A THEORETICAL STUDY OF SPACE CHARGE AND HYDROMAGNETIC WAVES IN SOLIDS

by

JOHN GEORGE COOK
B.Sc., University of British Columbia, 1960

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE

in the Department
of
PHYSICS

We accept this thesis as conforming to the standard
required for candidates for the degree of
MASTER OF SCIENCE

THE UNIVERSITY OF BRITISH COLUMBIA
February, 1962

ABSTRACT

This thesis is a theoretical study of some aspects of space charge waves and hydromagnetic waves in solids. Dispersion equations obtained in the hydrodynamic approximation are studied to gain information concerning transverse waves propagating along an applied magnetic field, and the conditions for which space charge waves may grow.

For the hydromagnetic waves various assumptions are made as to the ratio of the electron and hole masses and electron and hole number densities. Particular attention is paid to the extrinsic and intrinsic cases. It is shown that often waves which are apparently different from waves previously studied, may be considered as simple extensions or special cases of the type of wave motion that are well established.

In studying growing space charge waves it is assumed that the solid is intrinsic, the hole mass equals the electron mass, and the plasma found in the solid is cold. Recombination and damping of the carriers is taken into account at all times. For this model exact conditions are given for which growth of space charge waves propagating along an applied electric field may occur.

ACKNOWLEDGEMENTS

I take pleasure in expressing my gratitude to my thesis supervisor, Professor R. E. Burgess, for advice and encouragement given during the preparation of the material in this report.

I am indebted to the National Research Council of Canada for a Studentship.

TABLE OF CONTENTS

I. INTRODUCTION.

This thesis investigates some aspects of wave motion in the gas of electrons and holes which may be found in solids. Section II is devoted to waves propagating along a constant magnetic field; section III studies space charge waves propagating in the direction of an applied electric field. For the hydromagnetic waves various assumptions are made as to the ratios of electron and hole masses and number densities; this discussion is thus valid also for gaseous plasmas. For the third section, which is essentially independent of the second, it is assumed that the holes and electrons are identical except for the sign of the charge.

We shall first study plane transverse waves propagating along an applied magnetic field. Transverse unperturbed magnetic fields and longitudinal perturbations are assumed to be non-existent. The discussion was motivated by a recent paper by T. Watanabe (1961) which seeks the frequency ranges in which undamped Alfven waves may propagate in a gas of any degree of ionization, the mass of neutral and positively charged particles being equal and large compared with the electron mass. In this thesis we shall discuss a gas which corresponds to the gas of electrons and holes in solids: neutral and charged particles with both signs are present with a mass so heavy that at the frequencies considered these particles may be considered to be immobile, and light positively charged particles and negatively charged particles are present in such numbers that overall charge neutrality is maintained. We shall not restrict ourselves to the damped or undamped Alfvén waves as commonly known, but instead seek those conditions for which the magnetic field is a dominant influence on the waves. Alfvén waves are such waves, and the oscillations in Sodium which have recently been found at low temperatures by Bowers et al (1961) are included in this scheme. We shall find that the magnetic field dependence is not the same in extrinsic (monopolar) and intrinsic (bipolar) cases.

In the third section we study a paper by Groschwitz (1957) in which a dispersion equation is derived for space charge waves which propagate in a solid along the direction of an applied electric field. The dispersion equation contains the diffusion and recombination coefficients, important in solids. In its derivation it was assumed that electrons and holes have equal mass and are present in equal numbers. The equation was solved for the complex wave number $K = K_1 + iK_2$. We show in the third section that the solutions given are not correct, and that thermal velocity fluctuations are not taken into account correctly. The Groschwitz paper is extended to a study of the conditions for "double-stream amplification" in intrinsic semiconductors for which recombination is not ignored but the velocity distribution functions are delta functions centred on the drift velocities of the electrons and holes.

Throughout this thesis dispersion equations are studied which are obtained from hydrodynamic equations using a linearization method. It is well to consider briefly the limits to the validity of this derivation. We consider the holes and electrons in the effective mass approximation. The characteristic times and distances over which macroscopic variables change (such as the period and wavelength of any wave motion present) must be large compared with the mean free time - not mean collision time - and mean free path of the collision mechanisms: the particles are "locked" together and local equilibrium is attained. The system may then be described by such macroscopic variables as mean density, mean mass flow,[1] and a parameter such as temperature describing the fluctuations about the mean. Finally a linearization method is used to make the equations more amenable to study and a superposition theorem is called upon to make the approach worthwhile. Again, it must be kept in mind that the linear approximation loses meaning if higher order terms become important as the waves grow.

II. TRANSVERSE WAVES PROPAGATING ALONG AN APPLIED MAGNETIC FIELD.

1. Introductory Comments and Organization.

On the following pages we study transverse hydrodynamic waves propagating along an external magnetic field. Two charged carrier species, electrons of charge -e and positive ions or holes as we shall call them of charge +e, and a neutral or charged background are present. Electron-hole, electron-background and hole-background scattering is taken into account, and no restrictions are placed on the hole mass nor on the electron and hole number densities. For most

1. It should be noted that a mean velocity may not be defined for low mobility semiconductors, as the mean free path may be of the order of or smaller than the interatomic spacing or the electron wavelength.

of the discussion the background medium is assumed to be immobile. Much of the work is thus valid for such plasmas as semiconductor plasmas and gaseous plasmas; the principal assumption is the validity of the hydrodynamic approximation.

The discussion is an extension of several published papers. Hines in 1953 treated the case of an unspecified number of carrier species interacting only with a mobile neutral medium. Oster in 1960 discussed wave motion in general, devoting a few pages to the type of wave we are interested in but assumes no background is present. Watanabe in 1961 studies these waves with our assumptions except that in his treatment the background medium is neutral and mobile and the ion mass is large. Tanenbaum (1961) has discussed plane waves for any angle between the direction of propagation and the magnetic field but again the ion mass is taken to be much larger than the electron mass. As this author is interested in solid state plasmas this assumption is too restrictive and hence not made.

In section II.2. the dispersion equation is derived and briefly discussed. Section II.3. is devoted to a discussion of "general Alfvén waves", and a study of the dispersion equation for the extrinsic and near intrinsic cases. The approach we use is compared with Watanabe's approach in II.4, and a recipe for including the effects of a mobile neutral background is given in section II.5. Finally, matters for future investigation are suggested in II.6.

2. The Dispersion Equation.

Consider an ionized gas consisting of electrons of mean velocity $\vec{U}_n$, number density N_n, mass M_n and charge -e; ions $(\vec{U}_p, N_p, M_p, e)$, and a third species of particles which are immobile (more exact criteria are given in section II.5.) but may be neutral or charged. In the presence of an electric field $\vec{E}$ and magnetic field $\vec{B}$ the equations of motion then are (Watanabe (1961) and Tannenbaum (1961)):

$$\rho_n \frac{D\vec{U}_n}{Dt} + \rho_n \nu_{np} \frac{M_p}{M_n + M_p} (\vec{U}_n - \vec{U}_p) + \rho_n \nu_{nb} \vec{U}_n = -e N_n (\vec{E} + \vec{U}_n \times \vec{B}) \qquad 2.1$$

and

$$\rho_p \frac{D\vec{U}_p}{Dt} + \rho_p \nu_{pn} \frac{M_n}{M_n + M_p} (\vec{U}_p - \vec{U}_n) + \rho_p \nu_{pb} \vec{U}_p = e N_p (\vec{E} + \vec{U}_p \times \vec{B}) \qquad 2.2$$

The collision frequency of an electron with the holes is denoted by ν_{np}; it is assumed to be constant. $\nu_{np}, \nu_{nb}, \nu_{pb}$ are defined similarly; $\rho_n = N_n M_n$, $\rho_p = N_p M_p$. The Lagrangian time derivative is denoted by $\frac{D}{Dt}$.

Due to conservation of momentum $\nu_{np} = \frac{N_p}{N_n} \nu_{pn}$. Define $\beta = \frac{N_p}{N_n}$, and $\alpha = \frac{M_p}{M_n}$.

Let plane waves propagate along the z- axis which lies along the primary magnetic field $\vec{B}_o$. Let no primary electric field be present, and assume $\vec{B} - \vec{B}_o$ and all other vectors lie in the x-y plane and vary as exp $(i(\omega t - kz))$. By standard perturbation techniques it follows (see Appendix A) from equations 2.1, 2.2, and the Maxwell equations that

ω and k obey

$$\frac{\mu\varepsilon\omega_{en}^2}{\omega\left(\frac{k^2}{\omega^2}-\mu\varepsilon\right)} = - \frac{(\omega\pm\omega_p-i\nu_{pb})(\omega\pm\omega_n-i\nu_{nb})-i\frac{\alpha}{1+\alpha}\left[(\omega\pm\omega_p-i\nu_{pb})+\frac{1}{\alpha\beta}(\omega\pm\omega_n-i\nu_{nb})\right]\nu_{np}}{\omega\pm\omega_p-i\nu_{pb}+\frac{\beta}{\alpha}(\omega\pm\omega_n-i\nu_{nb})-i\,\frac{(\beta-1)^2}{\alpha\beta}\,\frac{\alpha}{1+\alpha}\,\nu_{np}} \qquad 2.3$$

The signs are coupled. The lower signs give the Ordinary (O) wave (the field vectors rotate with the holes), the upper signs give the Extraordinary (E) wave (rotation with the electrons). This convention will be followed throughout the following discussion. These waves are very familiar from ionospheric studies.[2] The plasma frequency ω_{ec} for the complete ionized gas has been defined as $\sqrt{\left(\frac{N_n}{M_n}+\frac{N_p}{M_p}\right)\frac{e^2}{\varepsilon}}$. The plasma frequency ω_{en} for the electrons has been defined as $\sqrt{\frac{N_n e^2}{M_n\varepsilon}}$, and similarly for ω_{ep} . The permittivity ε and permeability μ will be assumed to be real. The gyrofrequencies ω_n and ω_p are defined as $-\frac{eB_0}{M_n}$ and $\frac{eB_0}{M_p}$ respectively.

