Theory of a Configurational emf*

MARVIN CHESTER
Physics Department, University of California, Los Angeles, California
(Received 24 September 1963)

A configuration emf refers to an electrical manifestation in the electron gas of the analog of the Bernoulli effect for fluid flow. A theoretical account of this phenomenon is given. The object of the theory is to present a description of the effect from first principles which in this instance is taken as the Boltzmann transport equation. The transport equation is solved for a two-band (holes and electrons) Fermi particle system. The solutions are carried beyond first order so as to exhibit the effect. It is shown that the form of the emf is, indeed, extractable from the Boltzmann equation. And its magnitude is obtained in terms of a coefficient C for which a formula is given in terms of familiar current carrier parameters.

I. INTRODUCTION

RECENT experiments[1] indicate the existence of an effect which may be called a "configurational emf." This phenomenon is the analog, in electrical current flow, of the Bernoulli effect in conventional fluid flow. Suppose that a current-carrying medium is physically constricted in cross section as in Fig. 1. Continuity demands that the total incoming current be equal to the total outgoing current. But the two regions differ in cross-sectional area. It follows that the current density in the narrow region must be larger than that in the wide region. Since the electron gas is relatively incompressible this difference in current density is accounted for by a difference in drift velocity of the carriers in the two regions. The incompressibility of the electron gas is due to the fixed density of background oppositely charged ions. This tends to hold the density of electrons constant.

This property of increased flow velocity due to a constriction is common to any incompressible fluid. In a hydrodynamic fluid the flow rate increase gives rise to a pressure decrease. For the electron gas case it has been postulated[1–5] that the increased momentum in the constricted region arises from electrical origins—in particular an electric field $\mathcal{E}_c$ confined to the region where the cross-sectional area is changing. Hence, instead of a pressure difference being correlated to the current density change, an electrical potential difference is expected. This state of affairs is illustrated in Fig. 1 for the case of positive carriers. The electrical potential step at the intersection of the two regions would be opposite in direction if the current was due to negative carriers. The term configurational emf is employed because the postulated electric field (and, hence, emf) arises from the shape or configuration of the specimen.

Since it is to account for an impulsive momentum change that this configurational field is invoked one may expect, for positive carriers of charge e and mass m, that

$$\int_{\text{region 1}}^{\text{region 2}} e\mathcal{E}_c dx = \int_1^2 \frac{d(mv)}{dt} dx = \tfrac{1}{2}mv^2\Big|_1^2. \tag{1}$$

Hence,

$$V_c\Big|_1^2 = \Delta V_c = -\frac{1}{2}\frac{m}{e}\left(\frac{1}{ne}\right)^2 j^2\Big|_1^2, \tag{2}$$

where the current density is j, n is the volume density of carriers, and ΔV_c is the configurational voltage step. For carriers of opposite sign we must replace e by $-e$ so that ΔV_c will change sign.

Equation (2) is based entirely upon an intuitive postulate. It has not been founded upon "first principles." In the following, a more fundamental theory is presented for a configurational emf. It is based on the Boltzmann transport equation for the electron gas. The theory results in a more exact expression than Eq. (2). This expression includes the case of both positive and negative carriers acting simultaneously. Furthermore, it is demonstrated that a configurational emf term can, indeed, be extracted from solutions of the transport equation.

II. MACROSCOPIC TRANSPORT

The program that will be followed entails solving the Boltzmann transport equation. The solutions must be

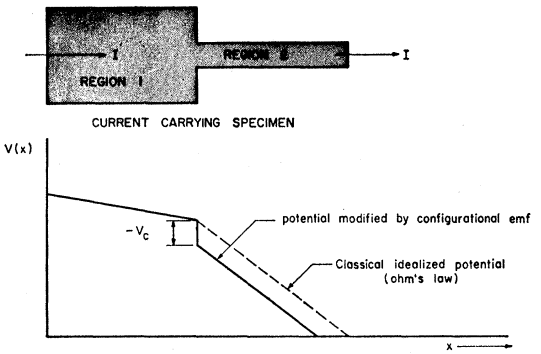

FIG. 1. Idealized electrical potential versus distance for a constricted specimen.

* This research was supported in part by the U. S. Office of Naval Research.

[1] M. Chester, Phys. Rev. Letters **5**, 91 (1960).
[2] A. Perrier, Bull. Soc. Vaudoise Sci. Nat. **56**, 39 (1925).
[3] N. V. Ivashchenko, Zh. Eksperim. i Teor. Fiz. **9**, 892 (1939).
[4] Y. G. Dorfman and A. S. Kagan, Zh. Eksperim. i Teor. Fiz. **10**, 358 (1940).
[5] R. Jaggi, Phys. Rev. **122**, 448 (1961).

carried to a sufficiently high order of approximation to exhibit the effect. The configurational emf is not a first-order effect. It is small and, therefore, it is expected to appear only in a higher order solution of the transport equation.

In order to relate the Boltzmann equation solutions to experimental macroscopic parameters it is well to have a macroscopic formulation of the problem with which the integrated Boltzmann solutions may be compared. The macroscopic formulation can be exemplified by the following[6]:

$$\mathcal{E}-(1/\sigma)\mathbf{j}+R\mathbf{j}\times\mathbf{B}-G\nabla T-2C\mathbf{j}\cdot\nabla\mathbf{j}=0. \quad (3)$$

Equation (3) exhibits the macroscopic relationship between electric fields, magnetic fields ($\mathbf{B}$), thermal gradients (∇T), and current densities ($\mathbf{j}$) in a current-carrying medium. It is a more accurate statement than the bare Ohm's law. The latter is contained in the first two terms of the equation. Ohm's law holds only in the steady state when no magnetic fields are present and there are no thermal gradients nor spatial current gradients. In fact, of course, Eq. (3) is not exact either. Depending upon the measurements to be made and the conditions that exist, other terms, not exhibited in (3), may manifest themselves requiring further additions to that equation.

Equation (3) is to be interpreted in the following way. Each separate term in it constitutes a separate emf-generating electric field. The field $\mathcal{E}$ is that applied by some external source of emf. The field $-(1/\sigma)\mathbf{j}=\mathcal{E}_{\text{Ohm}}$ is that field which generates an ohmic "emf." The field $R\mathbf{j}\times\mathbf{B}=\mathcal{E}_{\text{Hall}}$ generates a Hall effect emf (R=Hall coefficient). The field $\mathcal{E}_T=-G\nabla T$ generates a thermoelectric emf with amplitude G (G=thermoelectric power). And the field $\mathcal{E}_C=2C\mathbf{j}\cdot\nabla\mathbf{j}$ generates a configurational emf. To obtain the emf (measurable potential difference) associated with any of the fields in Eq. (3), it is merely necessary to integrate the field along a path connecting the two points between which the potential difference is desired.

