THE DESIGN OF LINEAR ACCELERATORS

J. C. SLATER

TECHNICAL REPORT NO. 47

SEPTEMBER 2, 1947

RESEARCH LABORATORY OF ELECTRONICS
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

The research reported in this document was made possible through support extended the Massachusetts Institute of Technology, Research Laboratory of Electronics, jointly by the Army Signal Corps, the Navy Department (Office of Naval Research), and the Army Air Forces (Air Materiel Command), under the Signal Corps Contract No. W-36-039 sc-32037.

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
Research Laboratory of Electronics

Technical Report No. 47 September 2, 1947

THE DESIGN OF LINEAR ACCELERATORS

by

J. C. Slater

- Table of Contents and Abstract -

groups of particles travel faster or slower than the wave, and are not locked to it. With a wave traveling with the velocity of light, some particles are locked to the wave, asymptotically approaching a fixed phase, in which they continually gain energy. In a wave whose velocity increases to match the acceleration of particles, stable bunches are formed, as in the first case, but now about positions corresponding to acceleration of the particles. As the velocity approaches that of light, the bunches decrease in size.

In the M.I.T. accelerator, electrons will be injected at two million electron-volts energy, into a tube where the velocity equals that of light, so that many particles will be locked to the wave in asymptotic phases, continually gaining energy. In some other project, including the positive-ion accelerator, particles are injected at lower velocity in a tube whose velocity increases, resulting in stable accelerated bunches.

In case stable bunches of particles are formed, a focussing instability is necessarily present. This defocussing is not met in the M.I.T. accelerator, for it vanishes at the velocity of light, as the stability of bunches also vanishes. In electron accelerators with injection at less than the velocity of light, and subsequent acceleration, defocussing is present while the velocity is increasing, but can be easily counteracted by a longitudinal magnetic field. In positive-ion accelerators, however, the defocussing effect can be very serious. In addition to the defocussing of particles displaced from the axis, there can be spreading of the beam on account of the spread of directions in the injected beam. It is shown that for an electron accelerator this spreading is not serious.

THE DESIGN OF LINEAR ACCELERATORS

The linear accelerator is a device for accelerating electrons or positively charged particles to a high energy, by the application of an alternating rather than a direct field, and in a straight line, rather than in an orbit curved by a magnetic field as in the cyclotron, betatron, or synchrotron. Its essential principle is some form of loaded waveguide, in which an oscillating field of high amplitude can be set up, which can be analyzed into traveling waves, one of which travels with the velocity of the particle to be accelerated. Particles in the correct phase can then remain always in the phase of the traveling wave corresponding to acceleration, and can pick up energy continually, just as if they were in a constant field. If the properties of the field are arranged to vary, as the particle goes along the tube, so as to keep step with the acceleration of the particle, energies of any amount can in principle be acquired.

The main advantages and disadvantages of the linear accelerator are obvious from its form. As advantages, we may list the saving of the expense of the large magnet, which is necessary in the circular machines; the fact that the size and expense of the machine is roughly proportional to the final energy of the particle, rather than to a higher power of the energy, as in circular machines, suggesting that at least for very high energies it may be a more economical device than the circular machines; and the fact that the particles will automatically emerge in a well-collimated beam, whereas beam ejection is one of the principal difficulties in the circular machines. The principal disadvantage of the linear accelerator is the fact that an individual particle, instead of passing through the same alternating field again and again, and using the same power source and accelerating gaps many times, as in the cyclotron and other circular machines, must pass through a succession of alternating fields and a succession of power sources. This multiplicity of sources and fields results in expense, in a large duplication of high-frequency equipment; in complication, in the construction of a very long and elaborate tube with its adjacent power sources; and in a design difficulty, in the adjustment of the phase of the oscillating field over a great length of accelerator, so as to insure that the field will stay in step with the particles.

There is another design difficulty not shared by the circular machines, which does not appear until one makes a little mathematical analysis. In any accelerator, the particles must travel a very long distance, either in a straight line or a circle or a spiral, before they acquire the energies about which one talks in present discussion of hundreds of millions or billions of electron-volts. In this very long path, they are likely to spread from their ideal path, and become lost to the beam. The spreading can be of two sorts. First, they can spread laterally, or be defocussed. Secondly, since the operation of all these devices except the betatron depends on having the particles bunched longitudinally, in bunches a wavelength apart, in the proper phase of a traveling wave to be accelerated, the particles can spread longitudinally, getting ahead or behind their bunches, and can be debunched. Now in the cyclotron and synchrotron, those particles which are in the correct phase to be bunched and accelerated are also in the right phase to be focussed, so that stability is automatic. On the other hand, in the linear accelerator, particles in the phase for bunching and acceleration are defocussed. This defocussing imposes a design problem of serious proportions for the linear accelerator.

It is too early to evaluate completely the advantages versus the disadvantages and difficulties of the linear accelerator. If the circular machines in the billion-volt range were perfectly easy to build, there would be no question butthat they would be the preferred devices, because the disadvantages and difficulties of the linear accelerator, as just enumerated, are formidable. But all the recent thinking about circular machines has shown that their difficulties are formidable as well. Among electron accelerators, present thinking indicates that the synchrotron is preferred to the betatron for energies more than a few hundred million electron-volts, but the synchrotron is believed to face a practical limit, on account of radiative losses by the electrons, at about a billion electron-volts.[1] No device except the linear accelerator has been suggested to go beyond this range. As for proton and other positive ion accelerators, the ordinary cyclotron meets relativistic difficulties at about a hundred million electron-volts. The frequency-modulated or synchro-cyclotron is ideally adapted for energies of several hundred million electron-volts.[2] Attempts at design of higher energy machines of this type, however, show that the dee structure and the magnet design begin to face problems at about six hundred million electron-volts which become very serious indeed at a billion electron-volts, and probably are insurmountable much above that energy.

For the range of several billion electron-volt positive ions, present thinking suggests only one alternative to the linear accelerator: the proton synchrotron, a device in which both magnetic field and frequency are simultaneously modulated, during the cycle of acceleration of the particles, in such a way as to keep the radius of the particle's orbit constant, and yet to keep the particle in step with the radio-frequency field. This has the advantage that only an annular magnet is required, as in the synchrotron, so that the magnet cost does not go up so rapidly with energy as in the synchro-cyclotron, with its solid magnet. Furthermore, the accelerating electrodes, or dees, can be relatively small, and the frequencies encountered are so low that the frequency modulation is simple. Its difficulties are nevertheless very great. The magnet, even though annular, is exceedingly large and expensive. The mutual adjustment, in time during the accelerating cycle, of frequency of oscillator and magnetic field, to keep the particles in step with the accelerating field, must be very accurate, and is probably in principle as difficult an adjustment as that of the field in the linear accelerator, to keep it in step with the particles. And there is the further difficulty, not so obvious at first glance, that the circumference of the orbit is very large, and in all present thinking there is only one accelerating unit around the circumference, which does not supply a very great increment of energy per revolution of the particle. Hence the particle receives much less energy, per unit length of its own path, than in the linear accelerator, where every effort is made to concentrate the accelerating units and make them as powerful as possible, so as to cut down the total length of the machine. In building up its total energy, then, the particle in the proton synchrotron travels very much further than in the linear accelerator, and as a result the chances of spreading, straggling, and loss of particles from the beam are much greater, so that it is likely that the number of particles lost from the beam will be even greater in the proton synchrotron than in the linear accelerator, even in spite of the inherently defocussing nature of the latter device, which we have mentioned above.

When we consider these difficulties in the way of attaining very high energies by other means, we are likely to feel, and the writer feels, that the linear accelerator cannot be dismissed without careful discussion, though it is by no means obvious that there will be any range of energies in which it is superior to other types of machines. All the work carried on so far has been preliminary and exploratory, in the range of a few million electron-volts. The writer has no doubt but that the linear accelerator can be made to work successfully in this range, but it is also not clear that it has any advantage over circular machines for low-energy particles. For this reason, the primary emphasis of the design considerations suggested in the present paper is on the possibility of extending the machine to very high energies. We shall constantly be looking for difficulties that will appear at high energies, though they will be absent at low energy. We may as well state at the outset that, while some of these difficulties appear serious, we can at least suggest possible means of overcoming them. It is obvious that, before starting the very expensive construction of a high-energy linear accelerator, all these difficulties should be examined as far as they can be in preliminary tests, so that there is a reasonable assurance of success. Many of these tests have already been made, but not all, and some of them will be taken up in the body of the paper.

The linear accelerator is not a new device. Sloan[3] and others in the decade before the war worked, with moderate success, on similar instruments. The production of high-energy, high-frequency power sources such as the magnetron[4] during the war, however, has given the possibility of feeding enough power into a linear accelerator, operating as a pulsed device, so as to promise really high accelerations, and to make it a promising field for further work. Two main lines of development have resulted since the war, closely related to each other, and both really stemming from radar work. First, Alvarez,[5] at the University of California, has been working on the design of a positive ion accelerator, working at a frequency of about 200 megacycles, using radio-frequency equipment designed for radars of that frequency. Secondly, a number of laboratories have been working on electron accelerators, using high-power magnetrons, generally at about 10 cm wavelength, developed for radar purposes. It is such a program that has been carried out at the Massachusetts Institute of Technology, under the auspices of the Research Laboratory of Electronics.[6] This program has been supported by the Joint Service Contract No. W-36-039 sc-32037 of the Signal Corps with the Research Laboratory of Electronics. It is a fairly direct outgrowth of thinking that was prevalent in the M.I.T. Radiation Laboratory during the latter days of the war, but the actual work and detailed thinking about design have all been carried on since the establishment of the Research Laboratory of Electronics. Other similar programs in this country are at Stanford University,[7] the General Electric Co., the University of Virginia, Purdue University,[8] Yale University,[9] and a number of others. Outside the country, there are two ambitious programs in England, at the Telecommunications Research Establishment (TRE),[10] and at the Cavendish Laboratory in Cambridge, a program at the Polytechnic Institute of Mexico, and various other projects. As far as the writer is aware, the general thinking in all these projects is along similar lines, with few differences of opinion of any moment. Reports have been available to the writer from Stanford[7] where Hansen and his collaborators have made important advances, and from TRE[10] where a theoretical group under Walkinshaw, and experimental groups, have made elaborate and complete surveys of the problem. In many respects similarities will appear between the treatment of the present paper and those of the groups just mentioned. The present work has been in almost all respects carried out independently at M.I.T., in the course of the last two years, and since the work of Stanford and TRE is only available in the form of private reports, it has seemed wise to make the present paper complete, even though there might appear to be duplication between it and those of the other groups.

The project at M.I.T. is an experimental as well as a theoretical one, devoted to exploration of the design factors to be encountered in a large accelerator, and to the building of an experimental section to accelerate electrons to energies of the order of twenty million electron-volts. It is still far from this goal, and its progress is described in the Quarterly Progress Reports of the Research Laboratory of Electronics.[6] Nevertheless, the general theory seems to be in good enough shape to warrant publishing the present paper at this time. On completion of the M.I.T. project, an experimental paper will be published by the group concerned, discussing the results of the project. It can be stated in the meantime that many of the points mentioned in the present paper have been checked experimentally, and many more will be before the completion of the project.

Naturally the writer is indebted to many colleagues and friends for discussions of the problem. First comes the group at M.I.T., consisting of Prof. A. F. Kip, Dr. Winston H. Bostick, a group of Research Associates consisting of Messrs. R. J. Debs, P. T. Demos, L. C. Maier, S. J. Mason, and J. R. Terrall, and several members of the mechanical and technical troup, including Mr. I. J. Polk and Mr. M. Labitt. Former members of the group were Dr. Jules Halpern, now of the University of Pennsylvania, who made important contributions during its earlier stages, Mr. E. Everhart, and Mr. R. A. Rapuano. We have profited by discussions with other members of the staff of M.I.T., including Prof. J. G. Trump, and are particularly indebted to Prof. J. A. Stratton and Prof. A. G. Hill, Director and Associate Director of the Research Laboratory of Electronics, for their constant friendly interest. The Technical Advisory Committee of the services associated with the Research Laboratory of Electronics, consisting of Messrs. Harold A. Zahl, E. R. Piori, and John E. Keto, of the Signal Corps, Office of Naval Research, and Army Air Forces respectively, has taken a lively and much more than formal interest in the project. Dr. R. Q. Twiss, of TRE, assigned to M.I.T. for liaison purposes, has taken a personal interest in the problem, and has contributed to discussions, as well as acquainting us with the progress of the TRE project. We have had useful discussions with Drs. W. C. Hahn and L. Tonks and their associates from the General Electric Company, with Prof. J. W. Beams and colleagues from the University of Virginia, with Prof. L. Alvarez and his colleagues of Berkeley, and Prof. W. W. Hansen and his colleagues from Stanford University, as well as with various other workers in the field. The writer is particularly indebted to Stanford University for the opportunity of spending some weeks there in the summer of 1947,

where most of this paper is being written, and of becoming acquainted with the linear accelerator programs at that University and at the University of California.

<u>1. Properties of Periodically Loaded Waveguides</u>. The phase velocity of a wave in a waveguide is greater than the velocity of light in free space, so that a particle cannot travel with the same velocity as such a wave, and some more complicated structure must be used to produce the linear accelerator field. The structures in use all consist of periodically loaded waveguides of various types. Accordingly, we shall take up in this section the properties of periodically loaded waveguides, emphasizing the general features which all such guides have in common, more than the peculiarities of individual structures. Much of the theory is of a type familiar to mathematical physicists from many problems of wave propagation in periodic structures, such as the theory of the weighted string, filter theory in electrical engineering, the electronic structure of metals and crystalline solids, and x-ray and electron diffraction. The interrelationships of these problems are described in "Wave Propagation in Periodic Structures," by L. Brillouin (McGraw-Hill Book Co., Inc., 1946). Similar methods were used by the writer during the war in discussing the resonant modes of the many-cavity magnetron (touched on briefly in Ref. 11). Since these general methods are so familiar, we shall give many results in the present section in rather general language, without proof or detailed discussion. The structures most used are shown schematically in Fig. 1, in which (a) shows the iris tube, used in most of the microwave

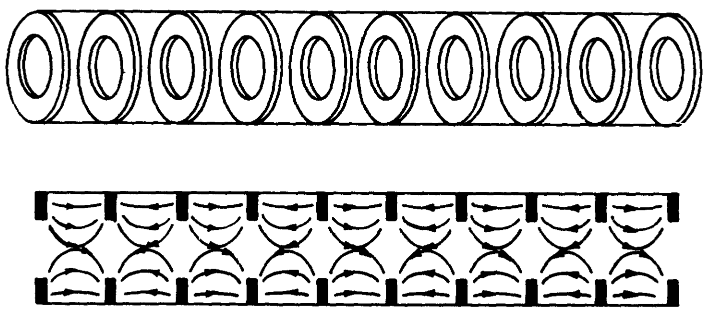

Figure 1 a.

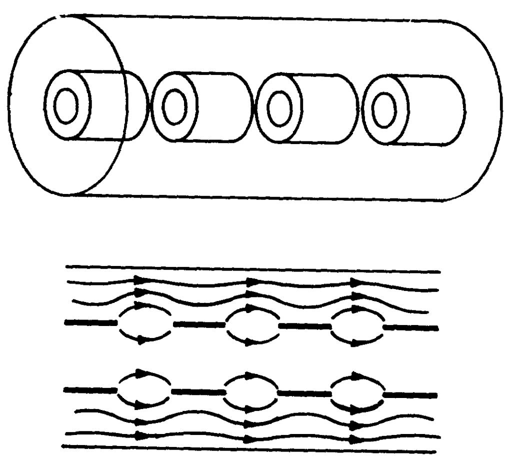

Figure 1b.

electron accelerators, including that at M.I.T., and in which (b) shows the structure being used by Alvarez for his positive ion accelerator. When we wish to discuss a specific example, it will be the iris tube, but most of our remarks will apply to Alvarez's structure as well.

The foundation of the study of wave propagation in periodic structures is a theorem called Floquet's theorem, which is very simple: in a given mode of oscillation of the structure, at a given frequency, the wave function (that is, in the electromagnetic case, the values of electric and magnetic field) is multiplied by a given complex constant when we move down the structure by one period. The proof of this theorem is not abstruse, and results from the fact that if the whole structure is displaced along its axis, which we take to be the z axis, by one period L, it coincides with the original structure, so that the new wave function can differ from the original one by only a constant factor. Let us write this factor in the form $e^{-\gamma L}$, where γ is a constant, in general complex. We now notice that a very simple function of z which has this property, namely that of multiplying the function by $e^{-\gamma L}$ when z is increased by L, is simply $e^{-\gamma z}$. In fact, the most general function of z with this property is the product of $e^{-\gamma z}$ with an arbitrary periodic function of z of period L, which will be unchanged when z increases by L. Such a periodic function can be expressed as a Fourier series, which can be written, in its complex form, as a sum of exponentials $\exp(-2\pi nj z/L)$, where n is an integer, positive or negative, each exponential with an appropriate coefficient. Thus the wave function can be written as a sum of exponentials $\exp(-\gamma - 2\pi nj/L)z)$, with appropriate coefficients.

The interpretation of this result depends on the nature of γ. In general, this constant can be complex, but it can be shown that in a structure without energy dissipation it must be real or pure imaginary. If it is real, each exponential decreases with increasing z, aside from its phase change, and we have an attenuated wave. Such waves are not appropriate for accelerator operation. In other cases γ is pure imaginary, and may be written $\gamma = j\beta_o$. Then if we define

$$\beta_n = \beta_o + \frac{2\pi n}{L} , \tag{1}$$

we can write our exponentials in the form $e^{-j\beta_n z}$. Combined with the factor $e^{j\omega t}$ expressing the complex time dependence of the sinusoidal field, a single Fourier component has the form $e^{j(\omega t - \beta_n z)}$. This represents a progressive wave, with angular frequency ω, wavelength $2\pi/\beta_n$, traveling along the z axis with velocity $v_n = \omega/\beta_n$. We now see that, using (1), the disturbance can be regarded as a superposition of a great many traveling waves, with a variety of velocities v_n, which get numerically smaller for large values of n. Each traveling wave will have an appropriate coefficient or amplitude.

Some one of these traveling-wave components may well be appropriate for use in a linear accelerator. The velocities of many components will be less than the velocity of light, so that the velocity of one component can be arranged to equal the velocity of the particle being accelerated. This component will then resonate with the particle, in the sense that the particle stays in a fixed phase relationship to the wave (until, of course, the particle speeds up enough to get ahead of the wave; we shall take up such questions later). If a particle resonates with one component of the Fourier expansion of the field, it will not resonate with any other, for the others all travel with different velocities. In fact, as seen from the moving particle, the resonating component acts like a field independent of time, which can then exert effects felt over many periods; but all other components move with respect to the particle with large velocities, so that as far as the particle is concerned they are rapidly oscillating fields. Their effect on the particle will be a rapidly alternating one, which will produce almost no net motion; and for almost all purposes they can be completely neglected. We have arrived, then, at a very important result: only one component of the field in the periodic structure will travel with the same speed as the particle; it acts like a sinusoidal traveling wave; and in considering the particle's motion, only this traveling wave need be considered. There is, of course, a corollary to this. Each of the components in the Fourier expansion has a finite amplitude, and therefore stores energy, and results in energy loss in case the walls of the system have a finite conductivity. This expenditure of energy is quite useless for purposes of accelerating particles, except for the particular Fourier component that resonates with the particle. Therefore we must look for a type of excitation in which the resonant component has as large an amplitude as possible, in comparison with the other, non-resonant components.