The magnetic field affects both numerator and denominator of the right-hand side. The denominator and coefficient of ν_{np} in the numerator involve a term in B_0 which vanishes as $\beta=\frac{N_p}{N_n}$ tends to one. Generally the numerator includes a term in B_0^2 , but for fixed $\alpha=\frac{M_p}{M_n}$ and large enough β (or small enough β) the numerator and denominator have a common factor and only a B_0 dependence remains. Varying α for β different from one has a similar effect.

2. Any book on the physics of the ionosphere may be consulted. A very detailed study is given by Ratcliffe in his book on the magneto-ionic theory (1959). In much of the discussion he assumes the ion mass is large compared to the electron mass.

These matters are discussed in greater detail on pages following.

3. Discussion of the dispersion equation.

(i) General Alfvén Waves.

Usually Alfvén waves are considered to be transverse waves propagating through a two-carrier gas along the direction of an applied magnetic field such that the equations

$$\rho_c \frac{\partial \vec{u}_c}{\partial t} = \vec{J} \times \vec{B}_o \qquad 3.1$$

and

$$\vec{E} + \vec{u}_c \times \vec{B}_o = 0 \qquad 3.2$$

are valid (Watanabe (1961)). The mean mass density and mean velocity are defined as $\rho_c = \rho_n + \rho_p$, $\vec{u}_c = \dfrac{\rho_n \vec{u}_n + \rho_p \vec{u}_p}{\rho_n + \rho_p}$.

For these waves the phase velocity is given by

$$\frac{\omega}{K} = \frac{B_o}{\sqrt{\mu N \rho_c}} ,$$

if this velocity is small compared to the velocity of light.

In general

$$\frac{1}{\frac{K^2}{\omega^2} - \mu\epsilon} = \frac{B_o^2}{\mu N \rho_c} \qquad 3.3$$

In the following sections the descriptions "Alfvén Waves" shall be restricted to these waves. They are one form of a more general type of wave which is essentially governed by the applied magnetic field. For example, consider those waves which satisfy the dispersion equation (III.2.):

$$\frac{\mu\epsilon}{\frac{K^2}{\omega^2} - \mu\epsilon} = -\frac{2\omega(\omega \pm \omega_n - i\nu_{nb})}{\omega_{en}^2}$$

These waves are not essentially dependent on the magnetic field for $\omega \gtrsim |\omega_n|$ or $\nu_{nb} \gtrsim |\omega_n|$, and even if $\omega << |\omega_n|$ and $\nu_{nb} << |\omega_n|$ the right hand side varies not with B_0^2 , but B_0. However, in this discussion a sufficient condition for these waves to be in a class called "General Alfvén Waves" shall be that $\omega << |\omega_n|$. These waves may be heavily damped.

This convention is suggested as plasma physics is overburdened with technical terms and descriptions.

(ii) The Extrinsic Case.

If $\beta << 1$, $\alpha\beta << 1$, $\beta << \alpha$, $\frac{\beta}{\alpha}\nu_{nb} << \nu_{pb}$, and $\frac{1}{\alpha\beta}\nu_{nb} >> \nu_{pb}$ the holes cease to exert any influence[3] and the dispersion equation becomes the familiar Appleton-Hartree formula for a purely longitudinal magnetic field:

$$\frac{\mu\epsilon}{\frac{K^2}{\omega^2} - \mu\epsilon} = - \frac{2\omega(\omega \pm \omega_n - i\nu_{nb})}{\omega_{en}^2} \qquad 3.4$$

If the inequalities given are reversed a similar equation for the holes is of course found. This equation has been studied in great detail (see Ratcliffe (1959) and Oster (1960), for example). We briefly discuss some of the important features.

(a) High Frequency Waves.

If $\omega >> |\omega_n|, \nu_{nb}$ equation 3.4 yields the familiar equation

$$\omega^2 = \frac{1}{2}\,\omega_{en}^2 + K^2c^2,$$

putting $c^2 = 1/\mu\epsilon$. The group velocity is less than c and the product of the group and phase velocities is exactly equal

3. These inequalities may be unnecessarily restrictive. At high frequencies, for example, the collision frequencies are not important.

to c^2 , a characteristic of high-frequency radio waves.

(b) General Alfvén Waves.

For $|\omega_n| >> \omega$ we have

$$\frac{\mu\varepsilon}{\frac{K^2}{\omega^2}-\mu\varepsilon} = -\frac{2\omega(\pm\omega_n - i\nu_{nb})}{\omega_{en}^2}$$

which describes general Alfvén waves. If $\nu_{nb} \lesssim \omega$, ν_{nb} should also be ignored. We may demand that ω be real, thus restricting K^2 , or we may, depending on the problem, demand that K be real. If ω is real and $K = K_1 + i K_2$, then $K_1 K_2$ and ω have opposite sign and

$$\frac{2K_1K_2}{K_1^2 - K_2^2} = -\frac{\omega_{en}^2 \nu_{nb}}{\pm\omega_{en}^2|\omega_n| + \omega(\nu_{nb}^2 + \omega_n^2)}$$

Thus the waves are damped and K tends to be real or imaginary as $|\omega_n|/\nu_{nb}$ becomes large. This may be readily ascertained from 3.5; ignoring ν_{nb} it becomes .

$$\omega\left(|\omega_n| \pm \tfrac{1}{2}\omega_{en}^2\right) = |\omega_n| K^2$$

Let ω be positive and real. Then for the E-wave K^2 is positive and hence K is real; for the O-wave K is real for $\omega > \omega_{en}^2/2|\omega_n|$ and imaginary for $\omega < \omega_{en}^2/2|\omega_n|$. One of the latter cases may not be possible. If $\omega_{en} \sim \nu_{nb}$ and $\omega >> \nu_{nb}$ for example, we cannot have

$$\nu_{nb} << |\omega_n| < \frac{\omega_{en}^2}{2\omega} \sim \frac{\nu_{nb}}{2}\frac{\nu_{nb}}{\omega} << \frac{1}{2}\nu_{nb}$$

Hence K must be real.

For low frequencies (specifically, $|\omega(\pm\omega_n + i\nu_{nb})| << \frac{1}{2}\omega_{en}^2$) equation 3.4 yields

$$\frac{1}{2}\mu\epsilon\omega\omega_{en}^2 = K^2(\pm|\omega_n| + i\nu_{nb})$$

and, for real ω ,

$$\frac{2K_1K_2}{K_1^2 - K_2^2} = \mp \frac{\nu_{nb}}{|\omega_n|} \; .$$

If $\nu_{nb}/|\omega_n| >> 1$ then K is imaginary for the O-wave . For real K and $\omega = \omega_1 + i\omega_2$, ω_2 is positive and

$$\frac{\omega_2}{\omega_1} = \pm \frac{\nu_{nb}}{|\omega_n|} \qquad 3.6 \; ,$$

$$\omega_1 = \pm \frac{B_0}{\mu N_n e} K^2 \qquad 3.7$$

Equation 3.7 has recently been derived by Bowers et al (1961) who find that it gives good agreement with a new oscillatory effect in Sodium which they found experimentally at low temperatures. These authors do not discuss damping theoretically. Equation 3.6 predicts damping for the Bowers experiment which is smaller by a factor of ten than the damping actually observed. The discrepancy is probably due to the fact that we have treated media infinite in extent.

(c) Heavily damped waves.

For $|\omega_n|, \omega << \nu_{nb}$,

$$\frac{\mu\epsilon}{\frac{K^2}{\omega^2} - \mu\epsilon} = \frac{2i\omega\nu_{nb}}{\omega_{en}^2} \, .$$

If ω is real and $\omega_{en}^2/\omega\nu_{nb} << 1$, K_1^2 and K_2^2 both tend to

$$\frac{\mu\epsilon|\omega|\omega_{en}^2}{2\nu_{nb}} \; .$$

Let us consider what we have found for the extrinsic case. In general, the Appleton-Hartree formula is valid. We have the familiar behaviour of the high frequency waves and, for general Alfvén waves $(\omega \ll |\omega_n|)$ the O-wave may only propagate under severely restricted conditions. For very low frequencies the Bowers formula is found. Thus the oscillatory effect in Sodium is a special case of the type of wave originally discussed for the rare ionospheric gases; it may also be described as a low frequency general Alfvén wave.

(iii) The Near Intrinsic Case.

