noise temperature, it still remains to be explained why an even greater change does not take place. An expression for g_m is given by:[3]

$$g_m = \frac{\mu e N a}{L} f$$

where

μ = carrier mobility

N = donor density

L = gate length

a = channel thickness

f = factor dependent on channel and depletion-layer geometry and pinchoff voltage

In the temperature range under investigation, the carrier mobility varies as T^{-2}. One would therefore expect g_m to increase by a factor of $(313/208)^2 = 2{\cdot}26$. A smaller increase in g_m is due to a change in the depletion-region geometry brought about, in turn, by variation in the saturation field E_s with ambient temperature. To avoid 'hot-electron' effects, which would increase the noise temperature and reduce the amplifier gain, the drain voltage V_d is adjusted so that the electric field at the drain end of the channel reaches the saturation value E_s. As the temperature is lowered, E_s decreases and V_d is adjusted accordingly, for minimum T_e. However, a lower V_d changes the depletion-region geometry, and it becomes less tapered, reducing the factor f in the g_m expression. A similar conclusion can be drawn from the analysis of Hower and Bechtel:[4] with decreasing temperature, the ratio of the pinchoff voltage to the saturation voltage increases, resulting in a smaller increase in g_m than that expected from the increase in the carrier mobility.

P. BURA

RCA Ltd.
Ste-Anne-de-Bellevue
Montreal, Que., Canada

9th April 1974

References

1 JIMENEZ, J. J., OLIVA, J. A., and SEPTIER, A.: 'Very low noise cryogenic mesfet amplifier' Digest of European microwave conference, 1973
2 LORIOU, B., BELLEC, M., and LEROUZIC, M.: 'Performance à basse température d'un transistor hyperfréquences faible bruit à effet de champ', *Electron. Lett.*, 1970, **6**, pp. 819–820
3 WOLF, P.: 'Microwave properties of Schottky-barrier field-effect transistors', *IBM J. Res. Develop.*, 1970, **14**, pp. 125–141
4 HOWER, P.L., and BECHTEL, N. G.: 'Current saturation and small-signal characteristics of GaAs field-effect transistors', *IEEE Trans.*, 1973, **ED-20**, pp. 213–220
5 BAECHTOLD, W.: 'Noise behavior of Schottky-barrier transistors at microwave frequencies', *ibid.*, 1971, **ED**–18, pp. 97–104
6 ROBINSON, F. N. H.: 'Noise in common-source f.e.t. amplifiers at moderately high frequencies', *Electron. Eng.*, 1969, **41**, pp. 77–79

POWERFUL 2-PART PROGRAM FOR SOLUTION OF NONLINEAR SIMULTANEOUS EQUATIONS

Indexing terms: Nonlinear equations, Numerical methods

A composite program that has proved very effective in solving sets of highly nonlinear equations from a wide range of starting values is outlined. Either first, or first and second, partial derivatives can be used in the first part of the program, and two Tables of results are shown for a problem of eight equations in eight unknowns.

Introduction: By a 2-part program for the solution of sets of nonlinear simultaneous equations is meant a composite package in which a final damped Newton–Raphson section, suitably modified to cope with overdetermined sets of equations, if necessary, is preceded by an initial section in which the sum of squares, say, of the residuals, is reduced, e.g. by the conjugate-gradient method of Fletcher and Reeves.[1] In this way, it has been found possible to obtain accurate solutions to some problems of engineering interest that hitherto neither section of the composite program had been able to locate on its own.[2–4] After considerable numerical experimentation, it was found that the best strategy to employ for difficult problems was to enter the Newton-Raphson section prior to every iteration performed within the first section, returning to the first section of the program if certain empirically based criteria indicated that convergence within the second section was unlikely to be obtained. However, it was only possible to obtain an accurate solution from the range of starting values shown in Tables 1 and 2 for the test problem defined by eqns. 8 and 9 by employing a more powerful type of gradient-descent method in the first section, the essential features of which will now be indicated.

Some details of powerful gradient-descent method: If the equations to be solved are written as

$$\left.\begin{array}{l} f_i(x_j) = 0 \\ i = 1, 2, 3, \ldots, m \\ j = 1, 2, 3, \ldots, n \\ m > n \end{array}\right\} \quad \ldots \ldots \quad (1)$$

and a function F is defined as

$$F = \sum_{i=1}^{m} f_i^2 \quad \ldots \ldots \quad (2)$$

the conventional method of steepest descent can be written as

$$\delta_j = -\mu \frac{\partial F}{\partial x_j} \quad \ldots \ldots \quad (3)$$

