A Weighted Möbius Function

Derek Garton

Follow this and additional works at: https://pdxscholar.library.pdx.edu/mth_fac

Part of the Discrete Mathematics and Combinatorics Commons

Let us know how access to this document benefits you.

Citation Details

This Pre-Print is brought to you for free and open access. It has been accepted for inclusion in Mathematics and Statistics Faculty Publications and Presentations by an authorized administrator of PDXScholar. Please contact us if we can make this document more accessible: pdxscholar@pdx.edu.
A weighted Möbius function

Derek Garton

September 10, 2015

1 Introduction

Fix an odd prime \( \ell \) and let \( \mathcal{G} \) be the poset of isomorphism classes of finite abelian \( \ell \)-groups, with the relation \([A] \leq [B]\) if and only if there exists an injective group homomorphism \( A \to B \). (For notational simplicity, from this point forward we will conflate finite abelian \( \ell \)-groups and the equivalence classes containing them.) In 1984, Cohen and Lenstra [CL84] proved that the function

\[
\nu : \mathcal{G} \to \mathbb{R}_{\geq 0}
\]

\( A \mapsto |\text{Aut } A|^{-1} \prod_{i=1}^{\infty} (1 - \ell^{-i}) \)

is a discrete probability distribution on \( \mathcal{G} \). (This fact had already been proved by Hall in [Hal38], who used a different method). They then conjectured that if \( A \in \mathcal{G} \), then \( \nu(A) \) is the probability that the \( \ell \)-Sylow subgroup of the ideal class group of an imaginary quadratic number field is isomorphic to \( A \). Since then, mathematicians have defined various probability distributions on \( \mathcal{G} \) and conjectured that these distributions describe various phenomena, both number-theoretic (e.g., [FW89], [CM90], [EVW09], [Mal10], [Gar15]) and combinatorial (e.g., [Mat14], [CKL+]).

Given any discrete probability distribution \( \xi : \mathcal{G} \to \mathbb{R}_{\geq 0} \) and any \( A \in \mathcal{G} \), define the \( A \)th moment of \( \xi \) to be

\[
\sum_{B \in \mathcal{G}} |\text{Surj}(B, A)| \xi(B),
\]

where for any \( B, A \in \mathcal{G} \), we define \( \text{Surj}(B, A) \) to be the set of surjective group homomorphisms from \( B \) to \( A \).

This terminology, which is becoming standard in the literature related to the Cohen-Lenstra heuristics (see, for example, [EVW09] and [Mat14]), is meant to evoke an analogy with the \( k \)th moment of a real-valued random variable \( X \): just as the \( k \)th moment of \( X \) is the expected value of \( X^k \), the \( A \)th moment of \( \xi \) is the expected value of \( |\text{Surj}(B, A)| \), where \( B \) is a \( \mathcal{G} \)-valued random variable. Moreover, under certain favorable conditions, the set of \( A \)th moments of a distribution on \( \mathcal{G} \) determines the distribution, making the analogy even stronger.

A precise description of these “favorable conditions”, however, is still elusive. In [EVW09], [Mat14], and [Gar15], for example, the moments of the particular discrete probability distributions on \( \mathcal{G} \) in question completely determine the distribution. In [Gar15], a Möbius inversion-type procedure transforms closed formulas for moments of certain distributions on \( \mathcal{G} \) into closed formulas for the distribution itself. Some natural questions are:

- what is this Möbius-type function?,
- in what ways does it behave like the classical Möbius function?, and
- in what conditions can it transform formulas for moments into formulas for distributions?

In this paper, we focus on the first two questions, leaving the third for later work. In Section 2, we begin by addressing the first question. That is, we define this new Möbius-type function associated to the poset \( \mathcal{G} \), which we denote \( S : \mathcal{G} \times \mathcal{G} \to \mathbb{Z} \). We also compare it to the case of the poset of subgroups of a group \( G \), which we denote \( \mathcal{P}_G \), and its associated Möbius function, which we denote \( \mu_G : \mathcal{P}_G \times \mathcal{P}_G \to \mathbb{Z} \). In particular,
we state a result relating these two functions; see Remark 2.2. We then state the main results of the paper, Theorems 3.8 and 3.9, which we prove in Section 3. As an example application of Theorems 3.8 and 3.9, we remark that they immediately imply:

**Corollary 3.10.** If $A, C \in \mathcal{G}$, then $S(A, C) = 0$ unless there exists an injection $\iota : A \to B$ with $\coker(\iota)$ elementary abelian.

