may be viewed as k-dimensional "measures" on $\mathrm{Par}(d)$, with $\mu_d(P)$ being the standard volume formula for a parallelotope. Furthermore, if P is k-dimensional, then $\mu_{k-1}(P)$ is equal to one half the surface area of P. Just as the elementary symmetric polynomials generate all symmetric polynomials, the intrinsic valuations μ_k generate all invariant valuations that are continuous on $\mathcal{P}^d$. In fact, a stronger result is true.

The proof of the following result is similar to [KR97, Theorem 4.2.4, Theorem 4.2.5].

Theorem 3.7 (Basis Theorem for $\mathrm{Par}(d)$)**.** *The intrinsic valuations on* $\mathrm{Par}(d)$ *are a basis of the vector space of all invariant valuations on* $\mathrm{Par}(d)$ *that are continuous* *on* $\mathcal{P}^d$.

Proof. Induction on d. For $d = 0$, the result is trivial.

Next, suppose $d > 0$ and let μ be an invariant valuation that is continuous on $\mathcal{P}^d$. Denote by $\mu|_{\mathrm{Par}(d-1)}$ the valuation on $\mathrm{Par}(d-1)$ defined by

$$\mu|_{\mathrm{Par}(d-1)}\left(\prod_{i=1}^{d-1}[0,\ell_i]\right) = \mu\left(\prod_{i=1}^{d-1}[0,\ell_i] \times \{0\}\right).$$

By induction, there are scalars $\alpha_0, \dots, \alpha_{d-1}$ such that

$$\mu|_{\mathrm{Par}(d-1)} = \sum_{i=0}^{d-1} \alpha_i \mu_i.$$

Since μ is invariant under permutations of coordinates, the valuation

$$\nu := \mu - \sum_{i=0}^{d-1} \alpha_i \mu_i$$

vanishes on all parallelotopes of dimension strictly less than d. It remains to show that ν is a scalar multiple of μ_d.

Let $\ell_1, \dots, \ell_d \in \mathbb{R}$ be arbitrary. First, observe that since ν vanishes on parallelotopes of dimension $< d$, we have $\mu\left(\{1\} \times \prod_2^d [0,\ell_i]\right) = 0$ so that by induction and the translation invariance of ν, we have for each integer $p \geq 0$,

$$\nu\left([0,p] \times \prod_{i=2}^{d} \ell_i\right) = \sum_{k=1}^{p} \nu\left([p-1,p] \times \prod_{i=2}^{d} \ell_i\right) = p\nu\left([0,1] \times \prod_{i=2}^{d} \ell_i\right).$$

Similarly, for each integer $q \geq 1$,

$$\nu\left(\left[0,\frac{1}{q}\right] \times \prod_{i=2}^{d} \ell_i\right) = \frac{1}{q}\sum_{k=1}^{q} \nu\left(\left[\frac{k-1}{q},\frac{k}{q}\right] \times \prod_{i=2}^{d} \ell_i\right) = \frac{1}{q}\nu\left([0,1] \times \prod_{i=2}^{d} \ell_i\right).$$

It follows that

$$\nu\left([0,r] \times \prod_{i=2}^{d} \ell_i\right) = r\nu\left([0,1] \times \prod_{i=2}^{d} \ell_i\right) \qquad (r \in \mathbb{Q}).$$

By continuity, the above equality also holds for all $r \in \mathbb{R}$. Since ν is invariant under permutations of coordinates,

$$\nu\left(\prod_{i=1}^{d}[0,\ell_i]\right)=\ell_1\nu\left([0,\ell_2]\times[0,1]\times\prod_{i=3}^{d}[0,\ell_i]\right)=\cdots=\ell_1\cdots\ell_d\nu\left([0,1]^d\right).$$

Thus, if we set $\alpha_d=\nu([0,1]^d)$, then

$$\alpha_d\mu_d=\nu=\mu-\sum_{i=1}^{d-1}\mu_i.$$

Lastly, the independence of the μ_i follows from the fact that $\mu_k(P)=0$ whenever $\dim P<k$. □

A valuation μ on $\mathrm{Par}(d)$ that satisfies $\mu(P)=0$ for all P with $\dim P<d$ is called **simple**. Notice that the proof of the preceding theorem also shows that if μ is a simple invariant valuation that is continuous on $\mathcal{P}^d$, then

$$\mu(P)=\mu\left([0,1]^d\right)\mathrm{vol}(P)\qquad(P\in\mathrm{Par}(d)),$$

where $\mathrm{vol}(P)=\mu_d(P)$ is the volume of P. This is known as the volume theorem. The definitions of each μ_k illustrate that whenever $\dim P<k$, we have

$$\mu_k(P)=0.$$

As such, the μ_k may be viewed as the unique k-dimensional invariant valuation on $\mathrm{Par}(d)$ that is continuous on $\mathcal{P}^d$. In particular, this describes the manner in which the Euler characteristic $\chi=\mu_0$ can be viewed as a 0-dimensional "measure" on $\mathrm{Par}(d)$.

The next goal will be to translate these results to the much broader class of all polyconvex sets in $\mathbb{R}^d$.

3.2 Polyconvex Sets

Let $\mathcal{K}^d$ denote the collection of all compact convex subsets of $\mathbb{R}^d$. A set is **polyconvex** if it is a finite union of compact convex subsets of $\mathbb{R}^d$. The lattice of all polyconvex sets is denoted by $\mathrm{Polycon}(d)$. As the intersection of two compact convex sets is again a compact convex set, we see that $\mathrm{Polycon}(d)$ is generated by $\mathcal{K}^d$. It will be convenient to view $\mathrm{Par}(d)$ as a sublattice of $\mathrm{Polycon}(d)$.

As in the case of Par(d), we will define the **dimension** of a polyconvex set $K_1 \cup \cdots \cup K_n \in \text{Polycon}(d)$ to be the smallest integer k for which there exist a finite number of k-planes whose union contains $K_1 \cup \cdots \cup K_n$.

We will say that a valuation μ is **continuous** on $\mathcal{K}^d$ if the restriction of μ to $\mathcal{K}^d$ is continuous with respect to the Hausdorff metric.

Theorem 3.8 (Groemer's Theorem for Polycon(d))**.** *Let $\nu : \mathcal{K}^d \to \mathbb{R}$ be a continuous map that satisfies*

$$\nu(K_1 \cup K_2) = \nu(K_1) + \nu(K_2) - \nu(K_1 \cap K_2)$$

for all $K_1, K_2 \in \mathcal{K}^d$ for which $K_1 \cup K_2 \in \mathcal{K}^d$. Then ν extends to a valuation on Polycon(d)*.*

We refer the reader to [KR97, Theorem 5.1.1], [Gro78, Theorem 3], or [SW08, Theorem 14.4.2] for a proof of the above theorem. The idea behind the proof is not too different from the proof in the case of parallelotopes (or polytopes). The assumption of continuity is required because the argument now involves: writing convex sets as countable unions of closed half-spaces, showing that an equation analogous to (3.5) holds for all m, and then using continuity to conclude that equality holds for countable intersections as well.

By the previous theorem, we are justified in defining the **Euler characteristic** to be the unique valuation on Polycon(d) satisfying

$$\chi(\varnothing) = 0 \qquad \text{and} \qquad \chi(K) = 1 \qquad (K \in \mathcal{K}^d).$$

Notice that the definition of the Euler characteristic on Polycon(d) agrees with the definition of the Euler characteristic on the sublattice Par(d). We will follow the convention of [KR97] and define polytopes as follows. A **convex polytope** is an intersection of a finite number of closed half-spaces, and a **polytope** is a finite union of convex polytopes. We will now set out to prove that χ satisfies the alternating face formula for all bounded polytopes in $\mathbb{R}^d$. Our treatment of this subject is similar to the one given in [KR97, Section 5.2].

Lemma 3.9. *Let σ be a d-simplex in $\mathcal{K}^d$. Then $\chi(\partial\sigma) = 1 + (-1)^{d+1}$.*

Proof. Denote the i-th facet of σ by $\sigma_i = \Delta[x_0, \ldots, x_{i-1}, x_{i+1}, \ldots, x_d]$. Since

$$|\{k\text{-faces of } \sigma\}| = \binom{d+1}{k+1} \qquad \text{and} \qquad \sum_{k=0}^{d+1} (-1)^k \binom{d+1}{k} = 0,$$

it follows that

$$\begin{aligned}
\chi(\partial\sigma) &= \chi\left(\bigcup_{i=1}^{d+1} \sigma_{i-1}\right) \\
&= \sum_{k=1}^{d+1} (-1)^{k+1} \sum_{1 \le i_1 < \cdots < i_k \le d+1} \chi\left(\sigma_{i_1 - 1} \cap \cdots \cap \sigma_{i_k - 1}\right) \\
&= \chi(\varnothing) + \sum_{k=1}^{d} (-1)^{k+1} \left|\{(d-k)\text{-faces of } \sigma\}\right| \\
&= \sum_{k=1}^{d} (-1)^{k+1} \binom{d+1}{d+1-(d-k+1)} \\
&= \sum_{k=1}^{d} (-1)^{k+1} \binom{d+1}{k} \\
&= 1 + (-1)^{d+1} - \sum_{k=0}^{d+1} (-1)^k \binom{d+1}{k} \\
&= 1 + (-1)^{d+1}.
\end{aligned}$$

□

Let $\overline{\text{Polycon}(d)}$ denote the relative boolean closure of Polycon(d). Let $K \in \mathcal{K}^d$ with $\dim K = n$. Then we can find an n-dimensional hyperplane H such that $K \subseteq H$. We will define the **relative interior** of K, denoted $\operatorname{relint} K$ to be the interior of K in the subspace topology induced by H. Notice that if $\dim K = d$, then we may take $H = \mathbb{R}^d$ and thus $\operatorname{relint} K = \operatorname{Int} K$ in this case. We will similarly define the **relative boundary** of K, denoted $\partial_{\text{rel}} K$.

Notice that for all polytopes $P \in \mathcal{K}^d$, the relative boundary of P is a finite union of compact convex subsets and hence lies in $\mathcal{K}^d$. Thus,

$$\operatorname{relint} P = P \smallsetminus \partial_{\text{rel}} P \in \overline{\text{Polycon}(d)} \qquad \text{and} \qquad \chi(\operatorname{relint} P) = \chi(P) - \chi(\partial_{\text{rel}} P).$$

It then follows from Lemma 3.9 that for all simplices $\sigma \in \mathcal{K}^d$,

$$\chi(\operatorname{relint}(\sigma)) = \chi(\sigma) - \chi(\partial_{\text{rel}} \sigma) = 1 - \left(1 + (-1)^{1 + \dim \sigma}\right) = (-1)^{\dim \sigma}.$$

Theorem 3.10. *Let Δ be a simplicial complex with geometric realization $\|\Delta\|$ equal* *to some polytope $P \in \overline{\text{Polycon}(d)}$. Let $r_k(\Delta)$ denote the set of k-faces of Δ. Then*

$$\chi(P) = \sum_{k=1}^{d} (-1)^k |r_k(\Delta)|.$$

Proof. Recall that for all distinct faces $\sigma_i, \sigma_j \in \Delta$, we have $\text{relint}(\sigma_i) \cap \text{relint}(\sigma_j) = \varnothing$. Then since χ is a valuation, we have

$$\begin{aligned}
\chi(\|\Delta\|) &= \chi\left(\bigcup_{\sigma\in\Delta} \text{relint}\,\sigma\right) \\
&= \sum_{\sigma\in\Delta} \chi(\text{relint}\,\sigma) \\
&= \sum_{k=0}^{d} \sum_{\sigma\in r_k(\Delta)} \chi(\text{relint}\,\sigma) \\
&= \sum_{k=0}^{d} (-1)^k |r_k(\Delta)|.
\end{aligned}$$

□

We thus obtain the alternating face formula for all bounded finite-dimensional polyhedra. Notice that since χ is a well-defined valuation, the Euler characteristic of a polyhedron P does not depend on the choice of simplicial complex Δ.

