SIAM J. APPL. ALGEBRA GEOMETRY
Vol. 7, No. 1, pp. 29–48

Complex Links and Hilbert–Samuel Multiplicities*

Martin Helmer† and Vidit Nanda‡

Abstract. We describe a framework for estimating Hilbert–Samuel multiplicities $\mathbf{e}_X Y$ for pairs of projective varieties $X \subset Y$ from finite point samples rather than defining equations. The first step involves proving that this multiplicity remains invariant under certain hyperplane sections which reduce X to a point p and Y to a curve C. Next, we establish that $\mathbf{e}_p C$ equals the Euler characteristic (and hence the cardinality) of the complex link of p in C. Finally, we provide explicit bounds on the number of uniform point samples needed (in an annular neighborhood of p in C) to determine this Euler characteristic with high confidence.

Key words. intersection multiplicity, stratified Morse theory, complex links, Hilbert–Samuel multiplicity

MSC codes. 14B05, 57N80, 13H15, 32S60, 14J17

DOI. 10.1137/22M1475533

1. Introduction.

One of the most fundamental quantities of interest in intersection theory is the **Hilbert–Samuel multiplicity**, which associates an integer $\mathbf{e}_X Y \geq 0$ to each pair consisting of an irreducible subvariety X inside a pure-dimensional scheme Y. This integer serves, among other things, as a coarse measurement of the singularity type of X inside Y. When Y is reduced, $\mathbf{e}_X Y = 1$ holds if and only if X is nonempty and smoothly embedded in Y. The importance of Hilbert–Samuel multiplicities stems from their wide-ranging connections with several other intersection-theoretic invariants. For instance, $\mathbf{e}_X Y$ appears as the coefficient of $[X]$ in the Segre class $s(X, Y)$ [7, Chapter 4.3], in Fulton and MacPherson's intersection product [7, Chapter 12.3], and in Serre's Tor formula [22, Theorem 1, page 112]. Computing $\mathbf{e}_X Y$, either directly from its definition or as a consequence of these connections, requires serious algebraic manipulations of the defining equations for X and Y.

Our goal in this paper is to describe a new framework for estimating $\mathbf{e}_X Y$ from finite local point samples without recourse to any such equations. In this setting, we have no means to capture the scheme structure of Y and will therefore restrict ourselves to the case where Y is reduced, i.e., a pure-dimensional variety in some n-dimensional projective space $\mathbb{P}^n$. The methods developed here could also be applied to any data set which we would expect to have the structure of a complex variety, even if the variety is not known. Here is an informal version of our main result for estimating $\mathbf{e}_X Y$ for such pairs $X \subset Y$.

*Received by the editors February 2, 2022; accepted for publication (in revised form) November 10, 2022; published electronically March 10, 2023.
https://doi.org/10.1137/22M1475533
Funding: The work of the second author was supported by EPSRC grant EP/R018472/1 and by DSTL grant D015 funded through the Alan Turing Institute.
†Department of Mathematics, North Carolina State University, Raleigh, NC 27695 USA (mhelmer@ncsu.edu).
‡Mathematical Institute, University of Oxford, Oxford, UK, OX2 6GG (nanda@maths.ox.ac.uk).

Theorem. *Let $X \subset Y$ be projective varieties in $\mathbb{P}^n$, and let $L \subset \mathbb{P}^n$ be a linear space obtained by intersecting $(\dim Y - 1)$ hyperplanes which are general except for the requirement that they all pass through a generic point p of X. The Hilbert–Samuel multiplicity $\mathbf{e}_X Y$ can be determined with high confidence from a sufficiently large (but finite) uniform point sample S lying on the curve $Y \cap L$ in a local annular neighborhood around p.*

We provide explicit bounds on how large S must be in terms of the local geometry of $Y \cap L$ near p and the desired probability of successful estimation. Our proof has three basic steps, each involving a different key ingredient and producing an intermediate result. These steps are summarized below.

Step 1: Algebra. We first establish that $\mathbf{e}_X Y$ is invariant under the operation of slicing both X and Y by certain hyperplanes. The key ingredient here is a new *degree formula* for Hilbert–Samuel multiplicities [11, Theorem 5.3]. Using this formula, we prove the following result.

Theorem (**A**). *Given $X \subset Y$ as above, let L be the intersection of k general hyperplanes which all pass through some general point p of X; then the following hold:*

1. *if $k < \dim X$, then $\mathbf{e}_X Y = \mathbf{e}_{X\cap L}(Y \cap L)$; moreover,*
2. *if $k = \dim X$, then $\mathbf{e}_X Y = \mathbf{e}_p(Y \cap L)$; and finally,*
3. *if $k \leq \dim Y - 1$, then $\mathbf{e}_X Y = \mathbf{e}_p(Y \cap L)$.*

In fact, the first two assertions follow readily from basic properties of Segre classes, whereas the last one is new and makes essential use of the aforementioned degree formula from [11]. As a consequence of this third assertion (for $k = \dim Y - 1$), every $\mathbf{e}_X Y$ calculation can be reduced to the case where $X = p$ is a point and $Y = C$ is a curve in $\mathbb{P}^n$ containing p. In this special case, the degree formula for Hilbert–Samuel multiplicity simplifies to

$$\mathbf{e}_p C = \deg(C) - \deg((C \cap H) - p),$$

where H is a general hyperplane passing through p. While this is a convenient reformulation for algebraic computation of $\mathbf{e}_p C$, both degrees appearing on the right-hand side are global computations in the sense that they require checking for intersections far away from p. The purpose of the next step is to replace these with a local computation near p.

Step 2: Topology. The starting point for our second step is the observation that $\deg(C)$ equals the cardinality of $C \cap H'$, where H' is a general hyperplane in $\mathbb{P}^n$. Crucially, we let H' be parallel to the plane H when restricted to an affine chart of $\mathbb{P}^n$ containing p. The special ingredient here is *Thom's first isotopy lemma* [9, Chapter I.1.5], which allows us to relate $\mathbf{e}_p C$ to the Euler characteristic (and hence cardinality) of the zero-dimensional space

$$\mathscr{L}_p := C \cap \mathbf{B}_\epsilon(p) \cap H'.$$

Here $\mathbf{B}_\epsilon(p)$ denotes a small closed ball around p in a chart of $\mathbb{P}^n$. In particular, we employ a homological argument to show the following result.

Theorem (**B**). *If p is any (possibly singular) point on a curve $C \subset \mathbb{P}^n$, then its Hilbert–Samuel multiplicity satisfies $\mathbf{e}_p C = \chi(\mathscr{L}_p)$, where χ denotes Euler characteristic.*

The space $\mathscr{L}_p$ plays a fundamental role in (complex) stratified Morse theory—it provides *normal Morse data* for p with respect to a stratified Morse function defined on C and is called the **complex link** of p in C [9, Chapter II.2]. Variants of Theorem (**B**) have been assigned as exercises to the reader on several occasions, including the introduction to [9] and [15, Example 4.6]. However, we were unable to locate a proof in the literature; since it forms an essential part of our overall argument, we have included a proof here.

Step 3: Geometry. It remains to estimate the cardinality of $\mathscr{L}_p$ using a uniform finite point sample S chosen from $B := C \cap \mathbf{B}_\epsilon(p)$. The main difficulty here is that generically the intersection $S \cap H'$ will be empty even when the sample size is enormous. As such, we are compelled to search for points of S which lie within some small distance $\alpha_0 > 0$ of H' and hope that these points naturally organize into $\mathbf{e}_X Y$-many clusters. The key ingredient here is a suite of *geometric inference results*, which date back to the work of Niyogi, Smale, and Weinberger from [18]. Given a compact Riemannian submanifold $M \subset \mathbb{R}^d$ and a probability parameter $\gamma \in (0,1)$, these results give explicit bounds on the cardinality of a finite point sample $P \subset M$ required to estimate the homology of M with probability exceeding $(1-\gamma)$.

Recently, Wang and Wang have extended results of [18] to the case where $M \subset \mathbb{R}^d$ is a smooth submanifold with boundary [27]; they provide an explicit lower bound $N_M(\alpha, \gamma)$ on the size of a uniform point sample $P \subset M$ required to ensure, again with probability at least $(1-\gamma)$, that P is $\alpha/2$-dense in M. These results require α to be sufficiently small relative to the injectivity radii of the embeddings $M \hookrightarrow \mathbb{R}^d$ and $\partial M \hookrightarrow \mathbb{R}^d$ of the manifold and its boundary, respectively. Although the space B of interest to us is not a manifold (thanks to the singularity at p), it does become a manifold with boundary by removing the interior of a smaller ball $\mathbf{B}_{\epsilon_0}(p)$ with $\epsilon_0 < \epsilon$. After this excision, we can safely apply the density results from [27] and obtain the following result.

