CHAPTER 1

Elimination Theory

Elimination theory has important roles to play in both algebraic geometry and symbolic computation. The first two sections of the chapter take a historical approach so that you can see how elimination theory developed in the last 300 years. In the third section Carlos D'Andrea will survey recent developments.

1.1. Elimination Theory in the 18th and 19th Centuries

Algebra was well-established in Europe by the end of the 17th century. The full story of how this happened and what followed in the subsequent two centuries is far beyond the scope of this volume. Readers interested in the details should consult [**17**] (for algebra in general) and [**287**] (for elimination theory). We will instead focus on some vignettes that illustrate the evolution of elimination theory.

Newton and Tschirnhaus. In spite of the title of the section, we begin in the 17th century with extracts from Newton and Tschirnhaus. In a manuscript written sometime between 1666 and 1671, Newton writes:

> *Datis duabus curvis invenire puncta intersectionis.* this is rather a principle then a probleme. But rather propounded of y$^{\mathrm{e}}$ Algebraicall then geometricall solutions & y$^{\mathrm{t}}$ is done by eliminating one of the two unknown quantitys out of y$^{\mathrm{e}}$ equations. From whence it will appeare y$^{\mathrm{t}}$ there are soe many cut points as the rectangle of the curves dimensions. [**282**, p. 177]

The Latin says "Given two curves, to find their points of intersection." The English is archaic, but once you realize that "y$^{\mathrm{e}}$ = the" and "y$^{\mathrm{t}}$ = that", Newton's meaning is clear:

- He solves the geometric problem by working algebraically.
- The algebra is "done by eliminating one of the two unknown quantitys out of y$^{\mathrm{e}}$ equations."
- The degree of the resulting equation is the product of the degrees of the curves ("the rectangle of the curves dimensions").

This is a clear statement of Bézout's Theorem, along with the strategy of reducing to a single variable that was to dominate the early history of elimination theory.

Here is an application of Newton's "principle."

Example 1.1. Newton counts the number of tangent lines to a curve that go through a given point:

> a line drawn from a given point may touch a curve of 2 dimensions in 1×2 points, of 3 in 2×3 points, of 4 in 3×4 points, &c. [**282**, p. 179]

We explain this as follows. Let $U(x,y,z)=0$ define a smooth curve of degree of n in the projective plane $\mathbb{P}^2$. If the given point is $P=(\alpha,\beta,\gamma)$, then any solution of

$$U = 0$$
$$\alpha\frac{\partial U}{\partial x}+\beta\frac{\partial U}{\partial y}+\gamma\frac{\partial U}{\partial z}=0$$

is a point on the curve whose tangent line contains P. These curves have degrees n and $n-1$, giving $n(n-1)$ tangent lines by Bézout's Theorem. This explains the products 1×2, 2×3 and 3×4 in the quote. ◁▷

Another example of 17th century elimination can be found in the work of Tschirnhaus. In 1683, he published a paper on roots of a polynomial equation [**348**]. Given $y^3-qy-r=0$, Tschirnhaus introduces a new variable z via the *Tschirnhaus transformation* $y^2=by+z+a$, where a and b are unknown coefficients. He writes:

> following from this let there be a third equation (by proceeding according to the recognized rules of analysis) in which the quantity y is absent, and we shall obtain
>
> $$z^3+(3a-2q)z^2+(3a^2-4qa+q^2-qb^2+3rb)z+\cdots=0.$$

This leads to a solution of $y^3-qy-r=0$ as follows:

- Pick a to make the coefficient of z^2 vanish.
- Use the quadratic formula to pick b to make the coefficient of z vanish.
- The equation reduces to $z^3=$ constant, so z is a cube root.
- Solving $y^2=by+z+a$ for y gives a root of our original cubic.

Tschirnhaus's approach is based on elimination ("in which the quantity y is absent") using the "recognized rules of analysis."

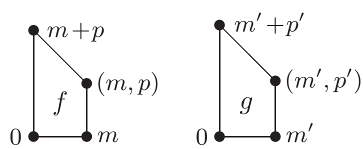
These examples show that elimination theory was an important part of algebra in the 17th century.

Cramer and Bézout. By the middle of the 18th century, Bézout's Theorem was well-known. Here is Cramer writing in 1750:

> If one has two variables and two indeterminate equations ... of which one is of order m and the other of order n, when, by means of these equations, one expels one of these variables, the one that remains has, in the final equation ... at most mn dimensions. [**116**]

In his 1764 paper *Sur le degré des équations résultantes de l'évanouissement des inconnues* [**27**], Bézout uses the terms "final equation" and "resultant equation" for the result of the elimination. The first main result of [**27**] is a version of Bézout's Theorem that draws on earlier work of Euler.

Example 1.2. Let $f(x,y)$ and $g(x,y)$ have Newton polytopes:



Thus f has total degree $m+p$ in x,y, but its degree in x is bounded by m, with a similar story for g.

To eliminate x from $f(x,y)=g(x,y)=0$, Bézout multiplies f,g by polynomials G,H of degree $m'-1,m-1$ in x with unknown coefficients. Then equating the coefficients of powers of x in the equation

$$Gf + Hg = h(y) \tag{1.1}$$

gives a linear system of $m + m'$ unknowns in $m + m'$ variables with a familiar coefficent matrix. We illustrate this for $m = m' = 2$, where

$$f(x, y) = A(y)x^2 + B(y)x + C(y),\ g(x, y) = A'(y)x^2 + B'(y)x + C'(y),$$

and $\deg(A(y)) \le p$, $\deg(B(y)) \le p+1$, $\deg(C(y)) \le p+2$, with similar bounds for the degrees of $A'(y), B'(y), C'(y)$. If we set

$$G = Mx + N,\ H = M'x + N'$$

for unknown coefficients M, N, M', N', then equating coefficients of x in (1.1) gives the system of equations

$$\begin{bmatrix} A & 0 & A' & 0 \\ B & A & B' & A' \\ C & B & C' & B' \\ 0 & C & 0 & C' \end{bmatrix} \begin{bmatrix} M \\ N \\ M' \\ N' \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ h \end{bmatrix} \tag{1.2}$$

whose coefficient matrix is the Sylvester matrix, over 75 years before Sylvester wrote it down in 1840. Bézout goes on to show that when x is eliminated, the "équation résultant" in y has degree $mm' + mp' + m'p$. ◁▷

These days, we use the term *sparse elimination* for Example 1.2, yet Bézout's paper dates from 1764. We will soon see that Bézout had even more exotic versions of his theorem.

We turn to a second example from Bézout's paper.

