

Cramer-type Formula for the Polynomial Solutions
of Coupled Linear Equations with Polynomial Coefficients

佐々木 建昭
Tateaki Sasaki

The Institute of Physical and Chemical Research (理研)
Wako-shi, Saitama 351, Japan

ABSTRACT

This paper derives a determinant form formula for the general solution of coupled linear equations with coefficients in $K[x_1, \dots, x_n]$, where K is a field of numbers, the number of unknowns is greater than the number of equations, and the solutions are in $K(x_1, \dots, x_{n-1})[x_n]$. The formula represents the general solution by the minimum number of generators, and it is a generalization of Cramer's formula for the solutions in $K(x_1, \dots, x_n)$. Compared with another formula which is obtained by a method typical in algebra, the generators in our formula are represented by determinants of quite small orders.

This paper is published in Publ. RIMS (Kyoto Univ.)

$$\left\{ \begin{array}{l} P_{11}(x_1, \dots, x_n)y_1 + \dots + P_{1s}(x_1, \dots, x_n)y_s = P_{1,s+1}(x_1, \dots, x_n), \\ \quad \quad \quad \cdot \quad \quad \quad \cdot \quad \quad \quad \cdot \quad \quad \quad \cdot \\ P_{r1}(x_1, \dots, x_n)y_1 + \dots + P_{rs}(x_1, \dots, x_n)y_s = P_{r,s+1}(x_1, \dots, x_n), \end{array} \right. \quad (1)$$

It is well known that coupled linear Diophantine equations with coefficients in Z are solved by the Euclidean algorithm and the general solution is represented by generators the number of which is not greater than $s-r$ if there exist solutions. Since the Euclidean algorithm applies also to polynomials in $S[x]$, we can solve (1) by applying the Euclidean algorithm successively and obtain $s-r$ generators. However, the procedure is quite tedious and often causes severe coefficient growth (see, for example, ref. 2).

As for coupled linear equations over fields, Cramer's formula gives the general solution in a determinant form. We may, therefore, well expect the existence of determinant form formula which gives all the s-r generators of the solutions of (1). With such a formula, we can calculate the solutions easily without introducing unnecessary coefficient growth. The purpose of this paper is to derive such a formula.

§2. Basic lemmas

In this paper, the variable $x_n = x$ is treated as the main variable, and the degree and the leading coefficient in the main variable x of polynomial P are represented by $\deg(P)$ and $\text{lc}(P)$, respectively. Furthermore, the resultant of polynomials F and G in x is represented by $\text{res}(F,G)$. The greatest common divisor, to be abbreviated to GCD, over the field K is defined by omitting numeric factor which is unit in K .

The following two lemmas are essential in solving (1).

[Lemma 1]. Given polynomials F, G, H in $K[x_1, \dots, x_n]$ satisfying

$$\begin{cases} \text{GCD}(F,G) = 1, \\ \deg(F) + \deg(G) > \deg(H), \end{cases} \quad (2.1)$$

there exist polynomials A and B in $K[x_1, \dots, x_n]$ such that

$$\begin{cases} AF + BG + RH = 0, \\ R \equiv \text{res}(F, G), \\ \deg(A) < \deg(G), \quad \deg(B) < \deg(F). \end{cases} \quad (2.2)$$

(Proof) Let $\deg(F) = \ell$, $\deg(G) = m$, $\deg(H) = k$, and represent F, G, H as

$$\begin{cases} F = f_\ell x^\ell + f_{\ell-1} x^{\ell-1} + \dots + f_0, & f_\ell \neq 0, \\ G = g_m x^m + g_{m-1} x^{m-1} + \dots + g_0, & g_m \neq 0, \\ H = h_k x^k + h_{k-1} x^{k-1} + \dots + h_0, & h_k \neq 0. \end{cases} \quad (2.3)$$

Multiplying F, G, and H to, respectively, the last columns of the following determinants A, B, and R, and considering determinants which are coefficients of x^i , $i=0, \dots, \ell+m-1$, we see that (2.2) is satisfied:

$$A = \begin{array}{ccccccccc} f_\ell & f_{\ell-1} & \cdot & \cdot & \cdot & f_0 & x^{m-1} \\ & f_\ell & f_{\ell-1} & \cdot & \cdot & f_0 & x^{m-2} \\ & & \cdot & \cdot & \cdot & \cdot & \cdot \\ & & & f_\ell & f_{\ell-1} & \cdot & \cdot & f_0 & x^0 \\ g_m & g_{m-1} & \cdot & \cdot & \cdot & g_0 & 0 \\ & g_m & g_{m-1} & \cdot & \cdot & g_0 & 0 \\ & & \cdot & \cdot & \cdot & \cdot & \cdot \\ & & & g_m & g_{m-1} & \cdot & \cdot & g_0 & 0 \\ & & & & h_k & h_{k-1} & \cdot & \cdot & \cdot & h_0 & 0 \end{array} \left. \begin{array}{l} \text{m rows} \\ \text{\ell rows} \end{array} \right\} \quad (2.4)$$

← $\ell+m+1$ columns →

$$B = [\text{replace the last column of } A \text{ by } (0 \dots 0 x^{\ell-1} \dots x^0 0)^T], \quad (2.5)$$

$$R = [\text{replace the last column of } A \text{ by } (0 \dots 0 0 \dots 0 1)^T]. //$$

Notes: Expanding R with respect to the rightmost column, we obtain famous Sylvester's determinant for the resultant. The above determinants and Lemma 1 were discovered in the process of generalizing the polynomial remainder sequence in ref. 1. See, also refs. 2 and 3. It is easy to prove that A and B in the above lemma are unique.

Removing the degree restriction in (2.1), we obtain the following lemma.

