Nonlinear Analysis of the Gunn Effect

B. W. KNIGHT* AND G. A. PETERSON
United Aircraft Research Laboratories, East Hartford, Connecticut
(Received 3 February 1966)

A model is analyzed for the growth and propagation of electrostatic waves in a medium with an arbitrary law of mobility. The equation of motion has been solved by the method of characteristics. Shock-wave criteria are developed. Special emphasis is given to the anomalous case where the average carrier velocity can decrease with field, which, as pointed out by Kroemer, is pertinent to the Gunn effect.

INTRODUCTION

RECENTLY, Gunn[1] observed an electronic front propagating in gallium arsenide. Pritor to this, Ridley[2] had shown that a bulk negative resistance leads to the formation of domains which move with the carriers. The intervalley transfer mechanism of Ridley and Watkins[3] and Hilsum[4] provided the microscopic basis of such an anomalous resistance. Kroemer[5] suggested that these considerations of electrostatic instabilities were pertinent to the Gunn effect. These ideas were subjected to numerical analysis and favorable agreement found with observation.[5,6]

Our intention here is to formulate and analyze a model for the evolution of an electrostatic pulse in any medium where the average drift velocity is field-dependent. In the fast relaxation limit of momentum and between species (relative to dielectric relaxation) our model is a precise description except for the effects of diffusion. The influence of external circuitry is manifested through the current delivered to the ends of the diode. An equation for the electrostatic field has been derived from Poisson's equation and a law for the current. We have solved this nonlinear partial differential equation of motion exactly by utilizing the method of characteristics. From the solution a criterion is obtained for shock front formation.

FORMULATION

The fundamental equation of this theory,

$$\mathcal{E}_t(x,t) = -j(x,t) + f(t)\,, \qquad (1)$$

states that there are two sources of the displacement current: (a) that in the motion of the mobile charges themselves and (b) that due to the current $f(t)$, supplied to the ends of the diode by the external circuit. Subscripts denote partial differentiation. This equation is, of course, a direct consequence of current conservation and Poisson's equation, albeit in dimensionless form.

To derive Eq. (1) we note that the time rate of change of the total mobile charge density, $N(\mathrm{X,T})$, obeys the continuity relation

$$N_{\mathrm{T}}(\mathrm{X,T}) = -J_{\mathrm{X}}(\mathrm{X,T})\,, \qquad (2)$$

where J is the carrier flux, while Poisson's equation requires that

$$E_{\mathrm{X}}(\mathrm{X,T}) = 4\pi e/\kappa(N(\mathrm{X,T}) - N_0)\,, \qquad (3)$$

where κ is the dielectric constant and N_0 the fixed, uniform, neutralizing background charge. (We use capital letters to denote variables in dimensional form.) Differentiating Eq. (3) with respect to time and substituting from (2), we find

$$E_{\mathrm{XT}} = (4\pi e/\kappa) N_{\mathrm{T}} = -(4\pi e/\kappa) J_{\mathrm{X}}$$

which admits a first integral,

$$E_{\mathrm{T}}(\mathrm{X,T}) = 4\pi e/\kappa(-J(\mathrm{X,T}) + F(\mathrm{T}))\,. \qquad (4)$$

$F(\mathrm{T})$, the function of integration, is the average of the carrier fluxes across the ends of the sample. There are two limiting cases which provide insight into Eq. (4):

Case (a) $J(\mathrm{X,T}) = 0$, the extreme case of an insulator; then the diode behaves as a capacitor and $sF(\mathrm{T})$ represents the time rate of change of the charge density supplied to the capacitor plates by the external current such that

$$E_{\mathrm{T}} = 4\pi e/\kappa F(\mathrm{T})\,;$$

Case (b) uniform current leakage $J(\mathrm{X,T}) = F(\mathrm{T})$. Then there is no charge accumulation and $E_{\mathrm{T}} = 0$. Thus under conditions of charge neutrality the product of $eF(\mathrm{T})$ and the area is the current that follows in the external circuit.

Units are now chosen such that charge density is relative to N_0, i.e., $n(\mathrm{x,t}) \equiv N(\mathrm{X,T})/N_0$. Field strengths are in terms of a characteristic intensity, E_α,

$$\mathcal{E} \equiv E/E_\alpha\,.$$

Time is measured on a scale of the dielectric relaxation time

$$\tau = ((4\pi e/\kappa)\mu(E_\alpha)N_0)^{-1}\,,$$

---

* Permanent addresses: The Rockefeller University, New York, N. Y., and Cornell University Graduate School of Medical Sciences, New York, N. Y.

