JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS **92**, 397–409 (1983)

The Lagrange–Good Inversion Formula and Its Application to Integral Equations

N. G. DE BRUIJN

Department of Mathematics, Eindhoven University of Technology, 5600MB Eindhoven, The Netherlands

It is shown how Good's extension of the Lagrange inversion formula for n variables can be derived straightforwardly from the case of a single variable. Using this to attack the case of infinitely many variables (integral equations) in Fredholm's style, we get Fredholm's formulas, but surprisingly with permanents instead of determinants.

1. Introduction

It is the main purpose of this paper to show how the Lagrange–Good inversion formula for n variables can be derived straightforwardly from the case of a single variable. As a side product we show how the inversion formula can be applied to integral equations. This is similar to Fredholm's method for solving an integral equation by taking the limit of Cramer's solution for finite systems of ordinary linear equations. Now treating Fredholm's integral equation by means of the Lagrange–Good formula instead, we obtain Fredholm's formulas, but with the surprising fact that we get permanents instead of determinants.

We take the formulas in the form presented by Good [2], who gave detailed proofs based on the Jacobi transformation formula for n-fold contour integrals. We also refer to Good's paper for the history of the multiple Lagrange expressions, but we mention that some more historical details can be found in [8]. Instead of Good's reference to Goursat [3, 4], Osgood [8, p. 47] mentions Darboux's earlier formulation for the case of two variables (*C. R. Acad. Sci. Paris* **68** (1869), 324). Darboux's work was based on a 1777 paper by Laplace, printed in his Oeuvres, Vol. 9, p. 330. Stieltjes (*Ann. Sci. École Norm. Sup.* (3) **2** (1885), 93) extended Darboux's results to the case of n variables.

We also refer to Osgood [8] for his treatment of the Jacobi transformation. For an entirely different short proof of the Lagrange–Good formula and for further references we refer to Hofbauer [7].

In this paper, we shall not bother about convergence of series. All infinite series will be considered as formal power series, and the only thing that

0022-247X/83 $3.00

matters for us is to obtain coefficients by means of formal procedures. Nevertheless, we shall be quite informal in the subject of formal series. For a serious treatment of that subject we refer to [11], a very substantial paper that culminates in a proof of the multiple Lagrange inversion formula.

For the sake of completeness we have included (Section 2) a proof for the (single variable) Lagrange formula. This proof is due to Tschebyschef [10]. It seems attractive to try to prove the case of n variables along the same lines, but the author did not succeed in doing this.

We shall deal with the multivariable case (Good's formula) in the equivalent form (2.11). A funny aspect of that formula is that it also generalizes MacMahon's "Master Theorem."

2. Notation and Statement of the Inversion Formula

Although formal power series in terms of powers of z are not functions of z, we shall use a notational system in which they are treated as if they were functions. In particular, we shall use things like $\lambda_w f(z, w)$ in order to denote the power series we get if we consider $f(z, w)$ as a series of powers of w, where the coefficients are power series in terms of z. And we use things like $[\ ]_{w:=f(z)}$ for denoting the result of the substitution $w := f(z)$ applied to the expression inside the square brackets.

We consider n formal power series

$$f_1(w_1,..., w_n),..., f_n(w_1,..., w_n)$$

in the variables $w_1,..., w_n$, where n is a positive integer. We shall study the following set of equations:

$$w_1 = z_1 f_1(w_1,..., w_n),..., w_n = z_n f_n(w_1,..., w_n) \tag{2.1}$$

and it is required to solve (2.1) by indicating power series $\phi_1(z_1,..., z_n),...,$ $\phi_n(z_1,..., z_n)$ that satisfy (2.1) when substituted for $w_1,..., w_n$.

We shall use bold face symbols as abbreviations for vectors. So $\mathbf{z} = (z_1,..., z_n)$, and similar notation is used in other cases. With this convention system (2.1) becomes

$$\mathbf{w} = \mathbf{z} * \mathbf{f}(\mathbf{w}). \tag{2.2}$$

(The asterisk denotes component multiplication of vectors: $(a_1,..., a_n) * (b_1,..., b_n) = (a_1 b_1,..., a_n b_n)$.) The solution of (2.2) can be written as

$$\mathbf{w} = \boldsymbol{\phi}(\mathbf{z}). \tag{2.3}$$

We note that $\boldsymbol{\phi}$ is uniquely determined by $\mathbf{f}$. If we evaluate $\mathbf{w}^{(1)} = \mathbf{z} * \mathbf{f}(\mathbf{0})$,

$\mathbf{w}^{(2)} = \mathbf{z} * \mathbf{f}(\mathbf{w}^{(1)})$, etc., then $\mathbf{w}^{(1)}$ gives us the terms of $\boldsymbol{\phi}(\mathbf{z})$ with degree $\leqslant 1$, $\mathbf{w}^{(2)}$ gives the terms of degree $\leqslant 2$, etc.