[Lemma 2] Let F, G, H, and R be the same as those in Lemma 1 except that $\deg(H) \geq \deg(F) + \deg(G)$. Then, there exist polynomials A', B', A'' and B'' in $K[x_1, \dots, x_n]$ such that

$$\begin{cases} A'F + B'G + f_\ell^{k-\ell-m+1}RH = 0, \\ \deg(A') \leq \deg(H) - \deg(F), \quad \deg(B') < \deg(F), \end{cases} \quad (2.6)$$

$$\begin{cases} A''F + B''G + g_m^{k-\ell-m+1}RH = 0, \\ \deg(A'') < \deg(G), \quad \deg(B'') \leq \deg(H) - \deg(G). \end{cases} \quad (2.7)$$

(Proof) It is evident from the following determinants:

$$A' = \left(\begin{array}{ccccccccc} f_\ell & f_{\ell-1} & \cdot & \cdot & \cdot & f_0 & & & x^{k-\ell} \\ & f_\ell & f_{\ell-1} & \cdot & \cdot & \cdot & f_0 & & x^{k-\ell-1} \\ & & \cdot & \cdot & \cdot & \cdot & \cdot & & \cdot \\ & & & f_\ell & f_{\ell-1} & \cdot & \cdot & \cdot & f_0 & x^0 \\ g_m & g_{m-1} & \cdot & \cdot & \cdot & g_0 & & & 0 \\ & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & & \cdot \\ & & & g_m & g_{m-1} & \cdot & \cdot & \cdot & g_0 & 0 \\ h_k & h_{k-1} & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & h_0 & 0 \end{array} \right) \begin{array}{l} \left. \begin{array}{l} \\ \\ \\ \end{array} \right\} k-\ell+1 \text{ rows} \\ \left. \begin{array}{l} \\ \\ \end{array} \right\} \ell \text{ rows} \end{array}, \quad (2.8)$$

← k+2 columns →

$$B' = [\text{replace the last column of } A' \text{ by } (0 \dots 0 x^{\ell-1} \dots x^0 0)^T], \quad (2.9)$$

$$B'' = \pm [\text{exchange } \ell \text{ and } m, f_i \text{ and } g_i, i=0, \dots, \max\{\ell, m\}, \text{ in } A'], \quad (2.10)$$

$$A'' = \pm [\text{replace the last column of } B'' \text{ by } (0 \dots 0 x^{m-1} \dots x^0 0)^T]. // \quad (2.11)$$

The next lemma is essential in deriving a Cramer-type formula.

[Lemma 3] Let $D_{\xi\eta}$ be the following determinant of order $r+2$ with elements P_{ij} , $i=1, \dots, r+2$, $j=1, \dots, r, \xi, \eta$:

$$D_{\xi\eta} = -D_{\eta\xi} = \begin{vmatrix} P_{11} & \dots & P_{1r} & P_{1\xi} & P_{1\eta} \\ \vdots & & \vdots & \vdots & \vdots \\ P_{r+2,1} & \dots & P_{r+2,r} & P_{r+2,\xi} & P_{r+2,\eta} \end{vmatrix}, \quad (2.12)$$

where ξ and η are any two elements of $\{\alpha, \beta, \gamma, \delta\}$. Then, we have

$$D_{\alpha\beta}D_{\gamma\delta} + D_{\alpha\gamma}D_{\delta\beta} + D_{\alpha\delta}D_{\beta\gamma} = 0. \quad (2.13)$$

(Proof) For $r = 0$ and 1 , we can easily prove (2.13) by direct expansion of determinants.

Assume (2.13) is true for $r=0, 1, \dots, t-1$, and consider the case of $r=t$. Defining

$$(a, b) = (b, a) \equiv P_{1a}P_{1b} + P_{2a}P_{2b} + \dots + P_{t+2,a}P_{t+2,b},$$

we can represent $D_{\alpha\beta}D_{\gamma\delta}$, $D_{\alpha\gamma}D_{\delta\beta}$, and $D_{\alpha\delta}D_{\beta\gamma}$ as

$$D_{\alpha\beta}D_{\tau\delta} = \begin{vmatrix} (1,1) & \dots & (1,t) & (1,\tau) & (1,\delta) \\ : & & : & : & : \\ (t,1) & \dots & (t,t) & (t,\tau) & (t,\delta) \\ (\alpha,1) & \dots & (\alpha,t) & (\alpha,\tau) & (\alpha,\delta) \\ (\beta,1) & \dots & (\beta,t) & (\beta,\tau) & (\beta,\delta) \end{vmatrix}.$$

$$D_{\alpha\tau}D_{\delta\beta} = \begin{vmatrix} (1,1) & \dots & (1,t) & (1,\delta) & (1,\beta) \\ : & & : & : & : \\ (t,1) & \dots & (t,t) & (t,\delta) & (t,\beta) \\ (\alpha,1) & \dots & (\alpha,t) & (\alpha,\delta) & (\alpha,\beta) \\ (\tau,1) & \dots & (\tau,t) & (\tau,\delta) & (\tau,\beta) \end{vmatrix}.$$

$$D_{\alpha\delta}D_{\beta\tau} = \begin{vmatrix} (1,1) & \dots & (1,t) & (1,\beta) & (1,\tau) \\ : & & : & : & : \\ (t,1) & \dots & (t,t) & (t,\beta) & (t,\tau) \\ (\alpha,1) & \dots & (\alpha,t) & (\alpha,\beta) & (\alpha,\tau) \\ (\delta,1) & \dots & (\delta,t) & (\delta,\beta) & (\delta,\tau) \end{vmatrix}.$$

Let us expand these determinants with respect to the last two rows and columns, and consider the coefficient factor in the term proportional to $(\alpha,a)(\beta,b)(\tau,c)(\delta,d)$, i.e., the factors other than $(\alpha,a)(\beta,b)(\tau,c)(\delta,d)$ in the term.

Case 1: Terms proportional to $(\alpha,\beta)(\tau,\delta)$, $(\alpha,\tau)(\delta,\beta)$, and $(\alpha,\delta)(\beta,\tau)$. We easily see that except for the sign, the coefficient factors of these terms are the same, the top-left minor of order t . Hence, the sum of these terms is found to be zero.