$\mu > 0$ for descent, where δ_j is the correction to be added to the variable x_j, and μ is varied to yield a minimum value of F. Extending eqn. 3 to include the second partial derivatives of F, we have

$$\delta_j = -\mu \left\{ \frac{\partial F}{\partial x_j} + \left(\sum_{i=1}^{m} \delta_i \frac{\partial^2 F}{\partial x_i \partial x_j} \right) \right\} \quad \ldots \quad (4)$$

which, in vector–matrix form, becomes

$$\boldsymbol{\delta} = -\mu[\boldsymbol{G} + H\boldsymbol{\delta}] \quad \ldots \ldots \quad (5)$$

and, writing

$$\lambda = -\frac{1}{\mu} \quad \ldots \ldots \quad (6)$$

we have

$$\boldsymbol{\delta} = -(H - \lambda I)^{-1} \boldsymbol{G} \quad \ldots \ldots \quad (7)$$

where $\boldsymbol{G}$ is the gradient vector and H the Hessian of F; in terms of f, $\boldsymbol{G} = 2J^T \boldsymbol{f}$. Thus $\lambda = +\infty$ corresponds to the direction of steepest ascent, and $\lambda = -\infty$ corresponds to the direction of steepest descent. $\lambda = 0$ corresponds to the stationary value of F, given by $\boldsymbol{\delta} = -H^{-1}\boldsymbol{G}$, if F is a quadratic function; this stationary value will be a minimum if H is positive definite. $\lambda = \lambda_i$, where λ_i is an eigenvalue of H, yields an infinite vector of corrections, $\boldsymbol{\delta}$, and hence a discontinuity in the value of F.

The eigenvalues λ_i are now calculated, and, with a knowledge of the discontinuous behaviour with respect to the parameter λ of the function $F(\lambda)$, it is possible to calculate relatively efficiently all the significant minima of the multimodal function $F(\lambda)$ over the whole range $-\infty \leqslant \lambda \leqslant +\infty$. In practice, the components of the vector of corrections are scaled down, if necessary, to ensure that the modulus of the largest correction component does not exceed a prescribed value. The sequence of events that is given first preference is for this part of the program to proceed from the global minimum of all minima of $F(\lambda)$ found at one iteration to the global minimum of the following iteration. However, if it appears that following this sequence of first choice is leading to a petering out of progress in the reduction of the function F, provision is made for automatically following alternative paths by restarting from a minimum other than the global minimum at some previous iteration; the precise details of this operation are determined by the data read-in at run time.

Test problem and Tables of results: The test problem of eight equations in eight unknowns is defined by eqns. 8 and 9:

$$
\left.\begin{aligned}
f_i &\equiv x_3(1-x_1x_2)[\exp\{x_4(Y_{1i}-Y_{3i}x_6\times10^{-3}-Y_{5i}x_7\\
&\qquad\times10^{-3}\}-1]-Y_{5i}+Y_{4i}x_2=0\\
f_{i+4} &\equiv \{x_1x_3(1-x_1x_2)/x_2\}[\exp\{x_5(Y_{1i}-Y_{2i}\\
&\qquad-Y_{3i}x_6\times10^{-3}+Y_{4i}x_8\times10^{-3})\}-1]\\
&\qquad\qquad-Y_{5i}x_1+Y_{4i}=0\\
&\qquad\qquad i=1,2,3,4
\end{aligned}\right\} \quad (8)
$$

with

$$
\left.\begin{aligned}
&Y_{11}=0{\cdot}485\; Y_{21}=0{\cdot}369\; Y_{31}=\ \ 5{\cdot}2095\; Y_{41}=23{\cdot}3037\\
&Y_{12}=0{\cdot}752\; Y_{22}=1{\cdot}254\; Y_{32}=10{\cdot}0677\; Y_{42}=101{\cdot}779\\
&Y_{13}=0{\cdot}869\; Y_{23}=0{\cdot}703\; Y_{33}=22{\cdot}9274\; Y_{43}=111{\cdot}461\\
&Y_{14}=0{\cdot}982\; Y_{24}=1{\cdot}455\; Y_{34}=20{\cdot}2153\; Y_{44}=191{\cdot}267\\
&\text{and}\\
&Y_{5i}=Y_{4i}+Y_{3i}\qquad i=1,2,3,4
\end{aligned}\right\} \quad (9)
$$

Table 1 shows a set of results obtained when the Hessian matrix of second partial derivatives required in the first part of the program is calculated accurately, and Table 2 shows the