It is important to notice the tensor character of $\mathcal{E}_C$ as opposed, for example, to $\mathcal{E}_T$. It is this which makes it possible to measure the configurational effect in a single material. If $\mathcal{E}_C$ could be written as the gradient of a scalar, then the effect would not be measurable in a single specimen.[7] Like the Seebeck effect, it would require two specimens.

The coefficient C represents the strength of the configurational emf. The crude model of the previous section yields a value for C of

$$C=\tfrac{1}{2}m/e^3n^2. \quad (4)$$

The coefficient is introduced so that it is positive for positive carriers and negative for negative carriers.

[6] A. H. Wilson, *The Theory of Metals* (University Press, Cambridge, England, 1953), Sec. 8.5.
[7] M. Chester, thesis, California Institute of Technology, 1961 (unpublished).

In the light of the preceding, the program reduces itself to the following two objectives. First, it is necessary to show that a term of the form $-2C\mathbf{j}\cdot\nabla\mathbf{j}$ does, indeed, represent a proper extension of Ohm's law as indicated in Eq. (3). (The other terms in the equation are, of course, already familiar.) Secondly, the value of the coefficient C must be calculated. For both of these purposes we resort to the Boltzmann transport equation for Fermi particles (electrons).

III. MICROSCOPIC TRANSPORT

Neglecting the interactions between the electron spin and possible magnetic fields, we let $(1/4\pi^3)f(\mathbf{k},\mathbf{r})\times d^3kd^3r$ represent the number of electrons found in the volume of phase space d^3kd^3r, where $\mathbf{r}$ represents a position vector and $\hbar\mathbf{k}$ the momentum vector (the constant $h=2\pi\hbar$ is Planck's constant). The wave vector $\mathbf{k}$ has magnitude $2\pi/\lambda$, where λ is the electron wavelength. With this notation the undisturbed equilibrium distribution $f_0(\mathbf{k},\mathbf{r})$ is the familiar Fermi-Dirac one:

$$f_0(\mathbf{k},\mathbf{r})=f_0(E)=\{1+\exp[(E-E_F)/KT]\}^{-1}. \quad (5)$$

Here $E=E(\mathbf{k})$ is the energy associated with wave vector $\mathbf{k}$, E_F is the Fermi energy, K is Boltzmann's constant, and T is the absolute temperature.

When steady-state transport processes are occurring in the medium the distribution function $f(\mathbf{k},\mathbf{r})$ is not the equilibrium one but rather follows Boltzmann's transport equation[8,9]

$$\mathbf{v}\cdot\nabla f+(1/\hbar)\mathbf{F}\cdot\nabla_k f=(\partial f/\partial t)_c. \quad (6)$$

In this equation, $\mathbf{v}$ is the group velocity given by

$$\mathbf{v}=(1/\hbar)\nabla_k E \quad (7)$$

and the gradient ∇ without a subscript refers to spatial derivatives, whereas ∇_k is a gradient in k space. The force F for electrons is given by[8]

$$\mathbf{F}=-e(\mathcal{E}+\mathbf{v}\times\mathbf{B}), \quad (8)$$

where $\mathcal{E}$ is the local impressed electric field, $\mathbf{B}$ the impressed magnetic field, and e is the magnitude of the charge on the electron.

Magnetic Field Terms

The effect we are examining is not a magnetic one so we do not consider external magnetic fields. This, however, does not imply that $\mathbf{B}$ in Eq. (8) is identically zero. A self-magnetic field is always present due to the currents which are flowing in the specimen. Such a field arises by virtue of the relation

$$\nabla\times\mathbf{B}=\mu\mathbf{j}, \quad (9)$$

where μ is the permeability of the medium (essentially

[8] J. M. Ziman, *Electrons and Phonons* (Clarendon Press, Oxford, 1960), Sec. 7.3.
[9] G. V. Chester, Proc. Phys. Soc. (London) **81**, 938 (1963).

equal to μ_0, that of free space). And, in fact, the entire effect of this field is to generate a "self-Hall" voltage.[10,11] This effect competes with the configurational emf in an experimental situation because the voltage associated with the self-Hall effect exhibits many of the same properties as the configurational emf.[5] For example, the self-Hall voltage depends upon the square of the current and upon spatial current density gradients. It has the same tensor properties as does the configurational field. However, its magnitude is quite different; its size dependence is different; and it involves no microscopic parameters other than the Hall coefficient directly.

The fact that it competes becomes clear upon substitution of Eq. (9) into the Hall term of (3)

$$\mathcal{E}_{\mathrm{Hall}}=R\mathbf{j}\times\mathbf{B}=(R/\mu)[\mathbf{B}\cdot\nabla\mathbf{B}-\tfrac{1}{2}\nabla B^2]. \quad (10)$$

Although any voltage associated with the second term in the brackets of Eq. (10) would not be measurable in a single material,[7] an emf connected with the first term would be measurable. It is this term which gives rise to the self-Hall voltage. A short calculation applying (10) to the specimen of Fig. 1 yields the result

$$V_{\mathrm{SH}}|_1{}^2=\Delta V_{\mathrm{SH}}=-\gamma(\mu/4\pi)RI^2/A\,, \quad (11)$$

where A is the cross-sectional area in region 2 (considered to be much smaller than that of region 1), and γ is a geometrical factor equal to unity for a right circular cylindrical cross section. It is of the order of unity for any other cross-sectional shape.

Now, a careful examination indicates that the foregoing completely exhausts the ways in which a self-magnetic field may enter the problem.[7] Therefore, having dealt with it, we will set $B=0$ and concentrate on the electric effects. We proceed with the twofold object of demonstrating the form of the configurational term in (3) and of calculating C. Further discussion will not be given to demonstrate that no other effects exist which can compete with the one of interest here.