For $\beta = 1$ and $\frac{(\beta-1)^2}{\beta+\alpha\beta} \nu_{np} \ll | \omega - i\nu_{pb} + \frac{1}{\alpha}(\omega - i\nu_{nb}) |$, the dispersion equation is

$$\frac{\mu\epsilon}{\frac{K^2}{\omega^2} - \mu\epsilon} = \frac{2\omega}{\omega_{en}^2}\left[i\frac{\alpha}{1+\alpha}\nu_{np} - \frac{(\omega \pm \omega_n - i\nu_{nb})(\omega \pm \omega_p - i\nu_{pb})}{\omega - i\nu_{nb} + \frac{1}{\alpha}(\omega - i\nu_{pb})} \right] \qquad 3.8$$

Again, as in the following, the signs are coupled. For $\alpha = \frac{M_p}{M_n} > 1$ we must certainly have $|\omega_n| \gg \omega$ for the waves to be general Alfvén waves. If $\omega \gg \nu_{nb}, \nu_{pb}$,

$$\frac{\mu\epsilon}{\frac{K^2}{\omega^2} - \mu\epsilon} = \frac{2\omega}{\omega_{ec}^2}\left[i\nu_{np} \mp \frac{\omega_n(\omega \pm \omega_p)}{\omega} \right] \qquad 3.9$$

which are general Alfvén waves without further restrictions. In equation 3.9 and 3.8 as elsewhere, the lower signs correspond to the O-wave, the upper signs to the E-wave. If we have not high frequencies but $\omega \ll \nu_{nb}, \nu_{pb}$,

$$\frac{\mu\epsilon}{\frac{K^2}{\omega^2} - \mu\epsilon} = \frac{2\omega}{\omega_{en}^2}\left[i\frac{\alpha}{1+\alpha}\nu_{np} - \frac{\pm\omega_n \nu_{pb} + (\omega \pm \omega_p)\nu_{nb} + i[\pm\omega_n(\omega \pm \omega_p) - \nu_{nb}\nu_{pb}]}{\nu_{pb} + \frac{1}{\alpha}\nu_{nb}} \right]$$

These are general Alfvén waves if $\omega_p >> \omega$ or $|\omega_n|\nu_{pb} >> \omega\nu_{nb}$. Thus, approximately $\omega << \omega_p$; then

$$\frac{\mu\epsilon}{\frac{K^2}{\omega^2}-\mu\epsilon} = \frac{2\omega}{\omega_{ec}^2}\left[\frac{\pm\omega_n(\nu_{pb}-\frac{1}{\alpha}\nu_{nb})}{\nu_{pb}+\frac{1}{\alpha}\nu_{nb}} + i\left(\frac{\alpha}{1+\alpha}\nu_{np} + \frac{\nu_{pb}\nu_{nb}-\omega_n\omega_p}{\nu_{pb}+\frac{1}{\alpha}\nu_{nb}}\right)\right] \qquad 3.10$$

The real part of the right-hand side of this equation - which is always smaller in magnitude than the imaginary part - varies as B_0 while the imaginary part may vary as B_0^2 if $\nu_{nb}\nu_{pb} \geq |\omega_n|\omega_p$.

Interparticle collisions play an important role if $|\omega\pm\omega_p - i\nu_{pb}| \lesssim \frac{1}{1+\alpha}\nu_{np}$ and $|\omega\pm\omega_n - i\nu_{nb}| \lesssim \frac{\alpha}{1+\alpha}\nu_{np}$. If it is completely dominant,

$$\frac{\mu\epsilon}{\frac{K^2}{\omega^2}-\mu\epsilon} = i\,\frac{\omega\nu_{np}}{\omega_{ec}^2}$$

an equation similar to that obtained for heavily damped waves in the extrinsic case.

If $|\omega\pm\omega_p - i\nu_{pb}| >> \frac{1}{1+\alpha}\nu_{np}$ and $|\omega\pm\omega_n - i\nu_{nb}| >> \frac{\alpha}{1+\alpha}\nu_{np}$, interparticle scattering may be ignored and

$$\frac{\mu\epsilon}{\frac{K^2}{\omega^2}-\mu\epsilon} = -\frac{2\omega}{\omega_{ec}^2}\,\frac{(\omega\pm\omega_n - i\nu_{nb})(\omega\pm\omega_p - i\nu_{pb})}{\omega - i\nu_{pb} + \frac{1}{\alpha}(\omega - i\nu_{nb})} \qquad 3.11$$

Let us treat this equation in more detail.

For frequencies higher than the collision and gyro-frequencies,

$$\omega^2 = \frac{1}{2}\omega_{ec}^2 + K^2c^2 \ .$$

a similar equation was obtained under the same conditions for the extrinsic case. Thus at high frequencies the gas behaves essentially as a monopolar gas, the particle mass being equal to the reduced mass.

Consider 3.11 for the case of general Alfvén waves. Then from 3.9 for the high frequency waves ($\omega \gg \nu_{nb}, \nu_{pb}$)

$$\frac{\mu\epsilon}{\frac{K^2}{\omega^2} - \mu\epsilon} = \frac{2|\omega_n|(\omega_p \pm \omega)}{\omega_{ec}^2} \qquad 3.12$$

For ω real, K is real. If $\alpha \simeq 1$ we obtain the familiar dispersion equation for Alfvén waves:

$$\frac{\mu\epsilon}{\frac{K^2}{\omega^2} - \mu\epsilon} = \frac{B_o^2}{\mu N \rho_c} \quad , \; N = N_n = N_p . \qquad 3.13$$

For low frequency waves such that in addition to $\omega \ll \nu_{nb}, \nu_{pb}$ and $\omega \ll \omega_p$, $\omega(|\omega_n|\omega_p + \nu_{nb}\nu_{pb}) \ll \omega_{en}^2(\nu_{pb} + \frac{1}{\alpha}\nu_{nb})$, 3.10 and 3.11 show that for real ω ,

$$\frac{2K_1K_2}{K_1^2 - K_2^2} = \frac{|\omega_n|\omega_p + \nu_{nb}\nu_{pb}}{\pm|\omega_n|(\nu_{nb} - \frac{1}{\alpha}\nu_{pb})}$$

For real K, $\frac{\omega_2}{\omega_1}$ equals the same expression. Ignoring the trivial case $\nu_{nb} = \frac{1}{\alpha}\nu_{pb}$ the right-hand side is seen to be large but decreasing with increasing magnetic field for $|\omega_n|\omega_p < \nu_{nb}\nu_{pb}$, reach a minimum for $|\omega_n|\omega_p \sim \nu_{nb}\nu_{pb}$, and then increase with increasing magnetic field.

——— : ———

Thus for very high frequencies or collision frequencies larger than the wave frequency and gyro frequencies the behaviour of the bipolar gas is not more complicated than the behaviour of the monopolar gas. In general the behaviour is more complex, the magnetic field not only appearing in the real part of the right-hand side of the dispersion equation, as in the case of Alfvén waves, but also in the imaginary part. The real part may vary both as B_o^2 (3.13) but also as B_o (3.10, or 3.12).

4. Some Notes on Watanabe's approach.

In section II.3. the dispersion equation has been used to study the behaviour of the waves. As Watanabe (1961) has shown, the equations of motion may be used directly to shed some light on at least the Alfvén waves.

From 2.1 and 2.2,

$$\rho_c \frac{D\vec{u}_c}{Dt} + \rho_n \nu_{nb} \vec{u}_n + \rho_p \nu_{pb} \vec{u}_p = e(\beta-1) N_n \vec{E} + \vec{J} \times \vec{B}_o ,$$

$$\frac{D\vec{J}}{Dt} + \nu_{np} \vec{J} + e N_n (\beta \nu_{pb} \vec{u}_p - \nu_{nb} \vec{u}_n) =$$

$$\omega_{en}^2 \left(1 + \frac{\beta}{\alpha}\right)(\vec{E} + \vec{u}_c \times \vec{B}_o) - \omega_{en}^2 \frac{\beta(\alpha - 1/\alpha)}{1+\alpha\beta} (\vec{u}_p - \vec{u}_n) \times \vec{B}_o ,$$

the equation of motion for the plasma and the generalized form of Ohm's law. The conditions may be sought for which we obtain 3.1 and 3.2, the equations characterizing the undamped Alfvén waves. These are obtained approximately for example if $\alpha \simeq 1$, $\beta \simeq 1$, $\omega \gg \nu_{nb}, \nu_{pb}$ and $|i\omega + \nu_{np}| \ll \frac{\omega_n^2}{\omega}$. This agrees with II.3.

This method is more difficult to use for general α and β , however.[4] Furthermore, a study of general Alfvén waves would seem to be easier with the dispersion equation.

4. See also section II.5. and Appendix B.

5. The Background Medium.

In the equations of motion 2.1 and 2.2 a damping force per unit mass of the form $\nu_{nb}\vec{u_n}$ or $\nu_{pb}\vec{u_p}$ has been used. The cause of this damping force may be quite general; it may be due to neutral or ionized impurity scattering or phonon-scattering in solids, or it may be due to scattering between ions and heavy molecules in gaseous plasmas. Even though the model we use is somewhat ideal, such a damping term is used for a great variety of scattering mechanisms.

It is evident from the papers of Hines (1953) and Watanabe (1961) that for Alfvén waves, which are a special case of the waves we study, the mass of the background particles is very important. From Watanabe's table I, for example, it may be seen that many frequency ranges exist for which Alfven waves may propagate in highly ionized gas ($\frac{\rho_c}{\rho_b}$ large). Only for high frequencies may the effect of the finite mass of the background particles be ignored.

