1. Topic 1: Hilbert functions of 0-dimensional subschemes of \( \mathbb{P}^2 \)

Speaker: Brian Harbourne
Venue: Krakow, Poland, May, 2015
Date of Draft: May 21, 2015

1.1. Algebraic Background. We work over an algebraically closed field \( K \). Let \( R = K[x,y,z] \) and let \( M \) be the ideal \((x,y,z)\). For each integer \( t \), let \( R_t \) be the \( K \)-vector space span of all homogeneous elements of \( R \) of degree \( t \). Note that \( R = \bigoplus_t R_t \).

**Definition 1.1.1.** We say an ideal \( I \subseteq R \) is homogeneous if \( I \) has a homogeneous set of generators, or equivalently if \( I = \bigoplus_t I_t \), where \( I_t = I \cap R_t \). Note that each \( I_t \) is a \( K \)-vector space. We say a homogeneous ideal \( I \) is saturated if whenever \( FM^n \subseteq I \) for some \( n \) then \( F \in I \).

There is a bijective correspondence between 0-dimensional subschemes \( Z \subseteq \mathbb{P}^2 \) and ideals of the form \( Q_1 \cap \ldots \cap Q_r \), where \( p_1, \ldots, p_r \) are distinct points of \( \mathbb{P}^2 \) such that each \( Q_i \) is homogeneous and saturated with \( \sqrt{Q_i} = I(p_i) \) being the ideal generated by all forms vanishing at \( p_i \). We write \( I_Z \) for the ideal \( Q_1 \cap \ldots \cap Q_r \) corresponding to \( Z \). We call the points \( p_i \) the points of support of \( Z \).

**Definition 1.1.2.** Given a homogeneous ideal \( I \), we define the Hilbert function \( h_I \) of \( I \) as \( h_I(t) = \dim_K(I_t) \). If \( I = I_Z \) for a 0-dimensional subscheme \( Z \subseteq \mathbb{P}^2 \), we define the Hilbert function \( h_Z \) of \( Z \) to be \( h_Z(t) = h_R(t) - h_{I_Z}(t) = \binom{t+2}{2} - h_{I_Z}(t) \). Note that \( R/I_Z \cong \bigoplus_R/I_{I_Z} \), and \( h_Z(t) = \dim_K R_t/(I_t) \).

Given any function \( f : Z \to \mathbb{Z} \), we write \( \Delta f \) for the first difference function, so

\[
\Delta f(t) = f(t) - f(t-1).
\]

Recursively we then have the \( n \)th difference function

\[
\Delta^n f(t) = \Delta^{n-1} f(t) - \Delta^{n-1} f(t-1).
\]

**Exercise 1.1.3.** Let \( p \in \mathbb{P}^2 \) and let \( Q \subseteq R \) be a homogeneous ideal. Show that \( Q \) is \( I(p) \)-primary if and only if \( Q \) is saturated with \( \sqrt{Q} = I(p) \). Show by example that \( \sqrt{Q} = I(p) \) does not imply that \( Q \) is \( I(p) \)-primary.

**Exercise 1.1.4.** Let \( S \subseteq \mathbb{P}^2 \) be any finite set of points. Let \( Z \) be a 0-dimensional subscheme with support \( S \). If \( F \in R_t \) but \( F \) does not vanish at any point of \( S \), show that \( F \) is a non-zero divisor on \( R/I_Z \). Show that there is a linear form \( L \) that does not vanish at any point of \( S \); conclude that \( L \) is a non-zero divisor on \( R/I_Z \).

**Definition 1.1.5.** Given a nonzero homogeneous ideal \( I \subseteq R \), we define \( \alpha(I) \) to be the least \( t \) such that \( I_t \neq 0 \). When \( I = I_Z \), we denote \( \alpha(I) \) by \( \alpha_Z \).

It can be useful to use sheaf theoretic methods. Given a 0-dimensional subscheme \( Z \subseteq \mathbb{P}^2 \), let \( \mathcal{I}_Z \) be the sheafification of \( I_Z \). Then \( (I_Z)_t \) can be identified with the global sections \( \Gamma(\mathbb{P}^2, \mathcal{I}_Z(t)) \) of the twist \( \mathcal{I}_Z(t) = \mathcal{I} \otimes \mathcal{O}_{\mathbb{P}^2}(t) \). The structure sheaf of \( Z \) is \( \mathcal{O}_Z = \mathcal{O}_{\mathbb{P}^2}/\mathcal{I}_Z \).

Another useful fact is Bezout’s Theorem. We state a version of that now.

**Theorem 1.1.6 (Bezout’s Theorem).** Let \( F,G \in R \) be nonzero forms of degrees \( a \) and \( b \). If there are more than \( ab \) points on which both \( F \) and \( G \) vanish, then \( F \) and \( G \) have a common homogeneous factor of positive degree.

1.2. Some initial properties of Hilbert functions of 0-dimensional subschemes of \( \mathbb{P}^2 \). Now we want to study properties of the Hilbert function \( h_Z \) of a 0-dimensional subscheme \( Z \subseteq \mathbb{P}^2 \). Clearly \( h_Z(t) = 0 \) for \( t < 0 \). We begin with a result about the value of \( h_Z(t) \) for \( t \gg 0 \).
Lemma 1.2.1. Given a 0-dimensional subscheme \( Z \subset \mathbb{P}^2 \), then \( h_Z \) is nondecreasing and attains a maximum value, denoted \( \deg(Z) \) and equal to \( \dim_k \Gamma(\mathbb{P}^2, \mathcal{O}_Z) \). We denote the least \( t \geq 0 \) such that \( h_Z(t) = \deg(Z) \) by \( \tau_Z \). Then \( h_Z \) is strictly increasing on \( 0 \leq t \leq \tau_Z \) with \( h_Z(t) = \left( \frac{t+2}{2} \right) \) for \( 0 \leq t < \alpha_Z \).

Proof. Let \( L \) be a linear form not vanishing at any point of support of \( Z \). Then \( L \) is a non-zero divisor on \( R/I \) so the map \( R_t/(I_Z)_t \rightarrow R_{t+1}/(I_Z)_{t+1} \) given by multiplication by \( L \) is injective. Thus \( h_Z(t) \leq h_Z(t+1) \) for all \( t \).

Now consider
\[
0 \rightarrow \mathcal{I}_Z(t) \rightarrow \mathcal{O}_{\mathbb{P}^2}(t) \rightarrow \mathcal{O}_Z(t) \rightarrow 0.
\]
Note that \( h_Z(t) = \dim_K \text{Im}(\Gamma(\mathbb{P}^2, \mathcal{O}_{\mathbb{P}^2}(t)) \rightarrow \Gamma(\mathbb{P}^2, \mathcal{O}_Z(t))) \). But \( \mathcal{O}_Z(t) \cong \mathcal{O}_Z \) since \( Z \) is affine, and by Serre vanishing, \( h^1(\mathbb{P}^2, \mathcal{I}_Z(t)) = 0 \) for \( t \gg 0 \), so the sheaf sequence is exact on global sections for \( t \gg 0 \), hence \( h_Z(t) \leq \dim_K \Gamma(\mathbb{P}^2, \mathcal{O}_Z) \) for all \( t \) with equality for \( t \gg 0 \).

Since \( (I_Z)_t = 0 \) for \( t < \alpha_Z \), we see that \( h_Z(t) = \dim_K R_t = \left( \frac{t+2}{2} \right) \) for \( 0 \leq t < \alpha_Z \).

Finally, let \( \mu_{L,t-1} : R_{t-1}/I_{t-1} \rightarrow R_t/I_t \) be the homomorphism induced by multiplication by \( L \). Say \( h_Z(t) = h_Z(t-1) \); then \( \mu_{L,t-1} \) is onto (in fact bijective). But \( R_{t+1}/I_{t+1} = xR_t/I_t + yR_t/I_t + zR_t/I_t \) and for each \( G \in R_t/I_t \), there is an \( F_G \) such that \( G = LF_G \). Thus given any \( G \in R_{t+1}/I_{t+1} \), there are \( G,G',G'' \in R_t/I_t \) with \( G = xG' + yG'' + zG''' = L(F_G' + F_G'' + F_G''') \), hence \( \mu_{L,t} \) is onto (and hence bijective) so \( h_Z(t+1) = h_Z(t) \) and we get \( h_Z(t + i) = h_Z(t + 1 + i) \) for all \( i \geq 0 \), so \( t > \tau \).

\[ \square \]

Remark 1.2.2. It follows from the long exact sequence in cohomology applied to (1) (and the fact that \( h^1(\mathbb{P}^2, \mathcal{O}_{\mathbb{P}^2}(t)) = 0 \) for all \( t \)) that \( h^1(\mathbb{P}^2, \mathcal{I}_Z(t)) = \deg(Z) - h_Z(t) \). We can also write this as \( h_{I_Z}(t) = \left( \frac{t+2}{2} \right) - (\deg(Z) - h^1(\mathbb{P}^2, \mathcal{I}_Z(t))) \), and hence \( h_{I_Z}(t) \geq \left( \frac{t+2}{2} \right) - \deg(Z) \). We can think of \( \deg(Z) \) as the number of linear conditions an element of \( F \in \mathcal{F} \) must satisfy in order for \( F \) to be in \( (I_Z)_t \). Thus \( h^1(\mathbb{P}^2, \mathcal{I}_Z(t)) \) is the extent to which these conditions are dependent.

By Lemma 1.2.1 we can regard \( h_Z \) as a sequence \( (\ldots, 0, \ldots, 0, 1, \ldots, \deg(Z), \ldots, \deg(Z), \ldots) \), where the 1 occurs in degree 0, the first \( \deg(Z) \) in degree \( \tau_Z \) and in between the entries are strictly increasing. Thus \( \Delta h_Z \) is the sequence \( (\ldots, 0, \ldots, 0, 1, 2, 3, \ldots, \alpha_Z, \ldots, 0, \ldots) \), where again the 1 occurs in degree 0, the \( \alpha_Z \) in degree \( \alpha_Z - 1 \) and the first 0 after that occurs in degree \( \tau_Z + 1 \), after which every entry is 0. It is convenient and conventional to just show the nonzero entries of \( \Delta h_Z \).

Exercise 1.2.3. Let \( Z \) be a 0-dimensional subscheme of \( \mathbb{P}^2 \). Show that \( \sum_t \Delta h_Z(t) = \deg(Z) \).

Exercise 1.2.4. Let \( I_Z = I(p)^2 \cap I(q)^3 \), where \( p \) and \( q \) are distinct points of \( \mathbb{P}^2 \). Show that \( \Delta h_Z = (1, 2, 3, 2, 1) \).

Exercise 1.2.5. Let \( I_Z = (F,G) \), where \( F \) and \( G \) are nonconstant homogeneous forms with no common factors. Let \( a = \deg(F) \leq \deg(G) = b. \) (i.e., \( Z \) is a complete intersection.) Show that \( \deg(Z) = ab \) and that \( \Delta h_Z = (1, 2, 3, \ldots, a, \ldots, a, a-1, \ldots, 2, 1) \), where the number of times that \( a \) appears is \( b - a + 1 \).

Note in the preceding two exercises that the sequence is unimodal (i.e., first nondecreasing, then nonincreasing, hence having at most one peak) with maximum value equal to \( \alpha_Z \). This is not a coincidence.

Proposition 1.2.6. Let \( Z \subset \mathbb{P}^2 \) be a 0-dimensional subscheme. Then:

1. \( h_{I_Z}(t) \geq 2h_{I_Z}(t-1) - h_{I_Z}(t-2) + 1 \) for \( t \geq \alpha_Z \);
2. \( 0 \leq \Delta h_Z(t) \leq \Delta h_Z(t-1) \) for all \( t \geq \alpha_Z \) (i.e., \( \Delta^2 h_Z(t) \leq 0 \) for \( t \geq \alpha \)) with \( \Delta h_Z(t) = 0 \) if and only if \( t > \tau_Z \); and
3. \( \Delta h_Z(t) \leq \Delta h_Z(\alpha_Z - 1) = \alpha_Z \) for all \( t \).
Proof. Let $\alpha = \alpha_Z$.

(1) The statement is clearly true for $t = \alpha$, so let $t > \alpha$. Pick a point $p$ not in the 0-locus of $(I_Z)_t$. Pick two linear forms $A, B$ defining distinct lines through $p$ but not containing any point of support of $Z$. Also pick a linear form $C$ not vanishing at $p$ and a form $F \in (I_Z)_{t-1}$ not vanishing at $p$. Then $FC \notin A(I_Z)_{t-1} + B(I_Z)_{t-1} \subseteq (I_Z)_t$ but $A(I_Z)_{t-1} \cap B(I_Z)_{t-1} = AB(I_Z)_{t-2}$, so $h_{I_Z}(t) \geq \dim_k (A(I_Z)_{t-1} + B(I_Z)_{t-1} + KFC) = 2h_{I_Z}(t-1) - h_{I_Z}(t-2) + 1$.

(2) The first inequality and the if and only if statement follow from Lemma 1.2.1. The second inequality is equivalent to $\Delta h_{I_Z}(t) \geq 1 + \Delta h_{I_Z}(t-1)$, and this is (1).

(3) By Lemma 1.2.1 $\Delta h_{I_Z}(t) = t + 1$ for $t \leq \alpha - 1$ and by (2) $\Delta h_{I_Z}(t)$ is nonincreasing for $t \geq \alpha - 1$, so the maximum value of $\Delta h_{I_Z}(t)$ occurs for $t = \alpha - 1$. \hfill \Box

Given a 0-dimensional subscheme $Z \subset \mathbb{P}^2$, the minimal number of homogeneous generators of $I_Z$ in each degree $t$ can be obtained as follows: let $J = MI_Z$. Thus $J_t$ is that part of the ideal $I_Z$ which can be obtained from $(I_Z)_t$ from degrees $s < t$. Thus the number of homogeneous generators in each degree $t$ needed in a minimal set of homogeneous generators for $I_Z$ is $h_I(t) - h_J(t)$. Generators themselves are given by picking a $K$-basis of $I_t/J_t$ and lifting to $I_t$.

Given a 0-dimensional subscheme $Z$, the quantity $\tau_Z + 1$ is also known as the Castelnuovo-Mumford regularity of $I_Z$, which we denote by $\sigma_Z$. We saw above that $\Delta h_{I_Z}(t) = 0$ if and only if $t \geq \sigma_Z$. Another indication of its importance is the following result giving bounds on numbers of generators, which together with the improvement given in Corollary 1.3.9 are due in some form to Dubreil [10]. A generalization to codimension 2 arithmetically Cohen-Macaulay subschemes of projective space is given in [4].