Theorem (**C**). *Fix sufficiently small radii $0 < \epsilon_0 < \epsilon$. There exists, for all sufficiently small radii $\alpha > 0$ and probabilities $\gamma \in (0,1)$, an explicit bound $N(\alpha, \gamma)$ with the following property. Any uniformly sampled subset*

$$S \subset [C \cap (\mathbf{B}_\epsilon(p) - \mathbf{B}_{\epsilon_0}(p)^\circ)]$$

of cardinality $\#S > N(\alpha, \gamma)$ can be used to correctly compute the Euler characteristic $\chi(\mathscr{L}_p)$ with probability exceeding $(1-\gamma)$.

The value of α in the theorem above will change if a new ϵ and ϵ_0 are chosen. Moreover, the process of recovering $\chi(\mathscr{L}_p)$ requires us only to cluster together points of S which lie within a small distance $\alpha_0 < \alpha$ of a random hyperplane passing near (but not through) the point p. Combining Theorems (**A**), (**B**), and (**C**) gives the promised main result. This result may also be extended to affine varieties $X \subset Y \subset \mathbb{C}^n$ by passing to their projective closures; see section 3.1.

Towards Implementation and Applications. The methods developed here are meant as a natural addition to the growing body of work on using sampled data points to understand the geometry and topology of algebraic varieties [3] and of point cloud data which are well-modelled by such varieties [1]. Our main result leads to the following probabilistic algorithm for estimating $\mathbf{e}_X Y$ for pairs of complex varieties $X \subset Y$ in $\mathbb{C}^n$ from a sufficiently large uniform finite point sample $S \subset Y$ in a neighborhood of a randomly chosen point $p \in X$.

1. Select a random affine subspace $A_p \subset \mathbb{C}^n$ of dimension $(n - \dim Y - 1)$ which passes through p (so in particular $Y \cap A_p$ is a curve containing p).
2. Select a random affine hyperplane $H' \subset A_p$ which passes within some small distance $\delta > 0$ of p.
3. For small radii $\epsilon \gg \delta \gg \alpha_0$, let S' be the subset of $S \cap (Y - X)$ containing points which simultaneously lie within distance α_0 of H' and within distance ϵ of p.

Theorems 3.5, 4.1, and 5.4 of this paper guarantee that the number of clusters in S' is an accurate estimator for $\mathbf{e}_X Y$ for sufficiently large S provided that the parameters ϵ, δ, and α_0 are chosen appropriately. Figure 1 contains three instances of S' obtained from the same point sample S.

Even when one has direct access to the polynomials defining X and Y, it may be advantageous to avail oneself of the probabilistic strategy described above. For instance, computing Jacobian minors of the appropriate size to find the singular locus of Y, where one expects to find X, might already be prohibitive. If Y is a 7-dimensional variety defined by 14 polynomials of degree 5 in 13 variables, computing its singular locus requires considering more than 5 million polynomials [13], each of degree $4^6 = 4096$. In such cases, our sample-based method may provide the only way to compute Hilbert–Samuel multiplicities. To properly implement the probabilistic algorithm, one would require (localized versions of) techniques for densely sampling varieties [3] as well as techniques for detecting singular loci from such samples [25] without knowing the equations. Finally, we would require heuristic estimates for various condition numbers and radii from section 5 of this paper. Describing all of this carefully here would take us too far afield of our main goals, but we hope to address these topics in future work.

One natural application domain for Hilbert–Samuel multiplicities can be found in the discipline of *algebraic kinematics*. The configuration space of a wide variety of robotic mechanisms can be modelled with polynomial (or trigonometric polynomial) equations; and by using isotropic coordinates, it becomes reasonable in several cases of interest to assume that such a configuration space is a complex algebraic variety [26, section 3.2]. A singularity is defined

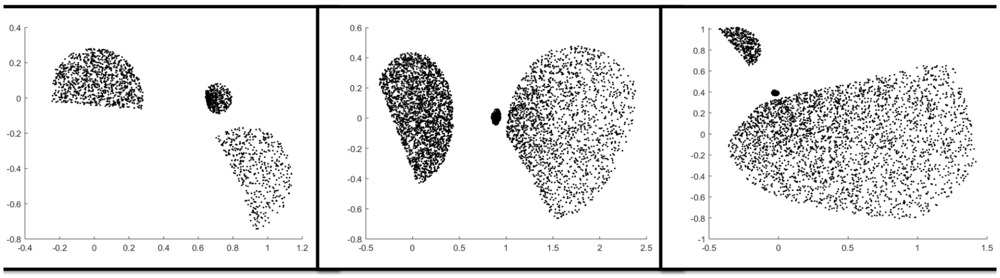


Figure 1. *A visualization of the output of our multiplicity inference procedure applied to the point $X = (0,0)$ inside the curve $Y \subset \mathbb{C}^2$ defined by $(y - x^2)(y + x^2)y = 0$. For this example, the multiplicity $\mathbf{e}_X Y$ equals 3. The visualization shows three different random choices of the the offset hyperplane H' which defines the complex link. The resulting point clusters in $\mathbb{C}^2 \simeq \mathbb{R}^4$ have been projected onto their real coordinates.*

in this context to be any configuration where the mechanism loses a degree of freedom. In an algebraic formulation of kinematics, these singularities are related to the singularities of the model algebraic variety; see [26, section 4.4.1] or [19]. The multiplicity of such singular points has a bearing on the kinematic properties the mechanism being considered.

The effect of multiplicities of singular points on the behavior of mechanisms was explored for the case of the 3-*RPR manipulator* in [17]. For this manipulator, the singular configurations of *cuspidal* type are of particular interest since these configurations allow for a nonsingular change of assembly mode. In [17, section 4], these singularities are identified by checking whether or not the Hilbert–Samuel multiplicity of a singular point on an algebraic variety exceeds three.[1] It is also interesting to note that the presence of the cuspidal configurations was first observed by a numeric approach based on sampling and graphing the associated joint space in [29], indicating to us that a full version of the procedures above would be a natural tool to apply to larger and more difficult problems of this type.

Organization. In section 2, we briefly review the definition of the complex link. Section 3 focuses on Step 1; here we give a brief overview of the Hilbert–Samuel multiplicity and prove Theorem (**A**). In section 4, we implement Step 2 by providing a proof of the folklore Theorem (**B**). And, finally, in section 5 we carefully state and establish Theorem (**C**) by describing not only the exact form of the bound $N(\alpha,\gamma)$ but also the precise constraints on α imposed by the local geometry of C near p.

2. Complex links.

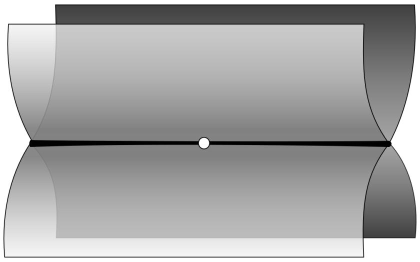
A *stratification* of a topological space $\mathbf{W}$ is a filtration

$$\emptyset = W_{-1} \subset W_0 \subset \cdots \subset W_k = \mathbf{W}$$

by closed subspaces so that each consecutive difference $W_i - W_{i-1}$ is a (possibly empty or disconnected) i-dimensional manifold called the *i-stratum*. Throughout this section, $\mathbf{W}$ will denote a Whitney-stratified complex analytic subspace of $\mathbb{C}^n$. We assume that each stratum $X \subset \mathbf{W}$ is a connected complex analytic manifold and write $Y > X$ to indicate that the closure of the stratum Y contains the stratum X. We further require all pairs of strata $X < Y$ to satisfy Whitney's Condition (B); see [28, section 19], [16, section 2], or [9, Chapter I.1.2]. Let T_pX denote the $(\dim X)$-dimensional linear subspace of $\mathbb{C}^n$ which corresponds to the tangent space of a stratum X at a point p in X.

Fix a (connected) stratum $X \subset \mathbf{W}$, and consider an arbitrary point p in X. Since it remains difficult to illustrate even 2-dimensional complex varieties, the following real picture (where $n = 3$ and $\dim \mathbf{W} = 2$ while $\dim X = 1$) will serve as a proxy for the local structure of $\mathbf{W}$ near p.

[1]The authors of [17] have considered the problem of computing the multiplicity of singular points on a surface in joint space, which is defined by the polynomials in [17, equation (3)]. They reduce this to the task of finding the multiplicity of a point in a zero-dimensional scheme (see [17, Definitions 5 and 6] for details). Their multiplicity is obtained from the *rational univariate representation* of [20], which agrees with the classical definition of the multiplicity of a point (see, e.g., [6, section 1] or [2, Definition 2.1]). This classical definition is in turn equivalent to the Hilbert–Samuel multiplicity of points in the case of local complete intersections [6, page 12], which subsume all varieties considered in [17].