Just as the Euler characteristic $\chi = \mu_0$ in $\text{Par}(d)$ may be viewed as part of the basis $\mu_0, \ldots, \mu_d$ of invariant continuous valuations, we may also view the Euler characteristic χ on $\text{Polycon}(d)$ as a part of a basis $\mu_0, \ldots, \mu_d$ of continuous invariant valuations on $\text{Polycon}(d)$. However, the construction of these valuations is considerably more involved in the polyconvex setting. A high level description of the intrinsic valuations is given in [Sch86], while a more rigorous one is presented in [KR97, Chapters 5-9].

4 Posets and Abstract Simplicial Complexes

The goals of this chapter are to define the Euler characteristic for finite simplicial complexes and for finite posets. We will aim to define the Euler characteristic of a finite simplicial complex in a way that mirrors its definition for parallelotopes, although the tools used will be different. We will also spend some time developing the connections between finite posets and finite lattices. This will help to make the transition to posets feel more natural. The chapter will then be concluded by defining the Euler characteristic for posets and ensuring it agrees with the lattice theoretic definition. This will be done in terms of the Möbius function of a poset. Our treatment of lattices draws heavy inspiration from [KR97, Chapters 2-3], [BBR86], and [Mun09], while our treatment of posets is inspired primarily by [Rot64].

We begin with some lattice theoretic notions that will be of use in this chapter. A nonzero element $p \in \Lambda$ is **irreducible** if

$$p = a \vee b \quad \rightsquigarrow \quad p = a \text{ or } p = b. \tag{4.1}$$

Similarly, we say that a nonzero $p \in \Lambda$ is **prime** if

$$p \leq a \vee b \quad \rightsquigarrow \quad p \leq a \text{ or } p \leq b. \tag{4.2}$$

Notice that the definitions of prime and irreducible do not require that Λ actually has a zero.

Example 4.1. Let $\Lambda = \mathbb{Z}_{>0}$, ordered by divisibility, i.e., $a \leq b \leftrightsquigarrow a|b$. Then Λ is a lattice with

$$n \wedge m = \gcd(n, m) \qquad \text{and} \qquad n \vee m = \operatorname{lcm}(n, m).$$

Notice that if p is prime (in the number-theoretic sense), then p^r is prime (in the lattice-theoretic sense) for all $r \geq 1$. As such, prime elements are in some sense more intuitively aligned with powers of primes than with primes themselves. This example also illustrates that it is possible to have distinct (lattice-theoretic) primes p, q with $p < q$. Furthermore, the integer 1 is the zero element of Λ, so that, in this case, the requirement of prime elements being nonzero agrees with the number-theoretic requirement that primes cannot be units. Next, let p^r and q^s be prime with $p \neq q$. Then both

$$p^r \vee q^s = p^r q^s \qquad \text{and} \qquad p^r \wedge q^s = 1$$

are not prime, illustrating that in general, the collection of prime elements does not form a sublattice.

In the previous example, the meet of two primes was either prime or the zero element. However, this is not true in general. For instance, consider the lattice depicted in Figure 3, in which $c = p \wedge q = a \vee b$.

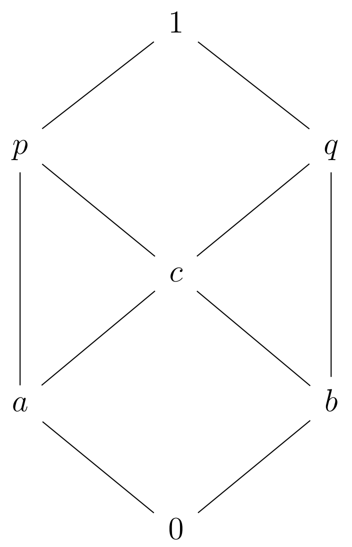


Figure 3: A lattice with primes p, q such that $p \wedge q$ is not prime.

Lemma 4.2. *In a distributive lattice, an element is prime if and only if it is irreducible.*

Proof. If p is prime and $p = a \vee b$, then without loss of generality $p \leq a$ so that $a \leq a \vee b = p \leq a$; this shows p is irreducible.

Next, suppose p is irreducible and $p \leq a \vee b$. Then by (2.1)

$$p = p \wedge (a \vee b) = (p \wedge a) \vee (p \wedge b),$$

so that by irreducibility, $p = p \wedge a$ or $p = p \wedge b$. □

Notice that the distributivity of Λ was only required to show that irreducible elements are prime, illustrating that in general, condition (4.2) is stronger than condition (4.1). The set of all prime elements of Λ is denoted by $\mathrm{Spec}(\Lambda)$. We will henceforth refer to prime and irreducible elements interchangeably. This is justified as we are only interested in distributive lattices.

We will now introduce a few poset definitions. Let P be a finite poset. A **chain** is a totally ordered finite subset of P. A chain may be written as a list of elements $a_0 < a_1 < \ldots < a_n$. The **length** of a chain is given by the number of $<$ symbols in such a list. An **antichain** A is a subset of P such that no two elements of A are comparable. An **ideal** $\mathfrak{a}$ of P is a subset $\mathfrak{a} \subset P$ such that if $a \in \mathfrak{a}$ and $x \in P$ with $x \leq a$, then $x \in \mathfrak{a}$. An ideal $\mathfrak{a}$ is **principal** if there exists $a \in P$ such that

$$\mathfrak{a} = P_{\leq a} := \langle a \rangle := \{x \in P : x \leq a\}.$$

We will denote by $\mathcal{J}(P)$ the set of all ideals of a poset P.

Notice that $\mathcal{J}(P)$ is closed under finite joins and meets. As such, it is a sublattice of the boolean algebra $\mathcal{P}(P)$. Consequently, $\mathcal{J}(P)$ is distributive. Hence, we may associate a distributive lattice to each poset. The statement of the following result was inspired by [BBR86].

Lemma 4.3. *Let P be a finite poset.*

(i) An ideal is principal if and only if it is $\cup$-prime.

(ii) $P \approx \mathrm{Spec}(\mathcal{J}(P))$.

Proof. (i) Suppose $\langle a_0 \rangle \leq \mathfrak{b} \cup \mathfrak{c}$. Then $a_0 \in \mathfrak{b} \cup \mathfrak{c}$ so that without loss of generality, $a_0 \in \mathfrak{b}$. Since $\mathfrak{b}$, is an ideal, we have $a \in \mathfrak{b}$ for all $a \leq a_0$, from which it follows that $\langle a_0 \rangle \leq \mathfrak{b}$ and thus $\langle a_0 \rangle$ is $\cup$-prime.

Conversely, suppose $\mathfrak{a}$ is $\cup$-prime. Since $\mathfrak{a}$ is nonempty, we may write $\mathfrak{a} = \{a_1, \ldots, a_n\}$. Then

$$\mathfrak{a} = \langle a_1 \rangle \cup \cdots \cup \langle a_n \rangle,$$

so that $\mathfrak{a} = \langle a_i \rangle$ for some i.

(ii) By (i), the order-preserving injection

$$P \hookrightarrow \mathcal{J}(P)$$
$$a \mapsto \langle a \rangle$$

has image equal to $\mathrm{Spec}(\mathcal{J}(P))$. □

Notice that (ii) does not necessarily hold when P is infinite. As an easy example, take $P = \mathbb{N}$ in the usual order and consider $\mathfrak{a} = \mathbb{N}$. If $\mathfrak{a} \leq \mathfrak{b} \cup \mathfrak{c}$, then one of $\mathfrak{b}$ or $\mathfrak{c}$ must be infinite, and hence equal to $\mathbb{N}$. Thus, $\mathfrak{a}$ is prime. However, $\mathfrak{a}$ is not principal.

From (ii), it follows that every poset may be viewed as the set of irreducible elements of a finite distributive lattice. As we will soon show, it is also the case that for every finite distributive lattice Λ, we have $\Lambda \approx \mathcal{J}(\mathrm{Spec}(\Lambda))$. We will utilize the following factorization theorem.

Theorem 4.4. *Every nonzero element of a finite distributive lattice Λ may be written uniquely as a join of pair-wise incomparable primes.*

Consider the lattice Λ from Example 4.1. While this lattice is certainly not finite, it does still possess unique factorization. By the fundamental theorem of arithmetic, any positive integer x may be written uniquely as a product of prime powers $x = p_1^{e_1} \cdots p_r^{e_r}$. In the language of Theorem 4.4, the pair-wise incomparable primes are then $p_1^{e_1}, \dots, p_r^{e_r}$. Notice that

$$p_1^{e_1} \vee \cdots \vee p_r^{e_r} = \mathrm{lcm}\,(p_1^{e_1}, \dots, p_r^{e_r}) = p_1^{e_1} \cdots p_r^{e_r} = x.$$

This shows that the finiteness assumption is not necessary for unique factorization to hold. And indeed, the proof of Theorem 4.4 will show that the hypothesis may be strengthened to the assumption that $\langle x \rangle$ is finite for all x. The following proof is due to [BBR86]

Proof of Theorem 4.4. **Existence** Let $x_0 \in \Lambda$ be nonzero. We will prove this by induction on the cardinality of $\langle x_0 \rangle$. If x_0 is itself irreducible, we are done. Otherwise, there exist $a, b \in \Lambda$ with $x_0 = a \vee b$ such that $a < x_0$ and $b < x_0$. Then by induction, we can write a and b as a join of irreducibles. We are thus able to write x_0 as a join of irreducibles $x_0 = p_1 \vee \cdots \vee p_n$. If p_i and p_j are comparable for some $i \neq j$, say $p_i \geq p_j$, then we have $p_i \vee p_j = p_i$ so that

$$x_0 = p_1 \vee \cdots \vee p_{j-1} \vee p_{j+1} \vee \cdots \vee p_n.$$

Continuing in this manner, we obtain a factorization of x_0 into pair-wise incomparable irreducible elements.

Uniqueness Now, suppose there exist pair-wise incomparable irreducibles $p_1, \dots, p_k$, $q_1, \dots, q_\ell$ such that

$$p_1 \vee \cdots \vee p_k = q_1 \vee \cdots \vee q_\ell.$$

From (2.1) it follows that $p_1 \leq q_1 \vee \cdots \vee q_\ell$. Since p_1 is prime, we have $p_1 \leq q_1$ without loss of generality. Similarly, $q_1 \leq p_{i_1}$ for some i_1. Since the p_i

are assumed to be pair-wise incomparable, and since $p_1 \leq q_1 \leq p_{i_1}$, we have $p_1 = p_{i_1}$ and hence $p_1 = q_1$.