[1] J. B. Gunn, in *Proceedings of the Symposium on Plasma Effects in Solids, Paris, 1964*, (Dunod Cie., Paris, 1965), p. 199.

[2] B. K. Ridley, Proc. Phys. Soc. (London) **82**, 954 (1963).

[3] B. K. Ridley and T. B. Watkins, Proc. Phys. Soc. (London), **78**, 293 (1961).

[4] C. Hilsum, Proc. IRE **50**, 185 (1962).

[5] H. Kroemer, Proc. IEEE **52**, 1736 (1964).

[6] D. E. McCumber and A. G. Chynoweth, IEEE Trans. Electron Devices **ED1 3**, 4 (1966).

such that $t\equiv \mathrm{T}/\tau$; $\mu(E_\alpha)$ is the mobility for the characteristic field E_α. Finally, the natural length is the drift distance in time τ of a particle with velocity $V(E_\alpha)$. Italic x is the dimensionless spatial coordinate. Scaling Eq. (4) in these units, Eq. (1) is derived; furthermore, Eq. (6) below is a dimensionless form of Eq. (3).

Features specific to the model appear when in turn we write

$$j(x,t)=n(x,t)v(\mathcal{E}), \tag{5}$$

where $v(\mathcal{E})$ is the average field-dependent drift velocity. v is assumed to be an instantaneous and local function of $\mathcal{E}$ alone. From Poisson's equation we further relate the space charge to the field gradient,

$$n(x,t)=\mathcal{E}_x(x,t)+1. \tag{6}$$

Substituting into Eq. (1) the expression for the current (5), we derive the equation of motion of the space-charge wave,

$$\mathcal{E}_t+v(\mathcal{E})\mathcal{E}_x=-v(\mathcal{E})+f(t). \tag{7}$$

It is this nonlinear partial differential equation which is central to all further discussion.

$v(\mathcal{E})$ will simply be taken to be a known functional form. For example, with a two-valley mechanism operative,

$$v(\mathcal{E})=\mathcal{E}[(n_1+\alpha n_2)/n],$$

and we can expect behavior as depicted in Fig. 1, since the fraction of carriers in the lower valley depletes with increasing $\mathcal{E}$. Here, α is the ratio of the two mobilities.

METHOD OF SOLUTION

The solution of Eq. (7) may be reduced to straightforward procedures, and the process of reduction lends some insight into the physical consequences of the equation. The general approach is by the method of characteristics.[7]

If a line $x=X(t)$ is chosen on the (x,t) plane, then along that line

$$d\mathcal{E}/dt=\mathcal{E}_t+\mathcal{E}_x dX/dt. \tag{8}$$

If $\mathcal{E}$ satisfies Eq. (7), then demanding that

$$dX/dt=v(\mathcal{E}) \tag{9}$$

gives, by use of (7) and (8), the ordinary differential equation

$$d\mathcal{E}/dt=-v(\mathcal{E})+f(t), \tag{10}$$

which is the equation that $\mathcal{E}$ satisfies along a "characteristic line" specified by (9). Notice that once $\mathcal{E}(t)$ has been obtained from Eq. (10), Eq. (9) may be used to determine the path of the characteristic line. If the solution of (10) initiates at $\mathcal{E}=\mathcal{E}_0$ when $t=0$, then the solution

$$\mathcal{E}=\mathcal{E}(t,\mathcal{E}_0) \tag{11}$$

may be re-expressed implicitly as

$$\mathcal{E}_0=\phi(\mathcal{E},t) \tag{12}$$

which will be the convenient form for subsequent use. Since $\mathcal{E}=\mathcal{E}_0$ when $t=0$,

$$\phi(\mathcal{E},0)=\mathcal{E}. \tag{13}$$

Several interesting features of the solution of (7) may be anticipated by considering the case of fixed external current ($f=$ const.). Equation (10) for $\mathcal{E}$ along a characteristic line then becomes

$$d\mathcal{E}/dt=-v(\mathcal{E})+f_0,$$

and the tendency of $\mathcal{E}$ will be determined by the signature of the right-hand side. The result will evidently be as shown in Fig. 1:

Thus on a characteristic line $\mathcal{E}$ will tend toward a value where

$$v(\mathcal{E})=f_0,$$

and where $v(\mathcal{E})$ has positive slope. Notice at the three intersection points Eq. (10) integrates trivially to

$$\mathcal{E}=\text{constant},$$

whereupon Eq. (9) integrates to

$$X(t)=X_0+f_0t.$$

If for example the initial condition $\mathcal{E}_0(x)$ is as shown in Fig. 2, then the values $\mathcal{E}_1$ and $\mathcal{E}_2$ will have loci on the (x,t) plane as shown in Fig. 3. Also shown in Fig. 3 are the curved characteristic lines originating from other initial values of x. The tendencies of these characteristic curves may be straightforwardly deduced by working between Fig. 1 and Eq. (9). In the range of $\mathcal{E}<\mathcal{E}_\alpha$, that is for characteristics not originating between x_α and x_α',

$$dX/dt\rightarrow f_0,$$

but in the inner (or "anomalous") region *the characteristic slopes diverge from that* of $X(t)=X_0+f_0t$.

It is further to be noted that the x_1' and x_2 characteristic lines have the interesting property that their *near neighbors converge toward them.* In the case of x_1' a collision seems possible; in the case of x_2 it seems inevitable. Where two characteristics touch there will be a discontinuity in $\mathcal{E}(x,t)$. This is the condition for

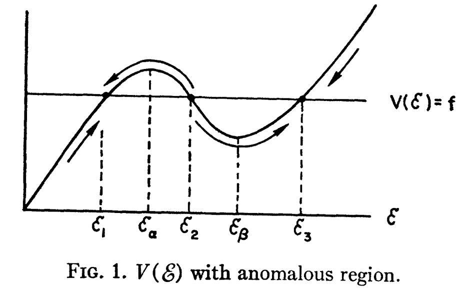

FIG. 1. $V(\mathcal{E})$ with anomalous region.

[7] F. B. Hildebrand, *Advanced Calculus for Engineers* (Prentice-Hall Publishers, Inc., New York, 1949), p. 368.

the formation of a shock front. Equation (7) will now be solved, and the criteria for shock formation will be explored quantitatively.

In the special case of $\mathcal{E}$ space-independent, Eq. (7) reduces to the ordinary differential Eq. (10) whose solution is (11). A second evident solution to (7) is

$$\mathcal{E}=-x+\int_0^t f(t')dt'. \tag{14}$$

From these two particular solutions a solution satisfying any specified initial conditions may be constructed in the following manner: A solution of (7) may be written implicitly as

$$\Psi(x,t,\mathcal{E})=0.$$

Using the implicit differentiation identities

$$\mathcal{E}_x=-(\Psi_x/\Psi_{\mathcal{E}}),\quad \mathcal{E}_t=-(\Psi_t/\Psi_{\mathcal{E}}),$$

Eq. (7) becomes

$$\Psi_t+v(\mathcal{E})\Psi_x+(-v(\mathcal{E})+f(t))\Psi_{\mathcal{E}}=0. \tag{15}$$

This equation is not only linear; it has the property

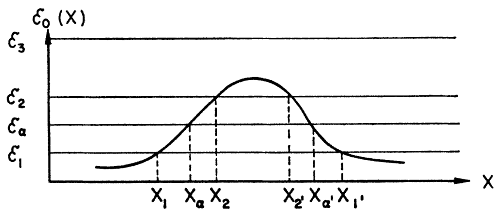

FIG. 2. Typical initial pulse $\mathcal{E}_0(x)$.

that if Ψ_1 and Ψ_2 are both solutions, then

$$\Psi=G(\Psi_1,\Psi_2)$$

is also a solution for any differentiable function G. This property is easily checked by direct substitition back into Eq. (15).

From Eq. (12)

$$\Psi_1=\phi(\mathcal{E},t)$$

is a solution; from Eq. (14)

$$\Psi_2=\mathcal{E}+x-\int_0^t f(t')dt'$$

is also a solution. Suppose the initial value of $\mathcal{E}(x,t)$ is $\mathcal{E}_0(x)$. Now let

$$G(\Psi_1,\Psi_2)=\Psi_1-\mathcal{E}_0(\Psi_2-\Psi_1)$$

which is also a solution to Eq. (15). An implicit solution to Eq. (7) will thus be given by

$$G=0,$$

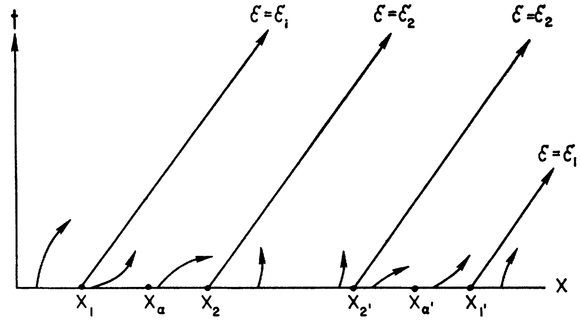

FIG. 3. Characteristic curves, f=constant.

