

## NUMERICAL CALCULATION OF THE MULTIPLICITY OF A SOLUTION TO ALGEBRAIC EQUATIONS

HIDETSUNE KOBAYASHI, HIDEO SUZUKI, AND YOSHIHIKO SAKAI

ABSTRACT. A method to calculate numerically the multiplicity of a solution to a system of algebraic equations is presented. The method is an application of Zeuthen's rule which gives the multiplicity of a solution as the multiplicity of a united point of an algebraic correspondence defined naturally by the system. The numerical calculation is applicable to a large scale system of algebraic equations which may have a solution that we cannot calculate the multiplicity by a symbolic computation.

### 1. INTRODUCTION

In this paper, we assume a polynomial has coefficients in the field  $\mathbb{C}$  of complex numbers. We suppose that a system

$$(1) \quad f_1(x_1, \dots, x_n) = 0, \dots, f_n(x_1, \dots, x_n) = 0$$

of algebraic equations have a finite number of solutions. There are some methods to obtain the multiplicity of a solution by using symbolic calculations, but a symbolic computation is not executable for a large scale system. Our aim is to present a numerical method to calculate the multiplicity of a solution to a large system of algebraic equations.

Our algorithm is an application of Zeuthen's rule in algebraic geometry (see [1], [2]). Since we need some notations related to an algebraic correspondence, we introduce them briefly.

We denote by  $C$  the curve defined by a two variable polynomial  $\sigma(u, v)$  in the affine plane  $\mathbb{C}^2$ :

$$C = \{(a, b) \in \mathbb{C}^2 \mid \sigma(a, b) = 0\}.$$

If a point  $(a, b)$  is on the curve  $C$ , we say the number  $a$  *corresponds* to the number  $b$ . Thus the curve  $C$  gives an *algebraic correspondence* between complex numbers which we denote by  $T_C$ . The adjective "algebraic" is used to imply the correspondence is defined by a polynomial.

Let the polynomial  $\sigma(u, v)$  be of degree  $\alpha$  in  $u$  and  $\beta$  in  $v$ . For almost any number  $a$ , the one variable equation  $\sigma(a, v) = 0$  with respect to  $v$  has  $\beta$  roots. Similarly  $\sigma(u, b) = 0$  has  $\alpha$  roots for almost any number  $b$ . Hence we call the correspondence  $T_C$  an  $(\alpha, \beta)$ -*correspondence*. If a number  $a$  corresponds to itself under the correspondence  $T_C$ , we call  $a$  a *united point* of  $T_C$ . The united points

---

Received by the editor July 13, 1995 and, in revised form, August 28, 1996.

1991 *Mathematics Subject Classification*. Primary 14Q99, 65H10; Secondary 14N05, 65H20.

*Key words and phrases*. Multiplicity, Zeuthen's rule, homotopy method, system of algebraic equations.

are determined by the one variable equation  $\sigma(u, u) = 0$ . Counting the number of roots of the equation  $\sigma(u, u) = 0$  with multiplicity, we see that the correspondence  $T_C$  has  $\alpha + \beta$  united points<sup>1</sup>.

In the next section, we make from the system (1) an algebraic correspondence between numbers, and define the multiplicity of a solution to the system (1) as the multiplicity of a united point of the correspondence. And then we introduce the theorem “Zeuthen’s rule” from which we derive a method to calculate numerically the multiplicity of a united point. In Section 3, we present a method to calculate the multiplicity using Zeuthen’s rule. We show a simple application of Zeuthen’s rule gives only an approximate value of the multiplicity. In Section 4, we show a procedure which gives in almost any case an interval containing the multiplicity value. Since the multiplicity value is an integer, if the interval is small enough and contains only one integer, then we can obtain the exact value of the multiplicity. Moreover, we discuss an exceptional case where the interval obtained by the procedure fails to contain the multiplicity.

Section 5 is devoted to a discussion on execution time. The preliminary part of our method is executed by a symbolic computation to avoid making unnecessary numerical errors. The machine we used has a CPU with a vector processor but whose MIPS value is not high. So, symbolic computation took relatively longer execution time compared to that on commonly used work stations. In the last section, we present several examples. They show that the intervals are small enough to determine the exact value of the multiplicity. An example in Section 4 shows even if we encounter the rare case that we have good approximations of the multiplicity (but from one side).

## 2. ALGEBRAIC CORRESPONDENCE DEFINED BY THE SYSTEM

We consider the system (1) in  $n$ -dimensional projective space  $\mathbb{P}_{\mathbb{C}}^n$  with the homogeneous coordinate system  $(X_0, X_1, \dots, X_n)$ , where  $x_i = X_i/X_0$ . Then we have a system

$$(2) \quad F_i(X_0, X_1, \dots, X_n) = X_0^{d_i} f_i(X_1/X_0, \dots, X_n/X_0) = 0 \quad (i = 1, \dots, n)$$

of homogeneous polynomials. Here  $d_i$  is the degree of the polynomial  $f_i$  in the system (1). The system (2) has  $d_1 \cdots d_n$  solutions in  $\mathbb{P}_{\mathbb{C}}^n$  by Bézout’s theorem. If a point  $(a_1, \dots, a_n) \in \mathbb{C}^n$  is a solution to the system (1), then a point  $(1, a_1, \dots, a_n) \in \mathbb{P}_{\mathbb{C}}^n$  is a solution to the homogeneous system (2), and the multiplicity of these two solutions have the same value.