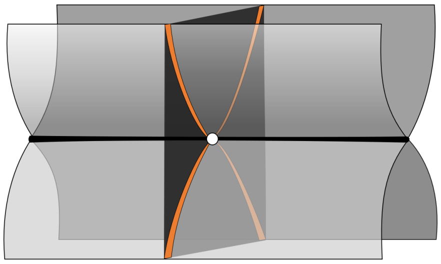
The stratum X is represented by the horizontal line along which the four sheets intersect, and the chosen point p is located near the center of X. We say that an affine subspace $A \subset \mathbb{C}^n$ containing p is *transverse* to X at p if the sum of tangent subspaces given by

$$T_pX + T_pA = \{v + w \mid v \in T_pX \text{ and } w \in T_pA\}$$

equals $T_p\mathbb{C}^n = \mathbb{C}^n$.

Definition 2.1. *A subset* $\mathbf{N} \subset \mathbf{W}$ *is called a* normal slice *to* X *at* p *if it equals the intersection* $\mathbf{W} \cap A$ *for some* $(n - \dim X)$*-dimensional affine subspace* $A \subset \mathbb{C}^n$ *which intersects* X *transversely at* p.

One possible choice of $\mathbf{N}$ for our example is shown below:



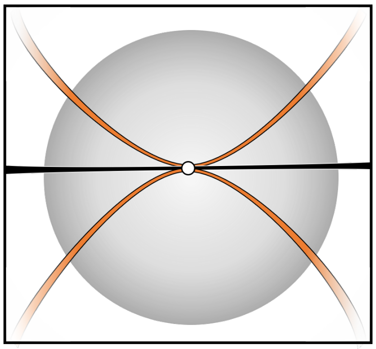
Here A is the plane which crosses X at p, while $\mathbf{N}$ is the union of four half-open arcs, all of which intersect at p. Evidently, $\mathbf{N}$ will not be a manifold in general; on the other hand, it follows from the definition of a Whitney stratification that A will remain transverse, at least in a small neighborhood around p, to all higher strata $Y > X$. Thus, $\mathbf{N}$ inherits a Whitney stratification from $\mathbf{W}$ near p as follows. Each $(\dim Y)$-dimensional stratum $Y > X$ carves out a (possibly disconnected, $(\dim Y - \dim X)$-dimensional) stratum $Y \cap A$ of $\mathbf{N}$.

Fix a radius $\epsilon > 0$ so that the intersection of $\mathbf{N}$ with the closed ball $\mathbf{B}_\epsilon(p)$ of radius ϵ around p inherits a Whitney stratification from $\mathbf{W}$ in the manner described above.[2] We write $\mathbf{N}_\epsilon(p) = \mathbf{N} \cap \mathbf{B}_\epsilon(p)$ to indicate this *restricted* normal slice. The next definition will make use of our chosen $\mathbf{N}$ and ϵ, and also of the usual inner product $\langle \bullet, \bullet \rangle$ defined on the ambient space $\mathbb{C}^n$.

[2]More precisely, two natural transversality constraints must hold for every radius $e \leq \epsilon$ and for every stratum Y of $\mathbf{W}$. First, the boundary of $\mathbf{B}_e(p)$ must be transverse to Y in $\mathbb{C}^n$, and, second, the boundary of $\mathbf{B}_e(p) \cap A$ must be transverse to $Y \cap A$ in A.

Definition 2.2. *A vector ξ in the affine space A is called* **nondegenerate** *for the pair $(\mathbf{N},\epsilon)$ if the following property holds for all strata $Y > X$. Given any sequence $\{(q_i,v_i)\}$ in the tangent bundle of $Y' = Y \cap \mathbf{N}_\epsilon(p)$, where q_i limits to p, if the v_i limits to some nonzero vector v, then $\langle \xi, v\rangle \neq 0$.*

If we restrict our pictorial example to the affine plane A, then the set of degenerate vectors will span the vertical line through p because the orthogonal complement of this vertical line (in A through p, as drawn below) shares a limiting tangent with all four arcs of $\mathbf{N}$. For any vertically aligned ξ, one can find a sequence (q_i,v_i) in the tangent bundle of each arc with $q_i \to p$ and $v_i \to v \neq 0$ lying along the horizontal line, which in turn forces $\langle \xi, v\rangle = 0$. Any ξ off the vertical line will be nondegenerate.



Fix a nondegenerate vector ξ for $(\mathbf{N},\epsilon)$, and consider the affine-linear map

$$\pi_\xi : A \to \mathbb{C}$$

given by $z \mapsto \langle z - p, \xi\rangle$. By nondegeneracy, there exists a $\delta > 0$ so that if the differential

$$(d\pi_\xi)_q : T_q Y' \to \mathbb{C}$$

at some point $q \neq p$ lying in a stratum $Y' \subset \mathbf{N}_\epsilon(p)$ annihilates a limiting tangent plane at q, then $|\pi_\xi(q)| > \delta$. Here is a summary of all the choices that have been made for the stratum $X \subset \mathbf{W}$ of dimension $\dim X$:

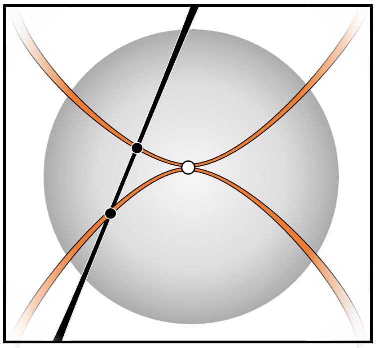
1. a point $p \in X$,
2. an $(n - \dim X)$-dimensional affine subspace $A \subset \mathbb{C}^n$ transverse to X at p,
3. a radius $\epsilon > 0$ so that $\mathbf{N}_\epsilon(p) = \mathbf{W} \cap A \cap \mathbf{B}_\epsilon(p)$ inherits a stratification from $\mathbf{W}$,
4. a nondegenerate vector $\xi \in A$, and, finally,
5. another radius $\delta \in (0,\epsilon)$ so that the restriction of π_ξ to $\mathbf{N}_\epsilon(p)$ has no critical points other than p valued in the closed ball of radius δ around $0 \in \mathbb{C}$.

The following definition makes provisional use of this tuple $(p,A,\epsilon,\xi,\delta)$.

Definition 2.3. *The* **complex link** *of the stratum $X \subset \mathbf{W}$ with respect to the choices $(p,A,\epsilon,\xi,\delta)$ is the intersection*

$$\mathscr{L}_X = \mathbf{N}_\epsilon(p) \cap \pi_\xi^{-1}(\delta).$$

Returning to our example one final time, the hyperplane $\pi_\xi^{-1}(\delta)$ is a nonhorizontal line in the plane A which passes near, but not through, the central point p. In a small ϵ-ball around p, this line generically intersects the arcs which form $\mathbf{N}_\epsilon(p)$ in two points, so the complex link $\mathscr{L}_X$ in this case is just the two-point space:



The (stratified homeomorphism type of the) complex link $\mathscr{L}_X$ depends only on the stratum X, and not on the auxiliary choices $(p, A, \epsilon, \xi, \delta)$ described above [9, Chapter II.2]. It is also interesting to note that the invariance of $\mathscr{L}_X$ to the chosen direction ξ is entirely a feature of complex analytic geometry—for real analytic Whitney stratified spaces, the stratified homeomorphism type of the intersection $\mathbf{N}_\epsilon(p) \cap \pi_\xi^{-1}(\delta)$ is liable to change as ξ is varied.

Remark 2.4. In this paper, we will be exclusively interested in the complex link of a point (i.e., a zero-dimensional stratum) within a complex projective variety. In this special case, one is not required to construct a normal slice, so the formula from Definition 2.3 reduces to $\mathscr{L}_X = \mathbf{W} \cap \mathbf{B}_\epsilon(p) \cap \pi_\xi^{-1}(\delta)$.

3. The Hilbert–Samuel multiplicity.

Let $R = \mathbb{C}[x_0, \ldots, x_n]$ denote the coordinate ring of affine space $\mathbb{C}^{n+1}$. Let X be an irreducible complex algebraic variety given by a prime ideal $I \lhd R$, and let Y be a scheme corresponding to a primary ideal $J \subset I$. The *local ring of Y along X*, usually written as $\mathscr{O}_{X,Y}$, is the localization of (R/J) at I. Following [23, Chapter 2, section 1.1], if X and Y are projective or quasi-projective, we define $\mathscr{O}_{X,Y}$ via a dense affine patch; i.e., we define $\mathscr{O}_{X,Y}$ as the local ring of $U \cap Y$ along $U \cap X$ for $U \subset Y$ any open affine variety where $U \cap X$ is nonempty. The following notion is due to Samuel [21].

Definition 3.1. *Let $\mathscr{M}$ be the maximal ideal of $\mathscr{O}_{X,Y}$, and let c be the codimension* $\dim Y - \dim X$. *The* **Hilbert–Samuel function** *of Y along X is*

$$\mathbf{HS}(t) = \text{length}\left(\mathscr{O}_{X,Y}/\mathscr{M}^t\right).$$

For all $t \gg 0$, this function is a polynomial in t of degree c whose leading coefficient is a strictly positive integer divisible by $c!$—and the **Hilbert–Samuel multiplicity** *of Y* **along** *X, written as $\mathbf{e}_X Y$, is the leading coefficient of the normalized polynomial* $(1/c!) \cdot \mathbf{HS}(t)$.

