The Euclidean Distance Degree of an Algebraic Variety

Bernd Sturmfels
UC Berkeley and MPI Bonn

joint work with
Jan Draisma, Emil Horobeţ,
Giorgio Ottaviani, and Rekha Thomas

Getting Close to Varieties

Many models in the sciences and engineering are the real solutions to systems of polynomial equations in several unknowns.

Such a set is an algebraic variety $X \subset \mathbb{R}^n$.

Given X, consider the following optimization problem:

for any data point $u \in \mathbb{R}^n$, find $x \in X$ that minimizes the squared Euclidean distance $d_u(x) = \sum_{i=1}^{n}(u_i - x_i)^2$.

Getting Close to Varieties

Many models in the sciences and engineering are the real solutions to systems of polynomial equations in several unknowns.

Such a set is an algebraic variety $X \subset \mathbb{R}^n$.

Given X, consider the following optimization problem:
for any data point $u \in \mathbb{R}^n$, find $x \in X$ that minimizes the squared Euclidean distance $d_u(x) = \sum_{i=1}^{n}(u_i - x_i)^2$.

What can be said about the *algebraic function*

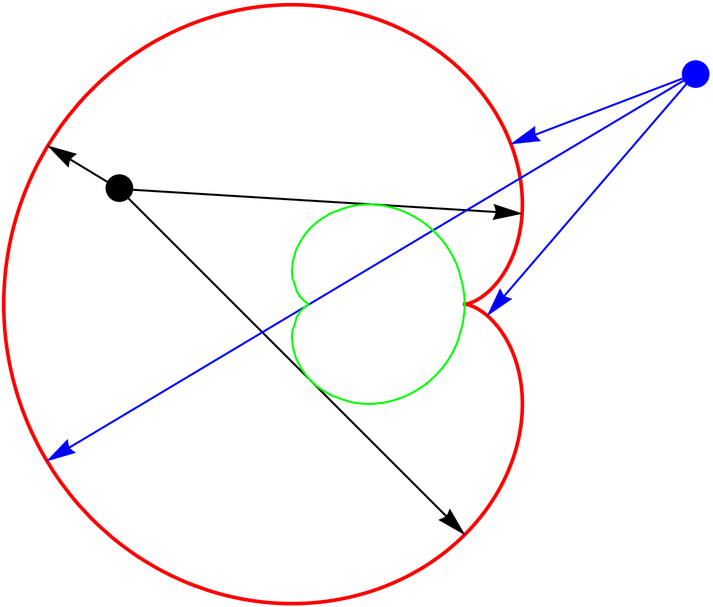
$$u \;\mapsto\; x(u)$$

from the data to the optimal solution?

Its branches are given by the complex critical points for generic u.

Their number is the Euclidean distance degree,
or short, the ED degree, of the variety X.

Logo



Plane Curves

Fix a polynomial $f(x, y)$ of degree d and consider the curve

$$X \ = \ \{(x, y) \in \mathbb{R}^2 \, : \, f(x, y) = 0\}.$$

Given a data point (u, v) we wish to find (x, y) on X
such that $(u - x, v - y)$ is parallel to the gradient of f.

Must solve two equations of degree d in two unknowns:

$$f(x, y) \quad = \quad \det \begin{pmatrix} u - x & v - y \\ \partial f/\partial x & \partial f/\partial y \end{pmatrix} \quad = \quad 0$$

By Bézout's Theorem, we expect d^2 complex solutions (x, y).

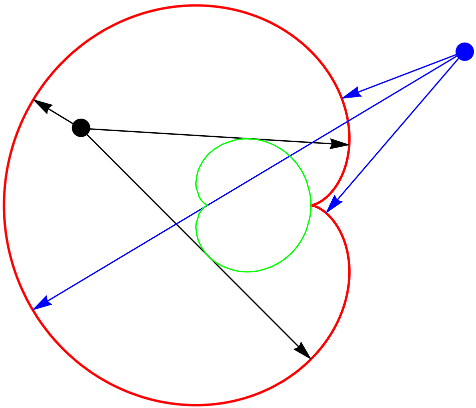
Proposition

A general plane curve X of degree d has $\mathrm{EDdegree}(X) = d^2$.

