X. PLASMA ELECTRONICS*

Prof. L. D. Smullin
Prof. H. A. Haus
Prof. A. Bers
Prof. W. D. Getty
Prof. D. J. Rose
Prof. T. H. Dupree
Prof. L. M. Lidsky
Prof. E. P. Gyftopoulos
Dr. T. Musha
R. R. Bartsch
T. S. Brown
J. F. Clarke
J. A. Davis
F. E. Dunn
P. H. Edmonds
S. A. Evans
T. J. Fessenden
R. W. Flynn
E. T. Gerry
J. N. Hamawi
C-F. G. Hsi
C. A. Kapentanakos
B. R. Kusse
S. H. Kyong
M. A. Lieberman
M. Lind
M. D. Lubin
R. W. Moir
P. E. Newton
A. A. Offenberger
R. R. Parker
C. S. Ribbeck
E. A. Robertson
H. M. Schneider
C. E. Speck
E. Thompson
C. E. Wagner
R. N. Wallace
J. C. Woo

A. BEAM PLASMA DISCHARGE: SYSTEM A

A 100-turn, 0.1-cm^2 area diamagnetic probe is placed outside the beam generated plasma[1] with the axis of the coil parallel to the beam and DC magnetic field. A sinusoidal signal of a few tens of millivolts appears on this probe[2] at a frequency of approximately 500 kc. Current collecting probes are inserted radially into the cylindrical drift tube to collect currents at different azimuthal angles. The electron current collected by these unbiased probes appears in 0.2-μsec pulses synchronously with the diamagnetic probe signal. The time-phase difference between the two current probe signals indicates that we are observing a right-handed rotating (about the DC magnetic field) macroinstability. This phenomenon has been studied by Hartenbaum.[3] Our rotation frequencies are approximately 5 times higher than his, however. The sinusoidal oscillation on the diamagnetic probe only appears at relatively low pressures (<1.4 μ Hg) and moderate magnetic field (250-600 gauss). Normally no radiofrequency is detected with the oscillation, but when it does occur, it appears as spikes of ~0.4-μsec duration. These X-band spikes occur synchronously with the oscillation, as do x-ray pulses. The x-rays are probably due to high-energy electrons that strike the walls and probes. It appears that the x-rays striking the walls at one azimuthal angle, and the radiofrequency and electron current at that same angle all occur at the same phase of the diamagnetic probe oscillation. The RF may be radiated more, however, by the plasma striking the probes than by interactions inside the plasma. This issue has not been conclusively resolved.

At pressures greater than ~1.4 μ Hg, the RF no longer appears in "spikes," or short bursts, but is radiated continuously. It is normally approximately 50 per cent amplitude-modulated by a noiselike envelope. The intensities of this RF vs axial distance have been measured, by using loop-type RF probes and wideband crystals. A plot taken at 2 μ Hg pressure is shown in Fig. X-1. Note the sharp drop in intensity as the point of minimum magnetic field is approached. The density could be expected to have a maximum at the

*This work was supported in part by the National Science Foundation (Grant GK-57).

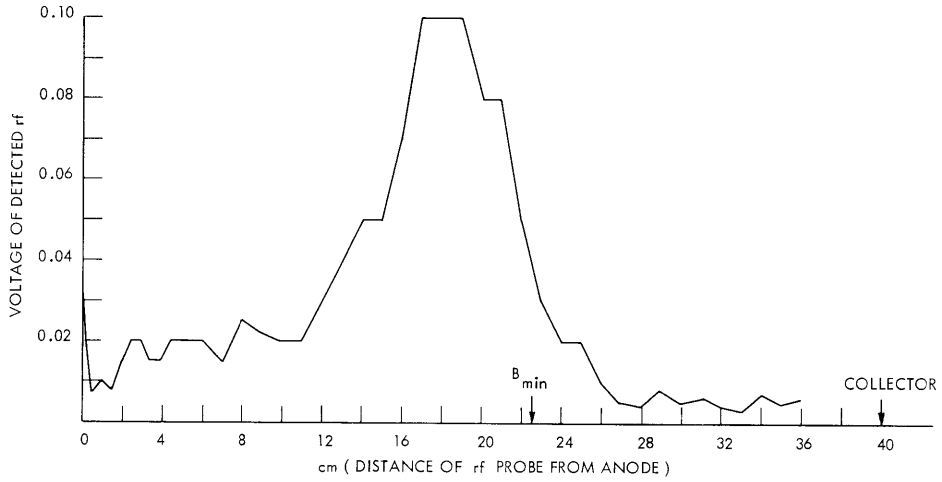

Fig. X-1. Voltage of detected X-band radiofrequency vs distance from the anode. (Magnetic field, 67 gauss; beam voltage, 8 kv; beam current, 0.65 amp; and gas is H_2 at 2 μ Hg pressure.)

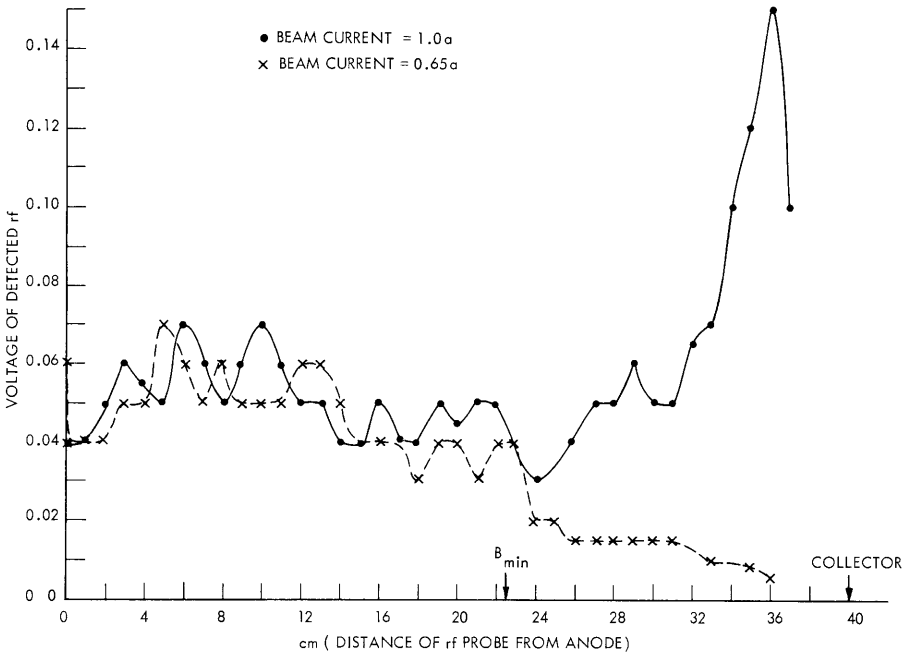

Fig. X-2. Voltage of detected K-band radiofrequency vs distance from the anode. (In both cases: pressure, 5.5 μ Hg; beam voltage, 8 kv; magnetic field for 0.65-amp beam current, 180 gauss; for the 1.0-amp beam current, 100 gauss; gas is hydrogen.)

minimum field of a magnetic mirror. Hence, these data may have significance if the radiation is due primarily to the coupling of longitudinal ($\bar{k}$ and $\bar{E}$ parallel) plasma waves to transverse waves by scattering off density gradients. At higher pressures (5.5 μ Hg), there is no sharp change near B_{min}. The RF may then either rise or fall in intensity as the collector is approached (Fig. X-2).

Sometimes, a very clean oscillation in the envelope of the continuously radiated RF is observed. At 5.5 μ Hg or H_2, 120 gauss, with an 8 kv, 0.65-amp beam, this amplitude modulation is 100 per cent. The envelope appears as 2-μsec pulses with a repetition rate of 1.6×10^5 pps. This is near the expected frequency of the fundamental standing ion acoustic wave. No azimuthal or longitudinal variation has been found in the phase of this modulation.

J. A. Davis

References

1. J. A. Davis, "Beam-Plasma Discharge: System A," Quarterly Progress Report No. 72, Research Laboratory of Electronics, M. I. T., January 15, 1965, pp. 102-104.
2. H. Y. Hsieh, "Experimental Study of Beam-Plasma Discharge," Ph. D. Thesis, Department of Electrical Engineering, M. I. T., September 1964.
3. A. H. Kritz and D. Mintzer, "Propagation of Waves Across a Density Discontinuity," Phys. Rev. 117, 382-386 (1960).

B. BEAM PLASMA DISCHARGE: SYSTEM C

1. HEATING OF IONS

Experiments are under way which have as their objective the heating of ions in a beam-plasma discharge. The proposed scheme involves modulating the beam collector with radiofrequency power of frequency equal to the ion cyclotron frequency in the center of the mirror. Conditions under which ion-cyclotron waves can be launched from the

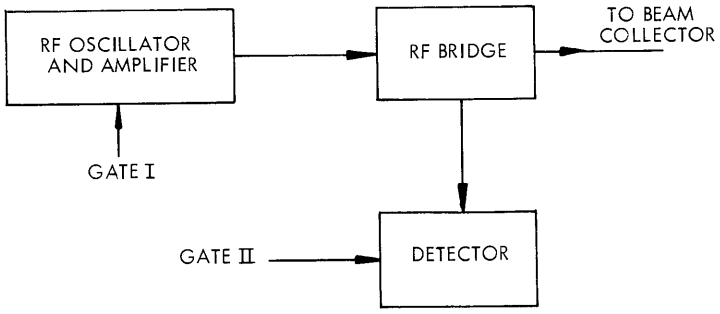

Fig. X-3. Diagram of apparatus.

collector and dissipated in the resonance region will be investigated.

As a first experiment, the impedance of the collector as a function of frequency and discharge parameters is being measured. For this measurement, the apparatus shown in Fig. X-3 has been constructed. In operation, at a variable time during the discharge, gate I produces an adjustable-frequency RF pulse of 200-μsec duration which serves as a source for the RF bridge. Approximately 50 μsec after initiation of the RF, gate II turns on a detector that stays on for 100 μsec. The detector consists of a tunable, R-C active filter that simulates a high-Q (~1000) tank circuit followed by a highpass filter. It represents the optimum linear, time-invariant filter for maximum signal-to-noise ratio.[1]

First results have shown that Re Z, the real-part of the impedance seen at the collector, has the following characteristics: (i) for $\omega < \omega_{ci}$, it is ~1.5 ohms and relatively independent of ω; (ii) in the vicinity of $\omega_{ci}/2$, it exhibits a small resonance; and (iii) it increases rapidly as ω approaches ω_{ci} (ω_{ci} is the ion-cyclotron frequency measured in the center of the mirror). At the present time, more complete data are being taken and work has begun on a theoretical model.

R. R. Parker

References

1. W. B. Davenport, Jr. and W. L. Root, An Introduction to the Theory of Random Signals and Noise (McGraw-Hill Book Company, New York, 1958).

2. ELECTRON TEMPERATURES

Following a method outlined by Griem,[1] the electron temperature in the beam plasma discharge System C was determined from the ratio of line to continuum intensities. This method was selected because it is accurate at the low electron densities encountered (~$10^{13}/cm^3$). Also, the ratio does not depend on the electron density, which may vary spatially and with time. A pure hydrogen plasma was used.

The continuum radiation is caused by Bremsstrahlung and by recombination of free electrons to bound hydrogen states with $n \geq 3$. The line-intensity formula presupposes a Maxwellian distribution. The intensity ratio is given by

$$\frac{I_\ell}{I_c} = \frac{4.2 \times 10^5 f g \exp((E_i - E)/kT)}{\lambda \Delta\lambda (g_{ff}/2)(kT/E_i) + \sum_{n=3}^{\infty} \left(g_{fb}/n^3\right) \exp\left(E_i/n^2 kT\right)},$$

where f is the line oscillator strength, g is the statistical weight of the lower state of the transition, E_i is the ionization energy for a ground-state atom, E is the emitted photon energy, wavelengths are given in angstroms, and g_{ff} and g_{fb} are the free-free

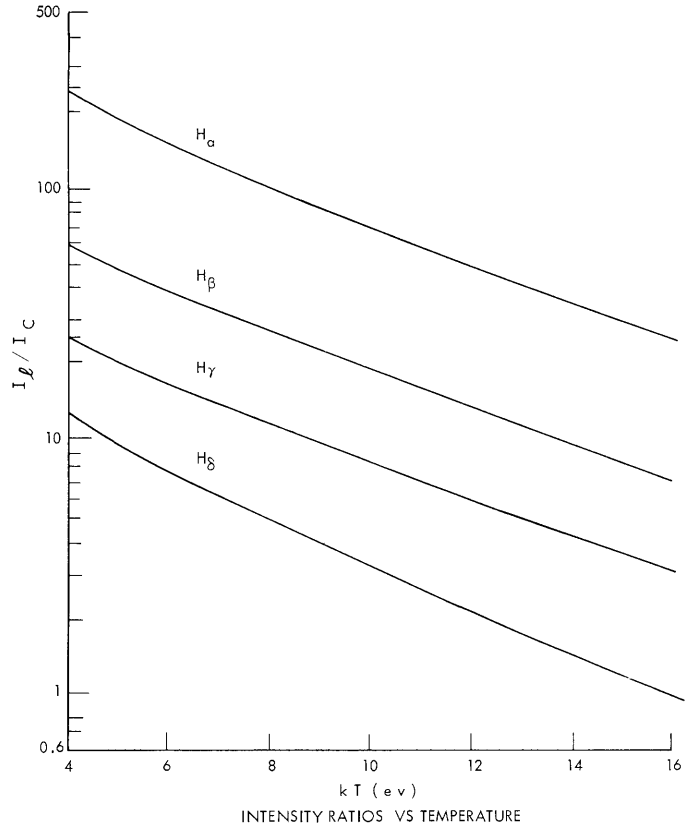

Fig. X-4. Intensity ratios for four Balmer lines.

and free-bound Gaunt factors[2] which are approximately unity. Figure X-4 is a plot of these ratios for four Balmer lines over the relevant temperature range.

A half-meter Ebert plane grating monochromator was used for the intensity measurements. The entrance slits were removed so that a wide band of the continuum could be observed. Continuum intensities were measured at two wavelengths above and below the line and were averaged. Any weak H_2 lines were avoided. The continuum intensities were subtracted from the line intensities when they were comparable in magnitude.

Results indicated an electron temperature of approximately 13 ev. Deviations from the Maxwellian distribution did not affect results greatly because the threshold energies for the relevant reactions were less than the temperature, kT.

Further details of this measurement and other spectroscopic measurements on System C will be found in the author's thesis.[3]

P. E. Newton

References

1. E. R. Griem, Plasma Spectroscopy (McGraw-Hill Book Company, New York, 1964), Chap. 13.

2. W. J. Karzas and R. Latter, Astrophys. J., Suppl. 6, No. 55, p. 167, 1961.

3. P. E. Newton, "Spectroscopic Determination of Electron Temperatures in a Hydrogen Plasma," S. B. Thesis, Department of Physics, M. I. T., May 21, 1965.

C. BEAM PLASMA DISCHARGE: SYSTEM D

From the beginning, System D has been something of a disappointment in that its performance was no better, and probably worse, than its smaller predecessor, System C. The problem centered mainly on what we called "axial breakdown." This manifested itself as a burst of current: electrons to the collector and ions to the cathode. The burst would appear soon after a beam-plasma discharge (BPD) had been established. Since this current burst could reach a peak value five to ten times the normal beam current of 10 amps, this burst would act as a "crowbar" and effectively shut off the discharge. The excess current usually limited the maximum beam power to less than 100 kw, and often to less than 50 kw.

After some time, we realized that the magnitude of the excess current corresponded very closely to the rate of generation of new ion-electron pairs caused by diffusion of neutral gas into the plasma. Thus, it seemed that the excess current could be eliminated if the ambient neutral density surrounding the plasma could be reduced. It appeared that a pulsed gas feed might help. In normal (previous) operation, gas is fed in continuously and pumped out to maintain an H_2 pressure of $\sim 10^{-4}$ torr. The new configuration with the pulsed gas feed is shown in Fig. X-5.