Case 2: Terms proportional to $(\alpha,i)(\beta,j)(\tau,\delta)$, $1 \leq i \leq t$, $1 \leq j \leq t$, hence $t \geq 1$, or term proportional to $(\alpha,i)(\beta,i)(\tau,c)(\delta,d)$, $1 \leq i,c,d \leq t$, hence $t \geq 2$. There are only two terms which are proportional to $(\alpha,i)(\beta,j)(\tau,\delta)$, and they come from $D_{\alpha\tau}D_{\delta\beta}$ and $D_{\alpha\delta}D_{\beta\tau}$. We easily see that, except for the sign, the coefficient factors of these terms are the same. Hence terms proportional to $(\alpha,i)(\beta,j)(\tau,\delta)$ cancel each other. The same is true for terms proportional to $(\alpha,i)(\beta,i)(\tau,c)(\delta,d)$. (Note that the top-left $t \times t$ submatrix is symmetric.) Similarly, the terms proportional to $(\alpha,i)(\tau,j)(\delta,\beta)$, etc. and terms proportional to $(\alpha,i)(\tau,i)(\delta,d)(\beta,b)$, etc. disappear.

Case 3: Only the remaining terms are those proportional to $(\alpha, a)(\beta, b)(\gamma, c)(\delta, d)$, $1 \leq a, b, c, d \leq t$, and a, b, c, d are different from each other, hence $t \geq 4$. There are three terms containing $(\alpha, a)(\beta, b)(\gamma, c)(\delta, d)$, and the coefficient factor of each term is the following determinant of order $t-2$:

$$\begin{vmatrix} (1, 1) & \dots & (1, t) \\ : & & : \\ (t, 1) & \dots & (t, t) \end{vmatrix} \leftarrow \text{no } (c', \cdot), (d', \cdot) \text{ rows}$$

$\uparrow$ no $(\cdot, a), (\cdot, b')$ columns

where $\{b', c', d'\} = \{b, c, d\}$. If we call the last two columns of the determinant in (2.12) additional columns of types ξ and η , the above determinant is nothing but the product of two determinants of the form (2.12), where the order of the determinants is now $t-2$ and the additional columns are of types a and b' for one determinant and of types c' and d' for the other. Hence, the problem reduces to the case of $r=t-4$. //

§3. Solutions of single equation

We first investigate the following single equation:

$$P_1 y_1 + \dots + P_s y_s = P_{s+1}, \quad (3.1)$$

$$P_i \in K[x_1, \dots, x_n], \quad i=1, \dots, s+1.$$

Without loss of generality, we may assume that

$$\deg(D) = 0 \quad \text{where } D \equiv \text{GCD}(P_1, \dots, P_s). \quad (3.2)$$

(If $\deg(D) \neq 0$, D must pseudo-divide P_{s+1} , i.e., P_{s+1}/D is a polynomial in $S[x]$ so far as (3.1) has solutions. Hence, we have only to divide (3.1) by D , satisfying (3.2).) Furthermore, without loss of generality, we may assume

$$\text{GCD}(P_1, P_2) = 1. \quad (3.3)$$

The reason is as follows: If condition (3.3) is not satisfied, we may construct

$$P'_2 = P_2 + \lambda_3 P_3 + \dots + \lambda_s P_s, \quad \lambda_i \in K[x_1, \dots, x_{n-1}],$$

such that $\text{GCD}(P_1, P'_2) = 1$ and consider the equation

$$P_1 y_1 + P'_2 y_2 + P_3 y'_3 + \dots + P_s y'_s = P_{s+1}, \quad y'_k \equiv y_k - \lambda_k y_2.$$

Suppose, for simplicity, that

$$\deg(P_i) < \deg(P_1) + \deg(P_2), \quad i=3, \dots, s+1. \quad (3.4)$$

Then, Lemma 1 tells that there exist polynomials A_i and B_i in $K[x_1, \dots, x_n]$ such that

$$\begin{cases} A_i P_1 + B_i P_2 + R P_i = 0, & i=3, \dots, s+1, \\ R \equiv \text{res}(P_1, P_2). \end{cases} \quad (3.5)$$

Multiplying R to (3.1) and using (3.5), we obtain

$$\begin{aligned} & P_1(Ry_1 - A_3 y_3 - \dots - A_s y_s + A_{s+1}) \\ & + P_2(Ry_2 - B_3 y_3 - \dots - B_s y_s + B_{s+1}) = 0. \end{aligned}$$

Since $\text{GCD}(P_1, P_2) = 1$, this equality is equivalent to

$$\begin{cases} -Ry_1 + A_3 y_3 + \dots + A_s y_s - A_{s+1} = -P_2 u_2, \\ -Ry_2 + B_3 y_3 + \dots + B_s y_s - B_{s+1} = P_1 u_2, \end{cases} \quad (3.6)$$

$$(3.7)$$

where $u_2 \in S[x]$. Since R is a unit in S , we can solve (3.6) and (3.7) for arbitrary $y_3, \dots, y_s$, and u_2 in $S[x]$. Therefore, representing the general solution of (3.1) as

$$\bar{y} = u_2 \bar{y}^{(2)} + \dots + u_s \bar{y}^{(s)} + \bar{y}^{(s+1)}, \quad (3.8)$$

$$\bar{y} \equiv (y_1, y_2, \dots, y_s),$$

$$u_j, j=2, \dots, s, \text{ are arbitrary elements in } S[x],$$

we obtain the generators $\bar{y}^{(2)}, \dots, \bar{y}^{(s)}$ and a particular solution $\bar{y}^{(s+1)}$ as

$$\bar{y}^{(2)} = (P_2/R, -P_1/R, 0 \dots 0), \quad (3.9)$$

$$\bar{y}^{(i)} = (A_i/R, B_i/R, 0 \dots 0, 1, 0 \dots 0), \quad i=3, \dots, s, \quad (3.10)$$

$$\bar{y}^{(s+1)} = (-A_{s+1}/R, -B_{s+1}/R, 0 \dots 0). \quad (3.11)$$

Since A_i, B_i and R are represented by determinants whose elements are coefficients of P_1, P_2 , and P_i , the above generators are of Cramer type.