This behaviour could be studied in solids also. The mass of neutral impurities could be introduced, or also the mass of ionized impurities, although the latter would create more difficulties as the charge of the background particles would have to be introduced. Phonon scattering seems to present difficulties. Shockley (1951) has shown that as far as energy and momentum transfer are concerned the scattering between electrons and acoustical phonons is equivalent to the scattering between two gases of hard spheres, one species of mass M_n, the other species of mass $\frac{KT}{S^2}$, S the velocity of sound for the mode considered. The number density of the heavy spheres may not be determined however, as the mean free path for hard sphere scattering - which must be compared to the actual mean free path for electron-acoustical phonon scattering - is a function of both this density and the cross-sectional area of the spheres.

It is a simple matter to introduce the mass of neutral background particles (density N_b, mass M_b) into the dispersion equation. We only substitute the expressions

$$N_n \nu_{nb} \frac{M_n M_b}{M_n + M_b} (\vec{u}_n - \vec{u}_b) \quad \text{for} \quad N_n M_n \nu_{nb} \vec{u}_n ,$$

$$N_p \nu_{pb} \frac{M_p M_b}{M_p + M_b} (\vec{u}_p - \vec{u}_b) \quad \text{for} \quad N_p M_p \nu_{pb} \vec{u}_p$$

in the equations of motion, and in addition use the equation of motion for the neutral particles:

$$\rho_b \frac{D\vec{u}_b}{Dt} + N_b \nu_{bn} \frac{M_b M_n}{M_b + M_n} (\vec{u}_b - \vec{u}_n) + N_b \nu_{bp} \frac{M_p M_b}{M_b + M_p} (\vec{u}_b - \vec{u}_p) = 0 .$$

$\vec{u}_b$ may then be eliminated from 3.1 and 3.2. It is found that

$$\frac{\alpha}{1+\alpha} \nu_{np} + \frac{\rho_n \nu_{nb} \nu_{pb}}{i\omega\rho_b + \rho_n \nu_{nb} + \rho_p \nu_{pb}} \quad \text{must be substituted for } \frac{\alpha}{1+\alpha}\nu_{np},$$

$$\frac{i\rho_b \nu_{nb}\, \omega}{i\omega\rho_b + \rho_n \nu_{nb} + \rho_b \nu_{pb}} \quad \text{for } \nu_{nb} ,$$

and

$$\frac{i\rho_b \nu_{pb}\, \omega}{i\omega\rho_b + \rho_n \nu_{nb} + \rho_b \nu_{pb}} \quad \text{for } \nu_{pb} .$$

ν_{nb}/ν_{bn} and ν_{pb}/ν_{bp} obey a relation similar to that given for ν_{np} and ν_{pn}.

From these substitutions criteria for the validity of ignoring effects due to the finite value of ρ_b may be obtained. The frequency is of prime importance as is to be expected. It must be noted that N_p must equal N_n.

These substitutions were carried out and the resultant equation compared with equation 3.11 of Watanabe (1961) and equation 19 of Hines (1953) for this condition under which these equations are valid. Complete agreement with Watanabe was found; as the algebra became involved consistency with Hines was tested only for several limiting cases.

The equations which Watanabe uses to discuss Alfvén waves may be readily extended so that they are valid for any value of the ratio $\frac{M_p}{M_n}$. The resultant equations are given for completeness in Appendix B. These equations may then be studied directly, following Watanabe's example closely. This was not done in general; it was only found that for $M_p \sim M_n$ the conditions for Alfvén waves to exist in weakly ionized plasmas are almost exactly those given by Watanabe for $M_p >> M_n$.

6. Discussion.

Let us summarize the results of the preceding pages. Although a dispersion equation for transverse waves propagating along an applied magnetic field is rarely given for general values of $\alpha = \frac{M_p}{M_n}$ and $\beta = \frac{N_p}{N_n}$ (the paper by Hines (1953) is an exception; he ignores electron-ion scattering), it was found that this dispersion equation is readily derived. We then treated this equation for the two cases $N_p \approx N_n$ and $N_p << N_n$, with $\alpha \sim 1$.

The extrinsic case ($N_p << N_n$) has been previously studied in the literature. It is interesting that the classic Appleton-Hartree dispersion equation which has been used extensively in ionospheric studies was seen to agree closely with the oscillatory effect that Bowers and his fellow workers discovered recently in sodium at low temperatures. As the ionospheric gaseous plasma and the plasma found in sodium are at first sight quite different, this agreement is striking. It was shown that although the sodium oscillations are not Alfvén waves as these are usually thought of, the similarities between these two wave motions are such that it is too restrictive to not describe the sodium waves as "general Alfvén waves".

The intrinsic case was found to be in general more complex. The inclusion of electron-hole scattering did not lead to any different type of wave motion. The dependence of the waves on the magnetic field was found to be quite different for different cases; for one frequency range it could be as that of the Alfvén waves (B_0^2 dependence) while at other frequencies the variation could be with B_0 as for the sodium oscillations.

There are several topics which could be studied in more detail. Boundary conditions could be introduced, although Sturrock (1958) points out that little new information would be introduced in so doing. The effect of the lattice upon the waves could be investigated. One method would be to use the dispersion equation with the mass of the background particles included, or the equations of Appendix C. A second method would be to treat the electron-lattice interaction as the neutron-lattice interaction has been dealt with in scattering or thermalization studies. Finally, the behaviour of the Ordinary and Extraordinary waves could be studied for the waves in the intrinsic case (with $\alpha \sim 1$) as it has been studied in the literature for the extrinsic case.

III. ON GROWING SPACE CHARGE WAVES IN SOLIDS.

1. Introductory Comments and Organization.

The following pages shall deal with unstable longitudinal space charge waves in solids. The dispersion equation we use, which has been previously derived by Groschwitz (1957), includes the effects of damping, recombination, and, supposedly, diffusion.

The word "unstable" requires careful definition. Sturrock (1960) has recently shown that the distinction between amplifying and evanescent waves (the latter are familiar from optics) requires some attention. In this thesis monochromatic waves are considered, and the term "growing wave" shall be used for a wave which has real frequency ω but complex wave number k such that the real and imaginary parts of k are of the same sign. Thus, putting $k = k_1 + i k_2$ in $\exp\,(i(\omega t - k z))$ and $z = \frac{\omega}{k_1} t$, we obtain $\exp \frac{k_2 \omega}{k_1} t$ which shows an instability for $\frac{k_2 \omega}{k_1} > 0$ or $k_2 k_1 \omega > 0$. This criterion avoids the pitfall of calling an attenuating wave a growing wave from the reverse direction. We shall assume throughout the discussion with only few exceptions that the frequency is real, i.e. the waves are excited with a constant amplitude at any given position.

We first determine the conditions for which waves obey this criterion while diffusion may be ignored. We shall find some errors in the Groschwitz analysis which, although only involving a sign, indicate that growing waves exist at all times. We then discuss the case for which diffusion is supposedly taken care of, and show that growing waves are indicated in the absence of any applied fields. The reason for this contradictory result is that spatial density gradients are not taken account of correctly in the Groschwitz equations. This matter is briefly dealt with and finally a paper discussing instabilities in semiconductors using the Boltzmann equation is discussed.

2. Derivation of the Dispersion Equation.

The dispersion equation we shall analyze has been derived by Groschwitz in 1957. He proceeds from the equations of motion

$$\frac{D\vec{u}_n}{Dt} + \nu\vec{u}_n = -\frac{e}{m}\left(\vec{E} + \vec{u}_n \times \vec{B}_0\right) \qquad 2.1$$

$$\frac{D\vec{u}_p}{Dt} + \nu\vec{u}_p = \frac{e}{M}\left(\vec{E} + \vec{u}_p \times \vec{B}\right) \qquad 2.2$$

in which it has been assumed that the damping coefficients and masses of the electrons and holes are the same. The electric field has been defined as a macroscopic average of the fields acting on the particles, and hence a pressure term is not used explicitly. We shall see that this matter requires further discussion. The equations of continuity are

$$\frac{\partial N_p}{\partial t} + \frac{1}{e}\nabla\cdot\vec{J}_p = R\left(N_nN_p - N_i^2\right), \quad \frac{\partial N_n}{\partial t} + \frac{1}{e}\nabla\cdot\vec{J}_n = -R\left(N_nN_p - N_i^2\right),$$

R being the recombination coefficient, $\vec{J}_p = N_p e\vec{u}_p - eD\nabla N_p$,

$\vec{J}_n = -N_n e \vec{u}_n + eD\nabla n$. D is the diffusion coefficient which is like R, assumed to be the same for electrons and holes.

Let plane space charge waves propagate along the z-axis, the direction of an applied electric field $\vec{E}_o$; no applied magnetic field is present. Thus $\frac{\partial}{\partial x} = \frac{\partial}{\partial y} = 0$, $\frac{\partial}{\partial z} = -iK$, $\frac{\partial}{\partial t} = i\omega$.

