On Elimination Theory *

W. Gröbner, Innsbruck

(Received 10 August 1949)

1. The task for which elimination theory was conceived is to find the complete solution of a system of algebraic equations:

$$f_1(x_1,\ldots,x_n)=0, f_2(x_1,\ldots,x_n)=0,\ldots,f_s(x_1,\ldots,x_n)=0. \quad (1)$$

The left-hand side of these equations are any given polynomials in the variables $x_1,\ldots,x_n$ with coefficients in a number field K. The task is clearly simpler the fewer variables are present. In a single variable, we may assume the solution is completely known. Hence, the fundamental idea of elimination theory is to reduce the system of equations (1) step-by-step to an equivalent system of equations in fewer variables. So first to "eliminate" the variable x_n, say, we derive a system of equations

$$g_1(x_1,\ldots,x_{n-1})=0, g_2(x_1,\ldots,x_{n-1})=0,\ldots,g_t(x_1,\ldots,x_{n-1})=0 \quad (2)$$

from (1) which has the following properties:

(a) Every solution $x_i=\xi_i$ $(i=1,\ldots,n)$ of (1) is also a solution of (2); here the ξ values come from K or an extension of K.

(b) Every solution $x_i=\xi_i$ $(i=1,\ldots,n-1)$ of (2) can be made a solution of (1) in at least one way by determining an appropriate $x_n=\xi_n$.

The most well-known method[1] for deriving such a system of equations (2) from (1) is Kronecker's method. It is probably the most convenient and transparent, but by no means the only one possible[2], indeed this cannot be, since by properties (a) and (b), the system (2) is far from unique. This is clearly a deficiency in the fundamental principles to which the usual form of elimination theory adheres, and which must be noted at some point. Of course as long as we ask only about the solutions in general, everything is

* *Über die Eliminationstheorie*, Monatshefte für Mathematik **5** (1950), 71-78. Translation by Michael Abramson.

[1] See [4, p. 12] and [11, p. 1]

[2] Compare for example Henzel's elimination method in [4, p. 14],

fine, but subtle differences pertaining to the multiplicity of the solutions are disregarded[3].

2. However, the means for constructing elimination theory free from this deficiency are provided by ideal theory. The polynomials $f(x)$ of the system (1) generate an ideal

$$\mathcal{A} = (f_1, f_2, \ldots, f_s) \tag{3}$$

in the polynomial ring $K[x_1, \ldots, x_n]$ whose zero manifold or "corresponding algebraic variety"[4] is to be determined. We want to adopt here the method outlined by elimination theory and ignore for now other possibilities offered by ideal theory. By Hilbert's Nullstellensatz, it follows from property (a) that the polynomials g of the system (2) must themselves, or a fixed power of them, be contained in the ideal $\mathcal{A}$: $g_i^\rho \equiv 0\ (\mathcal{A})$. For the polynomials g obtained by Kronecker's method, it is known that $\rho = 1$.

It is therefore reasonable to introduce the following definition: *All of the polynomials of the ideal $\mathcal{A}$ which do not contain the variable x_n form an ideal $\mathcal{B}$ in the polynomial ring $K[x_1, \ldots, x_{n-1}]$, whose basis polynomials set equal to zero produce the system of equations (2). The ideal $\mathcal{B}$ can be written as the intersection*

$$\mathcal{B} = \mathcal{A} \cap K[x_1, \ldots, x_{n-1}]$$

[3]This has actually been shown in a detailed discussion of an example given by Vahlen [10] on a famous theorem of Kronecker, in which every system of algebraic equations in n variables would be equivalent to a system of at most $n+1$ equations (so that both systems have the same solutions, see [3, p. 234]), or an algebraic variety in a linear n-dimensional space could be represented by at most $n+1$ equations (see [1, p. 232]). Since a precise, generally applicable definition of multiplicity had not been found at that time, and is still controversial today even though the construction of algebraic geometry is impossible without it, problems amalgamated with this were frequently treated and solved only intuitively, so that the results could not survive strong criticism. See [5-9].