In the special case that $\mathbf{f}$ satisfies

$$f_1(\mathbf{0}) \neq 0, ..., f_n(\mathbf{0}) \neq 0, \tag{2.4}$$

then $1/f_1(\mathbf{w}), ..., 1/f_n(\mathbf{w})$ are power series again, and we can write (2.2) in the form

$$\mathbf{z} = \boldsymbol{\psi}(\mathbf{w}). \tag{2.5}$$

However, assumption (2.4) plays no rôle in this paper.

If $h(\mathbf{w})$ is a single formal power series in n variables, we want to express $h(\boldsymbol{\phi}(\mathbf{z}))$ as a power series in terms of $\mathbf{z}$ by means of a formula that makes reference to $\mathbf{f}$ but not to $\boldsymbol{\phi}$. Such a formula was given by Good [2] in the form

$$h(\boldsymbol{\phi}(\mathbf{z})) = \sum_{\mathbf{m}} \frac{\mathbf{z}^{\mathbf{m}}}{\mathbf{m}!} \left[\left(\frac{\partial}{\partial \mathbf{w}} \right)^{\mathbf{m}} h(\mathbf{w})(\mathbf{f}(\mathbf{w}))^{\mathbf{m}} D(\mathbf{w}, \mathbf{z}; \mathbf{f}) \right]_{\mathbf{w}:=\mathbf{0}}, \tag{2.6}$$

where

$$D(\mathbf{w}, \mathbf{z}; \mathbf{f}) = \det \left(\delta_{\mu\nu} - z_\mu \frac{\partial f_\mu(\mathbf{w})}{\partial w_\nu} \right). \tag{2.7}$$

The notation needs some explanation. First $\sum_{\mathbf{m}}$ stands for the n-fold summation over $m_1, ..., m_n$, each running through $\{0, 1, 2, ...\}$. Next $\mathbf{z}^{\mathbf{m}} = z_1^{m_1} \cdots z_n^{m_n}$, and $\mathbf{m}! = m_1! \cdots m_n!$. Furthermore,

$$\left(\frac{\partial}{\partial \mathbf{w}} \right)^{\mathbf{m}} = \left(\frac{\partial}{\partial w_1} \right)^{m_1} \cdots \left(\frac{\partial}{\partial w_n} \right)^{m_n},$$

$$(\mathbf{f}(\mathbf{w}))^{\mathbf{m}} = (f_1(\mathbf{w}))^{m_1} \cdots (f_n(\mathbf{w}))^{m_n}.$$

Finally $\det(a_{\mu\nu})$ is the determinant of the $n \times n$ matrix with entries $a_{\mu\nu}$ $(\mu, \nu = 1, ..., n)$, and $\delta_{\mu\nu}$ is Kronecker's delta ($=1$ if $\mu = \nu$, otherwise 0).

The fact that (2.6) holds for all h can be used to transform it. First, we can replace $h(\mathbf{w})$ by a series $k(\mathbf{w}, \mathbf{y})$ in two sets of n variables, and we get

$$k(\boldsymbol{\phi}(\mathbf{z}), \mathbf{y}) = \sum_{\mathbf{m}} \frac{\mathbf{z}^{\mathbf{m}}}{\mathbf{m}!} \left[\left(\frac{\partial}{\partial \mathbf{w}} \right)^{\mathbf{m}} k(\mathbf{w}, \mathbf{y})(\mathbf{f}(\mathbf{w}))^{\mathbf{m}} D(\mathbf{w}, \mathbf{z}; \mathbf{f}) \right]_{\mathbf{w}:=\mathbf{0}}. \tag{2.8}$$

We now assume (2.4) and that identically (with the $\boldsymbol{\psi}$ of (2.5))

$$k(\mathbf{w}, \boldsymbol{\psi}(\mathbf{w})) = h(\mathbf{w}). \tag{2.9}$$

Then we have $k(\boldsymbol{\phi}(\mathbf{z}), \mathbf{z}) = h(\boldsymbol{\phi}(\mathbf{z}))$, whence

$$h(\boldsymbol{\phi}(\mathbf{z})) = \sum_{\mathbf{m}} \frac{\mathbf{z}^{\mathbf{m}}}{\mathbf{m}!} \left[\left(\frac{\partial}{\partial \mathbf{w}} \right)^{\mathbf{m}} k(\mathbf{w}, \mathbf{z})(\mathbf{f}(\mathbf{w}))^{\mathbf{m}} D(\mathbf{w}, \mathbf{z}; \mathbf{f}) \right]_{\mathbf{w}:=0}.$$

A possibility to satisfy (2.8) is

$$k(\mathbf{w}, \mathbf{y}) = h(\mathbf{w})\, D(\mathbf{w}, \boldsymbol{\psi}(\mathbf{w}); \mathbf{f})/D(\mathbf{w}, \mathbf{y}; \mathbf{f}),$$

and therefore we obtain

$$h(\boldsymbol{\phi}(\mathbf{z})) = \sum_{\mathbf{m}} \frac{\mathbf{z}^{\mathbf{m}}}{\mathbf{m}!} \left[\left(\frac{\partial}{\partial \mathbf{w}} \right)^{\mathbf{m}} h(\mathbf{w})(\mathbf{f}(\mathbf{w}))^{\mathbf{m}} D(\mathbf{w}, \boldsymbol{\psi}(\mathbf{w}); \mathbf{f}) \right]_{\mathbf{w}:=0}. \qquad (2.10)$$

The main reason for going into so much detail here is that Good [2] says that (2.6) and (2.10) are not obviously equivalent.