We would like to note the analogy between Corollary 3.10 and Hall’s result from 1934 [Hal34]: if $G$ is an $\ell$-group of order $\ell^n$, then $\mu_G(1, G) = 0$ unless $G$ is elementary abelian, in which case $\mu_G(1, G) = (-1)^n\ell(G)$.

In addition to implying Corollary 3.10, Theorems 3.8 and 3.9 are both integral to the inversion procedure deployed in [Gar15], and will be a useful tool in answering the third question mentioned above. In [Gar], we explore further properties of $S$, using it to expand on Cohen-Lenstra’s identities on finite abelian $\ell$-groups [CL84]. Moreover, Corollary 3.10 has applications to recent work in group theory: see Lucchini’s Theorem 2.3 [Luc07], below, and the discussion following it.

## 2 Definitions and results

Let $P$ be a locally finite poset. The M"{o}bius function on $P$, denoted by $\mu_P$, is defined by the following criteria: for any $x, z \in P$,

$$
\mu_P(x, z) = 0 \quad \text{if } x \not< z,
$$

$$
\mu_P(x, z) = 1 \quad \text{if } x = z,
$$

$$
\sum_{x \leq y \leq z} \mu_P(x, y) = 0 \quad \text{if } x < z.
$$

A classic reference for M"{o}bius functions is [Rot64]. Now, for any finite group $G$, let $P_G$ be the poset of subgroups of $G$ ordered by inclusion. (To ease notation, let $\mu_G$ be the M"{o}bius function on this poset.) For a history of the work on the M"{o}bius function on this particular poset, see [HIO89]. Recall that $\mathcal{G}$ is the poset of isomorphism classes of finite abelian $\ell$-groups.

**Definition 2.1.** For any $A, C \in \mathcal{G}$, let $\sub(A, C)$ be the number of subgroups of $C$ that are isomorphic to $A$. If $A \in \mathcal{G}$, an $A$-chain is a finite linearly ordered subset of $\{B \in \mathcal{G} \mid B > A\}$. Now, given an $A$-chain $\mathcal{C} = \{A_j\}_{j=1}^i$, define

$$
\sub(\mathcal{C}) := (-1)^i \sub(A, A_1) \prod_{j=1}^{i-1} \sub(A_j, A_{j+1}).
$$

Finally, for any $A, C \in \mathcal{G}$, let

$$
S(A, C) = \begin{cases} 
0 & \text{if } A \nleq C, \\
1 & \text{if } A = C, \\
\sum_{A\text{-chains } \mathcal{C}, \max \mathcal{C} = C} \sub(\mathcal{C}) & \text{if } A < C.
\end{cases}
$$

**Remark 2.2.** Though $S$ is defined on the poset $\mathcal{G}$, it is closely related to the classical work on the M"{o}bius function on the poset of subgroups of a fixed group. Indeed, by applying Lemma 2.2 of [HIO89], we see that

$$
S(A, C) = \sum_{B \leq C} \mu_C(B, C).
$$

There has been recent progress towards describing groups with non-zero M"{o}bius functions. For example, in 2007 Lucchini [Luc07] proved the following:

**Theorem 2.3.** Assume that $G$ is a finite solvable group and that $H$ is a proper subgroup of $G$ with $\mu_G(1, H) \neq 0$. Then there exists a family $M_1, \ldots, M_t$ of maximal subgroups of $G$ such that

- $H = M_1 \cap \cdots \cap M_t$, and
In the light of Remark 2.2, Corollary 3.10 implies that there exists an infinite family of pairs of finite abelian \( \ell \)-groups with trivial Möbius function:

**Corollary 3.11.** If \( A, C \in \mathcal{G} \) and \( C \) has exactly one subgroup isomorphic to \( A \), then \( \mu_C(A, C) = 0 \) unless there exists some \( \iota : A \to C \) with \( \text{coker} \, (\iota) \) elementary abelian.