The only remaining task is to remove the restriction (3.4), which is quite easy if we use Lemma 2. That is, if

$$d \equiv \max\{\deg(P_i) \mid i=3, \dots, s+1\} - \deg(P_1 P_2) + 1 \geq 1, \quad (3.12)$$

we generate A_i and $B_i, i=3, \dots, s+1$, in $K[x_1, \dots, x_n]$ such that

$$\begin{cases} A_i P_1 + B_i P_2 + \text{lc}(P_1)^d R P_i = 0, \\ \deg(A_i) \leq \max\{\deg(P_2)-1, \deg(P_i)-\deg(P_1)\}, \\ \deg(B_i) < \deg(P_1). \end{cases} \quad (3.13)$$

Then, (3.1) is reduced to the coupled equations

$$\begin{cases} -\text{lc}(P_1)^d R y_1 + A_3 y_3 + \dots + A_s y_s - A_{s+1} = -P_2 u_2, \end{cases} \quad (3.6')$$

$$\begin{cases} -\text{lc}(P_1)^d R y_2 + B_3 y_3 + \dots + B_s y_s - B_{s+1} = P_1 u_2, \end{cases} \quad (3.7')$$

which are directly solved to give the general solution.

§4. Solutions of coupled equations

Suppose equations in (1) are linearly independent over $K(x_1, \dots, x_n)$. Then, without loss of generality, we may assume

$$\Delta \equiv \begin{vmatrix} P_{11} & \dots & P_{1r} \\ \vdots & & \vdots \\ P_{r1} & \dots & P_{rr} \end{vmatrix} \neq 0. \quad (4.1)$$

With this assumption, we can rewrite (1) as

$$\begin{cases} \Delta y_1 = -\Delta_{1,r+1}y_{r+1} - \dots - \Delta_{1s}y_s + \Delta_{1,s+1}, \\ \cdot \quad \quad \quad \cdot \quad \cdot \quad \cdot \\ \Delta y_r = -\Delta_{r,r+1}y_{r+1} - \dots - \Delta_{rs}y_s + \Delta_{r,s+1}, \end{cases} \quad (4.2)$$

where Δ_{ik} , $1 \leq i \leq r$, $r+1 \leq k \leq s+1$, is the following determinant:

$$\Delta_{ik} = \begin{vmatrix} P_{11} & \dots & P_{1,i-1} & P_{1k} & P_{1,i+1} & \dots & P_{1r} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ P_{r1} & \dots & P_{r,i-1} & P_{rk} & P_{r,i+1} & \dots & P_{rr} \end{vmatrix}. \quad (4.3)$$

The following theorem is essential for proving the main theorem given later.

[Theorem 1-1] Let $\Delta \neq 0$, and for some i and j , $1 \leq i \neq j \leq r$, let $\Delta_{i,r+1} \neq 0$, $\Delta_{j,r+1} \neq 0$, and $\text{GCD}(\Delta, \Delta_{i,r+1}) = \text{GCD}(\Delta, \Delta_{j,r+1}) = 1$. For $t=i$ and $t=j$, define R_t as

$$\begin{cases} R_t = \text{l.c.}(\Delta)^{d_t} \text{res}(\Delta, \Delta_{t,r+1}), \\ d_t = \max\{0, \deg(\Delta_{tk}) - \deg(\Delta_{t,r+1}) + 1 \mid k=r+2, \dots, s+1\}. \end{cases} \quad (4.4)$$

For $k=r+2, \dots, s+1$, construct A_{tk} and B_{tk} satisfying

$$\begin{cases} A_{tk}\Delta + B_{tk}\Delta_{t,r+1} + R_t\Delta_{tk} = 0, \\ \deg(B_{tk}) < \deg(\Delta). \end{cases} \quad (4.5)$$

Then, we have

$$B_{ik}/R_i = B_{jk}/R_j. \quad (4.6)$$

[Proof] We note that the R_t , A_{tk} and B_{tk} satisfying (4.4) and (4.5) can be calculated by using Lemma 1 or 2. Moving the i -th and j -th columns in Δ etc. to the rightmost, and representing Δ , $\Delta_{i,r+1}$, Δ_{ik} , $\Delta_{j,r+1}$, Δ_{jk} as determinants of the form given in (2.12) with additional columns of types i , j , $r+1$, k , we obtain the following relation by Lemma 3:

$$\Delta \cdot \Delta_{ijk} + \Delta_{j,r+1} \Delta_{ik} + \Delta_{jk} (-\Delta_{i,r+1}) = 0, \quad (4.7)$$

where Δ_{ijk} is the following determinant:

$$\Delta_{ijk} = \begin{vmatrix} P_{11} & \cdots & P_{1,r+1} & \cdots & P_{1k} & \cdots & P_{1r} \\ \vdots & & \vdots & & \vdots & & \vdots \\ P_{r1} & \cdots & P_{r,r+1} & \cdots & P_{rk} & \cdots & P_{rr} \end{vmatrix}.$$

(i) (j) ← column number

Let $\text{GCD}(\Delta, \Delta_{ik}, \Delta_{jk}) = D$, then (4.5) implies $D \mid B_{ik}$ and $D \mid B_{jk}$. Hence, defining

$$\Delta = D\Delta', \quad \Delta_{ik} = D\Delta'_{ik}, \quad \Delta_{jk} = D\Delta'_{jk}, \quad B_{ik} = DB'_{ik}, \quad B_{jk} = DB'_{jk},$$

we can rewrite (4.5) and (4.7) as

$$A_{ik}\Delta' + B'_{ik}\Delta_{i,r+1} + R_i\Delta'_{ik} = 0, \quad \deg(B'_{ik}) < \deg(\Delta'), \quad (4.5')$$

$$A_{jk}\Delta' + B'_{jk}\Delta_{j,r+1} + R_j\Delta'_{jk} = 0, \quad \deg(B'_{jk}) < \deg(\Delta'), \quad (4.5'')$$

$$\Delta' \cdot \Delta_{ijk} + \Delta'_{ik}\Delta_{j,r+1} - \Delta'_{jk}\Delta_{i,r+1} = 0. \quad (4.7')$$

Eliminating Δ'_{ik} and Δ'_{jk} from (4.5') and (4.5''), we obtain

$$R_j\Delta'_{jk}(A_{ik}\Delta' + B'_{ik}\Delta_{i,r+1}) = R_i\Delta'_{ik}(A_{jk}\Delta' + B'_{jk}\Delta_{j,r+1}).$$

