

A GENERAL APPROACH TO NEWTON'S METHOD
FOR BANACH SPACE PROBLEMS
WITH EQUALITY CONSTRAINTS

BY R. A. TAPIA

Communicated by Hans Weinberger, August 6, 1973

1. Introduction. In this announcement we abstract and preview a theory recently developed by the author for applying Newton's method to constrained problems in infinite-dimensional normed linear spaces. The proofs of these results and a detailed account of the theory will appear elsewhere (see [10] and [11]).

Consider an operator $Q: X_1 \rightarrow X_2$ where X_1 and X_2 are normed linear spaces and functionals $\delta_1, \dots, \delta_m$ defined on X_1 . We are concerned with the constrained problem

$$(1.1) \quad Q(x) = 0; \quad \text{subject to } \delta_i(x) = 0, \quad i = 1, \dots, m.$$

More generally we consider the problem

$$(1.2) \quad Q(x) \in M; \quad \text{subject to } \delta_i(x) = 0, \quad i = 1, \dots, m,$$

where M is a finite-dimensional subspace of X_2 .

DEFINITION 1.1. By a normalization for problem (1.2) we mean any operator P defined from X_1 into a normed linear space such that $P(x) = 0 \Leftrightarrow x$ is a solution of problem (1.2).

The desirability of a normalization for problem (1.2) is obvious; namely it allows us to replace a somewhat unorthodox constrained problem with a standard-type unconstrained problem. The purpose of this paper is to present a theory for constructing a normalization for problem (1.2) which readily lends itself to Newton's method. In §3 we show that the following well-known methods for ordinary differential equations are actually special cases of our theory—the integral equation formulation of the initial value problem and the boundary value problem and the method of quasilinearization and superposition. An interesting feature of our general approach is that we do not consider Newton's method applied to a single equation, but construct a different equation at each iteration (i.e., a different normalization).

An obvious way to obtain a normalization for problem (1.1) is to consider the operator P from X_1 into $X_2 \times R^m$ defined by $P(x) = (Q(x), \delta_1(x), \dots, \delta_m(x))$. Indeed, this is the usual approach when X_1 and

AMS (MOS) subject classifications (1970). Primary 47H15, 49D15, 65J05.

Key words and phrases. Newton method, constrained problems.

X_2 are finite-dimensional. However when X_1 and X_2 are infinite-dimensional this approach is no more than a formal restatement of (1.1) and seldom, if ever, does this approach lead to the construction or implementation of usable algorithms.

Examples of problems which are naturally of the form (1.2) and not of the form (1.1) are the Euler-Lagrange equation and the Euler-Poisson equation from the calculus of variations. In these cases M is a space of polynomials. In §6 our general theory is applied to the Euler-Poisson equations.

2. Normalization of problem (1.2). Many of the operators we use in the construction of a particular normalization will depend on a point $x \in X$. However the majority of the analysis will be performed with this variable held fixed; hence its role is primarily that of an index. For this reason we write P_x (or S_x or T_x or Γ_x) to denote this dependence on x . We use the notation (α_{ij}) to denote a matrix and (α_i) to denote a column vector.

Suppose that $\delta_1, \dots, \delta_m$ and Q in problem (1.2) are Fréchet differentiable at $x \in X_1$ and $Q'(x) = T_x + S_x$, where

$$(2.1) \quad (i) S_x(X_1) \subset M, \quad (ii) \Gamma_x = T_x^{-1} \text{ exists.}$$

Suppose further that

$$(iii) \quad \dim(M) = m \text{ (number of constraints).}$$

When problem (1.2) satisfies (2.1) we will refer to it as problem (2.1).

Let $\{\Phi_1, \dots, \Phi_m\}$ be a basis for M and let $\alpha_1, \dots, \alpha_m$ be functionals (not necessarily linear) on X_1 . Consider

$$(2.2) \quad P(y) = Q(y) - \alpha_1(y)\Phi_1 - \dots - \alpha_m(y)\Phi_m, \quad y \in X_1.$$

PROPOSITION 2.1. *Given $x \in X_1$, if the matrix $(\delta'_i(x)\Gamma_x\Phi_j)$ is nonsingular and P_x is obtained from (2.2) by requiring that the functionals $\alpha_i, i=1, \dots, m$ in (2.2) satisfy*

$$(2.3) \quad (\alpha_j(y)) = (\delta'_i(x)\Gamma_x\Phi_j)^{-1}(\delta'_j(x)\Gamma_x Q(y) - \delta_j(y)) \text{ for } y \in X_1,$$

then P_x is a normalization of problem (2.1), P_x is Fréchet differentiable at x , $P'_x(x)^{-1}$ exists and $\Gamma_x = P'_x(x)^{-1}$.