A perturbation method is used as in section II to obtain the dispersion equation. The unperturbed values of N_n and N_p are assumed to equal N_i (the intrinsic case); the unperturbed drift velocities are of equal magnitude u_o and opposite in direction as $M_n = M_p = M$ and $B_o = 0$. Defining $\check{\vec{u}}_+$ as the vector sum of the perturbations in $\vec{u}_n$ and $\vec{u}_p$, $\check{\vec{u}}_+$ satisfies

$$\begin{aligned} &-u_o^2 D^2 \frac{\partial^6 \check{u}_+}{\partial z^6} + D^2 \frac{\partial^6 \check{u}_+}{\partial z^4 \partial t^2} + 2(\nu D^2 + u_o^2 D)\frac{\partial^5 \check{u}_+}{\partial z^4 \partial t} - 2D\frac{\partial^5 \check{u}_+}{\partial z^3 \partial t^2} + (\nu^2 D^2 + u_o^4 + 2RDN_i u_o^2)\frac{\partial^4 \check{u}_+}{\partial z^4} \\ &- 2(2\nu D + RDN_i + u_o^2)\frac{\partial^4 \check{u}_+}{\partial z^2 \partial t^2} + \frac{\partial^4 \check{u}_+}{\partial t^4} + (2RN_i + 2\nu)\frac{\partial^3 \check{u}_+}{\partial t^3} - 2(\nu^2 D + \nu u_o^2 + 2\nu RDN_i + \tfrac{1}{2}\omega_{ec}^2 \\ &+ RN_i u_o^2)\frac{\partial^3 \check{u}_+}{\partial z^2 \partial t} - (\nu \omega_{ec}^2 D + 2\nu^2 RDN_i - \omega_{ec}^2 u_o^2 + \nu^2 u_o^2)\frac{\partial^2 \check{u}_+}{\partial z^2} \\ &+ (\omega_{ec}^2 + \nu^2 + 4\nu RN_i)\frac{\partial^2 \check{u}_+}{\partial t^2} + (\nu\omega_{ec}^2 + 2\nu^2 RN_i + 2\omega_{ec}^2 RN_i)\frac{\partial \check{u}_+}{\partial t} + 2\omega_{ec}^2 \nu RN_i \check{u}_+ = 0 \end{aligned} \qquad 2.3$$

This equation differs from that given by Groschwitz in two coefficients. Letting $\check{u}_+$ vary as $\exp(i(\omega t - kz))$, the dispersion equation follows:

$$c_3 K^6 + (c_2 + i d_2)K^4 + (c_1 + i d_1)K^2 + c_0 + i d_0 = 0 \qquad 2.4$$

where

$$c_3 = D^2 u_o^2$$

$$c_2 = D^2(\nu^2 - \omega^2) + u_o^2(2RDN_i + u_o^2)$$

$$d_2 = 2\omega D(\nu D + u_o^2)$$

$$c_1 = (\nu^2 - \omega^2)(2RDN_i + u_o^2) + \nu D(\omega_{ec}^2 - 4\omega^2) - u_o^2(\omega^2 + \omega_{ec}^2)$$

$$d_1 = \omega D(\omega_{ec}^2 - 2\omega^2) + 2\omega u_o^2 (RN_i + \nu) + 2\nu\omega D(2RN_i + \nu)$$
$$c_o = 2\nu RN_i(\omega_{ec}^2 - 2\omega^2) - \omega^2(\nu^2 + \omega_{ec}^2 - \omega^2)$$
$$d_o = 2\omega RN_i(\nu^2 + \omega_{ec}^2 - \omega^2) + \omega\nu(\omega_{ec}^2 - 2\omega^2)$$

We have used the plasma frequency $\omega_{ec} = \sqrt{\frac{2N_i e^2}{M\varepsilon}}$ instead of $\omega_{en} = \omega_{ep} = \sqrt{\frac{Ne^2}{M\varepsilon}}$ which Groschwitz uses.

As this equation involves only powers of K^2 it is reversible in space. It is reversible in time only if $D=0$ and $\nu=0$. As a result of the exponential variation assumed for the perturbations the complex conjugates of ω and K may be inserted for ω and K simultaneously but not separately. If and only if $D \neq 0$ and $u_o \neq 0$ is the equation a cubic in K^2; if one of D and u_o is zero it is reduced to a quadratic equation. If u_o and D equal zero only c_o and d_o are non-zero and

$$(\omega - 2iRN_i)(\omega - i\nu)[\omega(\omega - i\nu) - \omega_{ec}^2] = 0$$

Thus the field vectors rotate over the whole space with $\omega(\omega - i\nu) = \omega_{ec}^2$, which for $\omega >> \nu$ yields the familiar Tonks and Langmuir result $\omega^2 = \omega_{ec}^2$ (Oster (1960)). It should be stressed that not even in this simple equation may damping be taken into account by replacing ω in the collision-free dispersion equation by $\omega - i\nu$.

3. Analysis of the Dispersion Equation.

Let us seek the conditions such that at least one root of the dispersion equation is growing, or the boundary between growing and attenuating waves. We shall first study the dispersion equation when one of E_o and D vanish, and then briefly when neither vanishes.

(i) Non-Zero Drift Velocity, Diffusion Neglected.

Consider the case $E_0 \neq 0$ but $D = 0$. Let us in addition put $\nu = 0$ and $R = 0$ for a preliminary study. For these undamped waves the dispersion equation is

$$u_0^4 K^4 - (2\omega^2 + \omega_{ec}^2) u_0^2 K^2 + \omega^2(\omega^2 - \omega_{ec}^2) = 0$$

or, equivalently,

$$\frac{\omega_{en}^2}{(\omega + u_0 K)^2} + \frac{\omega_{ep}^2}{(\omega - u_0 K)^2} = 1 \qquad 3.1$$

This equation is familiar from the theory of double stream amplification. The frequency and wavelength may be readily solved for; we obtain

$$u_0^2 K^2 = \tfrac{1}{2}\omega_{ec}^2 + \omega^2 \pm \tfrac{1}{2}\sqrt{8\omega^2\omega_{ec}^2 + \omega_{ec}^4}$$

$$2\omega^2 = \omega_{ec}^2 + 2u_0^2 K^2 \pm \sqrt{8u_0^2 K^2 \omega_{ec}^2 + \omega_{ec}^4}$$

From the first equation it follows that for $\omega > \omega_{ec}$, K is real for all four waves, but for $\omega < \omega_{ec}$ K is real for two waves and imaginary for the other two. If it is desired to keep K real it follows from the second equation that for $u_0 < \left|\frac{\omega_{ec}}{K}\right|$ two solutions in ω are real and two are imaginary; for $u_0 > \left|\frac{\omega_{ec}}{K}\right|$ all four waves are real. This type of evanescent but undamped wave is also found in lossless waveguides (for which the cut-off frequency is a function of the waveguide dimensions). The analogy may not be carried too far as we have an applied electric field and longitudinal waves.

These four waves have been the subject of much discussion in the literature. The dispersion equation has been used to study such devices as the travelling wave tube, but several errors have been made as to the exact waves which exhibit amplification. The situation has recently been clarified by Swift-Hook (1960) who finds the pair of solutions in ω for real K which may lead to amplification.

If the effects of damping but not recombination are included in 3.1 it becomes

$$\frac{\omega_{en}^2}{(u_oK+\omega)[u_oK+(\omega-i\nu)]}+\frac{\omega_{ep}^2}{(u_oK-\omega)[u_oK-(\omega-i\nu)]}=1 \qquad 3.2$$

Let us discuss this equation more fully with in addition the effects of recombination included. If only diffusion is ignored - the plasma is thus "cold" - we have

$$c_2K^4+(c_1+id_1)K^2+c_o+id_o=0 \qquad 3.3$$

a quadratic in K^2. Thus two roots in K^2 exist, and for each root K^2 two roots in K. The values of $K_1^2-K_2^2$ and K_1K_2 occur in pairs; growth or attenuation are displayed by an even number of roots. The roots in K may readily be found, but as we shall obtain results different from those of Groschwitz the method for solving 3.3 is given in appendix C.

The equation of interest is equation 4 of appendix C:

$$4c_2K_1K_2=$$

$$-d_1\pm\frac{1}{\sqrt{2}}\sqrt{-c_1^2+d_1^2+4c_oc_2+\sqrt{(c_1^2-d_1^2-4c_oc_2)^2+(2c_1d_1-4c_2d_o)^2}} \qquad 3.5$$

Define the right-hand side as $-d_1 \pm Z$. As d_1 has the same sign as ω and c_2 and D are positive, we may without loss of generality assume $\omega > 0$ (henceforth assume ω real). Hence two modes are attenuating at all times, and two modes are growing if and only if $Z > d_1$. Or,

$$\sqrt{(c_1^2 - d_1^2 - 4c_0c_2)^2 + (2c_1d_1 - 4c_2d_0)^2} > c_1^2 - 4c_0c_2 + d_1^2$$

Consequently two modes are growing if

$$4c_0c_2 > c_1^2 + d_1^2 \qquad 3.6$$

or, if 3.6 does not hold,

$$c_2^2 d_0^2 + c_0c_2d_1^2 > c_1c_2d_0d_1 \qquad 3.7$$

Substituting the values the c's and d's take on,

$$U_0^4 [(\nu^2 - \omega_{ec}^2)^2 + 4\omega^2\nu^2 + 8\omega^2\omega_{ec}^2 + RN_i\nu(24\omega^2 - 8\omega_{ec}^2) + 4\omega^2R^2N_i^2] < 0 \qquad 3.8$$

and

$$U_0^4 [2R^3N_i^3\nu(\omega_{ec}^2 - 2\omega^2) + 2R^2N_i^2(\omega_{ec}^4 + 3\nu^2\omega_{ec}^2 - 4\nu^2\omega^2 - \omega^2\omega_{ec}^2) + RN_i\nu(\tfrac{5}{2}\omega_{ec}^4 - 5\omega^2\omega_{ec}^2 + \tfrac{5}{2}\nu^2\omega_{ec}^2 - 4\omega^2\nu^2 - \nu^4) + \tfrac{1}{2}\nu^2\omega_{ec}^2(\tfrac{3}{2}\omega_{ec}^2 - 4\omega^2 - \nu^2) > 0 \qquad 3.9$$

As an example let $R = 0$; then 3.8 cannot hold and 3.9 becomes

$$\nu^2\omega_{ec}^2(\tfrac{3}{2}\omega_{ec}^2 - 4\omega^2 - \nu^2) > 0$$

Consequently all four modes are attenuating if $4\omega^2 + \nu^2 > \frac{3}{2}\omega_{ec}^2$, and for $4\omega^2 + \nu^2 < \frac{3}{2}\omega_{ec}^2$ two modes are attenuating and two are growing.