It is shown in [7, section 4.3] that $\mathbf{e}_X Y$ is also equal to the coefficient of $[X]$ in the *Segre class* $s(X, Y)$, which naturally lives in the *Chow group* of X (or in the *Chow ring* of an ambient smooth variety M via push-forward, $X \subset Y \subset M$; we will often work in the $M = \mathbb{P}^n$ setting).

3.1. Multiplicities from degrees. When $X \subset Y \subset \mathbb{P}^n$ are projective varieties, it is often algorithmically convenient to extract $\mathbf{e}_X Y$ from a choice $B_X = \{f_0, \ldots, f_r\}$ of homogeneous polynomials that generate the defining ideal $I \lhd R$ of X. We will assume here that all the f_i have the same degree d, which is always possible to arrange without loss of generality [11, section 2.1.4]. Let L_i be a generic $(n-i)$-dimensional linear subspace of $\mathbb{P}^n$, and let $V_i \subset \mathbb{P}^n$ be the varieties given by

$$V_i = \{x \in \mathbb{P}^n \mid F_1(x) = F_2(x) = \cdots = F_{\dim Y - i}(x) = 0\}, \tag{1}$$

where the F_j are homogeneous polynomials of degree d that have the form

$$F_j = \sum_{k=0}^{r} \lambda_k^j f_k$$

for general choices of $\lambda_k^j \in \mathbb{C}$. (Note that V_i contains X for all i by design.)

In [11], it is shown that the Segre class $s(X, Y)$, and hence the multiplicity $\mathbf{e}_X Y$, is determined by the numbers

$$\mathbf{\Lambda}_X^i Y = \deg(Y) \cdot d^{\dim Y - i} - \deg((Y \cap V_i \cap L_i) - X) \tag{2}$$

for each i between 0 and $\dim X$. In particular, [11, Theorem 5.3] establishes that

$$\mathbf{e}_X Y = \frac{\mathbf{\Lambda}_X^{\dim X} Y}{\deg X} \tag{3}$$

or, more explicitly,

$$\mathbf{e}_X Y = \frac{\deg(Y) \cdot d^{\dim Y - \dim X} - \deg((Y \cap V_{\dim X} \cap L_{\dim X}) - X)}{\deg X}. \tag{4}$$

The construction (3) requires that the varieties under consideration be projective. In order to apply the formula (3) to affine varieties $X \subset Y$ in $\mathbb{C}^n$, we must replace X and Y by their *projective closures* $PX \subset PY$ in $\mathbb{P}^n$; see [12, Exercise I.2.9] for a definition of the projective closure. Note that using the projective closures leaves the Hilbert–Samuel multiplicity unchanged, as, by definition, the local ring $\mathscr{O}_{PX,PY}$ is given by $\mathscr{O}_{X,Y}$.

Remark 3.2. We note that in [11] a different convention is used for the definition of the local ring $\mathscr{O}_{X,Y}$ for projective varieties $X \subset Y \subset \mathbb{P}^n$; in particular it is taken to be $(R/I_Y)_{I_X}$ *for* I_X and I_Y the homogeneous ideals of X and Y, respectively, in the coordinate ring $R = \mathbb{C}[x_0, \ldots, x_n]$ of $\mathbb{P}^n$. This corresponds to the local ring $\mathscr{O}_{\hat{X},\hat{Y}}$ in the convention used here, where $\hat{X}$ and $\hat{Y}$ are the respective affine cones in $\mathbb{C}^{n+1}$. However, from the point of view of Hilbert–Samuel multiplicity the two conventions yield the same result; i.e., the Hilbert–Samuel multiplicity of Y along X is equal to the Hilbert–Samuel multiplicity of $\hat{Y}$ along $\hat{X}$, since $\mathscr{O}_{X,Y}$ is the degree zero part of the graded ring $\mathscr{O}_{\hat{X},\hat{Y}}$.

3.2. Multiplicities of linear sections. Here we describe the behavior of the Hilbert–Samuel multiplicity $\mathbf{e}_X Y$ for complex projective varieties $X \subset Y$ when both X and Y are replaced by their intersections with (sufficiently generic) linear spaces. The proposition below can be seen as a direct consequence of standard properties of Segre classes (along with the relation between Segre classes and multiplicities).

Proposition 3.3. *Let Y be a pure-dimensional subscheme of the complex projective space $\mathbb{P}^n$, let X be an irreducible subvariety of Y, and let $L \subset \mathbb{P}^n$ be given by an intersection*

$$L = H_1 \cap H_2 \cap \cdots \cap H_\ell,$$

where each $H_i \subset \mathbb{P}^n$ is a generic hyperplane. If the codimension $\ell = n - \dim L$ is strictly less than $\dim X$, then the multiplicities $\mathbf{e}_X Y$ and $\mathbf{e}_{X\cap L}(Y \cap L)$ are equal. Further, if $\ell = \dim(X)$, then $\mathbf{e}_X Y = \mathbf{e}_p(Y \cap L)$, where p is any of the $\deg(X)$ points in $X \cap L$.

Proof. In this proof, we will work with the push-forward to the Chow ring of $\mathbb{P}^n$ of the Segre class $s(X,Y)$; in a slight abuse of notation, this will also be denoted as $s(X,Y)$. We recall that the Chow ring is $A^*(\mathbb{P}^n) \cong \mathbb{Z}[h]/\langle h^{n+1}\rangle$, where h is the rational equivalence class of a general hyperplane. Since each H_i is a general divisor on $\mathbb{P}^n$, the coefficient of $h^{\dim X - \ell}$ in the Segre class $s(X \cap L, Y \cap L)$ equals the coefficient of $h^{\dim X}$ in the Segre class $s(X,Y)$, i.e.,

$$\{s(X \cap L, Y \cap L)\}_{\dim X - \ell} = \{s(X,Y)\}_{\dim X} \cdot h^\ell.$$

A proof of the above property of Segre classes can be found, for example, in [10, Corollary 3.2]. First suppose that $\ell < \dim(X)$. Using the fact that $\mathbf{e}_{X\cap L}(Y \cap L)$ is the coefficient of $[X \cap L]$ in $s(X \cap L, Y \cap L)$, one obtains

$$\begin{aligned} \{s(X \cap L, Y \cap L)\}_{\dim X - \ell} &= \mathbf{e}_{X\cap L}(Y \cap L) \cdot [X \cap L] \\ &= \mathbf{e}_{X\cap L}(Y \cap L) \cdot \deg X \cdot h^{n - \dim X + \ell}, \end{aligned}$$

where the second equality follows from the fact that each H_i is a general divisor, so in particular, $\deg X = \deg(X \cap L)$. Now take $\ell = \dim(X)$. Then $X \cap L$ consists of $\deg(X)$ reduced points $p_1, \ldots, p_{\deg(X)}$; by [7, Example 4.3.4], we have that

$$\begin{aligned} \{s(X \cap L, Y \cap L)\}_{\dim X - \ell} &= \mathbf{e}_{p_1}(Y \cap L)[p_1] + \cdots + \mathbf{e}_{p_{\deg(X)}}(Y \cap L)[p_{\deg(X)}] \\ &= \mathbf{e}_p(Y \cap L)\deg(X)[p] \\ &= \mathbf{e}_p(Y \cap L) \cdot \deg X \cdot h^{n - \dim X + \ell}, \end{aligned}$$

where p is any point in $X \cap L$ (all of which are rationally equivalent since we work in the Chow ring of the ambient space $\mathbb{P}^n$). On the other hand, we also have

$$\begin{aligned} \{s(X,Y)\}_{\dim X} \cdot h^\ell &= \mathbf{e}_X Y \cdot [X] \cdot h^\ell \\ &= \mathbf{e}_X Y \cdot \deg X \cdot h^{n - \dim X + \ell}, \end{aligned}$$

which forces $\mathbf{e}_{X\cap L}(Y \cap L) = \mathbf{e}_X Y$ for $\ell < \dim(X)$ and $\mathbf{e}_p(Y \cap L) = \mathbf{e}_X Y$ when $\ell = \dim(X)$ as desired. ∎

Take L to be a general linear space of codimension equal to $\dim(X)$. The result of Proposition 3.3 tells us that if p is a point in $X \cap L$, we have that $\mathbf{e}_X Y = \mathbf{e}_p(Y \cap L)$. This leads us to consider $\mathbf{e}_p(Y \cap L)$ more closely. Hence in the next proposition we consider a distinguished p in Y and we will show, using (3), that intersecting with fewer than $\dim(Y)$ hyperplanes containing p will leave the multiplicity unchanged. This will mean, in particular, that $\mathbf{e}_p(Y \cap L)$ from the last proposition is equal to $\mathbf{e}_p Y$ (for this distinguished p which is some general point in X).