Figure X-6 shows x-ray decay with time. The curve was plotted with a 400-channel

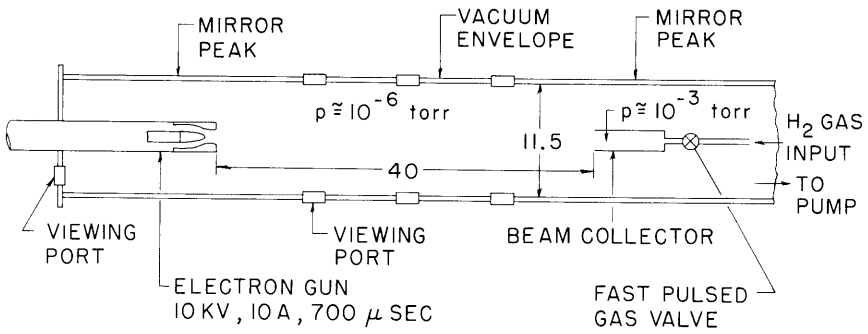

Fig. X-5. System D: showing location of pulsed gas valve.

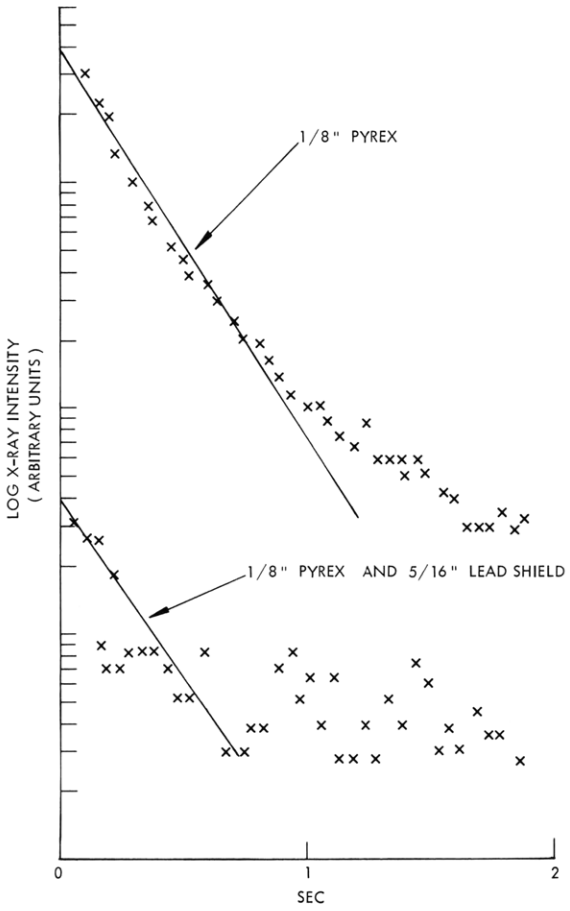

Fig. X-6.

X-ray decay with time. Vertical scale indicates number of counts (logarithmic scale) and horizontal scale, time. Major divisions are 0.2 sec. Upper curve is without lead shield and lower curve is with lead shield.

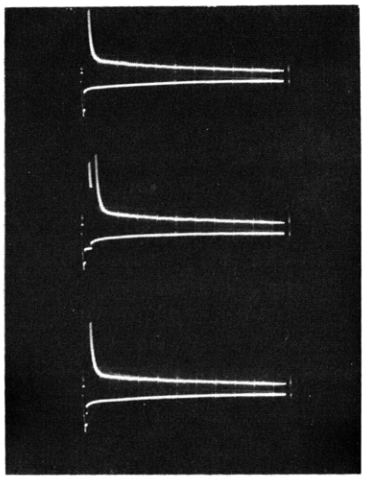

Fig. X-7.

In each frame, upper trace is the output of x-ray scintillator, and lower trace is the diamagnetic signal ΔB. Time scale: 20 msec/cm. Calibration of the ΔB trace is 16 gauss/cm.

analyzer operated in the scaling mode. The upper curve was taken with the scintillator looking at the 1/4 inch thick glass window. The lower curve was taken with a 5/17 inch thick lead shield placed in front of the scintillator. From this attenuation we infer an x-ray energy of 350 kev. The exponential decay time measured from the upper curve is 0.26 second. The neutral pressure at the beginning of the decay was 3×10^{-5} torr. This places the electron temperature at 60-70 kev if we assume that the decay is by electron-neutral scattering.

Figure X-7 shows some traces of integrated diamagnetic probe signals and x-ray scintillator outputs. The diamagnetic signal corresponds to the flux density (ΔB) change within the plasma because of its transverse energy. The uppermost curve shows a smooth decay, whereas the other two show sudden drops resulting from some instability. The maximum expelled flux (ΔB) was ~50 out of 660 gauss (the mirror ratio was 6, and most measurements were made with a center field of 660 gauss).

Figure X-8 shows light decay and diamagnetic signal with "simultaneous" instabilities. It has been observed that a burst of x-rays precedes the sudden drop in light or

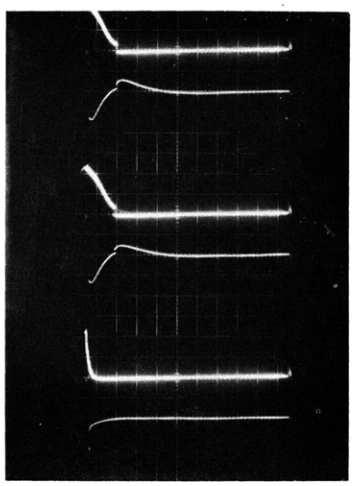

Fig. X-8. In each frame, upper trace is the light and lower trace, the diamagnetic signal, ΔB. Time scale: 50 msec/cm. The ΔB calibration is the same as in Fig. X-7.

ΔB signal by approximately 1 msec. After the burst, the x-ray level returns to its value just before the burst. The ΔB drop lags the light drop by a smaller interval. Both ΔB and light drop nearly to zero. We offer no explanation for the observed timing sequence or the nature of the instability.

The valve (a commercial unit) was placed just back of the collector and was driven

by a small pulser (2D21 thyratron, 10 μf, 150 volts). The beam pulse was delayed approximately 2-5 msec from the valve pulse. This delay was adjusted empirically for "best operation."

The presumed operation is as follows: The optimum delay presumably corresponds to a time when the gas is still concentrated inside the collector. Thus the beam initiates a BPD in the dense gas and the plasma so produced diffuses out ahead of the gas into the magnetic mirror where the electrons are further heated by the beam-plasma interaction. The neutral gas remaining in the collector is now forced to diffuse through the energetic plasma before entering the main drift tube. If the plasma is dense enough, very little neutral gas will escape without being ionized. If this model is correct, one would expect to find a much more energetic plasma, since there would be little cooling by the low-pressure ($<10^{-6}$ torr) ambient gas.

Experiments made within the last three weeks seem to confirm this model. The most dramatic effects are: (i) that the "mirror-shaped" plasma can be seen (visually) to glow during the entire interpulse period of 1-2 sec; (ii) the x-radiation has increased to more than 100 mR/hr; (iii) there is no sign of excess current to the cathode so the gun operates as if it were in vacuum.

Some of the oscillograms describing the behavior of the system are shown below. During operation, the ambient (base) pressure was $0.8\text{-}1.0 \times 10^{-6}$ torr. The valve was adjusted to admit approximately $1\text{-}5 \times 10^{17}$ molecules per pulse (this was estimated from the pump speed – 1000 ℓ/sec and the average H_2 pressure over the interpulse period $\approx 10^{-5}$ torr). We have not yet observed the pressure transient of the gas at the collector, but we suspect that a substantial amounts of gas drools out after the beam pulse ends (700 μsec), and serves no useful purpose.

L. D. Smullin, W. D. Getty, T. Musha, R. R. Bartsch

D. INSTABILITIES OF TRANSVERSE WAVES ALONG THE MAGNETIC FIELD

In a previous report plasma instabilities in transverse waves propagating along the magnetic field were presented.[1] The plasma system was assumed to be composed of stationary neutralizing ions penetrated by electrons with an unperturbed velocity-space distribution of the form

$$f_o(\bar{v}) = \frac{1}{2\pi v_{0\perp}} \delta(\bar{v}_\perp - \bar{v}_{0\perp}) \, \delta(v_{||} - v_{0||}), \tag{1}$$

where $v_\perp$ and $v_{||}$ are the velocities perpendicular to and along the magnetic field. Instabilities of transverse waves were first reported by Weibel.[2] Sudan has considered an anisotropic Maxwellian distribution employing the instability criterion of Penrose.[3-5] He obtained the real wave number range over which the instability exists as a function of $T_\perp/T_{||}$. In the present analysis the criterion of Bers and Briggs is used, in order

to classify the instability as convective or absolute.[6]

This report deals with some of the effects of finite electron temperature on the beam-type distribution of Eq. 1. Under the assumption of an exp j(ωt-kz) dependence of the field quantities, a simultaneous solution of the relativistic Vlasov equation and Maxwell's equations yield the dispersion equation

$$\frac{c^2k^2}{\omega^2} = 1 - \int_{-\infty}^{\infty} dv_{||} \int_0^{\infty} dv_\perp \, 2\pi v_\perp \frac{\omega_p^2}{\omega^2} f_0(v_{||}, v_\perp)$$

$$\times \left[\frac{\omega - kv_{||}}{(\omega - kv_{||} \mp \omega_b)} + \frac{v_\perp^2}{2} \frac{(k^2 - \omega^2/c^2)}{(\omega - kv_{||} \mp \omega_b)^2} \right], \tag{2}$$

where the minus and plus signs are for right- and left-polarized waves, respectively. The distribution function for the electrons is taken to be

$$f_0(\bar{v}) = \frac{1}{\sqrt{2\pi}\, v_{T||} I(v_{0\perp})} \exp\left[- \frac{(v_\perp - v_{0\perp})^2}{2v_{T\perp}^2} \right] \exp\left[- \frac{(v_{||} - v_{0||})^2}{2v_{T||}^2} \right], \tag{3}$$

where $v_{T||} = (T_{||}/m)^{1/2}$ is the thermal velocity along the field, $v_{T\perp} = (T_\perp/m)^{1/2}$ is the thermal velocity across the field,

$$I(v_0) = 2\pi \left\{ v_T^2 \exp\left(- \frac{v_0^2}{2v_T^2} \right) + \sqrt{\frac{\pi}{2}}\, v_{T\perp} v_{0\perp} \left[1 + \operatorname{erf}\left(\frac{v_{0\perp}}{\sqrt{2}\, v_{T\perp}} \right) \right] \right\}$$

is a normalization factor, and

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-u^2} du.$$

From Eq. 2 we then find the dispersion relation:

$$\frac{c^2k^2}{\omega^2} = 1 - \frac{\omega_p^2}{\omega^2} \left(1 - \frac{\omega_b}{\sqrt{2}\, kv_T} H(\xi) \right) + \frac{\omega_p^2}{\omega^2} \frac{\left(k^2 - \frac{\omega^2}{c^2} \right)(1 + \xi H(\xi))}{k^2 v_{T||}^2}$$

$$\times \left\{ \frac{v_{0\perp}^2}{2} + v_{T\perp}^2 + v_{0\perp} v_{T\perp}^3 (2\pi)^{3/2} \left[1 + \operatorname{erf}\left(\frac{v_{0\perp}}{\sqrt{2}\, v_{T\perp}} \right) \right] \right\}, \tag{4}$$

where

$$H(\xi) = \frac{1}{\sqrt{\pi}} \int_{+\infty}^{\infty} \frac{e^{-\xi^2}}{x - \xi} d\xi; \quad \text{Im } \xi < 0$$

$$\xi = \frac{\omega - kv_{0\parallel} - \omega_b}{\sqrt{2}\, kv_{T\parallel}}$$

and $H(\xi)$ is related to the plasma dispersion function, $Z(\xi)$, of Fried and Conte through analytic continuation.[7]

$$H(\xi) = Z(\xi) - j\xi\sqrt{\pi}\, e^{-\xi^2}.$$

Figure X-9 gives the locus of real k in the complex-frequency plane for various

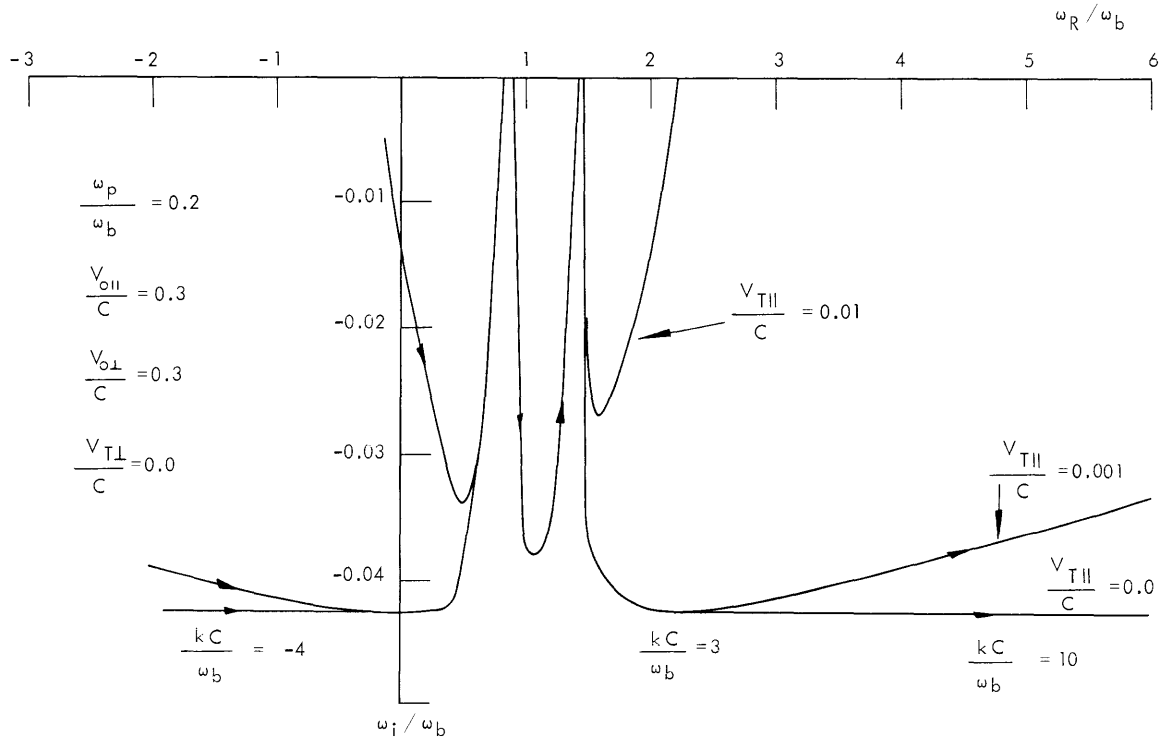

Fig. X-9. Real wave number locus in complex-frequency plane for various thermal velocities.

electron temperatures. The temperature across the field is zero so that the distribution function and dispersion equation reduce to

$$f_0(\bar{v}) = \frac{1}{(2\pi)^{3/2} v_{T\parallel} v_{0\perp}} \delta(v_\perp - v_{0\perp}) \exp\left[-\frac{(v_\parallel - v_{0\parallel})^2}{2v_{T\parallel}^2}\right] \tag{5}$$

$$\frac{c^2k^2}{\omega^2} = 1 - \frac{\omega_p^2}{\omega^2}\left(1 - \frac{\omega_b}{\sqrt{2}\,kv_{T\|}}H(\xi)\right) + \frac{\omega_p^2 v_{0\perp}^2(k^2-\omega^2/c^2)}{\omega^2 2k^2 v_{T\|}^2}(1+\xi H(\xi)). \quad (6)$$

The transcendental equation (5) was analyzed with the aid of the Project MAC time-sharing system. The method employed is similar to that reported by Lieberman.[8]

Figure X-9 demonstrates the damping of the short wavelength instability by longitudinal temperature. The long wavelength instabilities are undamped for the temperature ranges considered. Note that temperature across the field will not contribute to the damping. The location of the absolute instabilities was reported in the previous report and their temperature dependence is now under study.

