LECTURE 1

Basic definitions, the intersection poset and the characteristic polynomial

1.1. Basic definitions

The following notation is used throughout for certain sets of numbers:

\[
\begin{align*}
\mathbb{N} & \text{ nonnegative integers} \\
\mathbb{P} & \text{ positive integers} \\
\mathbb{Z} & \text{ integers} \\
\mathbb{Q} & \text{ rational numbers} \\
\mathbb{R} & \text{ real numbers} \\
\mathbb{R}^+ & \text{ positive real numbers} \\
\mathbb{C} & \text{ complex numbers} \\
[m] & \text{ the set } \{1, 2, \ldots, m\} \text{ when } m \in \mathbb{N}
\end{align*}
\]

We also write \([t^k] \chi(t)\) for the coefficient of \(t^k\) in the polynomial or power series \(\chi(t)\). For instance, \([t^2](1 + t)^4 = 6\).

A \textit{finite hyperplane arrangement} \(\mathcal{A}\) is a finite set of affine hyperplanes in some vector space \(V \cong K^n\), where \(K\) is a field. We will not consider infinite hyperplane arrangements or arrangements of general subspaces or other objects (though they have many interesting properties), so we will simply use the term \textit{arrangement} for a finite hyperplane arrangement. Most often we will take \(K = \mathbb{R}\), but as we will see even if we’re only interested in this case it is useful to consider other fields as well. To make sure that the definition of a hyperplane arrangement is clear, we define a \textit{linear hyperplane} to be an \((n - 1)\)-dimensional subspace \(H\) of \(V\), i.e.,

\[
H = \{v \in V : \alpha \cdot v = 0\},
\]

where \(\alpha\) is a fixed nonzero vector in \(V\) and \(\alpha \cdot v\) is the usual dot product:

\[
(\alpha_1, \ldots, \alpha_n) \cdot (v_1, \ldots, v_n) = \sum \alpha_i v_i.
\]

An \textit{affine hyperplane} is a translate \(J\) of a linear hyperplane, i.e.,

\[
J = \{v \in V : \alpha \cdot v = a\},
\]

where \(\alpha\) is a fixed nonzero vector in \(V\) and \(a \in K\).

If the equations of the hyperplanes of \(\mathcal{A}\) are given by \(L_1(x) = a_1, \ldots, L_m(x) = a_m\), where \(x = (x_1, \ldots, x_n)\) and each \(L_i(x)\) is a homogeneous linear form, then we call the polynomial

\[
Q_\mathcal{A}(x) = (L_1(x) - a_1) \cdots (L_m(x) - a_m)
\]

the \textit{defining polynomial} of \(\mathcal{A}\). It is often convenient to specify an arrangement by its defining polynomial. For instance, the arrangement \(\mathcal{A}\) consisting of the \(n\) coordinate hyperplanes has \(Q_\mathcal{A}(x) = x_1 x_2 \cdots x_n\).

Let \(\mathcal{A}\) be an arrangement in the vector space \(V\). The \textit{dimension} \(\dim(\mathcal{A})\) of \(\mathcal{A}\) is defined to be \(\dim(V) = n\), while the \textit{rank} \(\text{rank}(\mathcal{A})\) of \(\mathcal{A}\) is the dimension of the space spanned by the normals to the hyperplanes in \(\mathcal{A}\). We say that \(\mathcal{A}\) is \textit{essential} if \(\text{rank}(\mathcal{A}) = \dim(\mathcal{A})\). Suppose that \(\text{rank}(\mathcal{A}) = r\), and take \(V = K^n\). Let
Y be a complementary space in $K^n$ to the subspace $X$ spanned by the normals to hyperplanes in $A$. Define

$$W = \{v \in V : v \cdot y = 0 \ \forall y \in Y\}.$$ 

If $\text{char}(K) = 0$ then we can simply take $W = X$. By elementary linear algebra we have

$$\text{codim}_W(H \cap W) = 1$$

for all $H \in A$. In other words, $H \cap W$ is a hyperplane of $W$, so the set $A_W := \{H \cap W : H \in A\}$ is an essential arrangement in $W$. Moreover, the arrangements $A$ and $A_W$ are “essentially the same,” meaning in particular that they have the same intersection poset (as defined in Definition 1.1). Let us call $A_W$ the essentialization of $A$, denoted $\text{ess}(A)$. When $K = \mathbb{R}$ and we take $W = X$, then the arrangement $A$ is obtained from $A_W$ by “stretching” the hyperplane $H \cap W \in A_W$ orthogonally to $W$. Thus if $W^\perp$ denotes the orthogonal complement to $W$ in $V$, then $H' \in A_W$ if and only if $H' \oplus W^\perp \in A$. Note that in characteristic $p$ this type of reasoning fails since the orthogonal complement of a subspace $W$ can intersect $W$ in a subspace of dimension greater than 0.