EXAMPLE 1.3. Suppose we have cubic equations

$$\begin{aligned} f &= Ax^3 + Bx^2 + Cx + D = 0 \\ g &= A'x^3 + B'x^2 + C'x + D' = 0. \end{aligned}$$

Then Bézout proceeds as follows:

(1) Multiply the first equation by A', the second by A, and subtract to obtain

$$(A'B - AB')x^2 + (A'C - AC')x + A'D - AD' = 0.$$

(2) Multiply the first equation by $A'x + B'$, the second by $Ax + B$, and subtract to obtain

$$(A'C - AC')x^2 + (A'D - AD' + B'C - BC')x + B'D - BD' = 0.$$

(3) Multiply the first equation by $A'x^2 + B'x + C'$, the second by $Ax^2 + Bx + C$, and subtract to obtain

$$(A'D - AD')x^2 + (B'D - BD')x + C'D - CD' = 0.$$

Bézout then considers "each power of x as an unknown," so that (1)–(3) give a system of three equations in three unknowns. The determinant of this system is the final equation that eliminates x. ◁▷

Example 1.3 and its generalization to $\deg(f) = \deg(g) = m$ (also studied by Bézout) is the origin of what we now call the *Bézoutian* of f, g. Observe that

$$y^2 \cdot \text{RHS of (1)} + y \cdot \text{RHS of (2)} + \text{RHS of (3)} = -\frac{f(x)g(y) - f(y)g(x)}{x - y},$$

which means that up to a minus sign, the quantities $A'B - AB'$, $A'C - AC'$, etc. of Bézout are the coefficients b_{ij} in the expression

$$\frac{f(x)g(y) - f(y)g(x)}{x - y} = \sum_{i,j=1}^{3} b_{ij} x^{i-1} y^{j-1} \tag{1.3}$$

The *Bézoutian matrix* is $\mathrm{Bez}(f, g) = (b_{ij})$, and its determinant is the resultant $\mathrm{Res}(f, g)$. We will see later in the section that Cayley discovered this in 1857.

Bézout and Waring. We now turn to two important books:

- In 1779, Bézout published *Théorie Générale des Équations Algébriques.*
- In 1782, Waring published the second edition of *Meditationes Algebraicæ.*

We will go back and forth between the two books, beginning with a quote from Bézout's *Théorie Générale* (italics added):

> We conceive of each *given equation* as being multiplied by a *special polynomial.* Adding up all those products together, the result is what we call the *sum-equation.* This sum-equation will become the *final equation* through the vanishing of all terms affected by the unknowns to eliminate. [**28**, (224.)]

Here are the terms in italics, expressed in modern language:

- We have *given equations* $f_1 = \cdots = f_n = 0$ in $x_1, \ldots, x_n$ (for Bézout, the number of equations equals the number of unknowns).
- Multiply each by *special polynomials* $A_1, \ldots, A_n$.
- Adding then up gives the *sum-equation* $A_1 f_1 + \cdots + A_n f_n = 0$.
- This becomes the *final equation* when all terms involving the unknowns to eliminate vanish, e.g.,

$$A_1 f_1 + \cdots + A_n f_n \in k[x_1].$$

Thus the elimination ideal $\langle f_1, \ldots, f_n \rangle \cap k[x_1]$ is implicit in Bézout. But set theory lay over 100 years in the future, so objects like $\langle f_1, \ldots, f_n \rangle$ were not part of Bézout's mental landscape—he worked with polynomials finitely many at a time. Even writing $k[x_1]$ is problematic, for Bézout never specified a field. His definition of polynomial simply says "coéfficiens quelconques" ("any coefficients").

Now let's switch to Waring's *Meditationes Algebraicæ.* This book covers an impressive range of topics, from symmetric polynomials to the Waring problem about expressing positive integers as sums of kth powers. We will focus on Waring's comments about Bézout's *Théorie Générale*, which appeared three years before the second edition of *Meditationes Algebraicæ* [**358**].

Here is a quote from Waring, where we have added bullets and comments enclosed in [...] that indicate the relation to Bézout:

> If there are h equations [*given equations*], of degrees n, m, l, k, etc., respectively, in as many unknowns, and
>
> - if each of these h equations are multiplied by h assumed equations [*special polynomials*] of degrees $nmlk\cdots - n$, $nmlk\cdots - m$, $nmlk\cdots - l$, $nmlk\cdots - k$, etc., respectively, then
> - let the resulting equations be added together [*sum-equation*], and
> - equate to zero the coefficients of each term of degree $nmlk\cdots$ in the result (barring those involving only powers of y),
>
> whence the equation whose root is x or y or z, etc. [*final equation*], cannot be of degree greater than $nmlk\cdots$ [**358**, p. xxxv, English translation]

This is Bézout's Theorem for n polynomial equations in n unknowns, cast as a problem in elimination theory. We will state Bézout's version below.

Here is another of Bézout's results from *Théorie Générale*, again quoting from Waring with bullets added:

> If there are h equations respectively of degrees $n, m, l, k, \dots$,
> - all involving the same unknown quantities $x, y, z, v, \dots$, and
> - if $p, q, r, s, \dots, p', q', r', s', \dots, p'', q'', r'', s'', \dots$, etc., are the maximum degrees to which $x, y, z, v, \dots$ appear in the equations of degrees $n, m, l, k, \dots$ respectively,
> - then the equation whose root is x or y or z, etc., cannot be of degree higher than
>
> $$\begin{aligned} n \times m \times l \times k \times \cdots &- (n-p) \times (m-p') \times (l-p'') \times \cdots \\ &- (n-q) \times (m-q') \times (l-q'') \times \cdots \\ &- (n-r) \times (m-r') \times (l-r'') \times \cdots \\ &- \cdots \end{aligned}$$
> [**358**, p. 209, English translation]

The key feature here is that we fix the total degree of each polynomial and put separate bounds on the degree in each variable. This differs from Example 1.2, where in 1764 Bézout fixed the total degree and bounded the degree of just the first variable.

Bézout's *Théorie générale*. The results quoted from Waring are just the tip of the iceberg when it comes to the "Bézout theorems" in *Théorie Générale* [**28**]. This amazing book consists of 469 pages of text organized into sections (1.)–(561.).

Bézout had an elaborate classification of polynomials and systems of equations. He worked affinely with polynomials of degree T in n variables. A polynomial is *complete* when it uses *all* monomials of total degree $\leq T$. We are now ready for Bézout's first theorem:

> *The degree of the "équation finale résultant" of any number of complete equations containing the same number of unknowns, and of any degree, is equal to the product of the exponents of the degrees of these equations.* [**28**, (47.)].