The Cardioid

The cardioid is a special curve of degree 4. Its ED degree equals **3**.

$$X \;=\; \left\{(x,y) \in \mathbb{R}^2 \,:\, (x^2+y^2+x)^2 = x^2+y^2\right\}.$$



The inner cardioid is the *evolute* or *ED discriminant*. It is given by

$$27u^4 + 54u^2v^2 + 27v^4 + 54u^3 + 54uv^2 + 36u^2 + 9v^2 + 8u \;=\; 0.$$

Linear Regression

If X is a linear subspace of $\mathbb{R}^n$ then

$$\mathrm{EDdegree}(X) = 1.$$

Which *non-linear* varieties do arise in applications?

- Control Theory
- Geometric Modeling
- Computer Vision
- Tensor Decomposition
- Structured Low Rank Approximation
- .....

In many cases, X is given by *homogeneous* polynomials, so X is a cone. View it as a *projective variety* in $\mathbb{P}^{n-1}$.

Ideals

Let $I_X = \langle f_1, \ldots, f_s \rangle \subset \mathbb{R}[x_1, \ldots, x_n]$ be the ideal of X and $J(f)$ its $s \times n$ Jacobian matrix. The *singular locus* X_{sing} is defined by

$$I_{X_{\mathrm{sing}}} \;=\; I_X + \big\langle c \times c\text{-minors of } J(f) \big\rangle, \qquad \text{where } c = \mathrm{codim}(X).$$

The *critical ideal* for $u \in \mathbb{R}^n$ is

$$\left(I_X + \left\langle (c{+}1) \times (c{+}1)\text{-minors of } \begin{pmatrix} u - x \\ J(f) \end{pmatrix} \right\rangle \right) : \left(I_{X_{\mathrm{sing}}}\right)^{\infty}$$

Lemma

For generic $u \in \mathbb{R}^n$, the function d_u has finitely many critical points on the manifold $X \backslash X_{\mathrm{sing}}$, namely the zeros of the critical ideal.

$\longrightarrow \; \mathrm{EDdegree}(X)$

Ideals

Let $I_X = \langle f_1, \ldots, f_s \rangle \subset \mathbb{R}[x_1, \ldots, x_n]$ be the ideal of X and $J(f)$ its $s \times n$ Jacobian matrix. The *singular locus* X_{sing} is defined by

$$I_{X_{\mathrm{sing}}} \;=\; I_X + \big\langle c \times c\text{-minors of } J(f) \big\rangle, \qquad \text{where } c = \mathrm{codim}(X).$$

The *critical ideal* for $u \in \mathbb{R}^n$ is

$$\left(I_X + \left\langle (c{+}1) \times (c{+}1)\text{-minors of } \begin{pmatrix} u - x \\ J(f) \end{pmatrix} \right\rangle \right) : \big(I_{X_{\mathrm{sing}}}\big)^{\infty}$$

Lemma

For generic $u \in \mathbb{R}^n$, the function d_u has finitely many critical points on the manifold $X \backslash X_{\mathrm{sing}}$, namely the zeros of the critical ideal.

$\longrightarrow \ \mathrm{EDdegree}(X)$

If $f_1, \ldots, f_s$ are homogeneous, so that $X \subset \mathbb{P}^{n-1}$, we use instead

$$\left(I_X + \left\langle (c{+}2) \times (c{+}2)\text{-minors of } \begin{pmatrix} u \\ x \\ J(f) \end{pmatrix} \right\rangle \right) : \big(I_{X_{\mathrm{sing}}} \cdot \langle x_1^2 + \cdots + x_n^2 \rangle \big)^{\infty}$$

Bounds

Proposition

Let $X \subset \mathbb{R}^n$ be defined by polynomials $f_1, f_2, \ldots, f_c, \ldots$ of degrees $d_1 \geq d_2 \geq \cdots \geq d_c \geq \cdots$. If $\operatorname{codim}(X) = c$ then

$$\mathrm{EDdegree}(X) \;\leq$$

$$d_1 d_2 \cdots d_c \cdot \sum_{i_1+i_2+\cdots+i_c \leq n-c} (d_1-1)^{i_1}(d_2-1)^{i_2}\cdots(d_c-1)^{i_c}.$$

Equality holds when $f_1, f_2, \ldots, f_c$ are generic.

Example

If X is cut out by c quadratic polynomials in $\mathbb{R}^n$ then its ED degree is at most $2^c\binom{n}{c}$.

Similar bounds are available for projective varieties $X \subset \mathbb{P}^{n-1}$.

