J. A. Swanson

Diffusion Attenuation Part I

Abstract: Perturbation methods are applied to the problem of calculating the attenuation of signals consisting of compensated space charges moving in an electric field of general, but prescribed, form. Asymptotic formulas for attenuation and phase shift are derived which apply when the diffusion currents giving rise to attenuation are small compared to the field-induced currents. Alternate expansions of the continuity equation, e.g., in terms of the frequency, are discussed in a mathematical appendix.

Introduction

A mode of signal propagation which is peculiar to materials containing mobile charges of both polarities consists in the drift motion of electrostatically neutral, local concentration excesses of these mobile charges under the action of an electric field. Such an excess concentration is not dispersed by coulomb forces. The signal, however, is necessarily attenuated by thermal agitation, which produces a diffusion current tending to spread the concentration excess.

Aspects of this effect of diffusion have been rather thoroughly investigated in recent years by those interested in the applied physics of semiconductors. The present paper may serve the dual purpose of providing a simple, self-contained, and rather intuitive discussion of diffusion attenuation, while adducing results of somewhat more general nature than those customarily obtained.

Only the one-dimensional case is considered. The electric field, however, is allowed to be a general function of position. An iteration procedure is developed for obtaining the general solution, and asymptotic formulas for attenuation and phase change are derived, which apply when the change of electrostatic potential through which the signal passes is sufficiently large.

Formulation

• *General comments*

We shall regard the electric field as prescribed, i.e., as time independent, unmodified by the presence of the signal itself. This condition is approximately fulfilled if the excess concentration is sufficiently small. The electric field E causes a current j proportional to the local charge excess ρ:

$$j(x,t)=v(x)\rho(x,t), \tag{1}$$

where the coefficient of proportionality $v(x)$ is the local drift velocity of the excess concentration. The coordinates of position and time are x and t respectively. In Eq. (1) the current j and charge density ρ are those of either current carrier considered separately. The charge and current of the entire excess concentration is, of course, equal to zero.

An equation of continuity applies,

$$\frac{\partial\rho}{\partial t}=-\frac{\partial j}{\partial x}, \tag{2}$$

which combined with Eq. (1) yields the equivalent equations

$$\frac{\partial\rho}{\partial t}=-\frac{\partial(v\rho)}{\partial x}, \tag{3a}$$

$$\frac{\partial j}{\partial t}=-v(x)\frac{\partial j}{\partial x}. \tag{3b}$$

These equations apply only in the absence of diffusion.

It will prove more convenient to carry forward the discussion in terms of the current rather than in terms of the charge. This is true because the current is unaffected by changes of v with position, while the charge is not. The general solution of Eq. (3b) is given by

$$j(x,t)=j\left(x_0,\, t-\int_{x_0}^{x}\frac{dx'}{v(x')}\right). \tag{4}$$

This equation expresses the fact that the functional dependence of current on time is the same at all points x except for a translation in time equal to the propagation time between points. The current thus moves as an undistorted wave. The charge density, on the other hand, is inversely proportional to the local drift velocity.

Thermal agitation introduces a diffusion current j_D of value

$$j_D = -D \frac{\partial \rho}{\partial x}, \qquad (5)$$

where D is the diffusion coefficient. The new equation of continuity for the total current j is

$$\frac{\partial j}{\partial t} = -v \frac{\partial j}{\partial x} + D \frac{\partial^2 j}{\partial x^2}. \qquad (6)$$

Our problem is to calculate the effect of this extra term on the form of the current.

Intuitively one would expect the attenuation of a wave transmitted over a certain distance in a time T to be approximately equal to the attenuation of a stationary wave of similar shape, should the latter be acted upon by diffusion during the same time interval. The intuition is the basis of the following treatment.

• *Transformation to moving system*[1]

We shall now introduce a transformation to a new coordinate system in which the current, were it not for the effect of diffusion, would be stationary, i.e., to a coordinate system moving with the signal. The new distance coordinate y which, however, has the dimensions of time, is defined as

$$y = -t + \int_0^x \frac{dx'}{v(x')}. \qquad (7)$$

Points of constant y move with local velocity $v(x)$, and points unit y distance apart pass the point $x=0$ at unit time intervals. The new continuity equation can be found by using the partial derivative relations

$$\left(\frac{\partial}{\partial x}\right)_t = \frac{1}{v}\left(\frac{\partial}{\partial y}\right)_t, \qquad (8)$$

$$\left(\frac{\partial}{\partial t}\right)_x = \left(\frac{\partial}{\partial t}\right)_y - \left(\frac{\partial}{\partial y}\right)_t. \qquad (9)$$

It is

$$\left(\frac{\partial j}{\partial t}\right)_x = \frac{D}{v}\frac{\partial}{\partial y}\left(\frac{1}{v}\frac{\partial j}{\partial y}\right). \qquad (10)$$

As anticipated, this has a form similar to that of a pure diffusion equation and is well suited for solution by means of a power-series expansion in D.