To relate this to the number of solutions, recall that the "équation finale résultant" comes from eliminating all variables but one, say x_1, so its roots are the x_1-coordinates of solutions of the system. Hence, if the number of solutions is finite with distinct x_1 coordinates, then the number of solutions is bounded by the "product of the exponents of the degrees" since some solutions might be at ∞ (recall that we are working affinely).

For Bézout, a polynomial is *incomplete* when it is not complete. Incomplete polynomials fall into various *species*. For example, the second result of Bézout quoted from Waring is for polynomials of the *first species*. To indicate how deeply Bézout thought about polynomials, let's explore his third species in detail.

Example 1.4. A polynomial in u, x, y belongs to the third species considered by Bézout when it is satisfies the following conditions:

> 1.° that u does not exceed the degree a, x does not exceed the degree $a_{\prime}$, y does not exceed the degree $a_{\prime\prime}$;
>
> 2.° that u with x does not rise above the dimension b; u with y does not rise above the dimension $b_{\prime}$; x with y does not rise above the dimension $b_{\prime\prime}$;

3.° that u with x and with y cannot together rise to a dimension higher than t. [**28**, (82.)]

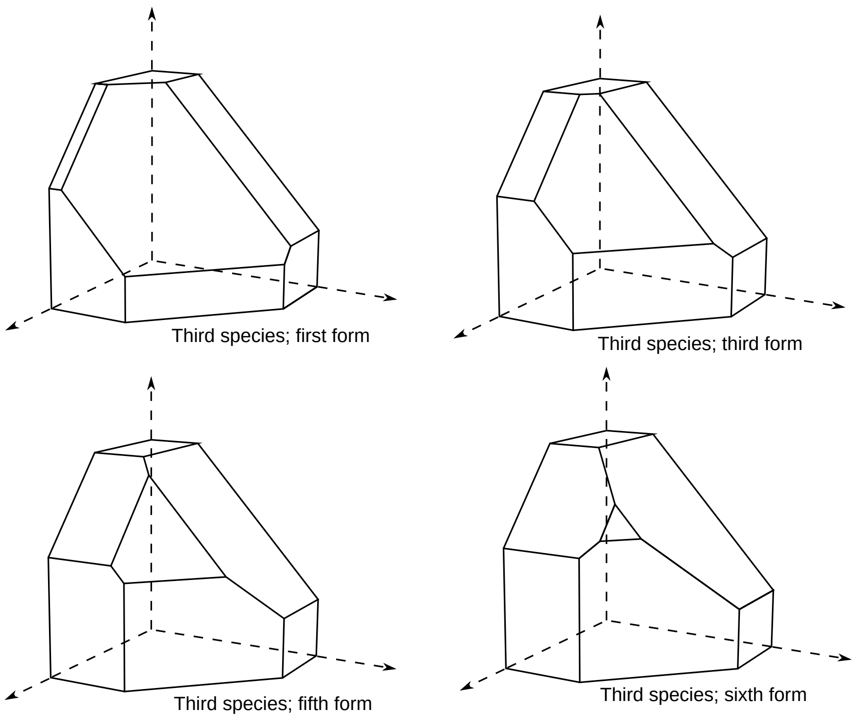
Then a monomial $u^{k_1}x^{k_2}y^{k_3}$ is of the third species considered by Bézout if there are constants $t, a, \underset{\prime}{a}, \underset{\prime\prime}{a}, b, \underset{\prime}{b}, \underset{\prime\prime}{b}$ (in Bézout's awkard notation) such that

$$
\begin{aligned}
&k_1 + k_2 + k_3 \le t \\
&0 \le k_1 \le a, \quad 0 \le k_2 \le \underset{\prime}{a}, \quad 0 \le k_3 \le \underset{\prime\prime}{a} \\
&k_1 + k_2 \le \underset{\prime\prime}{b}, \quad k_1 + k_3 \le \underset{\prime}{b}, \quad k_2 + k_3 \le b.
\end{aligned} \tag{1.4}
$$

Bézout notes that the third species has eight *forms*, according to the signs of

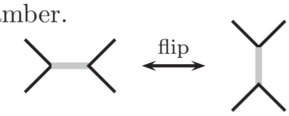
$$
t - b - \underset{\prime}{b} + \underset{\prime\prime}{a}, \quad t - b - \underset{\prime\prime}{b} + \underset{\prime}{a}, \quad t - \underset{\prime}{b} - \underset{\prime\prime}{b} + a. \tag{1.5}
$$

These days, we regard the inequalities (1.4) as describing the lattice points in a polytope, and the eight forms indicate give eight possibilities for what the polytope can look like. Here is a drawing of four of these, taken from [**288**]:



To a reader versed in toric geometry, this picture is astonishing. The polytopes share the same 10 inward-facing facet normals. These 10 vectors have a *secondary fan* whose chambers have some wonderful properties:

- Each polytope in the picture corresponds to a chamber.
- Two chambers that share a common wall give polytopes that differ by a *bistellar flip* of an edge:



In the picture, you can see a sequence of flips about edges:

$$
\text{first form} \overset{\text{flip}}{\longleftrightarrow} \text{third form} \overset{\text{flip}}{\longleftrightarrow} \text{fifth form} \overset{\text{flip}}{\longleftrightarrow} \text{sixth form.}
$$

This corresponds to a path in secondary fan where we traverse four chambers, crossing three walls in the process. See [**106**, §14.4] for more on the secondary fan.

However, Bézout was not thinking about the splendid geometry implicit in the eight forms of the third species. Rather, he wanted the degree of the "équation finale résultant" of a system $f = g = h = 0$ when the polynomials have the same form of the third species. In [**28**, (120.)–(127.)], he states eight Bézout Theorems when f, g, h satisfy (1.4) for parameters

$$(t, a, \underset{\prime}{a}, \underset{\prime\prime}{a}, b, \underset{\prime}{b}, \underset{\prime\prime}{b}),\ (t', a', \underset{\prime}{a}', \underset{\prime\prime}{a}', b', \underset{\prime}{b}', \underset{\prime\prime}{b}'),\ (t'', a'', \underset{\prime}{a}'', \underset{\prime\prime}{a}'', b'', \underset{\prime}{b}'', \underset{\prime\prime}{b}''). \tag{1.6}$$