Eliminating $\Delta'_{jk}\Delta_{i,r+1}$ from (4.7') and the above equation, we obtain

$$\begin{aligned} & \Delta'_{ik}\Delta_{j,r+1}(R_iB'_{jk} - R_jB'_{ik}) \\ & = \Delta'(R_j\Delta'_{jk}A_{ik} - R_i\Delta'_{ik}A_{jk} + R_iB'_{ik}\Delta_{ijk}). \end{aligned} \quad (4.8)$$

We consider only the case of $\deg(\Delta) > \deg(D)$, because if $\deg(\Delta) = \deg(D)$ then $B_{ik} = B_{jk} = 0$. Then, since $\text{GCD}(\Delta', \Delta_{j,r+1}) = 1$, (4.8) gives

$$\Delta' \mid \Delta'_{ik}(R_iB'_{jk} - R_jB'_{ik}).$$

Similarly, we obtain

$$\Delta' \mid \Delta'_{jk}(R_iB'_{jk} - R_jB'_{ik}).$$

Since $\text{GCD}(\Delta', \Delta'_{ik}, \Delta'_{jk}) = 1$, these relations imply

$$\Delta' \mid (R_iB'_{jk} - R_jB'_{ik}) \quad \text{or} \quad \Delta \mid (R_iB_{jk} - R_jB_{ik}).$$

Since $\deg(\Delta) > \deg(B_{ik}), \deg(B_{jk})$, the above relation leads to (4.6). //

Theorem 1-1 and (4.8) give the following corollary.

[Corollary] The A_{jk} , $j \neq i$, is calculated from A_{ik} , B_{ik} , R_i and R_j as

$$A_{jk} = (R_j\Delta_{jk}A_{ik} + R_iB_{ik}\Delta_{ijk})/R_i\Delta_{ik}. \quad (4.9)$$

This relation is quite useful in actual calculations because calculation of R_i , A_{ik} and B_{ik} is quite time consuming.

Theorem 1-1 is slightly generalized as follows:

[Theorem 1-2] Let $\Delta \neq 0$, and for $t=i$ and $t=j$, $1 \leq i \neq j \leq r$, let

$$\begin{cases} \text{GCD}(\Delta, \Delta_{t,r+1}, \dots, \Delta_{ts}) = D_t, \\ \text{GCD}(\Delta/D_t, \Delta_{t,r+1}/D_t) = 1, \\ \deg(\Delta/D_t) > 0, \quad \deg(\Delta_{t,r+1}/D_t) > 0. \end{cases} \quad (4.10)$$

Furthermore, let

$$\begin{cases} R_t = \text{lc}(\Delta/D_t)^{d_t} \text{res}(\Delta/D_t, \Delta_{t,r+1}/D_t), \\ d_t = \max\{0, \deg(\Delta_{tk}/D_t) - \deg(\Delta \cdot \Delta_{t,r+1}/D_t^2) + 1 \mid k=r+2, \dots, s+1\}. \end{cases} \quad (4.11)$$

For $k=r+2, \dots, s+1$, construct A_{tk} and B_{tk} satisfying

$$\begin{cases} A_{tk}(\Delta/D_t) + B_{tk}(\Delta_{t,r+1}/D_t) + R_t(\Delta_{tk}/D_t) = 0, \\ \deg(B_{tk}) < \deg(\Delta/D_t). \end{cases} \quad (4.12)$$

Then, we have

$$\tilde{\Delta} \mid (R_i B_{jk} - R_j B_{ik}) \quad \text{where} \quad \tilde{\Delta} \equiv \Delta / \text{GCD}(\Delta, D_i D_j). \quad (4.13)$$

(Proof) Putting $\Delta_t = \Delta/D_t$, $\tilde{\Delta}_{tk} = \Delta_{tk}/D_t$, $t=i, j$, we can rewrite (4.12) and (4.7) as

$$\begin{cases} (A_{ik} \Delta_i / \tilde{\Delta}) \tilde{\Delta} + B_{ik} \tilde{\Delta}_{i,r+1} + R_i \tilde{\Delta}_{ik} = 0, \\ (A_{jk} \Delta_j / \tilde{\Delta}) \tilde{\Delta} + B_{jk} \tilde{\Delta}_{j,r+1} + R_j \tilde{\Delta}_{jk} = 0, \\ \tilde{\Delta} \cdot \tilde{\Delta}_{ijk} + \tilde{\Delta}_{ik} \tilde{\Delta}_{j,r+1} - \tilde{\Delta}_{jk} \tilde{\Delta}_{i,r+1} = 0, \end{cases} \quad (4.14)$$

where $\tilde{\Delta}_{ijk} = \Delta_{ijk} \text{GCD}(\Delta, D_i D_j) / D_i D_j$. Note that $\tilde{\Delta}_{ijk} \in K[x_1, \dots, x_n]$ because (4.7) implies $D_i D_j \mid \Delta \cdot \Delta_{ijk}$. Since $\tilde{\Delta} \mid \Delta_i$ and $\tilde{\Delta} \mid \Delta_j$, equations in (4.14) are of the same forms as (4.5'), (4.5'') and (4.7'), and we obtain (4.13) by performing the same calculation as that in the proof of Theorem 1-1. //

[Corollary] If $\deg(\tilde{\Delta}) = \deg(\Delta_i) = \deg(\Delta_j)$ in Theorem 1-2 then we have

$$B_{ik}/R_i = B_{jk}/R_j. //$$

Now, we prove the main theorem.

[Theorem 2-1] Let $\Delta \neq 0$, and for $i=1, \dots, r$ let $\Delta_{i,r+1} \neq 0$ and $\text{GCD}(\Delta, \Delta_{i,r+1}) = 1$.

Representing the general solution of (1) as

$$\bar{y} = u_{r+1} \bar{y}^{(r+1)} + \dots + u_s \bar{y}^{(s)} + \bar{y}^{(s+1)}, \quad (4.15)$$

u_j , $j=r+1, \dots, s$, are arbitrary elements in $S[x]$,

the generators $\bar{y}^{(r+1)}$, ..., $\bar{y}^{(s)}$ and a particular solution $\bar{y}^{(s+1)}$ are given as

Therefore, $y_i \in S[x]$, $i=1, \dots, r-1$, and combining the above solution and $\bar{y}'$, we obtain the generators (4.16). //

Notes: We can calculate the solutions of (1) by successively solving each equation of (1) and substituting the solution into the yet unsolved equations. This method introduces extremely large factors which exactly cancel each other between the numerator and denominator. However, proof of the cancellation is quite tedious.