Boltzmann Equation Solution

The right-hand side of Eq. (6) represents the time rate of increase of the particle density at $(\mathbf{k},\mathbf{r})$ due to collisions. Let $W(\mathbf{k},\mathbf{q})$ represent the probability per unit time that an electron initially having momentum $\mathbf{q}$ will, upon collision, find itself with a momentum between $\mathbf{k}$ and $\mathbf{k}+d\mathbf{k}$. Then we have, for Fermi particles,[8,9]

$$(\partial f/\partial t)_c=[1-f(\mathbf{k})]\int W(\mathbf{k},\mathbf{q})f(\mathbf{q})d^3q - f(\mathbf{k})\int[1-f(\mathbf{q})]W(\mathbf{q},\mathbf{k})d^3q, \quad (12)$$

where the spatial dependence of f is not explicitly indicated but is meant to be understood in the foregoing.

The calculational procedure will be the following: We expand $f(\mathbf{k},\mathbf{r})$ in a sequence of terms of decreasing magnitude

$$f(\mathbf{k},\mathbf{r})=f_0(E)+f_1(\mathbf{k},\mathbf{r})+f_2(\mathbf{k},\mathbf{r})+\cdots. \quad (13)$$

Expressions for the functions $f_n(\mathbf{k},\mathbf{r})$ are obtained using the Boltzmann equation successively for each order.[12,13] The inhomogeneous equation for the nth order has as its "forcing" term the solution to the $n-1$ order equation. Next, the expanded solution for the transport-modified distribution function $f(\mathbf{k},\mathbf{r})$ is used to find the current density $\mathbf{j}$ created by the applied electric field

$$\mathbf{j}=-\frac{e}{4\pi^3}\int\mathbf{v}f(\mathbf{k},\mathbf{r})d^3k =-\frac{e}{4\pi^3}\int(\mathbf{v}f_1+\mathbf{v}f_2+\cdots)d^3k. \quad (14)$$

The emf associated with each of the various terms contributing to the current is then given by the following integral connecting two points between which the emf is desired.

$$\mathrm{emf}=\int_\sigma^1\mathbf{j}\cdot d\mathbf{r}. \quad (15)$$

The series of functions $f_n(\mathbf{k},\mathbf{r})$ are characterized primarily by their angular dependence. They have the angular form of the forcing function which drives the integral equations to which they are solutions. Therefore, to obtain the successive approximations, it is expedient to expand the scattering probability $W(\mathbf{k},\mathbf{q})$ in normalized surface spherical harmonics $Y_l{}^m(\mu,\Phi)$, where[14]

$$Y_l{}^m(\mu,\Phi)=\left[\frac{2l+1}{4\pi}\frac{(l-|m|)!}{(l+|m|)!}\right]^{1/2}P_l{}^m(\mu)e^{im\Phi}, \quad (16)$$

and the $P_l{}^m(\mu)$ are the associated Legendre spherical functions of the first kind. For these functions we follow the usage[15]

$$P_l{}^{-m}=P_l{}^m=\frac{(1-\mu^2)^{|m|/2}}{2^l l!}\frac{d^{|m|+l}}{d\mu^{|m|+l}}(\mu^2-1)^l, \quad (17)$$

and the normalization of the $Y_l{}^m(\mu,\Phi)$ is written as

$$\int_{-1}^{1}\int_0^{2\pi}Y_l{}^mY_\lambda{}^\nu d\Phi d\mu=\delta_{l\lambda}\delta_{m\nu}. \quad (18)$$

[10] R. Jaggi and R. Sommerhalder, Helv. Phys. Acta **32**, 167 (1959).

[11] G. Busch and R. Jaggi, Z. Angew. Math. Phys. 4, 425 (1953).

[12] H. Grad, in *Handbuch der Physik*, edited by S. Flügge (Springer-Verlag, Berlin, 1958), Vol. XII.

[13] K. Huang, *Statistical Mechanics* (John Wiley & Sons, Inc., New York, 1963), Chap. 6.

[14] L. I. Schiff, *Quantum Mechanics* (McGraw-Hill Book Company, Inc., New York, 1949), p. 73.

[15] H. Margenau and G. M. Murphy, *Mathematics of Physics and Chemistry* (D. Van Nostrand, Inc., New York, 1943), p. 109.

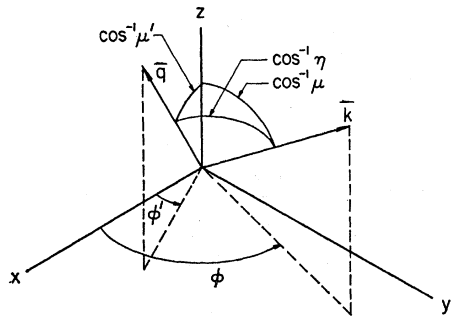

FIG. 2. Coordinate nomenclature for an electron scattering.

With reference to Fig. 2 it is clear that we may write down the expansion for the scattering probability as follows:

$$W(\mathbf{k},\mathbf{q})=\sum_{l=0}^{\infty}\left[\frac{2l+1}{4\pi}\right]^{1/2}w_l(k,q)Y_l^0(\eta),$$

$$=\sum_l\sum_{m=-l}^{+l}w_l(k,q)Y_l^m(\mu,\Phi)Y_l^{-m}(\mu',\Phi'). \quad (19)$$

The first line of Eq. (19) just reflects the idea that W depends only upon the angle between $\mathbf{k}$ and $\mathbf{q}$ and not upon angles relative to a fixed coordinate system. The second line of Eq. (19) results from a well-known addition theorem regarding spherical harmonics.[16] By virtue of the orthogonality of the Y_l^m and Eq. (19) we may always obtain the functions $w_l(k,q)$ from

$$w_l(k,q)=2\pi\int_{-1}^{1}W(\mathbf{k},\mathbf{q})P_l(\eta)d\eta. \quad (20)$$

To proceed to the solution of the Boltzmann equation, we must first take account of the relationship between $W(\mathbf{k},\mathbf{q})$ and $W(\mathbf{q},\mathbf{k})$. In equilibrium the net change in the distribution function due to collisions is zero. This idea, strengthened by the principle of detailed balance, yields the result

$$[1-f_0(k)]W(\mathbf{k},\mathbf{q})f_0(q)=[1-f_0(q)]W(\mathbf{q},\mathbf{k})f_0(k). \quad (21)$$

Hence the right-hand side collision term of the Boltzmann equation (6) becomes

$$\left(\frac{\partial f}{\partial t}\right)_c=-g(\mathbf{k},\mathbf{r})\int W(\mathbf{q},\mathbf{k})\frac{f_0(q)}{f_0(k)}d^3q+f_0(k)\int W(\mathbf{k},\mathbf{q})$$

$$\times g(\mathbf{q},\mathbf{r})d^3q+[1-f_0(k)]\int W(\mathbf{q},\mathbf{k})g(\mathbf{q},\mathbf{r})d^3q$$

$$+g(\mathbf{k},\mathbf{r})\int[W(\mathbf{k},\mathbf{q})-W(\mathbf{q},\mathbf{k})]g(\mathbf{q},\mathbf{r})d^3q. \quad (22)$$

[16] W. Magnus and F. Oberhettinger, *Formulas and Theorems for the Functions of Mathematical Physics* (Chelsea Publishing Company, New York, 1954), p. 55.