Some analysis has been done on determining the physical mechanism that gives rise to the instabilities. It has been determined that there are two distinct effects. The instability near zero wave number is due to a relativistic change in the cyclotron frequency of the electron as it interacts with the electromagnetic wave. For zero wave number, the magnetic field is unimportant and the electric field is space-independent and (right-hand) circularly polarized at a frequency ω. Figure X-10a shows an electron whose velocity vector is in the direction of the electric field at t = 0. This electron will give

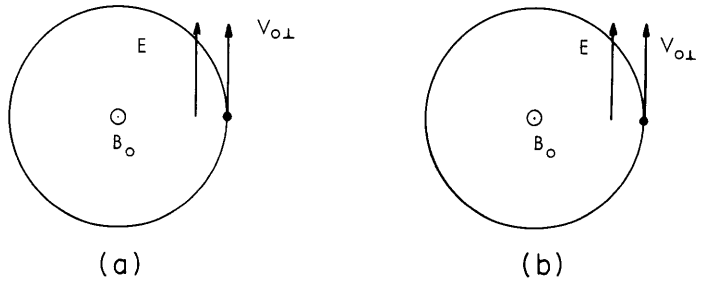

Fig. X-10. Relation between in-phase electron and electric field for $\omega > \omega_{bo}$: (a) t = 0; (b) t = $2\pi/\omega$.

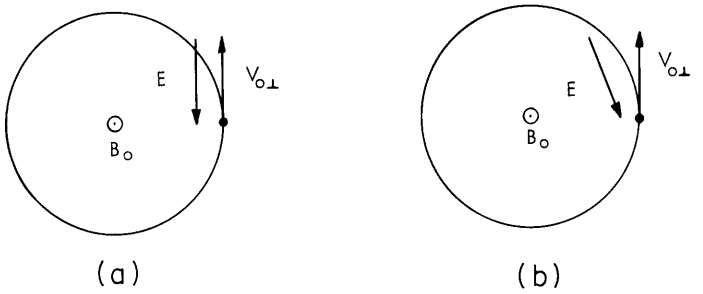

Fig. X-11. Relation between antiphase electron and electric field for $\omega > \omega_{bo}$: (a) t = 0; (b) t = $2\pi/\omega$.

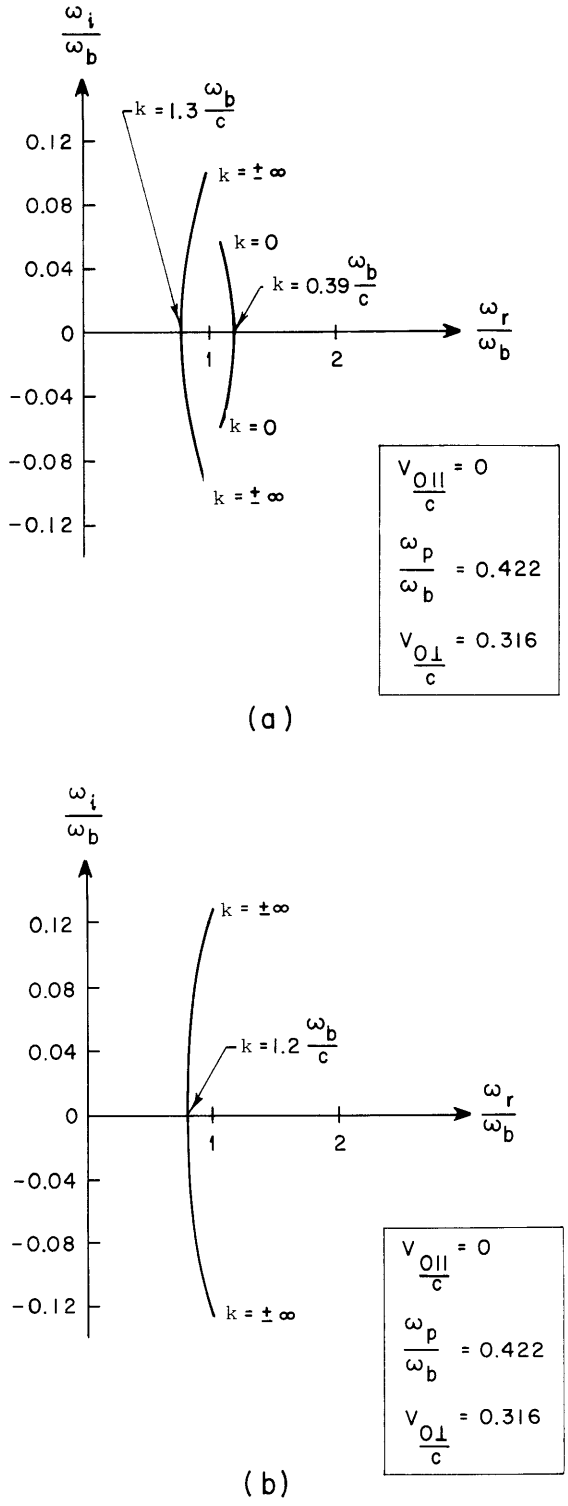

Fig. X-12. (a) Relativistic dispersion showing instability at k = 0.
(b) Nonrelativistic dispersion showing absence of zero k instability.

up energy to the field and relativistically its mass will decrease. Thus the cyclotron frequency of the electron is increased. If the frequency, ω, of the electric field is slightly greater than the unperturbed cyclotron frequency of the electron, ω_{bo}, then the electron will remain in phase with the field durings its orbit, as shown in Fig. X-10b. The electron gives up energy to the wave throughout its orbit. Figure X-11a shows an electron whose velocity vector is 180° out of phase with the electric field at t = 0. This electron will take energy from the field and its mass will increase. Thus its cyclotron frequency is decreased. If the electric field frequency is greater than the unperturbed cyclotron frequency, this electron will fall out of antiphase. Thus the antiphase electron gains less energy from the wave than the in-phase electron gives up, and the wave will grow in time. The computer solutions of the dispersion equation with the distribution function of Eq. 1 used indicate that the zero wave number instability is at $\omega > \omega_{bo}$. A nonrelativistic analysis fails to yield these small wave number instabilities as demonstrated by Fig. X-12.

Work is, at present, being done to establish the physical mechanism for the instabilities at large wave number. In this region the first-order magnetic field plays an important role.

E. A. Robertson, A. Bers

References

1. A. Bers, J. K. Hoag, and E. A. Robertson, "Instabilities in Transverse Waves along B_o for Beam-Type Distributions," Quarterly Progress Report No. 77, Research Laboratory of Electronics, M. I. T., April 15, 1965, pp. 149-152.
2. E. S. Weibel, Phys. Rev. Letters 2, 83 (1959).
3. R. N. Sudan, Phys. Fluids 6, 57 (1963).
4. R. N. Sudan, Phys. Fluids 8, 153 (1965).
5. O. Penrose, Phys. Fluids 3, 258 (1960).
6. A. Bers and R. J. Briggs, "Criteria for Determining Absolute Instabilities and Distinguishing between Amplifying and Evanescent Waves," Quarterly Progress Report No. 71, Research Laboratory of Electronics, M. I. T., October 15, 1963, pp. 122-131.
7. B. Fried and S. Conte, The Plasma Dispersion Function (Academic Press, New York, 1961).
8. M. A. Lieberman, "Dispersion Diagrams for Hot-Electron Plasmas," Quarterly Progress Report No. 77, Research Laboratory of Electronics, M. I. T., April 15, 1965, pp. 141-144.

E. INSTABILITIES OF LONGITUDINAL WAVES ACROSS THE MAGNETIC FIELD

We have been studying longitudinal waves in a uniform plasma. By solving the linearized, relativistic Vlasov equation self-consistently with Maxwell's equations, it has

been shown[1] that longitudinal waves with dependence exp $j(\omega t - \bar{k}\cdot\bar{r})$ satisfy the dispersion relation, $\bar{k}\cdot\bar{\bar{K}}\cdot\bar{k} = 0$:

$$1 - \sum_{\text{species}} \left\{ \int_0^\infty dv_\perp\, v_\perp \int_{-\infty}^{\infty} dv_{||}\, \frac{\omega_p^2}{k^2}\, 2\pi f_0 \times \right.$$

$$\left. \sum_{n=-\infty}^{\infty} \left[\frac{n\left(\frac{k_\perp}{v_\perp}\right)\frac{d}{dp}\left[J_n^2(p)\right]}{(\omega - k_{||}v_{||} - n\omega_b)} + \frac{\left[k^2 - \frac{(k_{||}v_{||} - n\omega_b)^2}{c^2}\right] J_n^2(p)}{(\omega - k_{||}v_{||} - n\omega_b)^2} \right] \right\} = 0, \tag{1}$$

where $f_0 = f_0(v_\perp, v_{||})$ is the unperturbed velocity distribution function for each species, $J_n(p)$ is the n^{th}-order Bessel function of the first kind and argument $p = (k_\perp v_\perp)/\omega_b$, $v_{||}$ is the component of the velocity parallel to the magnetic field, $v_\perp$ is the magnitude of the velocity transverse to the magnetic field, and $k_{||}$ and $k_\perp$ are the components of the propagation vector, parallel and perpendicular to the magnetic field. Note that ω_b and ω_p are the relativistic cyclotron and plasma frequencies, and thus are functions of velocity.

We shall consider a background of infinitely massive ions, free electrons that have an unperturbed velocity distribution of the form

$$f_0(\bar{v}) = \frac{1}{2\pi v_\perp}\, \delta(v_\perp - v_{0\perp})\, \delta(v_{||}), \tag{2}$$

and shall limit this study to waves propagating perpendicular to the magnetic field, $\bar{k} = \bar{i}_x k_\perp$. The dispersion relation then becomes

$$1 - \frac{\omega_p^2}{2\omega_b^2} \left\{ \sum_{n=-\infty}^{\infty} \left[\frac{J_{n-1}^2(p) - J_{n+1}^2(p)}{\left(\frac{\omega}{\omega_b} - n\right)} \right] - \frac{2\beta_\perp^2}{p^2} \sum_{n=-\infty}^{\infty} \left[\frac{n^2 J_n^2(p)}{\left(\frac{\omega}{\omega_b} - n\right)^2} \right] \right\} = 0, \tag{3}$$

where $\beta_\perp = v_0/\omega_b$. This dispersion relation is seen to differ from that reported by Harris, Dory, and Guest[2] by the addition of the second sum which is the relativistic correction.

In this report we shall discuss the interaction near $\omega = 0$ which Harris and his co-workers termed the "zero-frequency mode." Arguing from the nonrelativistic dispersion relation, they predicted that for $\omega_p > 4.13\,\omega_c$ instabilities should exist in which the real part of the frequency was identically zero. These were reported to appear approximately for bands of real values of p: $2.4 < p < 3.83$, $5.52 < p < 7.02$, $8.65 < p < 10.17$, etc. In this report we enlarge upon their findings by actually calculating the

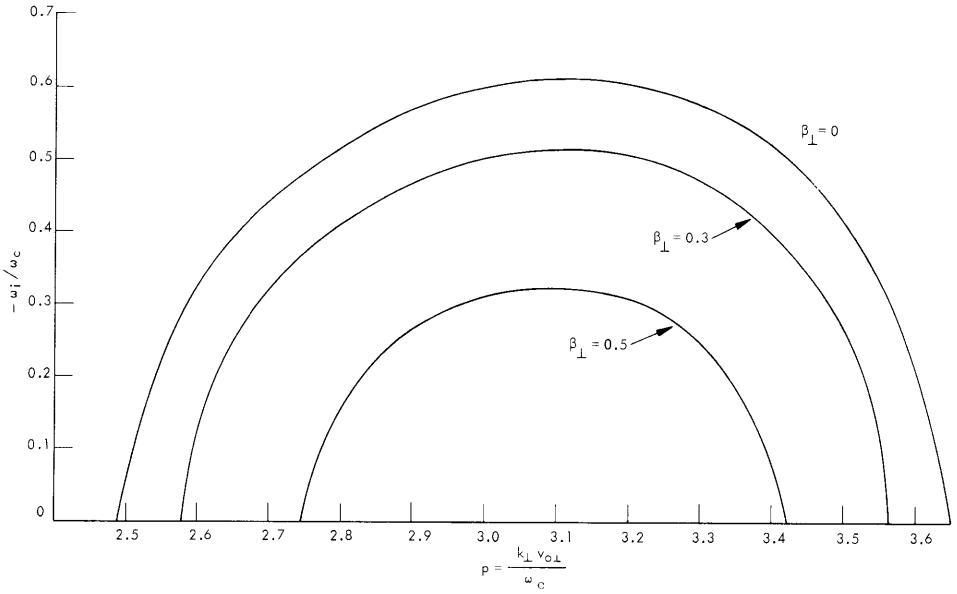

Fig. X-13. Growth rates for the first band of the "zero-frequency" instability $\omega_p/\omega_b = 7.3$.

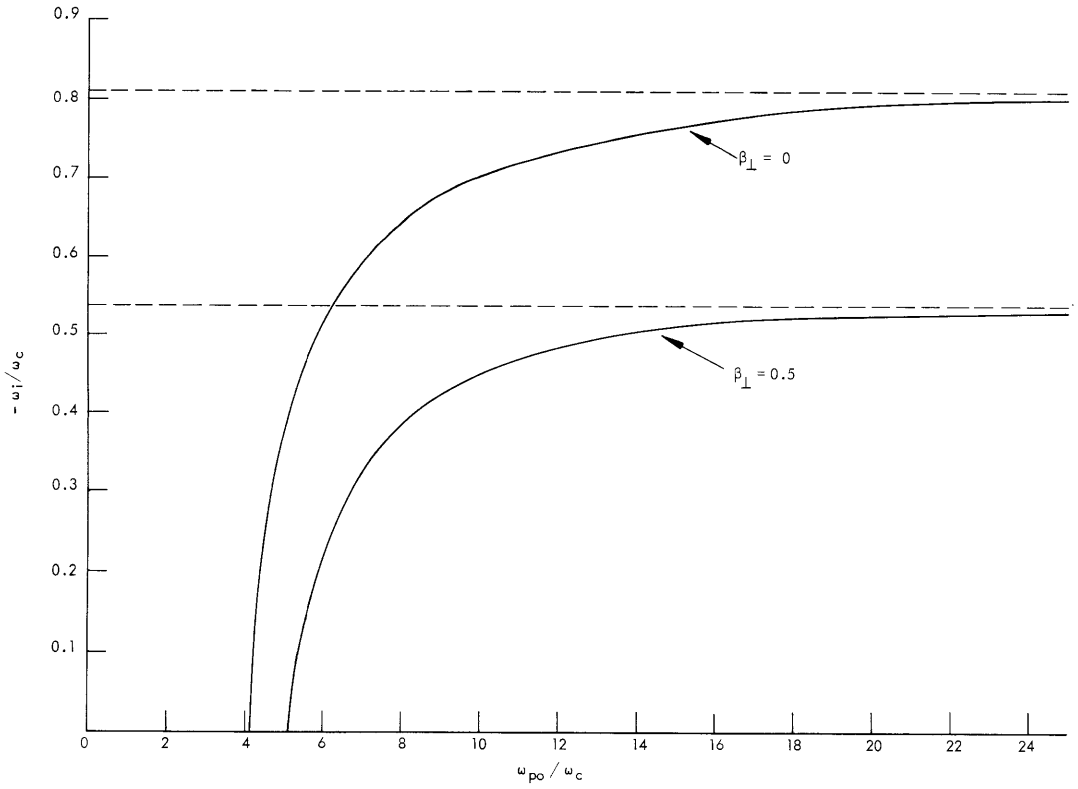

Fig. X-14. Maximum growth rate of the first band as a function of ω_{p0}/ω_c.

growth rates (Im ω) for the first band, and show how they appear in the relativistic formulation. Figure X-13 illustrates the results of a numerical investigation of the first band, under the assumption of a density much greater than the predicted threshold value. Note that the growth rate within the band is normalized to the nonrelativistic cyclotron frequency ω_c, and ω_b is related to ω_c by the expression

$$\omega_b = \omega_c\left(1-\beta_\perp^2\right)^{1/2}. \tag{4}$$

The region of instability for $\beta_\perp = 0$ lies entirely within the approximate band predicted by Harris, Dory, and Guest. The other two curves of Fig. X-13 show the effect of the relativistic correction. The magnitude of the growth rate decreases with increasing $\beta_\perp$, as does the width of the band of real p values. The value of p for maximum growth rate remains constant, however, at a value of $p \cong 3.1$. Figure X-14 shows how this maximum growth rate depends upon the ratio ω_{p0}/ω_c. Here ω_{p0} is the nonrelativistic plasma frequency related to ω_p by the expression $\omega_p = \omega_{p0}\left(1-\beta_\perp^2\right)^{1/4}$. Note that the $\beta_\perp = 0$ branch does indeed have a threshold at $(\omega_{p0}/\omega_c) = 4.13$. Increasing ω_{p0}/ω_c increases the growth rate until it saturates at a value given approximately as $\omega_i \cong 0.81\omega_c$. The effect of the relativistic correction is to increase the threshold of ω_{p0}/ω_c and decrease