**Proposition 1.2.7.** Let $Z \subset \mathbb{P}^2$ be a 0-dimensional subscheme. Let $\nu_Z(t)$ be the number of generators in degree $t$ in a minimal set of homogeneous generators for $I_Z$. Then:

1. $\nu_Z(t) = 0$ for $t < \alpha_Z$;
2. $\nu_Z(t) = h_{I_Z}(t) = 1 - \Delta^2 h_{I_Z}(t)$ for $t = \alpha_Z$;
3. $\nu_Z(t + 1) \leq -\Delta^2 h_{I_Z}(t + 1)$ for $\alpha_Z \leq t$; and hence
4. $\nu_Z(t) = 0$ for $t > \sigma_Z$.

Thus $h_{I_Z}(\alpha_Z) \leq \sum \nu_Z(t) \leq \alpha_Z + 1$; i.e., the number $\nu_Z$ of generators in a minimal set of homogeneous generators is at least $h_{I_Z}(\alpha_Z)$ and at most $\alpha_Z + 1$.

**Proof.** Let $\alpha = \alpha_Z$. First, (1) is clear since $(I_Z)_t = 0$ for $t < \alpha$. Similarly, for (2): $\nu_Z(\alpha) = h_{I_Z}(\alpha)$, and it is easy to check that $\Delta^2 h_{I_Z}(\alpha) = 1 - h_{I_Z}(\alpha)$, hence $h_{I_Z}(\alpha) = 1 - \Delta^2 h_{I_Z}(\alpha)$.

For (3), assume $t \geq \alpha$. Then arguing as in the proof of Proposition 1.2.6 (1), $h_{I_Z}(t + 1) \geq \nu_Z(t + 1) + 2h_{I_Z}(t) - h_{I_Z}(t - 1) + 1$, which is equivalent to $\nu_Z(t + 1) \leq -\Delta^2 h_{I_Z}(t + 1)$. Since $\Delta^2 h_{I_Z}(t) = \Delta h_{I_Z}(t) - \Delta h_{I_Z}(t - 1)$ and both terms are 0 for $t - 1 \geq \sigma_Z$, we see $\Delta^2 h_{I_Z}(t) = 0$ for $t \geq \sigma_Z + 1$, which gives (4). Summing over $t$ now gives $\sum \nu_Z(t) = 1 - \sum_{t \geq \alpha} \Delta^2 h_{I_Z}(t) = 1 + \Delta h_{I_Z}(\alpha - 1) = \alpha + 1$. \hfill \Box

**Exercise 1.2.8.** Give an example to show the bound $\sum \nu_Z(t) \leq \alpha_Z + 1$ in Proposition 1.2.7 is sharp. Give another example to show it is not always an equality.

**Exercise 1.2.9.** Give an example with $\nu_Z(\sigma_Z) > 0$.

1.3. Additional properties of Hilbert functions of 0-dimensional subschemes of $\mathbb{P}^2$.

Given a 0-dimensional subscheme $Z \subset \mathbb{P}^2$ and its Hilbert function $h_Z$, we can immediately determine $\alpha_Z$ and $\sigma_Z$. We now introduce two quantities that are more subtle and are not determined by $h_Z$ (see Exercise 1.4.3).

**Definition 1.3.1.** Let $Z \subset \mathbb{P}^2$ be a 0-dimensional subscheme. Then $\omega_Z$ denotes the degree of a homogeneous generator of $I_Z$ of maximum degree among a minimal set of homogeneous generators of $I_Z$; thus $\nu_Z(\omega_Z) > 0$ but $\nu_Z(t) = 0$ for $t > \omega_Z$. I.e., it is the least degree $t$ such that the
multiplication maps \((I_Z)_s \otimes R_1 \to (I_Z)_{s+1}\) are surjective for all \(s \geq t\) but \((I_Z)_{t-1} \otimes R_1 \to (I_Z)_t\) is not surjective. And \(\beta_Z\) denotes the least degree \(t\) such that there is no positive degree form \(F\) that divides every element of \((I_Z)_t\) (i.e., such that the base locus of \((I_Z)_t\) is not positive dimensional).

**Exercise 1.3.2.** Show that \(\nu_Z(\beta_Z) \geq 1\) and hence that \(\beta_Z \leq \omega_Z\). Give an example to show that \(\beta_Z < \omega_Z\) can occur.

We thus have \(\alpha_Z \leq \beta_Z \leq \omega_Z \leq \sigma_Z\). Of course, sometimes \(\beta_Z\) and \(\omega_Z\) can be determined just given \(h_Z\), such as when \(\alpha_Z = \sigma_Z\). But in general, \(\beta_Z\) and \(\omega_Z\) require information beyond just \(h_Z\) and thus we can obtain additional results if we know them. For example, we have \(\nu_Z(\beta_Z) \geq 1\) from Exercise 1.3.2 and hence \(-\Delta^2 h_Z(\beta_Z) \geq \nu_Z(\beta_Z) \geq 1\) if \(\beta_Z > \alpha_Z\), which gives \(\Delta h_Z(t) < \Delta h_Z(t-1)\) when \(t = \beta_Z\) and \(\beta_Z > \alpha_Z\).

**Exercise 1.3.3.** If \(\beta_Z = \alpha_Z\) and \(t = \beta_Z\), show that \(\Delta h_Z(t) < \Delta h_Z(t-1)\) follows directly from the definitions. Give an example with \(\beta_Z > \alpha_Z = t\) where \(\Delta h_Z(t) = \Delta h_Z(t-1)\).

**Exercise 1.3.4.** If \(t = \sigma_Z\), show that \(\Delta h_Z(t) < \Delta h_Z(t-1)\) again follows directly from the definitions.

More generally we have:

**Theorem 1.3.5.** Let \(Z\) be a 0-dimensional subscheme of \(\mathbb{P}^2\). Then \(\Delta h_Z(t) < \Delta h_Z(t-1)\) for \(\beta_Z \leq t \leq \sigma_Z\).

To extend this fact we need some results about ideals in \(K[x,y]\).

**Exercise 1.3.6.** For any \(f \in S = K[x,y]\), define \(\text{In}(f)\) to be 0 if \(f = 0\) and otherwise define it to be the monomial \(x^a y^b\) with \(a\) as small as possible among the monomial terms of least degree appearing in \(f\). Let \(I\) be any homogeneous ideal in \(S\) and define \(J = \text{In}(I)\) to be the ideal generated by the set \(\{\text{In}(f) : f \in I\}\). Show that \(h_J(t) = h_I(t)\).

**Exercise 1.3.7.** Let \(V\) be a nonzero vector subspace of \(S = K[x,y]\) spanned by a set \(\mathcal{M}\) of monomials of degree \(t\). Let \(W = xV + yV\). Show \(\dim W = \dim V + 1\) if and only if for some \(0 \leq i \leq j \leq t\) we have \(\mathcal{M} = \{x^i y^{j-i}, x^{i+1} y^{j-i-1}, \ldots, x^t y^{j-t}\}\) (i.e., if and only if the ideal generated by \(\mathcal{M}\) is \(x^i y^{t-j}(x,y)^{j-i}\)).

**Lemma 1.3.8.** Let \(S = K[x,y]\) and let \(V \subset S_t\) be a \(K\)-vector subspace. Let \(I\) be the ideal generated by \(V\), so \(I_t = V\). Let \(W = xV + yV\), so \(I_{t+1} = W\).

1. We have \(\dim W = \dim V\) (i.e., \(h_I(t+1) = h_I(t)\)) if and only if \(V = (0)\).
2. We have \(\dim W = \dim V + 1\) (i.e., \(h_I(t+1) = h_I(t) + 1\)) if and only if \(V = FS_r\) for some \(r \geq 0\) and some nonzero homogeneous \(F \in S_{t-r}\), (i.e., \(h_I(s+1) = s-t+r+2 = h_I(s)+1\) for all \(s \geq t\)).

**Proof.** (1) If \(V = (0)\), then certainly \(W = (0)\) so \(\dim W = \dim V\). Conversely, assume \(V \neq (0)\). Let \(F\) be a gcd for \(V\) so we have \(V = FU\) for some subspace \(U\). Let \(U' = xU + yU\) so \(\dim U' = \dim V\) and \(\dim U'' = \dim W\). Pick a basis \(\{u_1, \ldots, u_r\}\) for \(U\). Then \(\{xu_1, \ldots, xu_r, yu_1, \ldots, yu_r\}\) spans \(U'\). Clearly \(\{xu_1, \ldots, xu_r\}\) are linearly independent and so are \(\{yu_1, \ldots, yu_r\}\). If \(U'' = \dim U\), then each \(yu_i\) is in \(xU\) and hence each \(u_i\) is divisible by \(x\). But by choice of \(F\), no such common factor is possible. Thus \(\dim V = \dim U < \dim U'' = \dim W\).

(2) If \(V = FS_t\) for some \(t \geq 0\) and some nonzero homogeneous \(F \in S\), then \(W = FS_{t+1}\) and hence \(\dim W = t+2 = \dim V + 1\).

Conversely, let \(I\) be the ideal generated by \(V\). By Exercise 1.3.6 there is a monomial ideal \(J\) with the same Hilbert function. Since \(I_t = V\) and \(I_{t+1} = W\), we have \(h_J(t+1) = h_J(t+1) = h_I(t+1) = h_I(t) + 1\). By Exercise 1.3.7 we have \(h_J(s) = h_J(s) = s-t+h_J(t) = s-t+h_I(t)\) for all \(s \geq t\). If \(F\) is a gcd for \(I\) of some degree \(t-r\), then \(I \subseteq H\), where \(H = (F)\), and for \(s \gg 0\) we have \(I_s = FS_{s-t+r}\). Therefore \(s-t+h_I(t) = h_I(s) \leq h_H(s) = s-t+r+1\) for all \(s \geq t\) with
equality for \( s \gg 0 \). Thus \( h_1(t) = r + 1 \) and equality holds for all \( s \geq t \), hence \( I_s = H_s \) for \( s \geq t \); i.e., \( V = FS_r \).

**Proof of Theorem 1.3.5.** Assume \( \alpha_Z \leq \beta_Z \leq t \leq \sigma \) (so \( 0 < \Delta h_1(t) \) and \( 0 < \Delta h_2(t - 1) \)), and suppose on the contrary that \( \Delta h_2(t) = \Delta h_2(t - 1) = \delta > 0 \). Thus \( \Delta h_1(t) = \Delta h_1(t - 1) = \delta \). Now by Lemma 1.3.8, \( h_1(s) = s - \deg(F) + 1 \) for \( s \geq t - 1 \) and thus \( \nu_Z(s) = \alpha_Z + \beta_Z - \sigma_Z + 1 \). Since \( \nu_Z(s) \) is the number of generators in a minimal set of homogeneous generators of \( I_Z \), we know that \( I \) defines a 0-dimensional subscheme \( Y \) containing \( Z \), so \( h_1(s) = \binom{s + 2}{2} - \deg(Y) \) for \( s \gg 0 \) and hence \( \Delta h_2(t) = \Delta h_2(t - 1) = \delta \). Now by Lemma 1.3.8, \( h_1(s) = s - \deg(F) + 1 \) for \( s \geq t - 1 \) for some \( F \in S \). Thus \( h_1(s) = \deg(F) \) for all \( s \geq t - 1 \). But \( \delta = h_1(t - 1) = h_1(s) = \Delta h_1(t) = 0 \) for \( s \gg 0 \), contradicting our assumption that \( \Delta h_1(t) = \Delta h_2(t - 1) \). \[ \square \]

We can now draw some additional algebraic consequences.

**Corollary 1.3.9.** Let \( Z \) be a 0-dimensional subscheme of \( \mathbb{P}^2 \).

1. If \( \beta_Z \leq t \leq \sigma_Z \), then \( h_1(t) \geq 2h_1(t - 1) - h_1(t - 2) + 2 \).
2. If \( \beta_Z < t \leq \sigma_Z \), then \( \nu_Z(t) \leq -\Delta^2 h_2(t - 1) - 1 \), and thus \( \nu_Z \leq \alpha_Z + \beta_Z - \sigma_Z + 1 \), where \( \nu_Z \) is the number of generators in a minimal set of homogeneous generators of \( I_Z \).

**Proof.** (1) Note that \( \Delta h_2(t) < \Delta h_2(t - 1) \) is equivalent to \( \Delta h_1(t) = \Delta h_1(t - 1) + 1 < \Delta h_1(t) \), so this follows by Theorem 1.3.5.

(2) Let \( I \) be the ideal generated by all homogeneous elements of \( I_Z \) of degree at most \( t \). Let \( Y \) be the subscheme defined by \( I \). Since \( \beta_Z \leq t - 1 \), we see \( Y \) is 0-dimensional. Clearly, \( (I_Y)_s = I_s = (I_Z)_s \) for \( s \leq t - 1 \), but \( I = R(I_Z)_{t-1} \). Since \( t \leq \sigma_Z \), we see \( 0 < \Delta h_2(t) = \Delta h_2(t - 1) \), so \( \nu_Y \geq t \). Thus by part (1) we have \( \dim R(I_Z)_{t-1} = h_1(t) = h_1(t) - 2h_1(t - 1) - h_1(t - 2) + 2 \), and hence \( h_1(t) \geq \nu_Z(t) + 2h_1(t - 1) - h_1(t - 2) + 2 \), which is equivalent to \( \nu_Z(t) \leq -\Delta^2 h_2(t - 1) - 1 \). Summing over all \( t \geq \alpha_Z \) gives \( \nu_Z \leq 1 + \Delta h_2(\alpha - 1) = \delta_{Z - \beta_Z} = \alpha_Z + \beta_Z - \sigma_Z + 1 \). \[ \square \]

**Exercise 1.3.10.** Let \( Z \) be a 0-dimensional subscheme of \( \mathbb{P}^2 \). Show that \( \sigma_Z \leq \alpha_Z + \beta_Z - 1 \), with equality if and only if \( \nu_Z = 2 \) (i.e., if and only if \( Z \) is a complete intersection).

We can also draw a geometric consequence, due to Davis \[ \square \] and significantly extended by Bigatti, Geramita and Migliore \[ \square \].

**Corollary 1.3.11.** Let \( Z \) be a 0-dimensional subscheme of \( \mathbb{P}^2 \). If \( \Delta h_2(t) = \Delta h_2(t - 1) \) for some \( t \leq \sigma_Z \), then \( (I_Z)_t \) has a gcd \( F \) of degree \( \Delta h_2(t) \) and \( F \) defines a curve \( C \) such that \( Z_1 = C \cap Z \) has \( \Delta h_{Z_1}(s) = \min(\Delta h_Z(s), \Delta h_2(t)) \) and hence \( \deg(Z_1) = \sum_s \min(\Delta h_Z(s), \Delta h_2(t)) \).