If no damping occurs but recombination does occur, 3.8 cannot hold and 3.9 becomes $\omega^2 < \omega_{ec}^2$. If R tends to zero we obtain the evanescent waves previously discussed.

The conditions for which K is real or imaginary may also be obtained from equation 3.5. If $4c_2K_1K_2 = -d_1 - z$, K^2 is real if and only if $d_1 + z = 0$. Thus d_1 and z must both equal zero, as both are non-negative. As z is of the form $-V + \sqrt{V^2+W^2}$, $z = 0$ if $V = 0$ and $W = 0$; for $V<0, z \neq 0$. If $4c_2K_1K_2 = -d_1 + z$, we must have $d_1 = z$, or the inequality of 3.7 must become an equality.

Let R again equal zero, as an example. Then z may equal zero, as $V \geq 0$. Upon simplifying the inequalities given in the preceding paragraph it is found that two modes have K^2 real if $4\omega^2 + \nu^2 = \frac{3}{2}\omega_{ec}^2$, and K^2 is real for four modes if $\omega = 0$ or $\nu = 0$. It follows that all four modes have K real or imaginary on the axis of the (ω, ν) coordinate system, but two of these modes are of this nature on the axis only and are attenuating elsewhere, while two modes are real or imaginary on the axis and on the ellipse $4\omega^2 + \nu^2 = \frac{3}{2}\omega_{ec}^2$, are attenuating on the exterior of this ellipse and growing on the interior.

The situation for non-negligible recombination is similar. Two modes are always attenuating, and two modes are growing under the conditions given by 3.8 or 3.9 and the converse of 3.8. Growth in bands is not possible as 3.8 and 3.9 are both low frequency inequalities.

These results differ from those of Groschwitz, who finds that four modes are always growing. Our expressions for K_1 and K_2 (see equation 5 of appendix C) are almost identical to those of Groschwitz (ρ_j in his equation 11 should be changed to $-\rho_j$) but we disagree with the signs of K_1 and K_2. The actual signs may be obtained as in appendix C.

(ii) Zero Drift Velocity, Diffusion Not Neglected.

A somewhat disturbing result is obtained if the electric field is assumed to be completely absent, but diffusion is not ignored. We then find that, if in addition recombination may be ignored, the dispersion equation after some manipulation to make the leading coefficient real becomes

$$(\nu^2+\omega^2)D^2K^4 + D\left[\omega_{ec}^2\nu + i\omega(-\omega_{ec}^2+2\nu^2+2\omega^2)\right]K^2$$

$$+\omega^2(\omega_{ec}^2-\nu^2-\omega^2)+i\omega\nu\,\omega_{ec}^2 = 0 \qquad 3.10$$

Using the methods of section (i),

$$\frac{k_1k_2}{\omega} = \frac{-(\nu^2+\omega^2)+\frac{1}{2}\omega_{ec}^2\left(1\pm\frac{\omega}{|\omega|}\right)}{D(\nu^2+\omega^2)}$$

The two roots in K^2 correspond to the two signs possible; the two roots in K corresponding to the upper sign display attenuation for $\omega_{ec}^2 > \nu^2+\omega^2$ and growth for $\omega_{ec}^2 < \nu^2+\omega^2$. The lower sign gives $\frac{k_1k_2}{\omega} = -1/D$ at all times.

This result is contrary to expectations as in the absence of a primary electric field no growing waves are possible. It is suggested that the dispersion equations involving D (such as 3.10) are in error. This matter is discussed in III.4; we first treat the general case briefly for completeness.

(iii) The General Case.

Let us consider equation 2.2, making no restrictions on $R, D,$ or ε_0 . Although this equation could be studied to find the exact conditions for which $k_1 k_2 \omega > 0$, we shall not do so. The main reason as we have seen is that there is some doubt as to the correctness of this equation for $D \neq 0$; a second reason is that the algebra becomes quite involved. Instead of this we shall briefly discuss the conditions when K is real or imaginary.

As the left-hand side of equation 2.2 equals the product of all the terms $K^2 - K_s^2$, where K_s^2 is any solution, we find upon expanding this product that if not all three of d_0, d_1 , and d_2 are zero at least one root K_s^2 is complex. If $d_0 = d_1 = d_2 = 0$ and in addition all roots are real it must be true that (Burnside and Panton (1904), p.84)

$$(c_0 - 3c_1c_2 + 2c_2^2)^2 + 4(c_1 - c_2^2)^3 < 0 ,$$

a necessary condition. If at least one root is real, C.5 yields two equations from which we have

$$d_1^2 - d_0 d_2 \geq 0$$

and

$$d_1(d_0c_2 - c_0d_2)^2 + c_0d_1^2(c_2d_2 - d_1c_3) + d_0^2c_3(d_0c_3 - 2c_1d_2) + c_1d_2^2(d_0c_1 - c_0d_1)$$

$$+ 3c_0d_0d_1d_2c_3 + d_0d_1c_3(c_1d_1 - d_0c_2) - d_0c_1d_1c_2d_2 = 0 ,$$

again necessary conditions.

4. Discussion.

It is apparent from section 3.(ii) that there is reason to doubt the correctness of the complete dispersion equation. It is suggested that the error is due to incomplete equations of motion. It has been assumed that the pressure term, which is of course very important for longitudinal waves, may be absorbed into the alternating electric field by virtue of a collective effect or pseudo force. The effect of a concentration gradient has been at least partially taken into account through the diffusion current. It is apparent that we may not without some justification group the effect of a carrier concentration gradient or pressure under the electric field in the equation of motion, but take the effect of this gradient explicitly into account in the equation of continuity as a drift current. Equivalently, the temperature of the plasma may not be taken into account through the diffusion coefficient while it is arbitrarily left out through the pressure. A more rigorous procedure would be to use a suitable pressure term and then determining from the dispersion equation or the equations of motion themselves those conditions for which the pressure term may be ignored.

Oster has shown that the pressure term may cause the electron acoustic velocity to appear in the equations, and increase the number of modes possible. The complexity of the dispersion equation would thus be increased; six or more modes may exist in general. This matter is further discussed by Fried and Gould (1961).

The results of section 2.(i) are valid for those conditions for which the velocity fluctuations about the mean may be neglected; i.e., the plasma is "cold". This is further borne out by comparing our results and other published papers. It may be seen from equation 3.2 and 3.3 that the roots K vary inversely as u_0 : if u_0 decreases the wavelength decreases, but the amplification or attenuation factor $\frac{k_2}{k_1}$ remains fixed. The drift velocity may be reduced to any non-zero value and still amplifying waves may exist. Other authors have shown (see e.g. Pines and Schrieffer (1961) and their references) that no amplifying waves are possible for drift velocities less than a certain critical value which is a function of the mean thermal velocity. This critical velocity tends to zero as the momentum distributions of the carriers tend to delta functions.

Recombination may still occur, as instead of averaging the effect of recombination over the distribution function the recombination coefficient has that value which corresponds to the drift velocity at which the delta function has its sharp peak.

We have found the conditions for which growing waves may exist in a cold plasma. It remains to be determined which of the growing waves may serve for the amplification of an injected signal. This would perhaps be done using the criterion of Buneman which is given in J. E. Drummond's book (1960). Buneman determines those conditions for which power may be transferred to an external load by studying a dispersion equation obtained by matching admittances between the plasma and the field perturbations. This method has also been used by Swift-Hook (1960). It will not be used by this author.

The Boltzmann equation has recently been used by Pines and Schrieffer (1961) to study the possibility of observing high-frequency instabilities in InSb plasmas. The Boltzmann approach is more strictly valid, more elegant, and more powerful than the hydrodynamic approximation. These authors postulate a displaced Maxwellian distribution for both carriers and from derived dispersion equations find conditions such that the imaginary part of the frequency ω is positive, for several different electron and hole concentrations. The conditions are found for which the energy and momentum exchange for electron-electron interactions dominate the exchange for electron-lattice interactions and hence the postulate of a displaced Maxwellian distribution for the electrons is valid. Similar calculations are carried out for the hole-lattice interaction. The electron-hole interaction is not mentioned; the electron and hole gases probably exist in separate quasi-equilibria because of the large (~ 14) hole to electron mass ratio. This fact also makes InSb more suitable for observing the drift instability.

