Bulletin of the Seismological Society of America, Vol. 75, No. 3, pp. 779–790, June 1985

NONLINEAR EARTHQUAKE LOCATION: THEORY AND EXAMPLES

BY CLIFFORD H. THURBER

ABSTRACT

A fundamental modification to Geiger's method of earthquake location for local earthquakes is described which incorporates nonlinear behavior of travel time as a function of source position. The use of Newton's method rather than the usual Gauss-Newton method allows the inclusion of second-order partial derivatives of travel time with respect to source coordinates in the location algorithm. These second-order derivatives can be calculated quite easily for half-space and layered crustal models. Expected benefits are improved convergence and stability, as demonstrated in a series of examples, and more realistic assessment of solution uncertainty.

INTRODUCTION

Geiger's method (Geiger, 1910, 1912) is almost universally used for earthquake location [see Tarantola and Vallette (1983) for a notable exception]. Although the most common algorithms for local earthquakes currently in use (Lee and Lahr, 1975; Klein, 1978) are successful in most situations, several studies have drawn attention to deficiencies in these algorithms and possible means for improvement (Smith, 1976; Buland, 1976; Lee and Stewart, 1981; Anderson, 1982; Lienert and Frazer, 1983). These studies have addressed almost exclusively the step of computing the hypocenter adjustment vector given the partial derivative matrix and residual vector, discussing a variety of schemes for weighting, centering, scaling, transforming, and damping the same basic set of equations.

An issue which has received little attention is the linearization underlying Geiger's method. Lee and Stewart (1981) point out that Geiger's method is an example of the Gauss-Newton method for optimization, which is prone to failure for a variety of reasons. The studies listed above describe methods which are not fundamentally different from Geiger's original method, and which suffer some of the same problems. However Newton's method, which is nonlinear in that it requires second-order partial derivatives, holds promise for success in situations where the Gauss-Newton method is known to fail. Mellman (1980) describes one of the few seismological studies in which Newton's method is used; some efforts at using a nonlinear method for teleseismic location have also been made (IBM, 1965). Following a brief discussion of the Gauss-Newton and Newton's methods, we will present a derivation of the second-order partial derivatives necessary for Newton's method in half-space and layered velocity structures, indicating the situations in which these higher order terms are important. Simple examples will be presented to illustrate the utility of the nonlinear method. Detailed numerical studies will be required to assess fully the role of nonlinearity in the earthquake location problem.

MATHEMATICAL BACKGROUND

In Geiger's method modified for local earthquakes, a set of equations are constructed of the form

$$r_i = \frac{\partial t_i}{\partial x}\,\Delta x + \frac{\partial t_i}{\partial y}\,\Delta y + \frac{\partial t_i}{\partial z}\,\Delta z + \Delta t_0 \tag{1}$$

where $r_i = t_i^{obs} - t_i^{cal}$ is the arrival time residual for the ith station; $\partial t_i/\partial x$, $\partial t_i/\partial y$, and $\partial t_i/\partial z$ are partial derivatives of the travel time to station i with respect to the Cartesian coordinates of the hypocenter; and Δt_0, Δx, Δy, and Δz are the desired adjustments to the earthquake origin time and location. The set of equations can be written in matrix notation as

$$\boldsymbol{A}\ \delta\boldsymbol{x} = \boldsymbol{r} \tag{2}$$

where $\boldsymbol{A}$ is the matrix of partial derivatives, $\delta\boldsymbol{x}$ is the adjustment vector (Δx, Δy, Δz, Δt_0), and $\boldsymbol{r}$ is the residual vector. Published algorithms for computing the solution $\delta\boldsymbol{x}$ given $\boldsymbol{A}$ and $\boldsymbol{x}$ include least squares (Geiger, 1910, 1912), step-wise regression (Lee and Lahr, 1975), QR decomposition (Buland, 1976), generalized inverse (Bolt, 1970; Klein, 1978), and damped least squares (Herrmann, 1979). In all these cases, a purely linear equation is solved iteratively to derive the solution to this inherently nonlinear problem.

To find a significantly different approach to this problem, we must step back to the mathematics which underly such optimization methods. Following Lee and Stewart (1981) (but note our differing definitions of $\boldsymbol{A}$), we wish to minimize the objective function F given by

$$F(\boldsymbol{x}) = \boldsymbol{r}^T\boldsymbol{r} \tag{3}$$

where $\boldsymbol{x}$ is the vector of hypocentral parameters. Expanding F in a Taylor series about $\boldsymbol{x}$, we write

$$F(\boldsymbol{x} + \delta\boldsymbol{x}) = F(\boldsymbol{x}) + \boldsymbol{g}^T\delta\boldsymbol{x} + \tfrac{1}{2}\,\delta\boldsymbol{x}^T\boldsymbol{H}\delta\boldsymbol{x} + \cdots \tag{4}$$

where $\boldsymbol{g}$ is the gradient vector, and $\boldsymbol{H}$ is the Hessian matrix. It is straightforward to show that