PROPOSITION 2.2. *Suppose $W: X_1 \rightarrow X_2$ is such that $W(X_1) \subset M$. Then y is a solution of problem (1.2) if and only if it is a solution of problem (1.2) with Q replaced by $Q + W$.*

PROPOSITION 2.3. *The normalization for problem (2.1) obtained by using Proposition 2.1 is representation invariant, i.e., it depends only on the operator Γ_x given in (2.1) and not on the particular operator Q representing problem (2.1) or the particular basis representing M .*

3. Newton's method and local quadratic convergence for problem (2.1).

DEFINITION 3.1. By Newton's method for problem (2.1) we mean the construction of the sequence

$$(3.1) \quad x_{n+1} = x_n - \Gamma_n P_n(x_n), \quad n = 0, 1, 2, \dots,$$

where $x_0 \in X_1$, $\Gamma_n = P'_n(x_n)^{-1}$ and P_n is the normalization for problem (2.1) obtained from Proposition 2.1 by using x_n for x .

PROPOSITION 3.2. *If the functional δ_i , $1 \leq i \leq m$, in problem (1.2) is affine and P_x is the normalization of problem (2.1) obtained using Proposition 2.1 then the Newton iterates (3.1) satisfy the constraint $\delta_i(x) = 0$ beginning with any $x_0 \in X_1$.*

PROPOSITION 3.3. *Let x be a solution of problem (2.1). Suppose that the matrix $(\delta'_i(x)\Gamma_x\Phi_j)$ is nonsingular and the operators $\delta_1, \dots, \delta_m, T$ are of class C^1 . Then there exists a neighborhood of x in X_1 and a constant K such that the Newton sequence (3.1) may be constructed using Proposition 2.1 and converges to x for any x_0 in this neighborhood. Furthermore*

$$\|x - x_{n+1}\| \leq K \|x - x_n\|^2, \quad n = 0, 1, 2, \dots$$

4. **Normalization of problem (1.1).** Suppose we are given problem (1.2) with $\dim(M) = n \leq m$ (number of constraints). Consider a linear operator $D: X_1 \rightarrow X_2$ such that $\dim(N(D)) = m - n$. Let D_R be a linear right inverse for the operator D . The following proposition shows us how to construct an equivalent problem satisfying condition (iii) of (2.1).

PROPOSITION 4.1. *The following two statements are equivalent:*

(i) $Q(x) \in M$;

(ii) $\hat{Q}(x) \in \hat{M}$, where $\hat{Q} = D_R Q$ and $\hat{M} = N(D) + D_R(M)$.

Moreover, in (ii) the dimension of M is equal to m .

The following proposition suggests a method for normalizing problem (1.1).

PROPOSITION 4.2. *Suppose $D_R(x)$ is any linear right inverse for $Q'(x)$ and the dimension of the null space of $Q'(x)$ is equal to m . Then problem (1.1) can be put in the form of problem (2.1) by replacing $Q(y)$ by $D_R(x)Q(y)$, choosing T_x to be the identity map on X_1 and letting $M = N(Q'(x))$.*

5. **Applications to ordinary differential equations.** Consider the second order ordinary differential equation with generalized boundary conditions

$$(5.1) \quad Q(x)(t) = x''(t) - f(t, x, x') = 0; \quad \delta_1(x) = \alpha, \delta_2(x) = \beta.$$

In (5.1), $f \in C^1(R^3)$, $Q: C^2(a, b) \rightarrow C(a, b)$ and δ_1 and δ_2 are functionals on $C^2(a, b)$.