In Section 3, below, we prove the main results of this paper, mentioned in Section 1. (See Notation 3.4 for the definition of rank.)

**Theorem 3.8.** Suppose that \( A, C \in \mathcal{G} \) and \( \text{rank} \, A < \text{rank} \, C \). If there exists \( k \in \mathbb{Z}^{>0} \) and \( B \in \mathcal{G} \) such that \( A \leq B < C \), \( \text{rank} \, B = \text{rank} \, A \), and

\[
B \oplus (\mathbb{Z}/\ell) \oplus \cdots \oplus (\mathbb{Z}/\ell^k) = C,
\]

then \( S(A, C) = S(A, B) \cdot S(B, C) \). Otherwise, \( S(A, C) = 0 \).

**Theorem 3.9.** Suppose that \( A, C \in \mathcal{G} \), that \( \text{rank} \, A = \text{rank} \, C = r \), and that there does not exist an injection \( \iota : A \to C \) such that \( \text{coker} \, (\iota) \) is elementary abelian. Then \( S(A, C) = 0 \).

## 3 Proofs of main results

The combinatorics of the proofs that follow will rely on Lemmas 3.5 to 3.7, which follow immediately from Proposition 3.3 below. There are many descriptions of the quantity described in Proposition 3.3; one such can be found in Theorem 8 in a recent paper of Delaunay and Jouhet [DJ14]. The formula we present below is different than theirs; hopefully the ease with which it implies Lemmas 3.5 to 3.7 makes up for its unwieldiness. Before we begin, we introduce some notation.

**Notation 3.1.** Suppose \( A \in \mathcal{G} \). Let \( \Lambda(A) \) be the set of alternating bilinear forms on \( A \), with \( A \) thought of as a \((\mathbb{Z}/\exp(A))\)-module. Next, for any \( A, B \in \mathcal{G} \), let \( \text{Inj} \, (A, B) \) be the set of injective group homomorphisms from \( A \) into \( B \).

**Remark 3.2.** In Section 1, we defined moments in terms of surjections, which is standard, but there is an equivalent definition given in terms of injections; see Section 3 of [Gar15] for more details.

**Proposition 3.3.** Suppose \( A = \bigoplus_{i=1}^r \mathbb{Z}/\ell^{a_i} \) and \( B = \bigoplus_{i=1}^r \mathbb{Z}/\ell^{b_i} \), with \( a_i \geq a_j \) and \( b_i \geq b_j \) for \( i \leq j \). Then

\[
|\text{Inj} \, (A, B)| = |\Lambda(A)| \cdot \prod_{i=1}^r \left( \ell^{\Sigma_{j=1}^r \min\{a_i, b_j\}} - \ell^{\Sigma_{j=1}^r \min\{a_i-1, b_j\}} \right).
\]

so

\[
\text{sub} \, (A, B) = \prod_{i=1}^r \left( \frac{\ell^{\Sigma_{j=1}^r \min\{a_i, b_j\}} - \ell^{\Sigma_{j=1}^r \min\{a_i-1, b_j\}}}{\ell^{\Sigma_{j=1}^r \min\{a_i, b_j\}} - \ell^{\Sigma_{j=1}^r \min\{a_i-1, b_j\}}} \right).
\]

Before stating some consequences of Proposition 3.3, we need a bit more notation.

**Notation 3.4.** For any \( A \in \mathcal{G} \) and any \( i \in \mathbb{Z}^{>0} \), let

\[
A_{\oplus i} := A \oplus (\mathbb{Z}/\ell) \oplus \cdots \oplus (\mathbb{Z}/\ell).
\]

If \( i \geq 1 \), let

\[
\text{rank}_\ell A := \dim_{\mathbb{Z}/\ell^i} \left( \ell^{i-1} A/\ell^i A \right).
\]

We will abbreviate \( \text{rank}_\ell A \) by \( \text{rank} \, A \).