To determine the velocity of each component as a function of frequency, and hence to find which one is appropriate for use in the accelerator, we must find β_o, and hence β_n, as functions of ω. This is a considerable task, which we shall discuss later. Some general results are easy to prove, however. Let us consider ω as a function of β_o. Then we can show easily that the resulting curve is periodic in β_o, with period $2\pi/L$. For suppose β_o increases by $2\pi/L$. Then, from (1), the quantity which was previously β_{-1} will increase to become equal to the previous β_o, and similarly each β_n will change to equal the preceding β_{n+1}, changing the name of each β_n, but leaving the whole set of β's unchanged. This makes no change in the physical situation, for we must find the same coefficient for each of the β's that we did earlier for the β which had the same numerical value, and the frequency as determined from the field must be the same. We can also prove easily that the curve of ω as a function of β_o is an even function; that is, changing the sign of β_o leaves ω unchanged. This fact arises essentially because we can make two successive transformations without changing the physical situation: changing from a wave function to its conjugate, which changes each exponential from $\exp(j(\omega t - \beta_n z))$ to its conjugate $\exp(-j\omega t + j\beta_n z)$, then changing t to -t, resulting in $\exp(j(\omega t + \beta_n z))$, the net result being to change the sign of all β_n's, without changing the essential problem, or the frequency.

As a result of these general theorems, a plot of ω as a function of β_o will have the form shown in Fig. 2. Here it is convenient to plot $\omega/2\pi c = 1/\lambda_o$ instead of ω, and $\beta_o/2\pi = 1/\lambda_g$ instead of β_o, where c is the velocity of light in free space, λ_o is the wavelength of the wave in free space, and λ_g is the guide wavelength associated with the component of n = 0. We see that this guide wavelength cannot be uniquely defined, but that we can add any integral multiple of 1/L to $1/\lambda_g$, and have an equally valid guide wavelength. Otherwise stated, $1/\lambda_o$ is a periodic function of $1/\lambda_g$, with a period 1/L. Now we

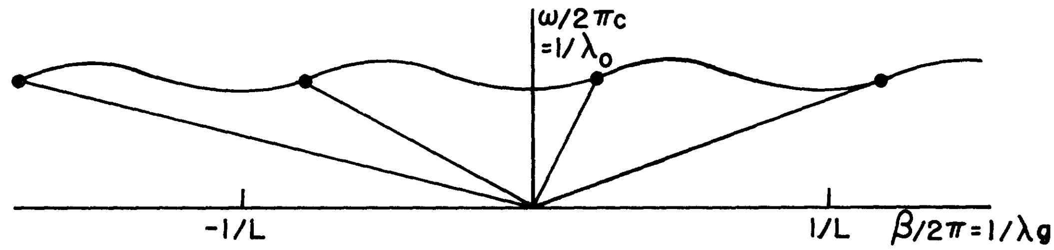

Figure 2. Frequency as function of reciprocal guide wavelength, for periodically loaded line. Slopes of radius vectors represent phase velocities of various Fourier components, divided by velocity of light.

see that the phase velocity of the component n = 0, divided by the velocity of light c, is $v_0/c = \omega/\beta_0 c = (1/\lambda_0)/(1/\lambda_g)$, or is the slope of the radius vector out to a point of the curve. If this slope is greater than that of the 45° line, the velocity is greater than that of light, and if it is less, the velocity is less than that of light. We can now see graphically the velocities of all the components of different n values associated with a given wave function, by setting up the values of $1/\lambda_g$ spaced by intervals $1/L$, and drawing the radius vector to each, as is done in Fig. 2. It is clear from this diagram how the components of higher n value have smaller phase velocities. Also it is clear that there are components with both positive and negative phase velocity, or components traveling both to the right and the left.

We shall now discuss in qualitative language the way in which curves like those of Fig. 2 arise, by using a specific example. We take the example of the iris tube, as shown in Fig. 1a, and ask how the curves change as the holes in the irises change their size, at constant iris spacing, from holes as large as the iris, (so that the iris does not exist) down through smaller and smaller holes, to the limiting case where the hole vanishes, and the tube consists of a set of cylindrical cavities, separated by conducting walls. This problem will be considered in detail by the writer in a forthcoming paper. We consider only the transverse magnetic mode in which the magnetic field runs in circles about the axis of the circular guide, and the electric field is in a plane containing the axis, and independent of the angle of rotation of that plane about the axis. This is the mode ordinarily met in a cylindrical cavity, in which the electric field runs along the z axis from one wall to the other, being proportional to $J_0(kr)$, where $k = 2\pi/\lambda_0$, and where the magnetic field is proportional to $J_1(kr)$. In the cavity, the resonant frequency is fixed by the condition that E_z must be zero at the outer wall of the cavity, or at r = R, where R is the radius of the cylinder; thus we must have $J_0(2\pi R/\lambda_0) = 0$. But the first zero of J_0, which is concerned in the mode we are interested in, comes for a value 2.405 of the argument of the Bessel function. Hence the resonance comes for $2\pi R/\lambda_0 = 2.405$. We now consider a succession of cases with fixed iris spacing, but varying iris openings, related to this mode of the circular cavity.

We start with the case where there are no irises at all, or where the radius a of the iris opening equals the radius R of the cylinder. Then we have an ordinary unloaded waveguide. It is well known that the free space wavelength λ_0 and the guide wavelength λ_g are related by the equation

$$\frac{1}{\lambda_o^2} = \frac{1}{\lambda_g^2} + \frac{1}{\lambda_c^2} , \qquad (2)$$

where λ_c is the cutoff wavelength. This latter quantity is determined as the free-space wavelength for which the guide wavelength becomes infinite, or in which the field is independent of z. In this case, the problem becomes just like that in the cylindrical cavity which we have just discussed, so that the cutoff wavelength is the free-space wavelength found in that case, or is given by

$$2\pi R/\lambda_c = 2.405. \qquad (3)$$

We now see that the relation between $1/\lambda_o$ and $1/\lambda_g$ given by (2) is the equation of a hyperbola, shown in Fig. 3. At values of $1/\lambda_o$ less than $1/\lambda_c$ (that is, for free-space wavelengths greater than the cutoff wavelength) the curve indicates no real value of λ_g; that is, no real propagation is possible, but there is a pure imaginary wavelength, resulting in attenuation. For all higher frequencies, the hyperbola lies above the 45° line,

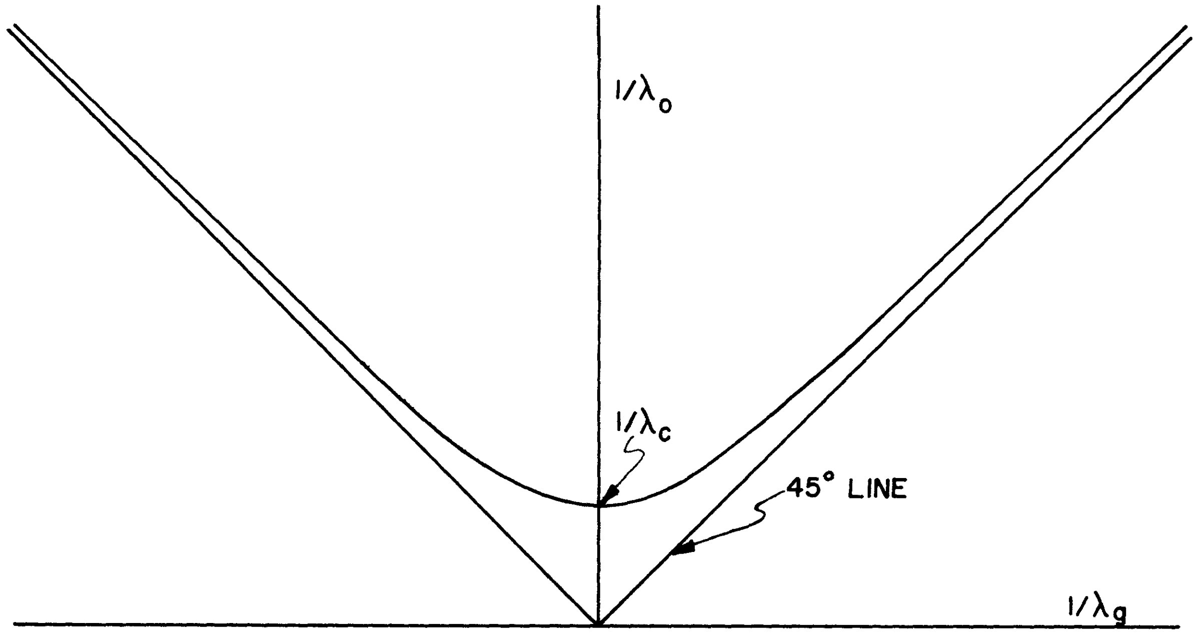

Figure 3. Hyperbola representing $1/\lambda_o$ vs. $1/\lambda_g$ for an unloaded waveguide.

indicating a phase velocity greater than the velocity of light, reducing asymptotically to the velocity of light as the frequency becomes very large. In this case, of course, there is no periodicity in the guide, and consequently no periodicity in the plot of Fig. 3.

A very small iris, however (that is, an iris whose hole radius a is nearly as large as the tube radius R), produces a change in the situation resulting in periodicity, and we see in Fig. 4 how this can come about. Over most of the range of β_o, the change of frequency is negligible, but just in the neighborhood of the value $\beta_o/2\pi = 1/2L$, the

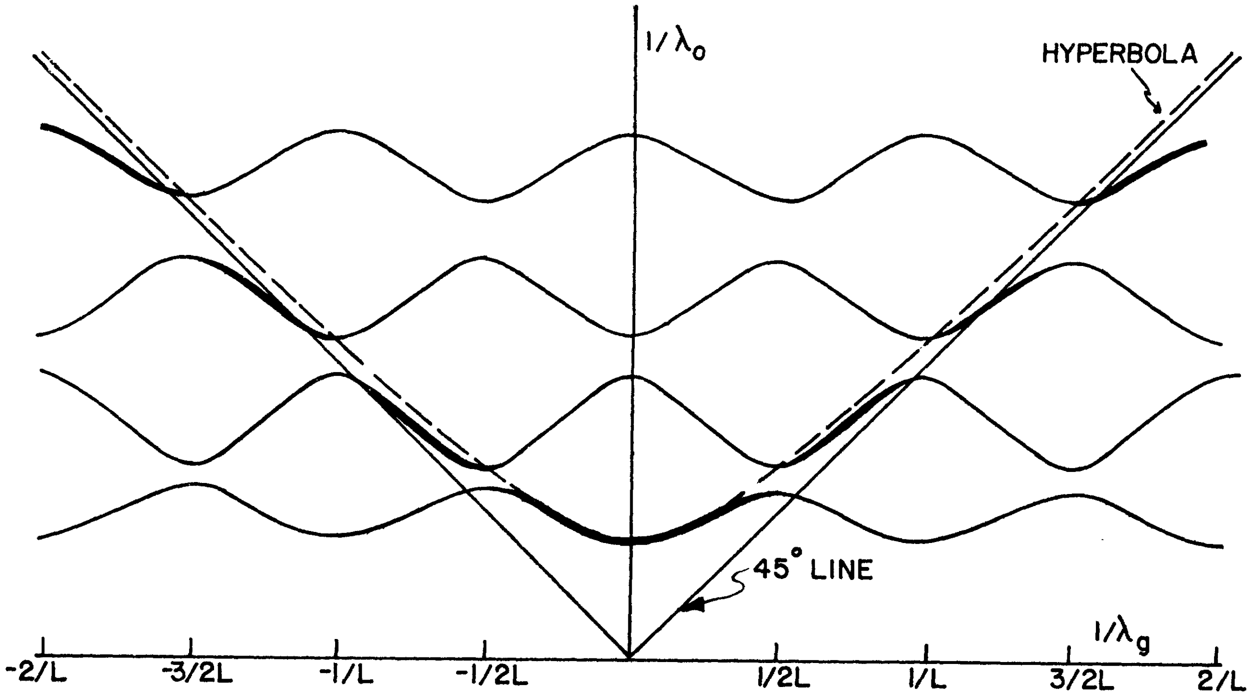

Figure 4. $1/\lambda_0$ vs. $1/\lambda_g$ for slight periodic loading of wavguide. Heavily shaded part of curve corresponds to Fig. 3.

change is considerable. This particular value has a simple significance: it corresponds to the situation $L = \lambda_g/2$, or the case where the iris spacing equals a half guide wavelength. Consider now the propagation of a wave of this wavelength down the guide. We may consider that each iris acts as a scatterer, so that a wave approaching an iris results in a reflected wave traveling backward, as well as a transmitted wave. With irises spaced a half guide wavelength apart, the reflected wave from each successive iris will be a whole period out of phase from the reflected wave from the preceding iris, since the direct wave will have traveled a half wavelength farther, and so will the reflected wave. Thus all these reflected waves can interfere with each other, and can produce a large effect. As a matter of fact, they interfere so strongly that no direct wave at all can be propagated through a lossless guide of infinite length of this sort; the reflected wave proves on analysis to equal the incident wave in amplitude, and a pure standing wave is set up. This is the phenomenon which, in the study of x-ray diffraction, becomes Bragg reflection.

The situation becomes clearer when we consider the amplitudes of the various Fourier components of propagation constant β_n, in this problem. Let us assume that $\beta_0/2\pi$ equals the value of $\beta_n/2\pi$ nearest to the reciprocal of the guide wavelength that a guide without irises would have for the corresponding frequency. Then for most frequencies or β's, only one Fourier component in the expansion of the field will be appreciable, that associated with β_0. We readily find that, for positive β_0's, the wave reflected from the irises is characterized by $n = -1$, and waves that have made multiple reflections are given by other values of n. Unless the interference conditions we have mentioned in the preceding paragraph are fulfilled, the reflected waves will be weak, for there will not be constructive interference between waves reflected from the various irises, and the Fourier component associated with n = -1 will be small, and all the others even smaller. As we approach our critical value of $\beta_0/2\pi = 1/2L$, however, the amplitude of the component n = -1 becomes large, equalling the amplitude of the component n = 0 at the critical value. As we pass beyond this value, the amplitudes of other components than n = 0 again become small. At the value $\beta_0/2\pi = 3/2L$, where the iris spacing is three half wavelengths, we again get interference, and the building up of a reflected wave, this time corresponding to n = -3; and so on.

We now consider the periodic nature of the cruves of Fig. 4. The part of the curve drawn with a heavy line refers to the frequency as a function of β_0 which we have described in the preceding paragraph, departing from the hyperbola characteristic of the waveguide without irises only in the neighborhood of the critical values given by $\beta_0/2\pi = m/2L$, where m is an odd integer. At each of these critical values, the curve proves to have a break, there being two possible frequencies associated with the same guide wavelength. These two frequencies can be shown to be associated with two modes, both standing rather than traveling waves (on account of the fact that a progressive wave under these conditions is completely reflected), and one of sine-like, the other of cosine-like, character, so that for one of the waves the irises come at nodes, for the other at antinodes; it is natural that the irises have different effects on the frequency in these two positions. For frequencies lying within these breaks, there is no possible propagated mode, but the value of β_0 becomes imaginary, or there is an attenuated wave. The range of frequencies is broken up, in other words, into certain pass bands of frequency, in which we can get propagation, with bands of attenuation between. The periodically loaded line acts like a bandpass filter, with an infinite number of pass bands; and in case the loading reduces to zero, and we get the case of Fig. 3, the bands coalesce, so that all frequencies greater than the cutoff frequency of the unloaded guide can be propagated.

On account of the periodicity of the curve of frequency vs β_0, of which we have already spoken, we see that the curves of Fig. 4 should be drawn in a periodic way, as we have done. The branch of the curve which is drawn heavily, as we have already seen, is the one which reduces to the hyperbola of Fig. 3 in the limiting case where the irises disappear, but the other branches are equally legitimate. They correspond to different conventions for the numbering of the β's. So far, we have named that component which had the largest Fourier component β_0. These other branches correspond to numberings in which the component for n = 0 has small amplitude, but some other component has a large amplitude. In this problem, one numbering, the one we used originally, is far more natural than any other; but when the iris becomes more important, the distinction is no longer clear, many Fourier components being large, and some of these other methods of numbering are equally reasonable.

When we look at the curves of Fig. 4, we see that now, in contrast to Fig. 3, we can have a phase velocity less than the velocity of light. The curves are of the form already discussed in Fig. 2, and we know that for any frequency in one of the pass bands, we can find an infinite number of velocities, one associated with each intersection of the curve with the straight line ω = constant. Each of these velocities is that of a particular

Fourier component, however, and the only Fourier component which is large is that represented by the heavy line in Fig. 4, which still, over most of its length, lies above the 45° line, and corresponds to a velocity greater than that of light. In other words, though the irises have introduced Fourier components of low velocity into the Fourier resolution of the field, these components have small amplitude, for a small iris, and would not be effective components to use in a linear accelerator.

The situation is different, however, when the iris projects further into the tube; that is, when the ratio a/R is smaller, where a is the radius of the hole in the iris. In this case, the curve is more like that of Fig. 2. The curve can become so depressed that the velocity of the Fourier component of largest amplitude can become less than that of light. Furthermore, the amplitudes of other Fourier components increase so much, on account of the large scattering by the irises, that other Fourier components of lower velocity have considerable amplitudes. These two effects are of course related; and while we now have components of considerable amplitude with velocities suitable for a linear accelerator, there is a compensating disadvantage, in that other components have considerable amplitude as well and they absorb power, which is useless as far as the accelerator is concerned. For instance, suppose we have loaded the guide so much that the so-called π-mode (that is, the mode in which the phase difference $\beta_o L$ from one section to the next is π; this corresponds to $\beta_o/2\pi = 1/2L$, or the critical value for reflection) has a velocity less than the velocity of light, and we use that mode. Then we have already seen that we have a reflected wave of amplitude equal to the direct amplitude, which is of no use, so that the power used to set up this reflected wave is thrown away. Or suppose we use the $\pi/2$-mode, for which $\beta_o L = \pi/2$, $\beta_o/2\pi = 1/4L$. In this case, as we see from the figure, we shall have to load even more heavily to bring the velocity of this mode down below that of light, the amplitudes of other Fourier components traveling in the same direction as the wave will become considerable, and we shall throw power away in them. In any case we pay a penalty for slowing the wave down, the penalty being greater the greater the slowing down.

The limiting case is that in which the hole in the iris vanishes altogether. Then the guide becomes a set of disconnected cylindrical cavities, each oscillating like a single cavity. To set up modes corresponding to different values of β_o, we merely arrange to have the phases of the oscillation differ from cavity to cavity by $\beta_o L$. The frequency will be independent of β_o, so that in this limit the curves of ω vs β_o degenerate to horizontal lines, and the pass bands have shrunk to zero. The allowed frequencies, or frequencies of these limiting pass bands, are easily found: they are just the resonant frequencies of the cylindrical cavity. The lowest one is given by Eq. (3), and corresponds to the lowest pass band; the higher ones correspond to the cases in which the disturbance varies sinusoidally along the length of the cavity, the tangential component of E having nodes at both plane walls of the cavity, so that the length L equals an integral number of half wavelengths. In this limiting case, we can find easily the Fourier resolution of the field, and the magnitude of the various Fourier components, and these limiting values hold approximately for irises with very small holes, $a \ll R$. Thus for instance in the lowest pass band, the value of the z component of field is approximately constant in each cavity, and equal to $e^{-j\beta_0 pL}$ in the pth cavity. All we have to do to get the Fourier components is to expand this simple step function in the way we have already described.

2. Fourier Resolution of the Field, Modulation Coefficient, and Transit Time Correction. We have stated in the preceding section that, to handle the operation of the loaded guide as a linear accelerator, we need only consider the Fourier component of the field which resonates with the particle, or has the same phase velocity. If we consider only this component, the particle remains in a fixed phase relation to it, is acted on by a constant force, and the dynamical problem is very simple. On the other hand, various writers on the subject proceed by a different method, familiar from conventional analyses of triodes and klystrons: they consider the effect of the finite transit time of particles between the two sides of the cavity. Suppose for instance that the iris hole is small; then the two irises are similar to two grids of a triode or klystron, and the field between them is approximately a constant, independent of z, but of course varying sinusoidally with time. As a particle moves from one to the other, the field cannot remain at its peak value for the whole transit; for it is varying with time. If we calculate the average force acting on the particle, we find that it is the peak value, multiplied by a certain coefficient, which in the language of klystron theory is called a modulation coefficient, equal to unity if the grids are so close together that the transit time is zero, but decreasing as the transit time increases, being a function of the form $(\sin x)/x$ of the transit time. This function can go to zero for certain transit times, or go negative, and such effects are encountered in the theory of high-frequency operation of triodes, where transit time becomes important.