It may be well to digress slightly at this point in order to indicate some general properties of $v(x)$ and D. The function $v(x)$ must be definite if signal propagation in the absence of diffusion is to exist; we shall adopt the convention that it be positive definite. The electric field is then either positive or negative definite, depending on whether excess concentrations move as positive or negative charges. We may define a mobility μ by the relation

$$v = \mu E. \qquad (11)$$

The Einstein relation[2] for Boltzmann statistics is

$$D = (kT/e)\mu, \qquad (12)$$

where k, T, and e are respectively Boltzmann's constant, absolute temperature, and charge of the electron. The quantity (kT/e) will be called "thermal voltage" and denoted by V_T. It is clear that expansion in powers of D is essentially equivalent to expansion in powers of V_T, and that one may expect convergence conditions to involve the ratio of V_T to some other voltage. It is also worthwhile to note that D and μ may in general be functions of position. Then it is especially convenient to regard the expansion as one in terms of V_T. We shall, however, assume D constant for simplicity, since any resultant loss of generality can be regained with small effort.

Solution

We shall next make use of the simplification introduced by taking the time variation of the signal to be harmonic. Because of the linearity of the governing equation, such a signal can be attenuated but not otherwise distorted. An unattenuated signal harmonic in time takes the form

$$j = \exp(i\omega y), \qquad (13)$$

in the new coordinate system. We can represent an attenuated signal by the form

$$j = \exp[i\omega y + \phi(y+t)], \qquad (14)$$

since the attenuation must behave similarly with respect to y and t in order that the solution have constant amplitude (in time) at every point x of the original coordinate system. The equation for ϕ resulting from substitution of Eq. (14) in Eq. (10) is

$$\phi' = \frac{D}{v^2}\left[\phi'' + \phi'^2 + \phi'\left(2i\omega - \frac{v'}{v}\right)\omega^2 - \frac{v'}{v} i\omega\right], \qquad (15)$$

a Ricatti equation for ϕ'. Except for the single case $v =$ constant, ϕ will be a complex quantity, yielding information about phase changes as well as about amplitude changes. The advantage of dealing with Eq. (15) rather than directly with Eq. (10) is evident. A change in ϕ small compared to ωy, and thus calculable by treating the diffusion term as a perturbation, may lead to extremely large changes in j, particularly at large y, which could hardly be treated as a perturbation upon j itself.

Expanding ϕ in powers of D, and denoting the solution containing the factor D^n by ϕ_n, we obtain the following equations:

$$\phi'_0 = 0, \qquad (16a)$$

$$\phi_1' = -\frac{D}{v^2}\left(\omega^2 + i\omega\frac{v'}{v}\right), \qquad (16b)$$

$$\phi'_{n+1} = \frac{D}{v^2}\left[\phi''_n + \left(2i\omega - \frac{v'}{v}\right)\phi'_n + \sum_{i+j=n}\phi'_i\phi'_j\right]. \qquad (16c)$$

The integral of ϕ_1' will yield the asymptotic formula for which we sought. However, if the asymptotic solution is to be meaningful, it is necessary that $|\phi_1'|$ be smaller than ω, i.e., that

$$\frac{D\omega}{v^2} < 1 \qquad (17)$$

and

$$\left| \frac{Dv'}{v^3} \right| < 1. \qquad (18)^3$$

The physical meaning of conditions (17) and (18) is not difficult to understand. Condition (17) states that the attenuation per radian must not be larger than $1/e$. Condition (18) requires that v not change so rapidly that the concentration change (which derives from the previously mentioned fact that the concentration is inversely proportional to the electric field) introduces diffusion currents comparable to the drift current. Condition (17) sets an upper limit on the frequency, while condition (18) is a frequency-independent restriction involving only the structure of the field. It may be more illuminating to write it in the form

$$V_T \frac{d(1/E)}{dx} < 1. \qquad (18a)$$

It should be noted that this condition need not apply at boundary points. An example of this is furnished by our later discussion of the case of uniform electric field.