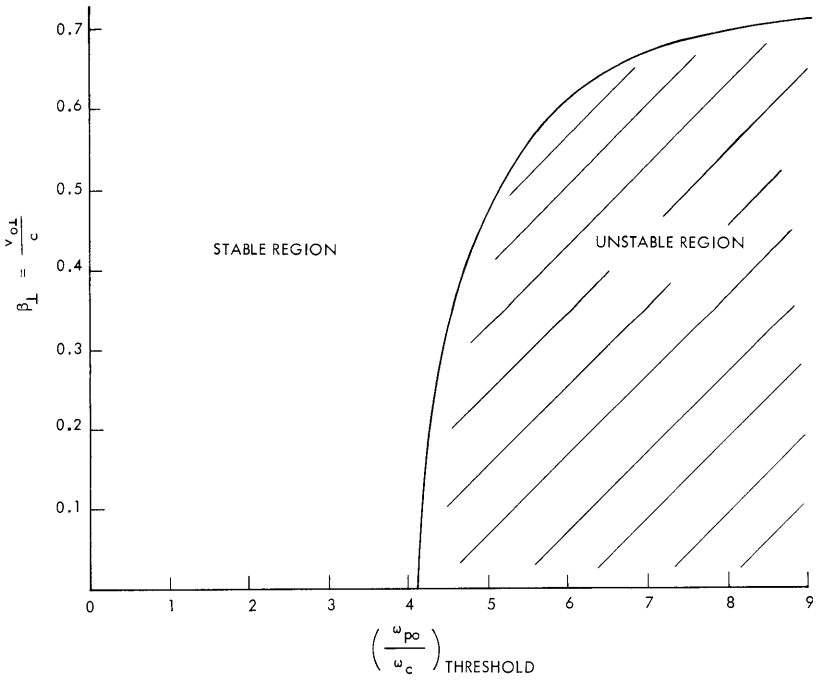

Fig. X-15. Relativistic dependence of "zero-frequency" density threshold $p = \frac{k_\perp v_{0\perp}}{\omega_b} = 3.1$.

the saturation value of the growth rate. Figure X-15 is a plot of $\beta_\perp$ against the threshold value of ω_{p0}/ω_c. This value of ω_{p0}/ω_c will be termed the density threshold. This curve separates the first band of the zero-frequency mode into a stable region and an unstable one. Note that the larger the value of $\beta_\perp$, the greater the density must be in order for the instability to occur.

The study of the dispersion relation (3) continues. This more complete study is not limited to the zero-frequency mode but also considers the interactions at the cyclotron harmonics.

A. Bers, C. E. Speck

References

1. A. Bers, Internal Memorandum, Research Laboratory of Electronics, M. I. T., August 1964 (unpublished).
2. E. G. Harris, R. A. Dory, and G. E. Guest, Phys. Rev. Letters 5, 131 (1965).

F. DYNAMICS OF THE PLASMA BOUNDARY

The motion of electrons near a boundary between a neutralized slab of plasma and free space is being investigated. This study has been undertaken to provide a more complete understanding of wave propagation in finite plasma systems.[1,2]

The geometry of a neutral plasma slab in its equilibrium state is shown in Fig. X-16. It consists of a region, which has thickness d in the x direction and is infinite in the y and z directions, over which electrons and ions are uniformly distributed with a number density n_o (particles per unit volume). The electrons have charge -e, and the ions +e, so that the system is neutral. If the electron cloud is uniformly displace a distance δ (in the x direction) from the ion cloud, an x-directed electric field will be present. The spatial variation of this field is shown in Fig. X-17. It can be seen that if $\delta \ll d$, most of the electrons experience the same electric field, $E_x = \frac{en_o}{\epsilon_o}\delta$. Under

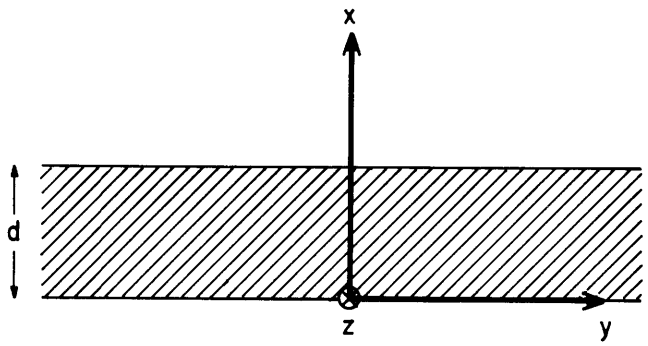

Fig. X-16. Neutral slab in equilibrium.

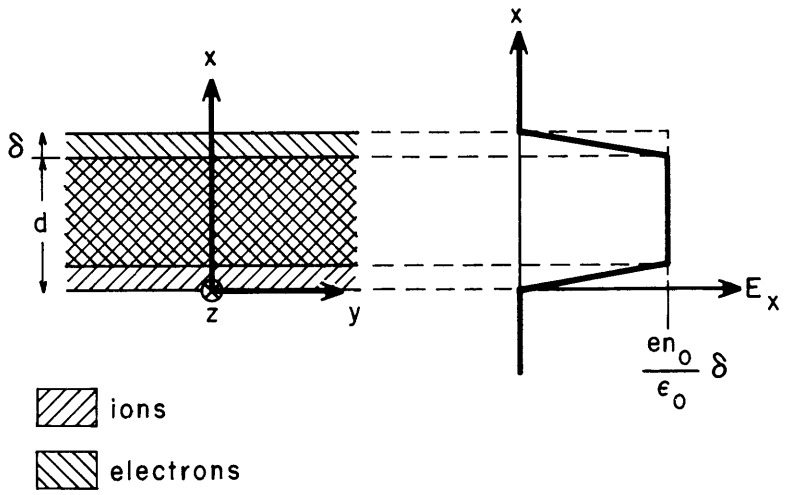

Fig. X-17. Uniform displacement of the electron cloud and generated electric field.

the assumptions that the ion cloud is stationary and the electron cloud moves as a rigid slab, the equation of motion for the electron slab ($\delta \ll d$) is

$$m \frac{d^2\delta}{dt^2} = - \frac{e^2 n_o}{\epsilon_o} \delta, \tag{1}$$

where δ is the displacement of the electron cloud from its equilibrium state, and m is the mass of an electron. According to Eq. 1, the slab oscillates at frequency $\omega = \sqrt{\frac{e^2 n_o}{\epsilon_o m}}$ which is the plasma frequency, $\omega = \omega_p$. It is clear from Fig. X-17, however, that the electrons near the top of the electron cloud (the part that has been displaced outside the ion background) do not experience the same restoring force as those that are inside. Since the electrons are not "tied" to each other as a rigid slab, but are separate particles, they will not move synchronously.

In order to follow the individual particles, a discrete model of the plasma slab has been studied. The continuous charge density of ions and electrons that was uniformly distributed over the thickness d is replaced by N ion sheets and N electron sheets parallel to the y - z plane and spaced uniformly in the x direction so that their separation is (d/N-1) in equilibrium. (In equilibrium, the ion and electron sheets are placed on top of each other pairwise so that the electric field is zero.) Each sheet contains S_o particles per unit area.

Perturbations from equilibrium may be studied by displacing the N electron sheets from the ion sheet background. The ion sheets are assumed to be stationary, while the electron sheets are allowed to move individually in the x direction. The electron sheets can pass freely through the stationary ion sheets, as well as through each other.

The equation of motion for the k^{th} electron sheet is

$$m \frac{d^2 x_k}{dt^2} = -eE_k, \tag{2}$$

where x_k is the position of the k^{th} sheet measured from x = 0, and E_k is the average of the x-directed electric field on either side of the k^{th} electron sheet. Figure X-18 shows a system of 5 electron and ion sheets subject to an initial perturbation of 2 electron

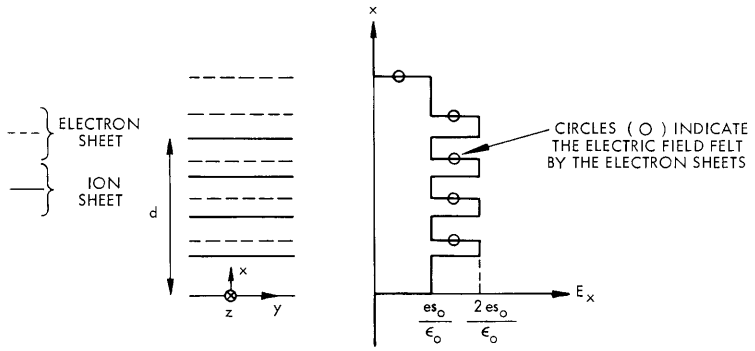

Fig. X-18. Sheet model of the plasma slab.

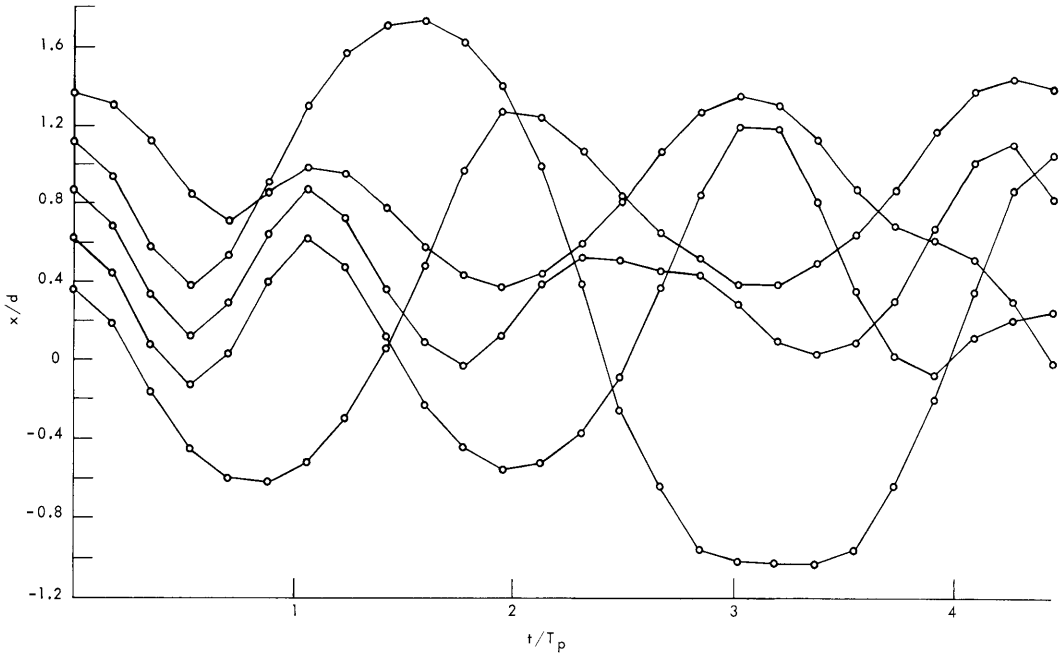

Fig. X-19. Slab dynamics.

$$T_p = \frac{2\pi}{\omega_p} \qquad \omega_p = \sqrt{\frac{5S_o}{\epsilon_o md}}$$

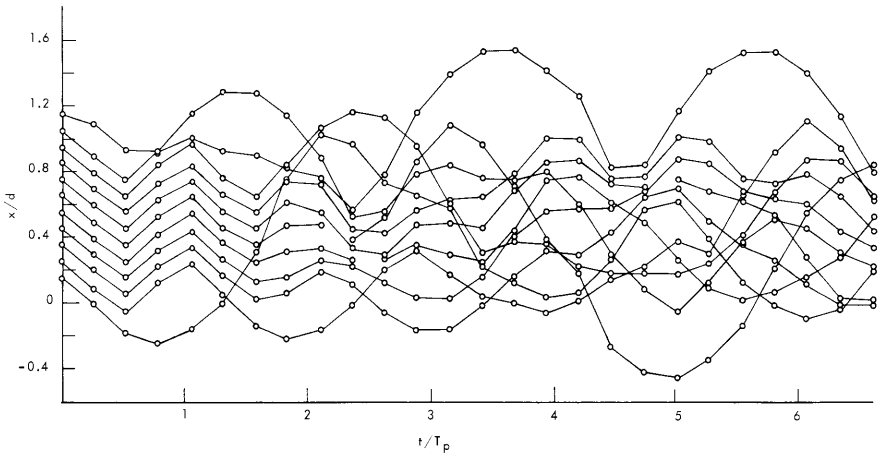

Fig. X-20. Slab dynamics.

$$T_p = \frac{2\pi}{\omega_p} \qquad \omega_p = \sqrt{\frac{11S_o}{\epsilon_o md}}$$

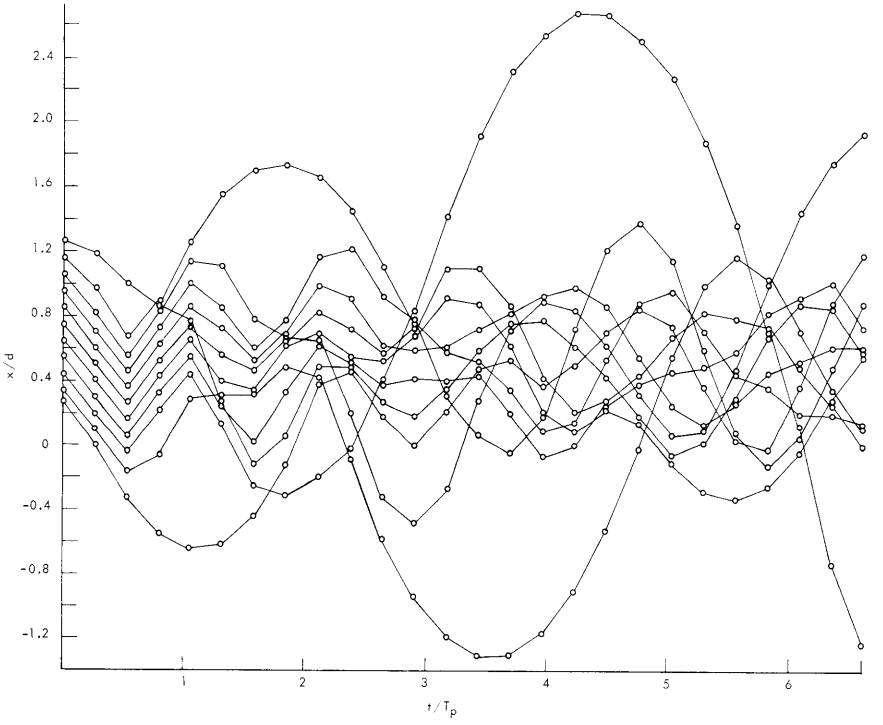

Fig. X-21. Slab dynamics.

$$T_p = \frac{2\pi}{\omega_p} \qquad \omega_p = \sqrt{\frac{11S_o}{\epsilon_o md}}$$

sheets displaced outside the ion sheet background. The electric field distribution in the presence of this perturbation is also shown in Fig. X-18. Note that the outermost electron sheet experiences a different electric field from the other 4 electron sheets. It can be seen that in this planar geometry, the electric field E_k acting on a particular sheet does not change until an electron sheet passes through an ion sheet or another electron sheet. At all times between crossings E_k is a constant, so that each sheet experiences a constant acceleration. The motion of each sheet is found by solving Eq. 2, the electric field being recalculated each time a crossing occurs.

