PROOF OF THE FUNDAMENTAL THEOREM ON IMPLICIT FUNCTIONS BY USE OF COMPOSITE GRADIENT CORRECTIONS

WILLIAM L. HART AND THEODORE S. MOTZKIN

1. Introduction. Many methods have been employed for establishing the classical result, Theorem 2.1, concerning the existence of functions $x_i(t)$ satisfying a system

$$f_j(x\,;\,t) = 0, \qquad (j = 1, 2, \cdots, n) \tag{1.1}$$

of n equations in n unknowns $(x_1, \cdots, x_n) = x$ with $(t_1, \cdots, t_p) = t$, where all variables and functions are real valued, and $f_j(\alpha\,;\,\beta) = 0$. The object of this article is to present a new proof of the theorem by a constructive method of successive approximations involving corrections related to the gradients in x-space of the functions $f_j(x\,;\,\beta)$.

To establish Theorem 2.1, a sequence $x^{(m)}(t)$ with $x^{(0)}(t) = \alpha$ will be defined, where $x^{(m)}(t)$ is obtained by adding to $x^{(m-1)}(t)$ a vector correction $\Delta x^{(m-1)}(t)$ which is equal to a certain constant, ρ, times the vector sum of corrections parallel to the gradients of the $f_j(x\,;\,\beta)$ at $x = \alpha$. The vector $\Delta x^{(m-1)}(t)$, for a fixed t, is a special case of the corresponding correction of an iterative process for solving a general system $g_j(x) = 0$, $(j = 1, \cdots, k)$, $k \geqq n$, introduced by the authors in a previous article [2].

For a particular system (1.1), the method of the present paper would be applicable to obtaining values of the $x_i(t)$ by use of a digital computing machine for any t sufficiently near $t = \beta$. Section 6 in [2] describes a related small arc method with the same objective; the two methods differ in the values of the arguments used in fundamental matrices which appear with similar roles in [2] and below. The method of [2] might be superior computationally to the method of the present paper. However, in § 6 in [2], Theorem 2.1 below was employed as a starting point. Thus the present paper shows that the composite gradient method is effective to establish the supporting Theorem 2.1 as well as the related small arc method of [2] for computing values of the implicit functions.

In connection with the present article, it is pertinent to mention the proof of Theorem 2.1 by E. Goursat, [1], extended by William L. Hart, [3] and [4], to various infinite systems. In the Goursat method for (1.1), a system

Received March 27, 1958. Presented to the American Mathematical Society, April 19, 1958. Sponsored in part by the Office of Naval Research; reproduction permitted for all purposes of the U.S. Government.

(1.2) $$x_j = \Phi_j(x\,;\,t) \qquad (j = 1, 2, \cdots, n)$$

equivalent to (1.1) is constructed by use of the inverse of the matrix[1] $A = (a_{ij})$, $a_{ij} = \partial f_j(\alpha\,;\,\beta)/\partial x_i$; then a solution $x(t)$ of (1.2) is defined by applying the method of successive substitutions to (1.2). In contrast, under the same hypotheses as those of Goursat, § 2 employs a system (1.2) constructed by direct use of A, without forming its inverse. This feature might be of computational advantage. In case $n = 1$, the present method with $\rho = 1$ is identical with the Goursat method.

Either Goursat's method or the present method can also be regarded as a constructive algorithm solving the problem of elimination of $n - 1$ variables $x_1, \cdots, x_{n-1}$ from n equations $f_j(x\,;\,t) = 0$ leading to a relation (such as $x_n = x_n(t)$) between the remaining variables $(x_n, t_1, \cdots, t_p)$.

The problem of solving $y = F(x\,;\,u)$, $F = (F_1, \cdots, F_n)$ and $y = (y_1, \cdots, y_n)$, by $x = \phi(y\,;\,u)$ (inversion with and without parameters), for nonzero Jacobian F_x, is only apparently more general than the solution of (1.1) (to subsume it set $t = (y, u), f = F - y$) and thus is equally amenable to our iterative procedure.

2. Construction of a system (1.2) equivalent to (1.1). We shall consider (1.1) subject to the following hypotheses:

(2.1) $\begin{cases}\text{The } f_j \text{ are continuous, and all derivatives } \partial f_j/\partial x_i \text{ exist and} \\ \text{are continuous in some open neighborhood } \Omega \text{ of } (x = \alpha\,;\,t = \beta).\end{cases}$

(2.2) $$f_j(\alpha\,;\,\beta) = 0 \qquad (j = 1, 2, \cdots, n).$$

(2.3) The matrix $A = (a_{ij})$, $a_{ij} = \partial f_j(\alpha\,;\,\beta)/\partial x_i$, is nonsingular.