Added in proof: This discussion and oral conversations during the mathematics congress in Innsbruck (29 Aug. to 2 Sept. 1949) have made clear that on the side of modern algebra no generally applicable notion of multiplicity exists presently (see [12, p. 211]: "But now the multiplicity of a non-isolated intersection point of 3 surfaces, to my knowledge, has never been defined."). On the other hand, it must be recognized that in the Italian school of algebraic geometry, Severi had long since developed a general notion of multiplicity, whose precise motivation and definition are undertaken in the work cited. It is also clear that all discussion on Kronecker's theorem and similar far-reaching theorems in algebraic geometry must be futile, as long as no accepted definition of multiplicity exists for both sides.

My previous work outlines the ideal theoretic notion of multiplicity which I developed in my book [11]. It is very simple and generally applicable, but does not agree exactly with Severi's notion which uses continuity considerations and passage to limits. In my opinion, the ideal theoretic definition may be preferable for the first construction of algebraic geometry because of its simplicity and generality. On the other hand, the treatment of Severi's equivalent systems brings to light the necessity of adding a notion of "virtual" multiplicity which agrees essentially with Severi's notion of multiplicity.

[4]The term "corresponding algebraic variety" is understood each time in the stricter sense of including multiplicity. See [2, p. 82]

and is called the "first elimination ideal"[5]. With this definition, all randomness in the choice and determination of the system (2) is eliminated, and it may be reasonably expected that the question of the multiplicity of the individual solutions can always be answered in a clear and easily understandable manner. Finally, to satisfy property (b), we need only assume, as in Kronecker's method, that the polynomials $f(x)$ be sufficiently generic, *i.e.* be regular[6] relative to the variable x_n.

3. The geometric meaning of the elimination of x_n is clear; namely, the algebraic variety corresponding to the ideal $\mathcal{B}$ is precisely the variety which is obtained by projecting the algebraic variety of $\mathcal{A}$ onto the coordinate hyperplane $x_n = 0$, if the center of projection is taken to be the point at infinity on the x_n-axis. The deficiency of Kronecker's method can now be expressed in that it sometimes produces this variety with too big a multiplicity, whereas, by the above definition, the multiplicity corresponds precisely to the theory and to intuition. This is illustrated by the following simple examples:

   (a) Example: $\mathcal{A} = (x_1^2, x_1 - x_2)$, $\mathcal{B} = (x_1^2)$. The algebraic variety corresponding to $\mathcal{A}$ consists of the point $x_1 = x_2 = 0$ and the neighboring[7] point lying in the direction $x_1 - x_2 = 0$, so the solution $x_1 = x_2 = 0$ has multiplicity 2. The elimination ideal $\mathcal{B}$ has the zero $x_1 = 0$ also with multiplicity 2, corresponding to the geometric idea that the two neighboring points on the x_1-axis are precisely the projection of the two neighboring points corresponding to $\mathcal{A}$. Kronecker's method would produce the same result here.

   (b) Example: $\mathcal{A} = (x_1, x_1 - x_2^2)$, $\mathcal{B} = (x_1)$. The zero $x_1 = x_2 = 0$ of $\mathcal{A}$ has multiplicity 2 again, and consists geometrically of this point and the neighboring point lying just off the x_2-axis. Under the projection, these two points collapse to a single point $x_1 = 0$, which is the simple zero of $\mathcal{B}$. Hence, the multiplicity of the corresponding zero in the elimination ideal has become smaller. We see immediately, however, that the one from $\mathcal{A}$ has multiplicity 2 because, in the complement, the quadratic equation $x_2^2 = 0$ must be solved, which has root $x_2 = 0$ with multiplicity 2. Kronecker's method would produce the result $x_1^2 = 0$ here, which would by chance lead to the correct result if we ignore the quadratic equation still to be solved.

[5] The most general and comprehensive resultant system (2) would be obtained by also allowing $\rho > 1$. We must then say: All of the polynomials of the polynomial ring $K[x_1, \ldots, x_{n-1}]$ for which a fixed power is contained in the ideal $\mathcal{A}$ form the elimination ideal $\mathcal{B}$ and whose basis is the resultant system (2). However, it is easy to see that this would suppress all of the existing multiplicities.