Figure X-19 shows the motion of each sheet as a function of time for a slab modeled by 5 electron and ion sheets which has the initial perturbation shown in Fig. X-18. The central group of electron sheets oscillates initially at a frequency $\omega = \omega_p$, where

$$\omega_p = \sqrt{\frac{e^2 N S_o}{\epsilon_o m d}}.$$

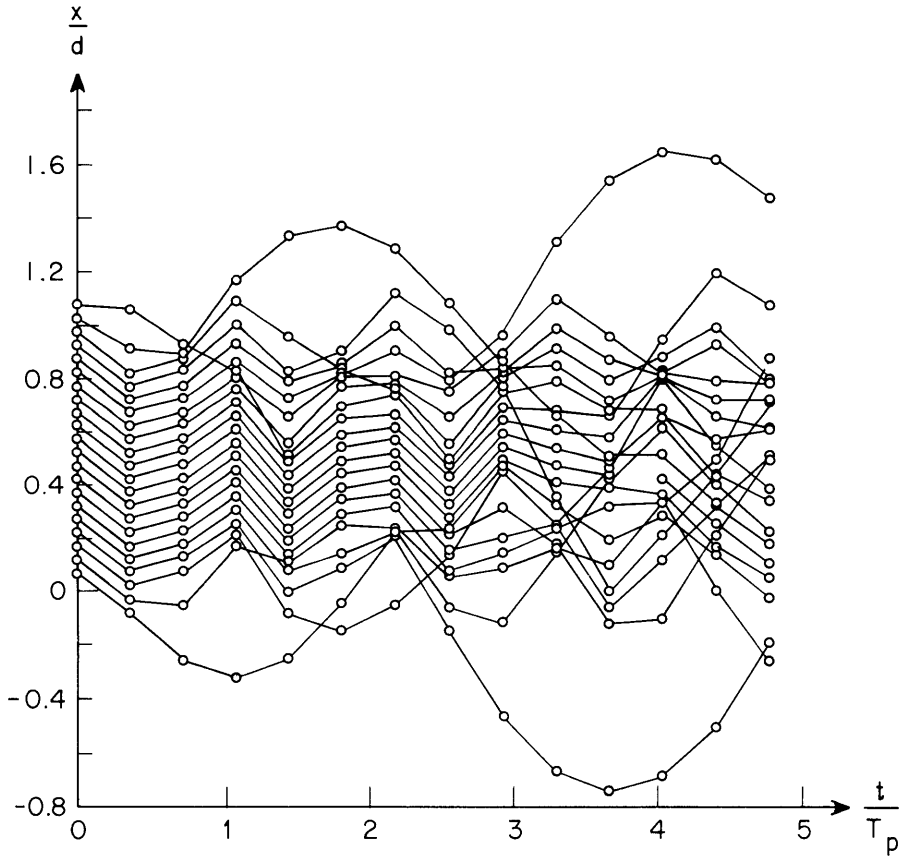

Fig. X-22. Slab dynamics.

$$T_p = \frac{2\pi}{\omega_p} \qquad \omega_p = \sqrt{\frac{21 S_o}{\epsilon_o m d}}$$

Overtaking occurs between the two top sheets after a time which is approximately $t = 2\pi/\omega_p$. Eventually, overtaking takes place between other sheets and the system appears to go toward a chaotic state.

Figures X-20 and X-21 show the dynamics of a slab modeled by 11 sheets. In Fig. X-20, the initial perturbation is taken with two electron sheets outside the ion background, while in Fig. X-21, three sheets are initially displaced. The amplitude of the oscillations is larger than in Fig. X-20 but again the core of the electron sheets begins to oscillate at ω_p. Then overtaking sets in and the slab becomes scrambled.

Finally, in Fig. X-22, the motion of 21 sheets in a slab is shown.

From comparison of Figs. X-19 through X-22, it appears that the coherent oscillation of the central core of electrons at frequency ω_p persists for a longer time as the number of sheets in the model is increased.

This sheet model of the plasma slab will be applied to a slab with nonuniform density. The work will be extended to include the ion-sheet motion, as well as a thermal spread of velocities. The effect of an applied AC electric field will also be investigated.

H. M. Schneider, A. Bers

References

1. W. P. Allis, S. J. Buchsbaum, and A. Bers, Waves in Anisotropic Plasmas (The M. I. T. Press, Cambridge, Mass., 1963).

2. D. L. Bobroff, H. A. Haus, and J. W. Klüver, "On the Small Signal Power Theorem of Electron Beams," J. Appl. Phys. 33, 2932 (1962).

G. ELECTRON CYCLOTRON RESONANCE DISCHARGE*

In the last quarter,[1] we reported a measurement of the plasma density in the interpulse period by a new microwave perturbation technique. In previous reports[2,3] measurements of long-time decay and the plasma electron "temperature" were reported. In this report, we give a physical explanation of the observed phenomena.

The dependence of the plasma density with time as shown in the last report[4] can be very well explained by the following ideas. The ionization time of hot electrons at 4.3×10^{-4} torr is approximately 2.5 μsec, according to Barnett, Ray, and Thompson.[5] Thus in the microwave-pulse time (1 μsec) the plasma electron density changes only slightly, and the principal effect produced by the pulse is to heat electrons already present to large energies. These hot electrons produce new electrons by ionizing the background gas in the interpulse period, thereby making up particle losses and providing a new group of cold electrons to be heated by the next microwave pulse. The decays of plasma measured at long times[2] are evidence of the escape of the hot electrons from the magnetic-mirror field.

Let us assume that after a pulse, all of the electrons are hot and that these electrons decay exponentially with a characteristic loss frequency a_h. We shall also assume that the hot electrons produce cold electrons through ionizing collisions with the background gas. The cold electrons do not produce further ionizations and are lost from the system at a frequency a_c. Writing a conservation equation for the cold electrons n_c gives

$$\frac{dn_c}{dt} = \nu_i n_o e^{-a_h t} - n_c a_c, \tag{1}$$

where ν_i is the ionization frequency of the hot electrons against the background gas. The solution of this equation is

$$n_c = \frac{\nu_i n_o}{a_c - a_h}\left(e^{-a_h t} - e^{-a_c t}\right). \tag{2}$$

The time dependence of the plasma electron density as given by (2) agrees well with that found experimentally.[4] Thus the experimental results can be used to find values for the parameters used in Eqs. 1 and 2. The initial slope of (2) is

$$\left.\frac{dn_c}{dt}\right|_{t=0} = \nu_i n_o. \tag{3}$$

After using the previous data[4] to find n_o and $\left.\frac{dn_c}{dt}\right|_{t=0}$, we find a value of the ionization

*This work was supported principally by the United States Atomic Energy Commission under Contract AT(30-1)-3221.

frequency ν_i which agrees within 20 per cent of that given by Barnett and his co-workers.[5]

The value of loss frequency of the cold electrons can be obtained from the point where the rate of change of plasma density with time is zero. This gives a characteristic loss time ($1/\alpha_c$) of approximately 30 μsec which may represent the time in which the average ion is lost from the system.

A full report on this experiment is available in the form of an Sc. D. thesis presented by T. J. Fessenden to the Department of Electrical Engineering, M. I. T., June 1, 1965.

L. D. Smullin, T. J. Fessenden

References

1. T. J. Fessenden, Quarterly Progress Report No. 77, Research Laboratory of Electronics, M. I. T., April 15, 1965, pp. 159-161.
2. T. J. Fessenden, Quarterly Progress Report No. 76, Research Laboratory of Electronics, M. I. T., January 15, 1965, pp. 125-128.
3. T. J. Fessenden, Quarterly Progress Report No. 75, Research Laboratory of Electronics, M. I. T., October 15, 1964, pp. 66-68.
4. See T. J. Fessenden, Quarterly Progress Report No. 77, op. cit., Fig. XI-18, p. 160.
5. C. F. Barnett, J. A. Ray, J. C. Thompson, "Atomic and Molecular Collision Cross Sections of Interest in Controlled Thermonuclear Research," ORNL-3113, August 1964, p. 137.

H. COMPUTER STUDY OF BEAM-PLASMA INTERACTIONS*

Theoretical investigations are being carried out with the intention of gaining a better understanding of the interactions between beam and plasma waves. With the aid of Project MAC and the Computation Center, M. I. T., dispersion equation solutions can be obtained which would be difficult, and in many cases impossible, to achieve by using a coupling-of-modes approach.[1] The following report describes some of the initial results of these beam plasma-computer studies.

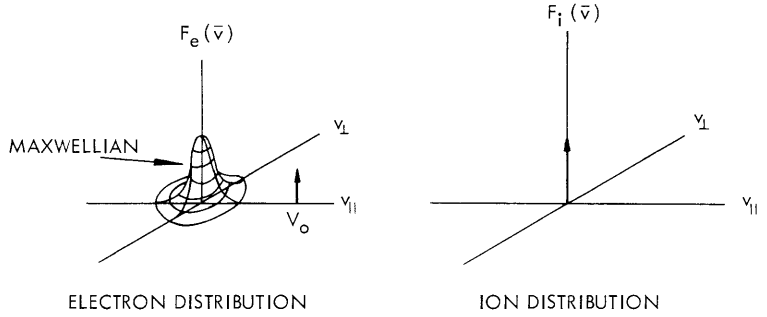

Fig. X-23. Distribution functions.

Under the assumption of an infinite plasma composed of warm electrons and cold ions (see Fig. X-23) and a z-directed magnetic field, quasistatics gives the following dispersion relation:

$$\Delta(\omega,k) = p^2 + k^2 - \omega_{pi}^2 \left\{ \frac{k^2}{\omega^2} + \frac{p^2}{\omega^2 - \omega_{ci}^2} \right\} - \omega_{pb}^2 \left\{ \frac{k^2}{(\omega - kV_o)^2} + \frac{p^2}{(\omega - kV_o)^2 - \omega_{cb}^2} \right\}$$

$$+ \omega_{pe}^2 \left\{ \frac{2\omega}{ka^3} Z\left(\frac{\omega}{ka}\right) + \frac{2}{a^2} + \frac{1}{2} \frac{p^2}{ka\omega_{ce}} \left[Z\left(\frac{\omega - \omega_{ce}}{ka}\right) - Z\left(\frac{\omega + \omega_{ce}}{ka}\right) \right] \right\}, \qquad (1)$$

where

p = component of propagation vector perpendicular to magnetic field

k = component of propagation vector parallel to magnetic field

ω_{pi} = ion plasma frequency

ω_{ci} = ion cyclotron frequency

---

* This work was supported principally by the United States Atomic Energy Commission under Contract AT(30-1)-3221.

ω_{pb} = beam plasma frequency

ω_{cb} = beam cyclotron frequency

V_0 = beam velocity

$a = \sqrt{2} \times$ thermal velocity

ω_{pe} = electron plasma frequency

ω_{ce} = electron cyclotron frequency

$v_{\parallel}$ = velocity parallel to magnetic field

$v_{\perp}$ = velocity perpendicular to magnetic field

and

$$Z(u) = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} \frac{e^{-x^2}}{x - u}\, dx.$$

In obtaining Eq. 1, it was assumed $\frac{pv_{\perp}}{\omega_b} \ll 1$, and consequently no cyclotron harmonics appear.

For $u \gg 1$,

$$Z(u) \rightarrow -\frac{1}{u}\left[1 + \frac{1}{2u^2} + \frac{3}{4u^4} \cdots\right].$$

Therefore, if it is assumed that $\frac{\omega \pm \omega_{ce}}{ka}$ and $\frac{\omega}{ka}$ are large compared with one, and only enough terms of the series above are used to give the lowest order temperature variation, the dispersion relation becomes

$$\Delta(\omega,k) = p^2 + k^2 - \omega_{pi}^2\left[\frac{k^2}{\omega^2} + \frac{p^2}{\omega^2 - \omega_{ci}^2}\right] - \omega_{pb}^2 \frac{k^2}{(\omega - kV_o)^2} + \frac{p^2}{(\omega - kV_0)^2 - \omega_{cb}^2}$$

$$- \omega_{pe}\left\{\frac{k^2}{\omega^2} + \frac{p^2}{\omega^2 - \omega_{ce}^2} + \frac{3}{2}\frac{k^4 a^2}{\omega^4} + \frac{1}{2}\frac{p^2 k^2 a^2}{\left(\omega^2 - \omega_{ce}^2\right)^3}\left(3\omega^2 + \omega_{ce}^2\right)\right\}.$$

This dispersion relation was programmed for the computer and complex roots of k were obtained for real ω. The results for two sets of plasma parameters are shown in Figs. X-24 and X-25. The dotted lines are beam asymptotic lines, and the solid lines represent solutions of real k for real ω. The axes are not drawn to scale but in such a way that the nature of the solutions can be easily seen.

For Fig. X-24, the parameters are as follows:

$V_o = 1 \times 10^8$ meters/sec

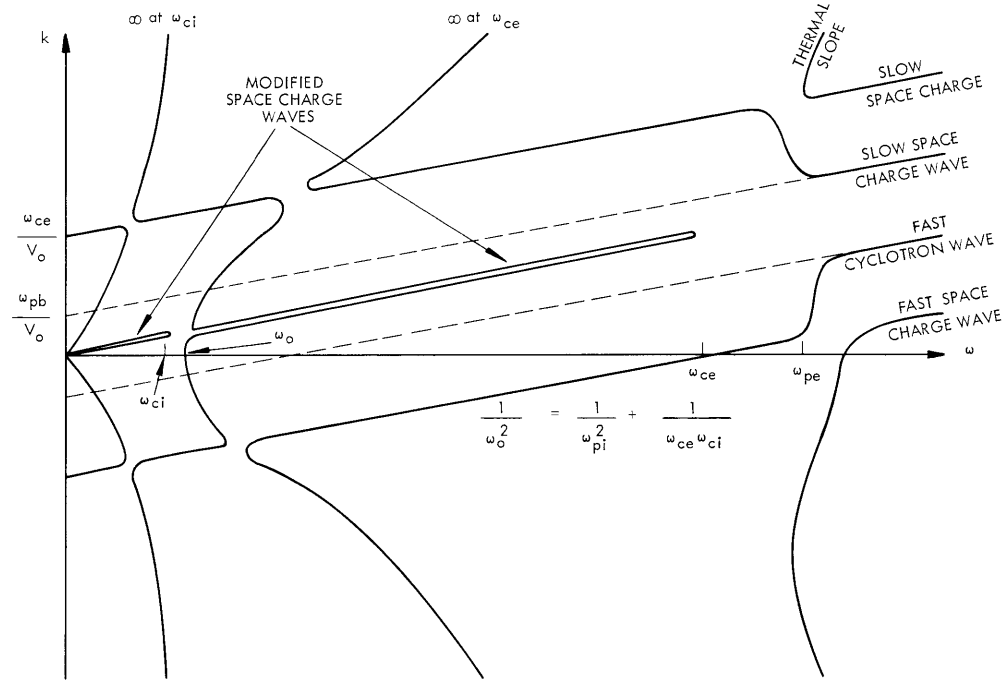

Fig. X-24. Weak-beam solution – $\omega_{pb} < \omega_{ce}$.

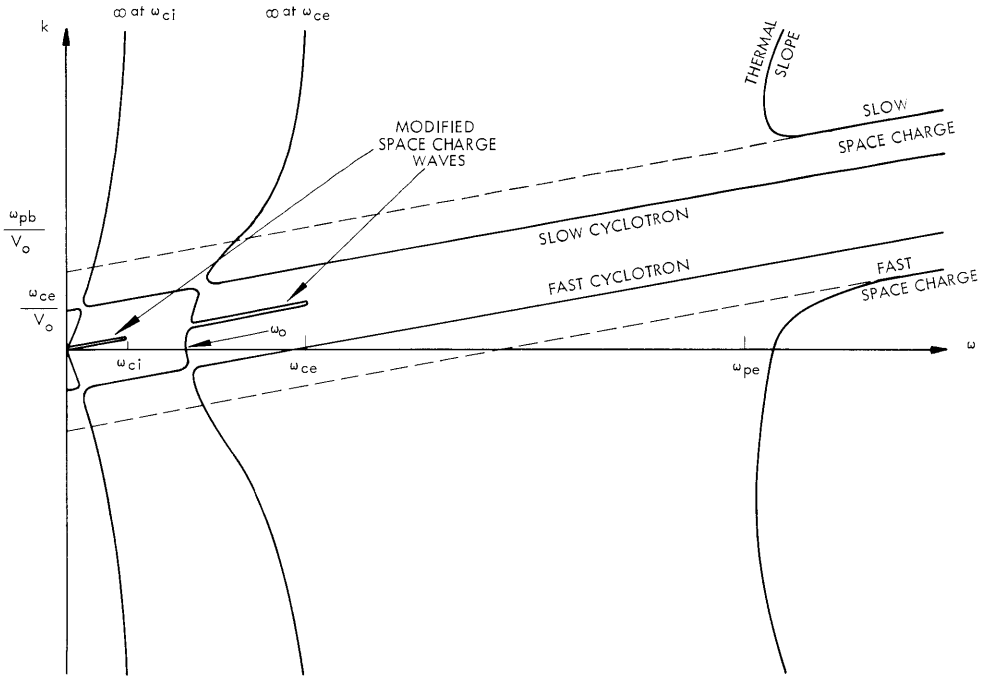

Fig. X-25. Strong-beam solution – $\omega_{pb} > \omega_{ce}$.