Asymptotic formulas

The solution for the real part of ϕ_1 is

$$Re\phi_1 = -D\omega^2 \int_0^W \frac{dx}{v^3}. \qquad (19)$$

Here the source is regarded as placed at $x=0$, and the attenuation "angle" $Re\phi_1$ is viewed at a position $x=W$. Use has been made of the fact that

$$d(y+t) = v^{-1}dx$$

to transform the integral to the original coordinate frame. The attenuation factor, which we shall call β, is given by

$$\beta = \exp\left(-D\omega^2 \int_0^W \frac{dx}{v^3} \right). \qquad (20)$$

The first-order phase change, $Im\phi$, which we shall denote by θ_1, is

$$\theta_1 = D\omega \left[\frac{1}{v^2} \right]_0^W. \qquad (21)$$

There is, of course, also a zero-order phase change

$$\theta_0 = \omega T, \qquad (22)$$

where T is the transit time,

$$T = \int_0^W \frac{dx}{v}. \qquad (23)$$

A distinction must now be made between expansions in terms of D and in terms of ω.[4] The reader may easily convince himself that $Re\phi$ contains only even powers of ω, while $Im\phi$ contains only odd powers. In order to discover how accurate the asymptotic expression for β is, one must compare $Re\phi_1$ with both the next higher term in ω^2 and with the next higher term in D of order ω^2. The latter term exists only if $v' \neq 0$. We find that the form

$$\beta = \exp(-k\omega^2), \qquad (24)$$

where k is a constant (at a given temperature), will certainly apply with an accuracy ε or better if

$$\left(\frac{D\omega}{v^2} \right)^2 < \varepsilon, \qquad (25)$$

i.e., it is sufficient that the square of the attenuation angle per radian be everywhere smaller than ε. Similarly if k is to agree with the asymptotic prediction of its value to within an accuracy δ, we should presumably require

$$V_T d(E^{-1})/dx < \delta. \qquad (26)$$

If v is constant we obtain the formula

$$\beta = \exp(-\theta_0^2/r), \qquad (27)$$

where r is the ratio of thermal voltage to the charge ΔV in potential occurring through the path of the signal

$$r = V_T/\Delta V, \qquad (28)$$

$$\Delta V = \int_0^W E dx = EW, \qquad (29)$$

$$\theta_0 = \omega T. \qquad (30)$$

If E is truly constant, $\theta_1 = 0$. However, a frequently employed boundary condition, that ρ be zero at $x=W$, can be represented in our framework by the requirement that $v=\infty$ at $x=W$. Although we have $v'=\infty$ at $x=W$, we may still obtain a meaningful solution. In fact we have

$$\theta_1 = -\frac{D\omega}{v(0)^2} = -r\theta_0. \qquad (31)$$

The integral of condition (18a) must always apply; i.e., we must have $r<1$ in order that the asymptotic solution be at all meaningful. One must not conclude from this, however, that the integrated condition is sufficient, since one could always satisfy it merely by taking $E(0)=E(W)$.

The case of uniform field can be solved exactly (cf. Part II), and the asymptotic formula for the attenuation factors is found to fit within the accuracy $(\theta/r)^2$ predicted by inequality (25). It is interesting to note that the accurate solution for β is always larger than the asymptotic value (see Fig. 6 of Part II). This fact gives us further insight into conditions (17) and (25). Evidently at the higher frequencies the ability of diffusion itself to propagate signals becomes important and it is no longer adequate to treat diffusion merely as the attenuator of another mode of propagation.

Summary

Let θ_0 be the zero-order phase change over some distance W. Then the attenuation β over this distance is given by $\beta = \exp(-\theta_0^2/r')$,
where

$$r' = \frac{1}{V_T} \frac{\left[\int_0^W \frac{dx}{E} \right]^2}{\int_0^W \frac{dx}{E^3}},$$

a quantity which for constant electric field E is equal to

$$r = \frac{\Delta V}{V_T},$$

the ratio of potential drop to thermal voltage.

A sufficient condition for the approximate validity of an expression of this form for β is that the attenuation in a distance equal to the wavelength be small throughout the region considered. However, in order that the above equation give the correct value of r' it is also necessary that $E(x)$ be a sufficiently smooth function, as discussed in the text.