**Proof.** If \( t \geq 0 \) and \( \Delta h_2(t) = \Delta h_2(t - 1) \), then \( t \geq \alpha_Z \), so any gcd for \( (I_Z)_t \) is well defined up to scalars. Since \( \Delta h_2(t) = \Delta h_2(t - 1) \) and \( t \leq \sigma_Z \), Theorem 1.3.5 ensures that \( t < \beta_Z \), so the gcd \( F \) of \( (I_Z)_t \) has positive degree, say \( \deg(F) = \delta \). Let \( I \) be the ideal generated by all homogeneous elements of \( I_Z \) of degree \( t \) or less. Thus \( \Delta h_1(s) = \Delta h_1(s) \) for \( s \leq t \), so \( \Delta h_1(t - 1) + 1 = \Delta h_1(t) \). Let \( A = I : F \). Then \( A \) defines a 0-dimensional subscheme \( Y \) (or the empty subscheme) so \( h_1(s + \delta) = h_A(s) = \binom{s + 2}{2} - \deg(Y) \) for \( s \gg 0 \), hence \( \Delta h_A(s + \delta) = \Delta h_A(s) = s + 1 \) for \( s \gg 0 \), so \( \Delta h_1(s) = s - \delta + 1 \) for \( s \gg 0 \).

Let \( L \) be a linear form not vanishing at the points of support of \( Z \). After a linear change of coordinates we may assume \( L = z \) and \( R/I = S = K[x,y,z] \), where \( R = K[x,y,z] \). Let \( J \) be the image of \( I \) in \( S \). Then \( h_J(s) = \Delta h_1(s) \) for all \( s \). Now by Lemma 1.3.8 we have \( h_j(s) = s - t + r + 1 \) for some \( 0 \leq r \leq t \) and all \( s \geq t \). Then \( s - t + r + 1 = \Delta h_j(s) = \Delta h_1(s) = s - \delta + 1 \) for \( s \gg 0 \) so \( t - r = \delta \) and \( \delta = t + 1 - h_j(t) = \Delta h_1(t) \).
For the rest, we mimic the proof of [2, Theorem 2.4]. We first show that $I_Z + (F)$ is saturated. In any case, its saturation defines the subscheme $Z_1 = Z \cap C$, where $C$ is the curve defined by $I_C = (F)$. First note that $(I_Z + (F))_s = (F)_s$ for $s \leq t$, since $(I_Z)_s \subset (F)_s$ in these degrees. If $(I_Z)_s = (I_Z + (F))_s$ for some $s \leq t$, then there would be a form $G$ and an $n$ such that $(L^n G \in (F))$ for every linear form $L'$ but $G$ is not in $(F)$, which is impossible (just take $L'$ not to be a factor of $L$). Thus $(I_Z + (F))_s = (I_Z)_s$ for all $s \leq t$.

Now we verify this for larger $s$. Let $I_{Z_2} = I_Z : F$. This is saturated since $I_Z$ is, and defines a 0-dimensional scheme $Z_2$ (possibly empty) since $F$ is the gcd of $(I_Z)_t$, so $(I_{Z_2})_{t-\delta}$ has no curves in its base locus. We have an exact sequence
\[ 0 \to I_{Z_2}(t-\delta) \to I_Z \otimes I_C \to I_Z + (F) \to 0, \]
where $I_{Z_2}(t-\delta)$ is the graded module with $(I_{Z_2}(t-\delta))_s = (I_{Z_2})_{s-\delta}$ and the map $I_{Z_2}(t-\delta) \to I_Z \otimes I_C$ is $G \mapsto GF \otimes GF$. Sheafifying gives
\[ 0 \to I_{Z_2}(t-\delta) \to I_Z \otimes I_C \to I_{Z_1} \to 0. \]
\[ (2) \]
Since $F$ is a gcd for $(I_Z)_t$, we have $F(I_{Z_2})_{s-\delta} = (I_Z)_s$ for $s \leq t$. Now $\Delta h_{Z_2}(t) = \Delta h_{Z}(t-1)$ and arithmetic shows that $\Delta h_{Z_2}(t-\delta) = \Delta h_{Z_2}(t-\delta-1)$. But $(I_{Z_2})_{t-\delta}$ has no curve in its base locus, so by Theorem 1.3.5 we must have $t-\delta > \sigma Z$. Thus $h^0(P^2, I_Z(s-\delta)) = 0$ for all $s \geq t-\delta - 2$, so (2) is exact on global sections in these degrees. Since $I_{Z_2}$ and $I_Z \otimes I_C$ are saturated, the exact sequence
\[ 0 \to \Gamma(P^2, I_{Z_2}(s-\delta)) \to \Gamma(P^2, I_Z(s) \otimes I_C(s)) \to \Gamma(P^2, I_{Z_1}(s)) \to 0 \]
is just
\[ 0 \to (I_{Z_2})_{s-\delta} \to (I_Z)_s \otimes (I_C)_s \to \Gamma(P^2, I_{Z_1}(s)) \to 0 \]
and hence $(I_{Z_1})_s = (I_Z + (F))_s$ for $s \geq t - 2$ and thus for all $s$; i.e., $I_Z + (F)$ is saturated.

Finally we must check that $\Delta h_{Z_1}(s) = \min(\Delta h_{Z_2}(t), \Delta h_{Z_2}(s))$. For $s \leq t$, we have $(I_{Z_1})_s = (I_Z + (F))_s = (F)_s$ contains $(I_Z)_s$, so $\Delta h_{Z_1}(s) = s + 1 = \Delta h_{Z_2}(s)$ for $s < \delta$ and $\Delta h_{Z_1}(s) = \delta$ for $\delta \leq s \leq t$. From (3) and arithmetic we get $\Delta h_{Z_2}(s) = \Delta h_{Z_1}(s) + \Delta h_{Z_2}(s-\delta)$ for $s \geq t - 2$, but as we saw above $\Delta h_{Z_2}(s-\delta) = 0$ since $s-\delta \geq t-\delta-2 \geq \sigma Z$. Thus $\Delta h_{Z_1}(s) = \Delta h_{Z_2}(s)$ for $s \leq \delta - 1$ and $s \geq t$, with $\Delta h_{Z_1}(\delta - 1) = \Delta h_{Z_1}(t)$. By unimodality, we must have $\Delta h_{Z_1}(s) \leq \delta \leq \Delta h_{Z_2}(s)$ for $\delta - 1 \leq s \leq t$.

\[ \Box \]

**Exercise 1.3.12.** Typically the sum of ideals is not saturated, in contrast to $I_Z + (F)$ in the proof of Corollary 1.3.11. Give an easy example of a 0-dimensional subscheme $Z \subset P^2$ and a form $F$ such that $I_Z + (F)$ is not saturated.

**Exercise 1.3.13.** Suppose a 0-dimensional subscheme $Z \subset P^2$ has
\[ \Delta h_Z = (1, 2, 3, 4, 5, 6, 7, 5, 3, 2, 1, 0), \]
Determine $\text{deg}(Z), \alpha_Z, \tau_Z, \sigma_Z$, and indicate what you can say about $\beta_Z, \nu_Z$, and about $\nu_Z(t)$ and about curves in the base locus of $(I_Z)_t$, for each degree $t$.

**1.4. Fat points.** A special but important case of a 0-dimensional subscheme $Z \subset P^2$ is that of fat points. We say $Z$ is a fat point subscheme if $I_Z = (p_1)^{m_1} \cdot \cdots \cdot (p_r)^{m_r}$; i.e., if the primary ideals $Q_i$ are powers of the ideals of the points. It is convenient to write $m_1 p_1 + \cdots + m_r p_r$ for the fat point subscheme $Z$ defined by $(p_1)^{m_1} \cdot \cdots \cdot (p_r)^{m_r}$. Note that a form $F$ is in $(I_Z)_t$ if and only if $\text{deg}(F) = t$ and $F$ has multiplicity $\text{mult}_p(F) \geq m_i$ at each point $p_i$. [Tautologically, $\text{mult}_p(F)$ is the largest $m$ such that $F \in (I(p))^m$]. Alternatively, if $U, V$ are linearly independent linear forms that vanish at $p$, and $W$ a linear form that does not vanish at $p$, we can write $F$ in terms of $U, V, W$. Then $\text{mult}_p(F)$ is the least degree of a monomial in $U, V$ appearing in $F(U, V, 1)$. As yet another alternative, if $p = [a : b : c]$, pick (for specificity) the first nonzero coordinate, say $b$. Then we can consider $F_p(u, v) = F(u + a/b, 1, v + c/b)$ and expand in terms of $u$ and $v$. Then $\text{mult}_p(F)$ is the
least degree of a monomial in $u, v$ appearing in $F_p(u, v)$. If we define a truncation map $\text{tr}_m(F_p(u, v))$ as removing all terms of degree $m$ or more in $u$ and $v$ (e.g., $\text{tr}_2(3 + 2u + 4v + 2uv + 7u^2v + 8v^4) = 3 + 2u + 4v + 2uv$), then the last map in the exact sequence

$$0 \rightarrow (I_Z)_t \rightarrow R_t \rightarrow \Gamma(\mathbb{P}^2, \mathcal{O}_Z(t))$$

is $F \mapsto \oplus_i \text{tr}_{m_i-1}(F_p)$. There are $(m_i+1)$ monomials in $u$ and $v$ of degree less than $m_i$. For $F \in R_t$ to actually be in $(I_Z)_t$, $F$ must satisfy these $\sum_i (m_i+1)$ linear conditions; i.e., $\deg(Z) = \sum_i (m_i+1)$. As noted in Remark 1.2.2, $h^1(\mathbb{P}^2, I_Z(t))$ is the extent to which these conditions are dependent.

**Example 1.4.1.** Let $Z = 2p + 2q$ for distinct points $p, q \in \mathbb{P}^2$. By change of coordinates we may assume $p = [0, 0, 1]$ and $q = [0, 1, 0]$. A form $F = az^2 + bxz + cyz + dxy + ex^2 + fy^2 \in R_2$ is in $I(p)^2$ if and only if $a = b = c = 0$ (so $2p$ imposes 3 conditions on the 6 dimensional space $R_2$) and $F$ is in $I(q)^2$ if and only if $c = d = f = 0$ (so $2q$ imposes 3 conditions), hence $Z$ imposes 6 conditions; i.e., $\deg(Z) = 6$. But these conditions are dependent ($c = 0$ is in both); i.e., $h^1(\mathbb{P}^2, I_Z(2)) = 1$ and so $h_{I_Z}(2) = \binom{4}{2} - (\deg(Z) - h^1(\mathbb{P}^2, I_Z(2))) = 6 - 5 = 1$, and indeed we have $(I_Z)_2 = (x^2)$.

We can obtain this last fact using Bezout’s Theorem. Suppose $F \in (I_Z)_2$. Thus $F$ has multiplicity at least 2 at both $p$ and $q$ (indicated heuristically in the following figure):

![Diagram of the curve D and the line H](image)

The curve $D$ defined by $F = 0$ and the line $H$ defined by $x = 0$ meet at least 4 times (at least twice at $p$ and twice at $q$), which is more than $\deg(D) \deg(H) = 2$, so they have a component in common, which must be $H$. Removing $H$ from $D$ gives a curve $C$ of degree 1 whose multiplicity at $p$ and $q$ goes down by 1 (since $H$ is no longer there but $\text{mult}_p(x) = \text{mult}_q(x) = 1$). Thus $C$ meets $H$ at least twice (once at each point $p$ and $q$) hence $C$ and $H$ have a component in common, so $C = H$. I.e., $D = 2H$ so $F = x^2$ (up to a nonzero scalar).

**Exercise 1.4.2.** Let $Z = m_1p_1 + \cdots + m,rp_r \subset \mathbb{P}^2$ be a fat point subscheme. Use Bezout’s Theorem to show that $\alpha Z \beta Z \geq \sum_i m_i^2$. Give an example to show that $\beta Z > \left\lceil \frac{\sum_i m_i^2}{\alpha Z} \right\rceil$ can occur.

**Exercise 1.4.3.** Give two different fat point subschemes $Z_1, Z_2$ such that $h_{Z_1} = h_{Z_2}$ but $\beta_{Z_1} \neq \beta_{Z_2}$. Give another such example with $\omega_{Z_1} \neq \omega_{Z_2}$.

Fat points allow us to bring in more geometry. Given $I_Z = I(p_1)^{m_1} \cap \cdots \cap I(p_r)^{m_r}$, let $\pi : X \rightarrow \mathbb{P}^2$ be the birational morphism obtained by blowing up the points $p_i$. Let $L$ be the pullback to $X$ of a general line, and let $E_i = \pi^{-1}(p_i)$. The divisor class group $\text{Cl}(X)$ of divisors on $X$ modulo linear equivalence is a free abelian group generated by the divisor classes $[L], [E_1], \ldots, [E_r]$. It also comes with a bilinear form, defined by $-1 = -L^2 = E_i^2$, with $L \cdot E_i = E_i \cdot E_j = 0$ for all $i \neq j$. 
Given nonnegative integers $m_1, \ldots, m_r$ and $Z = m_1p_1 + \cdots + m_rp_r$, we can make the following identifications:

$$H^0(X, \mathcal{O}_X(tL - m_1E_1 - \cdots - m_rE_r)) = H^0(\mathbb{P}^2, \mathcal{I}_Z(t)) = (I_Z)_t.$$  

We also have

$$h^1(X, \mathcal{O}_X(tL - m_1E_1 - \cdots - m_rE_r)) = h^1(\mathbb{P}^2, \mathcal{I}_Z(t)) = \deg(Z) - h_Z(t).$$

The quantity $a_Z$ is thus the least $t$ such that $h^0(X, \mathcal{O}_X(tL - m_1E_1 - \cdots - m_rE_r)) > 0$ and $\tau_Z$ is the least $t \geq 0$ such that $h^1(X, \mathcal{O}_X(tL - m_1E_1 - \cdots - m_rE_r)) > 0$.

**Example 1.4.4.** In Example 1.4.1 if we take $p = p_1$ and $q = p_2$, then $D$ corresponds to a divisor $D'$ on $X$ with $|D'| = |2L - 2E_1 - 2E_2|$ and $H$ corresponds to a divisor $H'$ on $X$ with $|H'| = |L - E_1 - E_2|$. Now $D' \cdot H' = -2 < 0$ so if $D'$ is effective it must be the case that $H'$ (being effective, reduced and irreducible) is a component, so $C' = D' - H'$ would be effective with $[C'] = [L - E_1 - E_2]$. Again $C' \cdot H' = -1 < 0$, so $C'$ must be $H'$. I.e., $D' = 2H'$, hence $D = 2H$ and $(I_Z)_2 = \langle \chi^2 \rangle$ as we found before.