**Example 1.1.** Let $A$ consist of the lines $x = a_1, \ldots, x = a_k$ in $K^2$ (with coordinates $x$ and $y$). Then we can take $W$ to be the $x$-axis, and $\text{ess}(A)$ consists of the points $x = a_1, \ldots, x = a_k$ in $K$.

Now let $K = \mathbb{R}$. A region of an arrangement $A$ is a connected component of the complement $X$ of the hyperplanes:

$$X = \mathbb{R}^n - \bigcup_{H \in A} H.$$ 

Let $\mathcal{R}(A)$ denote the set of regions of $A$, and let

$$r(A) = \#\mathcal{R}(A),$$

the number of regions. For instance, the arrangement $A$ shown below has $r(A) = 14$.

It is a simple exercise to show that every region $R \in \mathcal{R}(A)$ is open and convex (continuing to assume $K = \mathbb{R}$), and hence homeomorphic to the interior of an $n$-dimensional ball $\mathbb{B}^n$ (Exercise 1). Note that if $W$ is the subspace of $V$ spanned by the normals to the hyperplanes in $A$, then $R \in \mathcal{R}(A)$ if and only if $R \cap W \in \mathcal{R}(A_W)$. We say that a region $R \in \mathcal{R}(A)$ is relatively bounded if $R \cap W$ is bounded. If $A$ is essential, then relatively bounded is the same as bounded. We write $b(A)$ for
the number of relatively bounded regions of \( A \). For instance, in Example 1.1 take 
\( K = \mathbb{R} \) and \( a_1 < a_2 < \cdots < a_k \). Then the relatively bounded regions are the 
regions \( a_i < x < a_{i+1}, 1 \leq i \leq k - 1 \). In \( \varepsilon s( A) \) they become the (bounded) open 
intervals \( (a_i, a_{i+1}) \). There are also two regions of \( A \) that are not relatively bounded, 
viz., \( x < a_1 \) and \( x > a_k \).

A (closed) half-space is a set \( \{ x \in \mathbb{R}^n : x \cdot \alpha \geq c \} \) for some \( \alpha \in \mathbb{R}^n, c \in \mathbb{R} \). If 
\( H \) is a hyperplane in \( \mathbb{R}^n \), then the complement \( \mathbb{R}^n - H \) has two (open) components 
whose closures are half-spaces. It follows that the closure \( \bar{R} \) of a region \( R \) of \( A \) is 
a finite intersection of half-spaces, i.e., a (convex) polyhedron (of dimension \( n \)). A 
bounded polyhedron is called a (convex) polytope. Thus if \( R \) (or \( \bar{R} \)) is bounded, 
then \( \bar{R} \) is a polytope (of dimension \( n \)).

An arrangement \( A \) is in general position if 
\[
\{ H_1, \ldots, H_p \} \subseteq A, \; p \leq n \quad \Rightarrow \quad \dim(H_1 \cap \cdots \cap H_p) = n - p \\
\{ H_1, \ldots, H_p \} \subseteq A, \; p > n \quad \Rightarrow \quad H_1 \cap \cdots \cap H_p = \emptyset.
\]

For instance, if \( n = 2 \) then a set of lines is in general position if no two are parallel 
and no three meet at a point.

Let us consider some interesting examples of arrangements that will anticipate 
some later material.