$$\boldsymbol{g} = -2\ \boldsymbol{A}^T\boldsymbol{r} \tag{5}$$

and

$$\boldsymbol{H} = 2[\boldsymbol{A}^T\boldsymbol{A} - (\nabla\boldsymbol{A}^T)\boldsymbol{r}] \tag{6}$$

where ∇ is the vector gradient operator. Minimization (actually extremization) of F leads to

$$\delta\boldsymbol{x} = -\boldsymbol{H}^{-1}\boldsymbol{g}. \tag{7}$$

For the least-squares or Gauss-Newton solution, the second term on the right-hand side of (6) is neglected, yielding

$$\delta\boldsymbol{x} = (\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{A}^T\boldsymbol{r}. \tag{8}$$

In damped least squares, a small constant is added to the diagonal elements of $\boldsymbol{A}^T\boldsymbol{A}$ in (8). For Newton's method, the full nonlinear term is retained to give the solution

$$\delta\boldsymbol{x} = [\boldsymbol{A}^T\boldsymbol{A} - (\nabla\boldsymbol{A}^T)\boldsymbol{r}]^{-1}\boldsymbol{A}^T\boldsymbol{r}. \tag{9}$$

If we define the 4×4 matrix

$$N = (\nabla A^T)r \tag{10}$$

then the element N_{ij} is given by

$$N_{ij} = \sum_{k=1}^{m} \frac{\partial^2 t_k}{\partial x_i \partial x_j} r_k \tag{11}$$

for $i, j = 1, 2, 3$ (i.e., for the Cartesian coordinates) and where t_k and r_k are the travel time and corresponding residual at the kth station (we also note that $N_{ij} = 0$ for all i or $j = 4$, corresponding to the origin time). Thus, the solution given by Newton's method incorporates nonlinear behavior. We will demonstrate in the following sections that computation of the second-order partial derivatives required in (11) is simple in most common situations, and also will discuss some cases in which these nonlinear terms are dominant.

Travel-Time Derivatives in a Half-Space

The rather trivial example of second-order derivatives in a half-space of constant velocity V_0 serves to illustrate the nature and importance of these nonlinear terms. Given a source at (x_1, y_1, z_1) and a receiver at (x_0, y_0, z_0), the travel time and first-order partial derivatives are expressed as

$$t = \frac{1}{V_0}[(x_1 - x_0{}^2) + (y_1 - y_0)^2 + (z_1 - z_0)^2]^{1/2} = \frac{S}{V_0}$$

$$\frac{\partial t}{\partial x_1} = \frac{x_1 - x_0}{V_0 S} = -\frac{\sin\theta \cos\phi}{V_0 S}$$

$$\frac{\partial t}{\partial y_1} = \frac{y_1 - y_0}{V_0 S} = -\frac{\sin\theta \sin\theta}{V_0 S}$$

$$\frac{\partial t}{\partial z_1} = \frac{z_1 - z_0}{V_0 S} = -\frac{\cos\theta}{V_0 S} \tag{12}$$

where S is the path length, and θ and ϕ are take-off angle and azimuth (see Figure 1). Note that the first-order derivatives are directly proportional to the corresponding components of the ray vector at the source. This relation holds true in essentially any medium (Lee and Stewart, 1981).

The second-order derivatives can also be calculated analytically, yielding

$$\frac{\partial^2 t}{\partial z_1{}^2} = \frac{1}{V_0 S}\left[1 - \frac{(z_1 - z_0)^2}{S^2}\right]$$

$$\frac{\partial^2 t}{\partial y_1 \partial z_1} = -\frac{(y_1 - y_0)(z_1 - z_0)}{V_0 S^3}$$

$$\frac{\partial^2 t}{\partial x_1 \partial z_1} = -\frac{(x_1 - x_0)(z_1 - z_0)}{V_0 S^3} \tag{13}$$

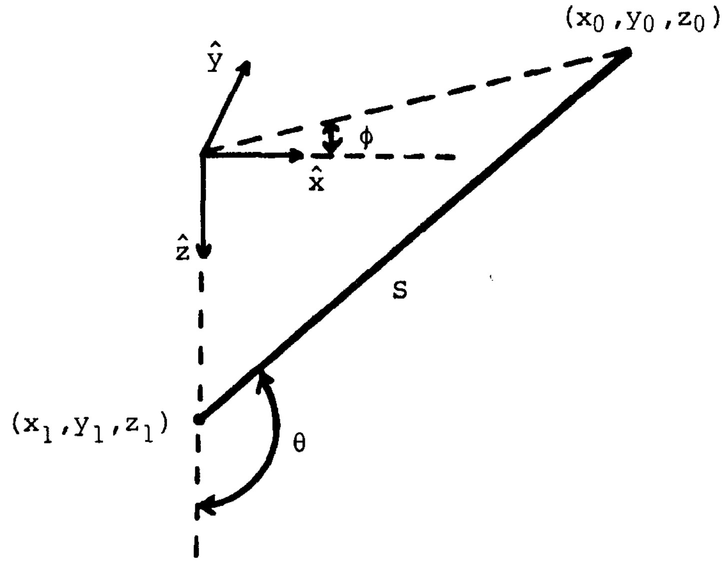

FIG. 1. Ray path geometry for a half-space with a source at (x_1, y_1, z_1), take-off angle θ, azimuth ϕ, and a receiver at (x_0, y_0, z_0).