As an example, consider the group \( A = \mathbb{Z}/\ell^4 \oplus \mathbb{Z}/\ell^4 \oplus \mathbb{Z}/\ell \). Then \( \text{rank}_\ell A = 0 \), \( \text{rank}_{\ell^2} A = \text{rank}_{\ell^3} A = \text{rank}_{\ell^2} A = 2 \), and \( \text{rank} \, A = 3 \). We will use the following three lemmas in the proofs of our main results.
Lemma 3.5. Suppose \( A, B \in \mathcal{G} \). If \( \text{rank} \, B = \text{rank} \, A + i \geq 0 \), then
\[
\sub(A, A_{\theta j}) \cdot \sub(A_{\theta j}, B) = \sub(A, B).
\]

Proof. Computation following from Proposition 3.3.

Lemma 3.6. Suppose \( A, B \in \mathcal{G} \) and \( \text{rank} \, A = \text{rank} \, B \). If
\[
j \leq \max \{ i \mid \text{rank}_i A = \text{rank} A \},
\]
then
\[
\sub\left( A \oplus \mathbb{Z}/\ell^i, B \oplus \mathbb{Z}/\ell^j \right) = \sub(A, B).
\]

Proof. Computation following from Proposition 3.3.

Lemma 3.7. Suppose \( A \in \mathcal{G} \). If \( i \in \mathbb{Z}^0 \), rank \( A = r \), and \( \bigoplus_{j=1}^r \left( \mathbb{Z}/\ell^i \right) \leq A \), then
\[
\sub\left( \bigoplus_{j=1}^r \left( \mathbb{Z}/\ell^i \right), A \right) = 1.
\]

Proof. Computation following from Proposition 3.3.

We now have the tools to prove Theorems 3.8 and 3.9. For any \( A, C \in \mathcal{G} \), Theorem 3.8 concerns the case where \( \text{rank} \, A < \text{rank} \, C \), and Theorem 3.9 concerns the case where \( \text{rank} \, A = \text{rank} \, C \).

Theorem 3.8. Suppose that \( A, C \in \mathcal{G} \) and \( \text{rank} \, A < \text{rank} \, C \). If there exists \( k \in \mathbb{Z}^0 \) and \( B \in \mathcal{G} \) such that \( A \leq B < C \), rank \( B = \text{rank} \, A \), and \( B_{\theta k} = C \), then \( S(A, C) = S(A, B) \cdot S(B, C) \). Otherwise, \( S(A, C) = 0 \).

Proof. By Definition 2.1, we know \( S(A, C) \) is a sum of products of subgroup data—one summand for every \( A \)-chain with maximum \( C \). Choose some such chain, say \( \mathcal{C} = \{ A_i \}_{i=1}^j \), where \( j \in \mathbb{Z}^0 \) and \( A = A_0 < \cdots < A_j = C \).

Consider the set
\[
M_{\mathcal{C}} = \{ j_0 \in \{1, \ldots, j \} \mid \text{there is no } k_{j_0} \in \mathbb{Z}^0 \text{ such that } A_{j_0} = A_{\theta k_{j_0}} \}.
\]

If \( M_{\mathcal{C}} \) is empty, then the theorem is trivially true since there is some \( k \in \mathbb{Z}^0 \) such that \( C = A_{\theta k} \). Thus, suppose it is not empty and let \( j' = \min(M_{\mathcal{C}}) \).

There are two possibilities for the ranks of \( A_{j'} \) and \( A_{j'-1} \): either \( \text{rank}(A_{j'-1}) = \text{rank}(A_{j'}) \) or \( \text{rank}(A_{j'-1}) < \text{rank}(A_{j'}) \). It turns out that summands in the former case cancel out those in the latter. Indeed, if
\[
\text{rank}(A_{j'}) = \text{rank}(A_{j'-1}) = k_0 > 0,
\]
then we know by Lemma 3.5 that
\[
\sub\left( (A_{j'-1})_{\theta k_0}, A_{j'} \right) \cdot \sub\left( (A_{j'-1})_{\theta k_0}, A_{j'} \right) = \sub(A_{j'-1}, A_{j'}).
\]

Thus, \( \sub(\mathcal{C}) \) cancels with another summand in \( S(A, B) \), one associated to a chain that is longer than \( \mathcal{C} \) by one subgroup; namely, the chain
\[
A_1 < \cdots < A_{j'-1} < (A_{j'-1})_{\theta k_0} < A_{j'} < \cdots < A_j = B.
\]

In contrast to \( \mathcal{C} \), the first subgroup in \( \mathcal{C}' \) that is not of the form \( A_{\theta k} \) for any \( k \in \mathbb{Z}^0 \) has the same rank as its predecessor (ie, \( \text{rank}(A_{j'}) = \text{rank}(A_{j'}) \)).