If in addition f, g, h are all of the first form, where the quantities in (1.5) are all negative, then the "équation finale résultant" has degree

$$\begin{aligned}
D = t t' t'' &- (t-a).(t'-a').(t''-a'') - (t-\underset{\prime}{a}).(t'-\underset{\prime}{a}').(t''-\underset{\prime}{a}'')\\
&- (t-\underset{\prime\prime}{a}).(t'-\underset{\prime\prime}{a}').(t''-\underset{\prime\prime}{a}'') + (t-b).(t'-b').(t''-b'')\\
&+ (t-\underset{\prime}{b}).(t'-\underset{\prime}{b}').(t''-\underset{\prime}{b}'') + (t-\underset{\prime\prime}{b}).(t'-\underset{\prime\prime}{b}').(t''-\underset{\prime\prime}{b}'')\\
&- (a+\underset{\prime}{a}-b).(t'-b').(t''-b'') - (a'+\underset{\prime}{a}'-b').(t-b).(t''-b'')\\
&- (a''+\underset{\prime}{a}''-b'').(t-b).(t'-b') - (a+\underset{\prime\prime}{a}-\underset{\prime}{b}).(t'-\underset{\prime}{b}').(t''-\underset{\prime}{b}'')\\
&- (a'+\underset{\prime\prime}{a}'-\underset{\prime}{b}').(t-\underset{\prime}{b}).(t''-\underset{\prime}{b}'') - (a''+\underset{\prime\prime}{a}''-\underset{\prime}{b}'').(t-\underset{\prime}{b}).(t'-\underset{\prime}{b}')\\
&- (\underset{\prime}{a}+\underset{\prime\prime}{a}-\underset{\prime\prime}{b}).(t'-\underset{\prime\prime}{b}').(t''-\underset{\prime\prime}{b}'') - (\underset{\prime}{a}'+\underset{\prime\prime}{a}'-\underset{\prime\prime}{b}').(t-\underset{\prime\prime}{b}).(t''-\underset{\prime\prime}{b}'')\\
&- (\underset{\prime}{a}''+\underset{\prime\prime}{a}''-\underset{\prime\prime}{b}'').(t-\underset{\prime\prime}{b}).(t''-\underset{\prime\prime}{b}'),
\end{aligned} \tag{1.7}$$

where we follow the typsetting used in [**28**, (120.)]. Bézout gives similar formulas when the polynomials all have the second form, the third form, etc.

Here's what's happening from the modern point of view. Assume that the parameters (1.6) of $f, g, h \in \mathbb{C}[u, x, y]$ are of the first form and that f, g, h are generic within their Newton polytopes. Then:

- The above formula for D is the *mixed volume* of the Newton polytopes of the polynomials f, g, h. We will verify this in Section 1.2.
- By Bernstein's Theorem (see Theorem 1.12 in Section 1.2), D is the number of solutions of $f = g = h = 0$ in the torus $(\mathbb{C}^*)^3$.
- The minimal generator of the elimination ideal $\langle f, g, h\rangle \cap \mathbb{C}[u]$ has degree D. This polynomial is Bézout's "équation finale résultant."

It is wonderful to see so much contemporary mathematics packed into an example from a book published in 1779. ◁▷

Bézout's Proofs. In [**288**], Penchèvre gives a careful reading of *Théorie générale* and discusses Bézout's proofs, most of which are incomplete. But [**288**] also shows how tools from commutative algebra (Koszul complexes) and algebraic geometry (toric varieties) can be used to verify many Bézout's results, including the formula for D given in Example 1.4.

To give you a hint of what Bézout did, let's say a bit more about Example 1.4. If we set $\mathsf{a} = (a, \underset{\prime}{a}, \underset{\prime\prime}{a})$, then the data (1.6) can be written

$$(t, \mathsf{a}, \mathsf{b}),\ (t', \mathsf{a}', \mathsf{b}'),\ (t'', \mathsf{a}'', \mathsf{b}'')$$

To find the "équation finale résultant," Bézout uses *special polynomials* F, G, H so that $Ff + Gg + Hg$ lies in the elimination ideal. He picks F to have parameters

$(T-t, \mathsf{A}-\mathsf{a}, \mathsf{B}-\mathsf{b})$ so that Ff has parameters $(T, \mathsf{A}, \mathsf{B})$, and similarly for G, H. In modern language, this leads to the linear map

$$\phi : C_1 \oplus C_2 \oplus C_3 \xrightarrow{(f,g,h)} C_0,$$

where all possible F's lie in C_1, etc. Bézout is interested in $\dim \operatorname{coker} \phi$, which he describes by saying *useful coefficients* and *useless coefficients*. Then:

- The dimension of C_0 is the number of lattice points in the polytope (1.4) for parameters $(T, \mathsf{A}, \mathsf{B})$, and similarly for C_1, C_2, C_3.
- This number is a quasi-polynomial in the parameters, with a separate polynomial for each of the eight forms. Bézout computes these in [**28**, (92.)–(99.)].
- Bézout's formula for $\operatorname{coker} \phi$ implicitly assumes that the Koszul complex of f, g, h is exact, which means that the only cancellations that occur are the obvious ones. Bézout explains this by saying "the number of terms cannot be made lower by introducing fictitious terms." [**28**, (112.)]

All of this is discussed in detail in [**288**, Section 6]. There is a lot of math going in Bézout's book!

Poisson. While a student at the École Polytechnique, Poisson wrote the paper *Mémoire sur l'élimination dans les équations algébriques*, published in 1802. Here is how his paper begins:

> The degree of the "équation finale résultant" from the elimination, among a number of equations equal to that of the unknowns, of all the unknowns except one, cannot be greater than the product of the exponents of these equations; and it is precisely equal to this product, when the given equations are the most general of their degrees. This important theorem is Bézout's, but the way he proves it is neither direct nor simple; nor is it devoid of any difficulty. [**293**]

Poisson's goal is to prove Bézout's Theorem in the case of complete polynomials, and he observes that Bézout's bound is sharp in the generic case ("most general of their degrees"). The final sentence of the quote is telling—like many of his contemporaries, Poisson found Bézout's book to be difficult reading. This happened so often that in 1907, Brill and Noether [**45**, p. 143] would comment that Bézout's *Théorie générale* is "ebenso viel berühmten als wenig gelesen" ("as well known as little read").

Because of this, Poisson provides his own proof, which proceeds by induction on n, the number of equations in the system $f_1 = \cdots = f_n = 0$. The inductive step uses the product

$$\prod_{f_1(p)=\cdots=f_{n-1}(p)=0} f_n(p),$$

which is now part of the *Poisson formula* for the multivariable resultant (see, for example, [**105**, Exercise 3.3.8]). For the base case $n = 2$ of the induction, Poisson cites Cramer's 1750 paper [**116**], where Cramer uses the product $\prod_{f_1(p)=0} f_2(p)$.

Poisson's paper brings us to the end of our discussion of elimination in the 18th century. Readers interested in a fuller picture of what happened in this century should consult Penchèvre's article [**286**].