If in (2.8) we replace $k(\mathbf{w}, \mathbf{y})$ by $h(\mathbf{w})/D(\mathbf{w}, \mathbf{y}; \mathbf{f})$, and later replace $\mathbf{y}$ by $\mathbf{z}$, we get a formula which is equivalent to (2.6), that is,

$$\left[\frac{h(\mathbf{w})}{D(\mathbf{w}, \mathbf{z}; \mathbf{f})} \right]_{\mathbf{w}:=\boldsymbol{\phi}(\mathbf{z})} = \sum_{\mathbf{m}} \frac{\mathbf{z}^{\mathbf{m}}}{\mathbf{m}!} \left[\left(\frac{\partial}{\partial \mathbf{w}} \right)^{\mathbf{m}} h(\mathbf{w})(\mathbf{f}(\mathbf{w}))^{\mathbf{m}} \right]_{\mathbf{w}:=0}. \qquad (2.11)$$

It is in this form that we shall prove the Lagrange–Good inversion formula in Section 4.

It should be remarked that (2.11) turns exactly into MacMahon's "Master Theorem" (for a short proof, based on the Jacobi transformation, we refer to [3]) if we specialize by taking

$$h(w) = 1, \qquad f_\lambda(w_1, ..., w_n) = 1 + \sum_{\lambda=1}^{n} a_{\lambda\mu} w_\mu .$$

In other words, (2.11) provides a wealth of generalizations of the Master Theorem.

3. The Case of a Single Variable

In the case of a single variable there is another well-known form of the inversion formula, namely,

$$g(\phi(z)) = \sum_{m=1}^{\infty} \frac{z^m}{m!} \left[\left(\frac{d}{dw} \right)^{m-1} g'(w)(f(w))^m \right]_{w:=0}, \qquad (3.1)$$

where $w = \phi(z)$ is the solution of the equation $w = zf(w)$. From (3.1) we can get to the case $n = 1$ of (2.11) in a few lines (cf. [7, Part 3, Problem 207]).

This is done as follows: Taking derivatives in (3.1), using the derivative of the identity $\phi(z) = zf(\phi(z))$, and putting $g'(w)f(w) = h(w)$, we get

$$\frac{h(\phi(z))}{1 - zf'(\phi(z))} = \sum_{m=0}^{\infty} \frac{z^m}{m!} \left[\left(\frac{d}{dw}\right)^m h(w)(f(w))^m \right]_{w:=0}. \tag{3.2}$$

This is the case $n = 1$ of (2.11).

Formula (3.1) is usually proved by contour integration, of course restricted to power series f with a positive radius of convergence. An interesting alternative proof was given by Tschebyschef [10], in the form of a finite development with remainder term in the form of an integral. His idea was based on a generalization of the method of integration by parts. Integration by parts depends on $(f(x)\,g(x))' = f(x)\,g'(x) + f'(x)\,g(x)$. Tschebyschef's generalization has $\Phi(u, v)$ instead of uv, and

$$\frac{d}{dx}\,\Phi(f(x), g(x)) = \frac{\partial \Phi}{\partial u} f'(x) + \frac{\partial \Phi}{\partial v} g'(x).$$

He applied it to

$$\Phi_m(u, v) = \left(\frac{\partial}{\partial v}\right)^{m-1} \frac{(zf(v) - u)^m}{m!} h(v), \tag{3.3}$$

which satisfies

$$\frac{\partial \Phi_{m+1}}{\partial u} + \frac{\partial \Phi_m}{\partial v} = 0 \qquad (m = 1, 2, \ldots). \tag{3.4}$$

By integration from 0 to q, where q is such that $zf(q) = q$, he got

$$\int_0^q \left[\frac{\partial \Phi_m}{\partial u}\right]_{u:=x, v:=x} dx = -\Phi_m(0, 0) + \int_0^q \left[\frac{\partial \Phi_{m+1}}{\partial u}\right]_{u:=x, v:=x} dx.$$

By summation from 1 to N it follows that

$$\int_0^q h(x)\,dx = \sum_{m=1}^{N} \frac{z^m}{m!} \left[\left(\frac{d}{dv}\right)^{m-1} h(v)(f(v))^m\right]_{v:=0}$$
$$+ \int_0^q \left[\left(\frac{\partial}{\partial v}\right)^N h(v) \frac{(zf(v) - u)^N}{N!}\right]_{u:=x, v:=x} dx.$$

This proof was also presented in [1].

4. The Case of Several Variables

We shall prove (2.11) for n variables by induction with respect to n. Since the case $n = 1$ has been settled already, it suffices to show that if (2.11) holds for n variables and also for N variables, then it holds for $n+N$ variables.