Table 1 RESULTS WITH HESSIAN MATRIX IN FIRST PART OF PROGRAM

Initial input values of x_i, $i = 1, \ldots, 8$	Sum of squares of residuals $\sum_{i=1}^{8} f_i^2$		Number of restarts	Computing time
	Initial	Final		min s
0·1	$1{\cdot}10\times10^{5}$	$6{\cdot}77\times10^{-19}$	0	3 25
0·3	$6{\cdot}72\times10^{4}$	$2{\cdot}63\times10^{-17}$	0	2 12
0·5	$3{\cdot}50\times10^{4}$	$4{\cdot}71\times10^{-17}$	0	1 32
0·7	$1{\cdot}36\times10^{4}$	$2{\cdot}91\times10^{-17}$	0	1 9
0·9	$3{\cdot}24\times10^{3}$	$1{\cdot}31\times10^{-17}$	0	0 52
1·0	$2{\cdot}13\times10^{3}$	$1{\cdot}78\times10^{-17}$	0	0 36
2·0	$1{\cdot}33\times10^{5}$	$6{\cdot}09\times10^{-18}$	0	2 11
3·0	$5{\cdot}30\times10^{5}$	$1{\cdot}49\times10^{-18}$	0	1 39
4·0	$1{\cdot}77\times10^{6}$	$2{\cdot}54\times10^{-17}$	0	3 30
5·0	$1{\cdot}25\times10^{7}$	$3{\cdot}72\times10^{-17}$	0	2 32
6·0	$2{\cdot}51\times10^{8}$	$1{\cdot}52\times10^{-14}$	0	3 49
7·0	$9{\cdot}82\times10^{9}$	$8{\cdot}54\times10^{-17}$	0	4 7
8·0	$7{\cdot}27\times10^{11}$	$1{\cdot}72\times10^{-14}$	1	11 20
9·0	$1{\cdot}30\times10^{14}$	$1{\cdot}41\times10^{-16}$	1	19 35
10·0	$5{\cdot}63\times10^{16}$	$3{\cdot}97\times10^{-17}$	0	7 8
11·0	$5{\cdot}02\times10^{19}$	$2{\cdot}84\times10^{-17}$	0	4 51
12·0	$8{\cdot}58\times10^{22}$	$3{\cdot}69\times10^{-17}$	0	10 37

Table 2 RESULTS WITH APPROXIMATION TO HESSIAN IN FIRST PART OF PROGRAM

Initial input values of x_i, $i = 1, \ldots, 8$	Sum of squares of residuals $\sum_{i=1}^{8} f_i^2$		Number of restarts	Computing time
	Initial	Final		min s
0·1	$1{\cdot}10\times10^{5}$	$5{\cdot}65\times10^{-17}$	7	17 23
0·3	$6{\cdot}72\times10^{4}$	$2{\cdot}20\times10^{-17}$	3	7 34
0·5	$3{\cdot}50\times10^{4}$	$3{\cdot}18\times10^{-17}$	0	1 31
0·7	$1{\cdot}36\times10^{4}$	$5{\cdot}83\times10^{-17}$	0	2 7
0·9	$3{\cdot}24\times10^{3}$	$2{\cdot}55\times10^{-17}$	1	2 39
1·0	$2{\cdot}13\times10^{3}$	$1{\cdot}15\times10^{-16}$	0	2 1
2·0	$1{\cdot}33\times10^{5}$	$4{\cdot}98\times10^{-17}$	2	9 20
3·0	$5{\cdot}30\times10^{5}$	$1{\cdot}79\times10^{-14}$	0	3 23
4·0	$1{\cdot}77\times10^{6}$	$1{\cdot}84\times10^{-17}$	0	1 2
5·0	$1{\cdot}25\times10^{7}$	$2{\cdot}89\times10^{-17}$	0	1 50
6·0	$2{\cdot}51\times10^{8}$	$2{\cdot}89\times10^{-17}$	0	2 11
7·0	$9{\cdot}82\times10^{9}$	$1{\cdot}44\times10^{-16}$	0	7 24
8·0	$7{\cdot}27\times10^{11}$	$3{\cdot}80\times10^{-17}$	0	3 12
9·0	$1{\cdot}30\times10^{14}$	$1{\cdot}06\times10^{-14}$	0	6 4
10·0	$5{\cdot}63\times10^{16}$	$1{\cdot}53\times10^{-14}$	0	7 8
11·0	$5{\cdot}02\times10^{19}$	$6{\cdot}48\times10^{-17}$	1	20 9
12·0	$8{\cdot}58\times10^{22}$	$4{\cdot}64\times10^{-15}$	2	33 48

corresponding set of results when the Hessian matrix is approximated to by $2J^T J$, where J is the Jacobian matrix of first partial derivatives, i.e. $J_{ij} = \partial f_i / \partial x_j$. Apart from this, every run detailed in Tables 1 and 2 was obtained by running in the domain of the natural logarithms of the independent variables with an identical ALGOL program and with the same values for all preset parameters.* The same set of solutions was obtained in each case, namely that shown, to 4-decimal-figure accuracy, in eqns. 10:

$$\left.\begin{aligned} x_1 = 0{\cdot}9000 \quad x_2 = 0{\cdot}4500 \quad x_3 = 1{\cdot}0000 \quad x_4 = 8{\cdot}0000 \\ x_5 = 8{\cdot}0000 \quad x_6 = 5{\cdot}0000 \quad x_7 = 1{\cdot}0000 \quad x_8 = 2{\cdot}0000 \end{aligned}\right\} \qquad (10)$$

One other set of solutions to this problem, in which not all variables possess positive values, has been found. It is hoped that a more comprehensive paper will be published later.