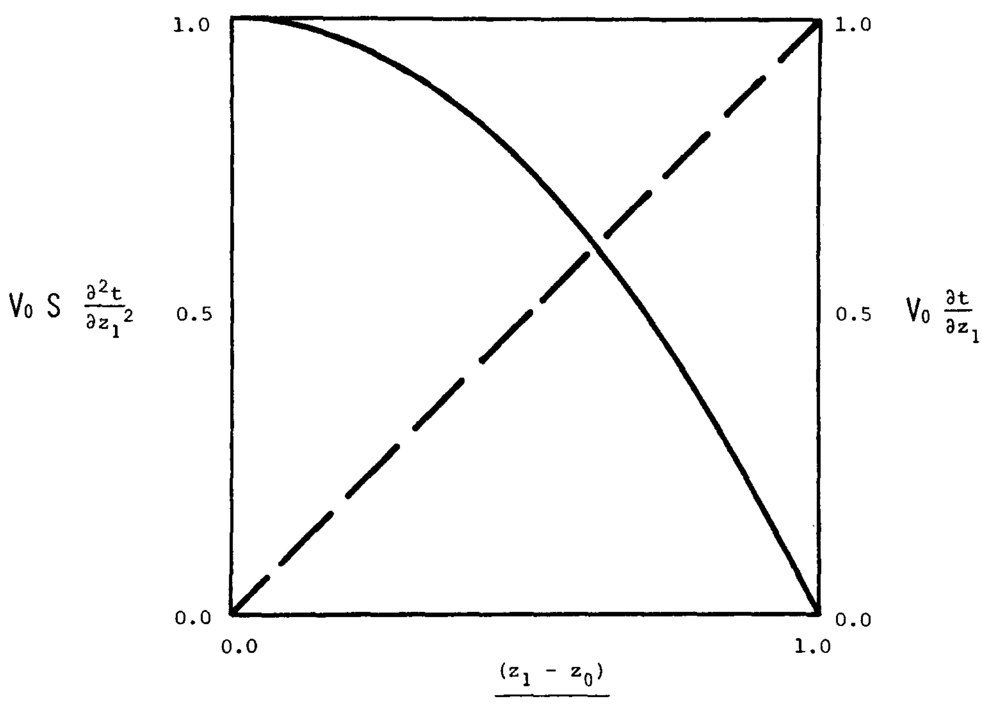

FIG. 2. Comparison of the nondimensionalized magnitudes of the first (dashed curve) and second (solid curve) order partial derivative terms for z_1 (source depth) as a function of the size of the ray vector component in the z direction.

with similar terms for x and y. It is immediately clear that the linear terms will not always dominate. Specifically, the first-order partial $\partial t/\partial z_1$ vanishes while the second-order partial $\partial^2 t/\partial z_1^2$ is maximized when $z_1 = z_0$. In fact the second-order term can be dominant over quite a range, especially for short path lengths (see Figure 2).

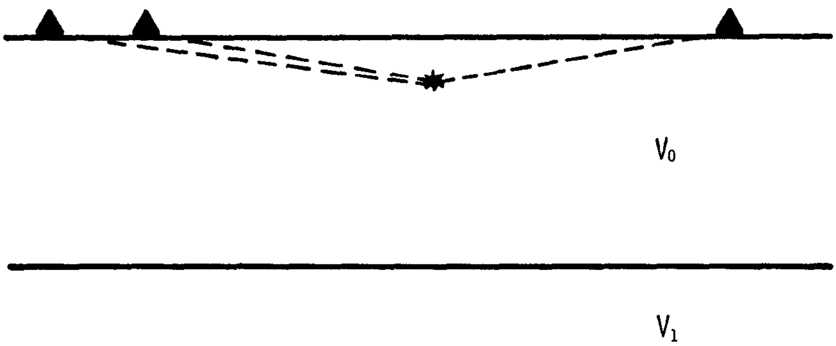

FIG. 3. Situation commonly leading to an "airquake," with a very shallow source recorded only at nearby stations.

This half-space example is relevant to the "airquake" problem—the tendency for very shallow earthquakes to have unstable depths in standard location algorithms if no station is close to the epicenter. If the event is small, only nearby stations will record the P-wave arrival time, and all the corresponding ray paths may be contained within the top layer of the crustal model (see Figure 3). In this case, $z_1 - z_0$ is nearly zero for all stations, so the derivatives $\partial t/\partial z_1$ will be very small positive numbers. If the residuals are all or mostly negative (e.g., if the top layer velocity is too high), then a standard location algorithm will perturb the source depth upwards a substantial distance, since the corresponding derivative is quite small, and an airquake results. Inclusion of second-order terms would prevent this instability, as we will show in a later example.