In x-space, let the positive gradient of $f_j(x\,;\,\beta)$ at $x = \alpha$ be defined as having the magnitude $(\sum_{i=1}^{n} a_{ij}^2)^{1/2}$, nonzero because of (2.3), and the direction angles Ψ_{ij} specified by

(2.4) $$\cos \Psi_{ij} = a_{ij} w_j^{-1}\,; \quad w_j = (\sum_{i=1}^{n} a_{ij}^2)^{\frac{1}{2}}\,.$$

For any $(x\,;\,t)$ and each j define, formally, a vector correction $\Delta_j x$ for x, where x is considered an approximation to a solution of $f_j(x\,;\,t) = 0$, by specifying the ith component $\Delta_j x_i$ of $\Delta_j x$ as follows:

(2.5) $$\Delta_j x_i = -\rho f_j(x\,;\,t) w_j^{-1} \cos \Psi_{ij}\,,$$

with a constant $\rho > 0$ to be restricted later. Then define the composite vector correction Δx for x, considered now as an approximation to a solution of (1.1), by specifying for Δx the ith component

[1] Capital italic letters represent n by n matrices. The transpose of a matrix A is denoted by A'. We treat x as a one-rowed matrix.

$$\Delta x_i = \sum_{j=1}^{n} \Delta_j x_i \text{ ; or, } \Delta x = \sum_{j=1}^{n} \Delta_j x \text{ .} \tag{2.6}$$

By use of (2.6) we introduce, formally, a sequence $x^{(m)}(t)$ of approximations to a solution of (1.1):

$$x^{(0)}(t) = \alpha \text{ ; } x^{(m)}(t) = x^{(m-1)}(t) + \Delta x^{(m-1)}(t), \ m > 0 \text{ .} \tag{2.7}$$

From (2.5), the ith coordinate $x_i^{(m)}(t)$ is given by

$$x_i^{(m)}(t) = x_i^{(m-1)}(t) - \rho \sum_{j=1}^{n} a_{ij} w_j^{-2} f_j(x^{(m-1)}(t) \text{ ; } t) \text{ .} \tag{2.8}$$

Let the components Φ_i of a vector $\Phi = (\Phi_1, \Phi_2, \cdots, \Phi_n)$ be defined by

$$\Phi_i(x \text{ ; } t) = x_i - \rho \sum_{j=1}^{n} a_{ij} w_j^{-2} f_j(x \text{ ; } t) \text{ .} \tag{2.9}$$

Then (2.7) is the sequence of approximations $x^{(m)}(t)$ arising if the method of successive substitutions, with $x^{(0)}(t) = \alpha$, is used to seek a solution of the system

$$x = \Phi(x, t) \text{ .} \tag{2.10}$$

By use of (2.1) and (2.3) we find that (1.1) and (2.10) are equivalent systems.

We remark that, in x-space, Δx of (2.6) is invariant under an orthogonal transformation of coordinates and under an alteration of $f_j(x \text{ ; } t)$ to $h_j f_j(x \text{ ; } t)$ if $h_j \neq 0$. Thus, before considering the existence and convergence of sequence (2.7), we may assume that (1.1) has been altered by dividing $f_j(x \text{ ; } t)$ by w_j of (2.4). Then, without change of notation, from (2.4) and (2.5) for all j we obtain

$$w_j = 1 \text{ ; } \Delta_j x_i = - \rho a_{ij} f_j(x \text{ ; } t) \text{ ; } \sum_{i=1}^{n} a_{ij}^2 = 1 \text{ .} \tag{2.11}$$

Note that AA' is symmetric and positive definite Hence there exist positive characteristic constants $\lambda_1, \cdots, \lambda_n$ and an orthogonal matrix $S = (s_{pq})$ such that

$$SAA'S' = (\delta_{ij}\lambda_i) = D, \text{ where } \delta_{ii} = 1 \text{ and } \delta_{ij} = 0 \text{ if } i \neq j. \tag{2.12}$$