A set

$$(3) \quad F = \{Q \in \mathbb{P}_{\mathbb{C}}^n \mid F_1(Q) = 0\}$$

of zero points of the first equation of the system (2) is a hypersurface. Similarly, a set

$$(4) \quad C = \{P \in \mathbb{P}_{\mathbb{C}}^n \mid F_2(P) = \dots = F_n(P) = 0\}$$

of common zeros of polynomials from the system (2) is a space curve. We take a point  $O$  remote from  $C$  and  $F$ . Let  $P$  be a point on the curve  $C$ , and let  $l$  be a line joining  $P$  and  $O$ . Then we have intersection points of  $l$  and the hypersurface  $F$ . We

<sup>1</sup>Precisely, we have to consider a curve  $\tau(u_0, u_1; v_0, v_1) = u_0^\alpha v_0^\beta \sigma(u_1/u_0, v_1/v_0) = 0$  in bi-projective space  $\mathbb{P}_{\mathbb{C}}^1 \times \mathbb{P}_{\mathbb{C}}^1$  as an algebraic correspondence between projective lines  $\mathbb{P}_{\mathbb{C}}^1$ .

denote by  $Q$  one of the intersection points. We call a pair  $(P, Q; l)$  a *fundamental form*, and we denote by  $\Sigma_1$  the set of all fundamental forms.

In the next subsection, we define an algebraic correspondence between projective lines determined by the system (2).

**2.1. Construction of the correspondence.** We take an  $(n - 2)$ -dimensional linear subspace  $S$  in the projective space  $\mathbb{P}^n_{\mathbb{C}}$  such that

- not including the point  $O$ ,
- not meeting the curve  $C$ ,
- not included in the hypersurface  $F$ .

Let  $P$  be a point in the projective space. We denote a hyperplane spanned by  $S$  and  $P$  as  $S \vee P$ . Then it is easy to see that the set of hyperplanes (called a pencil of hyperplanes)

$$\mathcal{H} = \{S \vee P; P \in \mathbb{P}^n_{\mathbb{C}}\}$$

is parametrized by the point of the projective line  $\mathbb{P}^1_{\mathbb{C}}$ ; or, by the projective duality, we can identify a hyperplane in  $\mathcal{H}$  with a point in  $\mathbb{P}^1_{\mathbb{C}}$ .

Let  $\Sigma_1$  be the set of fundamental forms defined by the system (2) as above.

**Proposition 1.** *Let  $(P, Q; l)$  be a fundamental form in  $\Sigma_1$ . The correspondence of  $S \vee P$  to  $S \vee Q$  defines an algebraic  $(d_1 \cdots d_n, d_1 \cdots d_n)$ -correspondence between projective lines.*

*Proof.* Let  $H$  be a hyperplane of the pencil  $\mathcal{H}$ . Then  $H$  meets the curve  $C$  at  $d_2 \cdots d_n$  points  $P_j$ . A line  $l_j$  joining  $P_j$  and  $O$  meets the hypersurface  $F$  at  $d_1$  points  $Q_{j,k}$ . We have  $(P_j, Q_{j,k}, l_j) \in \Sigma_1$ , and the hyperplanes  $S \vee P_j \in \mathcal{H}$  all coincide. Hence to one hyperplane in  $\mathcal{H}$ , we have  $d_1 d_2 \cdots d_n$  corresponding hyperplanes.

Conversely, let  $H$  be a hyperplane of the pencil. Then  $H \cap F$  is an  $(n - 2)$ -dimensional algebraic variety of order  $d_1$ . A cone  $V$  based on  $H \cap F$  with a vertex  $O$  meets the curve  $C$  in  $d_1 \cdots d_n$  points  $P_j$ . Then a line  $l_j$  joining  $P_j$  and  $O$  meets  $H \cap F$  at points  $Q_j$ . Hence  $(P_j, Q_j; l_j) \in \Sigma_1$  and we see that to a hyperplane  $H$  there are  $d_1 d_2 \cdots d_n$  hyperplanes of the form  $S \vee P_j$ . This shows that the correspondence is of type  $(d_1 d_2 \cdots d_n, d_1 d_2 \cdots d_n)$ .

We show how we parametrize the hyperplanes of the pencil of hyperplanes  $\mathcal{H}$ . Let  $H_0(X_0, \dots, X_n) = 0$  and  $H_1(X_0, \dots, X_n) = 0$  be two distinct hyperplanes containing the linear space  $S$ . We parametrize  $\mathcal{H}$  as  $c_0 H_0 + c_1 H_1 = 0$ , where  $c_0, c_1$  are complex numbers not both zero. Then we have a bijective map from the set of pairs  $(c_0, c_1)$  of the numbers to a point in the projective line  $\mathbb{P}^1_{\mathbb{C}}$ . Thus we can identify  $\mathcal{H}$  and  $\mathbb{P}^1_{\mathbb{C}}$  (this identification is called projective duality). Thus we obtained a  $(d_1 \cdots d_n, d_1 \cdots d_n)$ -correspondence between projective lines.

Finally, we prove that the correspondence is algebraic. We parametrize original  $\mathcal{H}$  as

$$u_0 H_0(X_0, \dots, X_n) + u_1 H_1(X_0, \dots, X_n) = 0$$

and destination  $\mathcal{H}$  as

$$v_0 H_0(X_0, \dots, X_n) + v_1 H_1(X_0, \dots, X_n) = 0,$$

respectively. We have  $n$  equations

$$(5) \quad \begin{aligned} &u_0 H_0(p_0, \dots, p_n) + u_1 H_1(p_0, \dots, p_n) = 0, \\ &F_i(p_0, \dots, p_n) = 0 \quad (i = 2, \dots, n), \end{aligned}$$

with the coordinates  $P = (p_0, \dots, p_n)$  of the fundamental form  $(P, Q; l) \in \Sigma_1$ , and two equations

$$(6) \quad \begin{aligned} v_0 H_0(q_0, \dots, q_n) + v_1 H_1(q_0, \dots, q_n) &= 0, \\ F_1(q_0, \dots, q_n) &= 0, \end{aligned}$$

with the coordinates  $Q = (q_0, \dots, q_n)$ . Since three points  $P, Q$ , and  $O = (o_0, \dots, o_n)$  are colinear, we have  $n - 1$  equations

$$(7) \quad \begin{vmatrix} o_0 & o_1 & o_i \\ p_0 & p_1 & p_i \\ q_0 & q_1 & q_i \end{vmatrix} = 0 \quad (i = 2, \dots, n).$$