Appendix I: Expansion in various parameters

The equation resulting from substitution of $j = J(x)e^{i\omega t}$ into Eq. (6).

$$D\frac{\partial^2 J}{\partial x^2} - v\frac{\partial J}{\partial x} - i\omega J = 0, \tag{32}$$

is a special case of the general form

$$\varepsilon^n \frac{\partial^2 u}{\partial x^2} + p(x)\frac{\partial u}{\partial x} + q(x)u = 0, \tag{33}$$

where ε is a small parameter, and where $p(x)$ and $q(x)$ are analytic functions of ε at $\varepsilon = 0$, at least one of which does not vanish at $\varepsilon = 0$. The dependence on ε of the solutions to this equation has been much studied, though mostly for special forms of $p(x)$ and $q(x)$. A simple transformation of dependent variable eliminates the term in the first derivative. The various special cases that arise can then be discriminated according to the dependence of the new "$q(x)$" on ε. However, the distinctions we make here are most simply discussed with retention of the original form.

The substitutions $u = e^{\phi}$, and $\phi' = y$ yield the equivalent Ricatti equation

$$\varepsilon^n(y' + y^2) + p(x)y + q(x) = 0. \tag{34}$$

Let us assume $q(x) \neq 0$ at $\varepsilon = 0$. Then there is an essential difference between the cases (a): $p(x) = 0$ and (b): $p(x) \neq 0$ at $\varepsilon = 0$, since a solution at $\varepsilon = 0$ exists only for case (b). If we expand all quantities in a power series in terms of ε and denote the terms of mth order in ε by a subscript m, this solution can be expressed as

$$y_0 = q_0(x)/p_0(x). \tag{35}$$

In case (a) on the other hand, the expansion of y commences with negative powers of ε, and u accordingly has an essential singularity with respect to ε at $\varepsilon = 0$. This case has received much attention.[5] The first equation resulting from the expansion is

$$\varepsilon^n y^2_{-(n/2)} + q_0(x) = 0. \tag{36}$$

If, further, $q(x) = q_0(x)$, it is evident that a single condition is all that is required to insure that the next higher term of the expansion of y is much smaller than the first. Taking $n = 2$ for simplicity, one easily finds the condition to be

$$|\varepsilon(q_0^{-\frac{1}{2}})'| << 2, \tag{37}$$

and since, as is evident from Eq. (36), $\varepsilon q_0^{-\frac{1}{2}}$ is the local "wavelength" λ, we can write equivalently

$$\frac{1}{2}\frac{\delta\lambda}{\delta x} << 1. \tag{37a}$$

(The time-independent Schroedinger equation has the form in question; then λ is the de Broglie wavelength.)

As we have seen, in case (b) the expansion of y commences with y_0. In the special situation parallel to that discussed above, i.e., $p(x) = p_0(x)$, $q(x) = q_0(x)$ (allowing us to take $n = 1$ without loss of generality), there will be two conditions necessary to imply $y_1 << y_0$, namely

$\varepsilon y'_0 << q(x)$, $\varepsilon y^2_0 << q(x)$, or

$$\varepsilon[q(x)/p(x)]' << q(x) \tag{38a}$$

$$\varepsilon[q(x)/p(x)^2] << 1. \tag{38b}$$

The fact that there are two conditions, rather than one, stems from the equal order with respect to ε of the lowest-order components of y' and y^2. Substitution of the particular forms of $p(x)$ and $q(x)$ contained in Eq. (32) yields the conditions found in the text.

A third case of interest to us is characterized by $n = 0$. Unless $q(x) = 0$ we cannot, in general, solve the equation for y_0, since the problem is equivalent to the general one before expansion. The case $p(x) = p_0(x)$, $q(x) = q_1(x)$ can, however, be solved, again by expansion of y in terms of ε. We have first

$$y'_0 + y^2_0 + p_0(x)y_0 = 0 \tag{39}$$

which can be solved in terms of quadratures, since the corresponding linear equation for u'_0 is soluble. Actually, the trivial solution $y_0 = 0$ turns out to be the proper one for the application of the method to Eq. (32). With this choice for y_0, our second equation becomes

$$y'_1 + p_0(x)y_1 + q_1(x) = 0, \tag{40}$$

which is easily solved in terms of quadratures. Continuing in this manner we next obtain

$$y'_2 + p_0(x)y_2 + y^2_1 = 0, \tag{41}$$

et cetera. The resulting set of equations is soluble by a sequence of quadratures. The constants of integration are of course determined by the type of solution it is desired to approximate. This method is directly applicable to solution of Eq. (32), by expansion in terms of ω, but the details will be omitted.