TRAVEL-TIME DERIVATIVES IN A LAYERED MODEL

Considering now a crustal model consisting of a layer of velocity V_0 over a half-space of velocity V_1, we must examine the cases of both direct and refracted rays (Figure 4, a and b). The refracted case can again be treated analytically. Given a source at $(x_1, y_1\, z_1)$ and a receiver at $(x_0, y_0, 0)$, the travel time and first-order partial derivatives are expressed as

$$t = \frac{\Delta}{V_1} + \frac{\sqrt{(V_1)^2 - (V_0)^2}}{V_0 V_1} \quad (2D_0 - z_1)$$

$$\frac{\partial t}{\partial x_1} = \frac{(x_1 - x_0)}{V_1 \Delta}$$

$$\frac{\partial t}{\partial y_1} = \frac{(y_1 - y_0)}{V_1 \Delta}$$

$$\frac{\partial t}{\partial z_1} = -\frac{\sqrt{(V_1)^2 - (V_0)^2}}{V_0 V_1} \tag{14}$$

where Δ is the epicentral distance, and D_0 is the layer thickness (Figure 4a). As before, the derivatives are proportional to the corresponding components of the ray vector at the source. Since the z_1 derivative is a constant, all second-order partials involving z_1 vanish. Those remaining are given by

$$\frac{\partial^2 t}{\partial x_1{}^2} = \frac{1}{V_1 \Delta}\left[1 - \frac{(x_1 - x_0)^2}{\Delta^2}\right]$$

$$\frac{\partial^2 t}{\partial y_1{}^2} = \frac{1}{V_1 \Delta}\left[1 - \frac{(y_1 - y_0)^2}{\Delta^2}\right]$$

$$\frac{\partial^2 t}{\partial x_1 \partial y_1} = \frac{\partial^2 t}{\partial y_1 \partial x_1} = \frac{-(x_1 - x_0)(y_1 - y_0)}{V_1 \Delta^3}. \tag{15}$$

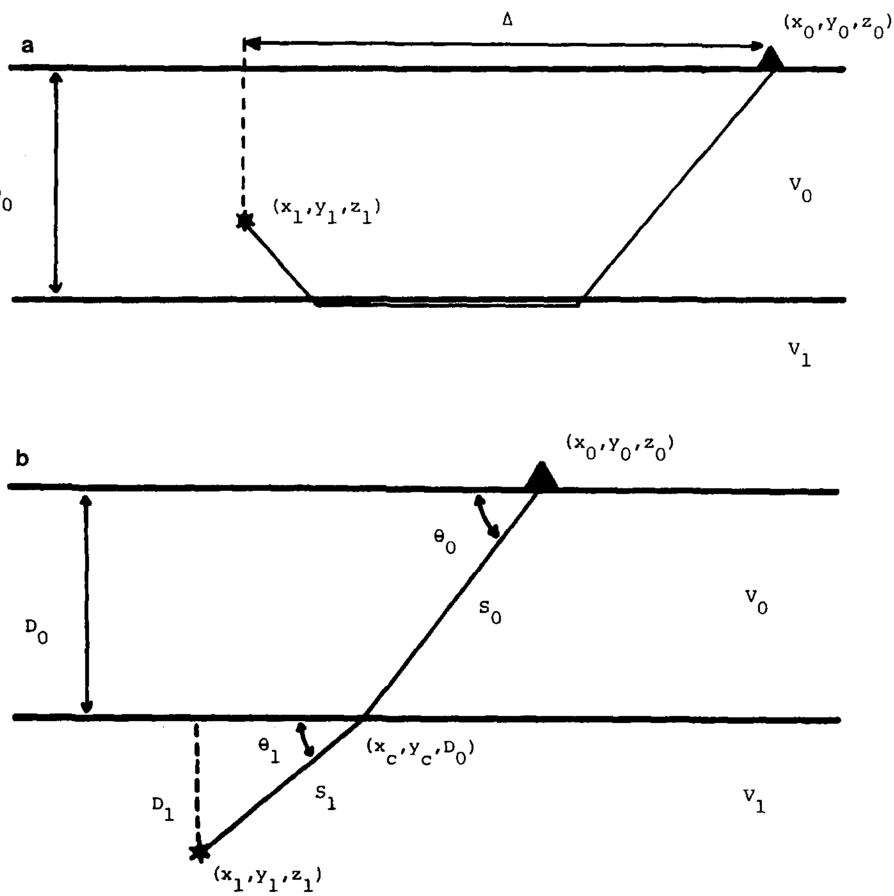

FIG. 4. Ray path geometry for a layer over a half-space. (a) Critically refracted arrival. (b) Direct arrival.