Now, in (1.1), let the coordinates be changed from $(x_1, \cdots, x_n)$ to $(z_1, \cdots, z_n) = z$ by the orthogonal transformation $x = zS$. Then, with $g_j(z \text{ ; } t) = f_j(x \text{ ; } t)$ when $x = zS$, and $\alpha = \gamma S$ or $\gamma = \alpha S'$, (1.1) becomes

$$g_j(z \text{ ; } t) = 0, \text{ where } g_j(\gamma \text{ ; } \beta) = 0 \text{ , } \quad (j = 1, 2, \cdots, n) \text{ .} \tag{2.13}$$

If we let $b_{ij} = \partial g_j(\gamma \text{ ; } \beta)/\partial z_i$ and $B = (b_{ij})$, we have

$$B = SA \text{ ; } \quad BB' = SAA'S' = D \text{ ; } \quad B'B = A'A \text{ .} \tag{2.14}$$

From $B'B = A'A$ on comparing main diagonal terms we obtain $\sum_{i=1}^{n} b_{ij}^2 = \sum_{i=1}^{n} a_{ij}^2 = 1$, for all j. Hence, if sequence (2.7) is formed for (2.13) by use of the analogue of (2.11) in the z-coordinates, from (2.8) we arrive at

$$(2.15)\quad \begin{cases} z^{(0)}(t) = \gamma \, ; \; z^{(m)}(t) = z^{(m-1)}(t) + \Delta^{(m-1)}(t), \; m > 0 \, ; \\ \Delta z_i^{(m-1)}(t) = - \rho \sum_{j=1}^{n} b_{ij} g_j(z^{(m-1)}(t) \, ; \, t) \, . \end{cases}$$

On account of the invariant features which were mentioned concerning the gradient corrections $\Delta x^{(m-1)}(t)$ of (2.7) for (1.1), it follows that the existence of all $z^{(m)}(t)$ for any t is equivalent to the existence of all $x^{(m)}(t)$ and that $x^{(m)}(t)$ and $z^{(m)}(t)$ represent the same point. We shall find it convenient to discuss $z^{(m)}(t)$ instead of $x^{(m)}(t)$.

We introduce the functions

$$(2.16)\qquad \phi_h(z \, ; \, t) = z_h - \rho \sum_{j=1}^{n} b_{hj} g_j(z \, ; \, t) \qquad (h = 1, \cdots, n) \, ,$$

and consider the following system, obtained as in (2.10), which is equivalent to (2.13):

$$(2.17)\qquad z_h = \phi_h(z \, ; \, t) \qquad (h = 1, \cdots, n) \, .$$

In (2.17) the ϕ_h and all derivatives $\partial \phi_h / \partial z_i$ are continuous when $(z \, ; \, t)$ is in Ω, now defined with coordinates $(z \, ; \, t)$. With $\phi = (\phi_1, \cdots, \phi_n)$, sequence (2.15) can be written

$$(2.18)\qquad z^{(0)}(t) = \gamma \, ; \quad z^{(m)}(t) = \phi(z^{(m-1)}(t) \, ; \, t) \, .$$

From (2.16) and $BB' = D$ we obtain $\dfrac{\partial \phi_h}{\partial z_i} = \delta_{hi} - \rho \sum_{j=1}^{n} b_{hj} \dfrac{\partial g_j}{\partial z_i}$;

$$(2.19)\qquad \frac{\partial \phi_h(\gamma \, ; \, \beta)}{\partial z_h} = 1 - \rho \sum_{j=1}^{n} b_{hj} b_{hj} = 1 - \rho \lambda_h \, ;$$

$$(2.20)\qquad \frac{\partial \phi_h(\gamma \, ; \, \beta)}{\partial z_i} = 0 \, , \qquad \text{if } h \neq i \, .$$

Let $\mu_i = 1 - \rho \lambda_i$ and $\sigma_\rho = \max_{i \leq n} | \mu_i |$. Since $(\lambda_1, \cdots, \lambda_n)$ are the characteristic constants of AA' and $\sum_{i=1}^{n} a_{ij}^2 = 1$ for all j, we have

$$\sum_{i=1}^{n} \lambda_i = n \, ; \qquad 0 < \lambda_i < n \, , \text{ if } n > 1 \, .$$

Then the following lemma can be proved easily as in [2][2].