Proposition 3.4. *Let Y be a pure-dimensional subscheme of the complex projective space $\mathbb{P}^n$, let p be any reduced point in Y, and consider a linear space $L \subset \mathbb{P}^n$ given by the intersection of $m \geq 0$ general hyperplanes containing p. If $m \leq \dim Y - 1$, then $\mathbf{e}_p(Y \cap L)$ is independent of the choice of L (provided it is general) and equals $\mathbf{e}_p Y$.*

Proof. The generating ideal of p in the polynomial ring $\mathbb{C}[x_0, \ldots, x_n]$ can be chosen to consist of n linear forms $\{\ell_1, \ldots, \ell_n\}$. It follows that $d = 1$ and $\deg X = 1$ in (4). Since $\dim p = 0$, we have

$$\mathbf{e}_p Y = \deg(Y) - \deg((Y \cap V_{\dim Y}) - p),$$

where $V_{\dim Y}$ is the variety defined by the polynomials $\{P_1, \ldots, P_{\dim Y}\}$, with each P_j being a linear combination of the form

$$P_j = \sum_{i=1}^{n} \lambda_i^j \ell_i \quad \text{for general } \lambda_i^j \in \mathbb{C}.$$

Without loss of generality, we may take L to be the variety defined by the first m of these, say $\{P_1, \ldots, P_m\}$. Thus, L is a linear system with base locus p. It follows from Bertini's theorem (see, for example, [4, Theorem 0.5] or [24, Theorem A.9.2]) that the linear system L forms a smooth complete intersection outside of p so that the intersection $Y \cap L$ is transverse in the expected dimension, i.e., in dimension $\dim(Y \cap L) = \dim Y - m > 0$ and, moreover, $\deg(Y \cap L) = \deg(Y)$. Letting $V_{>m}$ be the variety defined by $\{P_{m+1}, \ldots P_{\dim Y}\}$, we have

$$\begin{aligned} \mathbf{e}_p Y &= \deg(Y) - \deg((Y \cap V_{\dim Y}) - p) \\ &= \deg(Y \cap L) - \deg((Y \cap L \cap V_{>m}) - p) \\ &= \mathbf{e}_p(Y \cap L). \end{aligned}$$

This argument fails when $m = \dim Y$, since the intersection $Y \cap L$ may not be transverse in this case. ∎

Consider a pair of complex projective varieties $X \subset Y$ in $\mathbb{P}^n$, and pick a general point p in X. Let L be a general linear space defined by the intersection of some number of general hyperplanes containing p; Proposition 3.3 tells us that $\mathbf{e}_X Y = \mathbf{e}_p(Y \cap L)$ when L has codimension $\dim(X)$, while Proposition 3.4 gives $\mathbf{e}_p Y = \mathbf{e}_p(Y \cap L)$ when L has codimension between 0 and $\dim(Y) - 1 \geq \dim(X)$. Thus, we have $\mathbf{e}_X Y = \mathbf{e}_p(Y \cap L)$ for any choice of L generated by intersections of general hyperplanes through p (note that this includes the case $L = \mathbb{P}^n$ of zero hyperplanes). We collect these observations in Theorem 3.5 below.

Theorem 3.5. *Let $X \subset Y$ be a pair of complex projective subvarieties of $\mathbb{P}^n$, and let $L \subset \mathbb{P}^n$ be a linear space given by the intersection of $k \geq 0$ general hyperplanes $H_1, \ldots, H_k$ containing a point p of X:*

1. *If $k < \dim X$, then $\mathbf{e}_X Y = \mathbf{e}_{X \cap L}(Y \cap L)$.*
2. *If $k \leq \dim Y - 1$, then $\mathbf{e}_X Y = \mathbf{e}_p(Y \cap L)$.*

This is Theorem (**A**) from the introduction. We note that we may allow Y in the statement above to be any pure-dimensional subscheme of $\mathbb{P}^n$, but we have restricted ourselves to the case where Y is a variety, as this will be the only case we employ in later sections. Assertion 2 of this result (for $k = \dim Y - 1$) implies that the evaluation of $\mathbf{e}_X Y$ for arbitrary projective varieties $X \subset Y$ in $\mathbb{P}^n$ can be reduced to the computation of $\mathbf{e}_p C$, where p is a point lying on the curve $C = Y \cap L$; this scenario will be the central focus of the next section.

4. Point-curve multiplicities via complex links. Our goal here is to provide a stratified Morse-theoretic proof of Theorem (B) from the introduction.

Theorem 4.1. *If p is any (possibly singular) point on a curve $C \subset \mathbb{P}^n$, then we have*

$$\mathbf{e}_p C = \chi(\mathscr{L}_p),$$

where $\mathbf{e}_p C$ is the Hilbert–Samuel multiplicity (from Definition 3.1) and $\chi(\mathscr{L}_p)$ is the Euler characteristic of p's complex link in C (from Definition 2.3).

Since p can be defined as the zero set of n linear polynomials, by (4) we have

$$\mathbf{e}_p C = \deg(C) - \deg((C \cap H_p) - p),$$

where H_p is a generic hyperplane in $\mathbb{P}^n$ passing through p. We will work throughout in a generic chart $\mathbb{C}^n \subset \mathbb{P}^n$ containing p.

Let ξ be the unit normal to H_p, and let $\pi_\xi : C \to \mathbb{C}$ be the inner product map given by $z \mapsto \langle z - p, \xi \rangle$. Since the hyperplane H_p is generic, we may safely assume that the vector ξ is nondegenerate for C at p in the sense of Definition 2.2. Consequently, there exists a small $\delta > 0$ which satisfies the following requirements:

1. There are no singular points of C other than p valued in the closed half-disk

$$D_\delta^+ := \{x + iy \in \mathbb{C} \mid x \geq 0 \text{ and } x^2 + y^2 \leq \delta\}.$$

2. The restriction of π_ξ to the subset $\pi_\xi^{-1}(D_\delta^+) \subset C$ is a surjection onto D_δ^+.
3. The derivative of π_ξ does not vanish at any point of $C - \{p\}$ valued in D_δ^+.

Identifying δ with the point $\delta + 0i$ in $\mathbb{C}$, we note that the level set $\pi_\xi^{-1}(\delta)$ intersects C in a set of cardinality $\deg C$ since $\pi_\xi^{-1}(\delta)$ is a sufficiently generic hyperplane in $\mathbb{C}^n$. On the other hand, the level set $\pi_\xi^{-1}(0)$ is precisely $C \cap H_p$, which may contain fewer than $\deg C$ points because it is forced to pass through p. Thus, the quantity of interest to us here is

$$\mathbf{e}_p C = \#\pi_\xi^{-1}(\delta) - \#\pi_\xi^{-1}(0) + 1, \tag{5}$$

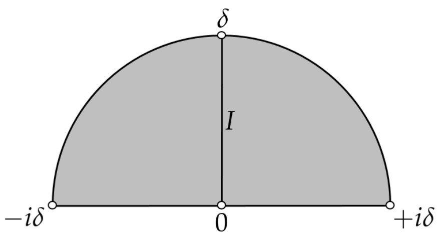
where the last $+1$ term comes from the fact that we are required to discard p from $C \cap H_p$ in the $\mathbf{e}_p C$ formula. The main tool in our argument here is one of Thom's celebrated isotopy lemmas; see [16, Proposition 11.1] or [9, Chapter I.1.5].

Lemma 4.2 (Thom's first isotopy lemma). *Let M and N be smooth manifolds and $Z \subset M$ a Whitney stratified subset. If $f : M \to N$ is a smooth proper map whose restriction $f|_X$ to each stratum $X \subset Z$ is a submersion (i.e., the derivative $df_p : T_pX \to T_{f(p)}N$ is surjective for all p in X), then $f|_X : X \to f(X)$ is a (locally trivial) fiber bundle.*

By our choice of δ, the function π_ξ has a nonzero derivative when restricted to any subset of the form $\pi_\xi^{-1}(S)$, where S is contained in $B_\delta^+ - \{0\}$. Reducing δ further if necessary, Lemma 4.2 therefore guarantees the existence of a local trivialization

$$\pi_\xi^{-1}(S) \simeq S \times \pi_\xi^{-1}(\delta),$$

whenever $S \subset D_\delta^+$ is a Whitney stratum not containing 0. Now consider the Whitney stratification of D_δ^+ into four 0-strata, five 1-strata, and two 2-strata depicted below; we are interested in the closure $\overline{I}$ of the 1-stratum labelled I.