$a = 1 \times 10^6$ meters/sec

p = 100 /meter

$\omega_{pe} = 4$ Gc

$\omega_{pi} = 1$ Gc

$\omega_{pb} = 0.75$ Gc

$\omega_{ce} = 3$ Gc

$\omega_{ci} = 0.5$ Gc

For frequencies above the plasma frequency, $\sqrt{\omega_{pe}^2 + \omega_{pi}^2}$, the beam waves are easily distinguished. For frequencies below the plasma frequency, the cyclotron waves behave as coupling of modes would predict. The space-charge waves become modified, however, by the dielectric constant of the plasma.

For Fig. X-25, the parameters are as follows:

$V_o = 5 \times 10^7$ meters/sec

$a = 0.2 \times 10^7$ meters/sec

p = 1000 /meter

$\omega_{pe} = 90$ Gc

$\omega_{pi} = 2$ Gc

$\omega_{pb} = 10$ Gc

$\omega_{ce} = 3$ Gc

$\omega_{ci} = 0.001$ Gc

For this case, the beam plasma frequency was greater than the electron cyclotron frequency. The cyclotron waves behave as might be expected. The modification of the beam space-charge waves by the plasma below the plasma frequency is more evident here than in Fig. X-24.

A preliminary investigation of the convective growth rates predicted by this dispersion relation indicates that they are in the neighborhood of 1 db/cm in the region between ω_{ce} and $\sqrt{\omega_{pe}^2 + \omega_{pi}^2}$. Further work is being done to investigate the growth rate for different plasma parameters.

B. R. Kusse

References

1. R. J. Briggs, Electron-Stream Interaction with Plasmas (The M.I.T. Press, Cambridge, Mass., 1964); Special Technical Report Number 10, Research Laboratory of Electronics, M.I.T., Cambridge, Mass., 1964.

I. NONADIABATIC DIFFUSION IN TOROIDAL GEOMETRY*

Work continues on the construction of a race track-shaped magnetic field device for trapping an electron beam.[1]

In order to cancel drifts of the electrons, we impart a rotation to the field line in the "U" bends of the race track. These fields are produced by helical conductors that cause the field lines to rotate around the axis as they traverse the U bend; however, they leave at a different radius at the exit than at the entrance. To correct this, we add a field correction at the entrance and exit of the U bend which will make all field lines that enter on a circle leave on a circle, that is, circularizers. Figure X-26 shows the results of

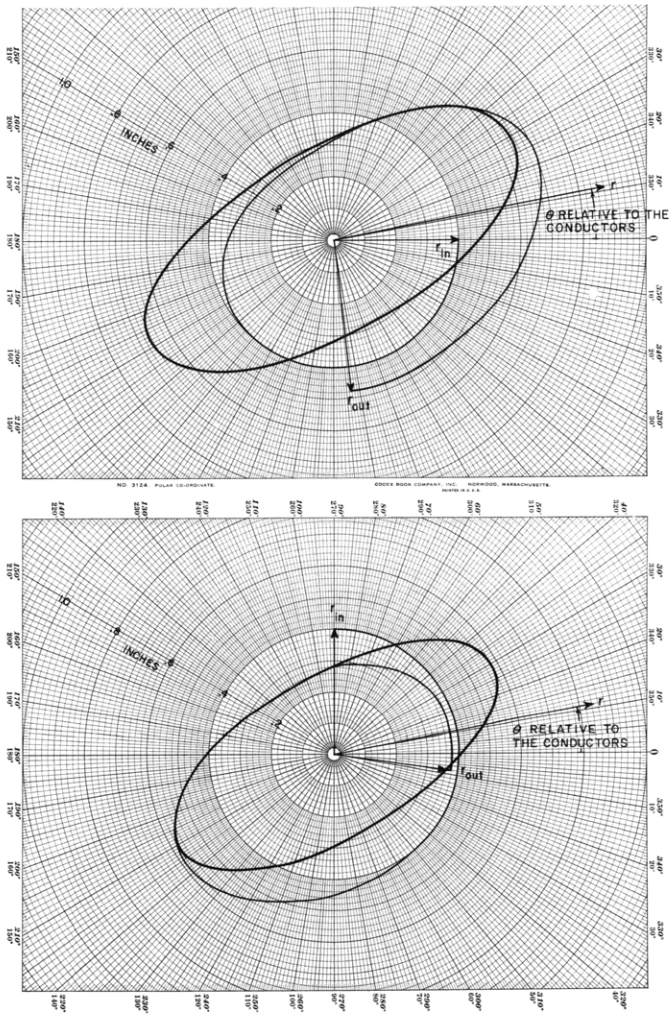

Fig. X-26. Preliminary results of circularizer calculations for two input conditions.

*This work was supported principally by the United States Atomic Energy Commission (Contract AT(30-1)-3285).

a circularizer calculation.[2] The field lines are plotted in a coordinate system rotating with the conductors; this takes advantage of the symmetry of the field. R_{in} is the radius of the field line at the entrance, and R_{out} is the radius at the exit. An ideal circularizer would have $R_{in} = R_{out}$.

We shall continue the development of a Rogowsky coil for measuring trapped beam current. A fluorescent screen and a fluorescent grid are being made for beam detection.

R. W. Moir, L. M. Lidsky

References

1. R. W. Moir and L. M. Lidsky, Quarterly Progress Report No. 77, Research Laboratory of Electronics, M. I. T., April 15, 1965, pp. 164-167.

2. These calculations were made possible by invaluable discussions and correspondence with K. Wakefield and U. Christensen of the Plasma Physics Laboratory, Princeton University, Princeton, New Jersey.

J. MAGNETIC NONADIABATIC SCATTERING IN THE MAGNETOSPHERE*

This report is an attempt to relate the studies of nonadiabatic scattering phenomena in magnetic-mirror machines to a possible scattering mechanism in the earth's magnetosphere. In a number of papers[1-4] electromagnetic waves have been proposed as scattering centers for particles in the earth's Van Allen belts. Such scattering is necessary, since the effect of particle-particle interactions is not sufficient to explain the observed density distribution in this region.

Helliwell[1] has proposed a synchronous acceleration of electrons trapped in phase with waves in the whistler mode. Parker[2] discounted the relative importance of this mechanism on energy grounds. He substituted a much smaller instantaneous nonadiabatic acceleration of the trapped particle as the whistler frequency passed through local cyclotron resonance. Since the whistler electric field is small and the energy transfer depends on phase which is essentially random, Parker proposed a diffusion coefficient to describe the mechanism. Later Dragt[3] presented similar arguments to arrive at the conclusion that protons are lost by scattering from hydromagnetic waves.

In general, then, the scattering mechanisms that have been proposed depend on some sort of cyclotron resonance of the particle with the wave's electric field. This type of interaction may be called an inelastic collision, since energy is added or subtracted from the particle. The elastic scattering from the wave's magnetic field is ignored. In the synchronous mechanism of Helliwell, the electron gains several Mev per interaction; under these conditions, neglecting the wave magnetic field is perhaps justifiable.

Under the conditions considered by Parker and Dragt, such an assumption is not justified. Plasma waves can have phase velocities much less than the speed of light, and the ratio of the magnetic force to the electric force on a particle is

$$vB/cE = v/v_p,$$

where v is the particle velocity, and v_p is the wave's phase velocity. This ratio can be large for conditions prevailing in the magnetosphere. One consequence of this is that even if the wave frequency does satisfy the cyclotron resonance condition, the energy transfer to the particle is bounded. The wave's magnetic field pushes the particle out of phase when the particle's velocity becomes the order of the wave phase velocity.

An estimate of the maximum energy change caused by the wave's electric field can easily be obtained.

If the electron's laboratory velocity is given by

$$\vec{v}_L = \vec{v}_z + \vec{v},$$

*This work was supported principally by the United States Atomic Energy Commission (Contract AT(30-1)-3285).

its velocity in a frame moving with the wave phase velocity is $\vec{v}_R = (\vec{v}_z - \vec{v}_p) + \vec{v}$. In this frame the particle sees only a static magnetic field, and its energy, $1/2\ mv_R^2$, is conserved. The wave magnetic field is unchanged if $v_p \ll c$, and it can only rotate the vector $\vec{v}_R$. The inelastic electric scattering is contained in the subsequent transformation back to the laboratory frame. The maximum energy change is

$$(\Delta W)_{max} = \frac{m}{2}\left(v^2 - v_R^2\right),$$

or

$$(\Delta W)_{max} \cong m v_p v_z.$$

For a 50-kev electron moving along the earth's field in the L = 2 shell scattering from a 5-k/sec whistler, we find

$$(\Delta W)_{max} \cong 200 \text{ volts.}$$

The magnetic elastic scattering is at least as large. If the particle satisfies the nonadiabatic resonance condition

$$2\pi(v_z - v_p) = \lambda\omega_B,$$

where λ is the whistler wavelength, it will experience a maximum magnetic scattering.[5] An estimate of the magnitude of this scattering can be obtained from Laing and Robson[6] who show that

$$\frac{v_\perp^2}{v_R^2} \propto (nh)^2,$$

where n is the number of wavelengths traveled before the particle falls off the resonant peak, and h is the ratio of the wave's magnetic field to the local earth's field. Previous computations have shown[6] that for n = 10 and h = 0.05,

$$\frac{v^2}{v_R^2} = 0.6.$$

The rotation of the particle's velocity vector in the field of a whistler (B=1γ) at L = 2 ($B_E \cong 10^3\gamma$) is, then,

$$\frac{v^2}{v_R^2} = \left(\frac{n\ 10^{-3}}{10 \times .05}\right)^2 0.6.$$

If we assume that the particle interacts with approximately 100 wavelengths,

$$\frac{v^2}{v_R^2} = 0.24.$$

This is sufficient to trap the particle. Nonresonant particles would be affected to a lesser extent, but it is clear that if the scaling used above applies, scattering from the wave's magnetic field may be a significant factor in determining the population of the radiation belts. The applicability of the scattering theory, which has been developed to describe particle confinement in mirror machine nonadiabatic fields to the earth's magnetosphere is being considered.

J. F. Clarke

References

1. R. A. Helliwell and T. F. Bell, J. Geophys. Res. 65, 1839 (1960).
2. E. N. Parker, J. Geophys. Res. 66, 2673 (1961).
3. A. J. Dragt, J. Geophys. Res. 66, 1641 (1961).
4. J. W. Dungey, Planet. Space Sci., Vol. II, p. 591, 1963.
5. J. F. Clarke, Quarterly Progress Report No. 77, Research Laboratory of Electronics, M. I. T., April 15, 1965, p. 171.
6. E. W. Laing and A. E. Robson, J. Nucl. Energy, Part C, 3, 146 (1961).

K. INCOHERENT SCATTERING OF LIGHT FROM A PLASMA. I.*

Preliminary signal-to-noise calculations for incoherent scattering of a laser beam from a plasma have been made, with the object of determining the feasibility of using a modulated argon continuous wave (cw) gas laser in place of a pulsed ruby system for a small-angle scattering experiment.

Spectral measurements of the light emission from a Hollow Cathode Discharge[1] hydrogen plasma were made by using a Jarrell-Ash monochromator, RCA 7265 photomultiplier, and a black-body emitter for calibrating the absolute intensity. A schematic view of the experiment arrangement is shown in Fig. X-27. At the wavelengths of interest, 5145 Å for an argon gas laser and 6943 Å for a ruby laser, the measured emission

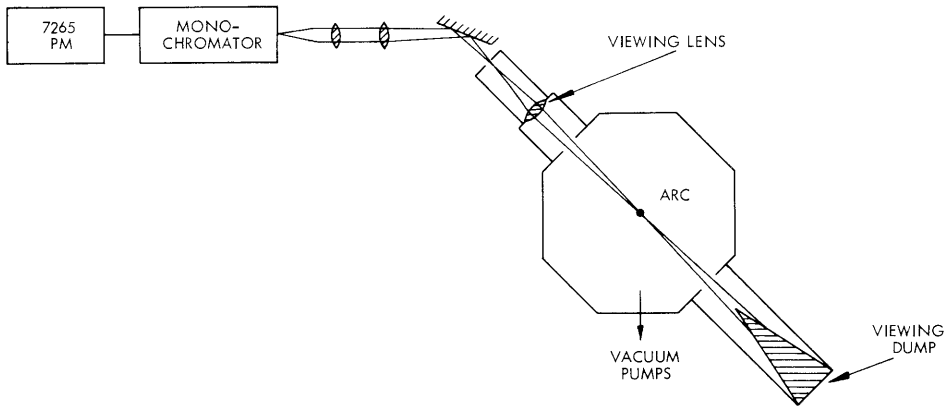

Fig. X-27. Observation of plasma luminosity.

is $\approx 10^{-8}$ watt Å^{-1} steradian^{-1} and $\approx 3 \times 10^{-8}$ watt Å^{-1} steradian^{-1}, respectively. Also, a 3 Å bandpass filter at 6940 Å was used with the photomultiplier output, either DC or AC coupled to an oscilloscope. With the geometry and photomultiplier properties known, the equivalent measured DC emission was 3×10^{-8} watt Å^{-1} steradian^{-1} at the ruby line (in agreement with the monochromator measurement); and the AC emission was $\approx 1.5 \times 10^{-11}$ watt Å^{-1} steradian^{-1}. With the detection system employed, the AC signal corresponds to the thermodynamic fluctuation level of the direct current. Since in fact the fluctuations in the light represent the noise of the system, the last figure is the important one.

One may calculate the expected emission from a hydrogen plasma, by computing the

*This work was supported principally by the United States Atomic Energy Commission (Contract AT(30-1)-3285).

Bremsstrahlung and radiative recombination[2] for the plasma of interest for scattering experiments, using the approximate electron density $n_e = 10^{14}$ cm^{-3}, and electron temperature $T_e = 2.4$ ev. We note that for recombination radiation to be centered near 5145 Å or 6943 Å, corresponding to 2.41 and 1.78 ev, respectively, only recombination to states $n \geq 3$ contributes. For $T_e = 2.4$ ev, it is readily determined that the free-bound radiation is of the order of the free-free radiation at 5145 Å. Calculating the free-free radiation, we get for the parameters above 2×10^{-8} watt Å^{-1} for a volume of 1 cm^3 at the argon line. Thus the expected emission for a fully ionized hydrogen plasma is $\approx 4 \times 10^{-8}$ watt Å^{-1} which compared with the measured value is nearly the same (the discrepancy depending on the effective solid angle of emission). For purposes of computing signal-to-noise we shall use the measured values.

A laser with its extremely coherent, monochromatic output enables measurement of wavelength shifts in the scattering of light waves from a plasma, and thus is receiving considerable attention for plasma diagnostics. The theory of scattering of electromagnetic radiation from a plasma indicates, that for a parameter $\alpha = \dfrac{1}{|\Delta \underline{k}| \lambda_D} = \dfrac{\lambda}{4\pi\lambda_D} \sin\dfrac{\theta}{2} > 1$ (where λ is the incident wavelength, λ_D is the Debye length, θ is the angle between incident and scattered wave vectors), the scattered power spectrum changes from one characteristic of free-electron scattering to one characterizing the density fluctuations in a plasma. As a result, an enhancement of scattering at a frequency shift $\Delta\nu \approx \omega_{pe}$ is predicted, and we are attempting to observe this scattered signal at $\theta = 2°$. This observation angle gives a value $\alpha = 2$ for $n_e = 10^{14}$ cm^{-3} and $T_e = 2.4$ ev. Neglecting the optical design to eliminate non plasma-scattered light from entering the detector, we can calculate the signal-to-noise ratio for plasma-scattered light versus plasma luminosity.