The 19th Century. The century that followed Poisson's 1802 paper witnessed an explosion of work on elimination theory. In 1892, Netto and Le Vavasseur wrote a 232 page review article about algebra in the 19th century for *l'Encyclopédie des* *sciences mathématiques pures et appliquées.* Their section on elimination theory is 97 pages long and begins as follows:

> The large number and variety of memoirs relating to elimination makes it difficult to classify these memoirs rationally. [**281**, pp. 73–169]

A more recent account of elimination theory in the 19th century can be found in Penchèvre's 318 page PhD thesis *Histoire de la théorie de l'élimination* [**287**, Chapters 12–16].

There is no way we can do justice to this vast amount of material. We will instead focus on a few selected topics, beginning with resultants.

The Emergence of Resultants. After Poisson, elimination theory had a low profile in the first few decades of the 19th century. In 1835, Jacobi [**218**] used elimination theory to prove a result that became the *Euler-Jacobi Formula* (see [**245**] for a modern account). But starting in 1839, multiple papers on elimination theory appeared in rapid succession.

The Sylvester matrix appeared twice during this period. In 1840, Sylvester [**340**] described a matrix whose rows are built from the coefficients of the polynomials, suitably shifted by successively adding more zeros. The result is

> a solid square $(m+n)$ terms *deep* and $(m+n)$ terms *broad*.

He takes the determinant of this matrix, though he never says "determinant."

The same matrix was discovered independently four years later by Hesse in 1844 [**204**]. He begins with polynomials

$$\begin{aligned} A_0 &= a_n\, x^n \ + a_{n-1}\, x^{n-1} \ + a_{n-2}\, x^{n-2} \ + \ldots . + a_1 x + a_0 \\ B_0 &= b_m x^m + b_{m-1} x^{m-1} + b_{m-2} x^{m-2} + \ldots . + b_1 x + b_0 \end{aligned}$$

(he uses four dots .... rather than the three dots ... commonly used these days). He multiplies A_0 by $x^{m-1}, x^{m-2}, \ldots, x, 1$ and B_0 by $x^{n-1}, x^{n-2}, \ldots, x, 1$. Expressing these in terms of the monomials $x^{m+n-1}, x^{m+n-2}, \ldots, x, 1$, he obtains the following $(m+n) \times (m+n)$ matrix:

	1	2	3	....	....	....	....	....	....	....	$m+n-1$	$m+n$
1	a_n	a_{n-1}	a_{n-2}	....	a_2	a_1	a_0	0	0	0	0	0
2	0	a_n	a_{n-1}	....	a_3	a_2	a_1	a_0	0	0	0	0
3	0	0	a_n	....	a_4	a_3	a_2	a_1	a_0	0	0	0
⋮												
$m-1$	0	0	0	0	a_n	a_{n-1}	a_{n-2}	....	a_2	a_1	a_0	0
m	0	0	0	0	0	a_n	a_{n-1}	....	a_3	a_2	a_1	a_0
$m+1$	b_m	b_{m-1}	b_{m-2}	....	b_2	b_1	b_0	0	0	0	0	0
$m+2$	0	b_m	b_{m-1}	....	b_3	b_2	b_1	b_0	0	0	0	0
$m+2$	0	0	b_m	....	b_4	b_3	b_2	b_1	b_0	0	0	0
⋮												
$m+n-1$	0	0	0	0	b_m	b_{m-1}	b_{m-2}	....	b_2	b_1	b_0	0
$m+n$	0	0	0	0	0	b_m	b_{m-1}	....	b_3	b_2	b_1	b_0

Like Sylvester, he arranged the coefficients in rows, so this is the transpose of what is commonly called the "Sylvester matrix." We saw an example of this in (1.2), though neither Sylvester nor Hesse seem to be aware of what Bézout did in 1764.

These papers from 1840 and 1844 worked affinely and said "result of elimination" rather than "resultant." But in 1847, Cayley published a paper on elimination that begins as follows:

> Designating $U, V, W, \ldots$ homogeneous functions of orders $m, n, p, \&c.$ and an equal number of variables respectively, and assuming that these functions are the most general possible, that is to say that the coefficient of each term is an indeterminate letter: we know that the equations $U = 0, V = 0, W = 0, \ldots$ offer a relation $\Theta = 0$, in which the variables no longer enter, and where the function Θ, which can be called the *complete Resultant of the equations*, is homogeneous and of the order $np\ldots$ relative to the coefficients of U, of the order $mp\ldots$ relative to those of V, and so on, while it is not decomposable into factors. [**71**]

There is a *lot* going in this quote:

- Cayley explicitly says "resultant."
- He uses homogeneous polynomials. We are no longer affine!
- He states some basic properties of the resultant (degree in the coefficients of each polynomial, irreducibility).

We have the beginnings of a general theory of resultants in 1847!

The switch to homogeneous polynomials was driven by two emerging areas of mathematics: projective algebraic geometry, where points in projective space are described by homogeneous coordinates, and invariant theory, where invariants are naturally homogeneous (an example is the determinant, invariant under the action of $\mathrm{SL}(n)$). To get a better sense of how the story unfolded, let's spend some more time with Cayley.

Cayley. In 1845 and 1846, Cayley wrote two papers on invariant theory that introduced *hyperdeterminants*. In the second paper, he observes a link with the resultant of two homogeneous polynomials of degree two and comments that

> Important results might be obtained by connecting the theory of hyperdeterminants with that of elimination, but I have not yet arrived at anything satisfactory upon this subject. [**70**]

We suspect that Cayley would have enjoyed the 1994 book *Discriminants, Resultants, and Multidimensional Determinants* [**165**], which devotes an entire chapter to hyperdeterminants.

Cayley was also familiar with Bézout's 1764 paper. In 1857, he published the short paper *Note sur la méthode d'élimination de Bezout* in Crelle's journal [**72**][1] about what he later called "Bezout's abbreviated process of elimination." We illustrated this in Example 1.3. Cayley and Sylvester wrote Bézout's name without the accent.

In his note, Cayley begins with homogenous polynomials

$$U = (a, \ldots \between x, y)^n = 0$$
$$V = (a', \ldots \between x, y)^n = 0,$$

[1] When Cayley published outside of England, he often wrote in French—in the 19th century, English was not the dominant mathematical language that it is today.

where $(a, b, c, \dots \between x, y)^n = ax^n + \binom{n}{1} bx^{n-1}y + \binom{n}{2} cx^{n-2}y^2 + \cdots$ is how Cayley writes a homogeneous polynomial of degree n in x, y. He refers to U and V as "quantics." This is part of the symbolic notation that Cayley developed for invariant theory.