Next, $p_2 \leq q_{j_2}$ for some j_2. We must have $j_2 > 1$, as otherwise $p_2 \leq q_1 = p_1$. We may thus suppose $j_2 = 2$. Then $q_2 \leq p_{i_2}$ for some i_2 and, as before, we must then have $i_2 = 2$ and $p_2 = q_2$. Continuing in this manner, we see that $p_i = q_i$ and $k \leq \ell$. By symmetry, we have $k = \ell$, proving uniqueness. $\square$

The statement of the following theorem was inspired by [BBR86, Mun09]; the proof follows the argument in [BBR86].

Theorem 4.5 (Birkhoff Representation Theorem)**.** *Let Λ be a finite distributive lattice. Then*

$$\Lambda \approx \mathcal{J}(\mathrm{Spec}(\Lambda)).$$

Proof. Consider the mapping

$$\begin{aligned} \Phi : \Lambda &\to \mathcal{J}(\mathrm{Spec}(\Lambda)) \\ x &\mapsto \{p \in \mathrm{Spec}(\Lambda) : p \leq x\} \end{aligned}$$

Note that for all $x, y \in \Lambda$, we see that for all primes p,

$$p \in \Phi(x \vee y) \quad \leftrightsquigarrow \quad p \leq x \vee y \quad \leftrightsquigarrow \quad p \leq x \text{ or } p \leq y \quad \leftrightsquigarrow \quad p \in \Phi(x) \cup \Phi(y).$$

This shows $\Phi(x \vee y) = \Phi(x) \cup \Phi(y)$. Similarly, $\Phi(x \wedge y) = \Phi(x) \cap \Phi(y)$. Hence, Φ is a lattice homomorphism. Next, observe that Φ restricts to a poset isomorphism

$$\begin{aligned} \Phi : \mathrm{Spec}(\Lambda) &\to \mathrm{Spec}(\mathcal{J}(\mathrm{Spec}(\Lambda))) \\ p &\mapsto \langle p \rangle. \end{aligned}$$

Thus, by unique factorization in Λ and $\mathcal{J}(\mathrm{Spec}(\Lambda))$, it follows that Φ is an isomorphism on all of Λ. $\square$

Birkhoff's theorem may be seen as stating that every finite distributive lattice is the set of ideals of some finite poset. Conversely, Lemma 4.3 says that every finite poset corresponds to the irreducible elements of some finite distributive lattice. This leads to a correspondence between finite posets and finite distributive lattices.

$$\mathsf{Posets} \underset{\mathrm{Spec}}{\overset{\mathcal{J}}{\rightleftarrows}} \mathsf{Lattices}$$

While the maps $\mathcal{J}$ and Spec are not proper inverses of each other, they are inverses up to isomorphism.

Example 4.6. Consider the case when $\Lambda = \mathcal{P}(S)$ is a boolean algebra. Then Spec(Λ) consists of all the singletons of S. Thus, $\mathcal{J}(\text{Spec}(\mathcal{P}(S)))$ consists of all the sets of the form

$$\{\{a_1\}, \{a_2\}, \ldots, \{a_n\}\},$$

where $a_i \in S$. The isomorphism then becomes

$$\begin{aligned} \mathcal{P}(S) &\to \mathcal{J}(\text{Spec}(\mathcal{P}(S))) \\ \{a_1, \ldots, a_n\} &\mapsto \{\{a_1\}, \ldots, \{a_n\}\} \end{aligned}$$

It's worth noting that, in this context, the unique factorization theorem amounts to the trivial statement that every finite subset may be written as a finite union of singletons.

4.1 Simplicial Complexes

We now shift our attention towards defining the Euler characteristic of a simplicial complex. Recall that a finite (abstract) simplicial complex Δ is a collection of subsets of some finite set S such that if $y \in \Delta$ and $x \subseteq y$, then $x \in \Delta$. The simplest example of a simplicial complex is given by $\Delta^d = \mathcal{P}(\{0, \ldots, d\})$; this is referred to as the **standard d-simplex**. We will say that a simplicial complex Δ is a **subcomplex** of a simplicial complex Δ' if $\Delta \subseteq \Delta'$.

Notice that if Δ is an arbitrary finite simplicial complex on a vertex set S. Then we may replace Δ with an isomorphic simplicial complex $\widetilde{\Delta}$ on the vertex set $\{0, \ldots, d\}$, where $d \geq |S|$. Furthermore, $\widetilde{\Delta}$ is by definition a subcomplex of the standard d-simplex Δ^d. It is thus sufficient to restrict ourselves to the case when the vertex set is given by $\{0, \ldots, d\}$ for some non-negative integer d.

Recall that an ideal of the lattice $\Delta^d = \mathcal{P}(\{0, \ldots, d\})$ consists of a collection $\mathfrak{a}$ of subsets of $\{0, \ldots, d\}$ such that if $y \in \mathfrak{a}$ and $x \subseteq y$, then $x \in \mathfrak{a}$. In other words, the ideals of Δ^d are exactly the subcomplexes Δ^d. Furthermore, the principal ideals of $\mathcal{P}(\{0, \ldots, d\})$ correspond to the faces of Δ^d. We may thus define the Euler characteristic of a simplicial complex as a valuation on the lattice $\mathcal{J}\Delta^d$.

Let us define the function $r : \Delta^d \to \mathbb{Z}_{\geq 0}$ by $r(x) = |x|$. Using the convention (or perhaps more accurately, the definition) $0 = \varnothing$, we see that the following hold:

(i) $r(0) = 0$;

(ii) if $x < y$, then $r(x) < r(y)$;

(iii) if y covers x, then $r(y) = r(x) + 1$.

Recall that "y covers x" means that $x < y$ and there does not exist a z for which $x < z < y$. For a general lattice Λ, a function $r : \Lambda \to \mathbb{Z}_{\geq 0}$ that satisfies (i)-(iii) is called a **rank function** on Λ. If such a function exists, then Λ is called a **graded lattice**. The elements of Λ with rank 1 are called **atoms**. Notice that all atoms are irreducible and thus prime.

Lemma 4.7. *Let Λ be a finite graded distributive lattice. Then every valuation on Λ is uniquely determined by its values on primes.*

Proof. Suppose μ is defined for every prime element of Λ. We will define an extension $\overline{\mu}$ of μ recursively.

For every $x \in \Lambda$ with $r(x) = 1$, we may define $\overline{\mu}(x) = \mu(x)$.

Now, suppose $x \in \Lambda$ and $\overline{\mu}(y)$ has been defined for all $y \in \Lambda$ with $r(y) < r(x)$. If x is irreducible, let $\overline{\mu}(x) = \mu(x)$. Otherwise, let $x = p_1 \vee \cdots \vee p_n$ be the unique factorization of x. Since x is reducible, we have $x \neq p_i$ for all i so that $r(p_i) < r(x)$ for all i. Thus, we may define

$$\overline{\mu}\,(x) := \sum_{k=1}^{n} (-1)^{k+1} \sum_{1 \leq i_1 < \cdots < i_k \leq n} \overline{\mu}\,(p_{i_1} \wedge \cdots \wedge p_{i_k})\,.$$

This defines an extension $\overline{\mu}$ of μ. It follows by induction that $\overline{\mu}$ is a valuation. □

Corollary 4.8. *Every valuation on $\mathcal{J}\Delta^d$ is determined by its values on the simplices.*

Proof. By Theorem 4.4, each $\mathcal{J}\Delta^d$ has unique factorization. By Lemma 4.3, the prime elements of $\mathcal{J}\Delta^d$ are exactly the principal ideals, i.e., the simplices. Since $\mathcal{J}\Delta^d$ can be graded by the rank function $r(x) = |x|$, the result follows from Lemma 4.7. □

Recall that every permutation σ of $\{0, \ldots, d\}$ induces a permutation of $\Delta^d = \mathcal{P}(\{0, \ldots, d\})$ and thus also on $\mathcal{J}\Delta^d$. We will say that a valuation μ on $\mathcal{J}\Delta^d$ is **invariant under permutations**, or just **invariant**, if for every permutation σ of $\{0, \ldots, d\}$ and every $\mathfrak{a} \in \mathcal{J}\Delta^d$, we have

$$\mu(\mathfrak{a}) = \mu\,(\sigma\mathfrak{a})\,.$$

Corollary 4.9. *Every permutation invariant valuation μ on $\mathcal{J}\Delta^d$ is determined by a function $f : \{0, \dots, d\} \to \mathbb{R}$ such that*

$$f(n) = \mu(\Delta^n). \tag{4.3}$$

The correspondence $f \mapsto \mu$ given by (4.3) defines an isomorphism between the vector space of functions $\{0, \dots, d\} \to \mathbb{R}$ and the vector space of invariant valuations on $\mathcal{J}\Delta^d$.

Proof. For $0 \le n \le d$ and $0 \le i_0 < \cdots < i_n \le d$, let $\sigma_{i_1,\dots,i_n}$ denote a permutation of $\{0, \dots, d\}$ that sends $\{i_0, \dots, i_n\}$ to $\{0, \dots, n\}$. Then for any n-simplex $x = \langle \{i_0, \dots, i_n\} \rangle$, the invariance of μ yields

$$\mu(x) = \mu(\sigma_{i_0,\dots,i_n} x) = \mu(\Delta^n).$$

In this manner, we may recover the value of μ on every n-simplex in $\mathcal{J}\Delta^d$ from the value of μ on the standard n-simplex. This proves the first statement. The second statement follows immediately from the definition of the map $f \mapsto \mu$. $\square$

We will use Corollary 4.9 to define the intrinsic valuations of the lattice $\mathcal{J}\Delta^d$. The definitions we provide are equivalent to the ones given in [KR97, Section 3.2]. For $0 \le i \le d$, define maps $f_i : \{0, \dots, d\} \to \mathbb{R}$ by

$$f_i(n) = \begin{cases} 1 & \text{if } i \le n \\ 0 & \text{if } i > n \end{cases}.$$

Notice that if $i < j$, then $f_i \cdot f_j = f_j$. For $0 \le k \le d+1$, let s_k denote the k-th symmetric polynomial. Then for $0 \le n \le d$ and $0 \le k \le d+1$,

$$\begin{aligned} s_k\left(f_0, \dots, f_d\right)(n) &= \sum_{0 \le i_1 < \cdots < i_k \le d} f_{i_1}(n) \cdots f_{i_k}(n) \\ &= \sum_{0 \le i_1 < \cdots < i_k \le n} 1 \\ &= \binom{n+1}{k}. \end{aligned}$$

For $0 \le k \le d+1$, we will define the **k-th intrinsic valuation**, denoted μ_k, to be the valuation corresponding to $s_k(f_0, \dots, f_d)$. These may be described more concretely as the valuations acting on standard simplices by

$$\mu_k(\Delta^n) = \binom{n+1}{k}. \tag{4.4}$$

For an arbitrary simplex $\langle x \rangle$, $\mu_k(\langle x \rangle)$ counts the number of $(k-1)$-faces of $\langle x \rangle$, under the convention that every non-empty simplex has a single face of rank -1. Notice that the μ_k do not depend on the ambient dimension d.

The **Euler Characteristic** is defined to be the valuation $\chi = \mu_0$. It assigns the constant value 1 to all non-empty simplices, and assigns the value 0 to the empty simplex.

The following is equivalent to [KR97, Theorem 3.2.4].