**Proposition 1.4.5.** Let $p_1, \ldots, p_r \in \mathbb{P}^2$ be distinct points with nonnegative integers $m_i$. Let $Z = m_1p_1 + \cdots + m_rp_r$. Then $\tau_Z \leq -1 + \sum_i m_i$.

**Proof.** If $i = 1$, then after a change of coordinates (so that $p_1 = [0, 0, 1]$), $\mathcal{I}_Z$ is a monomial ideal generated by the monomials $x^iy^jz^k$ with $i + j \geq m_1$. Thus $h_{\mathcal{I}_Z}(t)$ is the number of monomials of degree $t = i + j + k$ with $i + j \geq m_1$, and $h_{\mathcal{I}_Z}(t)$ is the number of monomials with $t = i + j + k$ and $i + j < m_1$. Thus $h_{\mathcal{I}_Z}(t) = \min\left(\binom{m_1+1}{2}, \binom{t+2}{2}\right)$. The least $t$ such that $h_{\mathcal{I}_Z}(t) = \deg(Z) = \binom{m_1+1}{2}$ is $t = m_1 - 1$. Let $X$ be the blow up of $\mathbb{P}^2$ at the points $p_i$ and take $F_r = (-1 + \sum_i m_i)L - m_1E_1 - \cdots - m_rE_r = ((m_1 - 1)L - m_1E_1) + m_2(L - E_2) + \cdots + m_r(L - E_r)$. We have just shown that $h^1(X, \mathcal{O}_X(F_1)) = 0$ for $F_1 = (-1 + m_1)L - m_1E_1$.

Now assume $r > 1$ and induct on $r$. It is enough to show that $h^1(X, \mathcal{O}_X(F_r)) = 0$, given that $h^1(X, \mathcal{O}_X(F_{r-1})) = 0$. Let $C$ be the proper transform of a line through $p_r$ that does not contain $p_i$ for any $i < r$. We have the short exact sequence

$$0 \to \mathcal{O}_X(F_{r-1} + (m_r - 1)(L - E_r)) \to \mathcal{O}_X(F_r) \to \mathcal{O}_C(-1 + \sum_i m_i) \to 0.$$  

When $m_r = 1$, we have $\mathcal{O}_X(F_{r-1} + (m_r - 1)(L - E_r)) = \mathcal{O}_X(F_{r-1})$ so by induction that $h^1(X, \mathcal{O}_X(F_{r-1} + (m_r - 1)(L - E_r))) = 0$. And $C \cong \mathbb{P}^1$ with $-1 + \sum_i m_i \geq -1$ so $h^1(C, \mathcal{O}_C(-1 \sum_i m_i)) = 0$. Thus $h^1(X, \mathcal{O}_X(F_{r-1} + (L - E_r))) = 0$. Induction on $m_r$ now gives the result for each $m_r \geq 0$. \hfill \Box

**Exercise 1.4.6.** Give an example with $\tau_Z = -1 + \sum_i m_i$ and another with $\tau_Z < -1 + \sum_i m_i$.

1.5. **Upper and lower bounds on $h_Z$ for a fat point subscheme** $Z \subset \mathbb{P}^2$. In section 1.3 we used properties of Hilbert functions to extract geometric information about curves containing subschemes of a 0-dimensional scheme $Z$. Now we show how to use information about curves containing subschemes of a 0-dimensional scheme $Z$ to get information about its Hilbert function. Specifically, given a fat point subscheme $Z$ here we show how to convert information about which points are colinear to obtain bounds on the Hilbert function $Z$. Let $Z = m_1p_1 + \cdots + m_rp_r \subset \mathbb{P}^2$. Let $L_1, \ldots, L_s$ be linear forms such that $\text{mult}_{p_i}(L_1 \cdots L_s) \geq m_i$ for each $i$ (i.e., at least $m_i$ of the lines pass through each point $p_i$).

We now define a sequence $Z = Z_0, Z_1, \ldots, Z_s$, for which we define $Z_j = m_1p_1 + \cdots + m_jp_j$ where, for $j > 0$, $m_{ij} = \max(0, m_i - \text{mult}_{p_i}(L_1 \cdots L_j))$. For $0 \leq j < s$, we also define a **reduction vector** $v = (v_0, \ldots, v_{s-1})$ where $v_j = \sum_{p_i \in L_{j+1}} m_{ij}$. Note that $v$ depends both on $Z$ and on the choice of lines $L_i$. 


In terms of \( v \) define two functions \( f_v \) and \( F_v \) as follows (where we regard \( \binom{m}{n} \) as 0 if \( m < n \)):

\[
f_v(t) = \sum_{i=0}^{s-1} \left( \min(t - i + 1, v_i) \right),
\]

\[
F_v(t) = \min_{0 \leq i \leq s} \left( \frac{t + 2}{2} - \left( t - i + 2 \right) + \sum_{i \leq j \leq s-1} v_j \right).
\]

**Exercise 1.5.1.** Given a sequence \( v = (v_0, \ldots, v_{s-1}) \) of nonnegative integers, arrange dots as follows: \( v_i \) dots on the line \( u = i + 1 \) with a dot at \((j, i + 1)\) if and only if \( i \leq j < v_i + i \). Show that \( f_v(d) \) is the number of dots on or to the left of the line \( t = d \) and thus \( \Delta f_v(d) \) is the number of dots on the line \( t = d \). The diagram below shows the dot diagram for \( v = (8, 6, 4, 3) \).

![Diagram showing dot diagram for v = (8, 6, 4, 3)](image)

The dot diagram in Exercise 1.5.1 looks a lot like the graph of \( \Delta h_Z \) for some 0-dimensional subscheme \( Z \subset \mathbb{P}^2 \). The following theorem, which is one of the main results of Section 2.5.1, shows that this is not a coincidence.

**Theorem 1.5.2.** Given \( Z, L_1, \ldots, L_s \) and \( v \) as above, then \( f_v(t) \leq h_Z(t) \leq F_v(t) \) for all \( t \geq 0 \) and both inequalities are equalities if either \( v \) is a decreasing sequence with nonzero entries or there is some \( j \) such that \( v \) is of the form \( v_0 > v_1 > \ldots > v_j = \cdots = v_s = 0 \).

An alternative description of \( f_v(t) \) can be given. Put dots at integer lattice points in the following way. Put a dot at the lattice point \((j, e)\) for each \( j = 0, \ldots, s - 1 \) and each \( 0 \leq e < v_j \). Then by a result of Section 2.5.1 we have that \( f_v(t) \) is the number of dots on or to the left of the line \( y = -x + t \). Thus if either \( v \) is a decreasing sequence with nonzero entries or \( v \) is of the form \( v_0 > v_1 > \ldots > v_j = \cdots = v_s = 0 \), then \( \Delta h_Z(t) = \Delta f_v(t) \), and this is just the number of dots on the line \( y = -x + t \).

The bounds given by Theorem 1.5.2 are most useful when there are a lot of large subsets of points on lines, as Example 1.5.4 suggests. They are not so useful for general points.

**Exercise 1.5.3.** Let \( Z = 2p_1 + \cdots + 2p_6 \) where \( p_1, \ldots, p_6 \) lie on an irreducible conic. Find \( f_v \) and \( F_v \) for some choice of lines \( L_i \) (so each line can be chosen to go through at most 2 points). Compare this to the result you obtain for \( Z \) assuming the points are general.

Another result of Section 2.5.1 (specifically Theorem 3.1.6) gives upper and lower bounds on numbers of generators. When the numbers \( v_i \) are strictly decreasing these bounds coincide and thus determine the numbers of generators exactly (see Section 1.6), but sometimes one can determine numbers of generators geometrically.

**Example 1.5.4.** Suppose \( Z = m(p_1 + \cdots + p_i) + n(q_1 + \cdots + q_j) \) where \( p_1, \ldots, p_i \) are on one line (call it \( A \)), \( q_1, \ldots, q_j \) are on another line (call it \( B \)), no point is on both lines, and \( i > n_j \). Let our lines \( L_1, \ldots, L_s \) be \( L_1 = \cdots = L_m = A \) and \( L_{m+1} = \cdots = L_n = B \), where \( s = m + n \). Then \( v = (mi, (m-1)i, \ldots, i, n_j, (n-1)j, \ldots, j) \), and we see that \( h_Z = f_v \). So for example if \( m = 3, i = 5, n = 2, j = 2, \) we have \( v = (15, 10, 5, 4, 2) \), and \( \Delta h_Z = \Delta f_v = (1, 2, 3, 4, 5, 4, 2, 2, 2, 1, 1, 1) \). Thus deg(\( Z \)) = 36, \( \alpha_Z = 5, \tau_Z = 14, \sigma_Z = 15 \). Moreover, \( \Delta^2 h_Z = (1, 1, 1, 1, 0, -1, -2, 0, 0, 0, -1, 0, 0, 0, -1) \), so \( \nu_Z(5) \leq 1, \nu_Z(6) \leq 1, \nu_Z(7) \leq 1, \nu_Z(11) \leq 1 \) and \( \nu_Z(15) \leq 1 \).
In fact, all of these are equalities. We have \( \nu_Z(5) = 1 \) since \( h_{I_Z}(5) = 1 \); note also by Bezout's Theorem that \( (I_Z)_5 \) has base locus \( A^3B^2 \) (i.e., every element \( F \in (I_Z)_5 \) is divisible by \( A^3B^2 \)).

Next, \( \nu_Z(6) \geq 1 \) (hence we have equality). To see this, note by Bezout the divisorial base locus is at least \( A^3B \). Factoring that we have \( (I_Z)_6 = (A^3B)(I(q_1 + q_2))_2 \). Since it is easy to see that \( (I(q_1 + q_2))_2 \) has no divisorial base locus, we conclude that the divisorial base locus of \( (I_Z)_6 \) is exactly \( A^3B \). Thus not every element of \( (I_Z)_6 \) is in \( M_1(I_Z)_5 \), hence \( \nu_Z(6) \geq 1 \).

And \( \nu_Z(7) \geq 2 \); here the divisorial base locus is by Bezout at least \( A^2 \). We have \( (I_Z)_7 = (A^3)(I(p_1 + \cdots + p_5 + 2q_1 + 2q_2))_5 \). It is easy to construct elements of \( (I(p_1 + \cdots + p_5 + 2q_1 + 2q_2))_5 \) with no common factors, so \( (I(p_1 + \cdots + p_5 + 2q_1 + 2q_2))_5 \) has no divisorial base locus. But it is not hard to find an element \( F \in (I(p_1 + \cdots + p_5 + 2q_1 + 2q_2))_5 \) whose divisorial base locus is \( B \) and another element \( G \) with no divisorial base locus. Clearly \( F \not\in M_1(I_Z)_6 \) and \( G \not\in M_1(I_Z)_6 + \langle F \rangle \), so \( \nu_Z(7) \geq 2 \) and hence equality holds again.

Similar arguments show that \( \nu_Z(11) \geq 1 \) and \( \nu_Z(15) \geq 1 \). Thus a minimal set of homogeneous generators for \( I_Z \) has \( 6 = \alpha_Z + 1 \) elements.

**Exercise 1.5.5.** In Example 1.5.4, write out the details for the argument that \( \nu_Z(11) = 1 \) and \( \nu_Z(15) = 1 \). Alternatively, use both the upper and lower bounds given in \([4]\) to determine \( \nu_Z(11) = 1 \) and \( \nu_Z(15) = 1 \), but show that these bounds give only \( 1 \leq \nu_Z(7) \leq 2 \).

**Exercise 1.5.6.** For colinear points \( p_1, \ldots, p_r \) and \( Z = m_1p_1 + \cdots + m_rp_r \), show that \( \tau_Z = -1 + \sum_i m_i \).

The underlying principle of the bounds is not complicated, but showing that the underlying principle leads to bounds \( f_v(t) \leq h_Z(t) \leq F_v(t) \) of the specific forms given above is rather involved. Thus we only explain the underlying principle. The idea is to give bounds on \( h_{I_Z} \), or equivalently on \( h^0(X, \mathcal{O}_X(F_Z(t))) \), for each \( t \), where, given \( Z = m_1p_1 + \cdots + m_rp_r \) and \( X \) the blow up of \( \mathbb{P}^2 \) at the points \( p_i \), we define \( F_Z(t) = tL - m_1E_1 - \cdots - m_rE_r \).

Given a curve \( C \subset X \), we have
\[
0 \to \mathcal{O}_X(F_Z(t) - C) \to \mathcal{O}_X(F_Z(t)) \to \mathcal{O}_C(C \cdot F_Z(t)) \to 0.
\]
If we have an upper bound \( B \geq h^0(X, \mathcal{O}_X(F_Z(t) - C)) \) and if we know \( h^0(C, \mathcal{O}_C(C \cdot F_Z(t))) \), then
\[
h^0(X, \mathcal{O}_X(F_Z(t) - C)) \leq B + h^0(C, \mathcal{O}_C(C \cdot F_Z(t))),
\]
with equality if and only if the \([4]\) is exact on global sections. In order to be sure we can determine \( h^0(C, \mathcal{O}_C(C \cdot F_Z(t))) \), we pick \( C \) to be the proper transform of a line. Since then \( C \) is a \( \mathbb{P}^1 \), we know the cohomology of any line bundle on \( C \) as long as we know the degree of the line bundle. If we have a lower bound \( b \leq h^0(X, \mathcal{O}_X(F_Z(t) - C)) \), then
\[
\min(b, h^0(X, \mathcal{O}_X(F_Z(t) - C)) - h^1(X, \mathcal{O}_X(F_Z(t) - C)) + h^0(C, \mathcal{O}_C(C \cdot F_Z(t)))) \leq h^0(X, \mathcal{O}_X(F_Z(t))),
\]
where equality holds if and only if \( h^0(X, \mathcal{O}_X(F_Z(t))) = 1 \). Note by Riemann-Roch we have \( h^0(X, \mathcal{O}_X(F_Z(t) - C)) - h^1(X, \mathcal{O}_X(F_Z(t) - C)) = \langle ((F_Z(t) - C))^2 - K_X \cdot (F_Z(t) - C) \rangle/2 + 1 \) so this is in practice easy to compute.