Then Cayley considers the quotient

$$\frac{(a, \dots \between x, y)^n\,(a', \dots \between \lambda, \mu)^n - (a, \dots \between \lambda, \mu)^n\,(a', \dots \between x, y)^n}{\mu x - \lambda y},$$

which we would write more simply as

$$\frac{U(x,y)V(\lambda,\mu) - U(\lambda,\mu)V(x,y)}{\mu x - \lambda y}. \tag{1.8}$$

As noted after Example 1.3, Bézout's method is encoded in the coefficients of (1.8) and the determinant of the resulting square matrix is the resultant of U and V. Also observe that (1.8) is the homogeneous version of what we did in (1.3).

A year later in his *Fourth Memoir Upon Quantics*, Cayley calls (1.8) the *Bezoutic emanant* and comments that

> the result of the elimination is consequently obtained by equating to zero the determinant formed with the matrix which enters into the expression of the Bezoutic emanant. In other words, this determinant is the Resultant of the two quantics. [**73**]

These days, we use the simpler term "Bézoutian," for which we should be grateful.

Sylvester also studied Bézout's work and in 1853 wrote a paper [**341**] that defines *Bezoutians*, *Bezoutics* and the *Bezoutian matrix*. This paper also introduced *Sylvester double sums*, a subject of current interest—see [**118**].

We next discuss a geometric example from Cayley that involves elimination.

Example 1.5. In Example 1.1, Newton counted tangent lines to a curve that go through a given point. In his 1847 paper, Cayley poses the problem of finding the *equations* of these tangent lines:

> Find the equation of the system of tangents drawn from a fixed point to a given curve. [**71**]

As before, we assume that $U(x, y, z) = 0$ defines a smooth curve of degree of n in $\mathbb{P}^2$ and that the given point is $P = (\alpha, \beta, \gamma)$. In order for a point $Q = (\xi, \eta, \zeta)$ to lie on a tangent line going through P, we must satisfy the system of equations

$$\begin{aligned} U &= 0 \\ \alpha\frac{\partial U}{\partial x} + \beta\frac{\partial U}{\partial y} + \gamma\frac{\partial U}{\partial z} &= 0 \\ \det\begin{bmatrix} \xi & \eta & \zeta \\ x & y & z \\ \alpha & \beta & \gamma \end{bmatrix} &= 0. \end{aligned} \tag{1.9}$$

The first two equations tell us that (x, y, z) is a point on the curve whose tangent line goes through P, and the third equation says that Q lies on this tangent.

For this system, Cayley regards the variables as x, y, z, while $\alpha, \beta, \gamma, \xi, \eta, \zeta$ are unknown coefficients. He then takes the resultant Θ of (1.9), where Θ is from the quote at the top page 10. You should reread this quote before proceeding further.

The resultant Θ is a polynomial in $\alpha, \beta, \gamma, \xi, \eta, \zeta$ and the coefficients of U. He notes that (1.9) has an obvious solution when (α, β, γ) is on the curve (courtesy of the Euler relation $xU_x + yU_y + zU_z = nU$), so that $U(\alpha, \beta, \gamma)$ is a factor of Θ. Thus

$$\Theta = U(\alpha, \beta, \gamma) \cdot \Phi,$$

where Φ is the defining equation of the tangent lines when regarded as a polynomial in ξ, η, ζ. In Cayley's terminology, Φ is a *reduced resultant.* ◁▷

Although the resultant is irreducible when the coefficients of the system are independent unknowns, *extraneous factors* can arise for the resultant of a particular system, such as (1.9). We will see some interesting examples of extraneous factors in Chapter 3.

We close our discussion of Cayley with a quote from a paper on elimination he wrote in 1864:

> In the problem of elimination, one seeks the relationship that must exist between the coefficients of a function or system of functions in order that some particular circumstance (or singularity) can occur. [**74**]

It is clear that for Cayley, elimination is at the heart of algebraic geometry.

Before we move on to the general theory of resultants, we need to discuss an interesting 1841 paper of Minding that illustrates a different flavor of elimination.

Minding. While the 19th century focused mostly on resultants of complete polynomials (in Bézout's terminology), we have seen that Bézout himself studied more general situations. But Minding's 1841 paper *Ueber die Bestimmung des Grades einer durch Elimination hervorgehenden Gleichung* [**275**] goes far beyond anything Bézout could have imagined. Minding considers a system of (affine) equations

$$\begin{aligned} f(x,y) &= A_0 y^m + A_1 y^{m-1} + \cdots + A_{m-1} y + A_m = 0 \\ \theta(x,y) &= B_0 y^n \;+ B_1 y^{n-1} \;+ \cdots + B_{n-1} y \;+ B_n \;= 0, \end{aligned}$$

where "the letters A and B with subscripts stand for arbitrary polynomials in x." The goal is to compute the degree of the "final equation" $\psi(x) = 0$ obtained by eliminating y. Following Poisson, Minding expresses the final equation as

$$\psi(x) = B_0^m f(x, y_1) \cdot f(x, y_2) \cdots f(x, y_n), \tag{1.10}$$

where $y_1, \ldots, y_n$ are the roots of $\theta(x, y)$. Minding expresses each y_i as a Puiseux series in x about ∞, so the exponents are decreasing. For example,

$$y_1 = y_1(x) = c_1 x^{h_1} + d_1 x^{h_1 - \alpha_1} + e_1 x^{h_1 - \alpha_2} + \cdots, \quad 0 < \alpha_1 < \alpha_2 < \cdots,$$

where $h_1, \alpha_1, \alpha_2, \cdots \in \mathbb{Q}$. Doing this for $y_2, \ldots, y_n$ and focusing on the highest power of x on each side of (1.10), we obtain

$$\deg(\psi(x)) = \deg\big(B_0(x)^m f(x, c_1 x^{h_1}) \cdots f(x, c_n x^{h_n})\big).$$

If we set $b = \deg\big(B_0(x)\big)$ (an integer) and $k_i = \deg\big(f(x, c_i x^{h_i})\big)$ (often a fraction), then the above equation simplifies to

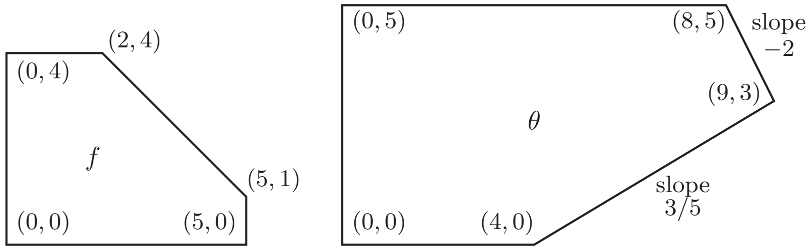
$$\deg(\psi(x)) = mb + k_1 + \cdots + k_n,$$

which is the major result of [**275**]. Here is an example from Minding's paper.