Equations (15) are quite similar to those for the half-space example above [equation (13)]. Again the second-order partials $\partial^2 t/\partial x_1{}^2$ and $\partial^2 t/\partial y_1{}^2$ are maximized when the corresponding first-order derivatives vanish. We also note these expressions can be generalized easily to multi-layer models.

Turning to the direct ray case (Figure 4b), we are faced with the absence of a direct equation for travel time in terms of the layer thickness, velocities, and endpoint coordinates only. However, we can write an implicit equation for travel time.

$$t = \frac{1}{V_1}\left[(x_1 - x_c)^2 + D_1{}^2\right]^{1/2} + \frac{1}{V_0}\left[(x_c - x_0)^2 + D_0{}^2\right]^{1/2} \tag{16}$$

where x_c is the x coordinate at which the ray crosses the layer boundary, determined generally by some iterative scheme, D_0 is the layer thickness, $D_1 = z_1 - D_0$, and assuming $y_1 = y_c = y_0$. The first-order partial derivatives are given by

$$\begin{aligned} \frac{\partial t}{\partial x_1} &= \frac{1}{V_1} \frac{(x_1 - x_c)}{(x_1 - x_c)^2 + D_1^2} \\ \frac{\partial t}{\partial z_1} &= \frac{1}{V_1} \frac{D_1}{(x_1 - x_c)^2 + D_1^2} \end{aligned} \tag{17}$$

(again we note, proportional to the ray vector components at the source). The second-order derivatives obviously require the evaluation of $\partial x_c/\partial x_1$ and $\partial x_c/\partial z_1$. The calculations are simplified by rewriting the first equation in (17), using Snell's

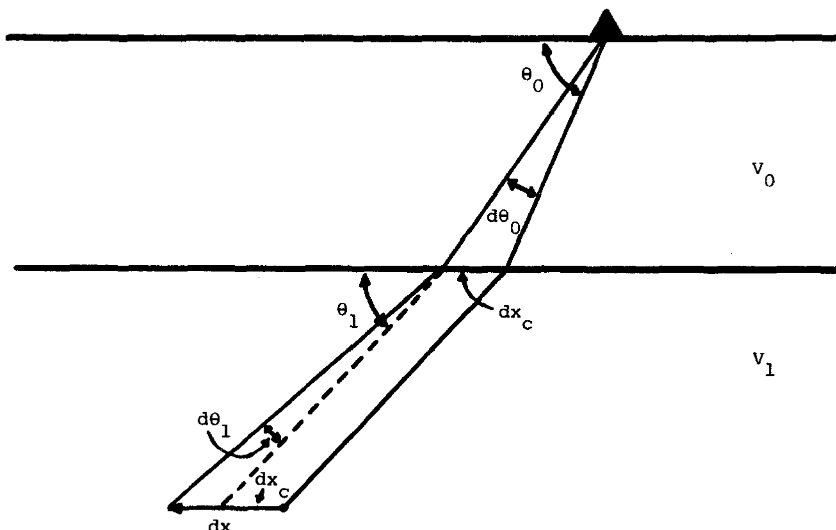

FIG. 5. Diagram for deriving the second-order derivatives for the direct ray case.

law, in terms of the fixed quantities x_0 and D_0 so that x_c is the only variable

$$\frac{\partial t}{\partial x_1} = \frac{1}{V_1} \frac{(x_c - x_0)}{(x_c - x_0)^2 + D_0^2}. \tag{18}$$

It is then straightforward (although tedious) to show

$$\begin{aligned} \frac{\partial^2 t}{\partial x_1^2} &= \frac{1}{V_0 S_0} \sin^2 \theta_0 \frac{\partial x_c}{\partial x_1} \\ \frac{\partial^2 t}{\partial z_1^2} &= \frac{1}{V_0 S_0} \cos \theta_0 \left(\cos \theta_1 - \sin \theta_1 \frac{\partial x_c}{\partial z_1} \right) \\ \frac{\partial^2 t}{\partial x_1 \partial z_1} &= \frac{1}{V_0 S_0} \sin \theta_1 \cos \theta_0 \frac{\partial x_c}{\partial x_1} \end{aligned} \tag{19}$$

where S_0 is the path length in the layer, θ_0 is the emergence angle at the surface, and θ_1 is the emergence angle at the layer boundary (Figure 4b).