Using the general elimination method or Gröbner basis method, we can eliminate  $q_1, \dots, q_n$  from a system of  $n + 1$  equations (6) and (7). Then we have an equation

$$(8) \quad g(u_0, u_1; v_0, v_1; p_0, \dots, p_n) q_0^s = 0$$

for some non-negative integer  $s$ , where  $g(u_0, u_1; v_0, v_1; p_0, \dots, p_n)$  is a polynomial which is homogeneous in three groups of variables  $(u_0, u_1)$ ,  $(v_0, v_1)$  and  $(p_0, \dots, p_n)$  respectively. If we choose a homogeneous coordinate properly, we have only to consider points  $(p_0, \dots, p_n)$  and  $(q_0, \dots, q_n)$  such that  $p_0 \neq 0$  and  $q_0 \neq 0$ . So (8) is equivalent to

$$(9) \quad g(u_0, u_1; v_0, v_1; p_0, \dots, p_n) = 0.$$

Similarly, we can eliminate  $p_1, \dots, p_n$  from a system of  $n + 1$  equations (5) and (9), and have an equation

$$\sigma(u_0, u_1; v_0, v_1) p_0^t = 0$$

for some non-negative integer  $t$ , where  $\sigma(u_0, u_1; v_0, v_1)$  is a polynomial with variables  $(u_0, u_1)$  and  $(v_0, v_1)$ . Therefore we get a defining polynomial equation

$$(10) \quad \sigma(u_0, u_1; v_0, v_1) = 0$$

of the correspondence. This shows that the correspondence is algebraic.  $\square$

Hereafter, we denote by  $T$ , the correspondence of the projective line to itself derived as above.

**2.2. United points of the correspondence.** The  $(d_1 \cdots d_n, d_1 \cdots d_n)$ -correspondence  $T$  has  $2d_1 \cdots d_n$  united points which come from the hyperplanes such that  $S \vee P = S \vee Q$  with  $(P, Q; l) \in \Sigma_1$ . Coincidences of the hyperplanes arise from the two cases:

- (1) coincidence of the points  $P$  and  $Q$ . This means  $P$  is a solution to the system of algebraic equations;
- (2) the line  $PQ$  meets the  $(n - 2)$ -linear space  $S$ .

From the theory of algebraic correspondence, we see that the number of united points arising from the case (1) is  $d_1 d_2 \cdots d_n$ . This fact is taken to be a proof of Bézout's theorem. Hence, we see the multiplicity of united points coincide with the multiplicity of a solution under additional conditions to the point  $O$  and the  $(n - 2)$ -dimensional subspace  $S$ :

**[Condition 1]**

- (1) For each solution  $A$  of the system (2), the hyperplane  $S \vee A$  does not contain the other solution of (2).
- (2) A hyperplane  $S \vee O$  does not include any solution  $A$  to the system.

**2.3. Zeuthen’s rule.** Since the multiplicity of a united point  $(U_0, U_1) \in \mathbb{P}^1_{\mathbb{C}}$  of the correspondence  $T$  is a local property, we have only to consider a united point within an affine plane  $\mathbb{C}^2$  including a pair  $(U, U)$  instead of  $\mathbb{P}^1_{\mathbb{C}} \times \mathbb{P}^1_{\mathbb{C}}$ . Here  $U = U_1/U_0$  if  $U_0 \neq 0$  or  $U = U_0/U_1$  otherwise. *Zeuthen’s rule* giving an algorithm to calculate the multiplicity (see [1]) is stated as:

**Theorem 1. (Zeuthen’s rule)** *Let  $U$  be a united point of an algebraic correspondence between numbers defined by  $\sigma(u, v) = 0$ . Given a number  $\alpha$ , let  $\beta_1, \dots, \beta_s$  be roots to  $\sigma(\alpha, v) = 0$  such that they tend to  $U$  according as  $\alpha$  tends to  $U$ . Then if the order of the infinitesimal of  $|\beta_k - \alpha|$  measured by  $|U - \alpha|$  is  $e_k$ , that is*

$$|\beta_k - \alpha| \approx c_k |U - \alpha|^{e_k},$$

with constant  $c_k$ , the multiplicity of  $U$  is

$$\sum_{k=1}^s e_k.$$

By virtue of this theorem, we can calculate the multiplicity by a numerical calculation. We have to use a symbolic calculation to obtain the defining polynomial  $\sigma(u, v) = 0$  of the correspondence by repeated elimination procedures or the Gröbner basis method. For a large system of algebraic equations, the elimination procedure requires a lot of memory and it happens that we cannot execute the symbolic computation.

To make it a convenient form to employ a numerical calculation, we express Theorem 1 in terms of the fundamental forms.