Theorem 2-1 is slightly generalized as follows.

[Theorem 2-2] Let $\Delta \neq 0$ and

$$\begin{cases} \text{GCD}(\Delta, \Delta_{i,r+1}, \dots, \Delta_{is}) = D_i, & i=1, \dots, r, \\ \deg(D_r) = 0, \end{cases} \quad (4.17)$$

$$\text{GCD}(\Delta/D_i, \Delta_{i,r+1}/D_i) = 1, \quad i=1, \dots, r, \quad (4.18)$$

$$D_i \mid \Delta_{i,s+1} \text{ over } S, \quad i=1, \dots, r. \quad (4.19)$$

Construct A_{ik} and B_{ik} , $k=r+2, \dots, s+1$, according to the formulas (4.12) with (4.11). Then, formula (4.16) with the following replacement gives the general solution of (1):

$$A_{ik}/R_i \rightarrow A_{ik}/R_i + \Delta_{i,r+1}(R_r B_{ik} - R_i B_{rk})/\Delta R_i R_r. \quad (4.20)$$

(Proof) Solving the last equation of (4.2) after dividing it by D_r , and substituting the solution into the i -th equation of (4.2) and dividing it by D_i , we have

$$\begin{aligned} \Delta_i y_i = & + (\tilde{\Delta}_{i,r+1} \Delta_r) u_{r+1}/R_r \\ & + \{ \Delta_i A_{i,r+2}/R_i + \tilde{\Delta}_{i,r+1} (R_r B_{i,r+2} - R_i B_{r,r+2})/R_i R_r \} u_{r+2} \\ & + \quad \cdot \quad \cdot \quad \cdot \\ & + \{ \Delta_i A_{is}/R_i + \tilde{\Delta}_{i,r+1} (R_r B_{is} - R_i B_{rs})/R_i R_r \} u_s \\ & - \{ \Delta_i A_{i,s+1}/R_i + \tilde{\Delta}_{i,r+1} (R_r B_{i,s+1} - R_i B_{r,s+1})/R_i R_r \}, \end{aligned}$$

where $\Delta_i \equiv \Delta/D_i$ and $\tilde{\Delta}_{ik} \equiv \Delta_{ik}/D_i$ as before. We have $\Delta_i \mid \Delta_r$ over S because $\deg(D_r) = 0$, and Theorem 1-2 with $j = r$ implies $\Delta_i \mid (R_r B_{ik} - R_i B_{rk})$, $k=r+2, \dots, s+1$. Hence, the y_i in the above equation is in $S[x]$ for arbitrary $u_{r+1}, \dots, u_s$ in $S[x]$. (Note that, because $D_r \in S$, we can delete the common factor D_r in the generators.) //

The conditions (4.18) may not be satisfied in many actual cases. In such a case, we have only to construct

$$\Delta'_{i,r+1} = \Delta_{i,r+1} + \lambda_{r+2} \Delta_{i,r+2} + \dots + \lambda_s \Delta_{is}, \quad \lambda_j \in K[x_1, \dots, x_{n-1}], \quad (4.21)$$

such that $\text{GCD}(\Delta/D_i, \Delta'_{i,r+1}/D_i) = 1$, $i=1, \dots, r$, and consider the equations

$$\begin{cases} \Delta y_1 = -\Delta'_{1,r+1}y_{r+1} - \Delta_{1,r+2}y'_{r+2} - \dots - \Delta_{1s}y'_s + \Delta_{1,s+1}, \\ \cdot \quad \quad \quad \cdot \quad \cdot \quad \cdot \\ \Delta y_r = -\Delta'_{r,r+1}y_{r+1} - \Delta_{r,r+2}y'_{r+2} - \dots - \Delta_{rs}y'_s + \Delta_{r,s+1}, \end{cases}$$

where $y'_k \equiv y_k - \lambda_k y_{r+1}$, $k=r+2, \dots, s$.

§5. General case (rare case)

If the conditions in Theorem 2-1 are satisfied, the general solution of (1) is given by the beautiful formula (4.16). If the conditions are not satisfied but the conditions in Theorem 2-2 are satisfied, we have less beautiful but still simple formula (4.16) with the replacements (4.20). The latter condition of (4.17) will be satisfied in most practical cases, and the condition (4.18) will also be valid in such cases so far as the transformation (4.21) is applied. We must, however, consider the rare cases in which

$$\deg(D_i) \neq 0 \quad \text{for all } i=1, \dots, r, \quad (5.1)$$

where $D_i \equiv \text{GCD}(\Delta, \Delta_{i,r+1}, \dots, \Delta_{is})$. In this case, if any of the conditions

$$D_i \mid \Delta_{i,s+1} \text{ over } S, \quad i=1, \dots, r, \quad (5.2)$$

is not satisfied, the coupled equations (1) have no solution. Hence, we assume (5.1) and (5.2) throughout this section.

Following Theorem 1-2, we introduce the following quantities:

$$\begin{cases} C_i \equiv \text{GCD}(\Delta, D_i D_r), \quad i=1, \dots, r-1, \\ C_r \equiv \text{GCD}(\Delta, D_r) = D_r. \end{cases} \quad (5.3)$$

Note that $C_r \mid C_i$. Defining $\tilde{\Delta}_i$ and $\tilde{y}_i$ as

$$\begin{cases} \tilde{\Delta}_i \equiv \Delta / C_i, \quad i=1, \dots, r, \\ \tilde{y}_i \equiv (C_i / D_i) y_i, \quad i=1, \dots, r, \end{cases} \quad (5.4)$$

we can rewrite (4.2) as

$$\begin{cases} \tilde{\Delta}_1 \tilde{y}_1 = -\tilde{\Delta}_{1,r+1}y_{r+1} - \dots - \tilde{\Delta}_{1s}y_s + \tilde{\Delta}_{1,s+1}, \\ \cdot \quad \quad \quad \cdot \quad \cdot \quad \cdot \\ \tilde{\Delta}_r \tilde{y}_r = -\tilde{\Delta}_{r,r+1}y_{r+1} - \dots - \tilde{\Delta}_{rs}y_s + \tilde{\Delta}_{r,s+1}, \end{cases} \quad (5.5)$$

where $\tilde{\Delta}_{ik} \equiv \Delta_{ik} / D_i$ as before. Note that $\tilde{y}_r = y_r$ because $C_r = D_r$ and $\tilde{\Delta}_i \mid \tilde{\Delta}_r$, $i=1, \dots, r-1$, because $C_r \mid C_i$. Furthermore, since $D_i \mid C_i$, $i=1, \dots, r$, we have $\tilde{y}_i \in S[x]$ if $y_i \in S[x]$.