Our notation will be $\mathbf{z}$, $\mathbf{w}$, $\mathbf{f}$, $\boldsymbol{\phi}$, and $\mathbf{w}$ for vectors of length n, and $\mathbf{Z}$, $\mathbf{W}$, $\mathbf{F}$, $\boldsymbol{\Phi}$, and $\mathbf{W}$ for vectors of length N. For the case of $n+N$ we start from the equations

$$\mathbf{w} = \mathbf{z} * \mathbf{f}(\mathbf{w}, \mathbf{W}), \tag{4.1}$$

$$\mathbf{W} = \mathbf{Z} * \mathbf{F}(\mathbf{w}, \mathbf{W}). \tag{4.2}$$

The solution can be obtained as follows: First we solve (4.1), and we get a solution of the type

$$\mathbf{w} = \boldsymbol{\phi}(\mathbf{z}, \mathbf{W}). \tag{4.3}$$

Next we consider the equation

$$\mathbf{W} = \mathbf{Z} * \mathbf{F}(\boldsymbol{\phi}(\mathbf{z}, \mathbf{W}), \mathbf{W}) \tag{4.4}$$

and write its solution as

$$\mathbf{W} = \boldsymbol{\Phi}(\mathbf{z}, \mathbf{Z}). \tag{4.5}$$

Since (4.3) solves (4.1) we have the identity

$$\boldsymbol{\phi}(\mathbf{z}, \mathbf{W}) = \mathbf{z} * \mathbf{f}(\boldsymbol{\phi}(\mathbf{z}, \mathbf{W}), \mathbf{W}). \tag{4.6}$$

In order to prove (2.10) for $n+N$ variables, we take an arbitrary power series $h(\mathbf{w}, \mathbf{W})$ and we have to show

$$\left[\left[\frac{h(\mathbf{w}, \mathbf{W})}{D(\mathbf{w}, \mathbf{W}, \mathbf{z}, \mathbf{Z}; \mathbf{f}, \mathbf{F})}\right]_{\mathbf{w}:=\boldsymbol{\phi}(\mathbf{z},\mathbf{W})}\right]_{\mathbf{W}:=\boldsymbol{\Phi}(\mathbf{z},\mathbf{Z})} \tag{4.7}$$
$$= \sum_{\mathbf{m}} \sum_{\mathbf{M}} \mathbf{z}^{\mathbf{m}} \mathbf{Z}^{\mathbf{M}} \left[\left(\frac{\partial}{\partial \mathbf{w}}\right)^{\mathbf{m}} \left(\frac{\partial}{\partial \mathbf{W}}\right)^{\mathbf{M}} h(\mathbf{w}, \mathbf{W})(\mathbf{f}(\mathbf{w}, \mathbf{W}))^{\mathbf{m}} (\mathbf{F}(\mathbf{w}, \mathbf{W}))^{\mathbf{M}}\right]_{\mathbf{w}:=\mathbf{0},\, \mathbf{W}:=\mathbf{0}}.$$

Here $D(\mathbf{w}, \mathbf{W}, \mathbf{z}, \mathbf{Z}; \mathbf{f}, \mathbf{F})$ is the $(n+N)\times(n+N)$ determinant built for the system (4.1) + (4.2) in the same way as (2.7) was built for (2.2). Let us use indices μ, ν running from 1 to n, and σ, τ running from 1 to N. Then $D(\mathbf{w}, \mathbf{W}, \mathbf{z}, \mathbf{Z}; \mathbf{f}, \mathbf{F})$ is the determinant of the block matrix $\left(\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right)$ with

$$(A)_{\mu\nu} = \delta_{\mu\nu} - z_\mu(\partial f_\mu/\partial w_\nu), \qquad (B)_{\mu\tau} = -z_\mu(\partial f_\mu/\partial W_\tau),$$
$$(C)_{\sigma\nu} = -Z_\sigma(\partial F_\sigma/\partial w_\nu), \qquad (D)_{\sigma\tau} = -\delta_{\sigma\tau} - Z_\sigma(\partial F_\sigma/\partial W_\tau).$$

In order to prove (4.7) we put

$$p(\mathbf{z}, \mathbf{W}) = \sum_{\mathbf{m}} \frac{\mathbf{z}^{\mathbf{m}}}{\mathbf{m}!} \left[\left(\frac{\partial}{\partial \mathbf{w}}\right)^{\mathbf{m}} h(\mathbf{w}, \mathbf{W})(\mathbf{F}(\mathbf{w}, \mathbf{W}))^{\mathbf{M}} \cdot (\mathbf{f}(\mathbf{w}, \mathbf{W}))^{\mathbf{m}} \right]_{\mathbf{w}:=\mathbf{0}},$$

and then right-hand side of (4.7) equals

$$\sum_{\mathbf{M}} \frac{\mathbf{Z}^{\mathbf{M}}}{\mathbf{M}!} \left(\frac{\partial}{\partial \mathbf{W}}\right)^{\mathbf{M}} p(\mathbf{z}, \mathbf{W}). \tag{4.8}$$