Theorem 4.10. *The intrinsic valuations $\mu_0, \ldots, \mu_d$ form a basis for the vector space of invariant valuations on $\mathfrak{J}\Delta^d$.*

Proof. By Corollary 4.9, the vector space of invariant valuations on $\mathfrak{J}\Delta^d$ has dimension $d+1$. It will thus suffice to show that $\mu_0, \ldots, \mu_d$ are linearly independent. Suppose $\alpha_0, \ldots, \alpha_d$ are such that

$$\sum_{k=0}^{d} \alpha_k \mu_k = 0.$$

We will begin by showing

$$\alpha_n = (-1)^n \alpha_0 \qquad (0 \leq n \leq d). \tag{4.5}$$

This will be done by induction on n. The case $n = 0$ is trivial. For $n = 1$, we have

$$0 = \sum_{k=0}^{d} \alpha_k \mu_k(\Delta^0) = \alpha_0 \binom{1}{0} + \alpha_1 \binom{1}{1}$$

and hence $\alpha_1 = -\alpha_0$. Now let n be given and suppose that $\alpha_k = (-1)^k \alpha_0$ for all

$k < n$. Then

$$\begin{aligned}
0 &= \sum_{k=0}^{d} \alpha_k \mu_k \left(\Delta^{n-1}\right) \\
&= \sum_{k=0}^{n} \alpha_k \binom{n}{k} \\
&= \sum_{k=0}^{n-1} (-1)^k \alpha_0 \binom{n}{k} + \alpha_n \\
&= \left(\sum_{k=0}^{n} (-1)^n - (-1)^n \right) \alpha_0 + \alpha_n \\
&= (-1)^{n+1} \alpha_0 + \alpha_n.
\end{aligned}$$

Thus, $\alpha_n = (-1)^n \alpha_0$. This concludes the proof of (4.5).

Next, observe that by (4.5),

$$0 = \sum_{k=0}^{d} \alpha_k \mu_k \left(\Delta^d\right) = \alpha_0 \sum_{k=0}^{d} (-1)^k \binom{d+1}{k} = \alpha_0 (-1)^{d+2}.$$

Thus, $\alpha_0 = 0$. Then by (4.5) again, we have $\alpha_n = 0$ for all n. This shows that $\mu_0, \ldots, \mu_d$ are linearly independent, as desired. □

We will now explore the similarities between the Euler characteristic on $\mathcal{J}\Delta^d$ and the Euler characteristic as defined in the previous chapter. Let $\Delta^n \in \mathcal{J}\Delta^d$. Then

$$\chi(\Delta^n) = 1 = \sum_{k=1}^{n+1} (-1)^{k+1} \binom{n+1}{k}.$$

This shows that our definition of Euler characteristic is equivalent to the alternating sum definition at the level of simplices. By (4.4), and since each μ_k is a valuation, we see that it holds at the level of simplicial complexes as well.

Given $x = \{i_0, \ldots, i_n\} \in \mathcal{P}(\{0, \ldots, d\})$, let us define the boundary of the n-simplex $\langle x \rangle$ by

$$\partial \langle x \rangle = \bigcup_{k=0}^{n} \langle \{i_0, \ldots, i_{k-1}, i_{k+1}, \ldots, i_n\} \rangle.$$

Then the following lemma can be proved using the same argument that was used to prove Lemma 3.9. It may thus be viewed as an abstract analogue of the formula given for the boundary of convex polytopes.

Lemma 4.11. *For any n-simplex $\langle x \rangle$, it holds that $\chi\left(\partial \langle x \rangle\right) = 1 + (-1)^{n+1}$.*

4.2 Posets

Recall that every simplicial complex Δ may be associated with a partially ordered set P whose elements are the faces of Δ, ordered by inclusion. And similarly, every poset P may be identified with a simplicial complex Δ in such a way that the chains of P are the faces of Δ. A more detailed treatment of this equivalence may be found in [Wac06, Lecture 1].

Since the Euler characteristic may be defined as an alternating sum of the k-faces in Δ, we may similarly define the Euler characteristic as an alternating sum of the k-chains in P and immediately obtain the definition of the Euler characteristic of a poset. However, we will instead follow the approach outlined in [Rot64] and define the Euler characteristic in terms of the Möbius function of a poset. This has the advantage of being easier to generalize to finite categories, which will be done in the next chapter. We will see later that both definitions coincide.

Before defining the Möbius function, we will develop some tools that will help to build an intuition that will transfer to the categorical setting. Consider the $\mathbb{R}$-algebra $R(P)$ of bivariate functions $P \times P \to \mathbb{R}$, with multiplication defined via

$$(f \cdot g)\,(a, c) = \sum_{b \in P} f(a, b) g(b, c),$$

and with addition and scalar multiplication defined pointwise. The identity of $R(P)$ is given by the **Kronecker delta**

$$\delta(a, b) = \begin{cases} 1 & \text{if } a = b \\ 0 & \text{if } a \neq b. \end{cases}$$

From this point forward, the inverse of a function $f \in R(P)$ will always refer to the function $g \in R(P)$ for which $f \cdot g = g \cdot f = \delta$ (if such a function g exists).

If we enumerate the objects of P as $p_1, \ldots, p_n$, then we may regard the maps in $R(P)$ as $n \times n$ matrices over $\mathbb{R}$. A consequence of this is that a function $f \in R(P)$ is invertible if and only if it has a left inverse if and only if it has a right inverse.

The **zeta function** is the function $\zeta \in R(P)$ defined by

$$\zeta(a, b) = \begin{cases} 1 & \text{if } a \leq b \\ 0 & \text{if } a \not\leq b. \end{cases}$$

The **Möbius function** is defined to be $\mu = \zeta^{-1}$, if it exists.

Suppose for the moment that μ exists. Let $f, g : P \to \mathbb{R}$ be arbitrary real-valued functions. Observe that

$$f(a) = \sum_{c} \delta(a,c)f(c) = \sum_{b,c} \mu(a,b)\zeta(b,c)f(c) = \sum_{\substack{b,c \\ b\leq c}} \mu(a,b)f(c),$$

and similarly,

$$g(c) = \sum_{a} g(a)\delta(a,c) = \sum_{a,b} g(a)\zeta(a,b)\mu(b,c) = \sum_{\substack{a,b \\ a\leq b}} g(a)\mu(b,c).$$

Now, fix $a_0, c_0 \in P$, and set

$$f(a) := \delta(a, c_0) \qquad \text{and} \qquad g(c) := \delta(a_0, c).$$

Then as special cases of the above equations, we have

$$\begin{aligned} \delta(a_0, c_0) &= \sum_{\substack{b,c \\ b\leq c}} \mu(a_0, b)\delta(c, c_0) = \sum_{\substack{b \\ b\leq c_0}} \mu(a_0, b) \\ \delta(a_0, c_0) &= \sum_{\substack{a,b \\ a\leq b}} \delta(a_0, a)\mu(b, c_0) = \sum_{\substack{b \\ a_0\leq b}} \mu(b, c_0). \end{aligned}$$

It follows from the above equations that for all $a_0, c_0 \in P$,

$$\mu(a_0, c_0) = \delta(a_0, c_0) - \sum_{\substack{b \\ a_0<b}} \mu(b, c_0) = \delta(a_0, c_0) - \sum_{\substack{b \\ b<c_0}} \mu(a_0, b). \tag{4.6}$$

Lemma 4.12. *For all $a, c \in P$, if $a \not\leq c$, then $\mu(a, c) = 0$.*

Proof. Let $c_0 \in P$ be arbitrary. We will show the result holds for all $a \not\leq c_0$. We will do this by induction on the size of $P_{<c_0} = \{b \in P : b < c_0\}$.

If $P_{<c_0} = \varnothing$, then by (4.6), we have for all $a \in P$ with $a \not\leq c_0$,

$$\mu(a, c_0) = 0.$$

Next, suppose $|P_{<c_0}| = n$ and the result holds for all $b \in P$ with $|P_{<b}| < n$. Let

$a_0 \in P$ with $a_0 \not\leq c_0$. Then for all $b < c_0$, we must have $a_0 \not\leq b$. Therefore, we have

$$\mu(a_0, c_0) = -\sum_{b \in P_{<c_0}} \mu(a_0, b) = -\sum_{b \in P_{<c_0}} 0 = 0. \qquad \square$$

It follows from (4.6) and the previous lemma that whenever μ exists, it satisfies

$$\mu(a_0, c_0) = \delta(a_0, c_0) - \sum_{\substack{b \\ a_0 < b \leq c_0}} \mu(b, c_0) = \delta(a_0, c_0) - \sum_{\substack{b \\ a_0 \leq b < c_0}} \mu(a_0, b). \tag{4.7}$$

In fact, we may use (4.7) to *define* μ.

Lemma 4.13. *The Möbius function exists for every finite poset. It may be defined recursively via*

$$\mu(a_0, c_0) = \begin{cases} 0 & \text{if } a_0 \not\leq c_0 \\ 1 & \text{if } a_0 = c_0 \\ \sum\limits_{a_0 \leq b < c_0} -\mu(a_0, b) & \text{if } a_0 < c_0. \end{cases}$$

Proof. Let μ be defined as above, and let $a_0, c_0 \in P$. Then

$$(\mu \cdot \zeta)(a_0, c_0) = \sum_b \mu(a_0, b)\zeta(b, c_0) = \sum_{\substack{b \\ a_0 \leq b \leq c_0}} \mu(a_0, b).$$

If $a_0 \neq c_0$, then by definition of μ,

$$(\mu \cdot \zeta)(a_0, c_0) = \sum_{\substack{b \\ a_0 \leq b \leq c_0}} \mu(a_0, b) = \mu(a_0, c_0) + \sum_{\substack{b \\ a_0 \leq b < c_0}} \mu(a_0, b) = 0.$$

For $a_0 = c_0$, we have

$$(\mu \cdot \zeta)(a_0, a_0) = \sum_{\substack{b \\ a_0 \leq b \leq a_0}} \mu(a_0, b) = \mu(a, a) = 1.$$

This shows that μ is a left inverse for ζ. As mentioned previously, by viewing μ and ζ as matrices in a finite-dimensional vector space, we have that μ must also be a right inverse for ζ. $\square$

The **Euler characteristic** of a poset P is defined to be

$$\chi(P) := \sum_{a,b \in P} \mu(a, b).$$

Let $C_n(P)$ denote the set of n-chains in P. That is,

$$C_n(P) := \left\{(a_0, \ldots, a_n) \in P^{n+1} : a_0 < \cdots < a_n\right\}.$$

Our statement and proof of the following theorem was inspired by [Rot64, Proposition 6].