Here the symbol $g(\mathbf{k},\mathbf{r})$ is simply a shorthand notation for the difference

$$f(\mathbf{k},\mathbf{r})-f_0(k)=g(\mathbf{k},\mathbf{r}). \quad (23)$$

First-Order Calculation

If we orient the z direction of Fig. 2 to be that of the field locally then, to first order, the left-hand side driving term of the Boltzmann equation becomes

$$\frac{1}{\hbar}\mathbf{F}\cdot\nabla_k f_0(k)=-\frac{e}{\hbar}(\mathcal{E}\cdot\nabla_k E)\frac{\partial f_0}{\partial E},$$

$$=-e\mathcal{E}v\mu\frac{\partial f_0}{\partial E},$$

$$=-e\mathcal{E}v\frac{\partial f_0}{\partial E}\left(\frac{4\pi}{3}\right)^{1/2}Y_1^0(\mu). \quad (24)$$

In the foregoing, we have assumed that the energy is isotropic and quadratic in $\mathbf{k}$ so that the velocity [Eq. (7)] is in the same direction as $\mathbf{k}$. Furthermore, we hold the temperature uniform throughout. We will maintain these simplifying assumptions in all that follows.

In response to the angular dependence of the driving term (24), we choose a form for the first-order change in distribution as follows:

$$f_1(\mathbf{k},\mathbf{r})=G_1(k,\mathbf{r})Y_1(\mu,\Phi). \quad (25)$$

Therefore, to first order, the Boltzmann equation becomes an integral equation for G_1

$$-e\mathcal{E}v\frac{\partial f_0}{\partial E}\left(\frac{4\pi}{3}\right)^{1/2}$$

$$=-G_1(k,\mathbf{r})\int_0^\infty w_0(q,k)\frac{f_0(q)}{f_0(k)}q^2dq$$

$$+f_0(k)\int_0^\infty w_1(k,q)G_1(q,\mathbf{r})q^2dq$$

$$+[1-f_0(k)]\int_0^\infty w_1(q,k)G_1(q,\mathbf{r})q^2dq. \quad (26)$$

We are interested only in the particular solution of this equation, not in the general one, because the latter is independent of the field. Since we have the "boundary condition" that G_1 is zero if $\mathcal{E}$ is zero, it follows that the particular (or driven) solution of (26) is the complete one.

As is usual in the computation to first order we now resort to an "elastic" scattering approximation. Since for impurity and for phonon scattering the electron exchanges only a very small fraction of its total energy with the scatterer we will neglect this energy exchange entirely where possible. It is possible in this order

but, such an approximation cannot always be implemented.[17–19] This will become apparent in the second-order calculation. The "elastic" approximation may be summarized in the statement

$$w_l(q,k)\simeq\delta(q-k)\omega_l(k), \tag{27}$$

where the delta function is a three-dimensional weighted one defined by

$$\int_0^\infty F(q)\delta(q-k)q^2dq=F(k). \tag{28}$$

With this approximation the integral equation (26) for $G_1(k,\mathbf{r})$ reduces to

$$-\left(\frac{4\pi}{3}\right)^{1/2}ev\mathcal{E}\frac{\partial f_0}{\partial E}=-[\omega_0(k)-\omega_1(k)]G_1(k,\mathbf{r}). \tag{29}$$

If we define[20] a relaxation time τ_1 by

$$\frac{1}{\tau_1}=\omega_0-\omega_1=2\pi\int_0^\infty q^2dq\int_{-1}^1(1-\eta)W(\mathbf{k},\mathbf{q})d\eta, \tag{30}$$

then the results of (25)–(30) may be recombined so that the first-order solution to the Boltzmann equation becomes the well known one[6]

$$f_1(\mathbf{k},\mathbf{r})=e\tau_1\frac{\partial f_0}{\partial E}\frac{1}{\hbar}\boldsymbol{\mathcal{E}}\cdot\boldsymbol{\nabla}_k E. \tag{31}$$

Second-Order Calculation

To carry the calculation to second order we inspect the left-hand driving term of (6) for its angular dependence. This driving term is

$$\left(\mathbf{v}\cdot\boldsymbol{\nabla}-\frac{e}{\hbar}\boldsymbol{\mathcal{E}}\cdot\boldsymbol{\nabla}_k\right)f_1(\mathbf{k},\mathbf{r})=\left(\mathbf{v}\cdot\boldsymbol{\nabla}-\frac{e}{\hbar}\boldsymbol{\mathcal{E}}\cdot\boldsymbol{\nabla}_k\right)\left(\frac{e}{\hbar}\tau_1\frac{\partial f_0}{\partial E}\boldsymbol{\mathcal{E}}\cdot\boldsymbol{\nabla}_k E\right). \tag{32}$$

The spatial dependence is contained in the electric field $\boldsymbol{\mathcal{E}}=\boldsymbol{\mathcal{E}}(\mathbf{r})$. Upon resolving (32) into spherical harmonics employing the notation of Fig. 2, it becomes clear that we must choose the following form for $f_2(\mathbf{k},\mathbf{r})$:

$$f_2(\mathbf{k},\mathbf{r})=G_0(k,\mathbf{r})+\sum_{m=-2}^{+2}G_{2m}(k,\mathbf{r})Y_2{}^m(\mu,\Phi). \tag{33}$$