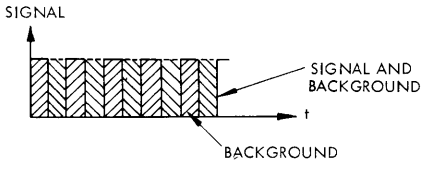

Fig. X-28. Calculation of signal-to-noise ratio.

Modulating the laser and detecting synchronously, we obtain ideally a signal-to-noise ratio from the following picture (Fig. X-28). We assume Poisson statistics to describe the shot-noise effect. In measuring the background B for a time t_b, we have a total count of

$$B = bt_p \pm \sqrt{bt_b},$$

where b = photons sec^{-1} emitted because of the self-luminosity of the plasma; and when detecting signal plus background for time t_s, the count is

$$S + B = (s+b)t_s \pm \sqrt{(s+b)t_s},$$

where s is scattered phonons $\sec^{-1}$. Subtracting, we find that the measured signal is

$$\langle S \rangle = [(s+b)t_s - bt_b] \pm \sqrt{(s+b)t_s + bt_b},$$

or for $t_s = t_b = \frac{t}{2}$, where t = integration time,

$$\langle S \rangle = s\frac{t}{2} \pm \sqrt{s\frac{t}{2} + 2b\frac{t}{2}}$$

for a signal-to-noise ratio

$$\left(\frac{S}{N}\right)_{eff} = \frac{s\frac{t}{2}}{\sqrt{\frac{st}{2} + bt}} = \frac{s/2}{\sqrt{s/2 + b}}\sqrt{t}\,.$$

For an angular resolution $\Delta\theta = 1/2°$, scattering length 1 cm, and incident power P_o at 5145 Å wavelength, the expected scattered power per angstrom, per unit solid angle, $P_s(\Omega, \lambda)$ is equal to or greater than 9×10^{-13} P_o watt Å^{-1} steradian^{-1}. Thus with an annular ring $\theta = 2°$, $\Delta\theta = 1/2°$ and $\Delta\lambda = 0.1$ Å (with a Fabry-Perot interferometer used for the required resolution) the scattered power detected is $P_s > 1.7 \times 10^{-16}$ P_o watt. The background luminosity detected is $P_L = 1.9 \times 10^{-12}$ watt. Since at 5145 Å the photon energy is 2.4 ev, the equivalent scattered signal and background are 450 P_o photons $\sec^{-1}$ and 5×10^6 photons $\sec^{-1}$ into the detection optics. If the transmission of a narrow bandpass filter is T_o, and that of a polarizer is T_1 for the polarized component and T_2 for the unpolarized (where $T_2 = T_1/2$, since 50 per cent of unpolarized light is in each polarization), and the quantum efficiency of a photocathode is ϵ; the scattered signal detected is 450 $P_o T_1 T_o \epsilon$ photons $\sec^{-1}$. Thus the signal-to-noise ratio is

$$\left(\frac{S}{N}\right) = \frac{s/2}{\sqrt{b}}\sqrt{t}$$

$$= \frac{225 P_o T_o T_1 \epsilon}{\sqrt{5 \times 10^6 T_2 T_o \epsilon}}\sqrt{t}$$

$$= 0.1 P_o T_1 \sqrt{\frac{\epsilon T_o}{T_2}}\,t,$$

since $s \ll b$. For $T_o = 0.5$, $T_1 = 0.8$, $T_2 = 0.4$, and $\epsilon = 0.1$ (PM response at 5100 Å) as typical values we have

$$\left(\frac{S}{N}\right) = 0.029 P_o \sqrt{t}\,.$$

Thus for P_o = 10 watts (at present an upper bound on the power available external to the cavity of an Argon gas laser), an integration time of 100 sec is necessary for $\left(\frac{S}{N}\right) = 2.9$. It thus becomes readily apparent that a cw laser at these relatively low powers would necessarily be operated intracavity; better than an order of magnitude in optical power is obtained, thereby reducing the integration time by two orders of magnitude.

The relative signal-to-noise calculations for pulsed ruby lasers may be calculated for (i) a 1-msec 100-joule pulse, and (ii) a Q-spoiled system with typically a 20-nsec 1-joule pulse.

The 1-msec pulse emits 3.5×10^{20} photons of energy at 1.8 ev, and the scattered number is $N_s = 9 \times 10^{-13}\ N_o(\Delta\Omega)$ photons $\text{Å}^{-1} = 6 \times 10^5$ photons Å^{-1} in 10^{-3} sec with the same angular factors. The background luminosity is $N_b = 1.6 \times 10^5$ photons Å^{-1}. For 0.1 Å bandwidth, the signal and noise are $N_s = 6 \times 10^4$ photons and $N_b = 1.6 \times 10^4$ photons. With the transmission factors T_o, T_1, T_2 and quantum efficiency ϵ, the signal is

$$\langle S \rangle = N_s T_o T_1 \epsilon \pm \sqrt{N_s T_o T_1 \epsilon + 2N_b T_o T_2 \epsilon}$$

for

$$\left(\frac{S}{N}\right) = \frac{N_s T_1 \sqrt{T_o \epsilon}}{\sqrt{N_s T_1 + 2N_b T_2}},$$

or since $T_o = 0.5$, $T_1 = 0.8$, $T_2 = 0.4$, $\epsilon \simeq 0.02$ (PM response at 6940 Å), $\left(\frac{S}{N}\right) = 19.5$.

Similarly, for a 20-nsec pulse of 1 joule, the scattered photons $N_s = 600$ and $N_b = 0$, whereby the signal-to-noise (leaving out the polarizer, since it is not required) is

$$\left(\frac{S}{N}\right) = \sqrt{N_s T_o \epsilon} = 2.5.$$

We may conclude, then, that a normally pulsed ruby or an argon cw gas laser is to be preferred over a Q-spoiled ruby laser, and that for reasonable integration times the argon and ruby lasers give comparable signal-to-noise.

In terms of beam divergence, however, the gas laser is distinctly superior to a pulsed ruby laser, and this is a significant factor when observing at small angles to the incident direction. Furthermore, the thermal shift in wavelength output of a ruby laser is an even more severe problem, since in fact 0.1 Å resolution is desired. Thus it may be concluded that a cw intracavity plasma scattering experiment is feasible, insofar as signal to noise is concerned (when the noise arises only from the luminosity of the plasma and not the stray scattered light). The desirability of the

cw experiment is increased when one considers the need for low-beam divergence and nonvariant wavelength. The attendant problems of beam divergence and scattering of laser light from surfaces in the arc are to be the subject of more intensive investigation.

A. A. Offenberger

References

1. E. T. Gerry, Quarterly Progress Report No. 77, Research Laboratory of Electronics, M.I.T., April 15, 1965, pp. 181-182.

2. M. J. Seaton, Reports on Progress in Physics, Vol. 23, p. 313, 1960.

L. AN INPUT-OUTPUT APPROACH TO THE PROBLEMS OF OPTIMAL CONTROL

1. Introduction

A new method has been developed for the solution of certain types of optimal control problems by taking the system description as an input-output relation. The system is assumed to be represented by[1]

$$\underline{y}(t) = \underline{H}(u(t)) = \sum_{j=1}^{J} \sum_{i=1}^{i_{j,r}} \int_{t_o}^{t} \cdots \int \underline{h}_{ij}(\sigma_1, \sigma_2, \ldots, \sigma_j) \cdot C_i(\underline{u}; \sigma_1, \sigma_2, \ldots, \sigma_j)\, d\sigma_1 d\sigma_2 \ldots d\sigma_j, \tag{1}$$

where the input $\underline{u}(t)$ and the output $\underline{y}(t)$ are r- and n-vectors, respectively, the n-vector $\underline{h}_{ij}(\sigma_1, \sigma_2, \ldots, \sigma_j)$ is the i^{th} component of the j^{th}-order kernel of the system, $C_i(\underline{u}; \sigma_1, \sigma_2, \ldots, \sigma_j)$ is the i^{th} possible j-tuple product of the components of the input vector u(t), each factor in the product having σ_1, σ_2, ... or σ_j as its argument, and $i_{j,r}$ is the total number of such products.

A typical statement of an optimal control problem may be: Given a system described by the input-output relation, Eq. 1, find the input $\underline{u}(t)$ that brings the output to some prescribed terminal value and, at the same time, minimizes the cost functional

$$\mathcal{J}(\underline{u}(t)) = \int_{t_o}^{t_1} \mathcal{L}(\underline{y}(t), \underline{u}(t))\, dt, \tag{2}$$

where $\mathcal{L}$ is a scalar-valued operator. The input $\underline{u}(t)$ belongs to some constrained or unconstrained space, and the terminal time, t_1, may be specified or left unspecified.

The significance of approaching an optimal control problem through an input-output type of representation lies mainly in the fact that a direct measurement of the characteristics of an unknown system results, in general, in an input-output relation. In some cases, however, even for the systems whose characteristics are known through a set of differential equations, the input-output approach to the optimal control problems significantly lessens the computational complexity, and more than offsets the trouble of converting the differential equations into an input-output form. It is well known that the systems described by a set of linear differential equations have an equivalent representation in the form of input-output relations. It has recently been reported that a class of nonlinear systems which can be represented exactly by a finite number of linear systems and multipliers also possess such a property.[2] Another potential significance of the present approach is that we can write the cost functional as an explicit function of the input

$$\mathcal{J}(\underline{u}(t)) = \int_{t_o}^{t_1} \mathcal{L}(\underline{H}(\underline{u}(t)), \underline{u}(t))\, dt. \tag{3}$$

The explicit dependence of Eq. 3 on $\underline{u}(t)$ may facilitate the optimal control problem at hand, particularly when $\underline{H}$ is nonlinear.

We shall develop the method of solving such optimal control problems. We shall consider the case in which $\underline{u}(t)$ is unconstrained and the cost functional is

$$\mathcal{J} = \int_{t_o}^{t_1} <\underline{u}(t), \underline{u}(t)>\, dt, \tag{4}$$

where $<\underline{a}, \underline{b}>$ is the scalar product of $\underline{a}$ and $\underline{b}$. Linear systems with constrained input will also be considered. The results are found to be the same as those derived by other workers by means of the maximum principle.[3]

2. Unconstrained Input Problems

We shall now consider only multiple-input linear systems and single-input nonlinear systems. The extension of the method to multi-input nonlinear systems is conceptually straightforward, except that we have to introduce much more involved algebraic and notational complexities.

a. Linear Systems

Let the dynamical system be represented by

$$\underline{y}(t) = \sum_{i=1}^{r} \int_{t_o}^{t} \underline{h}_i(t; \sigma)\, u_i(\sigma)\, d\sigma \tag{5}$$

We wish to bring the output vector to $\underline{\eta}$ at a given terminal time t_1. The problem is to find the input vector $\underline{u}(t)$, $t_o \leq t \leq t_1$, such that the terminal condition

$$\underline{y}(t_1) = \underline{\eta} \tag{6}$$

is satisfied and the cost functional

$$\mathcal{J} = \int_{t_o}^{t_1} <\underline{u}(t), \underline{u}(t)>\, dt = \int_{t_o}^{t_1} \sum_{i=1}^{r} u_i^2(t)\, dt \tag{7}$$

is minimized.

Writing Eq. 5 at the terminal time, we have

$$\underline{\eta} = \sum_{i=1}^{r} \int_{t_o}^{t_1} \underline{h}_i(t; \sigma)\, u_i(\sigma)\, d\sigma. \tag{8}$$

At the outset, let us assume that all of the n-elements of $h_i(t_1; \sigma)$ are of the same functional character. We shall say that the two functions a(t) and b(t) are of the "same functional character" if both can be written as a finite linear sum of the members of a consistent infinite set of linearly independent functions. For instance,

$$1, t, t^2, \ldots, t^n, \ldots$$

form a consistent infinite set, and the functions

$$a + bt^2$$

and

$$ct + dt^3 + et^4$$

are of the same functional character. We call such a consistent infinite set a "characteristic set" to a(t) and b(t), and the orthonormal set of functions $\{\phi_n(t)\}$ built up from this characteristic set a "characteristic orthonormal set." The range of orthonormality is $[t_o, t_1]$. With this nomenclature, let us expand each element of the vector $h_i(t_1; \sigma)$ and $u_i(\sigma)$ in terms of the characteristic orthonormal set.

$$h_{ij}(t_1; \sigma) = \sum_{\ell=1}^{L_j^i} A_\ell^{ij} \phi_\ell^i(\sigma) \tag{9}$$

$$u_i(\sigma) = \sum_{\ell=1}^{\infty} B_\ell^i \phi_\ell^i(\sigma), \tag{10}$$

where $\{\phi_n^i(\sigma)\}$ is the orthonormal set of functions characteristic to the index i.

Substituting Eqs. 9 and 10 in Eq. 8, we obtain

$$\eta_1 = \sum_{i=1}^{r} \sum_{\ell=1}^{L_1^i} A_\ell^{i1} B_\ell^i$$

$$\eta_2 = \sum_{i=1}^{r} \sum_{\ell=1}^{L_2^i} A_\ell^{i2} B_\ell^i \tag{11}$$

. . . . . .

The inputs of the form (10) will satisfy the system and the terminal conditions exactly if the coefficients B_ℓ^i, i = 1, 2, ..., r and ℓ = 1, 2, ..., L^i, satisfy the relations (11),

where L^i is the largest of the L^i_j for j = 1, 2, ..., n. Clearly, the coefficients B^i_ℓ for $\ell > L^i$, i = 1, 2, ..., r, may have arbitrary values. Therefore, it follows that the set of feasible inputs W is the set of all inputs of the form (10) whose coefficients satisfy (11). If W is nonempty, it contains an infinite number of elements.

Let us now consider Eq. 11. There are

$$L = \sum_{i=1}^{r} L^i \tag{12}$$

unknown coefficients B^i_ℓ to determine, while there are n-relations among them. Three situations may arise.

(i) $L < n$. Since the n-relations of (11) are linearly independent, we find that it is impossible to satisfy the system and the terminal conditions. In other words, the feasible set W is empty, and there does not exist an input that performs the prescribed task.

(ii) $L = n$. In this case, we can uniquely determine the coefficients.

(iii) $L > n$. This will be the most usual case encountered in practice. We determine the first n of the L coefficients in terms of the rest. Let us write this

$$B_k = f_k(B_{n+1}, B_{n+2}, \ldots, B_L) \qquad k = 1, 2, \ldots, n. \tag{13}$$

Here, we have renumbered the coefficients B^i_ℓ by affixing a single index k, k = 1, 2, ..., n.