Theorem 4.14. *For any finite poset P,*

$$\chi(P) = \sum_{n=0}^{\infty} (-1)^n |C_n(P)|.$$

Proof. We will define the **incidence function** $\gamma \in R(P)$ by

$$\gamma(a,c) = \zeta(a,c) - \delta(a,c) = \begin{cases} 1 & \text{if } a < c \\ 0 & \text{if } a \not< c \end{cases}.$$

Observe that

$$\gamma^2(a,c) = \sum_{b \in P} \gamma(a,b)\gamma(b,c) = \sum_{a<b<c} 1.$$

Thus, $\gamma^2(a,c)$ counts all the 2-chains starting at a and ending at c. Similarly,

$$\gamma^n(a,c) = \sum_{\substack{b_1,\ldots,b_{n-1} \\ a<b_1<\cdots<b_{n-1}<c}} 1$$

counts all the n-chains starting at a and ending at c. A consequence of this is that $\gamma^n = 0$ whenever $n > |P|$. It then follows that

$$\mu = \zeta^{-1} = (\delta - (-\gamma))^{-1} = \sum_{n=0}^{\infty} (-1)^n \gamma^n. \tag{4.8}$$

Furthermore, the above sum is finite. The result then follows from the observation that $|C_n(P)| = \sum_{a,b} \gamma^n(a,b)$. □

A consequence of the previous theorem is that if P is a finite poset and $\Delta(P)$ is the simplicial complex induced by P, then

$$\chi(P) = \chi(\Delta(P)).$$

5 Category Theory

5.1 Background

We now seek to generalize the Euler characteristic from finite posets to finite categories. We will begin with some elementary definitions. A **category** $\mathcal{A}$ consists of

(i) a collection of **objects**, denoted ob$\mathcal{A}$;

(ii) a collection of **morphisms** (or **arrows**) between the objects of $\mathcal{A}$, denoted mor$\mathcal{A}$. Given two objects $a, a' \in$ ob$\mathcal{A}$, we write $a \xrightarrow{f} a'$ to indicate that f is a morphism from a to a'. The collection of morphisms from a to a' is denoted $\mathcal{A}(a, a')$.

(iii) a **rule of composition**, denoted $\circ$, which associates to each pair of mappings $a \xrightarrow{f} b \xrightarrow{g} c$, a map $a \xrightarrow{g \circ f} c$.

The above data are required to satisfy the following conditions:

(i) For every $a \in \mathcal{A}$, there exists a unique morphism $a \xrightarrow{1_a} a$, called the **identity** of a, that satisfies

$$1_a \circ f = f \qquad \text{and} \qquad g \circ 1_a = g$$

for every possible choice of $a' \xrightarrow{f} a \xrightarrow{g} a''$, To simplify the notation, it is common to suppress the object in question when writing the identities (e.g. 1 instead of 1_a).

(ii) For all morphisms $a \xrightarrow{f} b \xrightarrow{g} c \xrightarrow{h} d$,

$$h(gf) = (hg)f.$$

Example 5.1. We present a few examples of frequently encountered categories.

(i) The category Set, where the objects are "small" sets (relative to some given Grothendieck universe), and the morphisms are functions between sets. The identity map of a set X is the function 1_X defined by $1_X(x) = x$, and composition of functions is defined in the usual manner: $(g \circ f)(x) = g(f(x))$. It is trivial to check that all the required axioms are satisfied.

(ii) The category Top, where the objects are topological spaces, and the morphismss are continuous maps between topological spaces. The identity map is again defined by $1_X(x) = x$, and it is straightforward to show that this map is continuous. The rule of composition again corresponds to the ordinary composition of functions, and this definition makes sense for because the composotion of two continuous functions is continuous.

(iii) The category Grp, where the objects are groups and the morphisms are group homomorphisms.

Notice how in all of the previous examples, the objects of a category were also sets and the morphisms were functions between sets. But this need not be the case in general, as the following examples illustrate.

Example 5.2. (i) A group G can be viewed as a category with a single object $*$. We will denote this category by G. The morphisms of G correspond to the elements of the group G. The rule of composition of morphisms in G is given by multiplication of the corresponding elements in G, and the identity map of the single object $*$ of G corresponds to the identity element of G.

(ii) A poset P can be viewed as a category, temporarily denoted by P. The objects of P are the elements of P, and the morphisms of P are defined by

$$\mathsf{P}(a, b) = \begin{cases} \{(a, b)\} & \text{if } a \leq b \\ \varnothing & \text{if } a \not\leq b \end{cases}.$$

In other words, if $a \leq b$, then there is a single morphism $a \to b$; if $a \not\leq b$, then there are no morphisms from a to b. Since the ordering on P is required to satisfy $a \leq a$ for all $a \in P$, we see that P has an identity element. The rule of composition is defined in the obvious way, which is possible because $\leq$ is transitive.

Before doing so, we will first define a few more concepts for a general category $\mathcal{A}$. We will cover them briefly for the sake of completeness and refer the interested reader to [Rie16].

(i) $\mathcal{A}$ is **finite** if mor$\mathcal{A}$ is a finite set. Notice that this implies ob$\mathcal{A}$ is also a finite set.

(ii) A morphism $a \xrightarrow{f} a'$ is an **isomorphism** if there exists a morphism $a' \xrightarrow{g} a$ satisfying

$$gf = 1_a \qquad \text{and} \qquad fg = 1_{a'}.$$

(iii) A **functor** is, intuitively, a "structure-preserving" map between categories. Formally, a functor F from $\mathcal{A}$ to $\mathcal{B}$, denoted $F : \mathcal{A} \to \mathcal{B}$, consists of maps $\mathrm{ob}\mathcal{A} \xrightarrow{F} \mathrm{ob}\mathcal{B}$ and $\mathrm{mor}\mathcal{A} \xrightarrow{F} \mathrm{mor}\mathcal{B}$ such that

(a) each morphism $a \xrightarrow{f} a'$ in $\mathcal{A}$ gets sent to a morphism $Fa \xrightarrow{Ff} Fa'$ in $\mathcal{B}$;

(b) $F(1_a) = 1_{Fa}$ for all $a \in \mathcal{A}$;

(c) $F(g \circ f) = F(g) \circ F(f)$ for all $a \xrightarrow{f} a' \xrightarrow{g} a''$.

(iv) Given two functors $F, G : \mathcal{A} \to \mathcal{B}$, a **natural transformation** from F to G, denoted $\alpha : F \Rightarrow G$, consists of a map $Fa \xrightarrow{\alpha_a} Ga$ for every $a \in \mathcal{A}$ such that for every map $a \xrightarrow{f} a'$,

$$\alpha_{a'} Ff = Gf \alpha_a.$$

This condition may also be expressed by stating that the following diagram commutes, meaning that both paths from Fa to Ga' are equal:

$$\begin{array}{ccc} Fa & \xrightarrow{\alpha_a} & Ga \\ {\scriptstyle Ff}\downarrow & & \downarrow{\scriptstyle Gf} \\ Fa' & \xrightarrow[\alpha_{a'}]{} & Ga' \end{array}$$

The maps α_a are called the **component maps** of α.

(v) A **natural isomorphism** is a natural transformation in which all the component maps are isomorphisms. We say that two functors $F, G : \mathcal{A} \to \mathcal{B}$ are **isomorphic** if there exists a natural isomorphism from F to G. In this case, we write $F \approx G$.

(vi) An **equivalence of categories** $\mathcal{A}$ and $\mathcal{B}$ consists of a pair of functions $\mathcal{A} \underset{G}{\overset{F}{\rightleftarrows}} \mathcal{B}$ such that $GF \approx 1_{\mathcal{A}}$ and $FG \approx 1_{\mathcal{B}}$. In this case, we say that $\mathcal{A}$ and $\mathcal{B}$ are **equivalent** and write $\mathcal{A} \simeq \mathcal{B}$.

(vii) The **skeleton** of a category $\mathcal{A}$, denoted $\mathsf{sk}(\mathcal{A})$, is a category that is equivalent to $\mathcal{A}$ and contains only one object in each isomorphism class. While $\mathsf{sk}(\mathcal{A})$

is not unique, it is unique up to isomorphism. A category is **skeletal** if it is isomorphic to its own skeleton. This means that no two distinct objects in $\mathcal{A}$ are isomorphic.

We will now introduce some definitions that will facilitate the process of extending the definition of the Euler characteristic to certain kinds of finite categories.

(i) We say that a category $\mathcal{A}$ is **EI** if all its endomorphisms are isomorphisms.

(ii) Given $a, b \in \mathcal{A}$ and $n \geq 0$, an n-**path** from a to b is a sequence of maps

$$a = a_0 \xrightarrow{f_1} a_1 \xrightarrow{f_2} \cdots \xrightarrow{f_n} a_n = b. \tag{5.1}$$

(iii) A **proper path** is an n-path of the form (5.1) where $n \geq 1$ and $a_i \neq a_{i+1}$ for all i.

(iv) An n-**cycle** is an n-path from a to a.

(v) $\mathcal{A}$ is **acyclic** if $\mathcal{A}$ does not admit any proper cycles. (Notice that every category has 1-cycles of the form $a \xrightarrow{1_a} a$.)

We will only concern ourselves with finite categories. Notice that every poset is an EI category. As the following lemma will help to illuminate, EI categories behave similarly to posets in many regards.

Example 5.2(ii) provides a motivation for attempting to extend the Euler characteristic to finite categories: the Euler characteristic has already been defined for certain finite categories, why not try to define it for a few more? This extension can be further motivated by considering the **classifying space**, or **geometric realization**, of a category. This may be defined as follows. For a given category $\mathcal{A}$, we may define the **nerve of** $\mathcal{A}$, denoted $N(\mathcal{A})$ to be the simplicial set with

(i) Objects of the form $C_n(\mathcal{A}) = \{n\text{-paths in } \mathcal{A}\}$ for each n.

(ii) Degeneracy maps given by sending

$$a_0 \xrightarrow{f_1} \cdots \xrightarrow{f_k} a_k \xrightarrow{f_k} \cdots \xrightarrow{f_n} a_n$$

to

$$a_0 \xrightarrow{f_1} \cdots \xrightarrow{f_{k-1}} a_{k-1} \xrightarrow{f_{k+1}f_k} a_{k+1} \xrightarrow{f_{k+2}} \cdots \xrightarrow{f_n} a_n,$$

for each $1 \leq k \leq n-1$. The degeneracy maps for $k = 0$ and $k = n$ are given by removing a_0 and a_n, respectively.

(iii) Simplicial identity maps given by inserting an identity map (of $\mathcal{A}$) into an n-chain.

We may then form the **geometric realization** of $N(\mathcal{A})$ by associating each $C_n(\mathcal{A})$ with an n-simplex and using the degeneracy and simplicial identity maps to glue these simplices together. This construction produces a CW complex and in effect generalizes the notion of the geometric realization of a simplicial complex. Some of these geometric realizations, being topological spaces, already have a notion of Euler characteristic. So there is in a very real sense, already a way of assigning an Euler characteristic to a finite category. As mentioned in [Lei06], however, the Euler characteristic may be defined for categories whose classifying spaces have infinitely many non-trivial homology groups. In [Lei06, Proposition 2.11], it is shown that the categorical definition of Euler characteristic agrees with the topological definition of the Euler characteristic for EI categories.

We will conclude this section with a characterization of finite EI categories which be useful in our studies of the Möbius function of a finite category. It is a restatement of [Lei06, Lemma 1.3].

Lemma 5.3. *Let $\mathcal{A}$ be a finite category. The following are equivalent.*

(i) $\mathcal{A}$ is an EI category.

(ii) Every idempotent in $\mathcal{A}$ is an identity.

(iii) For every $f : a \to a$, there exists a $k > 0$ such that $f^k = 1$.

(iv) Every cycle in $\mathcal{A}$ consists entirely of isomorphisms.

(v) $\mathsf{sk}(\mathcal{A})$ is EI and acyclic.

Proof.

(i) $\rightsquigarrow$ (ii): Let $f : a \to a$ be idempotent. Since f is an isomorphism by assumption,

$$f = f^2 f^{-1} = f f^{-1} = 1.$$

(ii) $\rightsquigarrow$ (iii): Let $f : a \to a$. By viewing f as an element of the finite monoid $\{f^n : n \in \mathbb{Z}_{\geq 0}\}$, it follows that there exists a $k > 0$ for which f^k is idempotent. By (i), we have $f^k = 1$.