Lemma 2.1. *In order that $\sigma_\rho < 1$, it is necessary that $0 < \rho < 2$, and it is sufficient that*

[2] See formulas (4.16)–(4.18) in **[2]** with $r = \omega = n$.

$$0 < \rho \leqq 2/n \qquad\qquad (0 < \rho < 2/n \text{ if } n = 1)\,.$$

Moreover, the minimum value of σ_ρ *occurs for a single value* $\rho = \rho_0$ *where*

$$\frac{2}{n} < \rho_0 < \frac{2(n-1)}{n} \quad \textit{if } n > 2, \textit{ and } \rho_0 = 1 \textit{ if } n \leqq 2\,.$$

For each t, and $i = 1, 2, \cdots, n$, let $(z = \xi^{(i)}\,;\, t)$ be a point in Ω and define

$$v_{hi}(t) = \frac{\partial \phi_i(\xi^{(i)}\,;\, t)}{\partial z_h} - \delta_{hi}(1 - \rho\lambda_i)\,. \tag{2.21}$$

Let $V(t) = (v_{hi}(t))$, and introduce the following matrices:

$$B_\rho = I - \rho D = (\delta_{hi}(1 - \rho\lambda_i))\,; \tag{2.22}$$

$$U(t) = B_\rho V'(t) + V(t) B_\rho + V(t) V'(t) = (u_{ij}(t))\,. \tag{2.23}$$

Note that $u_{ij}(t)$ is a polynomial with each term of degree 1 or 2 in the elements $v_{ij}(t)$ of $V(t)$. Let $H(t) = [\sum_{i,j=1}^n u_{ij}^2(t)]^{1/2}$.

LEMMA 2.2. *Select* $\rho > 0$ *so that* $\sigma_\rho < 1$, *and choose* $\theta > 0$ *with* $\sigma_\rho < \theta < 1$. *Then there exist* $\varepsilon > 0$ *and* $\delta \leqq \varepsilon$, $\delta > 0$, *such that, if*[3] $||t - \beta|| \leqq \delta$, $||z - \gamma|| \leqq \varepsilon$, *and all* $||\xi^{(i)} - \gamma|| \leqq \varepsilon$ in (2.21), *then the functions* $g_j(z\,;\, t)$, $\partial g_j(z\,;\, t)/\partial z_i$, *and* $z^{(1)}(t)$ *exist and are continuous, and*

$$||z^{(1)}(t) - \gamma|| \leqq \varepsilon(1 - \theta)\,; \tag{2.24}$$

$$0 \leqq H(t) \leqq \theta^2 - \sigma_\rho^2\,. \tag{2.25}$$

To establish Lemma 2.2 first notice that, if $t = \beta$ and all $\xi^{(i)} = \gamma$ in (2.21), then all $v_{ij}(t) = 0$ and thus all $u_{ij}(t) = 0$. Hence $\varepsilon > 0$ exists so that the specified conditions are satisfied by the g_j, $\partial g_j/\partial z_i$, and $H(t)$ if $||z - \gamma|| \leqq \varepsilon$, $||t - \beta|| \leqq \varepsilon$, and all $||\xi^{(i)} - \gamma|| \leqq \varepsilon$ in (2.21). From (2.18), $z^{(1)}(t) = \phi(\gamma\,;\, t)$ and thus $z^{(1)}(\beta) = \gamma$. Hence, if δ is sufficiently small and $0 < \delta \leqq \varepsilon$, we have (2.24) when $||t - \beta|| \leqq \delta$. This completes the proof of Lemma 2.2.

THEOREM 2.1. *Suppose that* $\rho > 0$ *and is such that* $\sigma_\rho < 1$. *Assume that* (2.1), (2.2), and (2.3) *are satisfied. Then there exist* $\varepsilon > 0$ *and* $\delta > 0$, $\delta \leqq \varepsilon$, *such that, if* $||t - \beta|| \leqq \delta$, *all* $x^{(m)}(t)$ *of* (2.7) *exist, are continuous, and satisfy* $||x^{(m)}(t) - \alpha|| \leqq \varepsilon$. *Also there exists, uniformly for* $||t - \beta|| \leqq \delta$,

---

[3] For any vector z we use $||z||$ for the length. Thus, $||z|| = (\sum_{i=1}^n z_i^2)^{1/2}$.

$$\lim_{m\to\infty} x^{(m)}(t) = x(t) ,$$

where $x = x(t)$ *satisfies* (1.1). *Moreover, if a point* $(x ; t)$ *with* $\|x - \alpha\| \leqq \varepsilon$ *and* $\|t - \beta\| \leqq \delta$ *satisfies* (1.1), *then* $x = x(t)$.