A fourth case of interest is that arising from the problem at hand when ω is considered large. Taking $\varepsilon^{-2} = i\omega/D$, and expanding J about $\varepsilon = 0$, we can easily obtain an asymptotic solution valid for large frequencies. This case is, of course, only a particular instance of case (a). The first and second terms of the expansion are easily found to be

$$y_{-1} = \sqrt{i\omega/D} = \frac{1}{\sqrt{2}}(1+i)\sqrt{\omega/D}$$

$$y_0 = \frac{v}{2D}.$$

The attenuation over a distance W is the exponential of the integral of the real part of these quantities over the distance W. It is seen that the principal, or -1st order, contribution to the attenuation is independent of the field, while the zero-order contribution depends only on the potential difference across the distance W, and not on the detailed shape of the potential function.

Appendix II: Minimal attenuation

A question of some interest is the following: given a potential difference ΔV between the points $x=0$ and $x=w$, what sort of potential function $V(x)$ between these points will minimize the attenuation? This question has an obvious answer if no continuity restrictions are imposed; for the field $E(x)$ may be made as large as desired, and the corresponding attenuation as small as desired, by making $\lim_{x\to 0} V(x)$ sufficiently large. But even requiring $V(x)$ to be continuous, introduction of a high maximum in $V(x)$ sufficiently close to $x=0$, followed by an approximately linear variation of $V(x)$ from this maximum to the endpoint $x=W$, will lower the attenuation factor to an arbitrarily small value. This fact, which underlies the advantage of the drift transistor,[6] is due to the efficiency of the diffusion mechanism over small distances. Thus, the question can meaningfully be asked only if the potential $V(x)$ assumes its highest value at $x=0$. Further, the value of the frequency at which the attenuation is to be minimized must be specified.

If the electric field is subject to the restriction (18a) over the entire interval $x=0$ to $x=W$, one can employ the asymptotic attenuation formula, and by equating to zero the change of attenuation caused by a variation $\delta v(x)$ of the drift velocity, one readily confirms the intuitively expected result that the attenuation has an extremum (minimum) for $v(x)=\text{const}$. However, this procedure is valid only if condition (18a) applies also to the variation. It is readily shown on the basis of qualitative arguments that, requiring only that $V(x)$ assume its largest value at $x=0$, the function $V_m(x)$ which minimizes the attenuation at a given frequency will have zero slope at $x=0$. However, if $\Delta V/V_T$ is large, the function $V_m(x)$ will not differ considerably from the straight line joining the prescribed values of V at $x=0$ and $x=W$. The distance near $x=0$ over which $V(x)$ is comparatively flat is approximately given by

$$W\left(\frac{2V_T}{\Delta V}\right)^3 \frac{\omega}{\omega_D},$$

where $\omega_D=W^2/2D$, and where ω is the frequency at which the attenuation is to be minimized. The fractional attenuation decrease over the attenuation occurring for the case of linear $V(x)$ is about $(2V_T/\Delta V)^3\omega/\omega_D$. As the frequency of minimization ω becomes large, this formula loses its validity, and a maximum fractional attenuation decrease, less than $2V_T/\Delta V$, occurs in the neighborhood of

$$\omega=\omega_D\left(\frac{2\Delta V}{V_T}\right)^2.$$

At higher frequencies the attenuation becomes substantially independent of the shape of the curve $V(x)$. Our conclusion is that the linear potential function is a good approximation to that function $V_m(x)$ which minimizes attenuation under the restriction that $V(x)$ attain its maximum at $x=0$.

References

1. Although this artifice is intuitively suggestive, it does not actually simplify the mathematical treatment (cf. Appendix).
2. This relation applies approximately in the case under discussion if one carrier is very much in the majority. In general, a factor $(n+p)/(n-p)$, where n and p are the concentrations of negative and positive carriers respectively, must be inserted on the right.
3. This condition is somewhat different from the analogous condition for the application of the B.W.K. method in quantum mechanics, which in our notation would read

   $$\frac{1}{\omega}\frac{dv}{dx}<1.$$

   In fact inequality (18) is the product of inequality (17) and the B.W.K. inequality. For this reason the B.W.K. inequality is not a necessary condition of convergence at low frequencies (cf. Appendix).
4. Cf. Appendix.
5. See, for example, A. Erdélyi, *Asymptotic Expansions,* Dover Publications, Inc., 1956, and the considerable literature upon the B.W.K. approximation.
6. H. Kroemer, *Archiv. Elek. Ubertr.* **8,** 223, 499 (1954).

Revised manuscript received June 26, 1958