(iii) $\rightsquigarrow$ (i): Let $f : a \to a$ and let k be the smallest positive integer such that $f^k = 1$. Then $f^{-1} = f^{k-1}$.

(i) $\rightsquigarrow$ (iv): Notice that all 1-cycles consist of isomorphisms by assumption. Next, suppose we are given a 2-cycle

$$a \underset{g}{\overset{f}{\rightleftarrows}} b$$

By (iii), we can find $p, q > 0$ such that

$$(gf)^p = 1 \qquad \text{and} \qquad (fg)^q = 1,$$

and thus

$$f \circ g(fg)^{q-1} = 1 \qquad \text{and} \qquad (gf)^{p-1}g \circ f = 1.$$

For $n \geq 3$, an n-cycle $a_0 \xrightarrow{f_1} a_1 \xrightarrow{f_2} \cdots \xrightarrow{f_n} a_0$ induces a 2-cycle

$$a_0 \xrightarrow{f_1} a_1 \xrightarrow{f_n \cdots f_2} a_0$$

to which we may apply the previous argument. The result follows by applying this argument to each a_i.

(iv) $\rightsquigarrow$ (v): It is a known result that two categories are equivalent if and only if there is a fully faithful and essentially surjective functor from one to the other. (See, for example, [Rie16, Theorem 1.5.9].) Let $F : \mathsf{sk}(\mathcal{A}) \to \mathcal{A}$ be such a functor. In this instance, this functor acts on objects by basically picking a representative of each equivalence class. Notice that every n-cycle

$$\overline{a} = \overline{a_0} \xrightarrow{\overline{f_1}} \cdots \xrightarrow{\overline{f_n}} \overline{a_n} = \overline{a}$$

in $\mathsf{sk}(\mathcal{A})$ gives rise to an n-cycle

$$a = a_0 \xrightarrow{f_1} \cdots \xrightarrow{f_n} a_n = a$$

in $\mathcal{A}$, where $a_i = F(\overline{a_i})$ and $f_i = F(\overline{f_i})$. By (iv), each of the maps f_i is an isomorphism. Since F is fully faithful, it follows that the $\overline{f_i}$ are also isomorphisms. Therefore, we must have $\overline{a_i} = \overline{a_{i+1}}$ for all i. This shows $\mathsf{sk}(\mathcal{A})$ is acyclic. Furthermore, since a 1-cycle is the same as an endomorphism, we see that $\mathsf{sk}(\mathcal{A})$ is EI.

(v) $\rightsquigarrow$ (i): Since $\mathsf{sk}(\mathcal{A})$ is EI, so is $\mathcal{A}$. □

Notice that it is trivial to prove directly that (i)-(v) of Lemma 5.3 hold for every finite poset.

5.2 The Möbius Function of a Finite Category

We will now generalize the machinery of Section 4.2 to arbitrary finite categories. Much of the material presented in the rest of this chapter was first introduced by Leinster in [Lei06]. However, we will present things in a slightly different manner, trying to stress first and foremost the similarities between the categorical definition and the combinatorial definitions.

Let $R(\mathcal{A})$ denote the collection of maps

$$\mathrm{ob}\mathcal{A} \times \mathrm{ob}\mathcal{A} \to \mathbb{R},$$

viewed as a $\mathbb{R}$-algebra with pointwise scalar multiplication and addition, and with multiplication given by

$$(f \cdot g)(a, c) := \sum_{b \in \mathcal{A}} f(a, b) g(b, c).$$

As in the case of posets, this multiplication is generally not commutative. The identity element is given by the Krönecker delta function:

$$\delta(a, b) := \begin{cases} 1 & \text{if } a = b \\ 0 & \text{if } a \neq b. \end{cases}$$

It will at times be conveninent to view a function in $R(\mathcal{A})$ as a matrix with coefficients in $\mathbb{R}$.

The **zeta function** of $\mathcal{A}$ is given by

$$\zeta(a, b) = |\mathcal{A}(a, b)|.$$

Notice that if $\mathcal{A}$ is a poset, then the above definition of ζ agrees with the definition given in Section 4.2. If ζ is invertible, then we denote the inverse by

$$\mu := \zeta^{-1}$$

and call it the **Möbius function** of $\mathcal{A}$. However, it is not always the case that the zeta function of an arbitrary category $\mathcal{A}$ is invertible. Indeed, if $\mathcal{A}$ is not skeletal, then we can find $a, a' \in \mathcal{A}$ with $a \neq a'$ and $a \approx a'$. Then the matrix representation of ζ will have two identical rows and will therefore not be invertible. However, as we will see, all finite skeletal EI categories have a Möbius function. We will define the **incidence function** γ by

$$\gamma(a, b) := \zeta(a, b) - |\mathrm{End}(a)|\, \delta(a, b).$$

It follows from the definitions that

$$\gamma(a, b) = \begin{cases} 0 & \text{if } a = b \\ |\mathcal{A}(a, b)| & \text{if } a \neq b, \end{cases}$$

and thus the incidence function of a category may be viewed as a generalization of the incidence function of a poset.

It will occasionally be useful to embed a single-variate function $f : \mathrm{ob}\mathcal{A} \to \mathbb{R}$ to a function in $R(\mathcal{A})$. We will do this by "embedding along the diagonal":

$$f(a, b) := f(a)\delta(a, b).$$

Under this convention, the incidence function γ may be written as $\gamma = \zeta - |\mathrm{End}|$. We also have

$$\gamma^2(a, c) = \sum_{b \in \mathcal{A}} \gamma(a, b)\gamma(b, c),$$

so that $\gamma^2(a, c)$ counts the proper 2-paths from a to c. And by induction, it follows that $\gamma^n(a, c)$ counts the proper n-paths from a to c.

Let us further adopt the notation

$$\|a\| := |\mathrm{End}(a)|\,.$$

The following is a (slight) generalization of [Lei06, Theorem 1.4]. The proof was obtained by adapting the proof of [Rot64, Proposition 6] to the categorical setting.

Theorem 5.4. *Suppose that $\mathcal{A}$ is a finite skeletal category such that the matrix $\left[\frac{\gamma(a,b)}{|\mathrm{End}(b)|}\right]_{a,b\in\mathcal{A}}$ has eigenvalues strictly less than 1. Then $\mathcal{A}$ has Möbius function given by*

$$\mu(a,b) = \frac{\delta(a,b)}{\|a\|} - \frac{\gamma(a,b)}{\|a\|\,\|b\|} + \sum_{n=2}^{\infty}(-1)^n \sum_{a_1,\dots,a_{n-1}} \frac{\gamma(a,a_1)\cdots\gamma(a_{n-1},b)}{\|a\|\,\|a_1\|\cdots\|a_{n-1}\|\,\|b\|} \tag{5.2}$$

The right side of (5.2) may be expanded as

$$\begin{aligned}\mu(a,b) = \frac{\delta(a,b)}{\|a\|} - \frac{\gamma(a,b)}{\|a\|\,\|b\|} + \sum_{a_1} \frac{\gamma(a,a_1)\gamma(a_1,b)}{\|a\|\,\|a_1\|\,\|b\|} \\ - \sum_{a_1,a_2} \frac{\gamma(a,a_1)\gamma(a_1,a_2)\gamma(a_2,b)}{\|a\|\,\|a_1\|\,\|a_2\|\,\|b\|} + \cdots .\end{aligned}$$

In other words, the n-th term in the sum corresponds to the proper n-paths $a = a_0 \to \cdots \to a_n = b$ from a to b, where each such path is inversely scaled by the number of endomorphisms at each a_i.

Proof. Since

$$|\mathrm{End}|^{-1}(b,c) = \frac{\delta(b,c)}{|\mathrm{End}(c)|},$$

it follows that

$$\left(\gamma \cdot |\mathrm{End}|^{-1}\right)(a,c) = \sum_b \gamma(a,b)\,|\mathrm{End}|^{-1}(b,c) = \frac{\gamma(a,c)}{|\mathrm{End}(c)|}.$$

Then by our hypothesis,

$$\begin{aligned}|\mathrm{End}|^{-1} \cdot \sum_{n=0}^{\infty}(-1)^n \left(\gamma \cdot |\mathrm{End}|^{-1}\right)^n &= |\mathrm{End}|^{-1} \cdot \left[\delta - \left(-\gamma \cdot |\mathrm{End}|^{-1}\right)\right]^{-1} \\ &= \left(\left[\delta - \left(-\gamma \cdot |\mathrm{End}|^{-1}\right)\right] \cdot |\mathrm{End}|\right)^{-1} \\ &= (|\mathrm{End}| + \gamma)^{-1} \\ &= \zeta^{-1} = \mu.\end{aligned}$$

Next,

$$\left(\gamma \cdot |\mathrm{End}|^{-1}\right)^2 (a,c) = \sum_b \left(\gamma \cdot |\mathrm{End}|^{-1}\right)(a,b) \left(\gamma \cdot |\mathrm{End}|^{-1}\right)(b,c)$$
$$= \sum_b \frac{\gamma(a,b)}{|\mathrm{End}(b)|} \frac{\gamma(b,c)}{|\mathrm{End}(c)|}.$$

So that by induction,

$$\left(\gamma \cdot |\mathrm{End}|^{-1}\right)^n (a_0,a_n) = \sum_{a_1,\ldots,a_{n-1}} \frac{\gamma(a_0,a_1)}{|\mathrm{End}(a_1)|} \cdots \frac{\gamma(a_{n-1},a_n)}{|\mathrm{End}(a_n)|}.$$

We conclude that

$$\begin{aligned}\mu(a,b) &= \left[|\mathrm{End}|^{-1} \cdot \sum_{n=0}^{\infty} (-1)^n \left(\gamma \cdot |\mathrm{End}|^{-1}\right)^n \right](a,b) \\ &= \frac{\delta(a,b)}{|\mathrm{End}(a)|} - \frac{\gamma(a,b)}{|\mathrm{End}(a)|\,|\mathrm{End}(a)|} \\ &\quad + \frac{1}{|\mathrm{End}(a)|} \sum_{n=2}^{\infty} (-1)^n \sum_{a_1,\ldots,a_{n-1}} \frac{\gamma(a,a_1)}{|\mathrm{End}(a_1)|} \cdots \frac{\gamma(a_{n-1},b)}{|\mathrm{End}(b)|}\end{aligned}$$ □

In the case when $\mathcal{A}$ is a poset, Theorem 5.4 reduces to the much simpler identity

$$\mu(a,b) = \sum_{n=0}^{\infty} (-1)^n \gamma^n(a,b),$$

with the understanding that $\gamma^0 = \delta$. It may thus be viewed as a generalization of the formula (4.8) for posets. If (5.2) holds, then we see at once that $\mathcal{A}(a,b) = \varnothing$ implies $\mu(a,b) = 0$, a property shared by the Möbius function of posets.

Notice that even for large values of n, the sums

$$\sum_{a_1,\ldots,a_{n-1}} \frac{\gamma(a,a_1)\gamma(a_1,a_2)\cdots\gamma(a_{n-2},a_{n-1})\gamma(a_{n-1},b)}{\|a\|\,\|a_1\|\cdots\|a_{n-1}\|\,\|b\|} \tag{5.3}$$

may be nonzero. Indeed, if $\mathcal{A}$ contains proper cycles, then it may very well be the case that the right side of (5.2) contains infinitely many nonzero terms. On the other hand, if $\mathcal{A}$ is acyclic, then for all $n > |\mathrm{ob}\mathcal{A}|$, the terms (5.3) will be equal to 0. By Lemma 5.3 we thus obtain the following result, which is equivalent to [Lei06, Theorem 1.4].