We shall now show that this modulation coefficient or transit time effect is not a new effect which we must handle separately, but that in fact our procedure of using a single Fourier component is itself an elegant way of making the transit time correction. Suppose our particle is traveling along the z axis with a velocity v, so that its position is given by $z = vt$. Let the value of E_z, the longitudinal component of electric field, as a function of z and t along the z axis, be

$$E_z = \sum_n F_n e^{j\omega(t - z/v_n)},$$

where F_n is the amplitude, v_n the velocity, of the nth Fourier component in the analysis of the field. As the particle moves, the field will change with time, and we can take account of this by substituting for the time t when the particle is found at position z the value $t = z/v$. Thus the field acting on the particle at the point z is

$$\sum_n F_n e^{j\omega[(1/v) - (1/v_n)]z} .$$

The average of the nth term, averaged over z (assuming the particle is traveling for a long distance) is zero if v is different from v_n, since the resulting sinusoidal function will average to zero. Thus the average field will be zero unless the velocity of the particle equals the velocity of one of the Fourier components, and in that case it will be just the F_n, or amplitude of that Fourier component. It is to be noticed that this result is only correct if the particle is traveling for a long distance at constant velocity in the field; this condition is approximately, though not exactly, met in the linear accelerator, and deviations from it do not enter in a way to affect our subsequent arguments.

We see, then, that our procedure of Fourier analysis is all that we need to do to take account of the transit time of the particles from cavity to cavity or iris to iris. This result, we notice, is quite general, and not dependent on the fact that we are using a waveguide loaded by irises, but applicable to any periodic structure. Any modification of the structure which results in an increase of the amplitude of the Fourier component which resonates with the particle will have an effect which can alternatively be described as improving the modulation coefficient or transit time situation. Our method, however, is much more general than the conventional argument in terms of transit time, in that that argument is often set up only for the case of a constant field between parallel grids or electrodes, whereas our method is correct for an arbitrary variation of longitudinal field with z. It is possible, as a matter of fact, to adapt our method, not for periodic structures, but for single grid systems, as in the triode or klystron, by the use of Fourier integrals instead of Fourier series. One can then prove familiar theorems about modulation coefficients and transit time, and as we shall see later about the transverse motions of particles and focussing. We shall not make further use of this fact, however.

<u>3. Group Velocity in the Loaded Guide</u>. In Section 1, we have shown that the ratio of the phase velocity of the wave in a loaded guide, to the velocity of light, is given by the slope of the radius vector out to the point representing the wave, in the graph of $(1/\lambda_o)$ vs $(1/\lambda_g)$. Now we shall consider the group velocity, which we shall call v_g, and show that it is given, not by the slope of the radius vector, but by the slope of the tangent:

$$\frac{v_g}{c} = \frac{d(1/\lambda_o)}{d(1/\lambda_g)} \quad . \tag{4}$$

We shall find the group velocity to be important in the study of power flow in the guide, and in the question of how much power dissipation we must have to set up a field of a given strength.

The conventional derivation of the formula for group velocity sets up a superposition of two waves, one with angular frequency ω and propagation constant β, the other with angular frequency $\omega + \Delta\omega$ and propagation constant $\beta + \Delta\beta$, and considers the velocity of the beats between these two waves. It is easy to show (see for instance Slater and Frank, "Mechanics", McGraw-Hill, New York, 1947, p. 168) that this velocity is given by $\Delta\omega/\Delta\beta$. This is closely related to the slope of the plot of the type given in Fig. 2. The two waves are represented by two points on the curve of that figure. Remembering that the abscissa in that figure is $1/\lambda_g = \beta/2\pi$, the ordinate $1/\lambda_0 = \omega/2\pi c$, we see that

$$\frac{\text{beat velocity}}{c} = \frac{\Delta(1/\lambda_0)}{\Delta(1/\lambda_g)} . \tag{5}$$

If the frequencies and propagation constants of the two waves are close together, we can replace the quantity (5), which is the slope of the chord, by the slope of the tangent to the curve, getting a common velocity of beat propagation for all waves of neighboring frequencies. This is the group velocity, as given by (4).

We may now set up a group of waves, such for instance as a wave train of finite length, by superposing sinusoidal waves of a variety of frequencies or wavelengths, the necessary band of wavelengths or frequencies being narrower, the longer the packet is in space or time. If the curve of $1/\lambda_0$ vs $1/\lambda_g$ can be considered to be straight over this range of wavelength or frequency, the beat velocity as defined in (5) will be the same for all pairs of waves in the wave train, and the whole pattern will move forward with a single velocity, the group velocity. If on the other hand the wave train is too short, so that its Fourier analysis extends over a wide range of frequencies or wavelengths, the beat velocities of different pairs of components will be different, the pattern will not be preserved, and the disturbance will spread in a complicated way.

The group velocity is the velocity with which energy is propagated in the guide. This is most easily seen by considering a finite wave train, of definite length, traveling say to the right. The energy crossing a given cross section of the guide, per unit time, will then clearly be the energy contained in a length v_g of the guide, since just this energy will pass the cross section in a second. Thus we have a relation between the Poynting's vector as integrated over a cross section of the guide (the energy flux) and the energy density per unit length in the guide: the energy flux is v_g times the energy density per unit length.

As we look at the curves of $1/\lambda_0$ vs $1/\lambda_g$, as in Fig. 2 and Fig. 4, we see that the slope of the curves is in every case less than unity, so that the group velocity is less than the velocity of light, as we expect from relativistic arguments. Furthermore, we see that as we approach the π-mode, the slope becomes horizontal, and the group velocity goes to zero. This is consistent with statements which we have made already about the π-mode. We saw earlier that as we approach that mode, the reflected wave becomes stronger, until the π-mode itself is a standing wave, with equal amplitudes for direct and reflected waves. In this case, there is no net flow of energy, so that v_g is zero. Near the π-mode, the reflected wave almost equals the direct wave, there is small energy flux, and consequently the group velocity must be very small. Furthermore, we see that if the holes in the irises are small, the whole band of frequencies will shrink, so that even in the $\pi/2$ mode, where the group velocity has its maximum value, the group velocity will still be very small. In the limit, as the holes shrink to zero, and the frequency becomes independent of the guide wavelength, the group velocity will approach zero as well. This is obviously consistent with the fact that, with vanishing holes in the irises, there can be no flow of power.

4. Attenuation in the Guide. So far we have been neglecting attenuation in the guide; but it plays an essential part in the operation of the linear accelerator, and we must next consider its effect. If we have a given energy density within a short section of loaded guide, therewill be a flow of energy into the walls of the guide, on account of the resistive losses in the walls. We can relate this loss of energy to the unloaded Q of the guide, which we shall Q_o. This is defined by the relation

$$\frac{1}{Q_o} = \frac{\text{energy dissipation per second in walls}}{\omega \times \text{stored energy}} . \qquad (6)$$

It is the Q which the guide would have if it were converted into a resonant cavity by perfectly reflecting end plates, so that no energy would be dissipated in the ends. Since the energy dissipation per second in the walls, and the stored energy, are both proportional to the length of the section of guide we are considering, we see that Q_o is independent of the length.

If we now have a traveling wave flowing down the guide, its power will be attenuated as it travels down the guide; for some of its energy will be drained off into the losses in the walls. We can easily find the rate of attenuation resulting from this. Let us set up an equation of continuity for the energy flow. If W is the energy per unit length in the guide, S the energy flux across a given cross section, D the power dissipation per unit length in the wall, then the equation of continuity states that the time rate of increase of W equals the negative of the divergence of S, minus the power dissipation D. That is, since S will depend only on z, the distance along the guide, we have

$$\frac{\partial W}{\partial t} + \frac{\partial S}{\partial z} + D = 0 . \qquad (7)$$

We have just seen from (6), however, that $D = \omega W/Q_o$; and from the preceding section we have seen that $S = v_g W$. Thus (7) becomes

$$\frac{\partial W}{\partial t} + v_g \frac{\partial W}{\partial z} + \frac{\omega}{Q_o} W = 0, \quad \frac{\partial S}{\partial t} + v_g \frac{\partial S}{\partial z} + \frac{\omega}{Q_o} S = 0 . \qquad (8)$$

Assuming a steady state, so that the time derivatives are zero, we find

$$S = S_o e^{-z/\ell_o}, \quad \ell_o = \frac{v_g Q_o}{\omega} .$$

Thus the flux is attenuated, with an attenuation constant $1/\ell_o$, where ℓ_o, the distance in which the flux falls to 1/e of its value, may be called the attenuation length. The attenuation represented by this expression must be applied to the fields as found in Section 1; the electric and magnetic fields separately will have attenuation coefficients half as great as the flux.

The attenuation length may be interpreted easily in terms of the time required for an oscillation in a resonant cavity to die down to 1/e of its initial value. In a resonant section of waveguide, there will be no variation of energy density with z; thus the derivatives with respect to z in (8) must be set equal to zero. Then we shall find

$$W = W_o e^{-(\omega/Q_o)t} ,$$

showing that the energy density falls to 1/e of its initial value in a time Q_o/ω. We now see that the attenuation length is just the distance which the energy will travel, moving with the group velocity, in the time in which the energy density falls to 1/e of its initial value, a time which we shall call T_o. The attenuation length, then, becomes greater as Q_o becomes greater, but it becomes smaller as the group velocity decreases.

We shall find that the attenuation length is a very important concept in the discussion of linear accelerators. We shall call a linear accelerator long if its length is large compared to the attenuation length, short if its length is short compared to the attenuation length. It is clear that long accelerators will involve us in difficulties. Any signal introduced at one end of the accelerator will be attenuated to negligible amounts by the time it reaches the other end, and if we start building up a signal at one end at a given time, this signal will not even have reached the other end by the time T_o, at which the field at the first end will have built up to its steady state value (for the build-up of energy naturally has the same time constant as the decay). This situation means that the two ends of a long accelerator are effectively decoupled from each other, both in space and in time. Yet we shall find that to set up a proper field in a long accelerator we must have definite phase relations between the fields at its two ends. This imposes problems of excitation, which we must consider; for it appears that if we wish accelerators long enough to produce particles with energies in the billions of electron-volts, the accelerators must be long in the sense in which we are using the word.

5. Power Input to the Guide. To operate a linear accelerator, we must build up a very strong electric field along the z axis, associated with the Fourier component traveling with the same velocity as the particle to be accelerated. If we know the whole pattern of the field within the accelerator, we shall find the integral of the square of the electric field per unit length of the accelerator, and hence the stored energy per unit length, to be proportional to the square of this field component. In fact, if the electric field along the z axis has the component

$$E\, e^{j\omega(t - z/v)}$$

associated with the resonant wave, we must have the stored energy per unit length equal to a constant times $\epsilon_o E^2$ times the cross section of the guide. For a given geometrical shape (that is, given iris spacings, dimensions, etc.) but arbitrary scale, the cross section will be proportional to the square of the free space wavelength at which the accelerator is operated, since all linear dimensions of the cavities will be proportional to the wavelength. Thus we have

$$W = A\,\epsilon_0 E^2 \lambda_0^{\,2} \quad ,$$

where A is a constant which can be determined when the field distribution is known. It is assumed above that W is the total stored energy, which of course is twice the electrostatic energy.

To produce a large value of E, then, we must have a large stored energy, and hence, in a steady state, a large dissipation of energy, and a large amount of power fed into the accelerator. During the build-up process, W will be less than in the steady state. Hence the accelerating field E will be less during buildup than in the steady state, and the proper way to operate the accelerator is to let the field build up first, then introduce the particles to be accelerated, when the field has reached substantially its final value. We shall consider in this section the amount of power which must be fed in in the steady state, to produce a given field E, and shall also consider the transient process of building up the steady state.

The results are somewhat different, depending on whether the accelerator is closed with perfectly reflecting walls at the ends, so that a standing wave is set up in it, or whether the ends are terminated with non-reflecting or matching resistances, so that traveling waves will not be reflected from the ends. They are also somewhat different depending on the distribution of power inputs along the length of the accelerator. Let us start with the case of reflecting end walls, and uniform distribution of power sources along the tube, as in the M.I.T. and Berkeley accelerators. In this case, all the power fed into unit length will be dissipated in power loss in unit length of walls. We have already seen that this dissipation is $\omega W/Q_0$. Hence, if the power fed in per unit length is P, we have

$$P = \frac{A\,\epsilon_0 E^2 \lambda_0^{\,2}\omega}{Q_0} \quad .$$

Remembering that $\omega/c = 2\pi/\lambda_0$, $c = 1/\sqrt{\epsilon_0\mu_0}$, and $\sqrt{\mu_0/\epsilon_0} = 377$ ohms, we find

$$E = \alpha\sqrt{\frac{377\ PQ_0}{\lambda_0}}\ , \quad \text{where } \alpha = \frac{1}{\sqrt{2\pi A}} \quad . \tag{9}$$

We note that the accelerating voltage per unit length is proportional to the square root of power per unit length; thus if we wish to have an accelerator producing a given voltage as short as possible, or to make E as large as possible, we wish to feed in as much power as possible per unit length. On the other hand, if the total power at our disposal is fixed, or if we wish to economize on power sources, we should go to the other extreme, and make the accelerator as long as possible. To see this fact, we rewrite (9) in a form to give the total voltage acquired by a particle, $E\ell$, where ℓ is the total length of the accelerator, in terms of $P\ell$, the total power input. We have

$$E\ell = \alpha\sqrt{\frac{377\ (P\ell)Q_0\ell}{\lambda_0}}\ ,$$

showing that the total voltage is proportional to the square root of the length, so that by making the length indefinitely large, we can in theory get any amount of acceleration from a fixed power input.

From the equation just written, we see that a given amount of total voltage can be produced either by using a large amount of total power and a short length, or a small power and large length, and that the total voltage is a function of the product of total power and length. There seems to be no scientific way of deciding what separate values of power and length to use, and therefore the decision will presumably be based on economic arguments. The cost of a linear accelerator, aside from the fixed cost of end terminations, will consist of two parts: the cost of the accelerator tube, pumping, housing, etc., which will be proportional to the length, and the cost of the power sources with their feeds into the accelerator tube, which will be proportional to the power input, and hence, for a definite total voltage, inversely proportional to the length. Hence the cost will be the sum of two parts, one proportional to the length, the other (for a fixed voltage) inversely proportional to the length. We choose the length to minimize the cost. Now the function $x + 1/x$ has its minimum for $x = 1$, where both terms are equal in magnitude. Thus the accelerator will cost least if the tube (that is, the parts whose cost is proportional to length) and the power sources cost equal amounts. If a decision has been made as to the type of power source to use, and if the sources are to be equally spaced along the line, the economic way to determine the spacing of sources is to build a length of line whose cost equals that of the power source. We shall see later that the spacing between oscillators must not be large compared to the attenuation length; therefore we must be sure that our design is such that the economic spacing of oscillators is less than the attenuation length. Presumably by the time the various preliminary accelerator projects now under way in the various institutions are completed, accurate enough cost estimates can be made to find these economic spacings between oscillators, but reliable figures are not yet available, so that it is not clear whether a close spacing of oscillators, as at M.I.T., or a wide spacing as at Stanford, or something intermediate, will eventually turn out to be most economical.

The situation is quite different if we have a traveling wave. Again we assume a fixed power input P per unit length. Now, however, power is lost in two ways: there is still a dissipation $\omega W/Q_0$ per unit length on account of resistive losses in the walls, but there is also a flux $v_g W$ out through the end of the tube, into the non-reflecting termination. Thus the total power $P\ell$ supplied to the tube must equal $(\omega\ell/Q_0 + v_g)\ W$. This amounts to replacing $1/Q_0$ by

$$\frac{1}{Q_L} = \frac{1}{Q_o} + \frac{v_g}{\omega \ell} = \frac{1}{Q_o}\left(1 + \frac{\ell_o}{\ell}\right) , \qquad (10)$$

where we may call Q_L the loaded Q, loaded by the resistive termination at the end of the guide. In place of (9), we then have

$$E = \alpha \sqrt{\frac{377\ PQ_o}{\lambda_o} \frac{\ell}{\ell + \ell_o}} . \qquad (11)$$

If the length ℓ of the tube is large compared to the attenuation length, the new factor $\sqrt{\ell/(\ell + \ell_o)}$ is approximately unity, and the voltage E is not much decreased by the loss in the end of the guide; but if the length is small compared to the attenuation length, the factor is much less than unity, and the voltage per unit length obtainable from a traveling-wave tube is seriously less than that of a standing-wave tube with the same geometry (same α) and same power per unit length. For a short accelerator tube, in other words, it does not seem reasonable to use a traveling-wave tube. For a long accelerator, the loss in the terminating ends becomes negligible compared to the losses in the side walls, and the difference in accelerating voltage is unimportant. In this case, the decision between standing and traveling waves must be made on other considerations, which we shall take up later.

Of the existing accelerator projects, at least two, those at Stanford and at TRE, are planning traveling-wave tubes, and at least some of their thinking has been in the direction of feeding the tubes with power from the input end. In this case we may modify (11) to express the total acceleration in terms of total power, finding

$$E\ell = \alpha \sqrt{\frac{377(P\ell)}{\lambda_o}\ Q_o \ell\ \frac{\ell}{\ell + \ell_o}} . \qquad (12)$$

In this formula, for ℓ small compared to ℓ_o, and for constant total power input $P\ell$, we observe that $E\ell$, the total acceleration, is proportional to the length of the accelerating tube, which suggests that we need only feed a traveling-wave tube at one end to get any acceleration we desire. Nevertheless Eq. (12) is equivalent to (11), and shows very definitely that though it is true that the acceleration obtained is proportional to the length of the accelerator for small ℓ/ℓ_o, still it is less, in the ratio $\sqrt{\ell/(\ell + \ell_o)}$, than the acceleration that would be obtained by putting the same power into a standing-wave tube of the same length. It is only because this factor becomes less disadvantageous for longer lengths that we get the rapid increase in acceleration with length in this case. Furthermore, there is another disadvantage connected with feeding at one end: as we have seen earlier, the energy density in the tube falls off exponentially with z, becoming small at a value of z equal to the attenuation length. Thus the acceleration of the particles will fall off after this distance. In other words, this method of end feeding into a traveling-wave tube is possible only for a tube which is not long compared to the attenuation length ℓ_0; and if its length is much smaller than ℓ_0, it is definitely inferior to using the same power in a standing-wave tube. We shall note presently that the traveling-wave tube has an advantage, in that it can have a larger α than a standing-wave tube, and as a result of this there is a small range of lengths, in the neighborhood of ℓ_0, for which the traveling-wave tube fed from the end can be slightly superior to the standing-wave tube in the matter of acceleration; this point has been noted by the workers both at Stanford and at TRE, and they propose to adjust v_g so that their actual length will be approximately equal to ℓ_0. For a long tube, however, power must be fed in along the length, and in this case, as we have already mentioned, the distinction between standing and traveling-wave tubes must be made on other grounds than power consumption.

It is interesting to consider the transient phenomena involved in the build-up of oscillations in the tube. In a standing-wave tube, the time required for the field to build up is of the order of Q_0/ω, as we have mentioned earlier. It is interesting to compare this time, which we have called T_0, with the time $T_1 = \ell/v_g$ required for a field, moving with the group velocity, to travel from one end of the guide to the other. We find that $T_1/T_0 = \ell/\ell_0$. In other words, as we have mentioned earlier, in a long accelerator tube, the time required for the field to go from one end to the other is large compared to the build-up time. During the time of build-up, one end of the tube does not know what the other end is doing; and to secure proper phase relations between the two ends of the tube, we must arrange some external phasing circuit, as discussed in a later section. On the other hand, in a short standing-wave tube, the field has time to be reflected back and forth many times from one end to the other in the time of build-up. Phase relations between the two ends can be easily established. Furthermore, in a short tube, it is immaterial where in the tube the power is introduced; on account of reflections, the energy will be properly distributed through the tube irrespective of the location of the power input. This is in contrast to the long tube, where evidently the power sources must be uniformly distributed along the tube, the spacing being no longer than the attenuation length, and preferably considerably shorter. It is interesting to note that, in the standing-wave tube, the whole phenomenon of build-up of oscillations will occur in just the same way, if reflecting walls (penetrated, of course, by small holes for the passage of the particles being accelerated) are introduced midway between the power sources. If these reflecting walls are absent, the field at a given point of the tube builds up by having the waves sent out by more and more distant power sources successively reach a given spot, each one weaker than the preceding ones on account of the attenuation. With the reflecting walls, these successive waves are the reflected waves, coming back again and again to a given spot after more and more reflections. We shall find later that in many ways it simplifies our thinking about the standing-wave tube, and may actually simplify its construction, to subdivide it in this way into short sections.