**Example 1.2.** Let \( A_m \) consist of \( m \) lines in general position in \( \mathbb{R}^2 \). We can compute 
\( r(A_m) \) using the sweep hyperplane method. Add a \( L \) line to \( A_k \) (with \( A_K \cup \{ L \} \) in 
general position). When we travel along \( L \) from one end (at infinity) to the other, 
every time we intersect a line in \( A_k \) we create a new region, and we create one new 
region at the end. Before we add any lines we have one region (all of \( \mathbb{R}^2 \)). Hence 
\[
r(A_m) = \#\text{intersections} + \#\text{lines} + 1 \\
= \binom{m}{2} + m + 1.
\]

**Example 1.3.** The braid arrangement \( B_n \) in \( K^n \) consists of the hyperplanes 
\( B_n : x_i - x_j = 0, \; 1 \leq i < j \leq n. \) 

Thus \( B_n \) has \( \binom{n}{2} \) hyperplanes. To count the number of regions when \( K = \mathbb{R} \), note 
that specifying which side of the hyperplane \( x_i - x_j = 0 \) a point \( (a_1, \ldots, a_n) \) lies 
on is equivalent to specifying whether \( a_i < a_j \) or \( a_i > a_j \). Hence the number of 
regions is the number of ways that we can specify whether \( a_i < a_j \) or \( a_i > a_j \) for 
\( 1 \leq i < j \leq n \). Such a specification is given by imposing a linear order on the 
a_i’s. In other words, for each permutation \( w \in S_n \) (the symmetric group of all 
permutations of \( 1, 2, \ldots, n \)), there corresponds a region \( R_w \) of \( B_n \) given by 
\[
R_w = \{ (a_1, \ldots, a_n) \in \mathbb{R}^n : a_{w(1)} > a_{w(2)} > \cdots > a_{w(n)} \}.
\]

Hence \( r(B_n) = n! \). Rarely is it so easy to compute the number of regions!

Note that the braid arrangement \( B_n \) is not essential; indeed, \( \text{rank}(B_n) = n - 1 \). 
When \( \text{char}(K) \neq 2 \) the space \( W \subseteq K^n \) of equation (1) can be taken to be 
\[
W = \{ (a_1, \ldots, a_n) \in K^n : a_1 + \cdots + a_n = 0 \}.
\]

The braid arrangement has a number of “deformations” of considerable interest.

We will just define some of them now and discuss them further later. All these 
arrangements lie in \( K^n \), and in all of them we take \( 1 \leq i < j \leq n \). The reader who
likes a challenge can try to compute their number of regions when \( K = \mathbb{R} \). (Some are much easier than others.)

- **generic braid arrangement**: \( x_i - x_j = a_{ij} \), where the \( a_{ij} \)'s are “generic” (e.g., linearly independent over the prime field, so \( K \) has to be “sufficiently large”). The precise definition of “generic” will be given later. (The prime field of \( K \) is its smallest subfield, isomorphic to either \( \mathbb{Q} \) or \( \mathbb{Z}/p\mathbb{Z} \) for some prime \( p \).)
- **semigeneric braid arrangement**: \( x_i - x_j = a_i \), where the \( a_i \)'s are “generic.”
- **Shi arrangement**: \( x_i - x_j = 0,1 \) (so \( n(n-1) \) hyperplanes in all).
- **Linial arrangement**: \( x_i - x_j = 1 \).
- **Catalan arrangement**: \( x_i - x_j = -1,0,1 \).
- **semiorder arrangement**: \( x_i - x_j = -1,1 \).
- **threshold arrangement**: \( x_i + x_j = 0 \) (not really a deformation of the braid arrangement, but closely related).

An arrangement \( A \) is **central** if \( \bigcap_{H \in A} H \neq \emptyset \). Equivalently, \( A \) is a translate of a **linear arrangement** (an arrangement of linear hyperplanes, i.e., hyperplanes passing through the origin). Many other writers call an arrangement central, rather than linear, if \( 0 \in \bigcap_{H \in A} H \). If \( A \) is central with \( X = \bigcap_{H \in A} H \), then \( \text{rank}(A) = \text{codim}(X) \). If \( A \) is central, then note also that \( b(A) = 0 \) (why?).

There are two useful arrangements closely related to a given arrangement \( A \). If \( A \) is a linear arrangement in \( K^n \), then **projectivize** \( A \) by choosing some \( H \in A \) to be the hyperplane at infinity in projective space \( P_K^{n-1} \). Thus if we regard

\[
P_K^{n-1} = \{(x_1, \ldots, x_n) : x_i \in K, \text{not all } x_i = 0\}/\sim,
\]

where \( u \sim v \) if \( u = \alpha v \) for some \( 0 \neq \alpha \in K \), then

\[
H = \{(x_1, \ldots, x_{n-1}, 0) : x_i \in K, \text{not all } x_i = 0\}/\sim \cong P_K^{n-2}.
\]