This choice of form is suggested by virtue of the identities

$$k_xk_z=k^2\left(\frac{2\pi}{15}\right)^{1/2}(Y_2{}^1+Y_2{}^{-1}), \tag{34}$$

$$k_yk_z=-ik^2\left(\frac{2\pi}{15}\right)^{1/2}(Y_2{}^1-Y_2{}^{-1}), \tag{35}$$

$$k_z{}^2=k^2[(4\pi)^{1/2}/3][Y_0+(2/5^{1/2})Y_2{}^0]. \tag{36}$$

That is, the new forcing function [expression (32)] is expandable into a spherically symmetric part plus second-order spherical surface harmonic parts. In examining the integral equation for the G_{2m} arising from the latter under the assumption of "elastic scattering" it is found that a new time τ_2 arises quite naturally[20] where

$$\frac{1}{\tau_2}=\omega_0-\omega_2=2\pi\int_0^\infty q^2dq\int_{-1}^1[1-P_2(\eta)]W(\mathbf{k},\mathbf{q})d\eta. \tag{37}$$

Because the scattering rate $W(\mathbf{k},\mathbf{q})$ depends upon $\mathbf{q}$, only relative to $\mathbf{k}$, and not upon the absolute directions of these vectors, the same time parameter τ_2 arises for each of the $G_{2m}(k,\mathbf{r})$ amplitudes independent of m.

In trying to adopt the same procedure used for the $l=1$ and $l=2$ cases to the spherically symmetric ($l=0$) integral equation a difficulty arises. From the form of $1/\tau_1=\omega_0-\omega_1$ and $1/\tau_2=\omega_0-\omega_2$ one might expect the spherically symmetric integral operator to yield $1/\tau_0=\omega_0-\omega_0=0$ or a relaxation time τ_0 which is infinite. Indeed, this is exactly what one obtains if the "elastic approximation" is applied to the spherically symmetric case. That this is so may be demonstrated if we write down the integral equation for the angle-independent part of f_2. This equation is

$$\tfrac{1}{3}e\tau_1\frac{\partial f_0}{\partial E}v^2\boldsymbol{\nabla}\cdot\boldsymbol{\mathcal{E}}-\frac{e^2}{\hbar^2}\tau_1\frac{\partial f_0}{\partial E}(\boldsymbol{\mathcal{E}}\cdot\boldsymbol{\nabla}_k)^2E-\tfrac{1}{3}e^2\mathcal{E}^2v^2\frac{\partial}{\partial E}\left(\tau_1\frac{\partial f_0}{\partial E}\right)-\tfrac{1}{3}e^2\mathcal{E}^2\tau_1v\frac{\partial f_0}{\partial E}\int_0^\infty[w_0(k,q)-w_0(q,k)]\tau_1(q)\frac{\partial f_0(q)}{\partial E(q)}v(q)q^2dq$$
$$=-G_0(k,\mathbf{r})\int_0^\infty w_0(q,k)\frac{f_0(q)}{f_0(k)}q^2dq+f_0(k)\int_0^\infty w_0(k,q)G_0(q,\mathbf{r})q^2dq+[1-f_0(k)]\int_0^\infty w_0(q,k)G_0(q,\mathbf{r})q^2dq. \tag{38}$$

The last term of the left-hand (forcing) side of (38) arises from the nonlinear aspect of the Boltzmann transport equation when Fermi-Dirac satistics obtain. It comes from the nonlinear part of the collision rate expression given in (22). From Eq. (38) it is immediately clear that the "elastic" approximation of (27) cannot be applied because it implies that the finite forcing function on the left is equal to zero if G_0 is finite.

[17] W. Shockley, Bell System Tech. J. **30**, 990 (1951).
[18] J. B. Gunn, Progr. Semiconductors **2**, 213 (1957).
[19] T. N. Morgan, Phys. Chem. Solids **8**, 245 (1959).
[20] R. E. Peierls, *Quantum Theory of Solids* (Clarendon Press, Oxford, 1955), p. 119.

Therefore, in order to obtain the function $G_0(k,\mathbf{r})$ we may not assume that the scattering is elastic.[19] We must solve the complete integral equation (38).

In order to obtain a tractable solution to the problem we proceed in the following way: Equation (38) is an inhomogeneous linear integral equation for $G_0(k,\mathbf{r})$. It has a particular solution (which we wish to find) and general solutions. We may symbolically write Eq. (38) in the form

$$F=IG_0, \tag{39}$$

where F represents all of the left-hand side of (38) and IG_0 represents the integral operator acting on G_0 exhibited by the complete right-hand side of (38). Now, if instead of (39), we examine the homogeneous equation

$$\alpha G_0=IG_0, \tag{40}$$

then there will be only a specific set, α_1, $\alpha_2\cdots$, of numbers α for which Eq. (40) will hold, and to each α there will correspond a specific function $G_0^{(\alpha)}$. These, of course, represent the eigenvalues and eigenfunctions of the homogeneous linear equation (40). We may expand the forcing function F in terms of these eigenfunctions $G_0^{(\alpha)}$. In general, all of the eigenfunctions will be needed to properly describe F by this expansion. However, it is not unreasonable to expect, by virtue of the close relationship between the driving function F and the response G_0, that one particular mode $G_0^{(\alpha_1)}$ will dominate all the others in the expansion for F. That is, we suppose that F will have approximately the same form as a particular one of the many eigenfunctions $G_0^{(\alpha)}$. Hence, F may be represented in large measure by just the one term containing this dominant eigenfunction. For that function which most closely approximates $F(k,\mathbf{r})$ we call the eigenvalue $\alpha=\alpha_1=-1/\tau_0$. It than follows from (40) and (39) that

$$G_0=-\tau_0 F. \tag{41}$$

Hence, in the spherically symmetric case, the form of the solution is the same as in the $l=1$ and the $l=2$ cases, but the meaning of the parameter τ_0 is slightly different. It must be obtained from the solution of the homogeneous integral equation (40). It should be mentioned here that a more exact, but less simple symbolic answer could easily have been obtained in place of (41) if we had not made the assumption that one mode of the $G_0^{(\alpha)}$ dominates. But this complication does not yield any better information and tends to obscure the interpretation and significance of the results.

Clearly, the relaxation time τ_0 must be very large[21,22] compared to τ_1 since in first approximation it was found to be infinite. It is also apparent, by virtue of the connection to the inapplicability of the "elastic" approximation, that this time is related to the inelasticity of collisions.[22] In fact the rate $1/\tau_0$ must pertain to the rate at which *energy* (rather than momentum) is transferred between electrons and scatterers on collision. In the elastic approximation no energy is ever transferred and hence this rate, quite reasonably, turned out to be zero.