In a traveling-wave tube, the situation is somewhat different. In a long traveling-wave tube, we may feed power in at uniform spacings along the length, much as in the long standing-wave tube. It is worth considering how to feed a traveling wave with an oscillator. A traveling wave can be built up of two standing waves, a cosine-like and sine-like wave, with a phase difference of 90° in time. To excite the traveling wave, we must excite both cosine-like and sine-like components separately. Thus we must have at least two points along the tube at which we feed in power, at the antinodes of both components, and by some form of phasing circuit we must feed these components in quadrature with each other. This phasing circuit can be as simple as a quarter-wave delay line, so that both components can be fed from the same oscillator.

If now we consider the build-up of a long traveling-wave tube, fed from uniformly spaced oscillators, the problem is substantially like that of feeding a standing-wave tube, except that the waves travel only one way. The field will again not be built up until the signals from distant oscillators have reached a given point of the tube, and this build-up time will again be T_o. There is an essential difference between the traveling-wave and the standing wave excitation of the long tube, however, near the input end. In the traveling-wave tube, at this end, with equally spaced oscillators, there are no reflected signals coming back, from the input end, as there were with the standing-wave tube; the field has no chance to build up by the successive signals of more and more distant oscillators, for these signals are all traveling away from the input end. The signal will build-up more promptly at the input end, but will not build up to as high a value as farther down the tube. To compensate this, we should have to add an extra power source at the input end, to supplement the uniformly distributed sources down the length of the tube. As the signal from this extra source was attenuated, in a distance comparable to the attenuation length, the signals from the uniformly spaced oscillators would build up to supplement it, and give a constant acceleration.

In a short traveling-wave tube, in which the time T_1 for signals to travel from one end to the other is small compared to T_o, it is clear that the field will have built up completely in a time of the order of T_1. For a signal sent from one end will have traveled the length of the tube, and will have been lost in the non-reflecting termination, in this length of time. Since power is being fed in only for a time T_1, short compared to T_o, before the field reaches its limiting value, it is clear that the tube will never build up the stored energy, and the accelerating voltage, of a standing-wave tube of the same length, in which the field can be reflected back and forth, remaining in the cavity for the whole time T_o. This is just another way of considering the same decrease of accelerating voltage in going from a short standing-wave tube to the short traveling-wave tube, which we have considered already.

<u>6. Geometrical Factors Affecting Acceleration</u>. In the preceding section, we have found a formula (9) for the amplitude of the accelerating field resonating with the particle, in terms of the power input to the tube per unit length, the unloaded Q of the cavity, and the wavelength. This held for a short standing-wave tube, and for any long tube, whether operating as a standing-wave or a traveling-wave tube. There are several quantities which we may vary in this formula, in an effort to get the maximum acceleration possible. These include the wavelength, the geometry, and the choice of what mode to operate in, including the question of standing or traveling wave. In the present section we shall consider these factors. We shall also consider the various accelerator projects, and the performance to be expected from their designs.

First we consider the effect of wavelength. It is well known that the Q_o of a resonant cavity is proportional to its volume, divided by the volume of a surface layer whose depth is the skin depth δ, the constant of proportionality being of the order of magnitude of unity. The skin depth is given, in terms of the conductivity σ (in mhos per meter) and the free-space wavelength λ_o (in meters) by the expression

$$\delta = \sqrt{\frac{\lambda_o}{377\pi\sigma}} .$$

For a series of geometrically similar cavities of different wavelengths, the volume of course will be proportional to the cube of the wavelength, and the surface to the square of the wavelength. Thus we shall have

$$Q_o = \text{constant} \frac{\lambda_o}{\delta} = B\sqrt{377\sigma\lambda_o} ,$$

where B is a numerical constant of the order of magnitude of unity. Substituting in (9), we find

$$E = \alpha\sqrt{B}(377P)^{1/2}(377\sigma/\lambda_o)^{1/4} . \qquad (13)$$

In other words, for a given power input per unit length, the voltage per unit length of acceleration is inversely proportional to the fourth root of the wavelength. Other things being equal, this indicates that a short wavelength is better than a long one for a linear accelerator. However, the variation is so slow with wavelength that the mechanical convenience, availability of power sources, and such considerations are more important than wavelength. Also (13) shows that the acceleration is proportional to the fourth root of the electrical conductivity. This suggests the obvious step of using a good conductor, such as copper or silver, for the material of the accelerator. It also suggests that an improvement in acceleration could be obtained by going to low temperatures, at which the conductivity of metals will increase; but the improvement is so small that it seems likely that the expense of refrigeration would be greater than the expense of providing extra power.

Next we consider the effect of geometry. This effect is principally on the coefficient α, though it also affects the Q_o, and hence B. From its definition, α will be large if the field along the z axis, where the particles are to be accelerated, is large for a given stored energy, or conversely if the stored energy, and hence the fields elsewhere in the cavity, are small for a given field along the axis. As a first step in seeing how to accomplish this, we note that the components of electric and magnetic field in the traveling wave resonating with the particle will vary with the distance r from the axis according to a Bessel's function, J_o or J_1, of the argument

$$2\pi r\sqrt{\frac{1}{\lambda_o^2} - \frac{1}{\lambda_g^2}} .$$

Since the phase velocity of this wave is given by $v = c\lambda_g/\lambda_0$, and must be less than the velocity of light, we see that the argument of the Bessel function must be imaginary, or zero in the limiting case where $v = c$. Now the Bessel functions of an imaginary argument increase more or less exponentially, so that we see that the field components will increase rapidly as we go away from the axis, more rapidly as the velocity becomes smaller. In the limiting case of the velocity of light, we find that the z component of electric field is independent of r. This behavior is in contrast to what is found in the unloaded guide, where the phase velocity is greater than the velocity of light, the field components are given by Bessel functions of real arguments, and decrease as we go out from the axis.

If we had only the Fourier component which resonates with the particle, then, we should find a large amount of stored energy off the axis, and a correspondingly small α. However, the other Fourier components become more and more important as we go away from the axis, and they have two effects: first, to allow the boundary conditions at the surface of the guide to be satisfied, which cannot be done with the resonant component alone; and secondly, to cancel part of this large field off the axis, and hence to improve α. It would be very hard to give a general discussion of the best structure, from the point of view of large α and large Q_0. For this reason we shall proceed by specific example, speaking of the structure of the M.I.T. accelerator, and asking how it could be improved by various possible changes of geometry. In Fig. 5, we show lines of electric force for

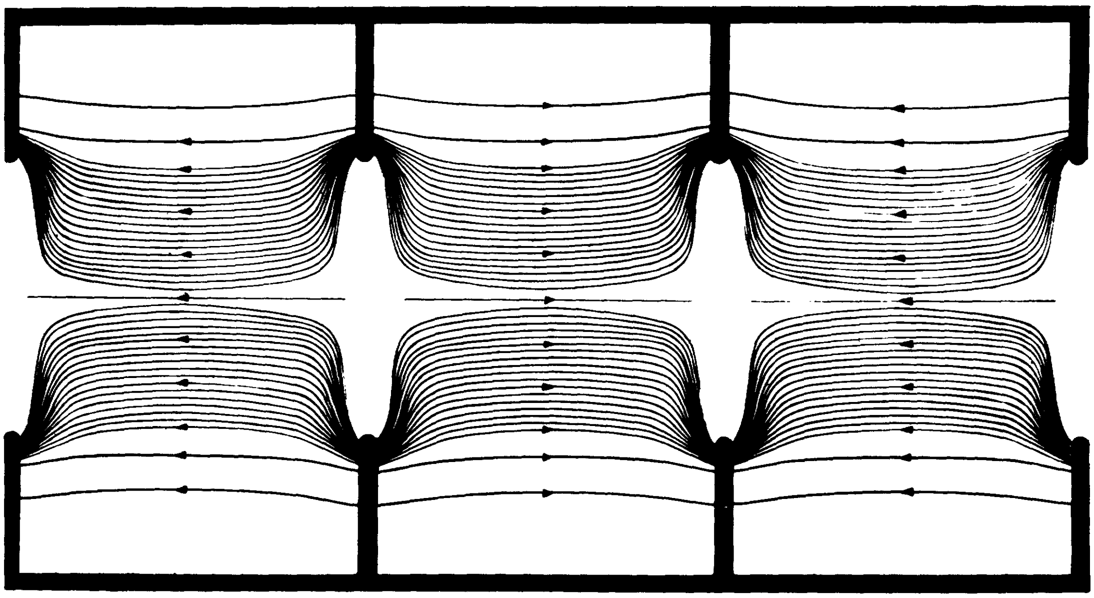

Figure 5. Lines of force in the M.I.T. accelerator.
Spacing of lines indicates field strength.

this case. It operates in the π-mode (that is, with a phase difference of π, or 180°, from one iris to the next), and is designed for $v/c = 1$. We see that the field midway between irises stays approximately constant with increasing r, out to about the radius of the hole in the iris, and then begins to decrease, very much as in a closed cylindrical cavity. This suggests that if the hole in the iris were smaller, the decrease of field would start earlier, and the stored energy would be decreased. This is in fact the case. The calculated value of α for the M.I.T. case is 0.48, while the limiting value when the hole in the iris is reduced to zero, and we have a closed cylindrical cavity, is 1.03. The holes in the irises in the M.I.T. case were made rather large so as to be sure to let the electron beam through. If it should be found that the beam was really more concentrated, a future design could use considerably smaller holes, with consequent improvement of α. This change of geometry makes comparatively small change in Q_0, or in B; the value of B for the M.I.T. case is about 0.45, compared to 0.39 for the limiting value of small holes. Thus the quantity $\alpha\sqrt{B}$, which is concerned in Eq. (12), increases from 0.32 for the M.I.T. case to 0.65 for the case of small holes. The maximum improvement in voltage which could be secured in this way is thus a factor of 2, and much less than this would be realizable in practice, since a considerable hole must certainly be left for the electrons. The Stanford project uses a considerably smaller hole than the M.I.T. project, and its geometry is probably closer to what would be used for construction of an actual accelerator.

The example we have just used is for $v = c$. For a velocity less than that of light, the situation changes. In the first place, as we have already seen, in this case the field tends to increase rapidly as we go away from the axis, instead of staying constant. This increase will presumably continue until we reach approximately the radius of the hole in the iris, after which it will start to decrease. Thus we may expect a considerably poorer α for large hole size, for small velocities. However, as the hole size is decreased this situation improves rapidly, and the limiting value of α for zero hole size is independent of the velocity of the wave. This is not true, however, of Q_0 and of B. As the velocity is decreased, the irises must be spaced closer together, and there is more surface loss in proportion to stored energy, so that Q_0 decreases, going in the limit proportionally to the velocity of the wave. Thus for very small velocities this type of geometry is inefficient. This is not serious for electron acceleration; the M.I.T. project proposes to use a Van de Graaff generator for injection of electrons at 2 million electron-volts, at which the velocity is essentially equal to the velocity of light, and most other electron acceleration projects will start the electrons with at least half the velocity of light. For positive ions, however, with their much smaller velocities, the iris type of cavity is unsuitable, until the ions are accelerated up to a fairly high speed. The structure shown in Fig. 1b, used in the Berkeley proton accelerator, is much better in this range. A rough calculation indicates that it has α equal to about 1.0, and B to 0.50, so that $\alpha\sqrt{B}$ for it is about 0.70, more than twice as good as for the M.I.T. case, and better than the case of iris cavities with small holes. This geometry, however, is not satisfactory for high velocities, for the diameter of the drift tubes becomes smaller as the velocity increases, and even on the very large scale of dimensions resulting from the 200-megacycle frequency the holes in the drift tubes are too small to admit the proton beam. It is likely that a change from this geometry to something more like the iris tube will be indicated, as the velocity of the ions increases.

We have so far been considering operation of the iris tube in the π-mode; now let us ask if we should have an improvement by operating in some other mode. For instance, we could introduce irises, just like those of Fig. 5, midway between them. Then just the same field pattern would still represent a solution of Maxwell's equations, for the electric lines will meet the new irises at right angles, and will satisfy the correct boundary conditions there. Thus we shall still have resonance, with the same guide wavelength, frequency, and phase velocity; but now the irises will be twice as close together, the phase change from one iris to the next is only 90° instead of 180°, and we have a $\pi/2$ mode. We now have two options. We may still set up a standing wave, or we may set up a traveling wave, a possibility which was impossible when this wave was a π-mode. The possibility of a traveling wave arises as follows. We have, with the intermediate irises, not only the solution of Fig. 5, but also another possible field, shifted along one iris spacing with respect to the first. These two possible solutions have the relation of a sine and cosine to each other, and if they are both excited together, with a 90° phase difference, they will represent a traveling wave, traveling to the right or the left depending on whether the 90° phase difference is positive or negative. From these traveling waves, we can get back to our standing-wave solution by superposing traveling waves in opposite directions, with equal amplitudes. Let us consider successively the standing-wave and the traveling-wave version of the $\pi/2$-mode.

In the standing wave with the $\pi/2$-mode, the stored energy, and field along the axis, are just as in the original π-mode tube. Hence α has the same value that we have already discussed. On the other hand, Q_o is seriously decreased. There are twice as many irises to carry current and produce loss, and yet just the same amount of stored energy as before. If all the losses were in the irises, this would lead to twice the loss, and a value of Q_o only half as great as for the π-mode. The situation is not so bad as this, since part of the loss is in the cylindrical walls of the cavity, and is the same in both cases. Calculation indicates that the Q_o for the $\pi/2$-mode should in fact be about 0.70 times that in the π-mode. The same factor is found in B. Thus for the $\pi/2$-mode the accelerating voltage, being proportional to $\alpha\sqrt{B}$, will be only 0.84 times as great as for the π-mode. It is not an advantageous mode to use, in other words, as far as power consumption is concerned. It is to be assumed that since, in going from the π- to the $\pi/2$-mode, there is this much decrease in operating efficiency, there will be some loss in going from the π-mode to any other mode, and the π-mode is presumably the best from this point of view, involving as it does the fewest irises per unit length, and hence the least circuit losses.

Nevertheless the $\pi/2$-mode has a compensating advantage, which is great enough so that in at least one project, that at the General Electric Company, it is being used in a standing-wave tube. For a $\pi/2$-mode, the group velocity v_g will have its maximum possible value, whereas for a π-mode the group velocity is theoretically zero. Thus the attenuation length ℓ_o can be considerable for a $\pi/2$-mode, whereas it is theoretically zero for a π-mode. An accelerator of considerable length can then be short in the $\pi/2$-mode, in the sense of having a length less than ℓ_o, whereas in the π-mode the situation is quite different. We shall discuss in a later section the peculiar properties of a π-mode, in the matter of its attenuation length, and shall show that one can in fact set up an accelerator of quite finite length in this mode, and still have it act as a short accelerator; but the corresponding length for the $\pi/2$-mode is much greater. Now we shall see later that the problem of feeding power into a short accelerator is much simpler than for a long accelerator.

In the G. E. project, their interest is in accelerators of a medium length; and it happens that by using the $\pi/2$-mode they can treat it as a short accelerator, with consequent ease in feeding, whereas if they had used a π-mode it would have been considered long, and their feeding problem would have been more severe. On the other hand, in the M.I.T. project, we are interested in facing the problem of feeding a long accelerator, and this advantage of the $\pi/2$-mode is not present.

The other possible use of the $\pi/2$-mode is to feed a traveling wave, and in this form it is being used in the Stanford project. Here there is still the decrease in B on account of the losses in the extra irises, but there is an increase in α which more than compensates for it. Let us consider the standing wave, which is equivalent to operation in the π-mode, to be set up by superposition of two traveling waves, in opposite directions. Only one of these traveling waves is resonant with the electrons. By eliminating the traveling wave in the opposite direction, which means eliminating half the stored energy, the traveling-wave tube still retains the same amount of acceleration as the standing-wave tube. The traveling-wave tube, in other words, produces the same acceleration for half the power, and hence from the derivation of Eq. (9), its α will be $\sqrt{2}$ times as great. This is not strictly true in case even the traveling wave contains an appreciable component of the reflected wave, as it will as we approach the π-mode; then the improvement becomes less and less, until as we approach the π-mode the factor α for the traveling wave approaches that of the π-mode continuously. For the $\pi/2$-mode, however, we can expect just about the improvement $\sqrt{2}$ in α. When we couple this with the decrease in B by the factor 0.7, we are led to a net improvement in acceleration of $\sqrt{2(0.7)} = 1.18$ in going from the π-mode to the traveling wave in the $\pi/2$-mode.

If there were no compensating disadvantages in traveling-wave operation, this would be an improvement worth having. In fact, in the Stanford accelerator, by the combination of going from standing to traveling wave, and by reducing the size of the holes in the irises, the factor $\alpha\sqrt{B}$ is improved with respect to the M.I.T. accelerator by about a factor of 1.68, from about 0.32 to about 0.54. This is balanced, however, by two disadvantages. First, we have already seen that a short traveling-wave accelerator, such as the model that is being built at present at Stanford, suffers with respect to a standing-wave accelerator by a factor $\sqrt{\ell/(\ell + \ell_0)}$. The present design of the Stanford accelerator has an attenuation length of about 47 feet (I am indebted to Prof. Hansen for this figure), and the total length will probably be about 20 feet. Thus this factor is $1/\sqrt{1 + (47)/(20)} = 0.54$. Multiplying by 0.54, the net value of the numerical coefficient for the Stanford accelerator should be only about 0.29, as opposed to 0.32 for the M.I.T. accelerator. This situation would be corrected, however, if the length were made of the order of magnitude of ℓ_0, or much longer, as in the eventual plans at Stanford. As we have pointed out earlier, there should be a range of length in the neighborhood of ℓ_0 where the performance of such a traveling-wave accelerator would be slightly better than that of the standing-wave model. The second disadvantage of the traveling-wave accelerator is encountered for long accelerators, and is associated with the difficulty of frequency stabilization of the oscillators. This will be taken up later, and in the writer's opinion furnishes a strong reason for using standing-wave methods for long accelerators, provided they are fed with self-excited oscillators.

It is interesting to compare overall performance to be expected from three projects, at M.I.T., Stanford, and Berkeley. Starting with Eq. (13), we may put in numerical values for the conductivity (we assume 5.5×10^7 mhos per meter for copper at 200 megacycles, and 0.8 of this value at the microwave frequencies; for it is almost always found that the observed unloaded Q of a microwave cavity at 10 cm wavelength is less than the theoretical value by about this amount). Then inserting numerical values, (13) leads to

$$E(\text{megavolts per meter}) = 7.3 \frac{\alpha \sqrt{B}\, (P(\text{megawatts per meter}))^{1/2}}{(\lambda_0(\text{meters}))^{1/4}}$$

where we are to insert the additional factor $\sqrt{\ell/(\ell + \ell_0)}$ for the traveling-wave case, and where the numerical factor 7.3, which involves the conductivity of copper, is to be replaced by about 7.0 in the microwave case. Then, inserting the values mentioned already in the text, we find the constant $C = 7.3\, \alpha \sqrt{B} \sqrt{\ell/(\ell + \ell_0)}/(\lambda_0)^{1/4}$ to be

$$\begin{aligned} C &= 3.9 \text{ for the M.I.T. accelerator,} \\ &= 3.6 \text{ for the Stanford accelerator,} \\ &= 4.6 \text{ for the Berkeley accelerator.} \end{aligned}$$

The figure for the Stanford accelerator would be 6.6, if it were not for the correction on account of its short length. We see from these figures, in other words, that the various accelerator projects attain similar values for acceleration for the same input power per unit length, in spite of their different structures. The differences in expected acceleration come largely from the differences in power input. Thus in the M.I.T. accelerator there will be one magnetron per unit length of 32 cm. This magnetron will probably deliver about 0.8 megawatts, of which about half will be fed into the accelerator, giving $P = 0.4/0.32 = 1.25$ megawatts per meter, resulting in $E = 4.4$ megavolts per meter. In the 6.4 meters of the accelerator, the total acceleration of the electrons should be about $4.4 \times 6.4 = 28$ million electron-volts. In the present preliminary model of the Stanford accelerator, the design calls for only one magnetron, in its length of about 20 feet, or 6.1 meters. Being a traveling-wave tube, if it used the same magnetron as the M.I.T. project, it might well feed the whole 0.8 megawatts into the accelerator, giving a power per meter of $0.8/6.1 = 0.131$ megawatts per meter, resulting in $E = 1.30$ megavolts per meter, or a total acceleration of about $6.1 \times 1.30 = 8.0$ megavolts. On the other hand, the Stanford project hopes eventually to feed with high-power klystrons of new design, operated as power amplifiers, periodically spaced with spacing comparable to ℓ_0, and it is too early to predict what acceleration can be expected from it. The Berkeley accelerator has sixteen oscillators, delivering about 130 kilowatts each, in its length of about 12.2 meters, so that the power is $0.13 \times 16/12.2 = 0.17$ megawatts per meter, resulting in $E = 1.9$ megavolts per meter, or about 23 million electron-volts in the 12.2 meters of the accelerator.