or

$$\phi(\mathcal{E},t)=\mathcal{E}_0\left(x+\mathcal{E}-\int_0^t f(t')dt'-\phi(\mathcal{E},t)\right). \tag{16}$$

When $t=0$, by Eq. (13) the above expression becomes

$$\mathcal{E}=\mathcal{E}_0(x).$$

Thus Eq. (16) is a solution satisfying the initial conditions of the problem.

SHOCK-WAVE CRITERIA

In this section the solution is analyzed for the appearance of discontinuities. On physical grounds there are three regimes in which such shocks are anticipated:

(1) The anomalous region $v_{\mathcal{E}}<0$. The lower amplitudes in the wave are running faster and we look for undercutting on the trailing edge of a pulse as, for example, in Fig. 2.

(2) The Ohmic region $v_{\mathcal{E}}>0$. The situation is reversed and consequently cresting is a possibility on the front of the wave since higher fields are now moving with greater velocity.

(3) When $\mathcal{E}_0(x)$ contains $\mathcal{E}_2$. The wave form "tears"; i.e., on the high-field side it grows towards $\mathcal{E}_3$ while below it relaxes back to $\mathcal{E}_1$.

To derive the criterion both sides of Eq. (16) are differentiated with respect to x,

$$\phi_{\mathcal{E}}\mathcal{E}_x=(-\phi_{\mathcal{E}}\mathcal{E}_x+\mathcal{E}_x+1)\mathcal{E}_{0x},$$

or

$$\mathcal{E}_x=\frac{1}{\phi_{\mathcal{E}}(1+1/\mathcal{E}_{0x})-1}. \tag{17}$$

A singularity sets in whenever the denominator vanishes, i.e., when

$$\mathcal{E}_{0x}=\phi_{\mathcal{E}}(t)/[1-\phi_{\mathcal{E}}(t)]. \tag{18}$$

This is interpreted to mean that along a characteristic line a given initial slope $\mathcal{E}_{0x}$ will turn into a shock when $\phi_{\mathcal{E}}(t)$ has unfolded to the point of satisfying the condition expressed by (18). (Physically, $\phi_{\mathcal{E}}$ measures the degree of steepening of the field gradient: it is the ratio of the initial gradient $\delta\phi/\delta x$ to that at time t, $\delta\mathcal{E}/\delta x$.)

Proceeding further, an expression can be obtained for $\phi_{\mathcal{E}}$. Since $\mathcal{E}$ obeys Eq. (10) and is continuous in ϕ, differentiation of Eq. (10) by ϕ gives

$$d\mathcal{E}_{\phi}/dt=-v_{\mathcal{E}}\mathcal{E}_{\phi}.$$

Equation (13) furnishes the initial condition $\mathcal{E}_{\phi}(0)=1$. Solving directly for $\mathcal{E}_{\phi}$,

$$\mathcal{E}_{\phi}=\exp\left(-\int_0^t v_{\mathcal{E}}dt'\right),$$

whence

$$\phi_{\mathcal{E}}=\exp\left(\int_0^t v_{\mathcal{E}}dt'\right). \quad (19)$$

Specializing to $f(t)=f_0$, $\phi_{\mathcal{E}}$ can be expressed in closed form by explicitly carrying out the integration and we find

$$\phi_{\mathcal{E}}(t)=[-v(\phi)+f_0]/[-v(\mathcal{E})+f_0]. \quad (20)$$

Two cases exist depending on the initial value, ϕ:

(1) $\mathcal{E}_{\alpha}<\phi<\mathcal{E}_{\beta}$: the anomalous or "negative slope" region. From Fig. 1 it is observed that the denominator of (20) first increases and then decreases. Therefore $\phi_{\mathcal{E}}$ starts at unity, takes a minimal value and then increases without bound. Moreover the minimal value of $\phi_{\mathcal{E}}$ is

$$(\phi_{\mathcal{E}})_m=[-v(\phi)+f_0]/[-v_m+f_0],$$

where v_m is the extremum value,

$$v_m=v(\mathcal{E}_{\alpha}) \quad \text{or} \quad v(\mathcal{E}_{\beta}).$$

Referring to the criterion Eq. (18), since for t small $\phi_{\mathcal{E}}<1$, for all initial field gradients such that

$$\mathcal{E}_{0x}\geqslant(\phi_{\mathcal{E}})_m/[1-(\phi_{\mathcal{E}})_m]$$

a shock is inevitable. Furthermore since $\mathcal{E}_{0x}$ is positive, when this shock appears it is located on the back side of the wave.

(2) $\mathcal{E}<\mathcal{E}_{\alpha}$ or $\mathcal{E}>\mathcal{E}_{\beta}$: the "positive slope" region. $\phi_{\mathcal{E}}$ monotonically increases. The criterion Eq. (18) implies that in order for a shock to develop $\mathcal{E}_{0x}$ must be sufficiently negative:

$$\mathcal{E}_{0x}\leqslant-1.$$

Equality is the criterion at $t\rightarrow\infty$. Since $\mathcal{E}_{0x}$ is negative, the shock is on the foreward side of the wave.

Finally it must be remarked that for $\phi=\mathcal{E}_2$ any initial form eventually develops a discontinuity on the trailing edge no matter how gentle the initial slope. This is the tearing phenomenon mentioned earlier. It follows directly from Eq. (19) that since $\mathcal{E}_2$ is a point of equilibrium, $\mathcal{E}(t)=\mathcal{E}_2$, a constant. Consequently,

$$\phi_{\mathcal{E}}=\exp[v_{\mathcal{E}}(\mathcal{E}_2)t],$$

which vanishes at $t\rightarrow\infty$ since $v_{\mathcal{E}}(\mathcal{E}_2)$ is intrinsically negative. Thus by (18) *any* positive initial gradient starting at field $\mathcal{E}_2$ will eventually develop into a trailing-edge shock.

The behavior of the electric field along the corresponding $\mathcal{E}=\mathcal{E}_2$ characteristic line on the *front* side of the wave is radically different. This may be seen by solving Eq. (17) for the condition that

$$\mathcal{E}_x=-1.$$

The result is

$$\phi_{\mathcal{E}}(1+1/\mathcal{E}_{0x})=0,$$

whence either

$$\phi_{\mathcal{E}}=0$$

or

$$\mathcal{E}_{0x}=-1.$$

Thus if $\mathcal{E}_x$ does not start at -1, the only way for it to achieve that value is $\phi_{\mathcal{E}}\rightarrow 0$. This happens only on the characteristic line $\mathcal{E}=\mathcal{E}_2$. If the slope initially is never less than -1, this slope will never be exceeded during the development of the wave, but will be approached asymptotically along the $\mathcal{E}=\mathcal{E}_2$ characteristic line.

It should be pointed out here that Eq. (6) is physically reasonable only for

$$n(x,t)\geqslant 0, \quad (21)$$

or equivalently for

$$\mathcal{E}_x(x,t)\geqslant-1.$$

Equality implies complete carrier depletion. In the paragraph above we have shown that if this physical limitation is not exceeded initially in our model, then it will never be exceeded, but will be approached asymptotically on the leading edge $\mathcal{E}=\mathcal{E}_2$ characteristic line. Also it should be noted that the limitation (21) prevents the occurrence of shock-fronts in ordinary Ohmic materials.

FURTHER REMARKS

A relationship exists between the $f=$const solutions of Eq. (7) and the problem of Gunn oscillations. With f constant, any characteristic of field $\mathcal{E}_1$ or $\mathcal{E}_3$ will maintain that field, and propagate at velocity f. Any characteristic at another $\mathcal{E}$ value will approach either $\mathcal{E}_1$ or $\mathcal{E}_3$ asymptotically. To within the approximations of Eq. (7), a wave form such as that shown in Fig. 2 will tend to a final configuration as shown in Fig. 4.

Since

$$j=(\mathcal{E}_x+1)v(\mathcal{E})$$

and

$$v(\mathcal{E}_1)=v(\mathcal{E}_3)$$

the current j will assume a constant value except where $\mathcal{E}$ is changing. The current flowing into and out of the ends of the diode subsequently will be constant until the pulse arrives at the end. The condition of constant current is equivalent to $f=$const in the formulation of Eq. (7).