To establish Theorem 2.1, we shall prove the corresponding facts for the sequence $z^{(m)}(t)$ of (2.15) and system (2.17). Let $\rho, \theta, \varepsilon$, and δ be determined by Lemma 2.2 and, hereafter, assume that $\|t - \beta\| \leqq \delta$. Then $z^{(0)}(t)$ and $z^{(1)}(t)$ exist in the region $\|z - \gamma\| \leqq \varepsilon$; by (2.24), since $z^{(0)}(t) = \gamma$, the following inequalities are true when $k = 1$:

$$\text{(2.26)} \qquad \|z^{(k)}(t) - \gamma\| \leqq \varepsilon ; \ \|z^{(k)}(t) - z^{(k-1)}(t)\| \leqq \varepsilon\theta^{k-1}(1 - \theta) .$$

Assume now, for $m > 1$, that $z^{(k)}(t)$ has been proved to exist, to be continuous, and to satisfy (2.26) when $k = 1, 2, \cdots, (m - 1)$. Then $z^{(m)}(t)$ exists and is continuous; also, by the mean value theorem with respect to $(z_1, \cdots, z_n)$ for fixed t,

$$\text{(2.27)} \qquad \begin{aligned} z_i^{(m)}(t) - z_i^{(m-1)}(t) &= \phi_i(z^{(m-1)}(t) ; t) - \phi_i(z^{(m-2)}(t) ; t) \\ &= \sum_{h=1}^{n} \frac{\partial\phi_i(\xi^{(m,i)}(t) ; t)}{\partial z_h} [z_h^{(m-1)}(t) - z_h^{(m-2)}(t)] , \end{aligned}$$

where $\xi^{(m,i)}(t)$ is a properly chosen point in z-space on the line segment joining $z^{(m-2)}(t)$ and $z^{(m-1)}(t)$. With $\xi^{(i)} = \xi^{(m,i)}(t)$ in (2.21), let $V(t)$ be the matrix with elements $v_{hi}(t)$, and let $U(t)$ be defined by (2.23). Note that $\|\xi^{(m,i)}(t) - \gamma\| \leqq \varepsilon$. Then, from (2.27),

$$\text{(2.28)} \qquad \begin{cases} z^{(m)}(t) - z^{(m-1)}(t) = (z^{(m-1)}(t) - z^{(m-2)}(t))(B_\rho + V(t)) ; \\ \|z^{(m)}(t) - z^{(m-1)}(t)\|^2 = (z^{(m)}(t) - z^{(m-1)}(t))(z^{(m)}(t) - z^{(m-1)}(t))' \\ = (z^{(m-1)}(t) - z^{(m-2)}(t))(B_\rho^2 + U(t))(z^{(m-1)}(t) - z^{(m-1)}(t))' . \end{cases}$$

On applying the Cauchy inequality twice[4] to the term involving $U(t)$ in (2.28), we find

$$\text{(2.29)} \qquad \begin{aligned} \|z^{(m)}(t) - z^{(m-1)}(t)\|^2 &\leqq \|z^{(m-1)}(t) - z^{(m-2)}(t)\|^2[\sigma_\rho^2 + H(t)] \\ &\leqq \theta^2\|z^{(m-1)}(t) - z^{(m-2)}(t)\|^2 . \end{aligned}$$

From (2.26) for $k = 1, 2, \cdots, (m-1)$ and (2.29), we obtain (2.26) for $k = m$. Thus, by induction, all $z^{(m)}(t)$ are defined and satisfy (2.26) if $\|t - \beta\| \leqq \delta$. From (2.26), the series $\sum_{m=1}^{\infty} [z_i^{(m)}(t) - z_i^{(m-1)}(t)]$ is termwise dominated by the series $\sum_{m=1}^{\infty} \varepsilon(1 - \theta)\theta^{m-1}$, and hence converges uniformly. Thus the sequence $z^{(m)}(t)$ approaches a limit, $z(t)$, uniformly for $\|t - \beta\| \leqq \delta$. Since all $z^{(m)}(t)$ are continuous, $z(t)$ is continuous. It follows from $z^{(m)}(t) = \phi(z^{(m-1)}(t) ; t)$ that $z = z(t)$ satisfies $z = \phi(z ; t)$.