7. Input Impedance of Standing-Wave and Traveling-Wave Tubes. We shall find the distinction between long and short accelerators, which we have taken up in preceding sections, to be a distinction of the greatest importance in the method of feeding power into the accelerator, as well as in the voltage to be expected from it. For considering power feeding, we must investigate the input impedance of the accelerator, looking into it through the input through which the power is fed. If we are using standing waves, there will be resonant modes, separated from each other in frequency, and the input impedance will go through a resonant peak at each of these resonant frequencies. The breadth of a resonant peak, in frequency units, will be to the whole frequency in the ratio $1/Q_o$, by familiar properties of resonances. If this breadth is small compared to the frequency separation between neighboring modes, the modes will be well separated. Then, as we can show later, we can use the cavity resonance to stabilize the frequency of the oscillator, obtaining what we might call a resonant feed. We shall show that this is the case when the accelerator is short compared to the attenuation length. On the other hand, we shall show that if the accelerator is long compared to the attenuation length, the modes will run together, and resonant methods of feeding cannot be used.

In a traveling-wave system, on the other hand, there is no trace of resonance in the input impedance, the cavity cannot be used to stabilize the oscillator frequency, and we shall have to use other methods of frequency control. The reason is that resonance, as far as the impedance is concerned, is produced by the interference of the direct and reflected waves, and there is no reflected wave in the traveling-wave system. There is only a pass band, with gradually changing input impedance, which between the pass bands becomes purely imaginary. Let us inquire how the resonant structure of the pass band, containing many resonant peaks, which we meet with the standing-wave system, becomes lost in the traveling-wave system. The mechanism is simple, and we can look at it in two ways. First, as we go to a long accelerator, the amplitude of the reflected wave from a reflecting termination will become small; thus even in the standing-wave case there can be only a small resonant effect on the input impedance, which will vanish in the limit of very long tubes. Secondly, as we change the termination of the tube from a reflecting termination (for a standing-wave tube) to a resistive termination (for a traveling-wave tube), we decrease the loaded Q of each resonant mode. This increases the breadth of the resonant peaks; and we shall show that the broadening on account of the resistive termination is just enough to make neighboring resonance peaks overlap each other, so that they can no longer be resolved, in a spectroscopic sense. This statement is essentially equivalent to that of the preceding paragraph. For if the frequency separation between adjacent lines is $\Delta\omega$, the frequency ω, then the condition for the modes to run together, in a standing-wave tube, is

$$\frac{\Delta\omega}{\omega} \sim \frac{1}{Q_o} \quad .$$

On the other hand, the condition that the contribution of the termination losses to the broadening should be enough to make the lines run together, from (10), is

$$\frac{\Delta\omega}{\omega} \sim \frac{v_g}{\omega\ell} = \frac{1}{Q_0}\frac{\ell_0}{\ell} . \tag{14}$$

We shall show that (14) gives approximately the actual frequency separation of modes in a standing-wave tube. We see at once, then, that for lengths less than the attenuation length the broadening of the lines on account of the unloaded Q, in a standing-wave tube, is not enough to make the lines run together, while for lengths greater than the attenuation length it is.

We wish, then, to consider the spacing of the resonant modes, and to show that that spacing is given approximately by (15). In a resonant cavity, with reflecting terminations at the ends, the normal modes of oscillation are determined by the condition that the total length, ℓ, must be a whole number of half guide wavelengths, in order that a boundary condition, such as the vanishing of the tangential component of electric field on the end surfaces, may be applied at both ends. Thus we have $\ell = m\,\lambda_g/2$, where m is an integer, or

$$\frac{1}{\lambda_g} = \frac{m}{2\ell} .$$

Thus the resonant modes come for equally spaced values of $1/\lambda_g$, spaced by amounts $1/2\ell$, or at equally spaced values of the abscissa in Fig. 2. We note that the π-mode, in Fig. 2, comes for $1/\lambda_g = 1/2L$; thus there will be N resonant modes, corresponding to $m = 1, 2, \ldots N$, for the terminated guide, the last one being the π-mode, if there are N sections in the guide, so that $\ell = NL$. Larger values of m, as is clear from the periodicity of Fig. 2, will merely repeat values already counted. We may now find the frequencies, or values of $1/\lambda_0$, of these modes as in Fig. 6, laying off the proper values of $1/\lambda_g$, and finding the

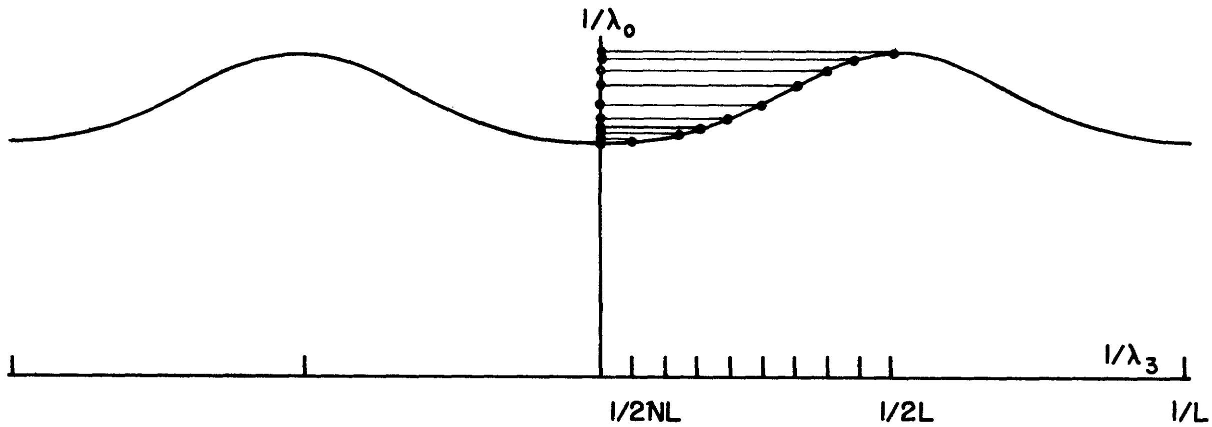

Figure 6. Calculation of resonant frequencies for resonant loaded line.

corresponding values of $1/\lambda_0$. It is clear that these modes form a group of N (or N+1) frequencies, in a restricted pass band, the modes being clustered in the neighborhood of the two edges of the pass band (or near the 0 and π-modes, corresponding to $m = 0$ and $m = N$), and more widely spaced between, (in the neighborhood of the $\pi/2$-mode, corresponding to $m \neq N/2$). Since there are N modes in the pass band, whose breadth is independent of N, it is clear that the average spacing of modes is equal to the breadth of the pass band, divided by N.

If now the spacing between two adjacent modes, in Fig. 6, is $\Delta(1/\lambda_o)$, which is $\Delta\omega/2\pi c$, then the difference of abscissas for the two points is $\Delta(1/\lambda_g) = 1/2\ell$. Then, using (5),

$$\frac{v_g}{c} = \frac{\Delta\omega/2\pi c}{1/2\ell} , \quad v_g = \frac{\ell\Delta\omega}{\pi} ,$$

and

$$\frac{\Delta\omega}{\omega} = \frac{\pi v_g}{\omega \ell} = \frac{\pi v_g}{\omega NL} .$$

Except for the factor π, which represents the numerical crudity of our methods, this agrees with (14), which we wished to prove. This also shows the spacing to go inversely proportionally to N, as we mentioned in the preceding paragraph. We have thus proved our statements made at the beginning of this section: that for a short accelerator tube, in the technical sense in which we are using it, using standing waves, the modes will be well separated, but that in a long tube they will run together; and that the broadening of the resonant lines produced by the resistive termination necessary to convert the tube from a standing-wave to a traveling-wave tube is in any case just enough to make the lines run together.

8. The Special Case of the π-Mode Tube. For the π-mode, the slope of the curve of $1/\lambda_o$ vs $1/\lambda_g$ is zero, and our formulas would tell us that the group velocity was zero, the mode separation zero, the attenuation length zero, and that any oscillator of any finite length would have to be regarded as long. The actual situation is not so extreme as this, and we take up in this section this special case, on account of its importance for the M.I.T. accelerator, and also for the Berkely accelerator, which uses essentially a 2π-mode, which has the same properties. The essential feature is that, since the slope of the curve of $1/\lambda_o$ vs $1/\lambda_g$ is zero, we can write the relation between them, near the π-mode, in the form

$$\frac{1}{\lambda_o} = \frac{1}{2L} - a\left(\frac{1}{\lambda_g} - \frac{1}{2L}\right)^2 ,$$

where $1/2L$ is the value of $1/\lambda_g$ for the π-mode, and where a is a constant. The π-mode comes for $m = N$, in the formula $1/\lambda_g = m/2NL$; the next nearest mode comes for $m = N-1$, and hence it has

$$\frac{1}{\lambda_g} = \frac{(N-1)}{2NL} = \frac{1}{2L} - \frac{1}{2NL} .$$

Substituting in the equation above, we have as the difference of $1/\lambda_o$ between the π-mode (m = N) and its nearest neighbor

$$\Delta(1/\lambda_o) = \frac{a}{(2NL)^2} \; .$$

This mode separation is not zero, as we should have concluded from our elementary theory and the fact that the group velocity is zero; but it decreases, with increasing length of tube, as $1/N^2$, rather than as $1/N$ as we have with the mode separations in the neighborhood of the π/2-mode.

For practical purposes, we may treat as an attenuation length of the π-mode cavity that length in which the frequency breadth of the π-mode resonance becomes comparable to the frequency difference between the π-mode and its nearest neighbor, as computed above. A cavity short compared to this length will have a well separated π-mode, so that it can be fed by resonant methods. The velocity of the beat phenomenon between the π-mode and its nearest neighbor will be given, in terms of this separation, by Eq. (5); we cannot speak of a group velocity, for we find easily that the beat velocity of the (N-k)th mode will be proportional to k, instead of being independent of k, so that a disturbance will not preserve its form as it spreads out in the tube. Nevertheless, in the time T_o necessary to build up energy in the cavity, the beats between the π-mode and its nearest neighbor will have traveled essentially the distance equal to the attenuation length as we have defined it, and beats between other pairs of modes will all have traveled by multiples of the length, so that the disturbance will have had a chance to spread itself over the length of the cavity fairly uniformly, though it will not have spread by regular multiple reflections as in the ordinary case. Thus for most of our purposes our definition of an attenuation length is adequate, and it is the best we can do in our present case. For the M.I.T. accelerator, the attenuation length as so defined will be of the order of magnitude of 4 or 5 meters, so that the 6-meter length contemplated will correspond to a long accelerator.

At one time in the development of the M.I.T. accelerator, when it seemed much easier to feed power into a short accelerator than into a long one, an effort was made to develop an auxiliary circuit which would convert the π-mode cavity effectively into a short accelerator, or make the attenuation length much longer than the value just discussed. This was to be done by providing a by-pass circuit, in the form of an ordinary non-loaded waveguide, connected to the accelerator tube at appropriate points, to speed the passage of power from one part to another of the accelerator tube, and hence facilitate the building up of oscillations in the correct phase relations. The by-pass circuit which was used was an ordinary waveguide, so adjusted that its guide wavelength was a rational multiple (in this case 3/2) of the wavelength in the accelerator tube. Thus there will be many points (in fact, every two wavelengths in the by-pass circuit, corresponding to every three wavelengths in the accelerator tube) where we can bridge across from one to the other with correct phase relations. The power is fed into the by-pass circuit, and from there into the accelerator tube. The properties of the connections from one circuit to the other can be chosen so that most of the excitation is in the accelerator tube, so that the losses in the by-pass circuit are not large. Thus the effective Q_o is substantially that of the accelerator tube, but the group velocity is substantially that of the by-pass circuit. The characteristic length $l_o = v_g Q_o/\omega$, which is the dividing length between short and long tubes, then becomes much greater than before, and in fact great enough so that an accelerator of very substantial size can still be considered as short. Another way of considering the effect of the by-pass circuit is its effect on the input impedance as seen from the oscillators, looking into the by-pass circuit. We find resonances only at those frequencies at which the by-pass circuit itself resonates. Being essentially an unloaded guide, these frequencies are relatively far apart. The by-pass circuit cannot be excited with appreciable intenisty at neighboring frequencies; thus it cannot feed power through to the other near-by modes of the accelerator tube, whose nearness was what made that tube itself act like a long tube.

Even though the by-pass circuit, of the type just described, can make a rather long tube effectively short, in the sense we have used, still this procedure cannot be carried far enough to make a really long tube, in the billion electron-volt range, seem short; so that in any case, for such high energies, we are forced to use long accelerator tubes. For this reason, the M.I.T. project is not planning actually to use the by-pass circuit just described, but prefers to face the problems of a long tube.

9. Feeding of Power into Linear Accelerators. In the preceding sections, we have seen that there are two very different problems of linear accelerators: the short and the long accelerator. The distinction is determined by whether the length is short or long compared to the characteristic length $l_o = v_g Q_o/\omega$, which is essentially the distance which a traveling wave will travel down the tube, moving with its group velocity v_g, in the time Q_o/ω in which the field has built up substantially to its maximum value. In a short accelerator, it is much more efficient to use a standing wave, so as to make the accelerator tube into a resonant cavity, and it then has its modes well separated from each other. The problem is then just that of feeding power into a resonant cavity. It is immaterial where in the cavity the input power is fed in. It is essential that the oscillator be operated at exactly the resonant frequency of the cavity; and if a number of oscillators are required in parallel to supply the necessary power, as will usually be the case, they must operate not only with the same frequency but in the same phase. We shall discuss this problem later. The problem is quite different, depending on whether we use self-excited oscillators, such as magnetrons, whose frequency of operation is determined by the resonant freqeuncy of the load into which they operate, or power amplifiers, in which the frequency is determined by the signal fed into the grid circuit of the amplifier. In either case, we shall find that the phasing of oscillators for a short accelerator is simple, and it has been successfully solved in practice in a number of laboratories.

The more difficult, but more important, problem is that of the long accelerator. In this case the distinction between traveling and standing-wave excitation is unimportant; the feeding problem is essentially the same in both cases. We have seen that we may subdivide the long accelerator into short sections, short compared to the characteristic length l_0, and that we must feed power separately into each of these sections. The sections may be physically separated, by reflecting walls (containing of course holes large enough for the particles to pass through), in the case of standing-wave excitation; the separations may, however, be purely imaginary walls in the standing-wave case, and must be in traveling-wave excitation, since walls would produce reflections. We may then think of the feeding of each of these sections like feeding a short accelerator; but there must be some external phasing circuit to make all the sections operate in phase. No internal phase locking, such as we shall describe for a short accelerator with self-excited oscillators, can produce the necessary phasing, for the accelerator is by hypothesis too long for a signal from one end to reach the other end before the field is built up. We must then phase externally, and this again is quite a different problem depending on whether we use self-excited oscillators or power amplifiers.

To phase a series of short accelerators requires first of all a phasing signal: we must start from one frequency standard, and amplify its signal if necessary, so that it is available in sufficient intensity at each of the short accelerators. If the short accelerators are fed with self-excited oscillators, each one can be locked in phase by feeding in a signal rather small in intensity compared to the power output of the oscillator, as we shall discuss in the next section. It should be possible, in any accelerator contemplated at present, to use a single one of the self-excited oscillators to produce the phasing signal. This oscillator would be operated separately from the accelerator tube, its frequency being stabilized by a stabilizing cavity, in the manner to be described in the next section. Its power would be divided, by suitably branching outputs, so as to flow to all the short sections of accelerator. Its pulse would start enough before the oscillators of the accelerating tube are turned on so that it would provide a signal as they build up their oscillations, and this signal would lock all the oscillators in phase as they build up their intensity. If necessary, protective switches (pre-TR's) could be put in the line, between the phasing circuit and the accelerator tube, to cut off the phasing circuit from the high power in the accelerator, once the power is built up in the accelerator. Preliminary experiments indicate that it should be possible to phase the self-excited oscillators feeding the separate short sections of accelerator in this way.

In case power amplifiers are used for operating the accelerator, a phasing signal of the type just described could be provided to each of the power amplifiers. Then by several stages of amplification this signal could be built up to the level needed for operating the accelerator. Power amplifiers are not at present available in the microwave region, though in principle it should be possible to build either high-power klystrons or resnatrons which could be operated as power amplifiers, and which might have power outputs comparable with the high-power magnetrons which are now available. The Stanford project, as mentioned earlier, hopes eventually to use klystron power amplifiers. With the gain which they would probably have, it would probably be necessary to go through several stages of amplification to raise the power level of the phasing signal to that needed for accelerator operation, and this would introduce both complication and possible distortion of the pulse shape, unless the bandwidth of the amplifier were sufficiently high. For these reasons it is not obvious that they would be more effective than self-excited oscillators for feeding the accelerator. It is possible to phase a magnetron by feeding in as little as 10^{-4} times as much power as it puts out. This makes it equivalent to an amplifier with a gain of 40 db, as far as our present purposes are concerned. No power amplifier is likely to approach this gain. It is significant that self-excited oscillators are being used in most of the existing linear accelerator projects, including not only the microwave projects, but also the Berkeley project, which, though operating at the flow frequency of 200 megacycles, where power amplifiers are available, nevertheless prefers to use self-excited oscillators, at present using triode circuits, but working on the development of a magnetron. The Yale project, however, is using power amplifiers at 568 megacycles, near the limit of frequency for which they are available.

It is obvious from the discussion we have just given that a principal part of the problem of feeding power into the accelerator from self-excited oscillators is the phasing of such oscillators, either by signals from one such oscillator to another, as in the locking of phase when many oscillators are fed into a short accelerator, or by an external phasing signal, as in a long accelerator. We now take up this problem, showing that under proper conditions a self-excited oscillator can be phased by a small external signal.

10. Power Feed from a Self-Excited Oscillator. The problems of feeding power from a self-excited oscillator into a resonant circuit, of locking many oscillators in phase, and of determining their phase by an external phasing signal, will be treated elsewhere in more detail by the writer, but we give here the essentials of the argument. In the first place, the frequency of operation of a self-excited oscillator is determined by the resonant frequency of the tank circuit or resonant cavity built into the oscillator. It is determined by a fundamental equation of operation,[11] which is

$$\frac{g + jb}{C\omega_0} = j\left(\frac{\omega}{\omega_0} - \frac{\omega_0}{\omega}\right) + \frac{1}{Q_0} + \frac{1}{Q_{ext}}\,(G + jB)\ . \qquad (15)$$

Let us describe the significance of this equation. There is a certain electronic discharge within the oscillator, which provides its driving power. This discharge is set up by a radio-frequency voltage (which is itself produced by the oscillator), and it has a radio-frequency current. The quantity $g + jb$ is the corresponding admittance, or ratio of current to voltage, g being the conductance, b the susceptance. Since the electronic discharge is non-linear, $g + jb$ is a function of voltage. In fact, self-excited oscillators in general have a component of current in phase with the voltage which decreases with increasing voltage, so that g decreases with increasing voltage, and is a definite function of voltage. Likewise b is a function of voltage, whose nature does not particularly concern us, but we note that since g and b both are determined when voltage is determined, we may eliminate the voltage from the relation and find a functional relation between b and g.