The idea now is to construct a sequence of such sheaf sequences where we always know \( h^0 \) of the right hand term and we eventually obtain a sequence where we know \( h^0 \) of the left hand term. We work through the details for Example 1.5.4. Here we want to compute \( h^0(X, \mathcal{O}_X(F_Z(t))) \),

\[
h^0(X, \mathcal{O}_X(tL - 3(E_1 + \cdots + E_5) - 2(E_6 + E_7)) = h^0(X, \mathcal{O}_X(tL - 3E - 2E')) \]

where for efficiency we write \( E \) for \( E_1 + \cdots + E_5 \) and \( E' \) for \( E_6 + E_7 \). We pick \( L_1 = \cdots = L_5 \) to be the line through \( p_1, \ldots, p_5 \) and \( C \) to be the proper transform of this line; thus \( [C] = [L - E] \). Then we pick \( L_6 = L_7 \) to be the line through \( p_6, p_7 \) and \( D \) to be the proper transform of this line, so \( [D] = [L - E'] \). We
now have a sequence of linked short exact sequences of sheaves:
\begin{align*}
0 & \to O_X((t-1)L - 2E - 2E') \to O_X(tL - 3E - 2E') \to O_C(t - 15) \to 0 \\
0 & \to O_X((t-2)L - E - 2E') \to O_X((t-1)L - 2E - 2E') \to O_C(t - 10) \to 0 \\
0 & \to O_X((t-3)L - 2E') \to O_X((t-2)L - E - 2E') \to O_C(t - 5) \to 0 \\
0 & \to O_X((t-4)L - E') \to O_X((t-3)L - 2E') \to O_P(t - 4) \to 0 \\
0 & \to O_X((t-5)L) \to O_X((t-4)L - E') \to O_P(t - 2) \to 0 
\end{align*}

Note that the middle sheaf in each sequence is the left sheaf in the sequence just above it. And we know \( h^0 \) for each right hand sheaf: it is \( \max(0, d + 1) \) where \( d \) is the degree. Also, we know \( h^0 \) for the left sheaf in the bottom row, since \( h^0(X, O_X((t-5)L)) = h^0(\mathbb{P}^2, O_{\mathbb{P}^2}(t - 5)) = \binom{t-5}{2} \) (where we set a binomial coefficient \( \binom{a}{b} \) to 0 if \( a < b \)). Thus \( h^0(X, O_X((t-4)L - E')) \leq h^0(X, O_X((t-5)L)) \) and hence of \( (5) \) is a subsheaf of the sheaf above it. And as explained above, each sheaf sequence in \( (5) \) is a subsheaf of the sheaf above it. And as explained above, each sheaf sequence in \( (5) \) is a subsheaf of the sheaf above it. And as explained above, each sheaf sequence in \( (5) \) is a subsheaf of the sheaf above it.

For the lower bound note that \( h^0(X, O_X((t-5)L)) \leq h^0(X, O_X((t-4)L - E')) \leq \cdots \leq h^0(X, O_X((t-1)L - 2E - 2E')) \leq h^0(X, O_X(tL - 3E - 2E')) \), because each sheaf on the left in \( (5) \) is a subsheaf of the sheaf above it. And as explained above, each sheaf sequence in \( (5) \) gives another lower bound for the dimension of the global sections of the middle term, and hence of \( h^0(X, O_X(tL - 3E - 2E')) \). For example, consider the middle sheaf sequence. We get \( h^0(X, O_X((t-3)L - 2E')) - h^1(X, O_X((t-3)L - 2E')) + h^0(C, O_C(t - 2)) \leq h^0(X, O_X((t-2)L - E - 2E')) \leq h^0(X, O_X(tL - 3E - 2E')). \) The maximum of these lower bounds is also a lower bound, and interprets to the bound \( h_Z(t) \leq \mu_v(t) \).

Under certain conditions (such as when the terms of \( v \) are strictly decreasing) the lower bound above equals the upper bound, so in these cases we get \( h_Z(t) = \mu_v(t) \).

1.6. Numbers of generator of ideals of fat point subschemes of \( \mathbb{P}^2 \) with strictly decreasing reduction vector. Consider a fat point subscheme \( Z = m_1p_1 + \cdots + m_sp_s \subset \mathbb{P}^2 \) and a sequence \( L_1, \ldots, L_n \) such that at least \( m_i \) lines pass through each point \( p_i \). Let \( Z = Z_0, \ldots, Z_s \) be the associated residual subschemes and \( v = (v_0, \ldots, v_{s-1}) \) the reduction vector. Assume that the entries \( v_i \) strictly decrease until they reach 0 (thus one or more of the entries at the end may be 0, but otherwise we have \( v_i > v_{i+1} \)).

Then as we saw above \( h_Z = \mu_v \) and hence \( h_Z(t) = \Delta h_Z(t) \) is the number of dots on the vertical line through \( t \) in the diagram in Exercise 1.5.1, from which it follows that \( \sigma_Z = v_0 \) (since \( v_0 \) is the first degree such that there are no dots on the vertical line for that degree in the diagram in Exercise 1.5.1). But in addition \( \beta_Z \) is determined, as is the divisorial base locus in each degree \( t \) and the number of generators \( \nu_Z(t) \) in each degree \( t \).

To see this, suppose \( v \) is the reduction vector for some \( Z \) and that the entries are strictly decreasing, such as is the case for example for \( v = (8, 6, 4, 3) \). In degree \( t \), the intersection of \( L_1 \) with \( Z \) has degree \( v_0 \), so if \( t < v_0 \), then \( L_0 \) is in the divisorial base locus of \( (I_Z)_t \). If \( v_1 > t - 1 \), then so is \( L_2 \), and likewise so is \( L_3 \) if \( v_2 > t - 2 \), etc. I.e., if we make a dot diagram as in Exercise 1.5.1 and as shown below (in this case for \( v = (8, 6, 4, 3) \)), the rows of dots, starting from the bottom going up, correspond the the lines \( L_1, L_2 \) etc. Each row with a dot on the vertical line at a given \( t \) corresponds to a line \( L_i \) in the divisorial base locus of \( (I_Z)_t \). So in the diagram below, we see the base locus contains \( L_1 \) and \( L_2 \). In particular, \( \beta_Z \) is the least \( t \) such that there are no dots on the vertical line for that \( t \); i.e., \( \beta_Z = \sigma_Z \).
The dot diagram for the residual scheme, in this case $Z_2$, is just the set of dots above the rows of dots for the lines $L_1$ and $L_2$ which get removed, but shifted down and left by 2 (to account for the removal of the two lines). Thus in the example $h_{L_2}(i) = h_{L_2}(i-2)$ for $i \leq 6$. More generally, if $L_1, \ldots, L_d$ are the lines forced as above to be in the base locus in some degree $t$, then we have $h_{L_2}(i) = h_{L_2}(i-d)$ for $i \leq t$. Also, the reduction vector for $Z_d$ is obtained from $v$ by truncation; i.e., it is $(v_d, \ldots, v_s)$. Note $t > v_j + j$ for $j \geq d$ since otherwise more of the lines would be in the base locus (recall only the first $d$ lines have dots on the vertical line through $t$). Thus $t - d > v_d = \sigma_{Z_d}$, so the base locus of $(I_{Z_d})_{t-d}$ is 0-dimensional and thus the base locus of $(I_Z)_t$ is just the lines $L_1, \ldots, L_d$ (taken with multiplicity). Moreover, we also have $\nu_{Z_d}(t-d +1) = 0$ hence $R_1(I_{Z_d})_{t-d} = (I_{Z_d})_{t-d +1}$ so $R_1(I_Z) = R_1(L_1 \cdots L_d)(I_{Z_d})_{t-d} = (L_1 \cdots L_d)(I_{Z_d})_{t-d +1}$, which means $\nu_Z(t+1) = h_{L_2}(t+1) - h_{L_2}(t-d +1)$.

**Exercise 1.6.1.** Apply \cite[Theorem 3.1.7]{6} (i.e., the methods of this section) to compute $\nu_Z(t)$ for all $t$ for Example 1.5.4 (Thus the bounds of \cite[Theorem 3.1.7]{6} are sometimes sharper than those of \cite{4}.)

**1.7. Characterizing the Hilbert functions of 0-dimensional subschemes of $\mathbb{P}^2$.** It is a theorem of Macaulay (see \cite{24, 28}) that a function $h : \mathbb{Z} \to \mathbb{Z}$ is the Hilbert function of $R/I$ for a homogeneous ideal $I \subseteq R = K[x_0, \ldots, x_n] = K[\mathbb{P}^n]$ for some $n$ if and only if it is what is called an O-sequence (where $O$ is the letter $O$ not the number 0). By \cite{15} (see also \cite[Proposition 6.3 and Theorem 6.4]{5}), $h = h_Z$ is the Hilbert function of $R/I$ for the homogeneous ideal $I \subseteq R = K[x_0, \ldots, x_n] = K[\mathbb{P}^n]$ of a 0-dimensional subscheme $Z$ if and only if $h$ is a differentiable 0-dimensional O-sequence.

Here is a version of this result for $n = 2$.

**Theorem 1.7.1.** Let $h : \mathbb{Z} \to \mathbb{Z}$ be a function and $\alpha$ a positive integer. The following are equivalent:

(a) $h = f_v$, where $v = (v_0, \ldots, v_{\alpha-1})$ for some $s \geq 0$ and $v_0 > v_2 > \ldots > v_{\alpha-1} > 0$;
(b) $h = h_Z$ for a 0-dimensional subscheme $Z \subseteq \mathbb{P}^2$ with $\alpha_Z = \alpha$; and
(c) $h(t) = 0$ for $t < 0$, $\Delta h(t) = t+1$ for $0 \leq t < \alpha$, $\Delta h(t) \leq \Delta h(t-1)$ for $t \geq \alpha$ and $\Delta h(t) = 0$ for $t \gg 0$.

**Proof.** We show (a) implies (b). Take $\alpha$ distinct linear forms $L_1, \ldots, L_\alpha$ with $v_{i-1}$ points on the line defined by $L_i$ and no points where two lines cross. Let $Z$ be the reduced union of the points. Then $v$ is the reduction vector and $h = h_Z$ by Theorem 1.5.2. Clearly $\alpha_Z \leq \alpha$, since the union of the $\alpha$ lines contains $Z$, and it is easy to check by Bezout that no nonzero form of degree $\alpha - 1$ can vanish on all of the points, so $\alpha_Z = \alpha$. (Note that $v_i \geq \alpha - i$. If $F_0 = F$ were a nonzero form of degree $\alpha - 1$ vanishing at all of the points, then since it vanishes at $v_0 \geq \alpha$ points on line $L_1$, we see $L_1$ is a factor of $F_0$. Let $F_1 = F_0/L_1$. Then $F_1$ vanishes at the $v_1 \geq \alpha - 1$ points on line $L_2$, hence $L_2$ divides $F_1$, etc. Thus $L_1 \cdots L_\alpha$ divides $F$ but this is impossible since deg $F = \alpha - 1$.)

The fact that (b) implies (c) follows from Lemma 1.2.1 and Proposition 1.2.6.

To see that (c) implies (a), consider the graph of $\Delta h$. Put a dot at each point $(i, j)$ such that $1 \leq i \leq h(i)$. Let $v_{i-1}$ be the number of dots on the horizontal line $L_i$ defined by $y = i$ for $1 \leq i \leq \alpha$. Note that the leftmost dot on line $L_{i+1}$ is one to the right of the left most dot on line $L_i$ (since $\Delta h(t) = t+1$ for $0 \leq t < \alpha$), but the rightmost dot on $L_{i+1}$ is not further right than the rightmost dot on $L_i$ (since a dot $(k, i+1)$ on $L_{i+1}$ means $h(k) \geq i+1$ and hence $(k, i)$ is one of
our dots), so \(v_{i-1} > v_i\). By Exercise 1.5.1 \(\Delta f_v(t)\) is the number of dots on the vertical line \(x = t\), which by construction of the dots is just \(\Delta h(t)\). Since \(f_v(t) = 0 = h(t)\) for \(t < 0\), \(h = f_v\) follows from \(\Delta f_v = \Delta h\).

**Exercise 1.7.2.** As in Exercise 1.3.13 suppose a 0-dimensional subscheme \(Z \subset \mathbb{P}^2\) has
\[
\Delta h_Z = (1, 2, 3, 4, 5, 6, 7, 5, 3, 2, 1, 0).
\]
Find some \(Z\) having this Hilbert function, and for this \(Z\) determine the curves in the base locus of \(I_Z\) in each degree and find the values of \(\beta_Z, \omega_Z\) and \(h_Z(t)\) for each \(t\) and for the total \(h_Z\).

**Exercise 1.7.3.** Show that \(h = (1, 3, 6, 10, 12, 15, 15, 15, \ldots)\) is not the Hilbert function of any 0-dimensional subscheme of \(\mathbb{P}^2\). Show however, that \(h = h_{R/I}\) is the Hilbert function of some homogeneous ideal \(I \subset R = K[\mathbb{P}^2]\), and if \(Z\) is the subscheme of \(\mathbb{P}^2\) defined by your ideal \(I\), determine \(h_Z\).

1.8. **Additional numerical characters.** Let \(p_1, \ldots, p_r\) be distinct points of \(\mathbb{P}^2\), and consider some subscheme \(Z = m_1 p_1 + \cdots + m_r p_r\). Let \(F_n = nL - m_1 E_1 - \cdots - m_r E_r\). We now define: the least degree \(\delta_Z\) such that \(F_n\) is numerically effective; the least degree \(\epsilon_Z\) such that \(F_n\) is effective and base point free; and the least degree \(\psi_Z\) such that the forms in \(I_Z\) of this degree generate the stalks of the sheaf of ideals \(I_Z\) (i.e., define \(Z\) scheme-theoretically).

Suppressing the subscript, it is clear that \(\delta \leq \beta \leq \epsilon \leq \psi \leq \omega \leq \sigma\).

**Exercise 1.8.1.** In the case that \(p_1, \ldots, p_r\) lie on a line, show that \(\sigma = \delta\), and therefore \(\delta = \beta = \epsilon = \psi = \omega\). In the case that \(p_1, \ldots, p_r\) lie on a conic, show that \(\sigma = \delta\) can fail.

In the case that \(p_1, \ldots, p_r\) lie on a conic, we nonetheless have \(\delta = \beta = \epsilon = \psi = \omega\) by [20, Theorem III.1.2]. The fact that \(\delta = \omega\) for \(p_1, \ldots, p_r\) on a curve of degree 2 or less, together with our being able to compute fixed components for effective classes and \(h^0\) for numerically effective classes, provides a way to determine the graded Betti numbers in a minimal homogeneous free resolution of \(I_Z\). It also means for points on a conic that if the forms in \(I_Z\) of degree \(n\) define \(Z\) scheme theoretically, then they actually generate \(I_Z\).

**Exercise 1.8.2.** Determine the graded Betti numbers for \(Z = 3p_1 + \cdots + 3p_5 + 2q_1 + \cdots + 2q_2\) where the \(p_i\) lie on one line and the \(q_j\) lie on another, and none of them are on both lines. Assume that the minimal free resolution of a ideal of fat points \(I_Z\) has the form \(0 \to F_1 \to F_0 \to I_Z \to 0\), where \(F_1\) and \(F_0\) are graded free modules. (I.e., there are no second syzygies since \(R/I_Z\) is Cohen-Macaulay.)