The remaining hyperplanes in \( A \) then correspond to “finite” (i.e., not at infinity) projective hyperplanes in \( P_K^{n-1} \). This gives an arrangement \( \text{proj}(A) \) of hyperplanes in \( P_K^{n-1} \). When \( K = \mathbb{R} \), the two regions \( R \) and \( -R \) of \( A \) become identified in \( \text{proj}(A) \). Hence \( r(\text{proj}(A)) = \frac{1}{2} r(A) \). When \( n = 3 \), we can draw \( P_\mathbb{R}^2 \) as a disk with antipodal boundary points identified. The circumference of the disk represents the hyperplane at infinity. This provides a good way to visualize three-dimensional real linear arrangements. For instance, if \( A \) consists of the three coordinate hyperplanes \( x_1 = 0, x_2 = 0, \) and \( x_3 = 0 \), then a projective drawing is given by

![Diagram](image)

The line labelled \( i \) is the projectivization of the hyperplane \( x_i = 0 \). The hyperplane at infinity is \( x_3 = 0 \). There are four regions, so \( r(A) = 8 \). To draw the incidences among all eight regions of \( A \), simply “reflect” the interior of the disk to the exterior:
Regarding this diagram as a planar graph, the dual graph is the 3-cube (i.e., the vertices and edges of a three-dimensional cube) [why?].

For a more complicated example of projectivization, Figure 1 shows proj($\mathcal{B}_4$) (where we regard $\mathcal{B}_4$ as a three-dimensional arrangement contained in the hyperplane $x_1 + x_2 + x_3 + x_4 = 0$ of $\mathbb{R}^4$), with the hyperplane $x_i = x_j$ labelled $ij$, and with $x_1 = x_4$ as the hyperplane at infinity.
We now define an operation which is “inverse” to projectivization. Let \( \mathcal{A} \) be an (affine) arrangement in \( K^n \), given by the equations
\[
L_1(x) = a_1, \ldots, L_m(x) = a_m.
\]
Introduce a new coordinate \( y \), and define a central arrangement \( c\mathcal{A} \) (the cone over \( \mathcal{A} \)) in \( K^n \times K = K^{n+1} \) by the equations
\[
L_1(x) = a_1y, \ldots, L_m(x) = a_my, \ y = 0.
\]
For instance, let \( \mathcal{A} \) be the arrangement in \( \mathbb{R}^1 \) given by \( x = -1, x = 2, \) and \( x = 3 \). The following figure should explain why \( c\mathcal{A} \) is called a cone.

\[\text{Diagram}\]

It is easy to see that when \( K = \mathbb{R} \), we have \( r(c\mathcal{A}) = 2r(\mathcal{A}) \). In general, \( c\mathcal{A} \) has the “same combinatorics as \( \mathcal{A} \), times 2.” See Exercise 1.

1.2. The intersection poset

Recall that a poset (short for partially ordered set) is a set \( P \) and a relation \( \leq \) satisfying the following axioms (for all \( x, y, z \in P \)):

(P1) (reflexivity) \( x \leq x \)
(P2) (antisymmetry) If \( x \leq y \) and \( y \leq x \), then \( x = y \).
(P3) (transitivity) If \( x \leq y \) and \( y \leq z \), then \( x \leq z \).

Obvious notation such as \( x < y \) for \( x \leq y \) and \( x \neq y \), and \( y \geq x \) for \( x \leq y \) will be used throughout. If \( x \leq y \) in \( P \), then the (closed) interval \([x, y]\) is defined by
\[
[x, y] = \{ z \in P : x \leq z \leq y \}.
\]

Note that the empty set \( \emptyset \) is not a closed interval. For basic information on posets not covered here, see [18].

**Definition 1.1.** Let \( \mathcal{A} \) be an arrangement in \( V \), and let \( L(\mathcal{A}) \) be the set of all nonempty intersections of hyperplanes in \( \mathcal{A} \), including \( V \) itself as the intersection over the empty set. Define \( x \leq y \) in \( L(\mathcal{A}) \) if \( x \supseteq y \) (as subsets of \( V \)). In other words, \( L(\mathcal{A}) \) is partially ordered by reverse inclusion. We call \( L(\mathcal{A}) \) the intersection poset of \( \mathcal{A} \).

**NOTE.** The primary reason for ordering intersections by reverse inclusion rather than ordinary inclusion is Proposition 3.8. We don’t want to alter the well-established definition of a geometric lattice or to refer constantly to “dual geometric lattices.”