EXAMPLE 1.6. Suppose that

$$\begin{aligned} f(x,y) &= (x^2)y^4 + (x^2)y^3 + (x^4)y^2 + (x^5)y + (x^5) \\ \theta(x,y) &= (x^8)y^5 + (x^6)y^4 + (x^9)y^3 + (x^4)y^2 + (x^3)y + (x^4), \end{aligned}$$

where (x^ℓ) means a polynomial of degree ℓ. These polynomials have the following Newton polygons:



The commentary to [**275**, English translation] explains how the negative reciprocals of the slopes of the polygon for θ give

$$h_1 = h_2 = \tfrac{1}{2},\ h_3 = h_4 = h_5 = -\tfrac{5}{3}.$$

This easily leads to

$$k_1 = k_2 = \tfrac{11}{2},\ k_3 = k_4 = k_5 = 5,$$

giving $\deg(\psi(x)) = 4 \cdot 8 + \frac{11}{2} + \frac{11}{2} + 5 + 5 + 5 = 58$. ◁▷

The commentary to [**275**, English translation] goes on to show that in general, Minding's formula for $\deg(\psi(x))$ equals the mixed volume (mixed area in this case) of the Newton polygons of f and θ. So Minding's paper from 1841 provides another splendid example of Bernstein's Theorem.

The Theory of Resultants. In the latter part of the 19th century, many papers on resultants were published, including Brill (1880) [**43**], Kronecker (1882) [**244**] and Mertens (1886) [**271, 272**]. This continued into the early part of the 20th century with Macaulay (1902) [**264**] and the books by Netto (1900) [**280**] and Macaulay (1916) [**265**].

These papers and books focus primarily on homogeneous polynomials where all monomials of a given degree are allowed. This is the homogeneous analog of Bézout's notion of a complete polynomial. Let $f_0, \cdots, f_n$ be homogeneous in $x_0, \dots, x_n$ of degrees $d_0, \dots, d_n$. Then the *classical* or *dense resultant*, denoted by $\mathrm{Res}_{d_0,\dots,d_n}(f_0, \dots, f_n)$, is an integer polynomial in the coefficients of the f_i with the property that over any algebraically closed field,

$$\text{(1.11)} \qquad \mathrm{Res}_{d_0,\dots,d_n}(f_0, \dots, f_n) = 0 \iff \begin{cases} f_0 = \cdots = f_n = 0 \text{ has a} \\ \text{solution in projective space } \mathbb{P}^n. \end{cases}$$

Kronecker, in his 1882 paper *Grundzüge einer arithmetischen Theorie der algebraischen Grössen* on the foundations of algebraic geometry, sketched a theory of resultants that gave elimination a central role. Here is a quote from his paper:

> A system of arbitrarily many algebraic equations for $\mathfrak{z}^0, \mathfrak{z}', \mathfrak{z}'', \mathfrak{z}^{(n-1)}$, in which the coefficients belong to the rationality domain $(\mathfrak{R}', \mathfrak{R}'', \mathfrak{R}'', \dots)$, define the algebraic relations between $\mathfrak{z}$ and $\mathfrak{R}$, whose knowledge and representation are the purpose of the theory of elimination. [**244**, §10]

Kronecker's notation and terminology are challenging for the modern reader—see [**289**] for a modern account of [**244**]. But what Kronecker is saying in the quote does not differ greatly from the 1864 quote from Cayley given earlier. Kronecker never published the details of his approach, leaving it to his student Mertens to complete the task.

The first book to cover resultants in detail was Volume II of Netto's *Vorlesungen über Algebra* [**280**], published in 1900. His treatment of resultants includes sections on Poisson, Minding, Kronecker, Cayley, Sylvester, and Bézout. He actually read Bézout's *Théorie Générale*!

In [**280**, pp. 146–147], Netto generalizes the matrix described by Sylvester and Hesse in the 1840s. Recall that when $f(x), g(x)$ have degree n, m, Hesse multiplied f by $x^{m-1}, \dots, x, 1$ and g by $x^{n-1}, \dots, x, 1$, and expressed these in terms of $x^{m+n-1}, \dots, x, 1$. This gave the Sylvester matrix whose determinant is $\mathrm{Res}(f, g)$.

Netto studies what happens in general. He works affinely, so we will modify his notation to work homogeneously. Consider $m+1$ homogeneous equations

$$f_\alpha(x_0, \dots, x_m) = 0, \quad \alpha = 0, \dots, m$$

of degree $n_\alpha = \deg(f_\alpha)$. We assume that f_α is the sum of all monomials of degree n_α with coefficients that are independent variables. This gives the universal version of the resultant in degrees $n_0, \dots, n_m$, which we denote by

$$\mathrm{Res}(f_0, \dots, f_m) = \mathrm{Res}_{n_0,\dots,n_m}(f_0, \dots, f_m) \in \mathbb{Z}[c].$$

Here, c is the vector of coefficients of $f_0, \dots, f_m$, so $f_0, \dots, f_m \in \mathbb{Z}[c, x_0, \dots, x_m]$.

To build a matrix that captures the resultant, set

$$n = n_0 + \cdots + n_m - m$$

and multiply f_α by all monomials x^β of degree $n - n_\alpha$. Then express the products $x^\beta f_\alpha$ as linear combinations of all monomials of degree n. The coefficients form the rows of a matrix M of size $\Sigma \times N$, where Σ is the number of products $x^\beta f_\alpha$ and N is the number of monomials z^β. Using the definition of n, Netto shows that

$$\Sigma = \sum_{\alpha=1}^{m+1} \binom{n - n_\alpha + m}{m} \geq N = \binom{n+m}{m}.$$

When $m = 1$, this is an equality and M is the Sylvester matrix of f_0 and f_1.

PROPOSITION 1.7. *Let Δ be an $N \times N$ maximal minor of M. Then:*

(1) *Δ is a multiple of* $\mathrm{Res}(f_0, \dots, f_m)$.
(2) $\Delta \in \mathbb{Z}[c] \cap \big(\langle f_0, \dots, f_m\rangle : \langle x_0, \dots, x_m\rangle^\infty\big)$.

PROOF. Let M_0 denote the $N \times N$ submatrix of M with $\Delta = \det(M_0)$. By construction,

$$M_0 \begin{bmatrix} \vdots \\ x^\beta \\ \vdots \end{bmatrix} = \begin{bmatrix} \vdots \\ x^\beta f_\alpha \\ \vdots \end{bmatrix}, \tag{1.12}$$

where the column vector on the right consists of those $x^\beta f_\alpha$ that give rows of M_0.