The other quantities in (15) are circuit parameters. We can show that the electronic admittance $g + jb$ is effectively in shunt with a parallel tuned circuit. Multiplying (15) by $C\omega_o$, where C is the capacity of this tuned circuit, ω_o the resonant frequency, we see that the first term on the right, $jC\omega$, is the admittance of the capacity, the next term $-jC\,\omega_o^2/\omega$ is the admittance of the inductance (it can be written $1/jL\omega$, if $\omega_o^2 = 1/LC$), and the next one is the admittance of the resistance, Q_o being the unloaded Q of the circuit. The final terms represent the admittance of the external load attached to the oscillator; $G + jB$ represents this admittance, normalized to be dimensionless, and Q_{ext}, which we may call external Q, represents the coupling of the oscillator to the load, Q_{ext} being small for tight coupling, large for loose coupling.

The operation of the magnetron is now determined in terms of the external load as follows. We separate (15) into real and imaginary parts. The real part then serves to determine the electronic conductance g, in terms of the load conductance G and the circuit parameters. As we have just seen, there is a functional relation between g and the r-f voltage, so that the voltage is determined by G. This in turn determines b. The imaginary part of (15) then gives the frequency, in terms of b and the susceptance B of the load. If for simplicity we assume that b is zero, this imaginary part gives, assuming that ω is not far from ω_o,

$$\frac{\omega - \omega_o}{\omega_o} = -\frac{B}{2Q_{ext}} \quad . \tag{16}$$

Thus the frequency of operation differs from the resonant frequency of the tank circuit by an amount proportional to the load susceptance, the change of frequency, or frequency pulling, as it is called, being smaller, the looser the coupling of the load, or the greater the external Q. In practice, the external Q of ordinary oscillators is rather small; for instance, for conventional 3000-megacycle magnetrons it is of the order of magnitude of 100. Thus a reasonable susceptance in the output load can produce a large amount of frequency pulling, or tuning, of the magnetron.

Next we consider the operation of a self-excited oscillator into a resonant load, such as we have with a short linear accelerator cavity operating in a standing wave. In this case, $G + jB$ represents the input admittance of the resonant load, suitably normalized, and reduced to a correct plane of reference for the magnetron. This will be a rapidly varying function of frequency, so that the frequency is concerned in both real and imaginary parts of (15), and we must use a more elaborate way to get the solution of that equation. The most useful method is to use a complex plane, in which we plot the locus of the right-hand side of (15) for all values of frequency, and the locus of the left-hand side for all values of voltage, and find the intersection of the two curves. This intersection determines both voltage and frequency. An example is shown in Fig. 7, representing a case actually met in linear accelerator design. This is the case where the load is a parallel resonant circuit, in series with a resistance. It can be realized in practice by taking a resonant cavity, which as seen across a suitable plane of reference is a parallel resonant circuit; inserting a suitable T in the output line of the cavity, at a suitable distance, and putting a resistance in one arm of the T; and connecting the

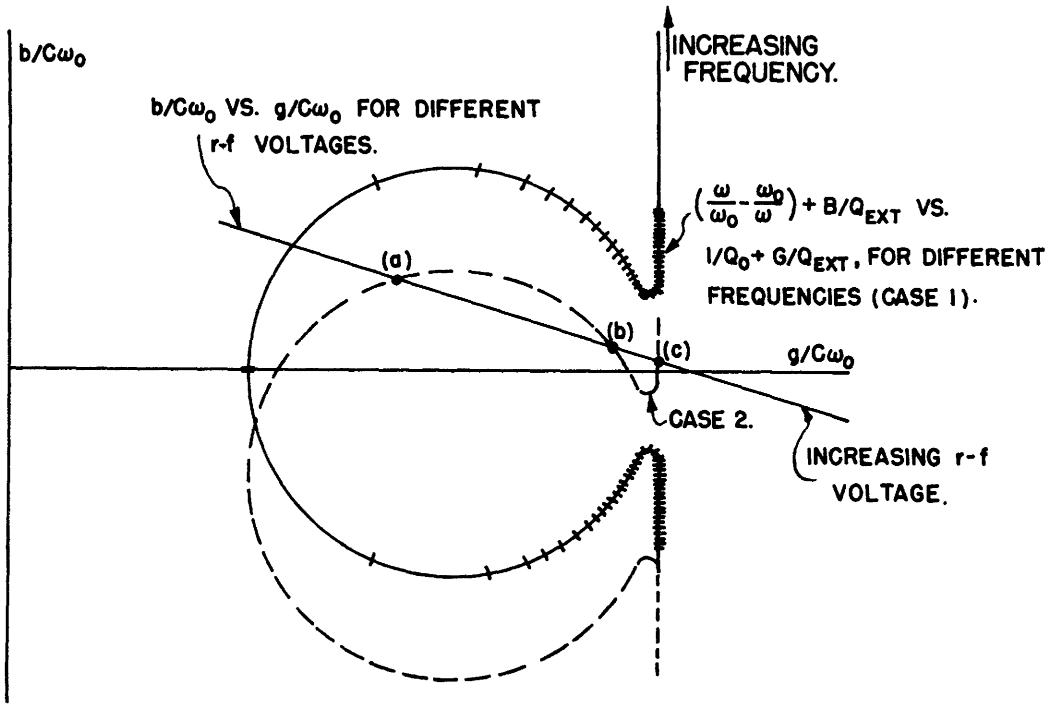

Figure 7. Determination of operating characteristics of magnetron operating into resonant load. On curve of $(\frac{\omega}{\omega_0} - \frac{\omega_0}{\omega}) + B/Q_{ext}$ vs. $1/Q_0 + G/Q_{ext}$, cross lines represent equal increments of frequency.

other arm of the T, with suitable line length, to the oscillator. (These properties of resonant cavities are taken up in Ref. 11, and the properties of T's will be discussed in a forthcoming paper by the present writer.) We show in Fig. 7 two cases, (1) where the resonant cavity is tuned to exactly the magnetron frequency, and (2) where it is tuned slightly away from it. We now see that in case (1) there is one intersection of the two curves representing the left and right sides of (15), but in case (2) there are three intersections, one corresponding to a frequency close to resonance, the others to displaced frequencies.

To understand the meaning of these three intersections, we must consider the process of starting the oscillation. One may examine a non-steady state by assuming that the frequency is complex, the real part being the ordinary frequency, the imaginary part representing a rate of build-up or decay of oscillation. The line representing the right-hand side of (15), in Fig. 7, is the locus of real frequencies; but every point of the plane corresponds to a complex frequency. As we go along the locus of real frequencies in the direction of increasing frequency, we find the region of building up of amplitude on our right, and the region of decay of amplitude on our left. We now can discuss the process of build-up or decay, for we have already noticed that an increase of voltage corresponds to a decrease of g, or corresponds to moving to the left on the g + jb curve. We thus start our oscillation, with small voltage, at a point far to the right on the g - b curve. The voltage increases, and we travel to the left on this curve. Finally, as we approach the locus of real frequencies, the rate of build-up diminishes to zero, and when we reach the locus of real frequencies we are in a steady state. Furthermore, at an intersection such as that in case (1), this is a state of stable equilibrium. For a small displacement to the left of the intersection, corresponding to an increase of voltage, carries us into a region of decreasing voltage, and the voltage is restored to its equilibrium value. Similarly a small displacement to the right results in a restoration to equilibrium. The same thing is true of the intersections (a) and (c) of case (2). On the other hand, the intersection (b) of case (2) is unstable: a displacement to the right, corresponding to a decrease of voltage, brings us into a region of voltage decrease, and this process continues, until we have displaced the point to intersection (c), a stable intersection; similarly a displacement to the left from (b) results in a motion to point (a). Thus, out of the three intersections, two represent stable modes of operation, and one an unstable one.

We next ask, which of the stable modes (a) and (c) will be set up under given circumstances? The answer is clear: if the oscillation is built up from a small amplitude, so that the point moves along the g - b curve from the right to the left, we shall build up the first mode which we meet, or (c). This, however, is not the mode representing energy in the resonant cavity; it is a useless mode, in which the resonant cavity is only slightly excited, and most of the excitation takes place in the oscillator itself. Now as we tune the cavity with respect to the oscillator, or vice versa, the effect is to shift one curve up or down with respect to the other, in Fig. 7, as from case (1) to case (2); the amount of tuning measures the vertical displacement. There will be only a small range of tuning, then, in which the desired, resonant mode is set up: that range in which the g - b curve passes through the narrow opening or neck of the approximately circular resonance curve. To widen this tuning range, we increase the series resistance; this has the effect of widening the neck. The neck closes as the series resistance goes to zero, if there is a finite length of line between magnetron and cavity, so that with too small a series resistance the system will not operate in the resonant mode, whereas with a considerable series resistance the tuning range for resonant operation is considerable; but of course with the disadvantage that the series resistance absorbs a considerable fraction of the power.

In the process of setting up this circuit, we have at the same time stabilized the frequency of the oscillator. We note that the marks of constant frequency difference, on the locus of real frequencies in Fig. 7, are very widely spaced in the resonant part of the curve. Thus a given amount of tuning of oscillator with respect to cavity, which shifts the locus up or down with respect to the g - b curve, will make a very small change in frequency. The frequency of operation must, in fact, be within the resonant band of the resonant cavity, and if this has a Q of the order of magnitude of many thousand, as it will if it forms the resonant line of a linear accelerator, the oscillator will have its frequency correspondingly stabilized. This is an important feature, for it is hard to tune most oscillators very accurately and hold them closely on a predetermined frequency, and a stabilizing cavity of this type automatically holds the frequency to the correct value for the cavity. The point is that, if the oscillator frequency is tuned slightly off from the resonator frequency, the resonator susceptance takes on such a large value that the frequency pulling resulting from this susceptance pulls the frequency of the oscillator back to the resonator frequency.

We have now considered the operation of a self-excited oscillator into a non-resonant and a resonant load. We have found that with a non-resonant load its frequency is determined by the resonant frequency of the tank circuit, but with a resonant load of high Q, tuned approximately to the same frequency as the tank circuit, the frequency will be stabilized by the resonant load rather than by the tank circuit. This is true only if we have correctly introduced a resistance into the circuit; for, as we usually find with two resonant circuits tuned to approximately the same frequency, we have two possible modes of operation, and one may be eliminated by proper use of resistance. Next we consider the operation of a self-excited oscillator in the presence of an external signal, of approximately its own frequency, fed in from outside.

An external signal of exactly the same frequency as the power being put out by an oscillator, fed in through the output line from the oscillator, cannot be distinguished from a reflected wave. It will then simulate a correction to the reflection coefficient, or to the impedance or admittance of the load, of magnitude proportional to the ratio of the amplitude of the external signal to the amplitude of the signal being produced by the oscillator, and of phase depending on the phase difference between signal and output power. In other words, it represents an addition to the last term of Eq. (15), of the order of magnitude of $(1/Q_{ext})$ times the ratio of external signal amplitude to output signal amplitude, and arbitrary phase. This term can have just the same effect on the oscillator that a resistive or reactive load would have, and in particular, it can have the effect of pulling the frequency, according to Eq. (16). If originally the oscillator were operating at a different frequency from the signal, the signal can pull the oscillator into synchronism with it, and into a definite phase relationship with it, provided there is some phase for which the reactive effect of the simulated admittance of the signal is large enough to provide the necessary frequency pulling. Examination of the transient situation when the external signal is turned on shows that such a phase locking will occur if it is possible. As soon as the external signal is present, the phase difference between external signal and oscillator will begin to approach just the right value to produce this frequency pulling, and after a very short time the oscillator will have settled down to a steady state, synchronous with the external signal, and with a phase angle between them which is greater, the greater the original frequency difference between signal and oscillator.

Thus we can lock an oscillator to an external signal, and so make its frequency exactly equal to that of the signal. A relatively small signal can produce this locking. Thus with an external Q of 100, and a signal voltage 1/10 of the voltage in the oscillator output, which means a signal power 1/100 of the oscillator power, Eq. (16) shows that the maximum frequency pulling of the oscillator which the signal can produce is about 1/2000 of the whole frequency. If the oscillator were originally tuned closer than this to the signal, there would be a correspondingly smaller phase difference between the signal and the oscillator when the steady state was reached. The situation will be altered if the oscillator were originally stabilized by an external resonant cavity, as we have discussed earlier in this section; for then as far as frequency pulling is concerned, the magnetron acts essentially as if it had the Q of the cavity, rather than its normal external Q. Thus in this case there will be a much smaller range over which locking can occur, but the oscillator is already tuned much more accurately to the desired frequency, by its stabilizing cavity, so that a large range is not needed. If the external signal lies well centered in the resonant band of the stabilizing cavity, it will still be able to lock the oscillator into synchronism.

Even with a signal smaller than that required to produce locking in the steady state, it may be possible to start an oscillator in phase with the external signal, and for it to continue in phase simply on account of having its frequency agree well enough with that of the signal. During the build-up of an oscillator, naturally its amplitude is less than when it has reached the steady state, and the signal voltage will be larger in proportion, and hence will correspond to a larger simulated admittance, capable of producing more frequency pulling. Thus it was found during the war [12] that a signal only 10^{-4} times as strong in power, or 1/100 as strong in voltage, as the final signal of a magnetron, was capable of starting it off in phase. Furthermore, it is possible to tune two oscillators, stabilized by stabilizing cavities, so closely together that they do not drift out of phase by more than a small fraction of a cycle in the time of a pulse of the length which we wish to use in a linear accelerator. Thus if relatively weak signals are used to start the oscillators in the correct phase, it should be possible to adjust them so that they stay in phase.

We have spoken about the phasing of an oscillator by external signal; let us consider also the phasing of an oscillator by another. We can have a resonant cavity, such as a short linear accelerator tube, with a number of feeds, and an oscillator feeding into each. Each oscillator then communicates with each other, through the cavity, and if the cavity is short, the time required for a signal to go from one to another is short compared to the time of build-up of the field in the cavity. Each oscillator then sends a signal to each other, which can be used to lock the oscillators in phase with each other. The coupling is so strong, and the phasing consequently so positive, that the oscillators under these circumstances will lock very completely together. Closer examination of the strength of locking shows that if each oscillator is tuned near enough to the frequency of the cavity so that it locks to the stabilizing cavity, it will have far more signal from the other oscillators than it needs to lock to them in phase. The only trouble with this system is a short period of confusion at the beginning of the pulse, when the different oscillators are competing to establish the phase in which they will all operate. This can be avoided, however, by introducing a phasing signal before they start, either from outside, or by starting one of the oscillators ahead of the others. In that case, this signal starts the oscillators in phase, and they continue to lock together, in a very positive and satisfactory way.

From what we have seen in this section, it then seems very likely that a long linear accelerator can be operated with self-excited oscillators in the following manner. It is subdivided, either by actual physical subdivisions or by imaginary surfaces, into a set of short sections. Into each of these sections a phasing signal is fed, from a pulsed oscillator which supplies all these sections, through lines whose lengths of course must be adjusted to give proper delays and phases. These phases may be checked by an auxiliary low-level control circuit, running parallel to the main accelerator tube, whose signal may be mixed with the signal generated in each short section, to test both its phase and frequency. Each of the short sections then will have a number of self-excited oscillators feeding into it, which are triggered a short interval after the phasing signal has started. They will start oscillation in synchronism with the phasing signal, will then lock to each other and be frequency stabilized by the accelerator cavity (if this is a standing-wave cavity), and if all short sections are tuned sufficiently closely to the same frequency, all the short sections will operate in phase with each other for the duration of the pulse.

It is clear that careful tuning of the various short sections to the same frequency is essential; in the next section we take up the question of tolerances, to see whether it is practicable to hold the necessary tolerances in a very long accelerator.

11. Tolerances in the Long Accelerator. In order that an electron or ion may resonate with the appropriate Fourier component of the field, that field must have the correct propagation velocity. For positive ion acceleration, where for most energies the velocity is small compared to that of light, we shall see in later sections that there is a stable bunching phenomenon, by which bunches of particles are formed, and automatically adjust their velocity to the velocity of the field. Thus in this case a very accurately determined velocity of the wave is not necessary. However, with electrons, they reach the velocity of light to all intents and purposes in the first few feet of the accelerator, and for the rest of its length they travel with exactly this velocity. Thus the wave in the tube must travel with the velocity of light as well. Furthermore, if at the end of the accelerator the phase of the wave is different from what it should be by an appreciable fraction of 2π, the electrons will not be in the correct phase relationship, will not be properly accelerated, and the accelerator will not operate properly. It is necessary, then, to have this phase correct within a small tolerance, and this demands accurate mechanical dimensions and accurate frequency control. We shall ask in this section what this demands, and whether the requirements can be met.

The situation is quite different, depending on what type of accelerator is being considered. The difficulties become severe only with long accelerators. Let us start, then, by considering the long accelerator made up of many short sections, each of these short sections being resonant, and the various sections controlled in phase by a phasing circuit. First we consider the conditions on an individual short section, then on their relative phasing. The short section will have a sharply defined resonant frequency, the breadth of the resonance peak being to the whole frequency in the ratio $1/Q_0$. We first note that it is unnecessary to control the frequency of the signal to much better than the breadth of this peak. The reason comes in the pulse length used. As we have stated, most of the time of the pulse will be occupied in building up the oscillation. The time required for the electron or ion to traverse the accelerator, after the field is built up, will not be large compared to the build-up time. Thus in the M.I.T. accelerator it will require about 2 microseconds to build up the field, and the electrons will be turned on for about 1 microsecond more. But the spectrum of a short pulse (in this case, a 3 microsecond pulse) of perfectly monochromatic vibration is broadened, the frequency breadth being to the total frequency in the ratio of the reciprocal of the number of waves in the train. In our case, the number of waves in the train is of the order of Q_o, since this is the number of periods required to build up the field in the cavity. Thus the spectrum width of the pulse will be comparable with the breadth of the resonance peak of the cavity. Since the cavity, operated as a stabilizing cavity, can hold the frequency of a self-excited oscillator to a tolerance well within the bandwidth of the cavity, it should be easy to hold the frequency within the required limits.

There is the possibility, however, that the cavity may not be constructed to have exactly the right resonant frequency; an incorrect relation of frequency wavelength will naturally lead to the wrong velocity. In practice, one can construct a short section having the correct resonant frequency, for a given length, to about two or three parts in ten thousand, and this is not accurate enough for our purpose, since with a Q_o of the order of 18,000, as in the M.I.T. accelerator, the bandwidth will be narrower than this. It seems, therefore, that it will be necessary to introduce a tuner into each short section, so as to adjust its frequency slightly, and correct errors in machining. It should not be hard to make such tuners, and the problem of testing the tuning is essentially simple. We have only to take a standard signal, whose frequency is just what is desired, and which will be an essential feature of the system; to beat this signal with the pulsed power in one of the short sections; and to observe whether there are any beats between the two signals in the time of a pulse. If the accelerator section and its oscillators are tuned so that there are no beats in this length of time, that establishes the correctness of the frequency, to the required accuracy. Furthermore, if each section is tuned in this way to the standard signal, it is clear that if all sections are started in phase with each other, they will remain in phase with each other during the pulse, to the required accuracy, and this is all that one can ask. It is not difficult in practice to tune stabilized oscillators to this degree of accuracy, and to maintain the tuning over periods of time.