Keeping in mind that the term in (38) containing the divergence of the electric field is zero because the divergence of the current density is zero for steady-state flow, we note that τ_0 is associated with $\mathcal{E}^2$ (or $\mathbf{j}\cdot\boldsymbol{\mathcal{E}}$). Furthermore, it arises only for the spherically symmetric part of the increase in the distribution function $f(\mathbf{k},\mathbf{r})$. These facts indicate that τ_0 is related to the heating of the electron gas due to the electrical power input $\mathbf{j}\cdot\boldsymbol{\mathcal{E}}$. This is exactly the rate at which energy is injected into the electron gas. And it is the rate at which energy must be removed through the inelasticity of collisions between electrons and scatterers if the temperature is to remain constant.

We may now present a reasonably approximate solution to (38). In this solution we have set $\nabla\cdot\boldsymbol{\mathcal{E}}=0$ for the reason of continuity of current flow ($\nabla\cdot\mathbf{j}=0$). We omit the integral term of the driving function F of (38) on the grounds that it is very small compared to the other terms since $w_0(k,q)-w_0(q,k)\approx 0$. This yields

$$G_0(k,\mathbf{r})=\tau_0 e^2\left[\tau_1\frac{1}{\hbar^2}\left(\frac{\partial f_0}{\partial E}\right)(\boldsymbol{\mathcal{E}}\cdot\nabla_k)^2 E + \frac{1}{3}\mathcal{E}^2 v^2\frac{\partial}{\partial E}\left(\tau_1\frac{\partial f_0}{\partial E}\right)\right]. \tag{42}$$

In view of (42) and the form (33) for $f_2(\mathbf{k},\mathbf{r})$ we may exhibit the entire expression for $f_2(\mathbf{k},\mathbf{r})$ simply as

$$f_2(\mathbf{k},\mathbf{r})=-\tau_2\left(\mathbf{v}\cdot\nabla-\frac{e}{\hbar}\mathcal{E}\cdot\nabla_k\right)\left(e\tau_1\frac{\partial f_0}{\partial E}\boldsymbol{\mathcal{E}}\cdot\mathbf{v}\right) + (\tau_0-\tau_2)e^2\mathcal{E}^2\left[\tau_1\frac{\partial f_0}{\partial E}\frac{1}{\hbar}\frac{\partial v_\mathcal{E}}{\partial k_\mathcal{E}} + \frac{1}{3}v^2\frac{\partial}{\partial E}\left(\tau_1\frac{\partial f_0}{\partial E}\right)\right], \tag{43}$$

where the subscript $\mathcal{E}$ indicates the component of the vector in the direction of the field. The entire angular dependence is contained in the various vector product relations. The relative simplicity of the result stems from the fact that the relaxation times depend only upon l and not upon m. One relaxation time $-\tau_2$ given in (37)—describes all of the five G_{2m}.

It should be remarked here that the f_2 correction to the distribution function cannot give rise to any current and hence produces no measurable voltages (emf's) in an integral unit specimen. This is because the current integrand (14) contains the function $f(\mathbf{k},\mathbf{r})$ multiplied by the velocity vector. Each component of this vector

[21] S. H. Koenig, Phys. Chem. Solids 8, 227 (1959).

[22] A. F. Gibson, J. W. Granville, and E. G. S. Paige, in *Proceedings of the International Conference on Semiconductor Physics, Prague, 1960* (Academic Press Inc., New York, 1961), p. 112.

is expandable in terms of $l=1$ spherical harmonics. Hence, only the $l=1$ parts of $f(\mathbf{k},\mathbf{r})$ can produce currents and emf's. The function $f_2(\mathbf{k},\mathbf{r})$ has no $l=1$ components but $f_3(\mathbf{k},\mathbf{r})$ does.

Configurational Current Density

In fact, inspection of the forcing function for $f_3(\mathbf{k},\mathbf{r})$ shows that it may be expanded in terms of surface spherical harmonics of orders $l=1$ and $l=3$. For all orders, other than zero, no difficulty arises from the elastic scattering approximation. And with this approximation a new relaxation time $\tau_3=(\omega_0-\omega_3)^{-1}$ enters quite naturally as in the case of τ_1 and τ_2. In response to the forcing function, $f_3(\mathbf{k},\mathbf{r})$ itself will have first-order and third-order surface spherical harmonic parts. The former will consist of the $l=1$ part of the forcing function multiplied by $(-\tau_1)$ and the latter will be the $l=3$ resolution of the forcing function multiplied by $(-\tau_3)$. That the solution will, indeed, follow this prescription is clear from the outcome of the first- and second-order cases. The forcing function under discussion has the form

$$\left(\mathbf{v}\cdot\nabla-\frac{e}{\hbar}\mathcal{E}\cdot\nabla_k\right)f_2(\mathbf{k},\mathbf{r})$$
$$=\text{third-order forcing function}, \quad (44)$$

where $f_2(\mathbf{k},\mathbf{r})$ is given in (43).

Now we are not really interested in the actual solution $f_3(\mathbf{k},\mathbf{r})$, but only in the current density $\mathbf{j}_3$ which arises from it. By virtue of the $l=1$ angular dependence of $\mathbf{v}$ in the integrand, and the orthogonality of the surface spherical harmonics we know that the τ_3 or $l=3$ part of $f_3(\mathbf{k},\mathbf{r})$ will integrate out to zero in the expression for the current—Eq. (14). Hence we are assured that the correct current will be given by the following integral, even though the integrand is not quite the right expression for $f_3(\mathbf{k},\mathbf{r})$. The $l=3$ part of the integral will automatically integrate out to zero.

$$\mathbf{j}_3=\frac{e}{4\pi^3}\int\mathbf{v}\tau_1\left(\mathbf{v}\cdot\nabla-\frac{e}{\hbar}\mathcal{E}\cdot\nabla_k\right)f_2(\mathbf{k},\mathbf{r})d^3k. \quad (45)$$