[4] As follows: $[\sum_{i=1}^{n} a_i \sum_{j=1}^{n} u_{ij}a_j]^2 \leqq (\sum_{i=1} a_i^2) \sum_{i=1}^{n} (\sum_{j=1}^{n} u_{ij}a_j)^2 \leqq (\sum_{i=1}^{n} a_i^2)^2 \sum_{i,j=1}^{n} u_{ij}^2$.

To prove that $z(t)$ is the unique solution of (2.13), suppose that $(\hat{z}\,;\,t)$ satisfies (2.13) for $||\hat{z}-\gamma|| \leqq \varepsilon$ and $||t-\beta|| \leqq \delta$ and assume that $\hat{z} \neq z(t)$. Then, from $\hat{z} = \phi(\hat{z}\,;\,t)$ and $z(t) = \phi(z(t)\,;\,t)$, by details duplicating the proof of (2.29), we have

$$||\hat{z} - z(t)|| \leqq \theta||\hat{z} - z(t)|| < ||\hat{z} - z(t)|| \,,$$

which contradicts the assumption that $\hat{z} \neq z(t)$. Hence the proof of Theorem 2.1 is complete, because the point $z^{(m)}(t)$ in n-space is the same point as $x^{(m)}(t)$, and the region $||x-\alpha|| \leqq \varepsilon$ is the same as the region $||z-\gamma|| \leqq \varepsilon$.

Note 2.1. With a different arrangement of details, we could arrive at Theorem 2.1 with rectangular neighborhoods $\{|t_i - \beta_i| \leqq \delta \text{ for all } i\}$ and $\{|x_i - \alpha_i| \leqq \varepsilon \text{ for all } i\}$ replacing the spherical neighborhoods $||t-\beta|| \leqq \delta$ and $||x-\alpha|| \leqq \varepsilon$.

Note 2.2. In use of the sequence $\{x^{(m)}(t)\}$ in any particular case to obtain approximate values of $x(t)$, flexibility is introduced through the presence of the somewhat arbitrary constant ρ. Greater flexibility could be introduced (as in § 5 of [2]) by permitting suitably restricted variation in ρ, with $\rho = \rho^{(m)}$ at the mth iteration; revised details would establish Theorem 2.1. with this change.

Note 2.3. Suppose that $(\alpha\,;\,\beta)$ is not a solution of (1.1.). With only (2.1) and (2.3) as hypotheses, there exists $\varepsilon > 0$ so that the region $(||x-\alpha|| \leqq \varepsilon,\ ||t-\beta|| \leqq \varepsilon)$ is in Ω and (2.25) is true when $||\,t-\beta|| \leqq \varepsilon$ and all $\xi^{(i)}$ of (2.21) satisfy $||\xi^{(i)} - \gamma|| \leqq \varepsilon$, as in Lemma 2.2. Now assume that

$$||\Phi(\alpha\,;\,\beta) - \alpha|| < \varepsilon(1-\theta)\,. \tag{2.30}$$

Then, with $x^{(0)}(t) = \alpha$, there exists $\delta \leqq \varepsilon, \delta > 0$, such that, if $||t-\beta|| \leqq \delta$, $x^{(1)}(t)$ exists and

$$||z^{(1)}(t) - \gamma|| = ||x^{(1)}(t) - \alpha|| = ||\Phi(\alpha\,;\,t) - \alpha|| \leqq \varepsilon(1-\theta)\,, \tag{2.31}$$

which is (2.24). Thus, with hypothesis (2.30) replacing (2.2) and δ defined as above, $\{x^{(m)}(t)\}$ converges as specified in Theorem 2.1 even when $(\alpha\,;\,\beta)$ is not a solution of (1.1)

References

1. É. Goursat, *Sur la théorie des fonctions implicites*, Bull. Soc. Math. France **31** (1903), 184–192. Also, cf. Gilbert A. Bliss, Princeton Colloquium Lectures on Mathematics, American Mathematical Society (1913), 16–18.
2. William L. Hart and Theodore S. Motzkin, *A composite Newton-Raphson gradient method*

for the solution of systems of equations, Pacific J. Math. **6** (1956), 691–707.
3. William L. Hart, *Differential equations and implicit functions in infinitely many variables*, Trans. Amer. Math. Soc. **18** (1917), 125–160.
4. ———, *Functions of infinitely many variables in Hilbert space*, Trans. Amer. Math. Soc. **22** (1922), 30–50.

UNIVERSITY OF MINNESOTA
UNIVERSITY OF CALIFORNIA, LOS ANGELES