We now ask about the relative adjustment of short sections of the accelerator, at long distances from each other, say at opposite ends of a long accelerator. If we are thinking of an accelerator many tens of thousands of wavelengths long, it is clear that an error of a small fraction of a wavelength will be a very small fractional error indeed. One should be able to hold manufacturing tolerances closely enough to make a tube of this length, with errors of a few parts in a hundred thousand in its total length, for there will be widespread cancellation of chance errors. If the over-all length is somewhat wrong, however, this can be compensated by an appropriate change in the resonance frequency from the nominal value, which can be done if each section is equipped with a tuner. For the sake of definiteness, suppose the accelerator is made, as the M.I.T. one is, to operate in the π-mode. Every half wavelength there is an iris. Therefore the total number of half wavelengths in the whole tube can be found merely by counting. By the method of operation which we are suggesting, it is certain that each short section will operate in the π-mode. Then the wavelength of operation is necessarily the total length, divided by the number of wavelengths. It is then a matter of arithmetic, together with a knowledge of the velocity of light, to determine the correct frequency, and by means of the separate tuners of the short sections, they can all be tuned to operate with this frequency. This demands, of course, accurate measurement of length, and accurate determination of the phase at which each section must operate, to fit correctly into the whole standing wave. These determinations can be handled easily if we have a control circuit, parallel to the accelerator, in which there is a signal fed from the standard signal generator, traveling with exactly the velocity of light, as it will if it travels in a coaxial line or other principal mode. Each short section can then be compared in phase, as well as in frequency, with the signal at the appropriate point of the control circuit.

By the means just described, it should not be hard to adjust an accelerator consisting of short resonant sections, each independently adjustable in frequency and phase, to the required tolerance. This circuit, as we have seen, automatically contains a stabilizing circuit for the oscillators, of the required sensitivity. When we consider a traveling-wave system, however, the situation is quite different. Here there is no resonant system for frequency stabilization. If we use self-excited oscillators, we must either rely on the frequency constancy of the separate oscillators as they stand, or must supply separate stabilizing cavities for them. If we do the latter, we add much complication, and lose much of the advantage of the traveling-wave system. If we do not, it is a grave question whether the oscillators can be kept in tune accurately enough to synchronize over a pulse, and to be locked in phase by the initial phasing signal. The frequency put out by a magnetron is a function of its temperature and operating conditions, and these are hard to regulate. The stabilizing cavity smooths out fluctuations from these causes, and it seems to the writer practically necessary for proper operation. It is significant that the British group at TRE, who are designing a traveling-wave accelerator to be operated by a magnetron, indicate in their reports that frequency variation is one of the features limiting the possible length, and the frequency variations which they contemplate are much larger than those which would be met in the case of frequency stabilization by a resonant accelerator, seriously limiting their estimate of the maximum acceleration which they can obtain. The other possibility with a traveling-wave tube is to feed from power amplifiers, in which the frequency is controlled from a well-regulated phasing signal. This seems to the writer the only method of feeding a long traveling-wave tube which is likely to be successful. This feeling is shared by the Stanford group, which are basing their plans on a high-power klystron amplifier, still to be designed.

An additional problem with a traveling-wave tube is that there is no way of dividing up the tube into physically separate short sections. Thus we lose the possibility of easily testing the resonant frequency of each section, and adjusting them to synchronism by separate tuners. We can, of course, terminate each section by reflecting ends during construction, test its resonant frequency, and adjust this with tuners inserted periodically down the line. After assembly, however, there is no way of testing by checking the resonant frequencies. It seems likely that the best method of testing would be to have a parallel phasing signal, in the form of a traveling wave in a coaxial line, and to make phase comparisons between this signal and that in the main accelerator, at various points down the line. This type of comparison would be more complicated than in the standing-wave case, but probably would not be impossible.

If high-power klystron power amplifiers were available at present, it is entirely possible that on account of the small advantage in the matter of acceleration possessed by the traveling-wave accelerator, it would be the most desirable type to build. On the other hand at present, when high-power magnetrons are available and high-power klystrons are not, the problem of frequency and tolerance control of the standing-wave accelerator is so much simpler than with the traveling-wave accelerator that the M.I.T. project is preferring to use a standing-wave system.

12. The Dynamics of Particles in the Accelerator. So far, we have been speaking entirely of the mechanism of setting up an accelerating field in the tube, and the nature of that tube. Now we consider the problem of the motion of particles in this accelerating field. We divide the problem into two parts: first, the longitudinal motion of particles traveling on the axis of the accelerator; secondly, the transverse motion of particles, and focussing problems. Furthermore, for considering the longitudinal motion, we divide the discussion into two parts: first, motion in an external field of constant velocity; secondly, motion in a field of varying velocity. So far, we have assumed throughout that the properties of the accelerator tube did not vary from point to point, so that the velocity of propagation of the resonant mode was constant. If the particles are electrons traveling with the velocity of light, this assumption of course is correct; but with positive ions, or with electrons as they are being accelerated up to the velocity of light, the velocity of the wave must increase to match the increase of the velocity of the particles.

For a uniform velocity, we have already seen that only one sinusoidal component of field is of importance, since the others have effects which cancel out over a period. The longitudinal, or z, component of electric field for this sinusoidal component may be written $E_z = E \sin \omega(t - z/v_0)$, where ω is the resonant frequency, v_0 the velocity of propagation. This is the value holding on the z axis, which alone we consider at the moment. In this field, a particle of rest mass m_0, charge e, will have an equation of motion

$$\frac{dp}{dt} = eE \sin \omega(t - z/v_0) \quad ,$$

where p, the momentum, is given by

$$p = \frac{m_0 v}{\sqrt{1 - v^2/c^2}} \ , \quad v = \frac{dz}{dt} \ .$$

It is convenient to introduce moving axes, moving with the velocity v_0 of the traveling wave. If z' is the displacement with respect to these moving axes, we then have

$$z' = z - v_0 t \ .$$

We can now write the equations of motion with respect to these moving axes in Hamiltonian form: essentially this possibility was pointed out to the writer by Mr. S. J. Mason, of M.I.T., and a similar method has also been used in the TRE reports. We set up a Hamiltonian function

$$H = \sqrt{m_o^2c^4 + p^2c^2} - pv_o - eE\frac{v_o}{\omega}\cos\frac{\omega z'}{v_o} \quad . \tag{17}$$

In terms of it, we verify immediately by substitution that the equation of motion may be rewritten in the form

$$\frac{dp}{dt} = -\frac{\partial H}{\partial z'} \; , \quad \frac{dz'}{dt} = \frac{\partial H}{\partial p} \; .$$

Since these are Hamilton's equations, with a Hamiltonian function not involving the time explicitly, they show that H remains constant during the motion. Thus if we set up a phase space in which z' is abscissa, p is ordinate, and draw lines H = constant in this phase space, these lines will give the relation between momentum and coordinate for a particle during its motion, determining the speed of the particle at an arbitrary point of its path.

In Fig. 8, we plot such a phase space for the case $v_o = c/2$, where the wave is traveling with half the velocity of light; later in Fig. 9 we shall consider for comparison the case $v_o = c$. As abscissa, we use the dimensionless quantity $\omega z'/v_o$, which increases by 2π when we go along the z axis by one wavelength of the wave. As ordinate, we use the dimensionless quantity p/m_oc. We also give a scale of $\sqrt{m_o^2c^4 + p^2c^2} - m_oc^2$, which is the ordinary kinetic energy of the particle, expressed in terms of the rest energy m_oc^2. Energy contours are drawn for constant values of the dimensionless quantity H/m_oc^2. As we see from (17), there is then one parameter in the Hamiltonian function, which in dimensionless form becomes

$$\frac{eE}{m_oc^2}\,\frac{v_o}{\omega} \, .$$

We can easily see the physical significance of this parameter. If λ_g is the wavelength of the resonant wave, which is traveling with velocity v_o, we have $v_o/\omega = \lambda_g/2\pi$. Thus our parameter gives the energy which the particle picks up in $1/2\pi$ wavelengths, if it is subject to the maximum acceleration of the field, divided by the rest energy of the particle. In. Fig. 8, this quantity is taken to be 0.10, which is of the order of magnitude of what can be realized in practice in the M.I.T. accelerator.

When we examine Fig. 8, we find first a set of closed oval curves, representing periodic orbits. These surround the points where the potential energy part of the Hamiltonian function (17) has a minimum, and where the momentum is that of a particle traveling with velocity v_o. These moving points of stable equilibrium are of course the positions where the force is zero, and where the particle will be pushed backward if it is

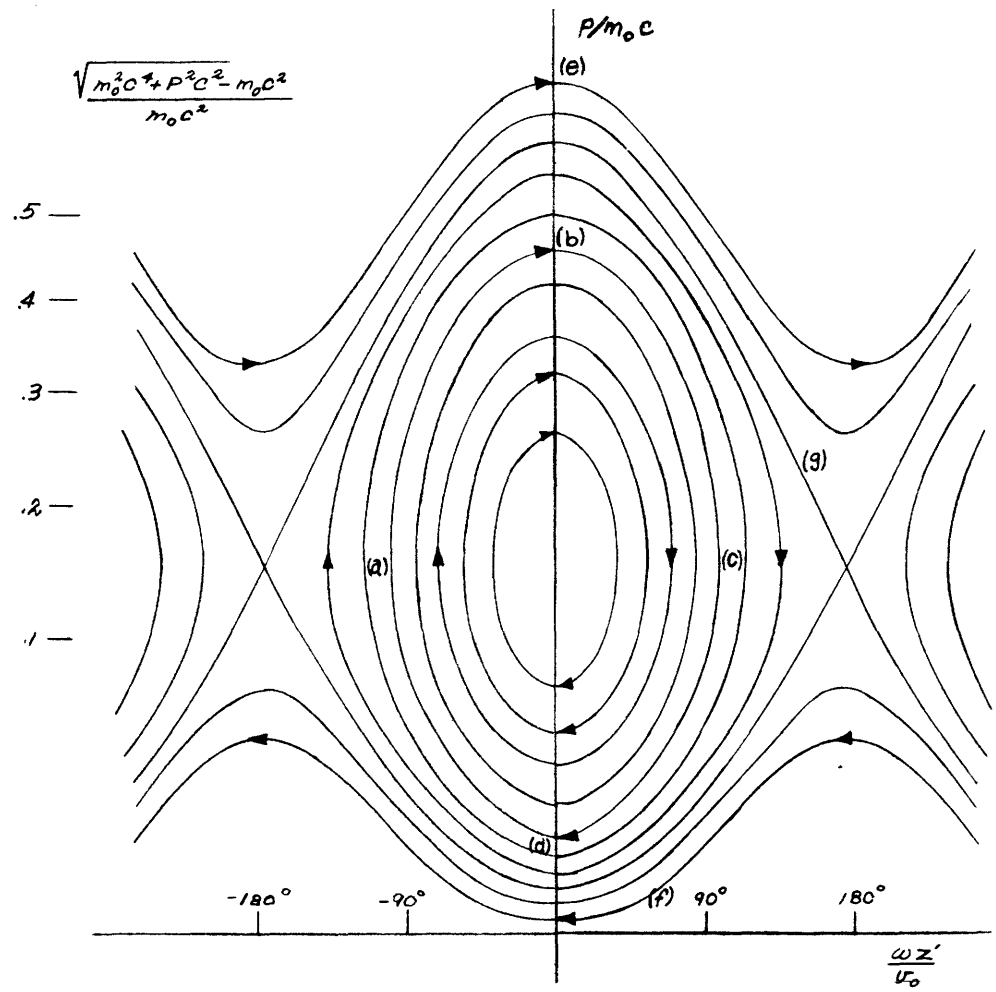

Figure 8. Phase space for $v_o = c/2$.

ahead of the position, forward if it is behind, or where the force at a fixed point of space is instantaneously zero, but is increasing as the traveling wave moves along. A particle which is lagging behind the stable position will then be speeded up, so that its velocity will get greater than v_o; it will then advance with respect to the wave, getting out of the accelerating into the retarding phase; will then slow down again to velocity v_o and below it; will lag behind, and so will repeat its cycle, as in the succession of points a, b, c, d in Fig. 8. In other words, the representative point of the particle will traverse the closed curve of Fig. 8 in a clockwise direction.

On the other hand, there can also be orbits like e in Fig. 8, in which the velocity of the particle is so much greater than than of the wave that, even though it is slowed down in the retarding phase of the wave, it still advances with respect to the wave, going from one potential minimum on to the next, and the next, and so on. Similarly there are retrograde orbits like f, in which the velocity is so much less than v_o that the particle continually slips behind the wave. The limiting orbit between the periodic and the advancing or retrogressing orbits is shown in g. It passes through a position of unstable equilibrium, corresponding again to motion with the velocity v_o, but 180^o out of phase from the position of stable equilibrium. At this unstable position, the force is again zero, as it is at the stable position; but the force is decreasing with time, so that a particle which has advanced in phase with respect to the wave finds itself in an accelerating field, and advances even more, while if it lags it is in a retarding field, and lags even more.

We recognize the physical similarity of this motion to the motion of a pendulum, a particle constrained to move in a vertical plane at a fixed distance from a pivot, with a uniform downward force. With a pendulum, motions of small amplitude are periodic, while with large energy the particle rotates round and round, slowing down at the top of its path, but never changing direction. Here also there is a limiting orbit passing through the position of unstable equilibrium: that corresponding to the case where the particle is just at rest at the top of its path, but is displaced an infinitesimal amount from this position, rolls down to the bottom, up the other side, just coming to rest again at the top, and as a matter of fact requiring an infinite time to do it. We shall show in a moment that the resemblance to the problem of the pendulum is fundamental, and not just superficial; we can use the analogy with the pendulum to derive results about periods and amplitudes of oscillation, and other properties of the motion. Thus for small amplitudes (that is, small ovals in Fig. 8), the period of oscillation is independent of the amplitude, but for larger amplitudes the period will get longer, becoming infinite in the limiting orbit which we have just discussed. As we then go into the non-oscillating orbits (which would be the rotating orbits in the pendulum problem), the time of rotation again starts infinite as we go away from the limiting orbit, but when the orbit is far from this value, corresponding in our case to a particle with velocity much greater or much less than v_o, the periodic field produces only a minor perturbation, and the motion has nearly uniform velocity, from which the period of rotation can be found at once.

We can go further with this analogy, and use it to compute periods and amplitudes, if we express our mathematics in proper form. The only thing keeping our problem from having exactly the form of the pendulum problem is the relativistic nature of the mass of the particle. During an oscillation, for instance, the energy of the particle goes from a minimum to a maximum. If this change of energy is large enough to produce an appreciable change of relativistic mass, then the problem is appreicably different from the pendulum problem, but if this change of mass can be neglected, we can reduce it to the pendulum problem. This can be expressed in mathematical form in the following way. Let us write $v' = dz'/dt = v - v_o$, and let us then expand the momentum p in powers of v'. We find for the beginning of this expansion

$$p = \frac{m_o v_o}{\sqrt{1 - v_o^2/c^2}} + \frac{m_o v'}{(1 - v_o^2/c^2)^{3/2}} + \cdots$$

$$= p_o + m_\ell v' + \cdots, \quad = p_o + p' + \cdots,$$

where $m_\ell = m_o/(1 - v_o^2/c^2)^{3/2}$ is the longitudinal mass. In a similar way we can expand the Hamiltonian function (17) in powers of v'. We find that in doing this we must carry the expansion of p one stage further than we have just done above; but the final answer is simply

$$H = m_o c^2 \sqrt{1 - v_o^2/c^2} + \frac{p'^2}{2m_\ell} - eE \frac{v_o}{\omega} \cos \frac{\omega z'}{v_o}$$

$$= H_o + H' , \tag{18}$$

where H_o is the first term above, H' the remainder. In terms of these quantities, the equations of motion become

$$\frac{dp'}{dt} = -\frac{\partial H'}{\partial z'}, \quad \frac{dz'}{dt} = \frac{\partial H'}{\partial p'} .$$

These are just the equations of motion for a particle of mass m_ℓ, oscillating in a field $-eE \sin \omega z'/v_o$, or the pendulum equations. We may now use a new phase space, in which z' and p' are abscissa and ordinate. The equilibrium position in Fig. 8 is now moved down to the axis of abscissas; but aside from this, the contours of constant H' in the new phase space will look like those of constant H in Fig. 8, with only small modifications on account of the fact that we have neglected higher powers of v'.

We may now use our knowledge of the pendulum problem to compute various properties of the motion, correct to our present approximation. The angular frequency of small oscillations is the square root of the force constant, over the mass. For small oscillations, the restoring force is $-eE\omega z'/v_o$, the sine being replaced by its argument. Thus the angular frequency of oscillation, which we shall call ω_o, is

$$\omega_o = \sqrt{\frac{e}{m_o} E \frac{\omega}{v_o} (1 - v_o^2/c^2)^{3/2}} . \tag{19}$$

It is, of course, an entirely different quantity from the angular frequency ω of the electromagnetic wave. We may also compute the amplitude of oscillation, in phase $\omega z'/v_o$, of a particle which has an oscillatory energy W in the moving coordinate system. We readily find this amplitude to be given by the equation

$$\frac{W}{m_o c^2} = \frac{1}{2} (\text{Amplitude})^2 \frac{eE}{m_o c^2} \frac{v_o}{\omega} .$$

This formula is correct for small amplitudes, where the energy is proportional to the amplitude. Its error for large amplitudes may be seen from the limiting case of the amplitude of π, where our formula would give $W/m_oc^2 = \pi^2/2\ (eEv_o/m_oc^2\omega)$, whereas the correct numerical coefficient should be 2, as we see at once from (18).

It is important to note that the quantity W, measuring the energy of oscillation in the moving coordinate system, does not represent the change in the kinetic energy of the particle in the fixed system. We can show easily that the total kinetic energy is

$$KE = KE_o + m_\ell v_o v' + \dots$$

where KE_o is the kinetic energy of a particle moving with velocity v_o. Then for instance the maximum and minimum kinetic energies of a particle moving with a given phase amplitude are

$$KE = KE_o \pm m_\ell v_o \sqrt{\frac{2W}{m_\ell}} = KE_o \pm (\text{Amplitude})\ m_oc^2 \frac{v_o}{c} \sqrt{\frac{eEv_o}{m_oc^2\omega(1 - v_o^2/c^2)^{3/2}}} \qquad (20)$$

and for the maximum and minimum kinetic energies of a particle in the limiting oscillatory orbit, the one which passes through the positions of unstable equilibrium, we have

$$KE = KE_o \pm 2m_oc^2 \frac{v_o}{c} \sqrt{\frac{eEv_o}{m_oc^2\omega(1 - v_o^2/c^2)^{3/2}}} \ . \qquad (21)$$

These formulas, we remember, hold only as long as the change of kinetic energy is small enough so that the change of mass over the oscillation is not serious; from the formula we have just written, we see that this approximation is not good for the case $eEv_o/m_oc^2\omega = 1/10$, used in Fig. 8, but even in this case it gives an answer which is not seriously in error. Formulas (20) and (21), of course, give changes of energy which could be found from the average force exerted by the accelerating field on the particle, during the half period when it is going from its minimum energy to its maximum, multiplied by the distance which it has traveled in the field, at its velocity of approximately v_o, during its acceleration.

We now consider how curves of the nature of Fig. 8 change when various parameters change. If the accelerating field E changes, the ovals become widened or flattened vertically. From (20) and (21), we see that their height is proportional to $\sqrt{E}$, for a fixed value of v_o. The case shown in Fig. 8 corresponds to a large acceleration, which can be attained with electrons. On the other hand, even with the same accelerating field, but with positive ions, (20) and (21) show that the ovals would become flattened in a ratio of the square root of the ratio of masses of electron and positive ion, corresponding to the fact that a given field can make a much smaller change in the energy of a positive ion, in proportion to its rest energy, in a given distance. The other parameter at our disposal is the velocity v_o of the wave. As this velocity is increased, to approach the velocity of light, a number of changes occur. In the first place, the position of stable equilibrium rises without limit; for its height measures the momentum of a particle of velocity v_o, and this momentum becomes infinite as v_o approaches c. At the same time the frequency of oscillation, given by (19), goes to zero, or the period becomes infinite. The reason is that even a very large change in energy, or momentum, of the particle gives it only a very small change of velocity, as v_o approaches c, so that it takes a very long time for this velocity difference between particle and wave to carry the particle back and forth with respect to the stable position of the wave. Associated with this is a great vertical elongation of the ovals; for as the period of oscillation becomes very great, the distance which the particle travels (with respect to fixed coordinates) in one of its oscillations becomes very great, and it can pick up or lose very large amounts of energy in a cycle. Clearly in this limit our approximation of replacing the problem by the pendulum problem becomes invalid.