This expression actually contains effects other than the configurational one. For example, a field-dependent resistivity may be extracted from it which depends upon the square of the applied field. Furthermore, Eq. (45) gives rise to current densities which are proportional to the divergence of the applied field $\nabla\cdot\mathcal{E}$ and to the gradient of the square of the field $\nabla\mathcal{E}^2$. The former is zero, to the order of approximation of this calculation, because $\nabla\cdot\mathbf{j}=0\approx\sigma\nabla\cdot\mathcal{E}$. The latter term constitutes an effect which, because of its symmetry properties, is not measurable in a configurational emf experiment.[7] There are other terms contained in (45) which also are of this nature. The only part of (45) which can give rise to a measurable "Bernoulli" effect is the one proportional to $\mathcal{E}\cdot\nabla\mathcal{E}$. This proportionality between the configurational current density $\mathbf{j}_c$ and $\mathcal{E}\cdot\nabla\mathcal{E}$ can be written in terms of the two integrals:

$$A=\frac{2}{3}\frac{e^3}{4\pi^3\hbar^4}\int\left(\frac{\partial E}{\partial k}\right)^2\left(\frac{\partial^2E}{\partial k^2}\right)\tau_1^2\tau_2\frac{\partial f_0}{\partial E}d^3k \quad (46)$$

and

$$B=\frac{1}{15}\frac{e^3}{4\pi^3\hbar^4}\int\left(\frac{\partial E}{\partial k}\right)^4\frac{1}{\tau_1\tau_2}\frac{\partial}{\partial E}(\tau_1^3\tau_2^2)\frac{\partial f_0}{\partial E}d^3k. \quad (47)$$

In terms of these two integrals the configurational field $\mathcal{E}_c$ is

$$\mathcal{E}_c=\frac{1}{\sigma}\mathbf{j}_c=-\frac{1}{\sigma}(A+B)\mathcal{E}\cdot\nabla\mathcal{E}, \quad (48)$$

from which it follows that the coefficient C of Eqs. (3) and (4) is given by

$$C=(1/2\sigma^3)(A+B). \quad (49)$$

It should be mentioned that there is a considerable amount of algebraic manipulation in the process of extracting the integrals A and B from (45). Although we limited the calculation to an isotropic parabolic dependence of the energy E on the wave number k, we have allowed an energy dependence of the τ's, and an integration by parts was performed to keep the integrand in terms of $\partial f_0/\partial E$ rather than higher derivatives of f_0. In performing the integration by parts, it must be recognized that $\partial f_0/\partial E$ goes to zero over a constant energy surface at large k.

Equations (46) through (49) effectively demonstrate the twofold object of the calculation. That is, a field dependence of the kind exhibited in (48) is, indeed, present and its strength C is given. However, the form of C is rather formidable; therefore, it is expedient now to evaluate C and write it in terms of more familiar macroscopic quantities.

Evaluation of C Integrals

We will allow for the possibility of both negative and positive carriers. This situation would obtain, in our isotropic model, only if two overlapping electronic bands were present in the medium. The upper band contains energy states which range upwards from a minimum. The minimum lies slightly below the Fermi level. We will label this the n band because it will give rise to negative carriers. The lower or p electronic band has available states ranging downward from a maximum energy E_M which lies slightly above the Fermi level. The word "slightly" here is merely meant to indicate that at the Fermi level the energy is still reasonably quadratic in k for both these bands. The curvature in k space (effective mass), however, need not be the same for the two bands.

With this overlapping-band picture we may let $N(E)$ represent the total number of electronic states available

between some zero of energy up to energy E. Then this number is the sum of the states available up to E in the n band plus those available in the p band. If N_M represents the total number of states out to E_M in the p band, and $N_p(E)$ represents the number of states in this band, counting downward in energy, between E_M and E then

$$N(E)=N_n(E)+[N_M-N_p(E)]. \tag{50}$$

Here $N_n(E)$ represents, of course, the number of states in the n band alone out to the energy E. In the foregoing no implication about occupations of states is intended. The N's merely record the presence of states.

Now each of the integrands of A and B may be written as functions of energy only. And the range of integration in k space is merely a manner of writing the density of states:

$$\int_{\substack{\text{surface}\\ E=\text{const}}} \frac{1}{4\pi^3}d^3k=\frac{dN(E)}{dE}dE =\frac{dN_n(E)}{dE}dE-\frac{dN_p(E)}{dE}dE. \tag{51}$$

Furthermore, we may take note of the relation,[23] for $KT\ll E_F$, that

$$\int\psi(E)\left(-\frac{\partial f_0}{\partial E}\right)dE=\psi(E_F)+\frac{\pi^2}{12}(KT)^2\frac{d^2\psi}{dE^2}\bigg|_{E_F} +\cdots, \tag{52}$$

where we keep only the first term in evaluating A and B. We employ the variable $E'=E_M-E$ to evaluate the integrals for the p band while using E for the n band. With this understanding the following identities are useful for p as well as for n if we replace n by p and E by E':

$$\frac{dN_n(E)}{dE}\bigg|_{E_F}=\frac{3}{2}\frac{N_n(E_F)}{E_F}. \tag{53}$$

If we represent the carrier mobility by μ then, for quadratic dispersion,

$$\mu_n=e\tau_{1n}\frac{1}{\hbar^2}\left(\frac{\partial^2E}{\partial k^2}\right)_n\bigg|_{E_F}=e\tau_{1n}\left(\frac{\partial E}{\partial k}\right)_n^2\frac{1}{2\hbar^2E_F}\bigg|_{E_F}, \tag{54}$$

where the subscript n throughout indicates that the evaluation should be done in the upper or n band. An analogous expression obtains, of course, for the p band.

In consequence of the foregoing, we may perform the A and B integrations. The result is that C may be written as the sum of two parts

$$C=C_p-C_n, \tag{55}$$

where

$$C_n=\frac{1}{\sigma^3}e\mu_n^2\tau_2N_n\left(1+\frac{1}{5}\frac{\partial\ln(\tau_1^3\tau_2^2)}{\partial\ln E}\right)\bigg|_{n,E_F}, \tag{56}$$

[23] A. H. Wilson, see Ref. 6, Sec. 1.7.