**Theorem 2.** *Let  $C$  be the curve (4), and let  $F$  be the hypersurface (3). Let  $S$  be an  $(n - 2)$ -dimensional subspace satisfying condition 1. Let  $A$  be a solution to the system (2). By a projective transformation, we transform  $S$  in the hyperplane at infinity and  $A$  still in the finite area. Let  $H_0$  be a hyperplane  $S \vee A$ , and let  $H_h$  be a hyperplane shifting  $H_0$  to its normal direction by a small number  $h$ . Among the intersection points of  $H_h$  and the curve  $C$ , we suppose that  $P_1, \dots, P_r$  tend to  $A$  according as  $h$  tends to zero. Similarly among intersection points of the hypersurface  $F$  and the line  $l_j$  joining  $P_j$  and  $O$ , we suppose that  $Q_{j,1}, \dots, Q_{j,s}$  tend to  $A$  according as  $P_j$  tends to  $A$ .*

We measure the order of the infinitesimal of  $P_j Q_{j,k}$  by  $h$  as

$$(11) \quad P_j Q_{j,k} = c_{j,k} h^{e_{j,k}},$$

where  $P_j Q_{j,k}$  is the Euclidean distance of  $S \vee P_j$  and  $S \vee Q_{j,k}$  and  $c_{j,k}$  are constants; then the multiplicity of  $A$  is equal to

$$(12) \quad \sum_{j=1}^r \sum_{k=1}^s e_{j,k}.$$

Here we check that this theorem is equivalent to Theorem 1. We represent the hyperplanes  $S \vee A$ ,  $S \vee P_j$  and  $S \vee Q_{j,k}$  by ratio of parameters  $a$ ,  $p_j$ , and  $q_{j,k}$  respectively (see the next section). Since  $S \vee P_j = H_h$ , we have  $p_j = a + h$ . Then  $\sigma(a + h, q_{j,k}) = 0$  by definition of  $T_C$ , and by the hypothesis  $q_{j,k}$  tend to  $a$  according as  $h$  tends to zero. Hence Theorem 2 is the same as Theorem 1.

By Theorem 2 we can calculate directly the multiplicity of a solution to the system of algebraic equations. We will show in Section 4 that for almost any choice

of  $S$ , repeating this procedure four times with different  $h$ 's, we can obtain an interval containing the multiplicity value.

### 3. CALCULATION OF THE MULTIPLICITY

Changing the coordinate system if necessary, we may assume the  $(n-2)$ -dimensional subspace  $S$  is given as the intersection of two hyperplanes  $X_0 = 0$  and  $X_1 = 0$ . Then we can parametrize  $\mathcal{H}$  as  $u_0X_0 + u_1X_1 = 0$ . Any hyperplane in  $\mathcal{H}$  different from the hyperplane  $X_0 = 0$  at infinity is parametrized by numbers  $x_1 = X_1/X_0 = -u_0/u_1$ . Therefore, hyperplanes  $S \vee A, S \vee P_j$  and  $S \vee Q_{j,k}$  in Theorem 2 are parametrized by the  $x_1$ -coordinate of points  $A, P_j$  and  $Q_{j,k}$  respectively.

We have some remarks on numerical calculation.

- (1)  $\sum_{j,k} e_{j,k}$  becomes a good approximation if the original  $h$  is sufficiently near to zero. To eliminate constants  $c_{j,k}$  in the equation (11), we take two numbers very close to zero  $h_\nu$  ( $\nu = 1, 2$ ), and calculate the ratio:

$$(13) \quad e_{j,k} = \frac{\log |p_{j,1}^{(1)} - q_{j,k,1}^{(1)}| - \log |p_{j,1}^{(2)} - q_{j,k,1}^{(2)}|}{\log |h_1| - \log |h_2|},$$

where  $P_j^{(\nu)} = (p_{j,1}^{(\nu)}, \dots, p_{j,n}^{(\nu)})$  and  $Q_{j,k}^{(\nu)} = (q_{j,k,1}^{(\nu)}, \dots, q_{j,k,n}^{(\nu)})$  are obtained from  $h_\nu$  ( $\nu = 1, 2$ ).

- (2) Let  $A = (a_1, \dots, a_n)$  be a coordinate representation. To distinguish clearly two united points  $S \vee A$  and  $S \vee O$ , we take a point  $O = (o_1, \dots, o_n)$  not included in  $F$  and  $C$  and  $|o_1 - a_1| > |h_\nu|$  (for both  $h_1$  and  $h_2$ ).
- (3) Since we need only points  $P_j^{(\nu)}$  on hyperplanes  $x_1 = a_1 + h_\nu$  close to  $A$ , we take points  $P_j^{(\nu)}$  within hypercubes

$$|x_1 - a_1| < ch_\nu, |x_2 - a_2| < ch_\nu^{1/(d_2 \cdots d_n)}, \dots, |x_n - a_n| < ch_\nu^{1/(d_2 \cdots d_n)},$$

( $\nu = 1, 2$ ). Here  $c$  is a positive number determined experimentally. These boundaries are given by the exponent of Puiseux expansions of  $C$  around  $A$ .

Now we present our method to calculate the multiplicity step by step.

**[Calculation of the multiplicity]**

- (1) Calculate a root  $A$  to the system (1) using a numerical method (e.g. Newton's method or the homotopy method).
- (2) Take a point  $O$  remote from  $F, C$  and  $A$ .
- (3) Take two small values  $h_\nu$  ( $\nu = 1, 2$ ).
- (4) Calculate all intersection points  $P_j^{(\nu)}$  of the curve  $C$  and the hyperplane  $x_1 = a_1 + h_\nu$  within the bound given above.
- (5) Calculate all intersection points  $Q_{j,k}^{(\nu)}$  of the hypersurface  $F$  and the line joining  $O$  and  $P_j^{(\nu)}$ .
- (6) Calculate the exponents  $e_{j,k}$  by the equation (13).
- (7)  $\sum_{j,k} e_{j,k}$  is the multiplicity of  $A$ .

**Note 1.** Since we choose  $S$  and  $O$  without testing whether **condition 1** of the previous section is satisfied or not, the multiplicity value may be wrong in very few cases. But if we calculate the multiplicity with some different choices of  $S$  and  $O$ , then we get true values of the multiplicity.

**Note 2.** If some roots of the system are very close to each other, we cannot see whether a root in the problem is one of the clustered roots or a multiple root. In such a case, we have to separate such close roots (see [3]).

4. AN INTERVAL CONTAINING THE MULTIPLICITY VALUE

We can obtain an approximate value of the multiplicity in the above way. In this section, we show we can determine the exact value of the multiplicity for almost any choice of  $O$  and  $S$ .