If (5.1) is an initial value problem, i.e., $\delta_1(x) = x(a)$ and $\delta_2(x) = x'(a)$, then the following normalization for problem (5.1) is both classical and well known.

$$(5.2) \quad P(x)(t) = x(t) - \alpha - \beta(t - a) - \iint_{aa}^{ts} f(r, x, x') dr ds.$$

PROPOSITION 5.1. *The normalization (5.2) for the initial value problem (5.1) is the special case of Proposition 2.1 and Proposition 4.1 where $D: C^2(a, b) \rightarrow C(a, b)$ and $D_R: C(a, b) \rightarrow C^2(a, b)$ are given by $D(x) = x''$ and $D_R(x)(t) = \int_a^t \int_a^s x(r) dr ds$.*

If (5.1) is the standard two-point boundary value problem i.e., $\delta_1(x) = x(a)$ and $\delta_2(x) = x(b)$, then the following normalization has been given by Antosiewicz and analyzed in terms of Newton's method (see [1]):

$$(5.3) \quad P(x)(t) = x(t) - \alpha - \frac{(t - a)}{(b - a)}(\beta - \alpha) - \int_a^b g(t, s)f(s, y, y') ds,$$

where

$$(5.4) \quad g(t, s) = (t - s)H(t - s) + (t - a)(s - b)/(b - a) \quad \text{for } a \leq t, s \leq b$$

and $H(t - s)$ denotes the Heaviside unit function.

PROPOSITION 5.2. *The normalization (5.3) for the boundary value problem (5.1) is the special case of Proposition 2.1 and Proposition 4.1 where $D: C^2(a, b) \rightarrow C(a, b)$ and $D_R: C(a, b) \rightarrow C^2(a, b)$ are given by $D(x) = x''$ and $D_R(x)(t) = \int_a^b g(t, s)x(s) ds$ with $g(t, s)$ given by (5.4).*

A well-known method for solving problem (5.1) is the so-called method of quasi-linearization and superposition, i.e., we obtain the solution of the linearized version of (5.1), as a linear combination of the solutions of various canonical initial value problems.

PROPOSITION 5.3. *The method of quasi-linearization and superposition for problem (5.1) is a special case of our generalized Newton's method where the normalizations are obtained via Proposition 2.1 and Proposition 4.2.*

6. Applications to the Euler-Poisson equation. Consider the functional J defined on $C^{n_1}(a, b) \times \cdots \times C^{n_m}(a, b)$ by

$$(6.1) \quad J(y_1, \cdots, y_m) = \int_a^b f(t, y_1, y_1^{(1)}, \cdots, y_m^{(m_1)}, \cdots, y_m, y_m^{(1)}, \cdots, y_m^{(n_m)}) dt,$$

subject to the boundary conditions (linear constraints)

$$(i) \quad y_i(a) = \alpha_{i1}, \quad y_i^{(1)}(a) = \alpha_{i2}, \dots, y_i^{(n_i-1)}(a) = \alpha_{in_i},$$

(6.2)

$$(ii) \quad y_i(b) = \beta_{i1}, \quad y_i^{(1)}(b) = \beta_{i2}, \dots, y_i^{(n_i-1)}(b) = \beta_{in_i}, \quad i = 1, \dots, m.$$

A classical and well-known argument using the second fundamental lemma of the calculus of variations shows that any $y_1 \times \dots \times y_m$ which gives an extrema of the functional (6.1) over the set defined by (6.2) must satisfy

$$(6.3) \quad \left(f_{y_1^{(n_1-1)}} - \int_a^t f_{y_1^{(n_1-1)}} + \dots + (-1)^{n_1} \int_a^t \dots \int_a^t f_{y_1} \right) \\ \times \dots \times \left(f_{y_m^{(n_m)}} - \int_a^t f_{y_m^{(n_m-1)}} + \dots + (-1)^{n_m} \int_a^t \dots \int_a^t f_{y_m} \right) \\ \in \Pi_{n_1-1} \times \dots \times \Pi_{n_m-1}$$

where π_i denotes the polynomial of degree i .