Generalizing the Theorem 2-2, we have the following theorem which gives the general

solution of (5.5).

[Theorem 2-3] Let $\Delta \neq 0$ and $\text{GCD}(\Delta, \Delta_{i,r+1}) = D_i \neq 1$, $i=1, \dots, r$. Construct A_{ik} and B_{ik} , $k=r+2, \dots, s+1$, according to the formulas (4.12) with (4.11). Then, the general solution $\bar{y} = (\tilde{y}_1, \dots, \tilde{y}_r, y_{r+1}, \dots, y_s)$ of (5.5) is given by (4.16) with the following replacements:

$$\begin{cases} \Delta_{i,r+1}/R_r \rightarrow \tilde{\Delta}_{i,r+1} C_i / D_i R_r, & -\Delta/R_r \rightarrow -\tilde{\Delta}_r/R_r, \\ A_{ik}/R_i \rightarrow A_{ik} C_i / D_i R_i + \tilde{\Delta}_{i,r+1} (R_r B_{ik} - R_i B_{rk}) / \tilde{\Delta}_i R_i R_r. \end{cases} \quad (5.6)$$

(Proof) Theorem 1-2 is true even if $\deg(D_r) \neq 0$. Hence, the same proof as for Theorem 2-2 applies to this case. //

Finally, we must solve coupled equations $(C_i/D_i)y_i = \tilde{y}_i$, $i=1, \dots, r-1$, with $\tilde{y}_i$ given as $C_{i,r+1}u_{r+1} + \dots + C_{is}u_s + C_{i,s+1}$, $C_{ik} \in S[x]$, $k=r+1, \dots, s+1$. That is, we must solve the coupled equations of the following form:

$$\begin{cases} P_1 y_1 + P_{1,r+1} y_{r+1} + \dots + P_{1s} y_s = P_{1,s+1}, \\ \cdot \quad \quad \quad \cdot \quad \quad \quad \cdot \\ P_r y_r + P_{r,r+1} y_{r+1} + \dots + P_{rs} y_s = P_{r,s+1}, \end{cases} \quad (5.7)$$

where $P_i, P_{ik} \in K[x_1, \dots, x_n]$, and $\deg(P_i) \neq 0$, $i=1, \dots, r$. (If $\deg(P_i) = 0$ then the i -th equation of (5.7) is satisfied by arbitrary $y_{r+1}, \dots, y_s$ in $S[x]$, hence we can delete it.) Note that if (1) is already of the form (5.7) then (5.5) is identical to (5.7).

The coupled equations (5.7) can be solved in the following way. We assume, as before, $\text{GCD}(P_i, P_{i,r+1}) = 1$, $i=1, \dots, r$. We construct R_i, A_{ik} and B_{ik} , $k=r+2, \dots, s+1$, such that

$$\begin{cases} R_i = \text{lc}(P_i)^{d_i} \text{res}(P_i, P_{i,r+1}), \\ d_i = \max\{0, \deg(P_{ik}) - \deg(P_i P_{i,r+1}) + 1 \mid k=r+2, \dots, s+1\}, \end{cases} \quad (5.8)$$

$$\begin{cases} A_{ik} P_i + B_{ik} P_{i,r+1} + R_i P_{ik} = 0, \\ \deg(B_{ik}) < \deg(P_i). \end{cases} \quad (5.9)$$

Using A_{ik} and B_{ik} , we can transform the i -th equation of (5.7) to the following equations

$$\begin{cases} y_i = (P_{i,r+1}u_i + A_{i,r+2}y_{r+2} + \dots + A_{is}y_s - A_{i,s+1})/R_i, \\ y_{r+1} = (-P_i u_i + B_{i,r+2}y_{r+2} + \dots + B_{is}y_s - B_{i,s+1})/R_i, \end{cases} \quad (5.10)$$

where u_i is an unknown in $S[x]$. These equations are such that the first equation is automatically satisfied by any u_i in $S[x]$. Hence, equating the right hand side of the second equation of (5.10) to that of equation for $i=r$, (5.7) is transformed to the following $r-1$ coupled equations:

$$\left\{ \begin{array}{l} R_r P_1 u_1 - R_1 P_r u_r + (R_1 B_{r,r+2} - R_r B_{1,r+2}) y_{r+2} + \dots \\ \quad + (R_1 B_{rs} - R_r B_{1s}) y_s = (R_1 B_{r,s+1} - R_r B_{1,s+1}), \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ R_{r-1} P_{r-1} u_{r-1} - R_{r-1} P_r u_r + (R_{r-1} B_{r,r+2} - R_r B_{r-1,r+2}) y_{r+2} + \dots \\ \quad + (R_{r-1} B_{rs} - R_r B_{r-1,s}) y_s = (R_{r-1} B_{r,s+1} - R_r B_{r-1,s+1}). \end{array} \right. \quad (5.11)$$

Coupled equations in (5.11) are the same form as those in (5.7) but simplified in that the number of equations is decreased by one. Note that the degrees of the coefficient polynomials in (5.11) are never greater than those in (5.7). Therefore, the problem is reduced to a simpler one. Continuing the above reduction, we can solve (5.7) with suppressing unnecessary coefficient growth.