The element \( V \in L(\mathcal{A}) \) satisfies \( x \geq V \) for all \( x \in L(\mathcal{A}) \). In general, if \( P \) is a poset then we denote by \( \hat{0} \) an element (necessarily unique) such that \( x \geq \hat{0} \) for all
$x \in P$. We say that $y$ covers $x$ in a poset $P$, denoted $x \lessdot y$, if $x < y$ and no $z \in P$ satisfies $x < z < y$. Every finite poset is determined by its cover relations. The (Hasse) diagram of a finite poset is obtained by drawing the elements of $P$ as dots, with $x$ drawn lower than $y$ if $x < y$, and with an edge between $x$ and $y$ if $x \lessdot y$. Figure 2 illustrates four arrangements $A$ in $\mathbb{R}^2$, with (the diagram of) $L(A)$ drawn below $A$.

A chain of length $k$ in a poset $P$ is a set $x_0 < x_1 < \cdots < x_k$ of elements of $P$. The chain is saturated if $x_0 < x_1 < \cdots < x_k$. We say that $P$ is graded of rank $n$ if every maximal chain of $P$ has length $n$. In this case $P$ has a rank function $\text{rk} : P \to \mathbb{N}$ defined by:

- $\text{rk}(x) = 0$ if $x$ is a minimal element of $P$.
- $\text{rk}(y) = \text{rk}(x) + 1$ if $x \lessdot y$ in $P$.

If $x < y$ in a graded poset $P$ then we write $\text{rk}(x, y) = \text{rk}(y) - \text{rk}(x)$, the length of the interval $[x, y]$. Note that we use the notation $\text{rank}(A)$ for the rank of an arrangement $A$ but $\text{rk}$ for the rank function of a graded poset.

**Proposition 1.1.** Let $A$ be an arrangement in a vector space $V \cong K^n$. Then the intersection poset $L(A)$ is graded of rank equal to $\text{rank}(A)$. The rank function of $L(A)$ is given by

$$\text{rk}(x) = \text{codim}(x) = n - \dim(x),$$

where $\dim(x)$ is the dimension of $x$ as an affine subspace of $V$.

**Proof.** Since $L(A)$ has a unique minimal element $\emptyset = V$, it suffices to show that (a) if $x \lessdot y$ in $L(A)$ then $\dim(x) - \dim(y) = 1$, and (b) all maximal elements of $L(A)$ have dimension $n - \text{rank}(A)$. By linear algebra, if $H$ is a hyperplane and $x$ an affine subspace, then $H \cap x = x$ or $\dim(x) - \dim(H \cap x) = 1$, so (a) follows. Now suppose that $x$ has the largest codimension of any element of $L(A)$, say $\text{codim}(x) = d$. Thus $x$ is an intersection of $d$ linearly independent hyperplanes (i.e., their normals are linearly independent) $H_1, \ldots, H_d$ in $A$. Let $y \in L(A)$ with $c = \text{codim}(y) < d$. Thus $y$ is an intersection of $c$ hyperplanes, so some $H_i$ ($1 \leq i \leq d$) is linearly independent from them. Then $y \cap H_i \neq \emptyset$ and $\text{codim}(y \cap H_i) > \text{codim}(y)$. Hence $y$ is not a maximal element of $L(A)$, proving (b). \qed
1.3. The characteristic polynomial

A poset $P$ is \textit{locally finite} if every interval $[x, y]$ is finite. Let $\text{Int}(P)$ denote the set of all closed intervals of $P$. For a function $f : \text{Int}(P) \rightarrow \mathbb{Z}$, write $f(x, y)$ for $f([x, y])$. We now come to a fundamental invariant of locally finite posets.

\textbf{Definition 1.2.} Let $P$ be a locally finite poset. Define a function $\mu = \mu_P : \text{Int}(P) \rightarrow \mathbb{Z}$, called the \textit{M"obius function} of $P$, by the conditions:

\begin{align*}
\mu(x, x) &= 1, \text{ for all } x \in P \\
\mu(x, y) &= -\sum_{x \leq z < y} \mu(x, z), \text{ for all } x < y \text{ in } P.
\end{align*}

This second condition can also be written

$$\sum_{x \leq z \leq y} \mu(x, z) = 0, \text{ for all } x < y \text{ in } P.$$ 

If $P$ has a $\hat{0}$, then we write $\mu(x) = \mu(\hat{0}, x)$. Figure 3 shows the intersection poset $L$ of the arrangement $\mathcal{A}$ in $K^3$ (for any field $K$) defined by $Q_{\mathcal{A}}(x) = xyz(x + y)$, together with the value $\mu(x)$ for all $x \in L$.

