—NOTES—

SOLUTION OF NON-LINEAR EQUATIONS

By ANDRÉ N. GLEYZAL (*U. S. Naval Ordnance Laboratory*)

A standard procedure for solving a set of n equations in n variables:

$$f_i(x_j) = 0, \qquad i, j = 1, \cdots, n, \tag{1}$$

where n is an integer and the f_i are real functions of the variables x_j may be described as follows. We start with a trial point P as a first approximation, and expand the function f_i in the vicinity of P. Retaining only the linear portions of the expansions we solve the resulting linear equations obtaining a point Q. The point P is now replaced by Q and the calculation is repeated. It has been shown under mildly restrictive conditions that this process yields a rapidly convergent series if the trial point P is sufficiently close to the solution.

We consider the procedure where instead of Q we select a point R on the line PQ which gives a minimum value to the sum of the squares of the function f_i .

Suppose $P \equiv (\xi_j)$ is an approximate solution. Then, assuming the functions (1) are analytic, let

$$\varphi = \varphi(dx_j) = \sum_i \left[f_i(\xi_j) + \sum_j \frac{\partial f_i}{\partial x_j}\, dx_j + \frac{1}{2!} \sum_{k,l} \frac{\partial^2 f_i}{\partial x_k\, \partial x_l}\, dx_k\, dx_l + \cdots \right]^2,$$

where

$$dx_j = x_j - \xi_j\,, \qquad i, j, k, l = 1, \cdots, n.$$

Let

$$\psi = \psi(dx_j) = \sum_i \left[f_i(\xi_j) + \sum_j \frac{\partial f_i}{\partial x_j}\, dx_j \right]^2.$$

Thus ψ differs from φ by terms of order two or higher in dx_j . Let

$$dx_j = du_j$$

be a solution* of the linear equations:

$$f_i(\xi_j) + \sum_j \frac{\partial f_i}{\partial x_j}\, dx_j = 0.$$

Thus $Q = (\xi_j + du_j)$. Now consider the two functions of λ:

$$\varphi = \varphi(\lambda\, du_j) = \varphi(\lambda),$$

$$\psi = \psi(\lambda\, du_j) = \psi(\lambda).$$

For $\lambda = 0$:

$$\varphi = \psi, \qquad d\varphi/d\lambda = d\psi/d\lambda.$$

Received November 4, 1957; revised manuscript received January 16, 1958.

*See Householder, *Principles of numerical analysis*, McGraw-Hill, New York, 1953, pp. 81-85.

Also, $\psi(\lambda)$ is a parabola which takes on its minimum value, zero, at $\lambda = 1$. The function $\varphi(\lambda)$ is always non-negative. Hence $\varphi(\lambda)$ must reach a relative minimum for a value $\lambda = \lambda_m$, where $0 < \lambda_m \leqq \infty$. The quantity λ_m may readily be calculated and the corresponding point:

$$R \equiv (\xi_i + \lambda_m \, du_i)$$

then used as a starting point for the calculation in place of ξ_i. Thus φ must decrease with every iteration. Moreover, the iterative method described above will converge in one step if Eqs. (1) are linear, and it has been found in many calculations that the method converges rapidly if the initial approximation is sufficiently close to the solution.

If the derivatives $\partial f_i/\partial x_j$ are not readily found one may approximate them by difference quotients $\Delta f_i/\Delta x_j$. In extremely non-linear problems where independent and dependent variables become very small or very large in the process of calculation the choice of suitable increments Δx_j may not be obvious. In this case one may take the increments Δx_j to be differences in successive approximations of x_j. It has been found in many calculations that for sufficiently smooth functions approximately maximum numerical accuracy is then obtained for the solution.

The method described here was successful in solving certain extremely non-linear chemical equilibrium equations where independent and dependent variables were unpredictably small or large. In other methods—such as the "steepest descent method"—which were tried the rate of convergence decreased rapidly and the method became impractical as the solution was approached, whereas in this method the rate of convergence increased rapidly. Combinations of the steepest descent method with this method suggest themselves. For example, let

$$x_j = \xi_j + \lambda \, du_j + \mu \, dv_j \, ,$$

where μ is a parameter and $dv_j = -\text{grad } \varphi$. The parameters λ and μ are then adjusted so that φ is minimum.

ON CONVERGENT PERTURBATION EXPANSIONS*

By RICHARD BELLMAN (*The Rand Corporation*)
and TOMLINSON FORT (*University of South Carolina*)

1. Introduction. In this paper, we wish to consider the Sturm-Liouville equation

$$\begin{aligned} u'' + \lambda(f(x) + \epsilon g(x))u &= 0, \\ u(0) = u(1) &= 0, \end{aligned} \tag{1}$$

and the problem of obtaining power series expansions for the first characteristic value and first characteristic function.

Let λ_0 and $u_0(x)$ be respectively the first characteristic value and associated characteristic function of the equation

$$\begin{aligned} u'' + \lambda f(x)u &= 0, \\ u(0) = u(1) &= 0. \end{aligned} \tag{2}$$

*Received February 24, 1958.