Singular Value Decomposition

Fix positive integers $r \leq s \leq t$ and $n = st$. Given an arbitrary $s \times t$-matrix U, we seek a matrix of rank r that is closest to U. Here X is the determinantal variety of $s \times t$-matrices of rank $\leq r$.

Proposition

$$\mathrm{EDdegree}(X) \;=\; \binom{s}{r}.$$

Proof. Compute the singular value decomposition

$$U \;=\; T_1 \cdot \mathrm{diag}(\sigma_1, \sigma_2, \ldots, \sigma_s) \cdot T_2.$$

with $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_s$. By the Eckart-Young Theorem,

$$U^* \;=\; T_1 \cdot \mathrm{diag}(\sigma_1, \ldots, \sigma_r, 0, \ldots, 0) \cdot T_2$$

is closest rank r matrix to U. All critical points are given by r-element subsets of $\{\sigma_1, \ldots, \sigma_s\}$.

Closest Symmetric Matrix

For symmetric $U = (U_{ij})$, consider two unconstrained formulations:

$$\mathrm{Min}_t \sum_{i=1}^{s} \sum_{j=1}^{s} \Big(U_{ij} - \sum_{k=1}^{r} t_{ik} t_{kj} \Big)^2 \quad \textbf{or} \quad \mathrm{Min}_t \sum_{1 \le i \le j \le s} \Big(U_{ij} - \sum_{k=1}^{r} t_{ik} t_{kj} \Big)^2.$$

Eckart-Young applies only in the first case:

$$\mathrm{EDdegree}(X) = \binom{s}{r} \qquad \textbf{or} \qquad \mathrm{EDdegree}(X) \gg \binom{s}{r}.$$

Here X is the variety of symmetric $s \times s$-matrices of rank $\le r$.

Closest Symmetric Matrix

For symmetric $U = (U_{ij})$, consider two unconstrained formulations:

$$\mathrm{Min}_t \sum_{i=1}^{s} \sum_{j=1}^{s} \Big(U_{ij} - \sum_{k=1}^{r} t_{ik} t_{kj} \Big)^2 \quad \textbf{or} \quad \mathrm{Min}_t \sum_{1 \leq i \leq j \leq s} \Big(U_{ij} - \sum_{k=1}^{r} t_{ik} t_{kj} \Big)^2.$$

Eckart-Young applies only in the first case:

$$\mathrm{EDdegree}(X) = \binom{s}{r} \qquad \textbf{or} \qquad \mathrm{EDdegree}(X) \gg \binom{s}{r}.$$

Here X is the variety of symmetric $s \times s$-matrices of rank $\leq r$.

For 3×3-matrices with $r = 1, 2$ we have

$$\mathrm{EDdegree}(X) = 3 \qquad \textbf{or} \qquad \mathrm{EDdegree}(X) = 13.$$

Fixing the Euclidean metric on $\mathbb{R}^6$, put rank constraints on either

$$\begin{pmatrix} \sqrt{2}x_{11} & x_{12} & x_{13} \\ x_{12} & \sqrt{2}x_{22} & x_{23} \\ x_{13} & x_{23} & \sqrt{2}x_{33} \end{pmatrix} \quad \textbf{or} \quad \begin{pmatrix} x_{11} & x_{12} & x_{13} \\ x_{12} & x_{22} & x_{23} \\ x_{13} & x_{23} & x_{33} \end{pmatrix}$$

Critical Formations on the Line

d'après [Anderson-Helmke 2013]

Let X denote the variety in $\mathbb{R}^{\binom{p}{2}}$ with parametric representation

$$d_{ij} \ = \ (z_i - z_j)^2 \quad \text{for} \quad 1 \leq i \leq j \leq p.$$

The points in X record the squared distances among p interacting agents with coordinates $z_1, z_2, \ldots, z_p$ on the real line. The ideal I_X is generated by the 2×2-minors of the *Cayley-Menger matrix*

$$\begin{bmatrix} 2d_{1p} & d_{1p}+d_{2p}-d_{12} & d_{1p}+d_{3p}-d_{13} & \cdots & d_{1p}+d_{p-1,p}-d_{1,p-1} \\ d_{1p}+d_{2p}-d_{12} & 2d_{2p} & d_{2p}+d_{3p}-d_{23} & \cdots & d_{2p}+d_{p-1,p}-d_{2,p-1} \\ d_{1p}+d_{3p}-d_{13} & d_{2p}+d_{3p}-d_{23} & 2d_{3p} & \cdots & d_{3p}+d_{p-1,p}-d_{3,p-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ d_{1p}+d_{p-1,p}-d_{1,p-1} & d_{2p}+d_{p-1,p}-d_{2,p-1} & d_{3p}+d_{p-1,p}-d_{3,p-1} & \cdots & 2d_{p-1,p} \end{bmatrix}$$