Now suppose that \( A_{j'} \) and \( A_{j'-1} \) had satisfied the other possibility; ie, that \( \text{rank}(A_{j'-1}) = \text{rank}(A_{j'}) \). If \( j' > 1 \), then the summand cancels with a summand whose chain is one shorter. Specifically, we know by Lemma 3.5 that it cancels with the summand associated to the chain \( \mathcal{C} \times \{ A_{j'-1} \} \). Thus, the only summands of \( S(A, B) \) that remain are those associated to chains with minimum element the same rank as \( A \). Using this fact, we can write
\[
S(A, C) = - \sum_{B_0 \in \mathcal{G}, \text{rank} \, B_0 = \text{rank} \, C} \sub(A, B_0) \cdot S(B_0, C).
\]
Note that if \( \{ B_0 \in G \mid A < B_0 < C, \text{rank} B_0 = \text{rank} A \} = \emptyset \), then the above sum vanishes and we are done. Thus, suppose it is not empty and let
\[
B = \max \{ B_0 \in G \mid A < B_0 < C, \text{rank} B_0 = \text{rank} A \}.
\]
We can repeat the argument above to see that
\[
S(A, C) = S(A, B) \cdot S(B, C).
\]
If there is some \( k \in \mathbb{Z}^{>0} \) such that \( C = B_{\#k} \), then we are done. If not, then the argument from the previous paragraphs and the definition of \( B \) imply that \( S(B, C) = 0 \), completing the proof.

In the light of Theorem 3.8, we now address \( S(A, C) \) in the case where \( \text{rank} A = \text{rank} C \).

**Theorem 3.9.** Suppose that \( A, C \in G \), that \( \text{rank} A = \text{rank} C = r \), and that there does not exist an injection \( \iota : A \to C \) such that \( \text{coker} \iota \) is elementary abelian. Then \( S(A, C) = 0 \).

**Proof.** Suppose that \( A < C \) (otherwise the result is trivial). We will induct on \( r \). To begin, suppose that \( r = 1 \), and define \( a, c \in \mathbb{Z}^{>0} \) by \( \ell^a = |A| \) and \( \ell^c = |C| \). Since \( A \) and \( C \) are cyclic, we see that for any \( i \in \{ 1, \ldots, c - a \} \),
\[
\{[A\text{-chains } \mathcal{C}] \mid \max \mathcal{C} = C, |\mathcal{C}| = i\} = \binom{c - a - 1}{i - 1}.
\]
Moreover, the fact that \( A \) and \( C \) are cyclic also implies that \( \text{sub} \mathcal{C} = (-1)^i \) for any \( A \)-chain in the above set. By assumption, we know that \( c - a - 1 > 0 \), so \( \sum_{i=1}^{c-a-1} (-1)^i \binom{c-a-1}{i-1} = 0 \). This completes the base case.

We split the general case into three cases. For the first case, suppose that \( \exp A = \exp C \). For any \( B \in G \), let \( \overline{B} \) denote \( B \setminus \{b\} \), where \( b \in B \) is any element of order \( \exp B \). Similarly, if \( \mathcal{C} \) is a \( B \)-chain, we define \( \overline{\mathcal{C}} \) to be \( \{ \overline{D} \mid D \in \mathcal{C} \} \). Now, since \( \exp A = \exp C \), we see that \( \{A\text{-chains } \mathcal{C} \mid \max \mathcal{C} = C\} \) is in bijection with \( \{A\text{-chains } \mathcal{C} \mid \max \mathcal{C} = C\} \) under the map \( \mathcal{C} \mapsto \overline{\mathcal{C}} \). Moreover, given any \( A \)-chain \( \mathcal{C} \) with \( \max \mathcal{C} = C \), Proposition 3.3 implies that \( \text{sub} \mathcal{C} = K \cdot \text{sub} \overline{\mathcal{C}} \), where \( K \) is a constant depending only on \( A \) and \( C \). The result now follows by induction.