We denote by  $m$  the multiplicity of a solution  $A$ . We take two small values  $h_1$  and  $h_2$  such that  $h_1 > h_2 > 0$ , and denote by  $m^+$  the approximate multiplicity value obtained from  $h_1$  and  $h_2$ . Similarly we denote by  $m^-$  the approximate value from  $-h_1, -h_2$ .

We assume that  $S$  is represented as the intersection of the hyperplane  $X_0 = 0$  and the hyperplane  $X_1 = 0$ ; and we remember that we are taking the hyperplane  $X_0 = 0$  as the hyperplane at infinity.

**Proposition 2.** *If we take a point  $O$  such that the line joining two points  $P_j$  and  $O$  does not meet a hypersurface  $F$  at infinity, we have almost always*

$$m^- \leq m \leq m^+ \quad \text{or} \quad m^+ \leq m \leq m^-.$$

*Proof.* For simplicity, we assume that the parametric value of the united  $S \vee A$  is zero. Then for each  $S \vee P_j^{(\nu)}$  the parametric value is  $h_\nu$  and the defining equation  $\sigma(u, v) = 0$  of the correspondence  $T$  satisfies  $\sigma(0, 0) = 0$ .

We represent  $\sigma(u, v)$  as a polynomial in a variable  $v$ ,

$$\sigma(u, v) = \tau_0(u)v^d + \tau_1(u)v^{d-1} + \dots + \tau_d(u),$$

where  $d = d_1 \dots d_n$  by Proposition 1. By the hypothesis that the line joining two points  $P_i$  and  $O$  does not meet  $F$  at infinity, we see that  $\tau_0(0) \neq 0$ . Therefore we can set

$$\sigma_0(u, v) = \frac{\sigma(u, v)}{\tau_0(u)} = v^d + \rho_1(u)v^{d-1} + \dots + \rho_d(u),$$

where  $\rho_1(u), \dots, \rho_d(u)$  are formal power series in a variable  $u$ .

Now we suppose that a polynomial  $\sigma_0(0, v)$  in a variable  $v$  is factorized as

$$\sigma_0(0, v) = \mu_1(v)\mu_2(v), \quad \text{s.t. } \mu_1(0) = 0 \text{ and } \mu_2(0) \neq 0.$$

By Hensel's Lemma (see [2]),  $\sigma_0(u, v)$  is factorized as

$$\sigma_0(u, v) = \sigma_1(u, v)\sigma_2(u, v), \quad \text{with } \sigma_1(0, v) = \mu_1(v) \text{ and } \sigma_2(0, v) = \mu_2(v),$$

where  $\sigma_1(u, v)$  and  $\sigma_2(u, v)$  are polynomials in  $v$  having formal power series in  $u$  as coefficients, and  $\sigma_1(u, v)$  converges for  $u$  sufficiently close to zero. When  $\sigma_0(0, v) = 0$  does not have roots different from zero, we set merely  $\sigma_0(u, v) = \sigma_1(u, v)$ .

Hence, roots  $v_k^{(\nu)}$  to  $\sigma(h_\nu, v) = 0$  satisfying

$$\lim_{h_\nu \rightarrow 0} v_k^{(\nu)} = 0$$

are also roots to  $\sigma_1(h_\nu, v) = 0$ . Since the equation (13) is represented as

$$e_k = \log_{|h_1/h_2|} \left| \frac{v_k^{(1)} - h_1}{v_k^{(2)} - h_2} \right|$$

using a logarithm function with the base  $|h_1/h_2|$ , we have

$$\sum_{k=1}^s e_k = \log_{|h_1/h_2|} \left| \prod_{k=1}^s \frac{v_k^{(1)} - h_1}{v_k^{(2)} - h_2} \right| = \log_{|h_1/h_2|} \left| \frac{\sigma_1(h_1, h_1)}{\sigma_1(h_2, h_2)} \right|.$$

Now, we represent  $\sigma_1(u, u)$  as a formal power series

$$\sigma_1(u, u) = cu^m(1 + c_1u + c_2u^2 + \dots)$$

in a variable  $u$ . Since  $h_\nu$  are sufficiently small, we have

$$\sum_{k=1}^s e_k \approx \log_{|h_1/h_2|} \left| \frac{ch_1^m(1 + c_1h_1)}{ch_2^m(1 + c_1h_2)} \right|$$

after ignoring the terms in  $h_\nu$  of order higher than  $m + 2$ . Namely,

$$m^+ \approx m + \log_{|h_1/h_2|} \left| \frac{1 + c_1h_1}{1 + c_1h_2} \right|, \quad m^- \approx m + \log_{|h_1/h_2|} \left| \frac{1 - c_1h_1}{1 - c_1h_2} \right|.$$

Therefore, we have only to show that

$$(14) \quad \left| \frac{1 + c_1h_1}{1 + c_1h_2} \right| \leq 1 \leq \left| \frac{1 - c_1h_1}{1 - c_1h_2} \right|$$

or

$$(15) \quad \left| \frac{1 - c_1h_1}{1 - c_1h_2} \right| \leq 1 \leq \left| \frac{1 + c_1h_1}{1 + c_1h_2} \right|.$$

Let us set  $h = h_1$ ,  $h_2 = dh$  ( $0 < d < 1$ ) and  $c_1 = \alpha + \beta i$ . Multiplying  $1 + dah - d\beta hi$  to both numerator and denominator and ignoring the terms of order  $O(h^2)$  or higher, we have

$$\left| \frac{1 + c_1h_1}{1 + c_1h_2} \right| = \left| \frac{1 + \alpha h + \beta hi}{1 + d\alpha h + d\beta hi} \right| \approx \left| \frac{1 + (1 + d)\alpha h - \alpha\beta hi}{1 + 2d\alpha h} \right|.$$

We have three cases  $\alpha > 0$ ,  $\alpha < 0$  and  $\alpha = 0$ . If  $\alpha > 0$ , since  $0 < 2d\alpha h < (1 + d)\alpha h$ , we have

$$\left| \frac{1 - (1 + d)\alpha h}{1 - 2d\alpha h} \right| < 1 < \left| \frac{1 + (1 + d)\alpha h}{1 + 2d\alpha h} \right|;$$

and since  $h$  is a very small positive number, we have  $m^- < m < m^+$ .