If we introduce the change of variable $z_i = y_i^{(n_i)}$, then we have

$$(6.4) \quad \begin{aligned} y_i^{(n_i)} &= z_i, \\ y_i^{(n_i-1)} &= \alpha_{in_i} + \int_a^t z_i, \\ &\vdots \\ y_i &= \alpha_{i1} + \alpha_{i2}(t-a) + \dots + \alpha_{in_i} \frac{(t-a)^{(n_i-1)}}{(n_i-1)!} + \int_a^t \dots \int_a^t z_i. \end{aligned}$$

Also, in terms of z_i the boundary conditions (6.2) (ii) become

$$(6.5) \quad \begin{aligned} \int_a^b z_i &= \beta_{in_i} - \alpha_{in_i}, \\ \int_a^b \int_a^t z_i &= \beta_{in_{i-1}} - \alpha_{in_{i-1}} - \alpha_{in_i}(b-a), \dots, \\ \int_a^b \int_a^t \dots \int_a^t z_i &= \beta_{i1} - \alpha_{i1} - \alpha_{i2}(b-a) \\ &\quad - \dots - \alpha_{in_i}(b-a)^{(n_i-1)} / (n_i-1)!. \end{aligned}$$

If we now make the change of variables (6.4) in (6.3) then we have an example of problem (2.1) where $X_1 = X_2 = C(a, b) \times \dots \times C(a, b)$, δ_i are given by (6.5), $M = \Pi_{n_1-1} \times \dots \times \Pi_{n_m-1}$ and the operator $Q(z)$ is given by (6.3). Observe that the number of linear constraints and the dimension of

M are both equal to $n_1 + \cdots + n_m$. We can now apply our general theory of Newton's method. The solution to the original problem (and its derivatives) is then obtained from (6.4).

7. An extension of the theory to optimization problems. Let $f, \delta_1, \cdots, \delta_m$ be functionals defined on a normed linear space X which are Fréchet differentiable. Consider the constrained optimization problem:

$$(7.1) \quad \text{minimize } f(x); \quad \text{subject to } \delta_i(x) = 0, \quad i = 1, \cdots, m.$$

It is a well-known result in the theory of Lagrange multipliers that under mild regularity conditions any solution of problem (7.1) must also be a solution of

$$(7.2) \quad f'(x) \in M(x); \quad \text{subject to } \delta_i(x) = 0, \quad i = 1, \cdots, m.$$

where $M(x)$ denotes the linear span of $\delta'_1(x), \cdots, \delta'_m(x)$. The theory of Newton's method developed in the previous sections will handle problem (7.2) if and only if the functionals $\delta_1, \cdots, \delta_m$ in (7.2) are linear. In [11] we develop the analogous theory of Newton's method for problem (7.2).

REFERENCES

1. H. A. Antosiewicz, *Newton's method for boundary value problems*, J. Comput. System Sci. **2** (1968), 177–202. MR **39** #5055.
2. J. E. Dennis, Jr., *Variations on Newton's method*, Doctoral thesis, University of Utah, Salt Lake City, Ut., 1966.
3. M. Golomb and R. A. Tapia, *The metric gradient in normed linear spaces*, Numer. Math. **20** (1972), 115–124.
4. J. M. Ortega and W. C. Rheinboldt, *Iterative solutions of nonlinear equations in several variables*, Academic Press, New York, 1970. MR **42** #8686.
5. L. B. Rall, *Computational solution of nonlinear operator equations*, Wiley, New York, 1969. MR **39** #2289.
6. S. M. Robinson, *Extension of Newton's method to mixed systems of nonlinear equations and inequalities*, Rep. 1161, Mathematics Research Center, University of Wisconsin, Madison, Wisc., 1971.
7. M. Stein, *On methods of obtaining solutions of fixed end point problems in the calculus of variations*, Doctoral thesis, University of California, Los Angeles, Calif., 1950.
8. R. A. Tapia, *An application of a Newton-like method to the Euler-Lagrange equation*, Pacific J. Math. **29** (1969), 235–246. MR **40** #3713.
9. ———, *The weak Newton method and boundary value problems*, SIAM J. Numer. Anal. **6** (1969), 539–550. MR **41** #9439.
10. ———, *Newton's method for problems with equality constraints*, SIAM J. Numer. Anal. (to appear).
11. ———, *Newton's method for optimization problems with equality constraints*, SIAM J. Numer. Anal. (to appear).