Evaluation of $\partial x_c/\partial x_1$ and $\partial x_c/\partial z_1$ can be accomplished through simple geometrical arguments. From Figure 5, we see that

$$dx_1 = dx_c + \frac{S_1}{\sin\,\theta_1}\,d\theta_1$$

$$dx_c = \frac{S_0}{\sin\,\theta_0}\,d\theta_0. \tag{20}$$

Since $d\theta_1$ and $d\theta_0$ are infinitesimal, it can also be shown that

$$\frac{d\theta_0}{d\theta_1} = \frac{V_0 \sin\,\theta_1}{V_1 sin\,\theta_0}. \tag{21}$$

Combining (20) and (21), we arrive at

$$\frac{\partial x_c}{\partial x_1} = \frac{V_0 S_0 \sin^2\,\theta_1}{V_0 S_0 \sin^2\,\theta_1 + V_1 S_1 \sin^2\,\theta_1} \tag{22}$$

and finally

$$\frac{\partial^2 t}{\partial {x_1}^2} = \left(\frac{V_0 S_0}{\sin^2\,\theta_0} + \frac{V_1 S_1}{\sin^2\,\theta_1}\right)^{-1}$$

$$\frac{\partial^2 \mathrm{t}}{\partial \mathrm{x}_1 \partial \mathrm{z}_1} = \cot\,\theta_1 \left(\frac{V_0 S_0}{\sin^2\,\theta_0} + \frac{\mathrm{V}_1 \mathrm{S}_1}{\sin^2\,\theta_1}\right)^{-1}. \tag{23}$$

By a similar procedure, we arrive at

$$\frac{\partial^2 t}{\partial {z_1}^2} = \cot^2\,\theta_1 \left(\frac{V_0 S_0}{\sin^2\,\theta_0} + \frac{V_1 S_1}{\sin^2\,\theta_1}\right)^{-1}. \tag{24}$$

The expressions in (23) and (24) can be evaluated quite easily given the initial ray path, and the generalization to three dimensions and extension to a multi-layer model is straightforward. Equations (23) and (24) correctly reduce to the corresponding half-space values if we set $V_0 = V_1$ and $\theta_0 = \theta_1$.

Examples

In addition to the "airquake" problem, another situation in which the nonlinear terms in Newton's method should add stability is for an event outside an array (Figure 6). Instabilities of hypocenter depth in both of these cases result from nearly

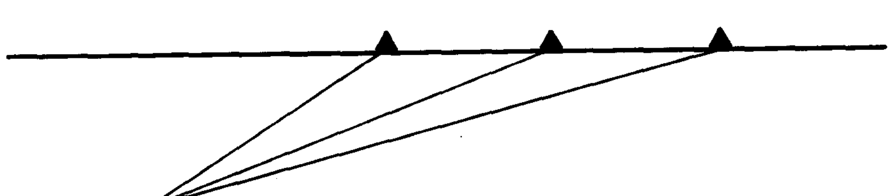

FIG. 6. Additional case in which Newton's method should lend stability to a location algorithm: shallow event outside an array.

parallel columns in the A matrix (origin time and z derivatives). From (13), we see that the relevant second-order derivatives are in fact more sensitive to differences in $(z_1 - z_0)$ than are the first derivatives, so that the former will reduce the degree of parallelism, and thus enhance stability. The second-order term remains nonzero when the event reaches $z_1 = 0$.

The ficticious array used in the calculations is diagrammed in Figure 7. For the two examples, a shallow event inside the array and an event outside the array, tests of Geiger's method (undamped and damped) versus the nonlinear method were performed using error-free data, in each case starting both with the true hypocenter and a displaced hypocenter. A half-space crustal model was used, and the velocity was perturbed from its true value to provoke instability. The test results comparing the undamped Geiger and nonlinear methods are summarized in Tables 1 and 2.

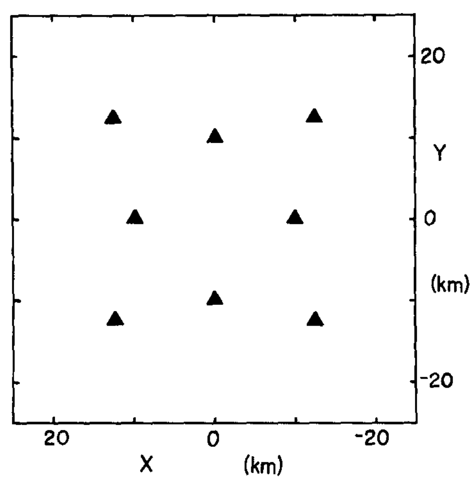

FIG. 7. Map of station network used in example calculations. All stations assigned $z_0 = 0$.

Note that because the model velocity was perturbed, the "best" solution will always differ from the "true" solution.

A clear pattern emerges from all the test cases. For Geiger's method (undamped), convergence is not achieved. From iteration to iteration the hypocenter adjustments are oscillatory, frequently leading to instability and divergence. In sharp contrast, the nonlinear method proves to be remarkably stable. The rms and average residuals are seen to decrease almost monotonically, and the global rms minimum is reached rapidly in each instance.