In case the points \(p_1, \ldots, p_r\) lie on the curve of degree 3, things are more complicated. Here, 
\[-K_X \cdot F_{\delta+1} \geq 3, \]
where 
\[-K_X = 3L - E_1 - \cdots - E_r\]
is the canonical divisor, so by [22] we have only \(\tau \leq \delta + 1\) (the example that \(p_1, \ldots, p_r\) are the base points of a pencil of cubics shows that this inequality is best possible). Moreover, \(\omega\) can be bigger than \(\delta\), and bigger even than \(\epsilon\). Consider the case that \(Z\) is the set of 9 points which are the base points of a pencil of cubics. Then \(\omega = 4\) but \(\psi = \delta = \epsilon = 3\). The fact that \(\omega\) can exceed \(\epsilon\) means that finding minimal free resolutions is more delicate than just determining base points, fixed components and \(h^0\)’s.

We now give a criterion for \(\psi\) in the case that \(p_1, \ldots, p_r\) are distinct points of \(\mathbb{P}^2\).

**Proposition 1.8.3.** Let \(L, E_1, \ldots, E_r\) be the divisors corresponding to the blowing up \(X\) of distinct points \(p_1, \ldots, p_r\) of \(\mathbb{P}^2\). For nonnegative integers \(m_i\) consider the fat point subscheme \(Z = m_1 p_1 + \cdots + m_r p_r\); then the following are equivalent:

(a) \(\psi_Z \leq n\);
(b) \(\epsilon \leq n\) and the restriction homomorphisms \(H^0(X, O_X(F_n)) \to H^0(E_i, O_X(F_n) \otimes O_{E_i})\) are surjective for every \(i > 0\); and
(c) the restriction homomorphisms \(H^0(X, O_X(F_n)) \to H^0(E_i, O_X(F_n) \otimes O_{E_i})\) are surjective for every \(i > 0\) and \(|F_n|\) has no base points off \(E_1 + \cdots + E_r\).
Proof. Note that (b) implies (c) by definition of $\epsilon$ while (c) implies (b) since surjectivity of the restriction homomorphisms imply there are no base points on any $E_i$. So now it suffices to show (a) and (c) are equivalent.

If $p \in \mathbb{P}^2$ but $p \not\in \{p_1, \ldots, p_i\}$, the stalk $\mathcal{I}_{Z,p}$ of $\mathcal{I}_Z$ at $p$ is just the local ring $\mathcal{O}_{\mathbb{P}^2,p_i}$, and a set of forms generate $\mathcal{O}_{\mathbb{P}^2,p_i}$ if and only if the forms do not all vanish there.

If $p \in \{p_1, \ldots, p_i\}$ (say $p = p_i$) the stalk of $\mathcal{I}_Z$ at $p$ is precisely $P_i^{m_i}$, where $P_i$ is the maximal ideal in the local ring $\mathcal{O}_{\mathbb{P}^2,p_i}$. By Nakayama’s lemma, a set of forms generates $P_i^{m_i}$ if and only if their images span $P_i^{m_i}/P_i^{1+m_i}$. But with respect to the identification $K[\mathbb{P}^2] = \bigoplus_{n \geq 0} H^0(X, \mathcal{O}_X(nE))$, the obvious map $H^0(\mathbb{P}^2, \mathcal{I}_Z(n)) \to P_i^{m_i}/P_i^{1+m_i}$ corresponds to $H^0(X, \mathcal{O}_X(F_n)) \to H^0(E_i, \mathcal{O}_X(F_n) \otimes \mathcal{O}_{E_i})$, keeping in mind that $H^0(E_i, \mathcal{O}_X(F_n) \otimes \mathcal{O}_{E_i}) = H^0(E_i, \mathcal{O}_{E_i}(m_i))$. Thus the forms $H^0(\mathbb{P}^2, \mathcal{I}_Z(n))$ generate $P_i^{m_i}$ if and only if $H^0(X, \mathcal{O}_X(F_n) \to H^0(E_i, \mathcal{O}_X(F_n) \otimes \mathcal{O}_{E_i})$ is onto. □

Example 1.8.4. Let $Z = 2p_1 + \cdots + 2p_6$ where the $p_i$ are six general points of $\mathbb{P}^2$; clearly, they lie on a cubic. It is not hard to show that $|4L - 2E_1 - \cdots - 2E_6|$ is empty, that $|5L - 2E_1 - \cdots - 2E_6|$ is non-empty and base point free and $\sigma_Z \leq 6$, and that $\nu_Z(6) \neq 0$. Thus we have $\epsilon = 5$ and $\omega = 6$. Since $5L - 2E_1 - \cdots - 2E_6$ is nef but $(5L - 2E_1 - \cdots - 2E_6) \cdot (5L - 2E_1 - \cdots - 2E_6 - E_i) = -1$ for $0 < i \leq 6$, we see that $h^0(X, \mathcal{O}_X(5L - 2E_1 - \cdots - 2E_6 - E_i)) = 0$. By Riemann-Roch, it follows that $h^1(X, \mathcal{O}_X(5L - 2E_1 - \cdots - 2E_6 - E_i)) = 0$, and thus that $H^0(X, \mathcal{O}_X(5L - 2E_1 - \cdots - 2E_6)) \to H^0(E_i, \mathcal{O}_X(5L - 2E_1 - \cdots - 2E_6) \otimes \mathcal{O}_{E_i})$ is surjective. Now from Proposition 1.8.3 we see that $\psi = 5$.

Example 1.8.5. We also recover the following example of Ein (see [11]). If $p_1, \ldots, p_{18}$ are general points of $\mathbb{P}^2$, with $X$ and $L, E_1, \ldots, E_{18}$ as usual, let $Z = p_1 + \cdots + p_{18}$; then $\tau = 5$ (since the conditions imposed by $r$ general points are independent on forms of degree $n$ if $r \leq \binom{n+2}{2}$), but $5 < \omega$ (since $9 = h^0(\mathbb{P}^2, \mathcal{O}(1)) h^0(\mathbb{P}^2, \mathcal{I}_Z(5)) < h^0(\mathbb{P}^2, \mathcal{I}_Z(6)) = 10$). Since $\omega \leq \tau + 1$, we see $\omega = 6$. Since a general quintic through 18 general points is smooth, of genus 3, we see that $5L - E_1 - \cdots - E_{18}$ is linearly equivalent to a smooth irreducible curve $C$ of genus $g = 3$. Since the sections of $\mathcal{O}_X(C)$ surject onto those of $\mathcal{O}_X(C) \otimes \mathcal{O}_C$, and the latter, having degree $2g + 1$, is very ample, we see $C$ is base point free, and, because $p_1, \ldots, p_{18}$ are general, it is easy to see that $h^1(X, \mathcal{O}_X(5L - E_1 - \cdots - E_{18} - E_i)) = 0$ for each $i > 0$. Thus, as in Example 1.8.4 we see that $\psi = 5$.

1.9. Historical Discussion. Most of the results presented here are well-known; see for example [5, 9, 23] and papers referenced in these papers. However, section 1.8 has not appeared in print before. The problem of characterizing Hilbert functions of ideals goes back to [24]; see also [28]. For the case of ideals of points see [12, 14, 15, 16]. The bounds $\nu_Z(t) \leq \alpha_Z + \beta_Z - \sigma_Z + 1 \leq \alpha + 1$ in Proposition 1.2.7 and Corollary 1.3.9 are due to Dubreil [10]. For Proposition 1.2.7 (4), see, for example, [9, Proposition 3.7]. See also [4, 9, 26] for exposition, refinements and extensions. The other results in Proposition 1.2.7 are special cases of results of Campanella [1]; see also [11], and, for refinements for fat points in $\mathbb{P}^2$, see [19]. Exercise 1.3.6 is a special case of a theorem of Macaulay [24], stating that for every homogeneous ideal in a polynomial ring, there is a monomial ideal with the same Hilbert function. See [9] and [23, Appendix C] for a different approach to Lemma 1.3.8 and Theorem 1.3.5. Lemma 1.3.6 is a special case of Gotzmann’s Persistence Theorem; see [17, 19, 21, 18]. I suspect that most if not all of the results in these notes through Corollary 1.3.9 were known to Dubreil, in some form at least. Corollary 1.3.11 is due to Davis; see [7, 8]. It has been greatly extended; see [2] and [25]. The now standard use of line bundles on blow ups of the plane to study fat points goes back to [21]. See [20] Lemma II.8 for a slightly more general version of Proposition 1.4.5. The results of section 1.5 are taken from [6] which in turn was motivated by [16, Theorem 1.7.1] showing how increasing numbers of points on a sequence of lines is related to Hilbert functions, helps to explain the motivation behind the definition of $k$-configurations. These were first introduced with that name in [13] but were implicit already in [15]; see also [27].
These notes benefitted from very helpful consultations with Susan Cooper and Juan Migliore.

REFERENCES

2. Topic 2: Aspects of Negativity on $\mathbb{P}^2$

Speaker: Brian Harbourne  
Venue: Krakow, Poland, May, 2015  
Date of Draft: May 18, 2015

2.1. Intersection Theory. Let $K$ be an algebraically closed field of arbitrary characteristic.

Question 2.1.1. Given a reduced curve $C \subset \mathbb{P}^2$, the general question is: how singular can $C$ be?

One way to measure this is with the so called excess intersection.

Definition 2.1.2. Let $C$ be a curve defined by a square free homogeneous polynomial $F \in R = K[x, y, z] = K[\mathbb{P}^2]$. Then the multiplicity of $F$ at $p \in \mathbb{P}^2$, denoted $\text{mult}_p(C)$ is the largest $m$ such that $F \in I(p)^m$.

Example 2.1.3. The multiplicity of $F = x^3y^4 + x^5z^2$ at $p = [0 : 0 : 1]$ is 5.

Definition 2.1.4. Given two curves $C$ and $D$ defined by nonconstant forms $F$ and $G$ with no common factors, we define the intersection multiplicity of $C$ and $D$ at $p$ by $I_p(C, D) = \dim_K(R/J)_t$ for $t \gg 0$, where $J = I(p)^m + (F, G)$ for any $m \geq \deg(F)\deg(G)$.

Theorem 2.1.5 (Bezout’s Theorem). Let $C$ and $D$ be curves defined by nonconstant forms $F$ and $G$ with no common factors. Then $\sum_p I_p(C, D) = \deg(C)\deg(D)$. Moreover, $I_p(C, D) \geq \text{mult}_p(C)\text{mult}_p(D)$ for each point $p \in \mathbb{P}^2$.

Definition 2.1.6. Given curves $C$ and $D$ and a finite set $S \subset \mathbb{P}^2$, define the excess intersection to be $E(C, D; S) = \deg(C)\deg(D) - \sum_{p \in S}\text{mult}_p(C)\text{mult}_p(D)$ and the excess self-intersection to be $E(C; S) = E(C, D; S)$.

We now have two Open Questions:

Question 2.1.7 (Bounded Negativity 1). For each $S$, is there a $b_S > 0$ such that $E(C; S) \geq -B_S$ for all reduced $C$?

Question 2.1.8 (Bounded Negativity 2). For each $S$, is there a $b_S > 0$ such that $E(C; S) \geq -b_S$ for all reduced and irreducible $C$?

Theorem 2.1.9. These two problems are equivalent.

We will do the proof later, using Zariski decompositions.

If $S$ is a generic set of $s \geq 3$ points, then it is conjectured (as a consequence of what is known as the SHGH Conjecture) that $b_S = 1$ and $B_S = s$.

Exercise 2.1.10. Let $S$ be 6 generic points of the plane. Let $D$ be the union of the 6 conics through subsets of 5 of these 6 points. Show that $E(D; S) = -6$.

2.2. An application.

Question 2.2.1. If $C$ is reduced, how negative can $E(C; S)$ be?

Definition 2.2.2. To measure negativity (or singularity) of a singular curve $C$ per point, define $H(C; S) = \frac{E(C; S)}{|S|}$ and $H(C)$ to be $H(C; S)$, where $S$ is the singular set of $C$.

Exercise 2.2.3. If $C$ is reduced and singular with $H(C) < -1$, show that $H(C; S)$ is least for some $S$ contained in the singular set of $C$, and $H(C; S)$ is least when taking $S$ to be the whole singular set if $H(C) > -4$. (Since no characteristic 0 examples are known with $H(C) \leq -4$, the current research focus is on $H(C)$ rather than on $H(C; S)$ for some subset $S$ of the singular locus.)

A significant open question is:
Question 2.2.4. What is $\inf \{ H(C) : C \text{ reduced and singular} \}$?

Exercise 2.2.5. Show that $\inf \{ H(C) : C \text{ union of distinct lines} \} = -\infty$ when $K$ is algebraically closed of positive characteristic. Thus this question as stated is of interest only in characteristic 0, but perhaps a version can be formulated which is of interest also in positive characteristics.

Another significant open question is:

Question 2.2.6. What is $\inf \{ H(C) : C \text{ reduced, irreducible and singular} \}$?

Example 2.2.7. We know $-2$ is an upper bound when $K$ is algebraically closed of any characteristic. No example of an irreducible $C$ is known with $H(C) \leq -2$.

Example 2.2.8. Over $\mathbb{R}$ we have $-3 = \inf \{ H(C) : C \text{ union of distinct lines} \}$, hence $-3 \geq \inf \{ H(C) : C \text{ reduced and singular} \}$.

Exercise 2.2.9. Over $\mathbb{Q}$ we thus have $-3 \leq \inf \{ H(C) : C \text{ union of distinct lines} \}$, but the value of $\inf \{ H(C) : C \text{ union of distinct lines} \}$ over $\mathbb{Q}$ is not known. How negative an example can you find for $H(C)$ for a curve defined over $\mathbb{Q}$?

Example 2.2.10. Over $\mathbb{C}$ we have $-4 < \inf \{ H(C) : C \text{ reduced and singular} \} \leq -225/67$.

2.3. Another formulation of the bounded negativity problem. Let $X \to \mathbb{P}^2$ be the blow up of a finite set $S$ of points $p_1, \ldots, p_r$. Let $E_1, \ldots, E_r$ be the corresponding exceptional curves and $L$ the pullback of a general line.

We can express $E(C; S)$ and $E(C)$ in terms of $L, E_1, \ldots, E_r$. In fact, $E(C; S) = F^2$, where $F = tL - m_1E_1 - \cdots - m_rE_r$ and $d = \deg(C)$ and $m_i = \text{mult}_{p_i}(C)$. Moreover, if $C'$ is the proper transform of $C$ to $X$, then $C' \in |F|$ so $(C')^2 = E(C; S)$.