[6] See the detailed exposition on this in my book [2, p. 39].

[7] For the definition of "neighboring", see my book [2, p. 86].

(c) Example (from Perron[8]): $\mathcal{A} = (x_1^2, x_2^2, x_3^2)$ with the successive elimination ideals $\mathcal{B} = (x_1^2, x_2^2)$ and $\mathcal{C} = (x_1^2)$. Clearly the last ideal $\mathcal{C}$ has zero $x_1 = 0$ with multiplicity 2. However, since we must solve the two quadratic equations $x_2^2 = 0$ and $x_3^2 = 0$ in the course of complementing, it eventually follows that the solution $x_1 = x_2 = x_3 = 0$ for $\mathcal{A}$ has multiplicity $2 \cdot 2 \cdot 2 = 8$, agreeing with the expected outcome, either out of geometric considerations or by using Bezout's theorem. Yet as Perron explains, Kronecker's method leads eventually to the resulting system

$$x_1^4 = 0,\ x_1^2 x_2^2 = 0,\ x_2^4 = 0$$

and finally to $x_1^{16} = 0$. Thus the original system of equations would have multiplicity 16, and if we take into consideration the two remaining quadratic equations, the multiplicity would be 64, which is clearly absurd. As Perron shows, an arbitrary bilinear transformation does not change this result.

4. The resource of modern ideal theory allows us to adopt other completely different ways to solve our problems which both open new insights for theoretical work, as well as develop new computational methods for practical implementation. I would now like to exhibit explicitly such an ideal theoretic method in the case of a zero-dimensional ideal $\mathcal{A}$, that is, in the case where the system of equations (1) has only a finite number of solutions. The general case can always be reduced to this in a well-known way.

   If $\mathcal{A}$ is zero-dimensional, then the residue class ring $\mathcal{O} = K[x_1, \ldots, x_n]/\mathcal{A}$ is a commutative algebra (hypercomplex system) over the field K, *i.e.* there are only finitely many residue classes which are linearly independent over K.[9] Now everything depends on this algebra, and in particular on determining its multiplication table explicitly.

   For this purpose we proceed in the following manner: we consider the power products of the variables $x_1, \ldots, x_n$ ordered in a suitable, mostly lexicographic way[10]; for the sake of clarity, polynomials would also be written in this order. Now we denote the residue classes of the variables $x_1, \ldots, x_n$ by the sequence $u_1, \ldots, u_n$; if linearity occurs among the polynomials $f(x)$, this causes a linear dependency of the residue classes u in the residue class ring, and we can express the corresponding one or more of these u linearly in terms of the remaining ones and then delete it. In any case, the remaining u with their power products generate the entire residue class ring $\mathcal{O}$; but we must still take into account the relations expressed by the polynomials $f(x)$. Thus we will form the power products of

[8] See the work of Perron [7, p. 656]

[9] See my book [2, p. 96].

[10] A power product $x_1^{i_1} x_2^{i_2} \cdots x_n^{i_n}$ precedes the power product $x_1^{j_1} x_2^{j_2} \cdots x_n^{j_n}$ if the degree $i_1 + i_2 + \ldots + i_n < j_1 + j_2 + \ldots + j_n$, or for equal degrees if $i_1 > j_1$, or if $i_1 = j_1$, $i_2 > j_2$, etc.

u in the chosen order, and these will be denoted by the successively numbered symbols u_i, as long as they are shown not to be linearly dependent, due to the polynomials $f(x)$, on the previous ones.

Since the multiplication is commutative and associative, every power product can be decomposed into two factors in different ways. The multiplication must then be carried out using the formulas already determined and the results compared with each other. It can happen that they do not agree with each other; then a linear dependency of the previously introduced quantities u holds, which must be used to express one of these quantities linearly in terms of the remaining ones, and to strike it out of the scheme of linearly independent residue classes. This adjustment must then be carried out similarly in every relation where the deleted quantities occur, before we can consider the next power product.