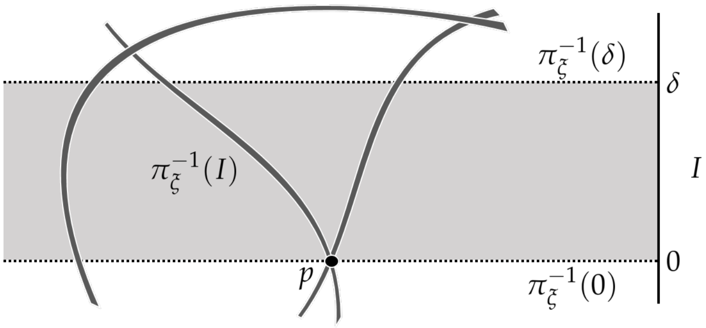
Since ξ is nondegenerate by assumption, we know that p is the only singular point of π_ξ valued in D_δ^+. Therefore, $\pi_\xi^{-1}(\overline{I})$ is a one-dimensional curve over $\mathbb{R}$, for which the decomposition

$$\pi_\xi^{-1}(\overline{I}) = \pi_\xi^{-1}(0) \cup \pi_\xi^{-1}(I) \cup \pi_\xi^{-1}(\delta)$$

constitutes a valid Whitney stratification. Our strategy is to examine the following zigzag diagram of inclusion maps:

$$\pi_\xi^{-1}(0) \hookrightarrow \pi_\xi^{-1}(\overline{I}) \hookleftarrow \pi_\xi^{-1}(\delta). \tag{6}$$

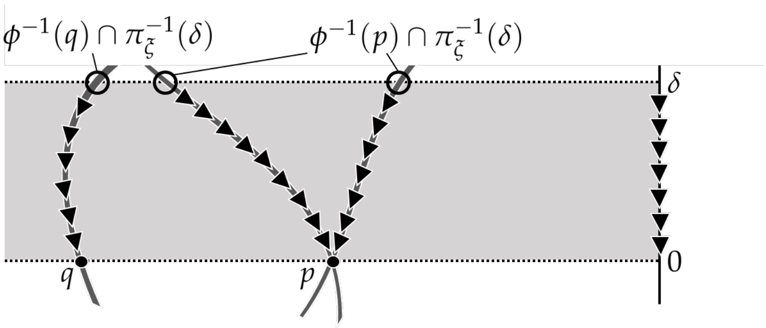
The three spaces involved are illustrated below. Here $\pi_\xi^{-1}(\overline{I})$ is the region of the curve lying within the shaded gray rectangle, while $\pi_\xi^{-1}(0)$ and $\pi_\xi^{-1}(\delta)$ consist of points lying in the intersection of this curve with the bottom and top edges of the rectangle.



Our next result is concerned with the first inclusion from (6).

Proposition 4.3. *The inclusion $\pi_\xi^{-1}(0) \hookrightarrow \pi_\xi^{-1}(\overline{I})$ is a homotopy equivalence, and in particular it admits a homotopy-inverse $\phi : \pi_\xi^{-1}(\overline{I}) \to \pi_\xi^{-1}(0)$.*

Proof. By Lemma 4.2, the unit vector field $U(x) = -1$ on $\overline{I} - \{0\}$ lifts to a vector field V on $\pi_\xi^{-1}(\overline{I}) - \pi_\xi^{-1}(0)$ so that at each point x the differential $(d\pi_\xi)_x$ sends the vector $V(q)$ to -1, as depicted below:



The desired map ϕ is obtained by flowing along the integral curves of V. ■

The second map from our zigzag (6) will be described via the corresponding relative homology group, namely

$$\mathrm{H}_\bullet\left(\pi_\xi^{-1}(\overline{I}), \pi_\xi^{-1}(\delta)\right),$$

where we have implicitly assumed rational coefficients throughout. The following result shows that this group only depends on local data pertaining to the fibers of ϕ over p.

Lemma 4.4. *Let $\phi : \pi_\xi^{-1}(\overline{I}) \to \pi_\xi^{-1}(0)$ be a homotopy inverse to the inclusion (as in the proof of Proposition 4.3). There is an isomorphism of relative homology groups:*

$$\mathrm{H}_\bullet\left(\pi_\xi^{-1}(\overline{I}), \pi_\xi^{-1}(\delta)\right) \simeq \mathrm{H}_\bullet\left(\phi^{-1}(p), \phi^{-1}(p) \cap \pi_\xi^{-1}(\delta)\right).$$

Consequently, the associated Euler characteristics satisfy

$$\chi\left(\pi_\xi^{-1}(\overline{I})\right) - \chi\left(\pi_\xi^{-1}(\delta)\right) = 1 - \chi\left(\phi^{-1}(p) \cap \pi_\xi^{-1}(\delta)\right).$$

Proof. The set $\pi_\xi^{-1}(\overline{I})$ decomposes as a disjoint union

$$\pi_\xi^{-1}(\overline{I}) = \coprod_q \phi^{-1}(q),$$

where q ranges over the points in $\pi_\xi^{-1}(0)$. By the additivity of homology, we have a direct sum decomposition

$$\mathrm{H}_\bullet\left(\pi_\xi^{-1}(\overline{I}), \pi_\xi^{-1}(\delta)\right) = \bigoplus_q \mathrm{H}_\bullet\left(\phi^{-1}(q), \phi^{-1}(q) \cap \pi_\xi^{-1}(\delta)\right).$$

It therefore suffices to show that the summands corresponding to $q \neq p$ are all trivial. Since no such q is a critical point of π_ξ by nondegeneracy of ξ, the vector field V on $\pi_\xi^{-1}(\overline{I}) - \pi_\xi^{-1}(0)$

which was used to construct ϕ in the proof of Proposition 4.3 extends nontrivially through q. Now, by Thom's isotopy lemma (Lemma 4.2) above, the stratified homeomorphism type of $\phi^{-1}(q)\cap\pi_\xi^{-1}(t)$ remains unchanged across all $t\in\overline{I}$, so in particular there is a homeomorphism of pairs

$$\left(\phi^{-1}(q),\phi^{-1}(q)\cap\pi_\xi^{-1}(\delta)\right)\simeq\left([0,\delta],\delta\right),$$

and hence the relative homology is trivial as desired. To extract the statement about the Euler characteristics from the statement about relative homology groups, one uses the observation that $\phi^{-1}(p)$ is homeomorphic to the cone at p over $\phi^{-1}(p)\cap\pi_\xi^{-1}(\delta)$. Since all cones are contractible, we obtain $\chi(\phi^{-1}(p))=1$. ∎

To conclude our proof of Theorem 4.1, we observe that

$$\begin{aligned}
\mathbf{e}_pC &= \#\left\{\pi_\xi^{-1}(\delta)\right\}-\#\left\{\pi_\xi^{-1}(0)\right\}+1 && \text{by (5)}\\
&=\chi\left(\pi_\xi^{-1}(\delta)\right)-\chi\left(\pi_\xi^{-1}(0)\right)+1 && \text{since } \dim_{\mathbb{R}}\pi_\xi^{-1}(t)=0 \text{ for } t\in[0,\delta]\\
&=\chi\left(\pi_\xi^{-1}(\delta)\right)-\chi\left(\pi_\xi^{-1}(\overline{I})\right)+1 && \text{by Proposition 4.3}\\
&=\chi\left(\phi^{-1}(p)\cap\pi_\xi^{-1}(\delta)\right) && \text{by Lemma 4.4}\\
&=\chi(\mathscr{L}_pC) && \text{by Definition 2.3}
\end{aligned}$$

as desired.

Remark 4.5. Here are two observations pertaining to our proof of Theorem 4.1.

1. The argument could have been considerably shortened by employing the sheaf-theoretic language of nearby and vanishing cycles [8, 14]. We have presented a longer and more elementary argument here in order to avoid stranding readers who are unfamiliar with this machinery.
2. Neither Proposition 4.3 nor Lemma 4.4 requires any constraint on the complex dimension of C, and both would work just as well when $\dim_{\mathbb{C}}C>1$. On the other hand, it is only when $\dim_{\mathbb{C}}C=1$ that one obtains $\dim_{\mathbb{R}}(\pi_\xi^{-1}(t))=0$ for t in $[0,\delta]$, and it is a miracle of zero-dimensionality that degree and Euler characteristic coincide. This accident is exploited only once in our argument, namely when transitioning from the first line to the second one in the string of equalities above.

Unfortunately, we do not anticipate any direct relationship between degrees and Euler characteristics of higher-dimensional projective varieties. Thus, this argument does not extend directly to the scenario where our curve C is replaced by a variety Y of dimension >1. In any event, Theorems 3.5 and 4.1 guarantee that all Hilbert–Samuel multiplicity computations can be reduced to Euler characteristic estimation for a finite collection of points. We now turn our attention to inferring such multiplicities from point samples.

5. Estimating multiplicities from finite samples.

The *reach* $\tau_M>0$ of a smooth compact submanifold $M\subset\mathbb{R}^n$ is the smallest radius $r>0$ for which the radius-r normal bundle around M self-intersects. This notion was first introduced by Federer in [5], and it serves as an important measure of the regularity of the embedding $M \hookrightarrow \mathbb{R}^n$. The reciprocal $1/\tau_M$, called the *condition number* of M, features prominently in the homological inference results of Niyogi, Smale, and Weinberger from [18]. These results have been extended to the case where M has a smooth boundary ∂M by Wang and Wang [27]. In this setting, the role of the reach is played by a new parameter

$$\Delta_M := \min\{\tau_M, \tau_{\partial M}, \rho_M\}, \tag{7}$$

where ρ_M is the largest radius $r > 0$ so that at each point x in $(M - \partial M)$ the exponential map $M \to T_x M$ is a diffeomorphism onto its image when restricted to the open ball $\mathbf{B}_r(x)^\circ \cap M$. The following result is [27, Theorem 3.3].