Theorem 5.5. *Every finite EI category has a Möbius function given by (5.2).*

In the case when $\mathcal{A}$ does contain cycles and has a Möbius function, Theorem 5.4 tells us that $\mathcal{A}$ cannot have "too many" distinct cycles, relative to the number of endomorphisms of the objects involved in those cycles. The following example illustrates this. It also reveals that Theorem 5.4 is a "proper" generalization of [Lei06, Theorem 1.4].

Example 5.6. Pick two positive integers $k < \ell$. We will construct a category $\mathcal{A}$ that satisfies the following.

(i) $\mathcal{A}$ is skeletal category with k objects.

(ii) $\mathcal{A}$ is not EI.

(iii) $\mathcal{A}$ has a Möbius function.

(iv) For any $a, b \in \text{ob}\mathcal{A}$,

$$|\mathcal{A}(a,b)| = \begin{cases} 1 & \text{if } a \neq b \\ \ell & \text{if } a = b. \end{cases}$$

This will be done as follows. Let $\mathcal{A}$ have objects $o_1, \dots, o_k$. For $i \neq j$, let there be a unique morphism $f_{i,j} : o_i \to o_j$. For each $o_i \in \text{ob}\mathcal{A}$, let $\text{End}(o_i)$ be a distinct copy of $\mathbb{Z}_\ell$. Composition will be defined as follows.

Case 1: For $m, n \in \text{End}(o_i)$, define

$$m \circ n := m \cdot n \in \mathbb{Z}_\ell.$$

Case 2: Given $o_i \xrightarrow{f_{i,j}} o_j$, define the composition of $f_{i,j}$ with respect to $m \in \text{End}(o_i)$ and $n \in \text{End}(o_j)$ by

$$f_{i,j} \circ m = f_{i,j} = n \circ f_{i,j}.$$

Case 3: Define the composition of $o_r \xrightarrow{f_{r,s}} o_s \xrightarrow{f_{s,t}} o_t$ to be

$$f_{s,t} \circ f_{r,s} = \begin{cases} 0 \in \text{End}(o_r) & \text{if } r = t \\ f_{r,t} & \text{if } r \neq t. \end{cases}$$

It is clear that composition is associative and that for all i, the map $1 \in \text{End}(o_i)$ is an identity. Thus, $\mathcal{A}$ is a category.

Next, observe that for all i, the map $0 \in \mathrm{End}(o_i)$ is not an isomorphism. Hence, no two objects of $\mathcal{A}$ are isomorphic, proving (i) and (ii). Furthermore, (iv) holds by definition. Lastly, the matrix $[|\mathrm{End}|^{-1} \circ \gamma]$ has zeros on the diagonal and entries $\frac{1}{\ell}$ elswhere; thus its eigenvalues are $\frac{-1}{\ell}$ and $\frac{k-1}{\ell}$ so that (iii) holds by Theorem 5.4 and the fact that $k-1 < \ell$.

It is also possible to use Theorem 5.4 to explicitly compute the Möbius function of this category. By (iv) and (5.2),

$$\begin{aligned}\mu(a,b) &= \frac{\delta(a,b)}{\ell} + \sum_{n=1}^{\infty}(-1)^n \sum_{a_1,\ldots,a_{n-1}} \frac{\gamma(a,a_1)\cdots\gamma(a_{n-1},b)}{\ell^{n+1}} \\ &= \sum_{n=0}^{\infty}(-1)^n \frac{\gamma^n(a,b)}{\ell^{n+1}}\end{aligned}$$

Since

$$\gamma(a,b) = \begin{cases} 1 & \text{if } a \neq b \\ 0 & \text{if } a = b, \end{cases}$$

we see that $\gamma^n(a,b)$ counts the number of ways of choosing $a_1, \ldots, a_{n-1}$ in such a way that $a \neq a_1$, $a_{n-1} \neq b$, and $a_i \neq a_{i+1}$ for all i. To simplify the notation a bit, let $P_n(\mathcal{A})$ temporarily denote the set of $\mathbf{a} = (a_1, \ldots, a_{n-1})$ with $a_i \neq a_{i+1}$. Notice that

$$|\{\mathbf{a} \in P_n(\mathcal{A}) : a \neq a_1\}| = (k-1)^{n-1},$$

as there are $(k-1)$ choices for a_1, then $(k-1)$ choices for a_2, and so on. Furthermore, for $n \geq 3$,

$$|\{\mathbf{a} \in P_n(\mathcal{A}) : a \neq a_1, a_{n-1} = b\}| = \gamma^{n-1}(a,b),$$

since both sets count the number of ways of choosing $a_1, \ldots, a_{n-2}$ with $a \neq a_1, a_{n-2} \neq b$, and $a_i \neq a_{i+1}$ for all i. It follows that for $n \geq 3$, we obtain the recursive formula

$$\begin{aligned}\gamma^n(a,b) &= |\{\mathbf{a} \in P_n(\mathcal{A}) : a \neq a_1, a_{n-1} \neq b\}| \\ &= |\{\mathbf{a} \in P_n(\mathcal{A}) : a \neq a_1\}| - |\{\mathbf{a} \in P_n(\mathcal{A}) : a \neq a_1, a_{n-1} = b\}| \\ &= (k-1)^{n-1} - \gamma^{n-1}(a,b).\end{aligned}$$

For $n = 2$, we have

$$\gamma^2(a,b) = |\{a_1 \in \mathcal{A} : a \neq a_1, a_1 \neq b\}| = \begin{cases} k-1 & \text{if } a = b \\ k-2 & \text{if } a \neq b \end{cases}.$$

Hence, the relation

$$\gamma^n(a,b) = (k-1)^{n-1} - \gamma^{n-1}(a,b)$$

holds for all $n \geq 2$ and all $a, b \in \mathcal{A}$. By unwinding the above equality, we see that for $n \geq 1$,

$$\begin{aligned} \gamma^n(a,b) &= (k-1)^{n-1} - (k-1)^{n-2} + \cdots + (-1)^{n-2}(k-1) + (-1)^{n-1}\gamma(a,b) \\ &= (-1)^{n-1}\gamma(a,b) + \sum_{i=0}^{n-2}(-1)^i(k-1)^{n-1-i} \\ &= (-1)^{n-1}\gamma(a,b) + \sum_{i=2}^{n}(-1)^{n-i}(k-1)^{i-1}. \end{aligned}$$

For $a \neq b$, we also have $\gamma(a,b) = (k-1)^0$ so that

$$\gamma^n(a,b) = \sum_{i=1}^{n}(-1)^{n-i}(k-1)^{i-1},$$

and hence

$$\mu(a,b) = \sum_{n=1}^{\infty}\sum_{i=1}^{n}\frac{(-1)^i(k-1)^{i-1}}{\ell^{n+1}}.$$

Meanwhile, for $a = b$, we have $\gamma(a,a) = 0$ so that

$$\gamma^n(a,a) = \sum_{i=2}^{n}(-1)^{n-i}(k-1)^{i-1},$$

and hence

$$\mu(a,a) = \frac{1}{\ell} + \sum_{n=2}^{\infty}\sum_{i=2}^{n}\frac{(-1)^i(k-1)^{i-1}}{\ell^{n+1}}.$$

We conclude this section with a lemma that generalizes (4.7). This begins with the observation that when P is a poset and $a \neq c$, we have

$$\mu(a,c) = -\sum_{a \leq b < c}\mu(a,b) = -\sum_{b}\mu(a,b)\gamma(b,c)$$

and similarly

$$\mu(a,c) = -\sum_{a<b\leq c} \mu(b,c) = -\sum_b \gamma(a,b)\mu(b,c)$$

Lemma 5.7. *Suppose that the conclusions of Theorem 5.4 hold. Then for all $a, c \in \mathcal{A}$ with $a \neq c$,*

$$\mu(a,c) = -\sum_b \mu(a,b)\frac{\gamma(b,c)}{\|c\|} = -\sum_b \frac{\gamma(a,b)}{\|a\|}\mu(b,c).$$

Proof. Observe that for all $a \neq c$,

$$\begin{aligned}
&\sum_b \mu(a,b)\frac{\gamma(b,c)}{\|c\|} \\
=&\sum_b \frac{\gamma(b,c)}{\|c\|}\left[\frac{\delta(a,b)}{\|a\|} - \frac{\gamma(a,b)}{\|a\|\,\|b\|} + \sum_{n=2}^{\infty}(-1)^n \sum_{a_1,\ldots,a_{n-1}} \frac{\gamma(a,a_1)\cdots\gamma(a_{n-1},b)}{\|a\|\cdots\|b\|}\right] \\
=&\frac{\gamma(a,c)}{\|a\|\,\|c\|} - \sum_b \frac{\gamma(a,b)\gamma(b,c)}{\|a\|\,\|b\|\,\|c\|} + \sum_{n=2}^{\infty}(-1)^n \sum_{a_1,\ldots,a_{n-1},b} \frac{\gamma(a,a_1)\cdots\gamma(a_{n-1},b)\gamma(b,c)}{\|a\|\cdots\|b\|\,\|c\|} \\
=&-\mu(a,c).
\end{aligned}$$

The proof of the second equality is similar. □

5.3 The Euler Characteristic of a Finite Category

The following definition of the Euler characteristic (and its justification) is due to [Lei06].

Let $\mathcal{A}$ be a finite category. A **weighting** on $\mathcal{A}$ is a function $\omega : \text{ob}\mathcal{A} \to \mathbb{R}$ such that

$$\sum_b \zeta(a,b)\omega(b) = 1 \qquad (a \in \text{ob}\mathcal{A}).$$

A **coweighting** is a weighting on $\mathcal{A}^{\text{op}}$; that is, a function $\sigma : \text{ob}\mathcal{A} \to \mathbb{R}$ such that

$$\sum_a \sigma(a)\zeta(a,b) = 1 \qquad (b \in \text{ob}\mathcal{A}).$$

Note that if $\mathcal{A}$ has both a weighting ω and a coweighting σ, then

$$\sum_b \omega(b) = \sum_b \left(\sum_a \sigma(a)\zeta(a,b)\right)\omega(b) = \sum_a \sigma(a)\sum_b \zeta(a,b)\omega(b) = \sum_a \sigma(a).$$

A consequence of this is that the value $\sum_b \omega(b)$ is independent of the weighting

ω (provided that a coweighting exists). We are therefore justified in adopting the following definition. If $\mathcal{A}$ has both a weighting and a coweighting, then we say that $\mathcal{A}$ **has an Euler characteristic**. We further define the **Euler characteristic** of $\mathcal{A}$ to be

$$\chi(\mathcal{A}) := \sum_b \omega(b),$$

where ω is a weighting on $\mathcal{A}$. The following is due to [Lei06].

Lemma 5.8. *The following are equivalent.*

(i) $\mathcal{A}$ has a Möbius function.

(ii) $\mathcal{A}$ has a unique weighting ω.

(iii) $\mathcal{A}$ has a unique coweighting σ.