By (2.10), applied to Eq. (4.1) and its solution (4.3), we have

$$p(\mathbf{z}, \mathbf{W}) = \left[\frac{h(\mathbf{w}, \mathbf{W})(\mathbf{F}(\mathbf{w}, \mathbf{W}))^{\mathbf{M}}}{D(\mathbf{w}, \mathbf{z}; \lambda_{\mathbf{w}} \mathbf{f}(\mathbf{w}, \mathbf{W}))}\right]_{\mathbf{w}:=\boldsymbol{\phi}(\mathbf{z}, \mathbf{W})}.$$

Now (4.9) becomes

$$\sum_{\mathbf{M}} \frac{\mathbf{Z}^{\mathbf{M}}}{\mathbf{M}!} \left[\left(\frac{\partial}{\partial \mathbf{W}}\right)^{\mathbf{M}} \left[\frac{h(\mathbf{w}, \mathbf{W})}{D(\mathbf{w}, \mathbf{z}; \lambda_{\mathbf{w}} \mathbf{f}(\mathbf{w}, \mathbf{W}))}\right]_{\mathbf{w}:=\boldsymbol{\phi}(\mathbf{z}, \mathbf{W})} (\mathbf{F}(\boldsymbol{\phi}(\mathbf{z}, \mathbf{W}), \mathbf{W})^{\mathbf{M}} \right]_{\mathbf{W}:=\mathbf{0}}.$$

By (2.10), applied to Eq. (4.4) and its solution (4.5), this equals

$$\left[\left[\frac{h(\mathbf{w}, \mathbf{W})}{D(\mathbf{w}, \mathbf{z}; \lambda_{\mathbf{w}} \mathbf{f}(\mathbf{w}, \mathbf{W}))}\right]_{\mathbf{w}:=\boldsymbol{\phi}(\mathbf{z}, \mathbf{W})} \frac{1}{D(\mathbf{W}, \mathbf{Z}; \lambda_{\mathbf{W}} \mathbf{F}(\boldsymbol{\phi}(\mathbf{z}, \mathbf{W}), \mathbf{W}))} \right]_{\mathbf{W}:=\boldsymbol{\Phi}(\mathbf{z}, \mathbf{Z})}.$$

In order to get to (4.7), it now suffices to show that

$$\begin{aligned} &[D(\mathbf{w}, \mathbf{W}, \mathbf{z}, \mathbf{Z}; \mathbf{f}, \mathbf{F})]_{\mathbf{w}:=\boldsymbol{\phi}(\mathbf{z}, \mathbf{W})} \\ &\quad = [D(\mathbf{w}, \mathbf{z}; \lambda_{\mathbf{w}} \mathbf{f}(\mathbf{w}, \mathbf{W}))]_{\mathbf{w}:=\boldsymbol{\phi}(\mathbf{z}, \mathbf{W})} \times D(\mathbf{W}, \mathbf{Z}; \lambda_{\mathbf{W}} \mathbf{F}(\boldsymbol{\phi}(\mathbf{z}, \mathbf{W}), \mathbf{W})). \end{aligned} \tag{4.9}$$

According to a formula of Schur, we have

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} A & 0 \\ C & I \end{pmatrix} \begin{pmatrix} I & A^{-1}B \\ 0 & D - CA^{-1}B \end{pmatrix},$$

whence

$$\det \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \det(A) \det(D - CA^{-1}B). \tag{4.10}$$

Since (4.3) solves (4.1), we have the identity $\boldsymbol{\phi}(\mathbf{z}, \mathbf{W}) = \mathbf{z} * \mathbf{f}(\boldsymbol{\phi}(\mathbf{z}, \mathbf{W}), \mathbf{W})$. By differentiation we obtain

$$[AL]_{\mathbf{w}:=\boldsymbol{\phi}(\mathbf{z}, \mathbf{W})} = -[B]_{\mathbf{w}:=\boldsymbol{\phi}(\mathbf{z}, \mathbf{W})},$$

where

$$(L)_{\nu\tau} = \partial \phi_\nu / \partial W_\tau.$$

Next we note that

$$D(\mathbf{W}, \mathbf{Z}; \lambda_{\mathbf{W}}\mathbf{F}(\boldsymbol{\phi}(\mathbf{z}, \mathbf{W}), \mathbf{W})) = \det(R),$$

where

$$(R)_{\sigma\tau} = \left[\delta_{\sigma\tau} - Z_\sigma \sum_{\nu} \frac{\partial F_\sigma}{\partial w_\nu} \frac{\partial \phi_\nu}{\partial W_\tau} + \frac{\partial F_\sigma}{\partial W_\tau}\right]_{\mathbf{w} := \boldsymbol{\phi}(\mathbf{z}, \mathbf{W})}.$$

It follows that

$$R = [D + CL]_{\mathbf{w} := \boldsymbol{\phi}(\mathbf{z}, \mathbf{W})} = [D - CA^{-1}B]_{\mathbf{w} := \boldsymbol{\phi}(\mathbf{z}, \mathbf{W})}.$$

So if in (4.10) we substitute $\mathbf{w} := \boldsymbol{\phi}(\mathbf{z}, \mathbf{W})$ we obtain (4.9), and that completes our proof.