If  $\alpha < 0$ , since  $(1 + d)\alpha h < 2d\alpha h < 0$ , we have

$$\left| \frac{1 + (1 + d)\alpha h}{1 + 2d\alpha h} \right| < 1 < \left| \frac{1 - (1 + d)\alpha h}{1 - 2d\alpha h} \right|,$$

i.e.  $m^+ < m < m^-$ .

If  $\alpha = 0$ , since the term in  $h$  of order one vanishes, we cannot see whether the equations (14, 15) are satisfied or not. But since  $c_1 = \alpha + \beta i$  and  $c_1$  depends on  $O$ , if we choose  $O$  properly, then we can get a non-zero  $\alpha$ . □

Summing up the calculation of an interval, we can write as

**[Calculation of an interval]**

- (1) Select two small positive numbers  $h_\nu$  ( $\nu = 1, 2$ ).
- (2) Repeat multiplicity algorithm for  $h_\nu$  ( $\nu = 1, 2$ ) and  $-h_\nu$  ( $\nu = 1, 2$ ).

**Note 3.** In the statement of Proposition 2, we used the words “almost always”. This means the statement is valid except for the case where  $\alpha$  (= the real part of the coefficient  $c_1$ ) in the above proof is zero. This case is very rare as illustrated in the following Example. Moreover, this example shows that even if this case happens, we have two good approximations of the multiplicity from one side of the multiplicity value.



**Example.** We show an example of one sided approximations:

$$(16) \quad \begin{cases} x - y = 0, \\ x - y^2 = 0. \end{cases}$$

We calculate the multiplicity of  $(0, 0)$  choosing the point  $O$  as  $(o_1, 0)$  with a variable coordinate  $o_1$ .

Multiplicity with $o_1 = 1$		
$h_1, h_2$	0.01, 0.001	-0.01, -0.001
multiplicity	$M^+ = 1.00009$	$M^- = 1.00008$

From this table we see that the multiplicity is 1, and this is apparently the correct value. To show the reason why this one sided approximation occurs, we calculate the homogeneous polynomial defining the image curve of the correspondence:

$$o_1^2 u_0^2 v_1^2 - u_1(o_1^2 u_0 v_0^2 - 2o_1 u_0 v_0 v_1 + u_0 v_1^2 + 2o_1 u_0 v_1^2 - u_1 v_1^2).$$

From this polynomial, by setting  $u = u_1, u_0 = 1, v = v_1$  and  $v_0 = 1$ , we have the affine defining polynomial:

$$-o_1^2 u + 2o_1 uv + (o_1^2 - u - 2o_1 u + u^2)v^2.$$

Calculating the Puiseux expansions of this curve at  $(0,0)$ , we have two branches:

$$(17) \quad \begin{aligned} b_1 &= u^{1/2} - 1/o_1 u + (1 + o_1)/o_1^2 u^{3/2} + \dots, \\ b_2 &= -u^{1/2} - 1/o_1 u - (1 + o_1)/o_1^2 u^{3/2} + \dots. \end{aligned}$$

Then the defining polynomial  $\sigma_1(u, v) = (v - b_1)(v - b_2)$  is expanded as

$$\frac{-o_1^4 u - 2o_1^3 u^2 + o_1^4 v^2 + 2o_1^3 uv - o_1^2 u^3 - o_1^2 u^2 - 2o_1 u^3 - u^3 + \text{higher order terms}}{o_1^4}.$$

Putting  $v = u$ , we have

$$\sigma_1(u, u) = \frac{-o_1^4 u - (o_1^2 - o_1^4)u^2 - (o_1^2 + 2o_1 + 1)u^3 + O(4)}{o_1^4}.$$

If we choose  $-1$  or  $1$  for  $o_1$  (roots of  $o_1^2 - o_1^4 = 0, o_1$  should not be equal to 0), the coefficient of  $u^2$  vanishes and we have

$$\sigma_1(u, u) = -u + O(4)$$

or

$$\sigma_1(u, u) = -u - 4u^3 + O(4).$$

The last expression shows that the approximation is only from one side and the former one shows that we cannot say whether we can obtain an interval or not.  $\square$

### 5. EXECUTION TIME

We discuss execution time required to calculate the multiplicity. If the scale of a system of algebraic equations is not so large, we can obtain all solutions with multiplicity by tracing the paths in the homotopy method. Since the homotopy method is well known, we take the execution time of the homotopy method as a standard, and we compare with the standard the execution time spent to calculate an interval containing the multiplicity value.

First we show the detail of our calculation of the homotopy method. We use both symbolic and numerical computation. Given a system of algebraic equations

$f_1(x), \dots, f_n(x)$  (resp.  $g_1(x), \dots, g_n(x)$ ) we denote by  $F(x)$  (resp.  $G(x)$ ) the vector  $(f_1(x), \dots, f_n(x))$  (resp.  $(g_1(x), \dots, g_n(x))$ ).