Theorem

The ED degree of the Cayley-Menger variety X equals

$$\mathrm{EMdegree}(X) \ = \ \begin{cases} \frac{3^{p-1}-1}{2} & \textit{if } \ p \equiv 1,2 \mod 3 \\ \frac{3^{p-1}-1}{2} - \frac{p!}{3((p/3)!)^3} & \textit{if } \ p \equiv 0 \mod 3 \end{cases}$$

Hurwitz Stability

A univariate polynomial with real coefficients,

$$x(t) \;=\; x_0 t^n + x_1 t^{n-1} + x_2 t^{n-2} + \cdots + x_{n-1} t + x_n,$$

is *stable* if each of its n complex zeros has negative real part.
Can express this using Hurwitz determinants

$$\bar{\Gamma}_5 \;=\; \frac{1}{x_5} \cdot \det \begin{pmatrix} x_1 & x_3 & x_5 & 0 & 0 \\ x_0 & x_2 & x_4 & 0 & 0 \\ 0 & x_1 & x_3 & x_5 & 0 \\ 0 & x_0 & x_2 & x_4 & 0 \\ 0 & 0 & x_1 & x_3 & x_5 \end{pmatrix}.$$

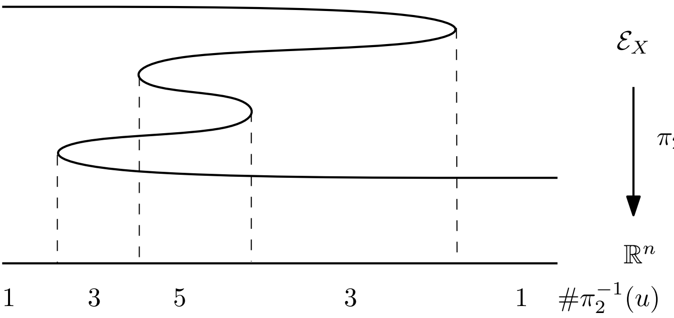
Theorem

The ED degrees of the Hurwitz determinants are

	$\mathrm{EDdegree}(\Gamma_n)$	$\mathrm{EDdegree}(\bar{\Gamma}_n)$
$n = 2m+1$	$8m-3$	$4m-2$
$n = 2m$	$4m-3$	$8m-6$

Here $\Gamma_n = \bar{\Gamma}_n\big|_{x_0=1}$

Average ED Degree



Equip data space $\mathbb{R}^n$ with a probability measure ω. Taking the standard Gaussian centered at 0 is natural when X is a cone:

$$\omega \;=\; \frac{1}{(2\pi)^{n/2}} e^{-||x||^2/2}\, dx_1 \wedge \cdots \wedge dx_n.$$

The *expected number* of critical points of d_u is

$$\mathrm{aEDdegree}(X, \omega) \;:=\; \int_{\mathbb{R}^n} \#\{\text{real critical points of } d_u \text{ on } X\} \cdot |\omega|.$$

Can compute this integral in some interesting cases.

Tables of Numbers

Hurwitz Determinants:

n	EDdegree(Γ_n)	EDdegree($\overline{\Gamma}_n$)	aEDdegree(Γ_n)	aEDdegree($\overline{\Gamma}_n$)
3	5	2	1.162...	2
4	5	10	1.883...	2.068...
5	13	6	2.142...	3.052...
6	9	18	2.416...	3.53...
7	21	10	2.66...	3.742...

ED degree can go up or down when replacing an affine variety by its projective closure. Our theory explains this ....