5. The Case of Infinitely Many Variables

Systems of n linear equations with n unknowns can be solved by means of Cramer's rule. That rule presents the solution in the form of a quotient of two determinants. Fredholm showed in 1900 how a linear integral equation can be considered as a limit of a sequence of linear equations, and how in the limit Cramer's solution leads to a solution of the integral equation. The solution gets the form of the quotient of two power series. In the denominator the mth coefficient is an m-fold integral over a determinant with m rows and m columns, and the numerator is built similarly.

Fredholm's equation and his solution are given in Section 6 below ; for detailed presentations the reader may consult [6, 12].

In this section we shall explain how a similar procedure can be applied to the Lagrange–Good inversion formula, and how this leads to solutions of linear and nonlinear integral equations.

In the special case of linear integral equations, one of course expects to get Fredholm's solution again. Very remarkably, however, it turns out that we get a modified form of Fredholm's solution: the modification is mainly that all determinants are replaced by permanents! Actually, it can be shown directly, without reference to the Lagrange–Good formula, that the permanents can be used as nicely as the determinants. At least formally, that is. The solution with determinants is superior to the one with permanents as far as convergence is concerned. The series in the case of permanents need not be always convergent. The success of the determinants depends on the use of Hadamard's inequality, and a similar estimation does not exist for the permanents.

We shall not try to be very formal, very precise, or very general when thinking of the following procedure:

(i) Describe an unknown function F, continuous on the interval $0 < x < 1$, say, by means of the n unknowns $F(1/n)$, $F(2/n)$,..., $F(n/n)$, where n is large.

(ii) Replace an integral equation with unknown function F by a set of ordinary equations for $F(1/n)$,..., $F(n/n)$.

(iii) Take the limit of the solution for $n \to \infty$, replacing sums by integrals in the obvious way.

Let us start from an equation

$$F(x) = z(x) \left\{ a^{(0)}(x) + \sum_{r=1}^{\infty} \int_0^1 a_r^{(1)}(x, y)(F(y))^r \, dy \right.$$
$$\left. + \sum_{r=1}^{\infty} \sum_{s=1}^{\infty} \int_0^1 \int_0^1 a_{rs}^{(2)}(x, y, z)(F(y))^r (F(z))^s \, dy \, dz + \cdots \right\}. \quad (5.1)$$

The functions z, $a^{(0)}$, $a_r^{(1)}$, $a_{rs}^{(2)}$,... are given, as continuous functions, say, and it is required to find a function F such that (5.1) holds for all x in $0 < x < 1$.

Carrying out steps (i) and (ii), we write the discretization of (5.1) in the form of Eq. (2.2). We take $w_i = F(i/n)$, $z_i = z(i/n)$, and we build the equation

$$w_i = z_i f_i(w_1, ..., w_n),$$

where

$$f_i(w_1, ..., w_n) = a^{(0)}(i/n) + \sum_{r=1}^{\infty} \sum_{j=1}^{n} n^{-1} a_r^{(1)}(i/n, j/n)(w_j)^r$$
$$+ \sum_{r=1}^{\infty} \sum_{s=1}^{\infty} \sum_{j=1}^{n} \sum_{k=1}^{n} n^{-2} a_{rs}^{(2)}(i/n, j/n, k/n)(w_j)^r (w_k)^s + \cdots. \quad (5.2)$$

Now taking any power series $h(w_1, ..., w_n)$, we consider solution (2.11). In order to eliminate the trouble of having to think about $D(w, z; f)$, we take two functions h_1 and h_2 instead of the single h, write down (2.11) for each of them, and form the quotient on both sides. This expresses

$$h_1(\phi(z))/h_2(\phi(z)) \quad (5.3)$$

as a quotient of two series, both of the type of the right-hand side of (2.11).

From now on we shall assume that $z_1 = \cdots = z_n = 1$. This is no essential restriction since it can be achieved by adjusting the a's.

We have to investigate what happens to the right-hand side of (2.11) if we pass to the limit for $n \to \infty$.

We have to sum over all vectors $\mathbf{m}$ of the form $m_1, ..., m_n$, where all m_i are nonnegative integers. If $\mathbf{m}$ is such a vector, we introduce the degree of $\mathbf{m}$ as the vector $d_1, d_2, \ldots$, where d_j is the number of i such that $m_i = j$. We shall group the terms of the right-hand side of (2.11) together according to their degree vectors.

For a given degree vector $\mathbf{d}$, the contribution to the right-hand side of (2.11) is obtained as follows: Let $m_1, \ldots, m_s$ be positive numbers such that the vector $m_1, \ldots, m_s, 0, \ldots, 0$ has length n and degree d. Then the contribution equals

$$(d_1!\, d_2! \cdots)^{-1} (\mathbf{m}!)^{-1} \qquad (5.4)$$
$$\times \sum \left[\left(\frac{\partial}{\partial w_{i_1}} \right)^{m_1} \cdots \left(\frac{\partial}{\partial w_{i_s}} \right)^{m_s} h(\mathbf{w}) (f_{i_1}(\mathbf{w}))^{m_1} \cdots (f_{i_s}(\mathbf{w}))^{m_s} \right]_{\mathbf{w} := \mathbf{0}},$$

where the sum runs over all s-tuples $(i_1, \ldots, i_s)$ of pairwise different elements of $\{1, \ldots, n\}$.