Since the ideal $\mathcal{A}$ is zero-dimensional by assumption, only finitely many linearly independent residue classes can occur. We can be sure that no new linearly independent residue class can be added, as soon as for some fixed degree m, every power product has been proved linearly dependent on the preceding ones. We will therefore be at the end of this procedure if first, every polynomial $f(x)$ is taken into account, and second, the complete multiplication table is formed from u which are known to be independent. The commutative and associative property of the multiplication table follows directly from the applied procedure, since the properties were verified each time during the formation of the power products of the generating quantities u.

We have therefore found an algebra $K[1, u_1, \ldots, u_m]$ which is isomorphic to the residue class ring $\mathcal{O}$ because every polynomial $f(x)$ in the ideal $\mathcal{A}$, and only these, reduce to zero under the projection $x_i \mapsto u_i$ [11].

5. We illustrate these methods in a simple example: $\mathcal{A} = (x_2 + x_1x_2 + x_2^2,\ x_2 - x_1x_2 + x_2^2,\ x_1^2 - x_1^3)$. First we set $x_1 \mapsto u_1$ and $x_2 \mapsto u_2$. Then $x_1^2 \mapsto u_1^2 = u_3$ since no linear dependencies among u_1, u_2, and u_1^2 have appeared yet in the relations in question. $x_1x_2 \mapsto u_1u_2 = u_4$ follows similarly. For $x_2^2 \mapsto u_2^2$, we obtain $u_2^2 = -u_2 + u_4 = -u_2 - u_4$ by using both of the first two basis polynomials, so it follows that $u_4 = 0$ and $u_2^2 = -u_2$. Thus we must also delete u_4 in the preceding row and set $u_1u_2 = 0$.

   The next power product is $x_1^3 \mapsto u_1^3 = u_1u_1^2 = u_1u_3 = u_3$ as a consequence of the third basis polynomial. Further power products produce no new

[11] If one of the residue classes u_i $(i = 1, \ldots, n)$ is shown to be linearly dependent, then under the mapping $x_i \mapsto u_i$ it is of course replaced by the linear form corresponding to u_i. The higher power products of the x_i will be mapped onto the analogous power products of the u_i, which are then computed using the multiplication table. The residue class 1 corresponds to the unit element 1 of the algebra.

independent quantities, so $u_1^2 u_2 = u_1(u_1 u_2) = u_2 u_3 = 0$ since $u_1 u_2 = 0$ already holds. Similarly $u_1 u_2^2 = 0$ and $u_2^3 = u_2 u_2^2 = u_2(-u_2) = -u_2^2 = u_2$. If we form $u_1^4 = u_1 u_3 = u_3^2 = u_3$, then the multiplication table is complete and the procedure is terminated.

The residue class ring $\mathcal{O} = K[x_1, x_2]/\mathcal{A}$ is an algebra $\mathcal{O} = K[1, u_1, u_2, u_3]$ with multiplication table

$$\begin{array}{lll} u_1^2 = u_3 & u_1 u_2 = 0 & u_1 u_3 = u_3 \\ & u_2^2 = -u_2 & u_2 u_3 = 0 \\ & & u_3^2 = u_3 \end{array}$$

In order to find the algebraic variety corresponding to $\mathcal{A}$, we must now represent $\mathcal{A}$ as a reduced intersection of primary ideals by the theorem of Lasker and E. Noether. If the number field K is sufficiently extended algebraically, the prime ideals corresponding to $\mathcal{A}$ have only a single zero at each step. These zeros form the solutions to the system of equations corresponding to $\mathcal{A}$, the length of the intersecting primary ideals specifying the multiplicity of the solutions. By general ideal theoretic methods, the reduced representation can be found without major difficulty using the multiplication table. We recommend the following way:

Using our multiplication table, we form the sequence of powers of u_1 until we can determine a linear dependency between them. Since $u_1^2 = u_3$ and $u_1^3 = u_3$, it follows that $u_1^3 - u_1^2 = 0$, an equation whose roots are 0, 1. Similarly $u_2^2 + u_2 = 0$ holds, so $u_2 = 0, -1$. Because $u_1 u_2 = 0$, we find three prime ideals in the ring $\mathcal{O}$

$$\mathcal{P}_1 = (u_1, u_2), \qquad \mathcal{P}_2 = (u_1, u_2 + 1), \qquad \mathcal{P}_3 = (u_1 - 1, u_2).$$