- (1) output  $F(x)$  and  $G(x)$  to a file for later numerical computation— symbolic computation
- (2) make the homotopy  $H(x, t) = tF(x) + (1 - t)G(x)$  and output it to a file — symbolic computation
- (3) make a matrix  $DH = [dH/dx \quad dH/dt]$  and output it to a file — symbolic computation
- (4) make a matrix  $DF = [dF/dx]$  and output it to a file — symbolic computation
- (5) tracing paths by the predictor-corrector method — numerical computation

To execute the calculation of an interval, we proceed as follows:

- (1) calculate the point  $A$  — numerical computation
- (2) select the point  $O$  — no computation
- (3) select the values  $h_\nu$  — no computation
- (4) calculate points  $P_j^{(\nu)}$ 
  - (a) substitute  $a_1 + h_\nu$  to  $x_1$  in  $f_2(x), \dots, f_n(x)$  — symbolic computation
  - (b) call the homotopy method above for  $f_2(x), \dots, f_n(x)$
- (5) calculate points  $Q_{j,k}^{(\nu)}$ 
  - (a) make a polynomial of single variable  $f_1(\alpha_1 x_1 + \beta_1, \dots, \alpha_{n-1} x_1 + \beta_{n-1}, x_1)$  and output its coefficients in a file — symbolic computation
  - (b) determine  $\alpha$ 's and  $\beta$ 's — numerical computation
  - (c) solve an algebraic equation to obtain the points  $Q$  — numerical computation
- (6) calculate the exponents  $e_{j,k}$  — numerical computation
- (7) calculate  $\sum_{j,k} e_{j,k}$  — numerical computation
- (8) repeat 3. - 7. for  $-h_\nu$

In the following table, we compare the execution time for the homotopy method and for the calculation of an interval containing the multiplicity value. Size means the product of degrees of polynomials in the given system of algebraic equations. The computation was executed on an FUJITSU VPX-120 with 17 MIPS, scalar calculation 17 MFLOPS and vector calculation Max, 170 MFLOPS CPU.

In Table 1, we see symbolic computation takes more than half of the total execution time in the calculation of an interval. But, if the system of algebraic equations is large, then tracing paths becomes heavy. And if we use a work station, we can expect that the symbolic calculations are executed in a shorter time (maybe less than 1/10 each). So we compare execution time of path tracing of the homotopy method and that of calculation of an interval. When we trace paths of the homotopy, we cannot predict after how many steps we reach the goal. So, we cannot give the complexity. However, we have a simple formula that gives approximately the ratio of execution time of homotopy and calculation of the interval as we show in the following.

When we solve the system (1) by the homotopy method, the number of paths are

$$d_1 \times d_2 \times \cdots \times d_n$$

where  $d_i$  is the total degree of the polynomial  $f_i$  in (1). As the mean value of path tracing in  $n$ -variable homotopy, we write  $\phi(n)$ . Then the total execution time will

TABLE 1. Execution time (seconds)

Size	method	Symbolic	Numeric (homotopy)	Numeric (others)	Total
4 × 4	homotopy	0.270	2.399	0	2.669
	interval	2.410	0.222 × 4	0.014	3.312
6 × 6	homotopy	0.240	7.033	0	7.273
	interval	2.800	0.676 × 4	0.062	5.566
10 × 10	homotopy	0.270	32.84	0	33.11
	interval	4.150	1.447 × 4	0.282	10.22
3 × 2 × 3	homotopy	0.240	5.482	0	5.722
	interval	3.600	1.226 × 4	0.042	8.546
5 × 5 × 5	homotopy	1.340	142.6	0	143.9
	interval	25.46	14.47 × 4	0.215	83.56
7 × 7 × 7	homotopy	1.670	938.0	0	940.0
	interval	45.11	64.24 × 4	0.440	302.5
3 × 3 × 3 × 3	homotopy	0.940	32.22	0	33.16
	interval	25.61	7.131 × 4	0.194	54.33
4 × 4 × 4 × 4	homotopy	1.810	284.3	0	286.1
	interval	51.47	43.43 × 4	0.482	225.7

be approximately

$$\phi(n) \times d_1 \times d_2 \times \dots \times d_n.$$

In the calculation of the interval, we eliminate  $x_j$  from  $x_1, x_2, \dots, x_n$  by substituting the number  $a_j + h_\nu$ . We introduce the symbol  $d_i^{(j)}$  defined as

$$d_i^{(j)} = \deg f_i(x_1, \dots, x_j = a_j + h_\nu, \dots, x_n),$$

where  $x_j = a_j + h_\nu$  means that we substitute a constant number  $a_j + h_\nu$  to  $x_j$ . So  $d_i^{(j)}$  is the degree of the polynomial  $f_i(x_1, \dots, x_j = a_j + h_\nu, \dots, x_n)$ . We have  $d_i^{(j)} \leq d_i$  for each  $j$ . Let  $D_r^{(j)}$  be the product

$$D_r^{(j)} = \min_{1 \leq r \leq n} \left\{ \prod_{i \neq r, i=1}^n d_i^{(j)} \right\}.$$

If we choose an index  $j$  satisfying

$$D_r^{(j)} = \min_{1 \leq i \leq n} D_r^{(i)},$$

then the number of paths to trace is minimal for the computation to obtain the interval containing the multiplicity value. In our program, we fix simply  $j = 1, r = 1$ , we eliminate  $x_1$  and we treat a system of algebraic equations  $f_2, \dots, f_n$  as follows. We trace paths for the new system of algebraic equations  $f_2, \dots, f_n$  with  $x_1$  eliminated. In this case the number of total paths is

$$d_2^{(1)} \times d_3^{(1)} \times \dots \times d_n^{(1)};$$

and since the number of variables is  $n - 1$ , the execution time of the homotopy method for this new system is

$$\phi(n - 1) \times d_2^{(1)} \times d_3^{(1)} \times \dots \times d_n^{(1)}.$$

In the calculation of the interval, since we repeat the homotopy method with four constants  $\pm h_1, \pm h_2$ , the execution time of the calculation of the interval is 4 times the above value:

$$4 \times \phi(n-1) \times d_2^{(1)} \times d_3^{(1)} \times \cdots \times d_n^{(1)}.$$

The ratio of calculation time is

$$r = \frac{4 \times \phi(n-1) \times d_2^{(1)} \times d_3^{(1)} \times \cdots \times d_n^{(1)}}{\phi(n) \times d_1 \times d_2 \times \cdots \times d_n}.$$

If  $d_i^{(1)} = d_i$ , then the above formula is expressed simply as

$$r = \frac{4\phi(n-1)}{\phi(n)d_1}.$$

If we set  $\phi(n) = n$ , then we have the formula

$$(18) \quad r = \frac{4(n-1)}{nd_1}.$$

In the following table, we see that a ratio given by the formula (18) is a rough approximation to an experimental value.