It may be added that the displaced Maxwellian carrier distribution is often postulated when instabilities or negative resistance are sought. Pines and Schrieffer have postulated it to study the twin-stream instability, which may occur when the distribution function for the composite gas departs sufficiently from the unperturbed single peaked distribution function. Adawi (1961) has shown that negative resistance cannot occur for extrinsic semiconductors for which the carrier distribution is displaced Maxwellian. This result may be interpreted as follows. According to the familiar low field theory, the conductivity varies as

$$-\int_0^\infty \varepsilon^{3/2}\,\tau(\varepsilon)\,\frac{\partial f_0}{\partial \varepsilon}\,d\varepsilon$$

where $\tau(\varepsilon)$ is the energy dependent relaxation time of the perturbations of the unperturbed carrier distribution $f_0(\varepsilon)$. Unlike other suggested forms for $f_0(\varepsilon)$ - such as the Davydov or Druyvestein distributions - the displaced Maxwellian distribution has the property that $\frac{\partial f_0}{\partial \varepsilon} > 0$ for a finite energy range. Thus Adawi has shown that $\tau(\varepsilon)$ varies with energy in such a manner that the positive contribution of $\frac{\partial f_0}{\partial \varepsilon}$ is not large enough to make the integral positive.

APPENDIX A. Derivation of the Dispersion Equation of Section II

Let $\vec{B} = \vec{B}_0 + \check{\vec{B}} \exp(i(\omega t - kz))$,
$\vec{E} = \check{\vec{E}} \exp(i(\omega t - kz))$, and similarly for $\vec{u}_n$ and $\vec{u}_p$. The perturbation symbol ($\check{}$) is used only in this appendix. Elsewhere its omission will cause no confusion. As the waves are purely transverse $\check{\vec{B}}$, $\check{\vec{E}}$, $\check{\vec{u}}_p$ and $\check{\vec{u}}_n$ lie in the x-y plane and are perpendicular to $\vec{B}_0$.

It is well-known (Oster (1960)) that under these conditions N_p and N_n are constant in time and space. For example, from the equations of continuity for semiconductors as given in section III it follows that

$$(i\omega + Dk^2)[i\omega + Dk^2 + R(N_{n_0} + N_{p_0})]$$

with R the recombination coefficient, D the diffusion coefficient and N_{n_0}, N_{p_0} the unperturbed carrier densities. This dispersion equation is completely independent.

The equations 3.1 and 3.2 may be linearized:

$$\rho_n \frac{\partial \check{\vec{u}}_n}{\partial t} + \rho_n \nu_{np} \frac{m_p}{m_n + m_p} (\check{\vec{u}}_n - \check{\vec{u}}_p) + \rho_n \nu_{nb} \check{\vec{u}}_n = -eN_n (\check{\vec{E}} + \check{\vec{u}}_n \times \vec{B}_0)$$

$$\rho_p \frac{\partial \check{\vec{u}}_p}{\partial t} + \rho_p \nu_{pn} \frac{m_n}{m_n + m_p} (\check{\vec{u}}_p - \check{\vec{u}}_n) + \rho_p \nu_{pb} \check{\vec{u}}_p = eN_p (\check{\vec{E}} + \check{\vec{u}}_p \times \vec{B}_0)$$

The assumption that $\vec{E}_0 = 0$ may be equivalently stated:
$|\vec{u}_{n_0} \times \check{\vec{B}}| << |\check{\vec{u}}_n \times \vec{B}_0|$ and $\frac{D}{Dt} \simeq \frac{\partial}{\partial t}$.

From Maxwell's equations it follows that
$\check{\vec{J}} = \check{\vec{J}}_n + \check{\vec{J}}_p = -i(\varepsilon\omega - k^2/\mu\omega)\check{\vec{E}}$. Hence the dispersion equation becomes (over)

$$\begin{vmatrix} i\omega+\nu_{nb}+\frac{\alpha}{1+\alpha}\nu_{np}+\Upsilon & -\omega_n & -\beta\Upsilon-\frac{\alpha}{1+\alpha}\nu_{np} & 0 \\ \omega_n & i\omega+\nu_{nb}+\frac{\alpha}{1+\alpha}\nu_{np}+\Upsilon & 0 & -\beta\Upsilon-\frac{\alpha}{1+\alpha}\nu_{np} \\ -\frac{\Upsilon}{\alpha}-\frac{1}{\beta+\alpha\beta}\nu_{np} & 0 & i\omega+\nu_{pb}+\frac{1}{\beta+\alpha\beta}\nu_{np}+\frac{\beta}{\alpha}\Upsilon & -\omega_p \\ 0 & -\frac{\Upsilon}{\alpha}-\frac{1}{\beta+\alpha\beta}\nu_{np} & \omega_p & i\omega+\nu_{pb}+\frac{1}{\beta+\alpha\beta}\nu_{np}+\frac{\beta}{\alpha}\Upsilon \end{vmatrix}=0$$

where $\Upsilon=\dfrac{N_n e^2}{iM_n(\varepsilon\omega-k^2/\mu\omega)}$, $\omega_n=-\dfrac{eB_0}{M_n}$, $\omega_p=\dfrac{eB_0}{M_p}=-\dfrac{1}{\alpha}\omega_n$

Expanding the determinant leads to

$$\left[\left(i\omega+\nu_{nb}+\frac{\alpha}{1+\alpha}\nu_{np}+\Upsilon\right)\left(i\omega+\nu_{pb}+\frac{1}{\beta+\alpha\beta}+\frac{\beta\Upsilon}{\alpha}\right)-\left(\frac{1}{\beta+\alpha\beta}\nu_{np}+\frac{\Upsilon}{\alpha}\right)\left(\frac{\alpha}{1+\alpha}\nu_{np}+\beta\Upsilon\right)-\omega_n\omega_p\right]^2$$

$$=\omega_n\omega_p\left[\frac{1}{\alpha}\left(i\omega+\nu_{nb}+\frac{\alpha}{1+\alpha}\nu_{np}+\Upsilon\right)-\left(i\omega+\nu_{pb}+\frac{1}{\beta+\alpha\beta}\nu_{np}+\frac{\beta\Upsilon}{\alpha}\right)\right]^2$$

The dispersion equation given is found upon taking the square root (which gives the double sign) and solving for Υ .

APPENDIX B. The Watanabe Equations for any Value of M_p/M_n.

Watanabe's formula numbers are given followed by the corrected equations. The notation is changed to conform to the notation of the present paper, which uses the notation common to solid state studies. The subscript 'b' is used to refer to variables of the background medium.

As

$\beta = 1,\ N_p = N_n \equiv N_c.$ Define $\rho = \rho_c + \rho_b$, $\vec{u} = \dfrac{\rho_c \vec{u}_c + \rho_b \vec{u}_b}{\rho_c + \rho_b}$.

Using M.K.S. units,

W2.8: $$\rho_b \frac{\partial \vec{u}_b}{\partial t} = N_c N_b (\alpha_{pb} + \alpha_{nb})(\vec{u}_c - \vec{u}_b + b\vec{J})$$

W2.9: $$\rho_c \frac{\partial \vec{u}_c}{\partial t} = -\rho_b \frac{\partial \vec{u}_b}{\partial t} + \vec{J} \times \vec{B}_o$$

W2.18: $$\frac{1}{\omega_{ec}^2} \frac{\partial \vec{J}}{\partial t} + \left[\frac{1}{\sigma} + b^2 N_c N_b (\alpha_{pb} + \alpha_{nb})\right] \vec{J} + \frac{\rho_p - \rho_n}{e N_c \rho_c} \vec{J} \times \vec{B}_o$$
$$= \vec{E} + \vec{u}_c \times \vec{B}_o + b N_c N_b (\alpha_{pb} + \alpha_{nb})(\vec{u}_b - \vec{u}_c)$$

or $$-b\sigma\rho_c \frac{\partial \vec{u}_c}{\partial t} + \rho \frac{\partial}{\partial t}\left(\frac{\vec{J}}{\omega_{ec}^2}\right) + \vec{J} = \sigma(\vec{E}^m + \vec{E}^e),$$ where

W2.4: $$\alpha_{pn} = \frac{\nu_{pn}}{N_n} \frac{M_p M_n}{M_p + M_n}$$

W2.5: $$\alpha_{nb} = \frac{\nu_{nb}}{N_b} \frac{M_n M_b}{M_n + M_b}$$

W2.6: $$\alpha_{pb} = \frac{\nu_{pb}}{N_b} \frac{M_p M_b}{M_p + M_b}$$

$$\frac{1}{\sigma} = \frac{\alpha_{pn}}{e^2} + \frac{N_b \alpha_{pb} \alpha_{nb}}{N_c e^2 (\alpha_{pb} + \alpha_{nb})}$$

$$b = \frac{\rho_n \alpha_{pb} - \rho_p \alpha_{nb}}{e N_c \rho_c (\alpha_{pb} + \alpha_{nb})}$$

$$\vec{E}^m = \vec{E} + \frac{\rho_n \vec{u}_p + \rho_p \vec{u}_n}{\rho_c} \times \vec{B}_o$$

$$\vec{E}^e = -b\rho \frac{D\vec{u}}{Dt}$$

Most of these equations are given in the references in Watanabe's paper. We add one equation and disagree in some minor points.