Exercise 2.3.1. Let $S$ consist of one point. Show that $E(C; S) \geq 0$ for every reduced plane curve $C$. Let $X$ be the blow up of the plane at that one point. Show that $\inf_D \{ D^2 \} = -1$, where the infimum is taken over all reduced curves $D \subset X$.

Exercise 2.3.2. Let $X$ be the blow up of the plane at a finite set $S$ of points. Show that $\inf_C \{ C^2 \}$ is bounded below for all reduced $C$ on $X$ (resp., prime divisors $C$ on $X$) if and only if $\inf_D \{ E(D; S) \}$ is bounded below for all reduced plane curves $D$ (resp., reduced and irreducible plane curves $D$).

Remark 2.3.3. So the Bounded Negativity Conjectures as stated above are equivalent to boundedness of self-intersections of reduced curves (resp. prime divisors) on the blow up $X$.

2.4. Zariski Decompositions. Zariski decompositions were first proved for effective divisors [19] on any smooth projective surface $X$. A more general version can be found in [11]. Here we prove them for any effective divisor $D$ on a blow up $X$ of the plane. It is not hard to see that it actually is enough to assume $tD$ is linearly equivalent to an effective divisor for some $t > 0$.

Theorem 2.4.1. Let $X$ be the blow up of a finite set of points of the plane. If $D = m_1N_1 + \cdots + m_rN_r$ where the $m_i$ are positive integers and each $N_i$ is a reduced irreducible curve on $X$, then we can write $D = P + N$ where $P = a_1N_1 + \cdots + a_rN_r$ is nef, the $a_i$ are nonnegative, $PN = 0$ and either $N = 0$ or $N = b_{i_1}N_{i_1} + \cdots + b_{i_s}N_{i_s}$ with the $b_{i_j}$ positive and the matrix $(N_{i_j}N_{i_k})$ negative definite. Moreover, if $D'$ is effective with Zariski decomposition $P' + N'$, and linearly equivalent to $D$, then $P'$ and $P$ are linearly equivalent and $N' = N$.

Exercise 2.4.2. Let $X$ be the blow up of $r$ points of the plane. If $N_1, \ldots, N_s$ are prime divisors such that the matrix $(N_iN_j)$ is negative definite, prove that the $N_i$ are linearly independent in the divisor class group. Conclude that $s \leq r$.

Exercise 2.4.3. Let $X$ be the blow up of the $r = \binom{d}{2}$ points of intersection of $d = 2t$ general lines in the plane for $t > 1$, and let $D = tL - E_1 - \cdots - E_r$. Show that $\{2D\}$ is nonempty, but $\{D\}$ is empty.
Exercise 2.4.4. In Exercise 2.4.3 assume \( d = 6 \). Let \( D \) be the sum of the proper transforms of the 6 lines (so up to linear equivalence \( D \sim 6L - 2E_1 - \cdots - 2E_{15} \)). Let \( L \) be the proper transform of a general line. Find a Zariski decomposition for each of the following divisors:

\[
D_3 = D/2 \sim 3L - E_1 - \cdots - E_{15}, \\
D_4 = L + D/2 \sim 4L - E_1 - \cdots - E_{15}, \\
D_5 = 2L + D/2 \sim 5L - E_1 - \cdots - E_{15}, \\
D_6 = D \sim 6L - 2E_1 - \cdots - 2E_{15}, \\
D_7 = 4L + D \sim 10L - 2E_1 - \cdots - 2E_{15}.
\]

Zariski decompositions provide an easy way to prove Theorem 2.1.9.

Proof of Theorem 2.1.9 By Remark 2.3.3 it’s enough to work on the blow up \( X \) of the points in \( S \). Certainly, if self-intersections of irreducible curves are bounded below, then so are the self-intersections of irreducible curves on \( X \). Conversely, let \( D \) be any reduced effective divisor. Write \( D = C_1 + \cdots + C_r \) where the \( C_i \) are distinct reduced, irreducible curves. Let \( D = P + N \) be a Zariski decomposition where \( N = n_1N_1 + \cdots + n_sN_s \) with \( 0 \leq n_i \leq 1 \) rational for all \( i \) and the \( N_i \) prime divisors of negative self-intersection. Then \( D^2 \geq N^2 \geq \sum (n_1N_1)^2 + \cdots + (n_sN_s)^2 \geq N_1^2 + \cdots + N_s^2 \geq (\rho(X) - 1) \min_i \{N_i^2\} \), where \( \rho(X) = |S| + 1 \) is the Picard number of \( X \) (i.e., the rank of the divisor class group). The last inequality comes from Exercise 2.4.2.

For the proof we will use a lemma and some exercises.

Exercise 2.4.5. Let \( N_1, \ldots, N_r \) be distinct reduced irreducible curves with \( N_i^2 < 0 \) for all \( i \) such that no nonzero nonnegative sum \( m_1N_1 + \cdots + m_rN_r \) is nef. Then there is an orthogonal basis \( N_1^*, \ldots, N_r^* \) where \( N_i^* = N_i \), \( N_2^* = c_2N_1^* + N_2 \), \( N_3^* = c_3N_1^* + c_2N_2^* + N_3 \), \ldots, \( N_r^* = c_{r1}N_1^* + c_{r2}N_2^* + \cdots + c_{r,r-1}N_{r-1}^* + N_r \) with each \( c_{ij} \) rational and \( c_{ij} \geq 0 \) (so each \( N_i^* \) is a nonnegative rational linear combination of the \( N_j \)) and \( (N_i^*)^2 < 0 \) for each \( i \).

Lemma 2.4.6. Let \( N_1, \ldots, N_r \) be reduced irreducible curves. Then the matrix \((N_iN_j)\) is negative definite if and only if no nonzero nonnegative sum \( m_1N_1 + \cdots + m_rN_r \) is nef.

Proof. Assume the matrix \((N_iN_j)\) is negative definite. Thus for any nonzero nonnegative linear combination \( N = m_1N_1 + \cdots + m_rN_r \), we have \( N^2 < 0 \) and hence \( N \) is not nef.

Conversely, assume no nonzero nonnegative sum \( m_1N_1 + \cdots + m_rN_r \) is nef. Thus \( N_i^2 < 0 \) for all \( i \). Now apply Exercise 2.4.5. Thus the span of \( N_1, \ldots, N_r \) has an orthogonal basis where each basis element has negative self-intersection, hence \((N_iN_j)\) is negative definite. \( \square \)

Exercise 2.4.7. Let \( V \) be a finite dimensional vector space with a positive definite inner product. Let \( v_1, \ldots, v_r \) be a basis such that \( v_i \neq v_j \) for all \( i \neq j \). If \( v \in V \) has \( \langle v, v \rangle \geq 0 \) for all \( i \), show that \( v = a_1v_1 + \cdots + a_nv_r \) where \( a_i \geq 0 \) for all \( i \).

Corollary 2.4.8. Let \( N_1, \ldots, N_r \) be reduced irreducible curves with \( N_i^2 < 0 \) for all \( i \) such that no nonzero nonnegative sum \( m_1N_1 + \cdots + m_rN_r \) is nef. Then there is a dual basis \( N_1', \ldots, N_r' \) where:

\[
N_i'N_j = 0 \quad \text{for all } i \neq j; \quad N_i'N_i = (N_i')^2 < 0 \quad \text{for all } i; \quad \text{and each } N_i' \text{ is a nonnegative rational linear combination of the } N_j.
\]

Proof. By Lemma 2.4.6 the intersection form on the span of \( N_1, \ldots, N_r \) is negative definite. The dual basis elements \( N_i' \) are solutions to the linear equations \( N_i'N_j = 0 \), which are defined over the integers, so the solutions are rational linear combination of the \( N_j \). Negative definiteness gives \((N_i')^2 < 0 \), and \( N_i'N_i = (N_i')^2 \) comes down to a choice of scaling. The fact that each \( N_i' \) is a nonnegative rational linear combination of the \( N_j \) comes from Exercise 2.4.7 (after converting the result to the negative definite case). \( \square \)

Proof of Theorem 2.4.1. Start with \( D = M + N \), where \( M = 0 \) and \( N = m_1N_1 + \cdots + m_rN_r \). If some nonzero nonnegative sum \( S = n_1N_1 + \cdots + n_rN_r \) is nef, let \( c \) be the minimum of the ratios \( m_i/n_i \) for which \( n_i > 0 \). Replace \( M \) by \( M + cS \) and replace \( N \) by \( N - cS \). Then \( D = M + N \) and \( M \) and \( N \) are still nonnegative sums of the \( N_i \), with \( M \) still nef but \( N \) having one fewer summand.
Repeat this process until either \( N = 0 \) or \( N \) is a sum \( N = b_i N_i + \cdots + b_j N_j \) such that \( b_j > 0 \) for all \( j \) but no nonnegative sum of the \( N_j \) is nef.

Thus we have \( D = M + N \) where \( M \) is a nef nonnegative rational sum of the curves \( N_i \), and \( N \) is either 0 (and we are done) or a positive rational sum \( N = b_i N_i + \cdots + b_j N_j \) where no nonzero nonnegative sum of the \( N_i \) is nef.

In the latter case, if \( MN_j = 0 \) for all \( i \) we take \( P = M \) and \( N \) as is, and we are done. So suppose \( MN_j > 0 \) for some \( i \). Consider the dual basis \( \{ N'_j \} \) given in Corollary 2.4.8. We can write \( N'_j = \sum_a a_{ij} N_i \) with nonnegative \( a_{ij} \). Choose the maximum \( t \) such that \( ta_{ij} \leq b_{ij} \) for all \( j \) and such that \( (M + tN'_j)N_j \geq 0 \), and replace \( M \) by \( M + tN'_j \) and \( N \) by \( N - tN'_j \). Then either the number of basis elements \( N_j \) meeting \( M \) positively has gone down by 1 of the number of terms in \( N \) has gone down by 1. Repeating this process eventually gives a \( P = M \) orthogonal to all terms (if any) of \( N \).

Moreover, if \( D' \) is effective with Zariski decomposition \( P' + N' \), and linearly equivalent to \( D \), then \( P' \) and \( P \) are linearly equivalent and \( N' = N \).

For the uniqueness assertion, pick an integer \( t > 0 \) such that \( tP, tP', tN \) and \( tN' \) are all integral. Then some component \( C_1 \) of \( tN \) has \( C_1 \cdot tN < 0 \), so \( C_1 \cdot tN' < C_1 \cdot tD' = C_1 \cdot tD = C_1 \cdot tN < 0 \). Thus \( C_1 \) is a component of \( tN' \). If \( tN \neq C_1 \), then for some component \( C_2 \) of \( tN - C_1 \), by negative definiteness we have \( C_2 \cdot (tN' - C_1) \leq C_2 \cdot (tN - C_1) < 0 \). Repeating this, we eventually see that \( tN' - tN \) is effective. Reversing the argument shows that \( tN' - tN \) is also effective, so \( tN' = tN \). Thus \( tP = tD' - tN' \) is linearly equivalent to \( tP = tD - tN \), as claimed. \( \square \)

2.5. Another formulation of bounded negativity. Let \( X \) be the blow up of the plane at a finite set of points \( S \). We say that \( X \) has bounded Zariski denominators if there is an integer \( d \) such that for each divisor \( D \) and integer \( t > 0 \) such that \( t \) is linearly equivalent to an effective divisor, there is an integer \( 0 \leq e \leq d \) such that the Zariski decomposition \( tD = P + N \) has integral divisors \( P \) and \( N \).

We now state a version of the main theorem of \([1]\).

**Theorem 2.5.1.** Let \( X \) be the blow up of the plane at a finite set of points \( S \). Then bounded negativity holds on \( X \) (i.e., the set of self-intersections \( C^2 \) of reduced curves on \( X \) is bounded below) if and only if \( X \) has bounded Zariski denominators.

An exercise will be helpful.

**Exercise 2.5.2.** Let \( X \) be a blow up of the plane at \( s \) points. Let \( C = dL - m_1E_1 - \cdots - m_nE_n \) be any divisor with \( C^2 < 0 \) and let the gcd of \( d, m_1, \ldots, m_n \) be \( g \). Then there is an ample divisor \( F \) such that \( FC \) and \( C^2 \) have gcd \( g \).

**Proof of Theorem 2.5.1.** Assume bounded negativity holds on \( X \); i.e., \( C^2 \geq -b \) for some \( b \) and every irreducible curve \( C \). Let \( D = d_1D_1 + \cdots + d_rD_r \) be effective (so each \( d_i \) is positive and each \( D_i \) is a prime divisor) with Zariski decomposition \( D = P + N \). Then \( P \) and \( N \) are sums of the \( D_i \) with nonnegative rational coefficients. Since \( D \) is integral, the largest denominator used for \( P \) is also the largest denominator used for \( N \), so it’s enough to look at \( N \). Say \( N = n_1N_1 + \cdots + n_sN_s \) where each \( n_i \) is positive rational and each \( N_i \) is a prime divisor of negative self-intersection. Note that \( DN_i = (n_1N_1 + \cdots + n_sN_s)N_i \) gives linear equations for the \( n_i \). The solution involves dividing by \( \text{det}(N_iN_j) \), so the largest possible denominator is \( \text{det}(N_iN_j) \), but \( |\text{det}(N_iN_j)| \leq \sqrt{|N_1^2 \cdots N_s^2|} \) (since the volume of a parallelepiped with edges of fixed length is most when the edges are orthogonal). By Exercise 2.4.2 we have \( s \leq |S| \). Thus the largest possible denominator is \( \sqrt{|b|^{19}} \), where \( b \) is a lower bound for self-intersections of irreducible curves on \( X \).

Conversely, assume \( X \) has bounded Zariski denominators, with bound \( b \). Let \( C \sim dL - m_1E_1 - \cdots - m_rE_r \) be any prime divisor with \( C^2 < 0 \). Let \( C = gD \) where \( g \) is the gcd of \( d, m_1, \ldots, m_r \). Now say \( C \) is a prime divisor with \( C^2 < 0 \), and pick \( D \) to be primitive (i.e., not linearly equivalent to \( tD' \))
for any integral divisor $D'$ with $t$ an integer bigger than 1) such that $C = gD$. By Exercise 2.5.2 we can pick an ample divisor $F$ such that $FC = g$. Since the Zariski decomposition of $D$ is $D = C/g$, we have $g \leq b$. But for large $m$, the Zariski decomposition of $F + mC$ is $P = F + (m-a)C$ and $N = aC$ for some $a$, so $a = (CF + mC^2)/C^2$, hence the denominator needed here is $C^2/gcd(CF,C^2) \leq b$, hence $C^2 \leq bgcd(CF,C^2) = bg \leq b^2$.