Important Application: Tensors of Rank One

Format	aEDdegree	EDdegree
$2 \times 2 \times 2$	4.2891...	6
$2\times2\times2\times2$	11.0647...	24
$2 \times 2 \times n,\ \ n \geq 3$	5.6038...	8
$2 \times 3 \times 3$	8.8402...	15
$2 \times 3 \times n,\ \ n \geq 4$	10.3725...	18
$3 \times 3 \times 3$	16.0196...	37
$3 \times 3 \times 4$	21.2651...	55
$3 \times 3 \times n,\ \ n \geq 5$	23.0552...	61

Duality

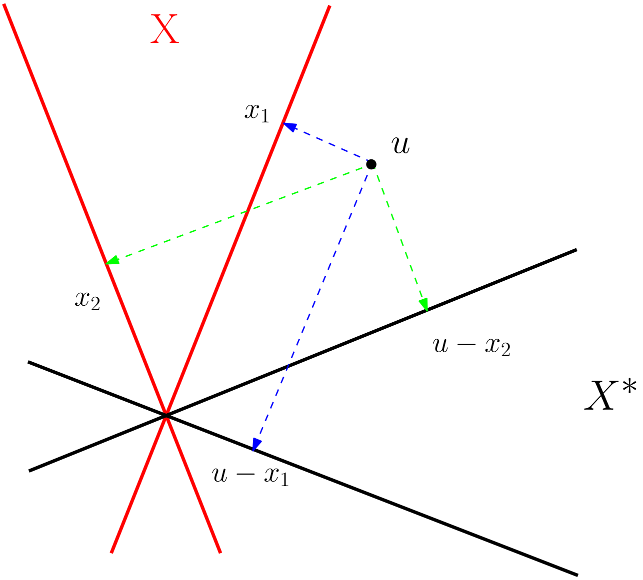


Figure: Bijection between critical points on X and critical points on X^*.

Duality

If X is a cone in $\mathbb{R}^n$ then its dual variety is

$$X^* \ := \ \overline{\{y \in \mathbb{R}^n \mid \exists x \in X \backslash X_{\mathrm{sing}} : y \perp T_x X\}}.$$

Theorem

Fix generic data $u \in \mathbb{R}^n$. The map $x \mapsto u - x$ gives a bijection from critical points of d_u on X to critical points of d_u on X^*, so

$$\mathrm{EDdegree}(X) \ = \ \mathrm{EDdegree}(X^*)$$

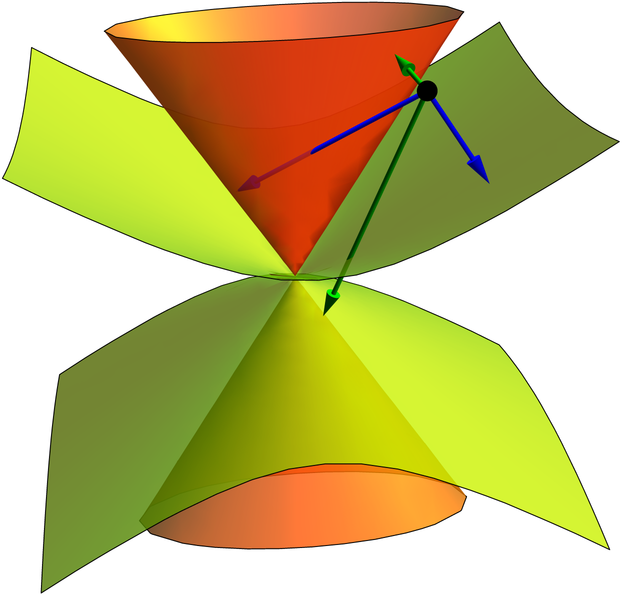
The map is proximity-reversing: the closer a real critical point x is to the data u, the further $u - x$ is from u.

Punchline: Solve the equation $x + y = u$ on the conormal variety.

Corollary

$\mathrm{EDdegree}(X)$ *is the sum of the polar classes of X, provided the conormal variety is disjoint from the diagonal in $\mathbb{P}^{n-1} \times \mathbb{P}^{n-1}$.*

Duality



Symmetric Matrices

If $\quad X = \{$ symmetric $s \times s$-matrices x of rank $\leq r\}$
then $\; X^* = \{$ symmetric $s \times s$-matrices y of rank $\leq s - r\}$.

Their conormal variety is defined by minors of x and y and entries of the matrix product xy.

Must solve $x + y = u$.

The polar classes give the *algebraic degree of semidefinite programming*, studied by von Bothmer, Nie, Ranestad, St.