A important application of the M"obius function is the \textit{M"obius inversion formula}. The best way to understand this result (though it does have a simple direct proof) requires the machinery of incidence algebras. Let $\mathcal{I}(P) = \mathcal{I}(P, K)$ denote the vector space of all functions $f : \text{Int}(P) \rightarrow K$. Write $f(x, y)$ for $f([x, y])$. For $f, g \in \mathcal{I}(P)$, define the product $fg \in \mathcal{I}(P)$ by

$$fg(x, y) = \sum_{x \leq z \leq y} f(x, z)g(z, y).$$

It is easy to see that this product makes $\mathcal{I}(P)$ an associative $\mathbb{Q}$-algebra, with multiplicative identity $\delta$ given by

$$\delta(x, y) = \begin{cases} 
1, & x = y \\
0, & x < y.
\end{cases}$$

Define the \textit{zeta function} $\zeta \in \mathcal{I}(P)$ of $P$ by $\zeta(x, y) = 1$ for all $x \leq y$ in $P$. Note that the M"obius function $\mu$ is an element of $\mathcal{I}(P)$. The definition of $\mu$ (Definition 1.2) is
equivalent to the relation $\mu \zeta = \delta$ in $\mathcal{I}(P)$. In any finite-dimensional algebra over a field, one-sided inverses are two-sided inverses, so $\mu = \zeta^{-1}$ in $\mathcal{I}(P)$.

**Theorem 1.1.** Let $P$ be a finite poset with Möbius function $\mu$, and let $f, g : P \to K$. Then the following two conditions are equivalent:

\[
    f(x) = \sum_{y \geq x} g(y), \text{ for all } x \in P
\]

\[
    g(x) = \sum_{y \geq x} \mu(x, y)f(y), \text{ for all } x \in P.
\]

**Proof.** The set $K^P$ of all functions $P \to K$ forms a vector space on which $\mathcal{I}(P)$ acts (on the left) as an algebra of linear transformations by

\[
    (\xi f)(x) = \sum_{y \geq x} \xi(x, y)f(y),
\]

where $f \in K^P$ and $\xi \in \mathcal{I}(P)$. The Möbius inversion formula is then nothing but the statement

\[
    \zeta f = g \iff f = \mu g.
\]

We now come to the main concept of this section.

**Definition 1.3.** The **characteristic polynomial** $\chi_A(t)$ of the arrangement $A$ is defined by

\[
    \chi_A(t) = \sum_{x \in L(A)} \mu(x) t^{\dim(x)}.
\]

For instance, if $A$ is the arrangement of Figure 3, then

\[
    \chi_A(t) = t^3 - 4t^2 + 5t - 2 = (t-1)^2(t-2).
\]

Note that we have immediately from the definition of $\chi_A(t)$, where $A$ is in $K^n$, that

\[
    \chi_A(t) = t^n - (\#A) t^{n-1} + \cdots.
\]

**Example 1.4.** Consider the coordinate hyperplane arrangement $A$ with defining polynomial $Q_A(x) = x_1x_2 \cdots x_n$. Every subset of the hyperplanes in $A$ has a different nonempty intersection, so $L(A)$ is isomorphic to the boolean algebra $B_n$ of all subsets of $[n] = \{1,2,\ldots,n\}$, ordered by inclusion.

**Proposition 1.2.** Let $A$ be given by the above example. Then $\chi_A(t) = (t - 1)^n$.

**Proof.** The computation of the Möbius function of a boolean algebra is a standard result in enumerative combinatorics with many proofs. We will give here a naive proof from first principles. Let $y \in L(A)$, $r(y) = k$. We claim that

\[
    \mu(y) = (-1)^k.
\]

The assertion is clearly true for $r(k) = 0$, when $y = 0$. Now let $y > 0$. We need to show that

\[
    \sum_{x \leq y} (-1)^{rk(x)} = 0.
\]
The number of $x$ such that $x \leq y$ and $rk(x) = i$ is $\binom{k}{i}$, so (5) is equivalent to the well-known identity (easily proved by substituting $q = -1$ in the binomial expansion of $(q + 1)^k \sum_{i=0}^{k} (-1)^i \binom{k}{i} = 0$ for $k > 0$. \qed