Furthermore, these functions are related via

$$\omega(b) = \sum_c \mu(b, c) \qquad \text{and} \qquad \sigma(b) = \sum_a \mu(a, b). \tag{5.4}$$

Proof. For the first part of the proof, fix a total ordering on the objects of $\mathcal{A}$, so that $\mathrm{ob}\mathcal{A} = \{a_0, \ldots, a_n\}$. We may view ζ as a matrix with entry (i, j) given by $\zeta(a_i, a_j)$. The existence of a unique (co)weighting amounts to stating that ζ is injective. This concludes the first part of the proof. Lastly, (5.4) follows from the fact that for all $a \in \mathrm{ob}\mathcal{A}$,

$$1 = \sum_c \delta(a, c) = \sum_c \sum_b \zeta(a, b)\mu(b, c) = \sum_b \zeta(a, b) \left(\sum_c \mu(b, c) \right).$$

□

A consequence of the previous lemma is that if $\mathcal{A}$ has a Möbius function, then it has Euler characteristic given by

$$\chi(\mathcal{A}) = \sum_{a,b \in \mathcal{A}} \mu(a, b).$$

Remark 5.9. Consider the case when $\mathcal{A} = P$ is a poset. Then, as mentioned previously, P has a categorical Möbius function, and by Theorem 5.4 and the subsequent remarks, it holds that

$$\mu(a, b) = \sum_{n=0}^{\infty} (-1)^n \gamma^n(a, b).$$

Note that since P is acyclic, the above expression is a finite sum. Recall from Section 4.2 that

$$\sum_{a,b} \gamma^n(a,b) = |C_n(P)|,$$

where $C_n(P)$ denotes the set of n-chains in P. Thus, we see that in this case, the definition of the Euler characteristic reduces to

$$\chi(P) = \sum_{a,b} \mu(a,b) = \sum_{n=1}^{\infty} (-1)^n |C_n(P)|.$$

We may thus recover the alternating sum formula from the categorical definition of the Euler characteristic, in the special case that P is a poset.

The following example is due to [Lei06]. It shows that $\chi(\mathcal{A})$ need not be an integer.

Example 5.10. Let G be a finite group, regarded as a category with a single object $*$. Then $\zeta(*,*) = |G|$ and hence

$$\chi(G) = \mu(*,*) = \frac{1}{|G|}.$$

We will now turn our attention to showing that the Euler characteristic is a categorical invariant. The content of the following three lemmas and corollary are due to [Lei06], but we have altered their presentation and proofs a bit.

Lemma 5.11. *Let $\mathcal{A}$ be a finite category. Then $\mathcal{A}$ admits a weighting if and only if* $\mathsf{sk}(\mathcal{A})$ *admits a weighting.*

Proof. Let $\overline{a}$ denote the equivalence class of an object $a \in \mathcal{A}$. Without loss of generality, we may suppose the object set of $\mathsf{sk}(\mathcal{A})$ consists of the equivalence classes $\overline{a}$ for $a \in \mathcal{A}$. We will use the fact that if $a \approx a'$ and $b \approx b'$, then $\zeta(a,b) = \zeta(a',b')$.

($\rightsquigarrow$): Suppose we are given a weighting ω on $\mathcal{A}$. Define a weighting ω_{sk} on $\mathsf{sk}(\mathcal{A})$ by

$$\omega_{\mathsf{sk}}\left(\overline{b}\right) = \sum_{b' \in \overline{b}} \omega(b').$$

Note that this is indeed a weighting since for every $\overline{a} \in \mathsf{sk}(\mathcal{A})$, we have

$$\begin{aligned}\sum_{\overline{b} \in \mathsf{sk}(\mathcal{A})} \zeta\left(\overline{a}, \overline{a}\right) \omega_{\mathsf{sk}}\left(\overline{b}\right) &= \sum_{\overline{b} \in \mathsf{sk}(\mathcal{A})} \sum_{b' \in \overline{b}} \zeta\left(\overline{a}, \overline{b}\right) \omega\left(b'\right) \\ &= \sum_{b' \in \mathcal{A}} \zeta\left(a, b'\right) \omega\left(b'\right) \\ &= 1.\end{aligned}$$

($\leftsquigarrow$): Conversely, suppose we are given a weighting ω_{sk} on $\mathsf{sk}(\mathcal{A})$. We can define a weighting ω on $\mathcal{A}$ by

$$\omega(b) = \frac{\omega_{\mathsf{sk}}\left(\overline{b}\right)}{\left|\overline{b}\right|}.$$

Then for every $a \in \mathcal{A}$,

$$\begin{aligned}\sum_{b \in \mathcal{A}} \zeta(a, b) \omega\left(b\right) &= \sum_{\overline{b} \in \mathsf{sk}(\mathcal{A})} \sum_{b' \in \overline{b}} \zeta(a, b') \frac{\omega_{\mathsf{sk}}\left(\overline{b}\right)}{\left|\overline{b}\right|} \\ &= \sum_{\overline{b} \in \mathsf{sk}(\mathcal{A})} \zeta(\overline{a}, \overline{b}) \omega_{\mathsf{sk}}\left(\overline{b}\right) \\ &= 1.\end{aligned}$$

□

Since equivalent categories have isomorphic skeletons, we also have the following

Corollary 5.12. *Let $\mathcal{A}$ and $\mathcal{B}$ be finite categories and suppose $\mathcal{A}$ is equivalent to $\mathcal{B}$. Then $\mathcal{A}$ admits a weighting if and only if $\mathcal{B}$ does.*

By duality, the same is also true of coweightings. We thus see that the existence of Euler characteristic is a categorical invariant. By closely examining the proof of Lemma 5.11, we may show that the value of the Euler characteristic is also a categorical invariant.

Lemma 5.13. *For any finite category $\mathcal{A}$ with Euler characteristic,*

$$\chi(\mathcal{A}) = \chi(\mathsf{sk}(\mathcal{A})).$$

Proof. Let ω be a weighting on $\mathcal{A}$. By the proof of Lemma 5.11, we have a weighting ω_{sk} on $\mathsf{sk}(\mathcal{A})$ given by

$$\omega_{\mathsf{sk}}(\overline{b}) = \sum_{b' \in \overline{b}} \omega(b').$$

Hence,

$$\chi(\mathcal{A}) = \sum_{b \in \mathcal{A}} \omega(b) = \sum_{\bar{b} \in \mathsf{sk}(\mathcal{A})} \sum_{b' \in \bar{b}} \omega(b') = \sum_{\bar{b} \in \mathsf{sk}(\mathcal{A})} \omega_{\mathsf{sk}}\left(\bar{b}\right) = \chi(\mathsf{sk}(\mathcal{A})). \qquad \square$$

Thus, any two equivalent categories have equal Euler characteristic. The above lemma helps motivate the reason for defining the Euler characteristic in terms of weightings and coweightings rather than the Möbius function. As previous mentioned, the existence of a Möbius function is not a categorical invariant, as only skeletal categories may have Möbius functions. However, the existence of weightings and coweightings is an invariant.

Observe that if $\mathcal{A}$ and $\mathcal{B}$ are finite categories that admit weightings $\omega_{\mathcal{A}}$ and $\omega_{\mathcal{B}}$, then we obtain a weighting ω on $\mathcal{A} + \mathcal{B}$, defined by

$$\omega(x) = \begin{cases} \omega_{\mathcal{A}}(x) & \text{if } x \in \mathcal{A} \\ \omega_{\mathcal{B}}(x) & \text{if } x \in \mathcal{B} \end{cases}.$$

The same is true of coweightings. This proves the following.

Lemma 5.14. *If $\mathcal{A}$ and $\mathcal{B}$ are categories with Euler characteristic, then $\mathcal{A} + \mathcal{B}$ also has Euler characteristic. Furthermore,*

$$\chi(\mathcal{A} + \mathcal{B}) = \chi(\mathcal{A}) + \chi(\mathcal{B}).$$

Example 5.15. Given any rational number, we may construct a category with Euler characteristic equal to that number. We will first define a category $\mathcal{A}_{-1}$ with objects $\{a, b\}$ and non-identity arrows given by

$$a \;\substack{\longrightarrow \\ \longrightarrow \\ \longrightarrow}\; b$$

The matrix representation of the zeta function on $\mathcal{A}_{-1}$ is given by

$$\zeta = \begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix}$$

and hence has inverse

$$\mu = \begin{bmatrix} 1 & 0 \\ -3 & 1 \end{bmatrix}.$$

Then by definition, $\chi(\mathcal{A}_{-1}) = -1$. By Lemma 5.14 and Example 5.10, it follows that for any positive integers k, m, n,

$$\chi\left(\sum_{i=1}^{m} \mathbb{Z}_k + \sum_{i=1}^{n} \mathcal{A}_{-1}\right) = \frac{m}{k} - n.$$

In this manner, we may construct categories whose Euler characteristic is equal to any given rational number.

Terms & Notation

References

[BBR86] Marilena Barnabei, Andrea Brini, and G.-C Rota. The theory of möbius functions. *Russian Mathematical Surveys - RUSS MATH SURVEY-ENGL TR*, 41:135–188, 06 1986. doi:10.1070/RM1986v041n03ABEH003326.

[Grä11] George Grätzer. *Lattice Theory: Foundation*. Birkhäuser Basel, 1 edition, 2011. doi:https://doi.org/10.1007/978-3-0348-0018-1.

[Gro78] Helmut Groemer. On the extension of additive functionals on classes of convex sets. *Pacific J. Math.*, 75, 04 1978. doi:10.2140/pjm.1978.75.397.

[KR97] Daniel A. Klain and Gian-Carlo Rota. *Introduction to Geometric Probability*. Cambridge University Press, Cambridge, UK, 1997.

[Lei06] Tom Leinster. The euler characteristic of a category, 2006. URL: https://arxiv.org/abs/math/0610260, arXiv:math/0610260.

[Mun09] Emanuele Munarini. On the euler characteristic of finite distributive lattices. In Ernesto Damiani, Ottavio D'Antona, Vincenzo Marra, and Fabrizio Palombi, editors, *From Combinatorics to Philosophy*, pages 145–166, Boston, MA, 2009. Springer US.

[Ric19] David S. Richeson. *Euler's Gem*. Princeton University Press, Princeton, 2019. URL: https://doi.org/10.1515/9780691191997 [cited 2025-01-15], doi:doi:10.1515/9780691191997.

[Rie16] Emily Riehl. *Category Theory in Context*. Dover Publications, 2016. URL: https://math.jhu.edu/~eriehl/context/.

[Rot64] Gian Carlo Rota. On the foundations of combinatorial theory i. theory of möbius functions. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 2(4):340–368, 1964. doi:10.1007/BF00531932.

[Sch86] Stephen H. Schanuel. What is the length of a potato? In *Categories in Continuum Physics*, pages 118–126, Berlin, Heidelberg, 1986. Springer Berlin Heidelberg.

[Sch13] Rolf Schneider. *Convex Bodies: The Brunn–Minkowski Theory*. Encyclopedia of Mathematics and its Applications. Cambridge University Press, 2 edition, 2013.

[SW08] Rolf Schneider and Wolfgang Weil. *Stochastic and Integral Geometry*. 2008. URL: https://link.springer.com/book/10.1007/978-3-540-78859-1, doi:https://doi.org/10.1007/978-3-540-78859-1.

[Wac06] Michelle L. Wachs. Poset topology: Tools and applications, 2006. URL: https://arxiv.org/abs/math/0602226, arXiv:math/0602226.