Use package `Schubert2` in `Macaulay2` to find these values for $\mathrm{EDdegree}(X)$:

$s =$	2	3	4	5	6	7
$r = 1$	4	13	40	121	364	1093
$r = 2$		13	122	1042	8683	72271
$r = 3$			40	1042	23544	510835
$r = 4$				121	8683	510835
$r = 5$					364	72271
$r = 6$						1093

Chern Class Formula

Theorem

Let X be a smooth irreducible variety of dimension m in $\mathbb{P}^{n-1}$. If X is transversal to the isotropic quadric $Q = V(x_1^2 + \cdots + x_n^2)$ then

$$\mathrm{EDdegree}(X) \;=\; \sum_{i=0}^{m} (-1)^i \cdot (2^{m+1-i} - 1) \cdot \deg(c_i(X)).$$

Corollary

Here, if X is a **curve** *of degree d and genus g then*

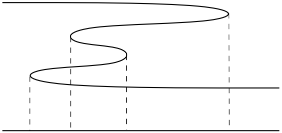
$$\mathrm{EDdegree}(X) \;=\; 3d + 2g - 2.$$

Corollary

Here, if X is **toric** *and V_j is the sum of the normalized volumes of all j-faces of the simple polytope P associated with X, then*

$$\mathrm{EDdegree}(X) \;=\; \sum_{j=0}^{m} (-1)^{m-j} \cdot (2^{j+1} - 1) \cdot V_j.$$

The ED Discriminant



is the variety in data space where two critical points come together.

Studied by [Catanese-Trifogli 2000]

Example

The quadric $X = V(x_0x_3 - 2x_1x_2) \subset \mathbb{P}^3$ has ED degree 6. Its ED discriminant Σ_X is a polynomial of degree 12 with 119 terms:

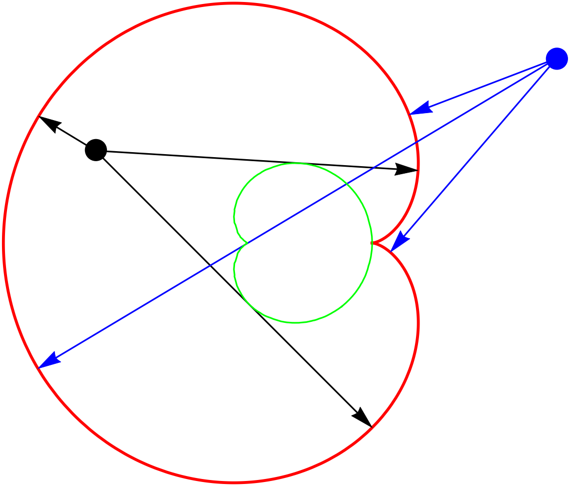
$$65536u_0^{12} + 835584u_0^{10}u_1^2 - 835584u_0^{10}u_3^2 + 9707520u_0^9u_1u_2u_3$$
$$+3747840u_0^8u_1^4 - 7294464u_0^8u_1^2u_2^2 + \cdots + 835584u_2^2u_3^{10} + 65536u_3^{12}.$$

Theorem (Trifogli 1998)

If X is a general hypersurface of degree d in $\mathbb{P}^n$ then

$$\mathrm{degree}(\Sigma_X) \;=\; d(n-1)(d-1)^{n-1} + 2d(d-1)\frac{(d-1)^{n-1}-1}{d-2}.$$

Conclusion



Optimization and Algebraic Geometry can be Friends.
All you need is an **Ideal**.

Epilogue

Chapter 1 in the 1932 *Anschauliche Geometrie* of Hilbert and Cohn-Vossen begins with: *The Simplest Curves and Surfaces*.

The first section, *Plane curves*, starts like this:

- The simplest plane curve is the line.
- Next comes the circle.
- Thereafter comes the parabola.
- And, finally, we get to the ellipse.

Why are these the simplest curves? And **why** in this order?

Epilogue

Chapter 1 in the 1932 *Anschauliche Geometrie* of Hilbert and Cohn-Vossen begins with: *The Simplest Curves and Surfaces*.

The first section, *Plane curves*, starts like this:

- The simplest plane curve is the line.
- Next comes the circle.
- Thereafter comes the parabola.
- And, finally, we get to the ellipse.

Why are these the simplest curves? And **why** in this order?

- The line has ED degree **1**.
- The circle has ED degree **2**.
- The parabola has ED degree **3**.
- The ellipse has ED degree **4**.