Exercise 2.5.3. Let $X$ be the blow up of the plane at a finite number of points such that there is a finite list $A = \{a_1, \ldots, a_r\}$ such that for every prime divisor $D$ with $D^2 < -1$ we have $D^2 \in A$. Assume that there are at most $n_i$ distinct divisors $D$ with $D^2 = a_i$ for each $i$ with $a_i < -1$. Show that no denominator bigger than $|a_1^{n_1} \cdots a_r^{n_r}|$ is ever needed for a Zariski decomposition on $X$.

Exercise 2.5.4. Determine the largest denominator needed for a Zariski decomposition when $X$ is the blow up of $r$ colinear points of the plane.

Exercise 2.5.5. Determine the largest denominator needed for a Zariski decomposition when $X$ is the blow up of $r+1$ points of the plane on a line $L_1$ and $s+1$ points on a different line $L_2$, where one of the points is the point of intersection of the two lines. Assume $r-1$ and $s-1$ are coprime and each is at least 2. [Hint: By “adjunction”, we have $C^2 \geq -2 + C(3L - E_1 - \cdots - Er + s + 1)$; see [1] Example 2.3].]

2.6. Waldschmidt constants. Let $p_1, \ldots, p_r$ be distinct points in the plane and let $I$ be the radical ideal of the points. Define the $m$th symbolic power of $I$ to be $I^{(m)} = I(p_1)^m \cap \cdots \cap I(p_r)^m$ and let $\alpha(I^{(m)})$ be the least degree $t$ such that $I^{(m)}$ has a nonzero form of degree $t$.

Definition 2.6.1. The Waldschmidt constant of the ideal of the points is defined to be the limit

$$
\lim_{m \to \infty} \frac{\alpha(I^{(m)})}{m}
$$

which we denote by $\tilde{\alpha}(I)$.

Exercise 2.6.2. Show that $\frac{\alpha(I^{(m+n)})}{m+n} \leq \frac{\alpha(I^{(m)})}{m} \leq \frac{\alpha(I^{(mn)})}{m}$ for all $m, n \geq 1$.

Exercise 2.6.3. Let $X$ be the surface obtained by blowing up points $p_1, \ldots, p_r$. Let $I$ be the ideal of the points and let $F_{t,m} = tL - m(E_1 + \cdots + E_r)$. Show that

$$
\tilde{\alpha}(I) = \inf \left\{ \frac{t}{m} : h^0(X, \mathcal{O}_X(F_{t,m})) > 0 \right\}
$$

Computing $\tilde{\alpha}(I)$ can be difficult, but is related to computing Zariski decompositions.

Proposition 2.6.4. Let $p_1, \ldots, p_r$ be distinct points in the plane and let $I$ be the radical ideal of the points. Let $X$ be the surface obtained by blowing up the points and let $F = dL - m_1E_1 - \cdots - m_rE_r$. If $F$ has a Zariski decomposition of the form $P + N$ where $P \neq 0$ and $N = aL - b(E_1 + \cdots + E_r) \neq 0$, then $\tilde{\alpha}(I) = \frac{a}{b}$.

Proof. Since $N$ is effective we have $\tilde{\alpha}(I) \leq \frac{a}{b}$. Let $E = E_1 + \cdots + E_r$. Since $P$ is nef, we have $(PL)b(\tilde{\alpha}(I)) - bPE = P(b(\tilde{\alpha}(I))L - bE) \geq 0 = PN = (PL)a - bPE$, so $b(\tilde{\alpha}(I)) \geq a$ or $\tilde{\alpha}(I) \geq a/b$.

Exercise 2.6.5. Compute $\tilde{\alpha}(I)$ for the ideal $I$ of the points of intersection of the lines in the following figure.
Exercise 2.6.6. Compute $\tilde{\alpha}(I)$ for the ideal $I$ of the points of intersection of the lines in the following figure.

Exercise 2.6.7. Compute $\tilde{\alpha}(I)$ for the ideal $I$ of the points of intersection of $d > 2$ general lines in the plane.

2.7. Historical discussion. The Bounded Negativity Conjecture extends to all smooth projective algebraic surfaces, but in this generality it is known not always to hold in positive characteristic (see for example [15, Exercise V.1.10]), although no counterexamples are known for a rational surface. It is not known who to attribute the conjecture to, but it goes back a long time (see the discussion in the introduction of [4]). It was in order to have a more accessible problem that might give insight to Bounded Negativity that $H$-constants were introduced (at a workshop at Oberwolfach in 2010) and discussed at several workshops since. The first occurrence in the literature seems to be in the extended abstracts by Pokora and by Harbourne in [3], followed by [2]. They have been referred to under a number of different names (see [2, 16, 17]) but always denoted by $H$.

Similarly, Zariski decompositions, first proved for effective divisors on smooth projective algebraic surfaces by Zariski [19], also exist for all pseudo-effective divisor classes (i.e., divisor classes in the closure of the effective cone) on such surfaces [14]. This thus includes, as noted here, the semi-effective divisor classes (i.e., divisor classes a multiple of which are equal to the class of an effective divisor).

The notion of semi-effectivity raises a problem related to bounded negativity. For each integral, primitive, semi-effective divisor class $D$, let $t_D$ be the least integer $t > 0$ such that $t_DD$ is the class of an effective divisor. Is the set of all such $t_D$ bounded above? (An open problem, posed as a question by D. Eisenbud and M. Velasco around 2009, is whether there’s an algorithm to determine whether a given divisor class is semi-effective.)

In the case that $D$ is semi-effective and primitive and the intersection matrix of the components of $t_DD$ is negative definite, then the equivalence of bounded negativity and bounded Zariski denominators implies that the set of $t_D$ for such negative definite $D$ is bounded above (since, as in
the proof of [1 Lemma 2.5], \( t_D \) is a Zariski denominator for the Zariski decomposition of \( D \) if the surface \( X \) on which \( D \) lives has bounded negativity. But the extent to which bounded negativity and bounded torsion of semi-effectivity modulo effectivity are equivalent is unclear.

Finally, Waldschmidt constants were introduced in [18] as complex analytic invariants related to work on transcendence theory in number theory and further studied in [6]. They arose again in work on a conjecture of Nagata and on Seshadri constants [14, Remark III.7] and on containment problems [5, Lemma 2.3.1]. See [12, 13, 9] for some recent work related to Waldschmidt constants.

**References**


3. Topic 3: Containment Problems

Speaker: Brian Harbourne
Venue: Krakow, Poland, May, 2015
Date of Draft: May 19, 2015

3.1. Background. Let \( p_1, \ldots, p_s \) be distinct points in the plane and let \( I \) be the radical ideal of the points. Recall that the \( m \)th symbolic power of \( I \) is \( I^{(m)} = I(p_1)^m \cap \cdots \cap I(p_s)^m \). It is interesting to compare this with the \( r \)th ordinary power \( I^r = (I(p_1) \cap \cdots \cap I(p_s))^r \) for various \( m \) and \( r \).

**Exercise 3.1.1.** If \( r \geq m \), show that \( I^r \subseteq I^{(m)} \). If \( m > r \), show that \( I^r \not\subseteq I^{(m)} \) (hint: considering the primary decomposition of \( I^r \), look at Hilbert functions and show that \( h_{rZ}(t) < h_{mZ}(t) \) for \( t \gg 0 \) where \( Z = p_1 + \cdots + p_r \)). Conclude that \( I^r \subseteq I^{(m)} \) if and only if \( r \geq m \).

A formerly open question was:

**Question 3.1.2.** Is \( \inf_{I \in S} \{ n : I^{(m)} \subseteq I^r \text{ for all } m \geq rn \} \) infinite (where \( S \) is the set of all ideals of reduced 0-dimensional subschemes of \( \mathbb{P}^2 \))?

Motivated by [13], the papers [9, 12] give a very general answer:

**Theorem 3.1.3.** Let \( I \subseteq K[\mathbb{P}^N] \) be a homogeneous ideal \( I \). Then \( I^{(r(m+N-1))} \subseteq I^{(m)} \), and if \( m \geq rN \), then (with an appropriate definition of symbolic power when \( I \) is not a radical ideal of a finite set of points) we have \( I^{(m)} \subseteq I^r \).

3.2. Sharpness of the bound. Optimality of the bound \( m \geq rN \) for the containment \( I^{(m)} \subseteq I^r \) was considered in [2], which showed in one sense that the bound of Theorem 3.1.3 is optimal:

**Theorem 3.2.1.** Taking \( S \) to be the set of all radical ideals in \( K[\mathbb{P}^N] \), we have

\[
\sup_{I \in S} \left\{ \frac{m}{r} : I^{(m)} \not\subseteq I^r \right\} = N.
\]

The proof uses a simple criterion for failure of containment.

**Exercise 3.2.2.** Let \( I \) and \( J \) be a nonzero homogeneous ideals in \( K[\mathbb{P}^N] \). Recall \( \alpha(I) \) is the least degree of a nonzero form in \( I \).

1. Show that \( \alpha(I^r) = r\alpha(I) \).
2. Show that \( J \not\subseteq I^r \) if \( \frac{\alpha(J)}{r} < \alpha(I) \).
3. If \( I \) is the radical ideal of a finite set of points with \( \frac{m}{r} < \frac{\alpha(I)}{\alpha(I)} \),

show that \( I^{(mt)} \not\subseteq I^{rt} \) for all \( t \gg 0 \).

To prove Theorem 3.2.1 in the case of \( N = 2 \), let \( Z \) be the reduced union of the points of intersection of \( d \) general lines in the plane. The number \( s \) of points is \( \binom{d}{2} \). Let \( p_1, \ldots, p_s \) be the points. Let \( X \) be the surface obtained by blowing up the points, with \( E_i \) being the blow up of \( p_i \) and \( L \) the pullback to \( X \) of a general line. Let \( E = E_1 + \cdots + E_s \) and let \( F_{t,m} = tL - mE \).

**Exercise 3.2.3.** Apply Exercise 3.2.2(3) and Exercise 2.6.7 to prove Theorem 3.2.1 when \( N = 2 \).

One can ask in what other senses Theorem 3.1.3 might be optimal.

Huneke asked:

**Question 3.2.4.** Let \( I_Z \) be the radical ideal of a finite set \( Z \subseteq \mathbb{P}^2 \). Is it always true that \( I_Z(3) \subseteq I_Z^2 \)?
Prompted by this, Harbourne conjectured for homogeneous ideals $I$ in $K[\mathbb{P}^N]$ that $I^{(rN-N+1)} \subseteq I^r$, but this turns out to be false [8]. A minor variation on the counterexample given in [8] was given in [5], outlined in the following exercise.

**Exercise 3.2.5.** Let $Z = p_1 + \cdots + p_{12}$ be the union of any 12 of the 13 points of $\mathbb{P}^2$ over the prime field of characteristic 3. Let $p_{13}$ be the remaining point.

1. Show that $\alpha_Z = 4$.
2. Show that $\beta_Z = 4$.
3. Using facts about Hilbert functions now show that the only possibility is $\Delta h_Z = (1, 2, 3, 4, 2)$.
4. Show that $p_{13}$ is not a base point of $(I_{3Z})_9$ but that it is a base point of $(I_Z)_4$ and hence of $(I_Z^2)_9$.
5. Conclude that $I_Z(3) \not\subseteq I_Z^2$.

Examples of points sets $Z$ giving $I_Z(3) \not\subseteq I_Z^2$ are rare. Most examples known come from line arrangements in which there are no points where exactly 2 lines cross. As shown in [6], at least some of a family of combinatorially interesting line arrangements [10] provide $Z$ (defined over the reals and hence forced to have simple crossings) with $I_Z(3) \not\subseteq I_Z^2$. At least one of these, shown in Figure 1 and taken from [7] but essentially the same one as given in [6] (just not as symmetrical), can be defined over the rationals. In this example, $Z$ consists of the crossing points of multiplicity more than 2.

![Figure 1. A configuration of 12 lines with 19 triple points.](image)

**Exercise 3.2.6.** Use Geogebra to draw other Füredi-Palásti arrangements. Can any others besides the one shown above be defined over the rationals? Use Macaulay2 to check whether the ideal $I_Z$ satisfies $I_Z(3) \not\subseteq I_Z^2$, where $Z$ is the set of singular points of multiplicity more than 2. Does $I_Z(3) \not\subseteq I_Z^2$ ever happen if one does not exclude all of the simple crossings?

### 3.3. Chudnovsky’s Conjecture

Another approach to studying optimality of Theorem 3.1.3 was conjectured in [11] and is still open. Here is a version of that conjecture:

**Conjecture 3.3.1** (Harbourne and Huneke). Let $I \subset K[\mathbb{P}^N] = K[x_0, \ldots, x_N]$ be the radical ideal of a finite set of points $Z$. Let $M$ be the ideal generated by the variables $x_i$. Then $I^{(rN)} \subseteq M^{rN-r}$. 

This was motivated by results of Waldschmidt [16, 17] and Skoda Sk3, by a conjecture of Chudnovsky [3] (the original statements were over the complex numbers) and by an observation of Harbourne:

**Theorem 3.3.2** (Waldschmidt and Skoda). *Let $I$ be the ideal of a finite set of points in $\mathbb{P}^N$. Then*

\[
\frac{\alpha(I^{(m)})}{m + N - 1} \leq \tilde{\alpha}(I).
\]

**Conjecture 3.3.3** (Chudnovsky). *Let $I$ be the ideal of a finite set of points in $\mathbb{P}^N$. Then*

\[
\frac{\alpha(I) + N - 1}{N} \leq \tilde{\alpha}(I).
\]

Harbourne observed that Theorem 3.3.2 is an almost immediate corollary of Theorem 3.1.3. Since $I^{(r(m+N-1))} \subseteq (I^{(m)})^r$, we have $r\alpha(I^{(m)}) = \alpha((I^{(m)})^r) \leq \alpha(I^{(r(m+N-1))})$ hence

\[
\frac{\alpha(I^{(m)})}{m + N - 1} \leq \frac{\alpha(I^{(r(m+N-1))})}{r(N + m - 1)}.
\]

Taking limits as $r \to \infty$ gives the result.

When Harbourne pointed this out to Huneke, the latter observed that a similar containment would imply Chudnovsky’s Conjecture (i.e., Conjecture 3.3.3), which led to Conjecture 3.3.1 in [11].

**Exercise 3.3.4.** Assuming Conjecture 3.3.1 prove Conjecture 3.3.3

### 3.4. Monomial ideals

Line arrangements have given a fruitful approach to studying containment conjectures. Monomial ideals give another good testbed. See for example [4, 13].

---

**References**