and C_p looks exactly like C_n except that all the quantities are evaluated for the Fermi level of the p band instead of for the n band and the derivative is taken with respect to $E'=E_M-E(E_F'=E_M-E_F)$ instead of E. The number $N_p=N_p(E_F)$ in the expression for C_p will refer to the total number of states between E_M and E_F (counting backwards in E). This, of course, corresponds to those states in the p band which are unoccupied by electrons. Hence, N_p is the total number of empty states in the p band or the total number of holes (positive carriers). Similarly, $N_n=N_n(E_F)$ is the total number of negative carriers. In terms of the quantities used in the foregoing, we also have the familiar relation

$$\sigma=e\mu_pN_p+e\mu_nN_n. \tag{57}$$

It is interesting to note that if only one carrier were present and the relaxation times were equal and varied as the $-\frac{1}{2}$ power of E, then Eq. (56) reduces to the original estimate in Eq. (4). This variation of τ_1 with energy is not unreasonable since $1/\tau_1$ is related to the probability per unit time that a collision increases the population of states with energy E. This probability is proportional to the density of states at E or, for our parabolic model, to $E^{1/2}$. Such an energy dependence corresponds simply to a fixed spatial mean free path model. In general, however, a variety of energy dependences of τ_1 are possible depending upon the electron scattering mechanisms.[24]

IV. CONCLUSIONS

The twofold purpose of the foregoing is embodied in Eq. (48)—where the form of the hypothesized configurational field is extracted from the transport equation—and in Eqs. (55) and (56) where the magnitude of the effect is presented. As pointed out above, the magnitude reduces to that yielded by a simple intuitive model under the appropriate limiting conditions. It is apparent from the final result that the configurational emf essentially measures properties of the relaxation time τ_2. Ordinary conductivity measures τ_1.

With regard to the approximations employed in the calculation, that of the parabolic dispersion relation is justified primarily on expediency. It is felt, however, that an allowance for crystalline anisotropy will not change the results basically, but does introduce exceedingly intractible computational complexities. Furthermore, crystalline anisotropy effects are quite distinguishable experimentally from the configurational effect. In the latter, directional properties are established externally by the impressed current or field. Crystalline anisotropy effects are locked to the specimen.

Perhaps a more serious approximational error is made

[24] C. Herring, in *Proceedings of the International Conference on Semiconductor Physics, Prague, 1960* (Academic Press Inc., New York, 1961), p. 60.

in neglecting electron-electron collisions. Although, in first order, these contribute nothing to electrical resistivity there may be a contribution to a configurational effect. The magnitude of the contribution remains to be seen. The foregoing calculation allowed only for collisions with nonelectronic scatterers.

ACKNOWLEDGMENTS

The author is indebted to Professor Richard P. Feynman for his zestful interest and his trenchant observations and comments. He also wishes to express his appreciation to Professor John R. Pellam for many helpful discussions.

PHYSICAL REVIEW VOLUME 133, NUMBER 4A 17 FEBRUARY 1964

Quantum Noise in a Parametric Amplifier with Lossy Modes

WILLIAM G. WAGNER
Hughes Research Laboratories, Malibu, California and California Institute of Technology, Pasadena, California

AND

R. W. HELLWARTH
Hughes Research Laboratories, Malibu, California

(Received 7 October 1963)

We derive the fundamental noise properties of a parametric amplifier (paramp) which is driven by a monochromatic pump but which has many lossy signal and idler modes. We use a quantum-mechanical approach similar to that employed by Louisell and co-workers in their treatment of paramps and lasers with lossless modes. Our results are not directly deducible from the lossless mode analysis, but they do give the same limiting noise temperature as that analysis. This minimum noise temperature is the same as that of an ideal laser. In fact, we find that one can always design a laser whose field correlation functions would be the same as those of any given paramp. Many results for laser noise in various configurations can be carried over directly and applied to parametric amplifiers by using correspondence substitutions which we set down. For example, in analogy to Pound's description of the maser (in terms of a Nyquist theorem extended to negative temperatures), the calculation of paramp noise is found to be equivalent to applying Nyquist's theorem to both the active and passive elements in the (classically computed) equivalent circuit of a signal mode of interest. From this one obtains the spectra of effective noise generators which, when amplified in the circuit, duplicate the "quantum" as well as thermal noise that is present in the output. The Nyquist theorem is extended to apply to each active element by simply taking the element temperatures to be minus the real temperature of the corresponding idler mode (whose parametric coupling into the signal circuit gives rise to that element) multiplied by the ratio of the pump minus the idler to the idler frequencies.

I. INTRODUCTION

SEVERAL papers have appeared recently on the quantum mechanics of parametric amplification. This aspect of the otherwise well-known parametric process has become interesting because of the development of sources of high-intensity electromagnetic fields with carrier frequencies in the infrared and optical regions. It is only in the high-frequency domain that the quantum aspects of the field should appear, and one suspects that their major effect would be to add a noise field to any signal and thus determine the lower limit for noise in a parametric amplifier (paramp).

Louisell and his co-workers,[1] in two papers, have presented a very interesting and instructive treatment of quantum fluctuations and noise in parametric processes. They have employed a model which treats the signal and idler modes as lossless, and have solved exactly the Heisenberg equations of motion for two modes coupled by a classical harmonically varying term. It is this classical treatment of the pump field which renders an exact solution of the problem possible, and we shall adopt it in our work. However, if the signal and idler modes are treated as lossless, the initial transients determine the behavior of the amplifier to a large extent. No steady state is established and the correlation properties of the noise cannot be discussed. LYS obtains an output field in the absence of any input fields, and draws the conclusion that the proper way to take into account the quantum effects inherent in the parametric process is to add to any real input noise energy initially present in the modes an effective $\frac{1}{2}$ photon into both the signal and idler modes. *A priori*, one cannot say whether this result is specific to lossless modes; it is therefore of real interest to investigate a quantum-mechanical model for parametric amplification employing lossy modes, in which model transients die out and the steady-state

[1] W. H. Louisell, A. Yariv, and A. E. Siegman, Phys. Rev. **124**, 1146 (1961); J. P. Gordon, W. H. Louisell, and L. R. Walker, *ibid.* **129**, 481 (1963). Hereafter these papers are designated by LYS and GLW, respectively.

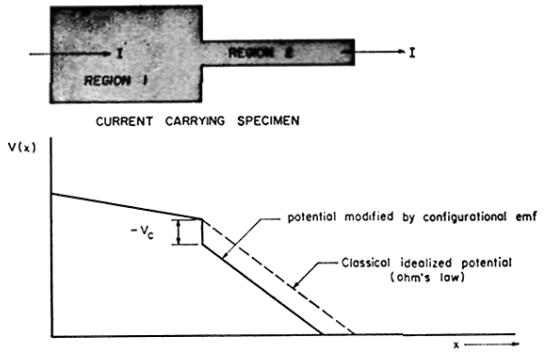

FIG. 1. Idealized electrical potential versus distance for a constricted specimen.