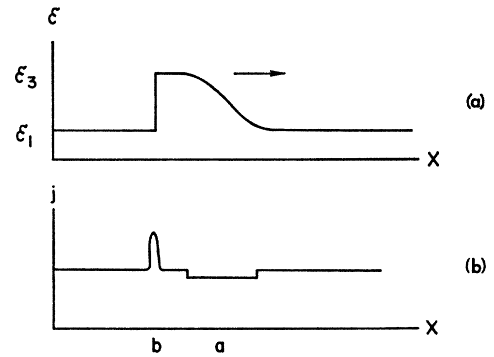

FIG. 4. (a) Steady propagating pulse; (b) associated current, j.

Once the pulse has arrived at the end of the Gunn diode, the subsequent history of $f(t)$ will be determined by external considerations (in particular the impedance characteristics of the voltage supply and the boundary condition at the input terminal of the diode) until the asymptotic pulse form is again established.

In regions of abruptly changing $\mathcal{E}$, as shown at a and b in Fig. 4, diffusion type processes may be expected to dominate, and the description furnished by Eq. (7) becomes inadequate. However, the general form of the charge and current profiles accompanying the pulse may be determined by using only conservation of charge. Given the traveling wave form

$$\mathcal{E}(x,t)=\mathcal{E}(x-ft)\,, \tag{22}$$

the charge density,

$$n(x,t)=\mathcal{E}_x(x,t)+1\,,$$

is similarly

$$n(x,t)=n(x-ft)\,,$$

whereupon the continuity equation

$$n_t+j_x=0\,,$$

gives

$$-fn_x+j_x=0\,,$$

or

$$\begin{aligned} j(x,t)&=fn(x,t)+\text{const}\\ &=f(\mathcal{E}_x(x,t)+1)+\text{const}\,, \end{aligned} \tag{23}$$

as shown in the figure. The situation is evidently one of two advancing walls of opposite charge. Thus asymptotically the system tends to the form of an advancing dipole layer.

It should be remarked that the streaming velocity of the pulse may in fact depart somewhat from the velocity f used in Eq. (22). This is because Eq. (10) permits $\mathcal{E}$ to change along a characteristic line. Assuming a streaming velocity c, and a streaming form

$$\mathcal{E}(x,t)=\mathcal{E}(x-ct)\,,$$

Eq. (7) becomes

$$(-c+v(\mathcal{E}))\mathcal{E}_x=-v(\mathcal{E})+f\,,$$

which has the immediate integral

$$x-x_1=-\int\frac{v(\mathcal{E})-c}{v(\mathcal{E})-f}d\mathcal{E}\,. \tag{24}$$

In the vicinity of $V(\mathcal{E}_i)=f$, the integral behaves as

$$x-x_1\simeq[(c-f)/v_{\mathcal{E}}]\ln|\mathcal{E}-\mathcal{E}_i|\,,$$

or

$$\mathcal{E}=\mathcal{E}_i\pm\exp[v_{\mathcal{E}}/(c-f)](x-x_1)\,, \tag{25}$$

in accordance with the previously known behavior in the vicinity of field strengths $\mathcal{E}_1$, $\mathcal{E}_2$, $\mathcal{E}_3$. In the vicinity of $V(\mathcal{E}_2)=c$ the integral behaves as

$$x-x_1\simeq-[v_{\mathcal{E}}/(c-f)]\tfrac{1}{2}(\mathcal{E}-\mathcal{E}_l)^2\,,$$

or

$$\mathcal{E}=\mathcal{E}_l\pm\{2[(c-f)/v_{\mathcal{E}}](x_1-x)\}^{1/2}\,. \tag{26}$$

As $x\to x_1$, $\mathcal{E}\to\mathcal{E}_l$, but the slope of $\mathcal{E}$ becomes infinite, and the solution Eq. (26) ceases to exist for $[(c-f)/v_{\mathcal{E}}]\times(x-x_1)>0$. Thus Eq. (26) furnishes the asymptotic condition for attachment between a slowly varying solution and a shock front. The details of attachment are beyond the scope of Eq. (7), and will be discussed in a subsequent paper.

ACKNOWLEDGMENT

One of the authors (GAP) is grateful for a discussion with E. C. Lary which sharpened his understanding.