For $M_p = \alpha M_n$, $M_n << M_p$ but not $M_p << M_b$,

$$\rho_b \frac{D\vec{u}_b}{Dt} = \frac{2\alpha N_c M_n \nu_2}{1 + \alpha M_n / M_b} \left[\vec{u}_c - \vec{u}_b - \frac{\nu_1}{eN_c(1+\alpha)\nu_2} \vec{J} \right]$$

$$\frac{1}{\omega_{ec}^2} \frac{D\vec{J}}{Dt} + \frac{\alpha M_n}{N_c e^2} \left(\frac{\nu_{pn}}{1+\alpha} + \frac{\nu_{pb}\nu_{nb}}{2\alpha\nu_2} \right) \vec{J} + \frac{\nu_1}{eN_c(1+\alpha)\nu_2} \rho_c \frac{D\vec{u}_c}{Dt}$$

$$= \vec{E} + \frac{\vec{u}_p + \alpha\vec{u}_n}{1+\alpha} \times \vec{B}_o + \frac{\nu_1}{eN_c(1+\alpha)\nu_2} \rho \frac{D\vec{u}}{Dt}$$

W3.2: $$\vec{u}_c = B_o \frac{\rho_b}{\rho_n} \frac{-\frac{1}{1+\alpha}\frac{\nu_1}{\omega_n}\vec{J} + \left(\frac{1}{2\alpha} - i\frac{\rho_n}{\rho_b}\frac{\nu_2}{\omega}\right)\vec{J} \times \vec{1}}{\nu_2 \rho + \frac{1}{2} i\omega \left(\frac{1+\alpha}{\alpha}\right)\left(1 + \alpha \frac{M_n}{M_b}\right)\rho_b}$$

W3.3: $$\vec{u}_b = -B_o \frac{-\frac{\nu_1}{\omega_n}\vec{J} + i\frac{\nu_2}{\omega}\vec{J} \times \vec{1}}{\nu_2 \rho + \frac{1}{2} i\omega \left(\frac{1+\alpha}{\alpha}\right)\left(1 + \alpha \frac{M_n}{M_b}\right)\rho_b},$$ where

$$\vec{1} = \frac{\vec{B}_o}{B_o},$$

W2.22 $$\nu_1 = \frac{1}{2}\left[\left(1 + \alpha\frac{M_n}{M_b}\right)\nu_{nb} - \nu_{pb}\right]$$

W2.23 $$\nu_2 = \frac{1}{2\alpha}\left[\left(1 + \alpha\frac{M_n}{M_b}\right)\nu_{nb} + \alpha\nu_{pb}\right]$$

Watanabe's equations follow from these for

$M_p = M_b >> M_n$.

For $M_p << M_b$,

W2.8 $$\rho_b \frac{\partial \vec{u}_b}{\partial t} = 2\rho_n \nu_2 (\vec{u}_c - \vec{u}_b) + \frac{2\rho_n}{1+\alpha} \frac{\omega_n \nu_1}{\omega_{en}^2 B_o} \vec{J}$$

W2.18 $$\frac{1}{\omega_{en}^2}\left(\frac{\alpha}{1+\alpha} \frac{\partial \vec{J}}{\partial t} + \nu_3 \vec{J} - \frac{\alpha-1}{\alpha+1} \omega_n \vec{J} \times \vec{1} \right) = \vec{E} + \vec{u}_c \times \vec{B}_o$$

$$- B_o \frac{2\alpha}{1+\alpha} \frac{\nu_1}{\omega_n} (\vec{u}_c - \vec{u}_b)$$

W3.2 $$\vec{u}_c = B_o \frac{\rho_b}{\rho_n} \frac{-\frac{1}{1+\alpha} \frac{\nu_1}{\omega_n} \vec{J} + \left(\frac{1}{2\alpha} - i \frac{\rho_n}{\rho_b} \frac{\nu_2}{\omega}\right) \vec{J} \times \vec{1}}{\nu_2 \rho + \frac{1}{2} i \omega \frac{1+\alpha}{\alpha} \rho_b}$$

W3.3 $$\vec{u}_b = B_o \frac{\frac{\nu_1}{\omega_n} \vec{J} - i \frac{\nu_2}{\omega} \vec{J} \times \vec{1}}{\nu_2 \rho + \frac{1}{2} i \omega \frac{1+\alpha}{\alpha} \rho_b}$$

W2.22 $$\nu_1 = \frac{1}{2}(\nu_{nb} - \nu_{pb})$$

W2.23 $$\nu_2 = \frac{1}{2\alpha}(\nu_{nb} + \alpha \nu_{pb})$$

W2.24 $$\nu_3 = \frac{\alpha}{1+\alpha} \nu_{pn} + \left(1 - \frac{2}{1+\alpha}\right) \nu_{nb} + \frac{\alpha \nu_{pb} + \nu_{nb}}{(1+\alpha)^2}$$

These equations are of the same form as those of Watanabe's paper.

APPENDIX C. The Quartic Equation With Complex Coefficients.

Consider the equation

$$c_2 K^4 + (c_1 + i d_1) K^2 + c_0 + i d_0 = 0 \qquad \text{C.1}$$

Hence

$$2c_2 K^2 = -c_1 - i d_1 \pm \sqrt{c_1^2 - d_1^2 - 4c_0 c_2 + i(2c_1 d_1 - 4c_2 d_0)}$$

$$\equiv -c_1 - i d_1 \pm \sqrt{V + iW} \qquad \text{C.2}$$

Let $X + iY = \pm\sqrt{V + iW}$; as $2XY = W$,we must have:

$$W>0: \ \pm\sqrt{V+iW} = \pm\frac{1}{\sqrt{2}}\left[\sqrt{V+\sqrt{V^2+W^2}} + i\sqrt{-V+\sqrt{V^2+W^2}}\right]$$

$$W<0: \ \pm\sqrt{V+iW} = \pm\frac{1}{\sqrt{2}}\left[\sqrt{V+\sqrt{V^2+W^2}} - i\sqrt{-V+\sqrt{V^2+W^2}}\right] \qquad \text{C.3}$$

The outer sign (which we shall call δ_1) gives the two solutions K^2. If δ_1 is the same for both signs of W, then as W becomes zero and changes sign the value of $\delta_1\sqrt{V+iW}$ changes smoothly for $V>0$, but to obtain a smooth transition with $V<0$, δ_1 must change sign at $W=0$. As it is reasonable that a solution in K changes smoothly as such parameters as ω, ℓ, ν vary, we adopt the following:

	V	W	δ_1	δ_2
First solution in K^2	>0	>0	$+1$	$+1$
	>0	<0	$+1$	-1
	<0	<0	$+1$	-1
	<0	>0	-1	$+1$

	V	W	δ_1	δ_2
Second solution in K^2	>0	>0	-1	$+1$
	>0	<0	-1	-1
	<0	<0	-1	-1
	<0	>0	$+1$	-1

The inner sign in $\sqrt{V+iW}$ has been defined as δ_2. As V changes from $-|V|$ to $+|V|$ with $W>0$, the first solution must become the second and the second the first, and similarly for the reverse change.

Thus,

$$2c_2(k_1^2-k_2^2+2ik_1k_2)$$
$$=-c_1-id_1+\frac{\delta_1}{\sqrt{2}}\Big[\sqrt{c_1^2-d_1^2-4c_0c_2+\sqrt{(c_1^2-d_1^2-4c_0c_2)^2+4(c_1d_1-2c_2d_0)^2}}$$
$$+i\delta_2\sqrt{-c_1^2+d_1^2+4c_0c_2+\sqrt{(c_1^2-d_1^2-4c_0c_2)^2+4(c_1d_1-2c_2d_0)^2}} \qquad \text{C.4}$$

From this equation k_1k_2 may be obtained directly. The sign of $c_2k_1k_2$ equals the sign of

$$-d_1+\frac{\delta_1\delta_2}{\sqrt{2}}\sqrt{-c_1^2+d_1^2+4c_0c_2+\sqrt{(c_1^2-d_1^2-4c_0c_2)^2+4(c_1d_1-2c_2d_0)^2}}$$

Our convention does not alter the fact that if this expression with $\delta_1\delta_2=+1$ is positive, at least two roots have $c_2k_1k_2$ positive.

From equation C.4, which may be put in the form $K^2=P+iQ$, it follows that

$$k_1=\delta_3\frac{1}{\sqrt{2}}\sqrt{P+\sqrt{P^2+Q^2}} \qquad \text{C.5}$$

$$k_2=\delta_4\frac{1}{\sqrt{2}}\sqrt{-P+\sqrt{P^2+Q^2}} \qquad \text{C.6}$$

where δ_3 and δ_4 equal ±1 and $\delta_3\delta_4$ has the same sign as Q.

BIBLIOGRAPHY

I. Adawi, J. Appl. Phys. 32, 1101 (1961)

R. Bowers et al, Phys. Rev. Lett. 7, 339 (1961)

O. Buneman, Plasma Physics ed. J. E. Drummond (McGraw-Hill, New York, 1960) Ch. 4

W. S. Burnside and A. W. Panton, The Theory of Equations (Dublin, Hodges, and Figgis, London, 1904)

B. O. Fried and R. W. Gould, Phys. Fluids 4, 139 (1960)

E. Groschwitz, Z. Naturforsch. 12a, 529 (1957)

C. O. Hines, Proc. Cambr. Phil. Soc. 49, 299 (1953)

L. Oster, Revs. Mod. Phys. 32, 141 (1960)

D. Pines and J. R. Schrieffer, Phys. Rev. 124, 1387 (1961)

J. A. Ratcliffe, The Magneto-Ionic Theory (Cambridge Univ. Press, 1959)

W. Shockley, Bell System Tech. J. 30, 990 (1951)

P. A. Sturrock, Phys. Rev. 112, 1488 (1958)

D. T. Swift-Hook, Phys. Rev. 118, 1 (1960)

B. S. Tanenbaum (to be published)

T. Watanabe, Can. J. Phys. 39, 1197 (1961)