Now, using (10) in (7), we find

$$\mathcal{J} = \sum_{\ell=1}^{L} B_\ell^2 + \sum_{\ell=L+1}^{\infty} B_\ell^2 \tag{14}$$

for L = n, and

$$\mathcal{J} = \sum_{\ell=1}^{n} f_\ell^2(B_{n+1}, B_{n+2}, \ldots, B_L) + \sum_{\ell=n+1}^{L} B_\ell^2 + \sum_{\ell=L+1}^{\infty} B_\ell^2 \tag{15}$$

for L > n, using the same renumbering of the coefficients. It is clear that the cost functional, Eq. 14 or Eq. 15, is continuously differentiable in all of the coefficients B_ℓ. Furthermore, the extremum of $\mathcal{J}$ with respect to B_ℓ for any given ℓ is attained at the values of B_ℓ given by

$$\frac{\partial \mathcal{J}}{\partial B_\ell} = 0. \tag{16}$$

This is unique by virtue of the simple quadratic dependence of $\mathcal{J}$ on B_ℓ. We also assert

that this unique extremum is a minimum by noting that

$$\frac{\partial^2 \mathcal{J}}{\partial B_\ell \partial B_m} = \begin{cases} 2 & \ell = m \\ 0 & \ell \neq m. \end{cases} \tag{17}$$

Let us consider the determination of the optimal input for the case $L > n$. The case of $L = n$ will become obvious. In order to find the minimum of (15), let us differentiate it with respect to B_p, $p \geq n + 1$, and set the derivatives to be zero.

$$\frac{\partial \mathcal{J}}{\partial B_p} = 2\left[\sum_{\ell=1}^{n} f_\ell(B_{n+1}, \ldots, B_L) \frac{\partial f_\ell}{\partial B_p}(B_{n+1}, \ldots, B_L) + B_p\right] = 0$$

$$p = n+1, n+2, \ldots, L \tag{18}$$

$$\frac{\partial \mathcal{J}}{\partial B_p} = 2B_p = 0 \qquad p > L \tag{19}$$

Equations 13, 18, 19 provide all of the necessary relations to completely determine the optimal input. If we denote the coefficients obtained by solving (13) and (18) by B_ℓ^{i*}, the optimal input is

$$u_i^*(t) = \sum_{\ell=1}^{L^i} B_\ell^{i*} \phi_\ell^i(t) \qquad i = 1, 2, \ldots, r. \tag{20}$$

EXAMPLE 1: Consider a system given by

$$x_1(t) = \xi_1 + t\xi_2 + \int_0^t (t-\sigma)\, u(\sigma)\, d\sigma$$

$$x_2(t) = \xi_2 + \int_0^t u(\sigma)\, d\sigma.$$

Let us find the input that brings the outputs to the origin at $t = T$, and minimizes

$$\mathcal{J} = \int_0^T u^2(\sigma)\, d\sigma$$

At the terminal time

$$-\xi_1 - T\xi_2 = \int_0^T (T-\sigma)\, u(\sigma)\, d\sigma$$

$$-\xi_2 = \int_0^T u(\sigma)\, d\sigma.$$

The characteristic set of orthonormal functions is

$$\phi_1(t) = \sqrt{\frac{1}{T}}, \quad \phi_2(t) = \sqrt{\frac{3}{T}} - \sqrt{\frac{12}{T^3}}\, t, \quad \ldots .$$

Expanding the kernels in terms of this set, writing

$$u(\sigma) = \sum_{\ell=1}^{\infty} B_\ell \phi_\ell(\sigma),$$

and using these in the equations written at t = T, we have

$$B_1 = -\frac{\xi_2}{\sqrt{T}}, \qquad B_2 = -\sqrt{\frac{12}{T^3}}\left(\xi_1 + \frac{T\xi_2}{2}\right), \qquad B_\ell = \text{arbitrary for } \ell \geq 3.$$

Minimization of $\mathcal{J} = \sum_{\ell=1}^{\infty} B_\ell^2$ gives $B_\ell = 0$ for $\ell \geq 3$. Therefore, the optimal input is

$$u^*(t) = B_1\phi_1(t) + B_2\phi_2(t)$$

$$= -\frac{2}{T^2}(3\xi_1 + 2\xi_2 T) + \frac{6}{T^3}(2\xi_1 + \xi_2 T)\, t.$$

b. Nonlinear Systems

A single-input, multioutput nonlinear system is represented by

$$\underline{y}(t) = \int_{t_o}^{t} \underline{h}_1(t;\sigma)\, u(\sigma)\, d\sigma + \int_{t_o}^{t}\int \underline{h}_2(t;\sigma_1,\sigma_2)\, u(\sigma_1)\, u(\sigma_2)\, d\sigma_1 d\sigma_2 + \ldots . \tag{21}$$

The m^{th}-order kernel $\underline{h}_m(t;\sigma_1,\sigma_2,\ldots,\sigma_m)$ can be written,[4] in general, as

$$\underline{h}_m(t;\sigma_1,\sigma_2,\ldots,\sigma_m) = \underline{k_{m1}(t;\sigma_1)\, k_{m2}(t;\sigma_2)\ldots k_{mm}(t;\sigma_m)}. \tag{22}$$

We assume, as we did for the linear systems, that all of the kernels are of the same functional character. In a multi-input problem, we shall only require that the kernels operating on the same component of the input be of the same functional character. Proceeding in the same manner as in the linear problem, we expand the kernels and the input in terms of the characteristic orthonormal set

$$k_{jk}^{i}(t;\sigma_k) = \sum_{\ell=1}^{L_{jk}^{i}} A_\ell^{ijk}\phi_\ell(\sigma_k), \tag{23}$$

where the superscript i denotes the index corresponding to the component of the output.

$$u(\sigma_k) = \sum_{\ell=1}^{\infty} B_\ell \phi_\ell(\sigma_k). \tag{24}$$

Using (23) and (24) in (21) written at the terminal time t = T, we obtain

$$\eta_j = \sum_{\ell=1}^{L_{11}^j} A_\ell^{1j1} B_\ell + \left(\sum_{\ell=1}^{L_{21}^j} A_\ell^{2j1} B_\ell \right) \left(\sum_{\ell=1}^{L_{22}^j} A_\ell^{2j2} B_\ell \right) + \ldots \qquad j = 1, 2, \ldots, n. \tag{25}$$

We have in Eq. 25 a set of simultaneous, nonlinear algebraic equations. It is not a simple matter to solve (25) for B_ℓ, but being a set of algebraic equations it is still much more easily amenable than the corresponding nonlinear differential or integral equations. In fact, closed-form solutions are possible for a large class of second- or third-order nonlinear systems, and numerical trial-and-error solution on a machine is relatively straightforward.

The rest of the procedure is the same as for the linear systems, except for some minor modifications arising in minimizing the cost functional, Eq. 7, such as the possibility of nonunique relations of the type Eq. 13. These modifications will not be discussed here.

EXAMPLE 2: Consider a nonlinear system given by

$$x_1(t) = \int_0^t (t-\sigma)\, u(\sigma)\, d\sigma + \int_0^t \int (t-\sigma_1)(t-\sigma_2)\, u(\sigma_1)\, u(\sigma_2)\, d\sigma_1 d\sigma_2$$

$$x_2(t) = -2 \int_0^t u(\sigma)\, d\sigma + \int_0^t \int u(\sigma_1)\, u(\sigma_2)\, d\sigma_1 d\sigma_2.$$

Suppose that we desire to find the input that brings the output to (1, -1) at t = 1, and minimizes

$$\mathcal{J} = \int_0^1 u^2(\sigma)\, d\sigma.$$

The characteristic orthonormal set is that of Example 1 with T = 1. Expanding all the kernels and the input in terms of this set, and using these expansions in the equation at t = 1, we have

$$\frac{1}{2}\left(B_1 + \frac{1}{\sqrt{3}} B_2\right) + \frac{1}{4}\left(B_1 + \frac{1}{\sqrt{3}} B_2\right)^2 = 1$$

$$-2B_1 + B_1^2 = -1.$$

Solving these equations for B_1 and B_2, we obtain

$$B_1 = 1, \quad B_2 = -1 \pm \frac{1}{2}\sqrt{4+\sqrt{12}}.$$

Minimization of $\mathcal{J} = \sum_{\ell=1}^{\infty} B_\ell^2$ gives the optimal input

$$u^*(t) = 1 + \left(-1 + \frac{1}{2}\sqrt{4+\sqrt{12}}\right)(\sqrt{3}-\sqrt{12}\,t).$$

3. Constrained Input Problems

We shall now consider the same system that was considered above, except that the input is constrained to be

$$\|\underline{u}(t)\|_p \leq \kappa, \tag{26}$$

where the norm is defined as

$$\|\underline{u}(t)\|_p = \left[\int_0^T \sum_{i=1}^{r} |u_i(t)|^p \, dt\right]^{1/p}. \tag{27}$$

The discussion will be limited to a linear time-optimal problem only, that is,

$$\mathcal{J} = \int_0^T dt = T. \tag{28}$$

The system is again given by Eq. 5. We desire to bring the output vector to $\underline{\eta}$ in a minimum possible time with an input satisfying Eq. 26.

At the terminal time,

$$\underline{\eta} = \sum_{i=1}^{r} \int_0^T \underline{h}_i(T;\sigma)\, u_i(\sigma)\, d\sigma. \tag{29}$$

Let us multiply both sides of (29) by a yet unspecified constant n-vector $\underline{g} = (g_1, g_2, \ldots, g_n)$. Then,

$$\langle \underline{g}, \underline{\eta} \rangle = \sum_{i=1}^{r} \int_0^T \langle \underline{g}, \underline{h}_i(T;\sigma) \rangle\, u_i(\sigma)\, d\sigma \tag{30}$$

or

$$\langle \underline{g}, \underline{\eta} \rangle = \int_0^T \langle \underline{g}, \underline{H}(T;\sigma) \rangle\, \underline{u}(\sigma)\, d\sigma, \tag{31}$$

where

$$\left\langle \underline{g}, \underline{H} \right\rangle = \left[\sum_{j=1}^{n} g_j h_{1j} \sum_{j=1}^{n} g_j h_{2j} \cdots \sum_{j=1}^{n} g_j h_{rj} \right]. \tag{32}$$

If it is assumed that all of the functions appearing in (29) possess a norm of the type given by (27), then

$$\begin{aligned} |\langle \underline{g}, \underline{\eta} \rangle| &\leq \int_0^T \left| \left\langle \underline{g}, \underline{H}(T; \sigma) \right\rangle u(\sigma) \right| d\sigma \\ &\leq \|\underline{u}(\sigma)\|_p \left\| \left\langle \underline{g}, \underline{H}(T; \sigma) \right\rangle \right\|_q, \end{aligned} \tag{33}$$

where

$$\frac{1}{p} + \frac{1}{q} = 1.$$

From this, we have

$$\|\underline{u}(t)\|_p \geq \frac{|\langle \underline{g}, \underline{\eta} \rangle|}{\left\| \left\langle \underline{g}, \underline{H}(T; t) \right\rangle \right\|_q}. \tag{34}$$

Combining Eqs. 26 and 33, yields

$$\kappa \geq \|\underline{u}(t)\|_p \geq \frac{|\langle \underline{g}, \underline{\eta} \rangle|}{\left\| \left\langle \underline{g}, \underline{H}(T; t) \right\rangle \right\|_q} \tag{35}$$

or

$$\left\| \left\langle \underline{g}, \underline{H}(T; t) \right\rangle \right\|_q \geq \frac{|\langle \underline{g}, \underline{\eta} \rangle|}{\kappa}. \tag{36}$$

For a given $\underline{g}$, the norm $\left\| \left\langle \underline{g}, \underline{H}(T; t) \right\rangle \right\|_q$ is a monotonically increasing function of T. Therefore, the minimum of $\left\| \left\langle \underline{g}, \underline{H}(T; t) \right\rangle \right\|_q$ over T implies the minimum of T. In order for $\left\| \left\langle \underline{g}, \underline{H}(T; t) \right\rangle \right\|_q$ to be the minimum in Eq. 36,

$$\left\| \left\langle \underline{g}, \underline{H}(T; t) \right\rangle \right\|_q = \frac{|\langle \underline{g}, \underline{\eta} \rangle|}{\kappa}. \tag{37}$$

This is true if (35) holds with the equalities,

$$\kappa = \|\underline{u}(t)\|_p = \frac{|\langle \underline{g}, \underline{\eta} \rangle|}{\left\| \left\langle \underline{g}, \underline{H}(T; t) \right\rangle \right\|_q}. \tag{38}$$

Since (31) holds for arbitrary $\underline{g}$, it must hold for $\underline{g} = \hat{\underline{g}}$ to satisfy

$$\langle \underline{g}, \underline{\eta} \rangle = 1 \tag{39}$$

and minimize $\left\| \left\langle \underline{g}, \underline{H}(T;t) \right\rangle \right\|_q$. Such a $\hat{\underline{g}}$ does exist if all components of $\left\langle \underline{g}, \underline{H}(T;t) \right\rangle$ vanish, at most, at some isolated points t_i, $0 \leq t_i \leq T$.[5,6] Equation 38 holds if

$$u_i(t) = C_i(t) \operatorname{sgn}\left(\sum_{j=1}^{n} \hat{g}_j h_{ij}(T;t) \right), \tag{40}$$

where the coefficients $C_i(t)$ are determined from (38). The use of (40) in (38) yields

$$\kappa = \left\| \underline{u}(t) \right\|_p = \left[\int_0^T \sum_{i=1}^{r} |C_i(t)|^p \, dt \right]^{1/p}. \tag{41}$$

Noting that

$$p = \frac{q}{q-1}, \qquad \frac{1}{p} = \frac{q-1}{q},$$

we find

$$C_i(t) = \kappa^q \left| \sum_{i=1}^{r} \hat{g}_j h_{ij}(T;t) \right|^{q-1}. \tag{42}$$

Hence the time-optimal input is

$$u_i^*(t) = \kappa^q \left| \sum_{i=1}^{r} \hat{g}_j h_{ij}(T;t) \right|^{q-1} \operatorname{sgn}\left(\sum_{j=1}^{n} \hat{g}_j h_{ij}(T;t) \right) \qquad i = 1, 2, \ldots, r. \tag{43}$$

We still must determine the vector $\hat{\underline{g}}$. The technique of nonlinear programming may be used to this end.[5] For a large class of problems, $\hat{\underline{g}}$ can easily be found by a direct computation, as illustrated in Example 3.

If the input constraint is of the form

$$\max_i |u_i(t)| \leq \kappa \tag{44}$$

that is, $p \to \infty$ and $q \to 1$ in Eq. 27, Eq. 43 becomes

$$u_i^*(t) = \kappa \operatorname{sgn}\left(\sum_{j=1}^{n} \hat{g}_j h_{ij}(T;t) \right) \qquad i = 1, 2, \ldots, r. \tag{45}$$

EXAMPLE 3: Consider the same system as in Example 1. We desire that the output of the system be brought to the origin in the minimum time by using an input whose magnitude is constrained by $|u(t)| \leq 1$. Direct use of (45) gives

$$u^*(t) = \operatorname{sgn}\left[\hat{g}_1(T-\sigma)+\hat{g}_2\right].$$

The vector $\hat{\underline{g}}$ is determined as follows:

$$-\hat{g}_1(\xi_1+T\xi_2) - \hat{g}_2\xi_2 = 1, \qquad \hat{g}_1 = -\frac{1+\hat{g}_2\xi_2}{\xi_1+T\xi_2}.$$

Minimizing the norm

$$\left\|\sum_{j=1}^{2}\hat{g}_j h_j(T;t)\right\|_1 = \int_0^T \left|\hat{g}_1(T-\sigma)+\hat{g}_2\right| d\sigma$$

or

$$\left\|\sum_{j=1}^{2}\hat{g}_j h_j(T;t)\right\|_1 = \int_0^T \left|\hat{g}_1 t+\hat{g}_2\right| dt$$

directly, we obtain

$$\hat{g}_1 = -\frac{1}{\xi_1+T\xi_2}\left[1+\xi_2\left(\frac{(T^2\xi_2+3T\xi_1)/2}{T^2\xi_2^2+3T\xi_1\xi_2+3\xi_1}\right)\right]$$

$$\hat{g}_2 = \frac{(T^2\xi_2+3T\xi_1)/2}{T^2\xi_2^2+3T\xi_1\xi_2+3\xi_1}.$$

Here, the terminal time T is found from

$$\left\|\sum_{j=1}^{2}\hat{g}_j h_j(T;t)\right\|_1 = \frac{1}{\kappa} = 1.$$

The results in this report are the same as the known results derived by means of maximum principle by using the corresponding differential equation description of the systems. The possibility of extending the concepts of the last part of this report to nonlinear systems is now under investigation.

S. H. Kyong, E. P. Gyftopoulos

References

1. S. H. Kyong, "Problems in the Theory of Optimal Control of Nonlinear Systems," Quarterly Progress Report No. 77, Research Laboratory of Electronics, M. I. T., April 15, 1965, pp. 202-203.

2. A. M. Bush, "On the Relation between Integral and Differential Characterization of Nonlinear Systems," Quarterly Progress Report No. 77, Research Laboratory of Electronics, M. I. T., April 15, 1965, pp. 265-269.

3. L. S. Pontryagin, et al., The Mathematical Theory of Optimal Processes (Interscience Publishers, New York, 1962).

4. D. A. Chesler, "Nonlinear Systems with Gaussian Inputs," Technical Report 366, Research Laboratory of Electronics, M. I. T., Cambridge, Mass., February 15, 1960.

5. F. M. Kirillova, "A Limiting Process in the Solution of an Optimal Control Problem," Prik. Mat. Mekh. 24, 2, 398-405 (1960).

6. L. W. Neustadt, "Optimization, A Moment Problem, and Nonlinear Programming," J. SIAM Control, Ser. A, Vol. 2, No. 1, pp. 33-53, 1964.