Formula	0.50	0.33	0.20	0.89	0.53	0.38	1.00	0.75
Experiment	0.37	0.38	0.18	0.89	0.41	0.27	0.89	0.61

## 6. EXAMPLES

We present some examples of calculation of the interval containing the multiplicity value.

**Example 1.** The multiplicity of a solution  $(x, y) = (0, 0)$  to the system

$$(19) \quad \begin{cases} (x^2 + y^2)^2 + 3x^2y - y^3 = 0, \\ (x^2 + y^2)^3 - 4x^2y^2 = 0 \end{cases}$$

is 14 (see [4]). Table 2 is the result of the calculation of the interval after setting  $O = (1, 1)$ . Since there is only one integer 14 in the interval  $[m^+, m^-]$ , we see that the multiplicity of  $(0, 0)$  is 14.

**Example 2.** A system<sup>2</sup>

$$(20) \quad f_i(x_1, \dots, x_5) = x_i^2 + \sum_{j=1}^5 x_j - 2x_i - 4 \quad (i = 1, \dots, 5)$$

has a solution  $(1, 1, 1, 1, 1)$  with the multiplicity *sixteen*. Using the homotopy

TABLE 2. Multiplicity to the system (19)

$h_1, h_2$	0.01, 0.001	-0.01, -0.001
multiplicity	$m^+ = 13.9679$	$m^- = 14.0321$

<sup>2</sup>This example is taken from [5]. But the original coordinate system does not satisfy *Condition 1*, so we use a new coordinate system.

TABLE 3. Multiplicity to the system (20)

$h_1, h_2$	0.001, 0.0009	-0.001, -0.0009
multiplicity	$M^+ = 15.9858$	$M^- = 16.0014$

TABLE 4. Multiplicity to the system (21)

$h_1, h_2$	0.001, 0.0009	-0.001, -0.0009
multiplicity	$M^+ = 638.3011$	$M^- = 637.6967$

method, we get the numerical solution

$$(1.000000004 - 0.1261539688 \times 10^{-14}i, 1.000000000 + 0.3228171677 \times 10^{-14}i,$$

$$1.000000006 - 0.8244469500 \times 10^{-15}i, 1.000000003 - 0.1568655907 \times 10^{-14}i,$$

$$0.9999999868 + 0.4264708798 \times 10^{-15}i).$$

To this numerical solution, Table 3 shows the bounds obtained by the calculation of the interval, where we take a point  $O$  as  $(-1.1, 1, 0, 0, 0)$ . We see the multiplicity is sixteen.

**Example 3.** A system

$$(21) \quad f_i(x_1, \dots, x_{10}) = x_i^2 + \sum_{j=1}^{10} x_j - 2x_i - 9 \quad (i = 1, \dots, 10)$$

has a solution  $(1, 1, 1, 1, 1, 1, 1, 1, 1, 1)$ . Using the homotopy method, we obtain the numerical solution

$$(1.000000002 - 0.5003286613 \times 10^{-08}i, 0.9999999991 + 0.4396894894 \times 10^{-08}i,$$

$$0.9999999956 + 0.5617517134 \times 10^{-08}i, 0.9999999992 + 0.4789822539 \times 10^{-08}i,$$

$$0.9999999982 + 0.4748659958 \times 10^{-08}i, 1.000000002 - 0.4640685652 \times 10^{-08}i,$$

$$1.000000003 - 0.5154107100 \times 10^{-08}i, 1.000000004 - 0.2181786934 \times 10^{-07}i,$$

$$0.9999999904 - 0.1821846643 \times 10^{-08}i, 1.000000003 - 0.4562487468 \times 10^{-08}i).$$

To this numerical solution, we have Table 4 by the computation of an interval with  $O = (-1, 1, 1, 0, \dots, 0)$ . We have the multiplicity 638.

REFERENCES

- [1] J. S. Semple and L. Roth : *Introduction to Algebraic Geometry*, Oxford Science Publications, Clarendon Press, Oxford, 1949. MR 11:535d
- [2] B. L. van der Waerden : *Einführung in die algebraische Geometrie*, Springer-Verlag, 1973. MR 49:8984
- [3] H. Kobayashi, H. Suzuki and Y. Sakai: *Separation of close roots by linear fractional transformation*, Proc. of ASCM, 1-10, Scientists INC, 1995.

- [4] W. Fulton: *Algebraic Curves*, The Benjamin/Cummings Publishing Company, 1969. MR **47**:1807
- [5] A. H. Wright: *Finding All Solutions to a System of Polynomial Equations*, Mathematics of Computation Vol. 44, No. 169, 1985, 125–133. MR **86i**:12001

DEPARTMENT OF MATHEMATICS, NIHON UNIVERSITY, 1-8 KANDA-SURUGADAI, TOKYO 101,  
JAPAN

*E-mail address*: `hikoba@math.cst.nihon-u.ac.jp`

TOKYO POLYTECHNIC COLLEGE, 2-32-1 OGAWA-NISHI KODAIRA, TOKYO, JAPAN

VISUAL SCIENCE LABORATORY, INC., 2-21 KANDA-AWAJICHO, CHIYODA, TOKYO, JAPAN

*E-mail address*: `sakai@vsl.co.jp`