With this fixed value of $\mathbf{d}$ we take the limit for $n \to \infty$. Let us discuss the simplest case, i.e., the one where $h(\mathbf{w}) = 1$ for all $\mathbf{w}$. The differentiation on the right leads to a finite number of terms of the form

$$\prod_{p=1}^{P} \left[\left(\frac{\partial}{\partial w_{i_1}} \right)^{m_1(p)} \cdots \left(\frac{\partial}{\partial w_{i_s}} \right)^{m_s(p)} f_{i_p}(\mathbf{w}) \right]_{\mathbf{w} := \mathbf{0}}, \qquad (5.5)$$

where

$$m_1(p) \geqslant 0, \ldots, m_s(p) \geqslant 0, \qquad \sum_{j=1}^{P} m_j(p) = m_j.$$

Each factor of (5.3) is expressed as a sum of coefficients a, multiplied by certain factorials and powers of n. Note that this involves a factor n^{-t}, where $t = \sum_{p=1}^{P} t_p$, and t_p is the number of j with $m_j(p) > 0$. And note that $t \geqslant s$, and that $t = s$ only if for all j we have either $m_j(p) = m_j$ or $m_j(p) = 0$.

If (5.3) is such that $t = s$, the summation over all $(i_1, \ldots, i_s)$ tends to an s-fold integral if $n \to \infty$. Note that the factor n^{-s} takes care of the fact that the integral over the s-dimensional unit cube is replaced by a sum over n^s points. And note that we can forget about the restriction that we only have to take s-tuples with pairwise different elements: the number of exceptions is of the order of n^{s-1}.

On the other hand, if (5.3) is such that $t > s$, its contribution tends to zero if $n \to \infty$.

Needless to say, the final solution of (5.1) becomes a notational monstrosity. We shall not try to present it here, apart from a very simple case to be treated in the next section.

6. Application to Fredholm's Integral Equation

If in (5.1) we specialize by taking all a's equal to zero apart from the $a^{(0)}$ and the $a_r^{(1)}$'s, and if we write these as g and K, respectively, we get Fredholm's equation

$$F(x) = g(x) + \int_0^1 K(x, y)\, F(y)\, dy. \tag{6.1}$$

Now (5.1) becomes

$$f_i(w_1,..., w_n) = g(i/n) + n^{-1} \sum_{j=1}^{n} K(i/n, j/n)\, w_j. \tag{6.2}$$

The procedure sketched at the end of Section 5 can now easily be carried out. We first take $h(\mathbf{w}) = 1$. The $(f(\mathbf{w}))^{\mathbf{m}}$ in (2.11) is a polynomial of degree $m_1 + \cdots + m_s$. Since this degree equals the number of differentiations, we can drop the constants $g(x_i)$ from the $f_i(\mathbf{w})$. And because of the factor n^{-1} in (6.2) we get a factor n^{-t}, where $t = m_1 + \cdots + m_s$. As explained in Section 5, the contribution of (5.4) tends to zero if $t > s$. Since all m_i are positive, this means that we only have to consider the case $m_1 = \cdots = m_s = 1$, whence $\mathbf{d} = (s, 0, 0,...)$.

The s-fold differentiation in (5.4) gives rise to $s!$ terms, in total

$$n^{-s} \sum_{\pi} \prod_{\nu=1}^{s} K(i_\nu/n, i_{\pi(\nu)}/n),$$

where π runs through all $s!$ permutations of the set $\{1,..., s\}$. Taking the sum over all s-tuples $(i_1,..., i_s)$ (as noted in Section 5, the restriction that the i's are pairwise different can be ignored), we get for (5.4)

$$(s!)^{-1} \int_0^1 \cdots \int_0^1 \sum_{\pi} \prod_{\nu=1}^{s} K(y_\nu, y_{\pi(\nu)})\, dy_1 \cdots dy_s.$$

By definition of the permanent of a matrix we have

$$\operatorname{per} \begin{pmatrix} K(y_1, y_1) & \cdots & K(y_1, y_s) \\ \vdots & \cdots & \vdots \\ K(y_s, y_1) & \cdots & K(y_s, y_s) \end{pmatrix} = \sum_{\pi} \prod_{\nu=1}^{s} K(y_\nu, y_{\pi(\nu)}). \tag{6.3}$$