Acknowledgments: The author is indebted to J. T. Henderson for his general assistance in connection with the work described above. The details for the test problem were kindly provided by J. K. Skwirzynski and G. A. Richards of the Research Division of the Marconi Co., Great Baddow.

* The computer used was an ICL (Elliott) 4130 running under DES 1 operating system

O. P. D. CUTTERIDGE *4th April 1974*

Department of Engineering
University of Leicester
Leicester LE1 7RH, England

References

1 FLETCHER, R., and REEVES, C. M.: 'Function minimization by conjugate gradients', *Comput. J.*, 1964, **7**, pp. 149–154
2 CUTTERIDGE, O. P. D.: 'Numerical experience with some two-part programmes for the solution of non-linear simultaneous equations'. Presented at conference on numerical methods for non-linear optimization, University of Dundee, 1971. University of Leicester Engineering Department Report 72–20
3 CUTTERIDGE, O. P. D.: 'Electrical network and general system synthesis using optimization techniques' *in* AVRIEL, M., RIJCKAERT, M. J., and WILDE, D. J. (Eds.): 'Optimization and design' (Prentice Hall, 1973), pp. 76–80
4 CUTTERIDGE, O. P. D.: 'Computer synthesis of lumped linear networks of arbitrary structure' *in* SKWIRZYNSKI, J. K., and SCANLAN, J. O. (Eds.): 'Network and signal theory' (Peter Peregrinus, 1973), pp. 105–111

REPLY TO COMMENT ON CALCULATION OF THE RADIATION PATTERNS OF REFLECTOR ANTENNAS, USING KIRCHHOFF INTEGRATION

The remarks made by Van Hoof in his comment on a recent letter[A] have led to some calculations of the radiation patterns of paraboloids excited by a number of mathematical models of antenna feeds. The results with a feed pattern producing uniform illumination over the aperture of the paraboloid have been discussed before.[A] We have extended the calculations to a feed producing a constant field strength within the solid angle subtended by the main reflector (Fig. 1*a*) and having a tapered cosine pattern. However, the tapered cosine patterns have been truncated at the angular aperture Ψ of the main reflector. The cosine patterns are then defined by

$$D_f(\psi) = \text{constant} \times \cos^n \psi \qquad 0 < \psi < \Psi$$

and

$$D_f(\psi) = 0 \qquad \psi > \Psi$$

As

$$\int_{4\pi} D(\psi, \xi)\, d\Omega = 4\pi$$

it is readily found that

$$\text{constant} = \frac{2(n+1)}{1 - \cos^{n+1} \Psi}$$

By taking different values for n, the edge illumination of the main reflector may be modified. The results are shown in Table 1.

The results with nontruncated primary cosine patterns have been mentioned before.[A] We have calculated the power distribution once more for $D = 50\lambda$, and have found that

$$\int_0^\pi D_E(\theta) \sin\theta\, d\theta = \int_0^\pi D_H(\theta) \sin\theta\, d\theta = 2$$

for $n = 2$ and $n = 4$.

A second experiment has been carried out with truncated patterns producing a uniform illumination over the aperture; the truncation was not at $\psi = \Psi$, but at larger angles (Table 2). The results from these experiments enable us to draw some conclusions. It appears that the edge illumination plays a rôle in the correctness of the power distribution if the truncation angle of the feed pattern coincides with the subtending

Table 1

	Ψ	Edge illumination	$\int_0^\pi D_E(\theta) \sin\theta\, d\theta$	$\int_0^\pi D_H(\theta) \sin\theta\, d\theta$
	deg	dB		
Uniform illumination over aperture	60	0	2·14	2·16
Constant field strength pattern	60	−2·5	2·12	2·14
Cosine pattern, $n = 2$	60	−8·5	2·06	2·07
Cosine pattern, $n = 4$	60	−14·5	2·02	2·03

Diameter of main reflector = 50λ

Table 2

D/λ	Ψ, reflector	Ψ, feed	$\int_0^\pi D_E(\theta) \sin\theta\, d\theta$	$\int_0^\pi D_H(\theta) \sin\theta\, d\theta$
	deg	deg		
50	60	60	2·14	2·16
50	60	70	1·99	2·01
50	60	80	1·99	2·00