Furthermore, we find that $\mathcal{P}_1^2 = (u_2, u_3) = \mathcal{P}_1^3 = \ldots, \mathcal{P}_2^2 = \mathcal{P}_2, \mathcal{P}_3^2 = \mathcal{P}_3$; hence the zero ideal $\mathcal{N}$ of the ring $\mathcal{O}$ is $\mathcal{N} = [\mathcal{P}_1^2, \mathcal{P}_2, \mathcal{P}_3]$ and the corresponding reduced representation of our ideal is

$$\mathcal{A} = [(x_1^2, x_2), (x_1, x_2 + 1), (x_1 - 1, x_2)].$$

¿From here we can immediately read off the zeros of our ideal, or the solutions of the system of equations. They are

$$\begin{array}{rl} x_1 = x_2 = 0 & \text{with multiplicity 2,} \\ x_1 = 0, x_2 = -1 & \text{with multiplicity 1,} \\ x_1 = 1, x_2 = 0 & \text{with multiplicity 1,} \end{array}$$

for a total of 4 zeros, agreeing with the fact that the residue class ring has 4 linearly independent quantities.

6. For about 17 years, I have applied and tested these methods in the most varied and complicated cases, and I believe I can say, on the basis of my experiences, that they in fact represent a useful and valuable tool for the solution of these and similar ideal theoretic tasks in every case. Because I have often been asked how one can most easily find the reduced representation of a polynomial ideal, I have now decided to publish the essential features of these methods, omitting the details. It is clear that in a few simple cases, the work, which lies in the determination of the multiplication table and further calculations, can become very large; but I believe we can be assured that this work is never bigger than that required by Kronecker's method. Most importantly, these methods produce a definitive explanation of the subtle structure and the multiplicity of the solutions.

References

[1] E. Bertini. *Introduzione alla geometria proiettiva degli iperspazi* [*Introduction to the Projective Geometry of Hyperspace*], 2nd ed. Messina, 1923.

[2] W. Gröbner. *Moderne algebraische Geometrie*. Vienna: Springer-Verlag, 1949.

[3] J. König. *Einleitung in die allgemeine Theorie der algebraischen Grössen* [*Introduction to the General Theory of Algebraic Quantities*]. Leipzig, 1903.

[4] W. Krull. *Theorie der Polynomideale und Eliminationstheorie* [*Polynomial Ideal Theory and Elimination Theory*]. Enzykl. math. Wiss., 2nd ed. Vol. 1, No. 5, 1939.

[5] O. Perron. *Über das Vahlensche Beispiel zu einem Satz von Kronecker* [*On Vahlen's Example for a Theorem of Kronecker*]. Math. Ann. **118** (1942): 441-448.

[6] O. Perron. *Über die Bedingungen, dass eine binäre Form n-ter Ordnung im R_n* [*On the Conditions that a Binary Form has Order n in R_n*]. Math. Ann. **118** (1941/3): 305-309.

[7] O. Perron. *Studien über den Vielfachheitsbegriff und den Bezoutschen Satz* [*Studies on the Concept of Multiplicity and Bezout's Theorem*]. Math. Z. **49** (1944): 654-680.

[8] F. Severi. Über die Darstellung algebraischer Mannigfaltigkeiten als Durchschnitte von Formen [*On the Representation of an Algebraic Variety as an Intersection of Forms*]. Abh. Math. Sem. Hamburg 15 (1943): 97-119.

[9] F. Severi. *Il concetto di molteplicità delle soluzioni pei sistemi di equazioni algebriche e la teoria dell'eliminazione* [*The Concept of Multiplicity of the Solutions of Systems of Algebraic Equations and Elimination Theory*]. Annali mat. pura appl. IV **26** (1947): 221-270.

[10] Vahlen. Crelles J. Math. **108** (1891).

[11] B.L. van der Waerden. *Moderne Algebra*, 2nd ed. Vol. 2. Springer-Verlag, 1940.

[12] B.L. van der Waerden. Zentralblatt für Math. **25** (1942).