(1) Since $\mathrm{Res}(f_0, \dots, f_m)$ is irreducible, it suffices to prove that for any choice of complex coefficients $\overline{c}$,

$$\mathrm{Res}(f_0, \dots, f_m)(\overline{c}) = 0 \Longrightarrow \Delta(\overline{c}) = 0.$$

This is easy, for by (1.11), the vanishing of the resultant for $\overline{c}$ implies that for these coefficients, the system has a nontrivial solution $\overline{x} \in \mathbb{P}^m$. If we evaluate (1.12) at $\overline{c}$ and $\overline{x}$, we get a nonzero element of the kernel of $M_0(\overline{c})$ since $\overline{x} \in \mathbb{P}^m$. This implies that $\Delta(\overline{c}) = \det(M_0(\overline{c})) = 0$.

(2) The classical adjoint M_0^{adj} of M_0 satisfies $M_0^{\text{adj}} M_0 = \Delta I_N$ and has entries in $\mathbb{Q}[c]$. If we multiply each side of (1.12) by M_0^{adj}, we obtain

$$\Delta x^\beta \in \langle f_0, \dots, f_m \rangle$$

for all x^β of degree n. This implies $\Delta \in \langle f_0, \dots, f_m \rangle : \langle x_0, \dots, x_m \rangle^n$, and then (2) follows easily. □

Part (1) of Proposition 1.7 says that when Δ is not identically zero, $\Delta = \text{Res}(f_0, \dots, f_m) E$ for some extraneous factor E. In 1902, Macaulay [**264**] showed how to find $m+1$ carefully chosen maximal minors $D_0, \dots, D_m$ of M so that

$$\text{Res}(f_0, \dots, f_m) = \gcd(D_0, \dots, D_m).$$

He also proved that for the D_i, the extraneous factor E_i is the determinant of an appropriately sized minor of D_i. The gives the formula

$$\text{Res}(f_0, \dots, f_m) = \frac{D_i}{E_i}$$

that expresses the resultant as a quotient of two determinants. Thus we have two formulas for the resultant, both lovely and both wildly impractical for computations due to the large size of the matrices involved.

Part (2) of the proposition is also important. Netto worked affinely, so his version of part (2) can be restated as

$$\Delta \in \mathbb{Z}[c] \cap \langle f_0, \dots, f_m \rangle.$$

Thus Δ is in the elimination ideal $\mathbb{Z}[c] \cap \langle f_0, \dots, f_m \rangle$ in the affine setting. Netto uses this to prove part (1).

When we switch to homogeneous polynomials, the usual elimination ideal is not helpful, since

$$\mathbb{Z}[c] \cap \langle f_0, \dots, f_m \rangle = \{0\}$$

when $f_0, \dots, f_m$ are homogenous of positive degree in $x_0, \dots, x_m$. This is where the *projective elimination ideal* $\mathbb{Z}[c] \cap \big(\langle f_0, \dots, f_m \rangle : \langle x_0, \dots, x_m \rangle^\infty\big)$ enters the picture. A key result is that the resultant generates the projective elimination ideal of $f_0, \dots, f_m$, i.e.,

$$\langle \text{Res}(f_0, \dots, f_m) \rangle = \mathbb{Z}[c] \cap \big(\langle f_0, \dots, f_m \rangle : \langle x_0, \dots, x_m \rangle^\infty\big). \tag{1.13}$$

For a modern proof, see [**226**, Proposition 2.3 and (4.1.2)], and for an introduction to projective elimination ideals, see [**104**, Chapter 8, §5]. Once we know (1.13), part (1) of Proposition 1.7 follows immediately from part (2).

We will encounter projective elimination ideals in Section 1.2 when we discuss van der Waerden's work.

Final Comments. Algebraic geometry in the 19th century was a vast enterprise. The 1892 review article by Brill and Noether [**45**] is over 450 pages long. There was a lot going on besides elimination theory.

One important topic we have not mentioned is enumerative geometry. Many famous results in the field—27 lines on a smooth cubic surface, 3264 conics simultaneously tangent to five general plane conics—date from the 19th century. In one sense, the classic version of Bézout's Theorem is a result in enumerative geometry, since it counts (with multiplicity) the number of intersections of n hypersurfaces in $\mathbb{P}^n$ when the number is finite. We saw this in Example 1.1, where Newton counted tangent lines using Bézout's Theorem. But researchers like Schubert, Chasles and others developed powerful geometric methods that had nothing to do with elimination. Writing in 1905, Brill comments that elimination theory is

> still outperformed by the well-known geometric counting methods that in the middle of the sixties [the 1860s] and ever since have flooded geometry with its enumerative results; I mean the characteristic method, the correspondence principle, and more recently the principle of conservation of number. [**44**, p. 278]

The last of these, the *principle of conservation of number*, is important to our story.

The "Princip von der Erhaltung der Anzahl" plays a central role in Schubert's 1879 book *Kalkül der abzählende Geometrie.* He explains the principle as follows:

> In algebraic interpretation, it says nothing else than that changing the constants of an equation either leaves the number of their roots unaffected or causes an infinite number of roots by making the equation an identity. [**310**, p. 12]

Though powerful, the principle of conservation of number lacked a rigorous foundation. In fact, at the 1900 International Congress of Mathematicians, Hilbert [**207**] posed the following problem:

> 15. Rigorous foundation of Schubert's enumerative calculus
>
> The problem is *to establish rigorously and with an exact determination of the limits of their validity those geometrical numbers which Schubert especially has determined on the basis of the so-called principle of special position, or conservation of number, by means of the enumerative calculus developed by him.*

Note that Hilbert refers to the principle of conservation of number. But what he says next is even more interesting:

> Although the algebra of today guarantees, in principle, the possibility of carrying out the processes of elimination, yet for the proof of the theorems of enumerative geometry decidedly more is required, namely, the actual carrying out of the process of elimination in the case of equations of special form in such a way that the degree of the final equations and the number of their solutions may be known in advance.

Hilbert is aware of the potential of elimination theory, but like Brill, he recognizes its limitations when applied to enumerative problems.

We will soon see that in the 20th century, the interaction between elimination theory and the principle of conservation of number led to some unexpected developments.

1.2. Elimination Theory in the 20th Century

After reaching a high point in the early 20th century, resultants and elimination theory went through some rough times in the middle part of the century, only to re-emerge stronger than ever in recent times.

A seminal event was the 1890 publication of Hilbert's great paper *Ueber die Theorie der algebraischen Formen* [**205**] that can lay claim to being the first paper in commutative algebra. Besides introducing the Hilbert Basis Theorem, free resolutions, and Hilbert polynomials, Hilbert's proofs also had a non-constructive aspect that was new to algebra. This was followed in 1893 with the Nullstellensatz in [**206**]. The ideas planted by Hilbert led directly to the work of Emmy Noether in the 1920s that revolutionized algebra.