As an additional experiment, Geiger's method with damping was tested for the two cases, starting with the true initial location. The results are naturally sensitive to the choice of the damping value, which was set successively at 0.005, 0.001, and 0.0005. Two types of behavior resulted: oscillatory and asymptotic. Table 3 summarizes the case by case behavior. In the oscillatory cases, the hypocenter adjustments continually bypassed the global minimum for one value of damping, giving no hint of convergence after 10 iterations (damping = 0.0005), but for another converging fairly quickly (damping = 0.001). The asymptotic cases (damping = 0.005) are mostly reminiscent of Zeno's paradox—the hypocenter adjustment would take it half-way (more or less) to the global minimum on each step, but never really reach it. With additional trials, it is likely that a damping parameter value could be found which would lead to fast but nonoscillatory convergence for these examples.

DISCUSSION

The examples above serve to illustrate the usefulness of the nonlinear method. It is quite stable in cases for which Geiger's method fails without damping, and it

TABLE 1

EXAMPLE 1—SHALLOW EVENT INSIDE ARRAY AT (0, 0, 1)*

Iteration	Initial Hypocenter Error (km)	Initial Origin Time Error (sec)	rms Residual (sec)	Average Residual (sec)	Hypocenter Displacement (km)	Origin Time Change (sec)
(a) True Initial Location						
			Geiger's Method			
1	0.00	0.00	0.11	0.10	8.84	0.21
2	8.84	0.21	0.45	−0.43	6.44	0.07
3	2.40	0.28	0.18	−0.18	6.72	−0.07
4	4.32	0.21	0.27	−0.26	4.89	0.02
5	0.57	0.23	0.13	−0.12	19.46	−0.02
			Nonlinear Method			
1	0.00	0.00	0.11	0.10	0.15	0.10
2	0.15	0.10	0.04	0.00	1.01	0.01
3	0.86	0.12	0.04	−0.01	0.09	−0.01
4	0.96	0.11	0.04	0.00	0.04	0.00
5	1.00	0.11	0.03†	0.00	0.00	0.00
(b) Displaced Initial Location (2, 4, 2)						
			Geiger's Method			
1	4.58	0.20	0.54	0.23	4.79	0.38
2	0.22	0.18	0.09	−0.08	6.97	0.09
3	2.53	0.30	0.20	−0.20	6.28	−0.09
4	3.75	0.21	0.24	−0.23	4.66	0.02
5	0.92	0.22	0.12	−0.11	99.18	−0.02
			Nonlinear Method			
1	4.58	0.20	0.54	0.23	6.12	0.36
2	2.90	0.16	0.22	−0.12	1.74	−0.08
3	1.63	0.08	0.05	−0.01	1.28	0.03
4	0.35	0.11	0.04	0.00	1.10	−0.01
5	‡					

* True velocity = 6.0 km/sec; perturbed velocity = 6.25 km/sec.

† Global rms minimum reached.

‡ Follows convergence path similar to nonlinear case in (a).

yields rapid convergence without the need for trial-and-error experimenting with damping. Further study is required to analyze more general test cases, including layered earth models.

Other related issues deserve investigation. The matrices of resolution $\boldsymbol{R}$ and covariance $\boldsymbol{C}$ (e.g., Jackson, 1972) are modified from their usual form in Newton's method. Whereas for least squares $\boldsymbol{R} = \boldsymbol{I}$, the identity matrix, we have instead

$$\boldsymbol{R} = [\boldsymbol{A}^T\boldsymbol{A} - (\nabla \boldsymbol{A}^T)\boldsymbol{r}]^{-1}\boldsymbol{A}^T\boldsymbol{A}. \tag{25}$$

TABLE 2

EXAMPLE 2—EVENT OUTSIDE ARRAY AT (20, 0, 5)*

Iteration	Initial Hypocenter Error (km)	Initial Origin Time Error (sec)	rms Residual (sec)	Average Residual (sec)	Hypocenter Displacement (km)	Origin Time Change (sec)
(a) True Initial Location						
			Geiger's Method			
1	0.00	0.00	0.34	0.31	6.79	0.50
2	6.89	0.50	0.19	−0.17	12.01	−0.01
3	5.28	0.49	0.55	−0.51	9.91	0.20
4	4.65	0.69	0.30	−0.28	51.30	−0.21
5	55.94	0.47	5.25	−5.16	87.50	6.85
			Nonlinear Method			
1	0.00	0.00	0.34	0.31	2.92	−0.04
2	2.92	−0.04	0.12	−0.01	6.08	0.54
3	4.20	0.50	0.13	−0.09	0.89	−0.19
4	4.93	0.31	0.08	0.00	0.62	−0.03
5	5.03	0.28	0.07†	0.00	0.07	0.00
(b) Displaced Initial Location (15, 0, 5)						
			Geiger's Method			
1	5.00	1.00	0.36	−0.15	5.93	−0.54
2	2.68	0.46	0.14	−0.11	9.40	0.02
3	12.08	0.48	0.36	−0.32	7.77	0.08
4	4.32	0.56	0.22	−0.20	27.11	−0.08
5	31.42	0.48	2.10	−2.02	31.61	1.73
			Nonlinear Method			
1	5.00	1.00	0.36	−0.15	3.50	−0.45
2	1.69	0.55	0.16	−0.03	2.46	−0.20
3	2.14	0.35	0.09	−0.02	1.52	−0.06
4	3.65	0.29	0.08	−0.01	1.05	−0.00
5	‡					

* True velocity = 6.0 km/sec; perturbed velocity = 6.5 km/sec.