For the second case, suppose that \( \ell \cdot \exp A = \exp C \). For any \( A \)-chain \( C \) with \( \max \mathcal{C} = C \), let \( \overline{\mathcal{C}} \) denote \( \min \{ B \in \mathcal{C} \mid \exp B = \exp C \} \). For any \( B \in G \) such that \( A < B \leq C \) and \( \exp B > \exp A \), let \( \mathcal{C}_B = B \setminus (\ell^{-1} \exp C) \cdot B \). For any such \( B \) with \( B_C \neq A \), we can partition the set \( \{A\text{-chains } \mathcal{C} \mid \max \mathcal{C} = C, \overline{\mathcal{C}} = B\} \) into two subsets: those chains that contain \( B_C \) and those that do not. We remark that these two subsets are in bijection under the following map: if an \( A \)-chain does not contain \( B_C \), then add it. The inverse to this map is simply the deletion of \( B_C \) from any \( A \)-chain. Now, by Lemmas 3.6 and 3.7, we know that \( B \) has exactly one subgroup isomorphic to \( B_C \). Thus, for any \( A_0 \) such that \( A \leq A_0 < B_C \), we know that
\[
\text{sub} (A_0, B_C) \cdot \text{sub} (B_C, B) = \text{sub} (A_0, B).
\]
But this means that any summand associated to a chain in the first subset cancels with the summand associated to the image of the chain under the above bijection. Thus,
\[
S(A, C) = \sum_{A\text{-chains } \mathcal{C} \mid (\overline{\mathcal{C}})_C = A} \text{sub} \mathcal{C}.
\]
Now, for any \( B \in G \) with \( A < B \leq C \), \( \exp B > \exp A \), and \( B_C = A \), note that
\[
\sum_{A\text{-chains } \mathcal{C} \mid \max \mathcal{C} = B, \overline{\mathcal{C}} \leq B} \text{sub} \mathcal{C} = S(B, C) \cdot \sum_{A\text{-chains } \mathcal{C} \mid \max \mathcal{C} = B, \overline{\mathcal{C}} \leq B} \text{sub} \mathcal{C}.
\]
But \( S(B, C) = 0 \) for all such \( B \), by the argument of the previous paragraph. Thus,
\[
S(A, C) = \sum_{A\text{-chains } \mathcal{C} \mid (\overline{\mathcal{C}})_C = A} \text{sub} \mathcal{C} = \sum_{A \leq B \leq C} \text{sub} \mathcal{C} (B_C = A) \sum_{\mathcal{C} \in \mathcal{C}_B} \text{sub} \mathcal{C} = 0,
\]
completing this case.

Finally, consider the case where \( \ell \cdot \exp A < \exp C \). As in the previous case, we have that

\[
S(A, C) = \sum_{\text{A-chains } \mathcal{C}} \text{sub}(\mathcal{C}).
\]

The difference in this case is that if \( \mathcal{C} \) is an \( A \)-chain with \( \max \mathcal{C} = C \), then it is impossible that \( (\mathcal{C})_C = A \), so the proof is complete.

Theorems 3.8 and 3.9 immediately imply the following corollary.

**Corollary 3.10.** If \( A, C \in \mathcal{G} \), then \( S(A, C) = 0 \) unless there exists an injection \( \iota : A \rightarrow B \) with \( \text{coker}(\iota) \) elementary abelian.

Finally, Remark 2.2 and Corollary 3.10 imply the following result.

**Corollary 3.11.** If \( A, C \in \mathcal{G} \) and \( C \) has exactly one subgroup isomorphic to \( A \), then \( \mu_C(A, C) = 0 \) unless there exists some \( \iota : A \rightarrow C \) with \( \text{coker}(\iota) \) elementary abelian.

**References**