Theorem 5.1. *Let M be a smooth, nonempty k-dimensional submanifold with boundary of $\mathbb{R}^n$. For any radius $r \in (0, \Delta_M/2)$ and probability $\gamma \in (0,1)$, there exists an explicit bound $N_M(r,\gamma)$ satisfying the following property. Any uniformly sampled[3] finite set $S \subset M$ of cardinality larger than $N_M(r,\gamma)$ is $(r/2)$-dense in M with probability exceeding $(1-\gamma)$.*

In other words, we can guarantee with high confidence that every point of M is no more than $r/2$ away from some point of S whenever $\#S > N_M(r,\gamma)$.

Remark 5.2. This bound $N_M(r,\gamma)$ has the form

$$N_M(r,\gamma) = \beta_M(r) \cdot \left[\beta_M\left(\frac{r}{2}\right) + \ln\left(\frac{1}{\gamma}\right)\right], \tag{8}$$

where $\beta_M : \mathbb{R} \to \mathbb{R}_{>0}$ is the function

$$\beta_M(x) := \frac{Vol(M)}{\frac{\cos^k(\theta)}{2^{(k+1)}} \cdot I_y\left(\frac{k+1}{2}, \frac{1}{2}\right) \cdot Vol(\mathbf{B}_x^k)}.$$

Here $Vol(\bullet)$ is standard k-dimensional Lebesgue volume and the auxiliary variables are

$$\theta := \arcsin\left(\frac{x}{4\Delta_M}\right) \quad \text{and} \quad y := 1 - \frac{x^2 \cdot \cos^2(\theta)}{16\Delta_M^2};$$

moreover, $I_y(a,b)$ denotes the regularized incomplete beta function

$$I_y(a,b) := \frac{B_y(a,b)}{B_1(a,b)}, \quad \text{with} \quad B_y(a,b) := \int_0^y t^{a-1}(1-t)^{b-1}dt.$$

And, finally, $\mathbf{B}_x^k$ is the ball of radius x in k-dimensional Euclidean space.

5.1. Setup and parameter choices. Let p be any (not necessarily singular) point on a curve $C \subset \mathbb{P}^n$. Passing to an affine chart of $\mathbb{P}^n$ containing p, we may as well work within $\mathbb{C}^{n+1} \simeq \mathbb{R}^{2n+2}$. In light of this identification, all dimensions of spaces mentioned henceforth are to be understood as dimensions over $\mathbb{R}$ rather than $\mathbb{C}$. Consider any choice of positive

[3] i.e., independent and identically distributed with respect to the uniform measure on M.

radii $\epsilon \gg \delta$ and (unit length) direction ξ so that the complex link of p in C is given by the intersection

$$\mathscr{L}_p = C \cap \mathbf{B}_\epsilon(p) \cap \pi_\xi^{-1}(\delta). \tag{9}$$

By Theorem 4.1, this is a set of (finite) cardinality $\ell := \mathbf{e}_p C$, so we may enumerate its points as $\{x_1, x_2, \ldots, x_\ell\}$. Let $\mu > 0$ be the smallest pairwise Euclidean distance between these points

$$\mu := \min\{\|x_i - x_j\|, \text{ where } 1 \le i \ne j \le \ell\}, \tag{10}$$

and let $\kappa > 0$ be the distance between $\mathscr{L}_p$ and the boundary of the closed ball $\mathbf{B}_\epsilon(p)$:

$$\kappa := \min\{\epsilon - \|p - x_i\|, \text{ where } 1 \le i \le \ell\}. \tag{11}$$

These new distances κ and μ are determined by the initial choices of ϵ, δ, and ξ. We also select a new parameter $\epsilon_0 \in (0, \delta)$, called the *inner radius*. Since ϵ_0 is smaller than δ, the open ball $\mathbf{B}_{\epsilon_0}(p)^\circ$ does not intersect the offset hyperplane $\pi_\xi^{-1}(\delta)$, and hence does not contain any of the points $\{x_1, \ldots, x_\ell\}$.

Proposition 5.3. *The intersection $C' := C \cap [\mathbf{B}_\epsilon(p) - \mathbf{B}_{\epsilon_0}(p)^\circ]$ forms a two-dimensional manifold with boundary embedded within $C_{\text{reg}} \subset \mathbb{R}^{2n+2}$.*

Proof. We recall that the radius ϵ from (9) satisfies the property that the boundary sphere $\partial\mathbf{B}_e(p)$ is transverse to C_{reg} for all $e \in (0, \epsilon]$. Thus, the smooth map $\mathbf{d}_p : \mathbb{R}^{2n+2} \to \mathbb{R}$ given by

$$x \mapsto \|x - p\|^2$$

has no critical points on C_{reg} valued in $(0, \epsilon]$. We know by Lemma 4.2 that the restriction $\mathbf{d}_p : C' \to [\epsilon_0, \epsilon]$ forms a trivial fiber bundle. Since all such e are regular values of $\mathbf{d}$, the implicit function theorem guarantees that each fiber $F_e := C_{\text{reg}} \cap \mathbf{d}_p^{-1}(e)$ is a smooth one-dimensional submanifold of C'. Therefore, the desired result now follows from the fact that C' is diffeomorphic to the product $F \times [\epsilon_0, \epsilon]$ where F is a smooth one-dimensional manifold. ∎

Recalling the fact that $\{x_1, \ldots, x_\ell\} \cap \mathbf{B}_{\epsilon_0}(p)$ is empty since $\epsilon_0 < \delta$, we have

$$C' \cap \pi_\xi^{-1}(\delta) = \{x_1, \ldots, x_\ell\}.$$

An immediate side effect of replacing C' by a finite point sample is that none of the sample points will lie exactly on $\pi_\xi^{-1}(\delta)$. Therefore, we require a final pair of *thickness* parameters $\alpha > \alpha_0 > 0$. These are sufficiently small positive radii for which the following property holds: the set of all points in C that lie within distance α_0 of the offset hyperplane $\pi_\xi^{-1}(\delta)$ is entirely contained within the union of closed radius-α balls around points of the complex link:

$$\left\{y \in C \mid \operatorname{dist}[y, \pi_\xi^{-1}(\delta)] < \alpha_0\right\} \subset \bigcup_{i=1}^{\ell} \mathbf{B}_\alpha(x_i). \tag{12}$$

The optimal radius α for a given thickness α_0 depends on the curvature of C at each x_i as well as the angles θ_i between the tangent spaces $T_{x_i}C$ and the offset hyperplane $\pi_\xi^{-1}(\delta)$. By sufficiently small here we mean

$$\alpha_0 < \alpha < \min\left\{(\epsilon-\delta), (\delta-\epsilon_0), \mu/4, \kappa, \Delta_{C'}/2\right\}. \tag{13}$$

This inequality encodes all of the geometric constraints required for our inference result. As atonement for introducing this deluge of parameters, we remind the reader that ϵ and δ were fixed in (9), while ϵ_0 is the inner radius used in Proposition 5.3; the quantities μ and κ are described in (10) and (11), respectively, and $\Delta_{C'}$ is from (7). In Figure 2, we have illustrated the typical local picture of C near p in the case where $\mathbf{e}_pC = 3$.

The first four terms within the minimum on the right-hand side of (13) are designed simply to ensure that balls of radius α around each of the x_i are well-separated from each other and fully contained within the annulus $[\mathbf{B}_\epsilon(p) - \mathbf{B}_{\epsilon_0}(p)^\circ]$; the final term is required for applying Theorem 5.1.

5.2. Inferring multiplicities with high confidence. Here we prove Theorem (**C**) from the introduction. The parameters encountered in its statement below were chosen in the previous subsection.