† Global rms minimum reached.

‡ Follows convergence path similar to nonlinear case in (a).

TABLE 3

SUMMARY OF DAMPED GEIGER'S METHOD TEST (PERTURBED VELOCITY, TRUE INITIAL LOCATION)

Damping Parameter	Example 1	Example 2
0.005	Asymptotic (slow)	Asymptotic (slow)
0.001	Oscillatory (convergent)	Oscillatory (convergent)
0.0005	Oscillatory	Oscillatory

In the case of uncorrelated errors, $\boldsymbol{C} = \sigma^2(\boldsymbol{A}^T\boldsymbol{A})^{-1}$ for least squares, but

$$\boldsymbol{C} = \sigma^2[\boldsymbol{A}^T\boldsymbol{A} - (\nabla \boldsymbol{A}^T)\boldsymbol{r}]^{-1}\boldsymbol{R} \tag{26}$$

for Newton's method. We expect that more realistic estimates of $\boldsymbol{R}$ and $\boldsymbol{C}$ will be obtained using (25) and (26).

Finally, we point out that estimates of geometric spreading for the direct ray case can be obtained from (23) and (24). These second-order derivatives are related to the wave-front curvature (Červeny and Psencik, 1984) and as a result could be used for the computation of amplitudes for direct rays.

Conclusions

For common earth models used in local earthquake location (half-space and layered), Geiger's method can be extended in a straightforward manner to incorporate the nonlinear behavior of travel time as a function of source position by way of Newton's method. Simple expressions for the required second-order partial derivatives have been obtained for the half-space and layer-over-half-space cases. As illustrated in a set of examples, employment of Newton's method for local earthquake location promises to improve the stability of the location algorithm. Additional benefits include a means for the direct evaluation of geometric spreading and possibly improved estimates of solution quality.

Acknowledgments

I thank Willie Lee and Jay Pulli for providing some of the motivation for this work, Rob Comer for pointing out an error in an earlier version of this manuscript, two anonymous reviewers for constructive suggestions, and Ann Lattimore for preparation of this manuscript. This material is based on work supported by NSF Grant EAR 82-06266.

References

Anderson, K. R. (1982). Robust earthquake location using *M*-estimates, *Phys. Earth Planet. Interiors* **30,** 119–130.

Bolt, B. A. (1970). Earthquake location for small networks using the generalized inverse matrix, *Bull. Seism. Soc. Am.* **60,** 1823–1828.

Buland, R. (1976). The mechanics of locating earthquakes, *Bull. Seism. Soc. Am.* **66,** 173–187.

Červený, V. and I. Psencik (1983). Gaussian beams and paraxial ray approximation in three-dimensional elastic inhomogeneous media, *J. Geophys.* **53,** 1–15.

Geiger, L. (1910). Herbsetimmung bei Erdbeben aus den Ankunfzeiten, *K. Gessell. Wiss. Goett.* **4,** 331–349.

Geiger, L. (1912). Probability method for the determination of earthquake epicenters from the arrival time only, *Bull. St. Louis Univ.* **8,** 60–71.

Herrmann, R. B. (1979). FASTHYPO—A hypocenter location program, *Earthquake* Notes **50,** 25–37.

IBM (1965). Large-aperture seismic array signal processing study. Final Report prepared for the Advanced Research Projects Agency, Contract SD-296.

Jackson, D. D. (1972). Interpretation of inaccurate, insufficient and inconsistent data. *Geophys. J. Roy. Astr. Soc.* **28,** 97–109.

Klein, F. W. (1978). Hypocenter location program HYPOINVERSE, Part 1: user's guide to versions 1, 2, 3 and 4, *U.S. Geol. Surv. Open-File Rept. 78-694.*

Lee, W. H. K. and S. W. Stewart (1981). Principles and applications of microearthquake networks, *Adv. Geophys. Suppl. 2.*

Lienert, B. R. and L. N. Frazer (1983). An improved earthquake location algorithm, *EOS, Trans. Am. Geophys. Union* **64,** 267.

Mellman, G. R. (1980). A method of body-wave waveform inversion for the determination of earth structure, *Geophys. J. Roy. Astr. Soc.* **62,** 481–504.

Smith, E. G. C. (1976). Scaling the equations of condition to improve conditioning, *Bull. Seism. Soc. Am.* **66,** 2075–2076.

Tarantola, A. and B. Valette (1982). Inverse problems = quest for information, *J. Geophys.* **50,** 159–170.

Department of Earth and Space Sciences
State University of New York
Stony Brook, New York 11794

Manuscript received 2 April 1984