§6. Comparison with another formula

Using the idea of Hermann[4] (see, also Seidenberg[5]), we can easily represent the generators of the solutions of (1) in a determinant form. Equations in (4.2) show that (1) has the following solutions

$$\left\{ \begin{array}{l} \bar{y}_{ap}^{(r+1)} = (\Delta_{1,r+1}, \dots, \Delta_{r,r+1}, -\Delta, 0 \dots 0), \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ \bar{y}_{ap}^{(s)} = (\Delta_{1s}, \dots, \Delta_{rs}, 0 \dots 0, -\Delta). \end{array} \right. \quad (6.1)$$

We call these solutions apparent solutions. Let

$$d = \max\{\deg(P_{ij}) \mid i=1,\dots,r; j=1,\dots,s\}. \quad (6.2)$$

Then $\deg(\Delta)$ and $\deg(\Delta_{ik})$ are less than or equal to rd . Following Hermann, it is easy to prove that every solution of (1) with $\deg(y_i) \geq rd$ for some i can be represented by apparent solutions. The remaining solutions can be represented as

$$\begin{aligned} y_i &= c_{i,rd-1} x^{rd-1} + c_{i,rd-2} x^{rd-2} + \dots + c_{i0}, \quad i=1,\dots,s, \\ c_{ij} &\in S, \quad j=0,1,\dots,rd-1. \end{aligned} \quad (6.3)$$

Substituting (6.3) into (1) and equating coefficients of x^k terms, $k=0,\dots,rd+d-1$, to zero we obtain $r(rd+d)$ coupled equations for srd unknowns c_{ij} . Since $c_{ij} \in S$, the solutions of these equations are given by Cramer's formula.

The method described above is given in refs. 4 and 5, and it is a typical method in algebra. The degree in x of a generator obtained by this method is less than or equal to

$rd-1$, which is the same as our formula. However, in this method, the order of numerator and denominator determinants is $rd(r+1)$ which is considerably greater than $2rd-1$, the order of determinants for A_{ik} and B_{ik} . Furthermore, the number of generators in the above method is as many as $rd(s-r-1)+(s-r)$, which is very inconvenient in actual applications. On the other hand, the number of generators in our method is only $s-r$. Hence, our formula is much more beautiful and useful than the formula obtained by the above method.

Finally, we present an example for the case of $r=2$ and $s=5$.

$$\begin{cases} (x^2+1)y_1 + (x^2+x+1)y_2 + (2x^2-x)y_3 + (2x^2+3)y_4 + (2x^2-3x+1)y_5 = 0, \\ (x^2+2x+2)y_1 + (3x^2-x+1)y_2 + (x^2+3x+5)y_3 + (2x^2-x+3)y_4 + (3x^2-1)y_5 = 0. \end{cases} \quad (6.4)$$

The Δ and Δ_{ik} , $i=1,2$, $k=3,4,5$, are calculated as

$$\begin{aligned} \Delta &= 2x^4 - 4x^3 - x^2 - 5x - 1, \\ \Delta_{13} &= 5x^4 - 9x^3 - 6x^2 - 9x - 5, \\ \Delta_{14} &= 4x^4 - 3x^3 + 7x^2 - 5x, \\ \Delta_{15} &= 3x^4 - 14x^3 + 6x^2 - 3x + 2, \\ \Delta_{23} &= -x^4 + 4x^2 + 5x + 5, \\ \Delta_{24} &= -5x^3 - 2x^2 - 7x - 3, \\ \Delta_{25} &= x^4 - x^3 + 3x^2 + 4x - 3. \end{aligned}$$

These polynomials give the resultants R_1 and R_2 as

$$\begin{aligned} R_1 &= \text{res}(\Delta, \Delta_{13}) = -396, \\ R_2 &= \text{res}(\Delta, \Delta_{23}) = -165. \end{aligned}$$

Hence, the conditions in Theorem 2-1 are satisfied. Using determinant representations (2.4) and (2.5), we can calculate A_{ik} and B_{ik} , $i=1,2$, $k=4,5$, as

$$\begin{aligned} A_{14} &= 12(-1135x^3 - 932x^2 - 1078x - 410), \\ B_{14} &= 12(+454x^3 + 282x^2 + 511x + 82), \\ A_{15} &= 12(+470x^3 + 604x^2 + 341x + 199), \\ B_{15} &= 12(-188x^3 - 204x^2 - 152x - 53), \\ A_{24} &= 5(+227x^3 + 595x^2 + 651x + 509), \\ B_{24} &= 5(+454x^3 + 282x^2 + 511x + 82), \\ A_{25} &= 5(-94x^3 - 290x^2 - 327x - 166), \\ B_{25} &= 5(-188x^3 - 204x^2 - 152x - 53). \end{aligned}$$

Formula (4.16) gives the generators of the general solution as

$$\begin{aligned}\bar{y}^{(3)} &= (\Delta_{13}, \Delta_{23}, -\Delta, 0, 0), \\ \bar{y}^{(4)} &= (A_{14}/R_1, A_{24}/R_2, B_{24}/R_2, 1, 0), \\ \bar{y}^{(5)} &= (A_{15}/R_1, A_{25}/R_2, B_{25}/R_2, 0, 1).\end{aligned}$$

That these generators satisfy (6.4) is easily checked. Furthermore, we can easily check the validity of relations (4.6) and (4.9).

Acknowledgement

The author would like to thank Mr. Akio Furukawa for valuable discussions.

References

- [1] Sasaki, T. and Furukawa, A., Secondary polynomial remainder sequence and an extension of subresultant theory, J. Inf. Proces., 7 (1984), 175-184.
- [2] Furukawa, A. and Sasaki, T., Multi polynomial remainder sequence and its application to linear Diophantine equations, Proceedings of '83 EUROCAL (Lecture Notes in Comp. Sci 162, Springer-Verlag, 1983), 24-35.
- [3] Sasaki, T. and Furukawa, A., Theory of multiple polynomial remainder sequence, Publ RIMS Kyoto Univ., 20 (1984), 367-399.
- [4] Hermann, G., Die Frage der endlich vielen Schritte in der Theorie der Polynomideale Math. Ann., 95 (1926), 736-788.
- [5] Seidenberg, A., Constructions in algebra, Trans. Amer. Math. Soc., 197 (1974), 273-313.