Theorem 5.4. *Let $S \subset \mathbb{R}^{2n+2}$ be a finite set of points sampled uniformly from the intersection $C' = C \cap [\mathbf{B}_\epsilon(p) - \mathbf{B}_{\epsilon_0}(p)^\circ]$. For any $\gamma \in (0,1)$, if the cardinality $\#S$ exceeds the bound $N_{C'}(\alpha,\gamma)$ from (8), then the following holds with probability exceeding $(1-\gamma)$: the set*

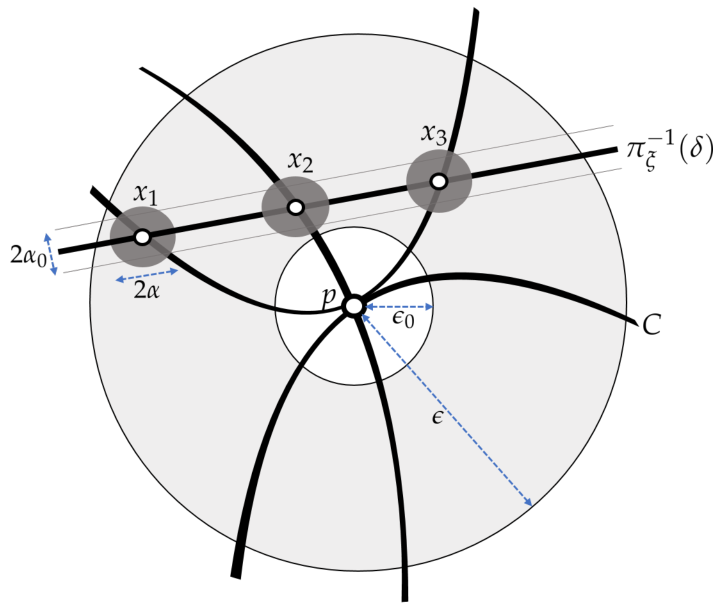


Figure 2. *The typical local picture near p when $\mathbf{e}_pC = 3$. The parameters $\epsilon_0 < \epsilon$ define the annulus of interest around p. The parameter α_0 serves to thicken the offset hyperplane $\pi_\xi^{-1}(\delta)$ while the radius α thickens the three points x_i of the complex link. The inequalities constraining α ensure, among other things, that the three α-balls around the x_i lie within the annulus.*

$$S' := \left\{ y \in S \mid \mathrm{dist}[y, \pi_\xi^{-1}(\delta)] < \alpha_0 \right\}$$

consists of exactly $\ell = \mathbf{e}_p C$ nonempty point clusters, each of diameter at most 2α, with the distance between distinct clusters exceeding $\mu - 2\alpha$.

Proof. The intersection C' is an embedded two-dimensional submanifold with boundary of $\mathbb{R}^{2n+2}$ by Proposition 5.3, and $\alpha < \Delta_{C'}/2$ holds by (13). Thus, we may safely apply Theorem 5.1 to conclude that the set S is α-dense in C' with probability exceeding $(1-\gamma)$. We will assume throughout the remainder of the argument that this density holds.

Since the inner radius ϵ_0 is smaller than δ, all points of the complex link $\{x_1, \ldots, x_\ell\}$ lie in C'. Let B_i denote the closed ball of radius α around each x_i. The inequalities which involve $\epsilon, \epsilon_0, \delta$, and κ in (13) guarantee that each B_i is entirely contained within the annulus $[\mathbf{B}_\epsilon(p) - \mathbf{B}_{\epsilon_0}(p)]$. It follows from the α-density of S in C' that the intersections $S_i := B_i \cap S$ are all nonempty; and, moreover, the diameter of each S_i is no larger than the diameter 2α of B_i. We claim that these S_i form the ℓ desired point clusters of S'.

To establish the claim, note from (12) that every point of S' must lie in one of the B_i, whence $S' = \bigcup_{i=1}^{\ell} S_i$. Thus, it remains to show that the S_i are separated from each other by a distance larger than $\mu - 2\alpha$. To this end, note from (10) that the points x_i and x_j are separated by distance at least μ whenever $i \neq j$. Thus, points in distinct B_i and B_j are at least $\mu - 2\alpha$ apart from each other, as desired. ■

We know from (13) that 2α is smaller than $\mu - 2\alpha$, so the desired number $\ell = \mathbf{e}_p C$ can be determined with high confidence by clustering together points of S' which lie within 2α of each other and then counting the clusters.

Acknowledgments. Mark Goresky kindly shared an advance copy of his survey *Morse theory, stratifications and sheaves* [8] with us; that paper served as our Polaris while we navigated the formidable waters surrounding these topics. We are grateful to Heather Harrington, Kate Turner, and Yossi Bokor for facilitating this work in its early stages. We also thank the Sydney Mathematics Research Institute (SMRI) at the University of Sydney for their generous hospitality. This paper benefited enormously from the suggestions and corrections provided by the two anonymous referees.

REFERENCES

[1] P. Breiding, S. Kališnik, B. Sturmfels, and M. Weinstein, *Learning algebraic varieties from samples*, Rev. Mat. Complut., 31 (2018), pp. 545–593.

[2] D. A. Cox, J. Little, and D. O'Shea, *Using Algebraic Geometry*, Grad. Texts Math. 185, Springer, New York, 2005.

[3] E. Dufresne, P. Edwards, H. Harrington, and J. Hauenstein, *Sampling real algebraic varieties for topological data analysis*, in Proceedings of the 18th IEEE International Conference on Machine Learning and Applications (ICMLA), 2019, pp. 1531–1536.

[4] D. Eisenbud and J. Harris, 3264 *and All That: A Second Course in Algebraic Geometry*, Cambridge University Press, Cambridge, UK, 2016.

[5] H. Federer, *Curvature measures*, Trans. Amer. Math. Soc., 93 (1959), pp. 418–491.

[6] W. Fulton, *Introduction to Intersection Theory in Algebraic Geometry*, CBMS Regional Conf. Ser. in Math. 54, American Mathematical Society, Providence, RI, 1984.

[7] W. Fulton, *Intersection Theory*, 2nd ed., Springer-Verlag, Berlin, 1998.

[8] M. Goresky, *Morse theory, stratifications and sheaves*, in Handbook of Geometry and Topology of Singularities, J. L. Cisneros-Molina, L. D. Tráng, and J. Seade, eds., Springer, Cham, 2020, pp. 275–319.
[9] M. Goresky and R. MacPherson, *Stratified Morse Theory*, Springer-Verlag, Berlin, 1988.
[10] C. Harris, *Computing Segre classes in arbitrary projective varieties*, J. Symbolic Comput., 82 (2017), pp. 26–37.
[11] C. Harris and M. Helmer, *Segre class computation and practical applications*, Math. Comp., 89 (2020), pp. 465–491.
[12] R. Hartshorne, *Algebraic Geometry*, Grad. Texts. Math. 52, Springer, New York, 2013.
[13] M. Helmer and V. Nanda, *Conormal spaces and Whitney stratifications*, Found. Comput. Math., (2022), https://doi.org/10.1007/s10208-022-09574-8.
[14] M. Kashiwara and P. Schapira, *Sheaves on Manifolds*, Grundlehren Math. Wiss. 292, Springer-Verlag, Berlin, 1990.
[15] D. Massey, *Characteristic cycles and the relative local Euler obstruction*, in A Panorama of Singularities: Conference in Celebration of Lê Dũng Tráng's 70th Birthday, Universidad de Sevilla, Sevilla, Spain, 2020, pp. 137–156.
[16] J. Mather, *Notes on topological stability*, Bull. Amer. Math. Soc., 49 (2012), pp. 475–506.
[17] G. Moroz, F. Rouiller, D. Chablat, and P. Wenger, *On the determination of cusp points of* 3*-RPR parallel manipulators*, Mech. Mach. Theory, 45 (2010), pp. 1555–1567.
[18] P. Niyogi, S. Smale, and S. Weinberger, *Finding the homology of submanifolds with high confidence from random samples*, Discrete Comput. Geom., 39 (2008), pp. 419–441.
[19] S. Piipponen, T. Arponen, and J. Tuomela, *Classification of singularities in kinematics of mechanisms*, in Computational Kinematics, Springer, Dordrecht, 2014, pp. 41–48.
[20] F. Rouillier, *Solving zero-dimensional systems through the rational univariate representation*, Appl. Algebra Eng. Commun. Comput., 9 (1999), pp. 433–461.
[21] P. Samuel, *Méthodes d'algèbre abstraite en géométrie algébrique*, *Ergebnisse der Mathematik*, Springer-Verlag, Berlin, Heidelberg, 1955.
[22] J.-P. Serre, *Local Algebra*, Springer-Verlag, Berlin, 2012.
[23] I. R. Shafarevich, *Basic Algebraic Geometry*, Vol. 2, Springer-Verlag, Berlin, 1994.
[24] A. J. Sommese and C. W. Wampler, *The Numerical Solution of Systems of Polynomials Arising in Engineering and Science*, World Scientific, Hackensack, NJ, 2005.
[25] B. J. Stolz, J. Tanner, H. A. Harrington, and V. Nanda, *Geometric anomaly detection in data*, Proc. Natl. Acad. Sci. USA, 117 (2020), pp. 19664–19669.
[26] C. W. Wampler and A. J. Sommese, *Numerical algebraic geometry and algebraic kinematics*, Acta Numer., 20 (2011), pp. 469–567.
[27] Y. Wang and B. Wang, *Topological inference of manifolds with boundary*, Comput. Geom., 88 (2020), 101606.
[28] H. Whitney, *Tangents to an analytic variety*, Ann. of Math. (2), 81 (1965), pp. 496–549.
[29] M. Zein, P. Wenger, and D. Chablat, *Singular curves in the joint space and cusp points of* 3*-RPR parallel manipulators*, Robotica, 25 (2007), pp. 717–724.