A similar argument can be applied to the slightly more difficult case where $h(\mathbf{w}) = w_q$ with some fixed index q. We skip the derivation and just state the result we get when taking the limit of (5.3) (where $h_1(\mathbf{w}) = w_q$, $h_2(\mathbf{w}) = 1$). It gives the solution of Fredholm equation (6.1) in the form

$$F(x) = \Phi_1(x)/\Phi_2(x), \tag{6.4}$$

where

$$\Phi_1(x) = \sum_{s=0}^{\infty} (s!)^{-1} Z_s(x),$$

$$\Phi_2(x) = \sum_{s=0}^{\infty} (s!)^{-1} W_s,$$

$$Z_s(x) = \int_0^1 \cdots \int_0^1 U(x, y_1, \ldots, y_s)\, dy_1 \cdots dy_s,$$

$$W_s = \int_0^1 \cdots \int_0^1 V(y_1, \ldots, y_s)\, dy_1 \cdots dy_s.$$

Here $U(x, y_1, \ldots, y_s)$ is the permanent of the matrix

$$\begin{pmatrix} g(x) & K(x, y_1) & \cdots & K(x, y_s) \\ g(y_1) & K(y_1, y_1) & \cdots & K(y_1, y_s) \\ \vdots & \vdots & \cdots & \vdots \\ g(y_s) & K(y_s, y_1) & \cdots & K(y_s, y_s) \end{pmatrix} \tag{6.5}$$

and $V(y_1, \ldots, y_s)$ the permanent of the matrix

$$\begin{pmatrix} K(y_1, y_1) & \cdots & K(y_1, y_s) \\ \vdots & \cdots & \vdots \\ K(y_s, y_1) & \cdots & K(y_s, y_s) \end{pmatrix}. \tag{6.6}$$

If $s = 0$, we define $Z_s(x) = g(x)$, $W_s = 1$.

Fredholm's solution has almost the same form. It can be written as

$$\sum_{s=0}^{\infty} (s!)^{-1}(-1)^s Z_s^*(x) \Big/ \sum_{s=0}^{\infty} (s!)^{-1}(-1)^s W_s^*,$$

where $Z_s^*(x)$ is the n-fold integral taken over the determinant of (6.5), and W_s^* is the n-fold integral taken over the determinant of (6.6). For an extensive account of Fredholm's formulas we refer to [6].

We can also verify directly that (6.4) satisfies (6.1), and the derivation can be given for determinants and permanents simultaneously. We just have to check that for $s = 1, \ldots$

$$Z_s(x) = g(x)\, W_s^* + s \int_0^1 K(x, y)\, Z_{s-1}(y)\, dy,$$

$$Z_s^*(x) = g(x)\, W_s^* - s \int_0^1 K(x, y)\, Z_{s-1}^*(y)\, dy.$$

We remark that the denominators in the determinant case and the permanent case are each other's inverses:

$$\sum_{s=0}^{\infty} (s!)^{-1} W_s = \left(\sum_{s=0}^{\infty} (s!)^{-1}(-1)^s W_s^*\right)^{-1}. \tag{6.7}$$

We get this result if we evaluate the multiple integrals termwise, splitting the permanent and the determinant in their $s!$ terms. Analyzing the cycles of the permutations, this leads to

$$\exp\left(\sum_{n=0}^{\infty} \operatorname{tr}(K^{(n)}/n)\right),$$

where $K^{(n)}$ is the n-fold convolution $K * \cdots * K$, and $\operatorname{tr}(K) = \int_0^1 K(y, y)\, dy$.

Note added in proof. The observation that Fredholm's equation can be solved by means of permanents instead of the usual determinants, was recently made by D. Kershaw [13].

References

1. N. G. de Bruijn, De reeks van Taylor en soortgelijke reeksen, *Simon Stevin* **33** (1959), 20–26.
2. I. J. Good, Generalizations to several variables of Lagrange's expansion, with applications to stochastic processes, *Proc. Cambridge Philos. Soc.* **56** (1960), 367–380.
3. I. J. Good, A short proof of MacMahon's "Master Theorem," *Proc. Cambridge Philos. Soc.* **58** (1962), 160.
4. E. Goursat, "Cours d'Analyse Mathématique," Vol. II, 7th ed., Gauthier-Villars, Paris, 1949.
5. E. Goursat, "A Course in Mathematical Analysis, I and II, Part I," translated by E. R. Henrick and O. Dunkel, Dover, New York, 1959.
6. G. Kowalewski, "Einführung in die Determinantentheorie," Chelsea, New York, 1948.
7. J. Hofbauer, A short proof of the Lagrange–Good formula, *Discrete Math.* **25** (1979), 135–139.
8. W. F. Osgood, Analytische Funktionen komplexer Grössen *in* "Enzyklopädie d. Math. Wissensch. II (2)," pp. 5–114, B. G. Teubner, Leipzig, 1901–1921.
9. G. Pólya and G. Szegö, "Aufgaben und Lehrsätze aus der Analysis," Vol. I, Dover, New York, 1945.
10. P. L. Tschebyschef, Sur la série de Lagrange *in* "Oeuvres," Vol. 1, Chelsea, New York, 1961.
11. W. T. Tutte, On elementary calculus and the Good formula, *J. Combin. Theory Ser. B* **18** (1975), 97–137.
12. E. T. Whittaker and G. N. Watson, "A Course of Modern Analysis," 4th ed., Cambridge Univ. Press, London/New York, 1927.
13. D. Kershaw, Permanents in Fredholm's Theory of Integral Equations. *J. Integral Equations* **1** (1979), 281–290.