When v_o becomes equal to c, the positions of equilibrium rise to infinity, and the ovals are no longer closed at all. In such a case we must go to our original, exact formulation for a proper answer. In Fig. 9, we give curves similar to Fig. 8 for this case, the only difference between this case and Fig. 8 being that Fig. 8 has $v_o = c/2$, Fig. 9 has $v_o = c$. In Fig. 9, the orbits are divided into two categories, rather than three. First, there are orbits which are essentially carried along with the field, as the periodic orbits of Fig. 8 are. Such an orbit is shown in a, b, c of Fig. 9. It starts with a finite momentum, therefore with a velocity less than that of light, and therefore it starts to lag behind the wave. It then gets into an accelerating part of the field, and starts to increase its momentum. This process can never speed up the particle to the speed of the wave, however, and as a result the path of the representative point in phase space approaches a vertical asymptote, so that it ends up in a given phase relation to the wave, which it preserves indefinitely. It continues to acquire momentum and energy without limit, the amount depending on the phase of the wave associated with its asymptotic position. If this asymptotic phase is near -90° (that is, the phase of the maximum acceleration), the particle will continue indefinitely to find itself in a field essentially of E, and will receive its maximum possible energy.

In addition to these orbits which are carried along with the traveling wave, and result in particles which eventually receive arbitrarily large amounts of energy, there are other orbits, such as d in Fig. 9, which retrogress, not having large enough velocity to be bound to the field. Naturally we lack the other class of orbits found in Fig. 8, those which go too fast for the field. The dividing orbit between the two classes is shown in e, Fig. 9, and is the orbit which approaches the phase -180° asymptotically. This limiting orbit comes to a definite minimum momentum, and hence kinetic energy, for zero phase. No particle with a kinetic energy less than this minimum can be bound into the accelerating orbits, but for any energy larger than the minimum, there will be some phases in which binding can occur. Although our formulas (20) and (21) will not hold quantitatively in this case, still we may assume by analogy with them that the greater E is, the lower will be this minimum kinetic energy for binding. We see that in the case of Fig. 9 the minimum comes at about three quarters of a million electron-volts.

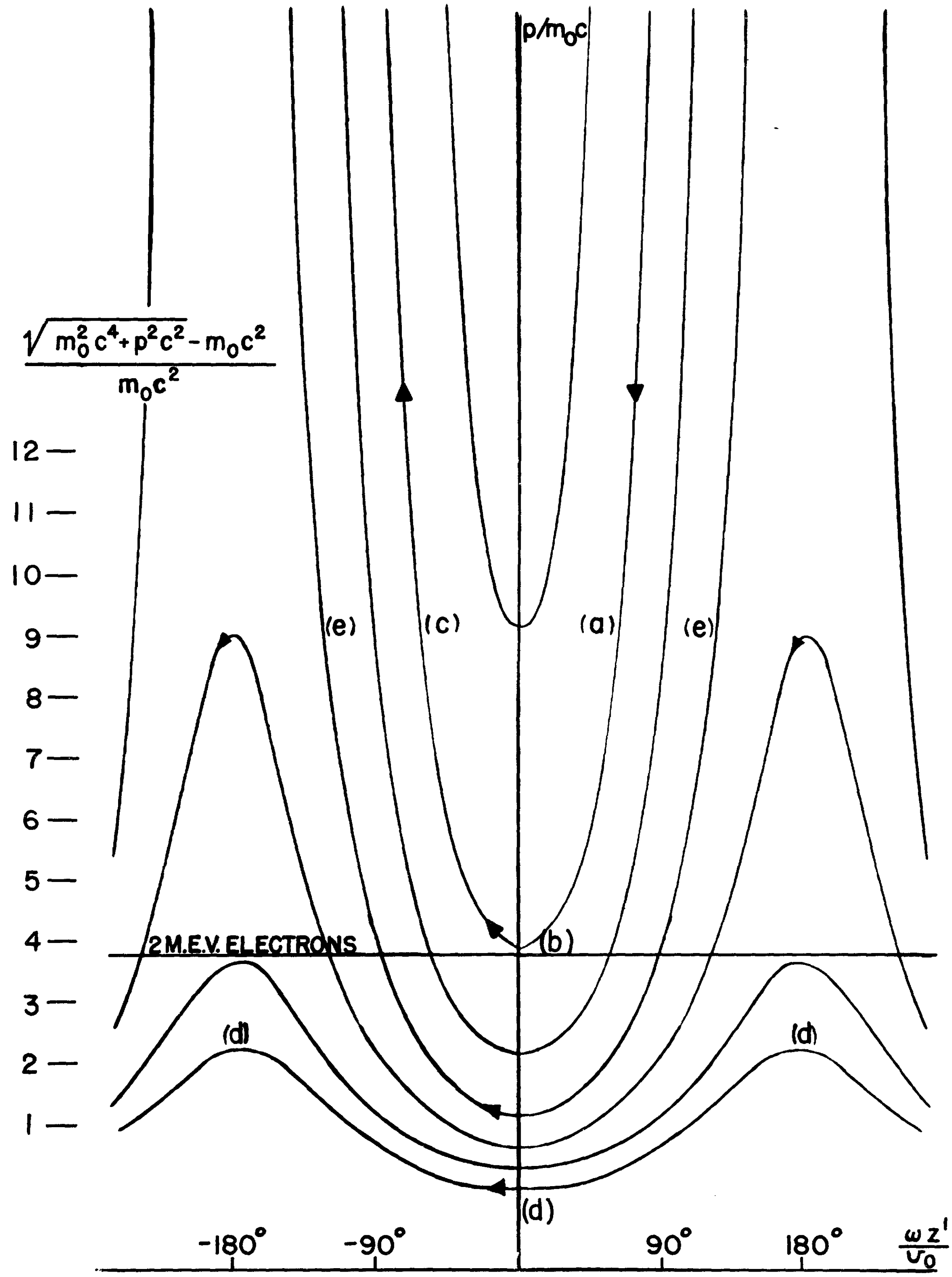

Figure 9. Phase space for $v_o = c$.

The case we have treated so far is that where the traveling wave has constant velocity. Next let us consider the case where the velocity gradually varies with position, a situation which we can secure, with our periodically loaded guide, by having the periodicity of the irises gradually change as we go along the guide, with appropriate changes of dimension so that the whole tube can be excited by a field of a given frequency. Thus we must have v_o a function of position. In place of our earlier expression, we must then assume a voltage $E_z = E \sin \omega(t - \int dz/v_o(z))$, where for simplicity we take only the case where E is independent of z. We now ask how to solve the dynamical problem of motion of particles in such a field. We cannot in this case carry through an exact solution such as that which led to the Hamiltonian (17), and to the curves of Figs. 8 and 9, but we can give a discussion equivalent to that of our pendulum approximation. Let us then start by making that approximation.

In our earlier case we measured our displacement z' of the particle relative to a moving point whose coordinate was $v_o t$, which moved with the wave; in fact, we chose our phase for the field in such a way that this was the point at which the field was always zero, so that the moving point z' = 0 represented a position of equilibrium, at which a particle could remain permanently at rest. In this case let us similarly find a moving point which represents a possible position of equilibrium, and let z' be the displacement from that point. Let $z_o(t)$ represent the displacement of this equilibrium position as a function of time, and let $p_o(t)$ be the momentum of a particle which travels along at this equilibrium position; since it now is an accelerated particle, p_o will depend on time. Evidently we shall have

$$p_o = \frac{m_o v_o}{\sqrt{1 - v_o^2/c^2}},$$

where v_o is to be computed at the point z_o which the particle has reached at time t. Then the equation of motion of the equilibrium particle is

$$\frac{dp_o}{dt} = eE \sin \omega\left(t - \int \frac{dz_o}{v_o(z_o)}\right) \quad .$$

We shall assume that the equilibrium particle is located at constant phase of the accelerating field; that is, that the velocity v_o is adjusted to correspond to the motion of a particle with constant force. Thus we must have

$$t - \int \frac{dz_o}{v_o(z_o)} = t_o \quad ,$$

where t_o is a constant. This equation defines z_o as a function of t. Furthermore, it allows us to find t_o from the known acceleration. If the spacings of irises are so arranged as to lead to a given acceleration, we find the force which would be necessary to produce this acceleration, and t_o is a constant such that $eE \sin \omega t_o$ is the necessary force. We note that this equation can be satisfied only if the assumed acceleration is less than eE, the maximum acceleration which the field can produce at the most favorable phase.

Now we consider a particle whose position is given by $z = z' + z_o$, so that z' measures its displacement from equilibrium, and $v' = dz'/dt$ the time rate of change of this displacement. We shall limit ourselves to small enough values of z' so that we may assume the velocity of the traveling wave at z to be the same as at z_o. Furthermore, we shall limit ourselves to short enough intervals of time so that the longitudinal mass of the particle and v_o will not change appreciably with time. Then the force may be written as $eE \sin \omega(t - \int dz/v_o(z)) = eE \sin \omega(t - \int dz_o/v_o(z_o) - z'/v_o) = eE \sin \omega(t_o - z/v_o)$. The time rate of change of the momentum of the particle will be

$$\frac{dp}{dt} = \frac{dp_o}{dt} + m_\ell \frac{dv'}{dt} = eE \sin\omega t_o + m_\ell \frac{dv'}{dt} .$$

Thus as our equation of motion we have

$$m_\ell \frac{dv'}{dt} = eE \left\{ \sin\omega(t_o - \frac{z'}{v_o}) - \sin\omega t_o \right\} .$$

This equation may be derived from the Hamiltonian function

$$H' = \frac{p'^2}{2m_\ell} - eE \frac{v_o}{\omega} \cos\omega(t_o - \frac{z'}{v_o}) + eEz' \sin\omega t_o \quad , \qquad (22)$$

where $p' = m_\ell v'$, and where m_ℓ is treated as a constant.

The potential energy of (22) is shown in Fig. 10, as a function of z'. It is like the cosine curve resulting from the pendulum problem, but tipped up, so as to be unsymmetrical. We can consider the motion, by the well-known method of the energy integral. We draw a horizontal line in Fig. 10, whose height is H'. The vertical distance between this curve and the potential energy gives the kinetic energy. In regions where this is positive, so that H' is greater than the potential energy, the kinetic energy is positive, and motion can occur. Thus we note that in a case like the energy H_1, we have an oscillatory motion, between points a and b, and also a motion, to the left of c, in which the particle moves to the right, reverses at c, and then moves to the left again. For a higher energy, as H_2, no oscillatory motion is possible. In other words, oscillatory motions are possible, about $z' = 0$, only for a narrow range of energies, narrower as the curve is more steeply tipped. We see from (22) that this steepness of tipping is increased, as ωt_o approaches $\pi/2$. This value is the phase angle to give the maximum acceleration to the equilibrium orbit. At that value, we find that the curve is so steep that the potential valleys have completely disappeared, and no stable orbits are possible any more.

These relations are also conveniently shown in a phase space, given in Fig. 11, similar to Figs. 8 and 9. Here we label paths of representative points connected with energies H_1 and H_2, as given in Fig. 10. The periodic character of the first case, and

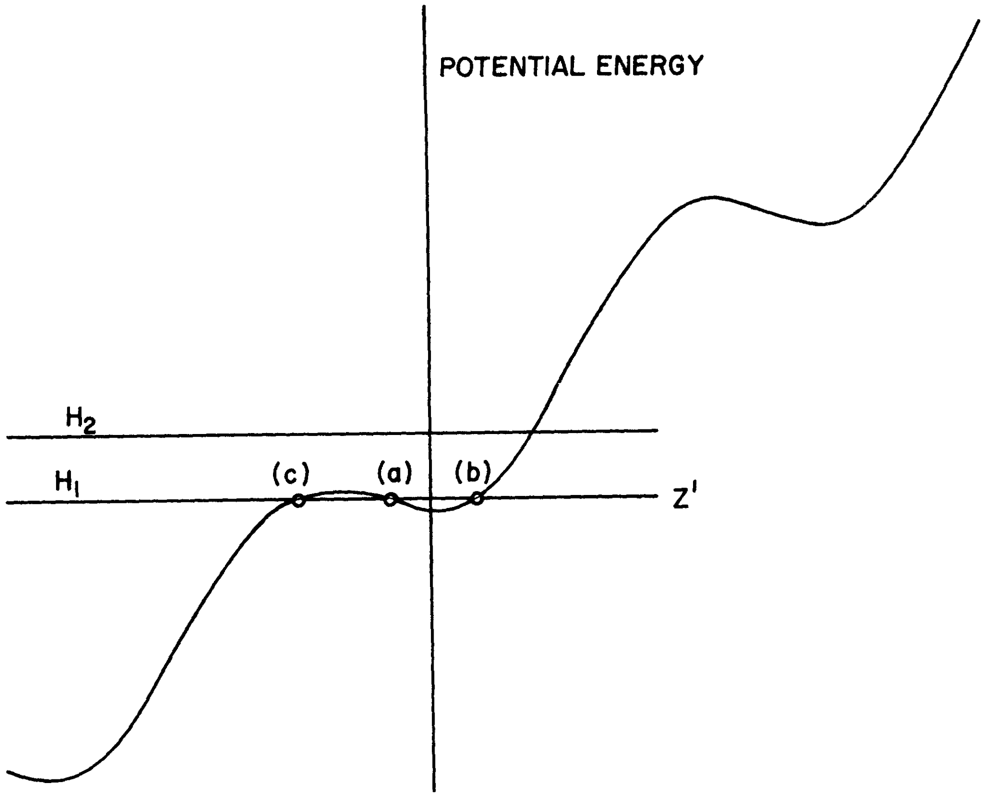

Figure 10. Potential energy as function of z', for accelerated particle.

the non-periodic character of the second, are clearly shown. Also the relation to Fig. 8 is clear. The non-periodic orbits, in the present case, do not move always forward or always backward with respect to the wave: for they are orbits in which there is no resonant phase relation between the particle and the wave, so that the particle is not accelerated by the wave, but travels, with respect to a fixed coordinate system, with a constant velocity, but with superposed periodic perturbations. If it starts traveling more rapidly than the wave, the wave will gradually catch up with it, on account of the acceleration of the wave. The wave will then go faster than the particle, and as seen from the frame of reference traveling with the wave, the particle starts with a positive velocity v' and positive momentum p', gradually stops, and then turns around and acquires negative velocity and momentum. It is clear that such a particle can never become captured by the wave, and follow along with it. The particles which execute periodic orbits, on the other hand, are effectively captured or bunched by the wave. Theytravel along with the wave, being continually accelerated, just as the equilibrium particle is, but oscillating back and forth with respect to the equilibrium particle.

This clear distinction between the particles which are captured by the wave, and resonate with it, continually gaining energy from it, and the other particles which are not captured, throws light on our assumptions of Section 1, in which we assumed that we can replace the effect of the complete field on a particle by the effect of the one resonant mode which travels with the same velocity as the particle. From our present analysis, we

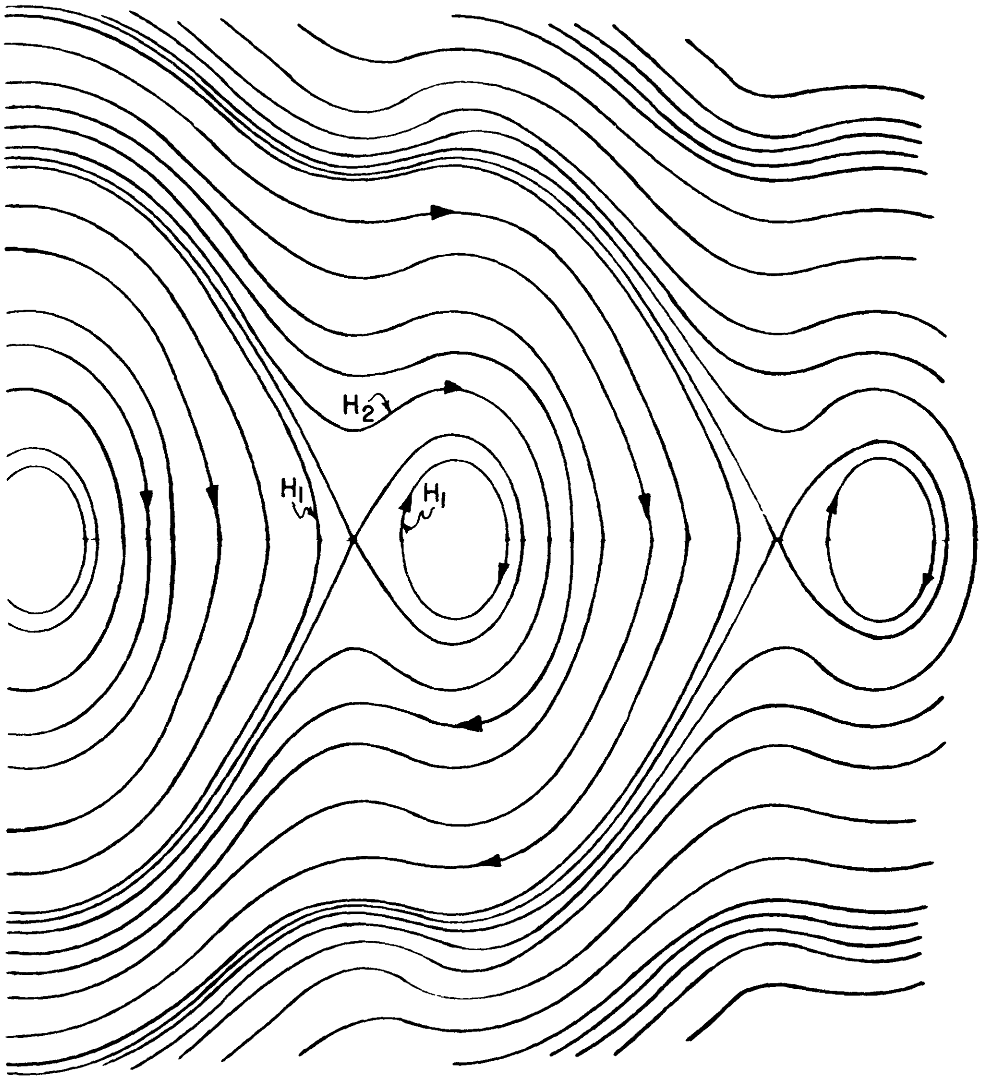

Figure 11. Phase space for accelerated particle.

see that there is a clear separation between those particles that will resonate with a wave, and those that will not. And it is quite clear that a particle which is captured by a particular Fourier component, and resonates with it, will not have any resonant effect with any other Fourier component, but only a periodic perturbation similar to that experienced by the non-periodic orbits in Fig. 11.

In the case of the traveling wave whose velocity changes with position, the various parameters of Eq. (22), the longitudinal mass, velocity v_0, and in many cases the amplitude E of the electric field as well, will vary with position, and hence with time, though this variation may be expected to be slow. It is thus important to ask how the motion of a particle will behave if there is a slow variation of these parameters. The equilibrium particle, of course, will remain at $z' = 0$, but if we have a particle oscillating about the equilibrium position with a given amplitude, it is interesting to know how its amplitude of oscillation changes with changes in the parameters. The answer to this question can be found from a familiar theorem of mechanics, which is known in the quantum theory as the adiabatic theorem. This theorem can be stated in terms of the phase space, as follows. If we have a periodic orbit, we compute its phase integral, or the area of the closed path of its representative point in phase space, which can be expressed as $\oint p\, dq$ if p is the momentum, q the coordinate, and the integral is to be extended around the path. Then if variations are made in parameters of the motion, slow enough so that the fractional change in any parameter in a period is small, the shape of the closed paths in phase space will of course change, and we might well ask which closed path in the revised phase space will actually represent the path of the particle. The adiabatic theorem states that it is that path which has the same phase integral as the original motion; in other words, the phase integral is a constant of a slowly varying motion.

This in general means that the energy will change as the parameters are changed. For instance, in a linear oscillator, the phase integral can be shown to equal the energy, divided by the frequency; thus as the mass or restoring force change, to change the frequency, the energy will change proportionally to the frequency. Now the energy of an oscillator, whose displacement is $A \cos\omega_0 t$, is the maximum value of the kinetic energy, or $\frac{1}{2} mA^2\omega_0^{\;2}$. Since this is proportional to ω_0, we see that A is proportional to $1/\sqrt{m\omega_0}$. This result can be used to discuss the variation of amplitude of small oscillations in the linear accelerator. The frequency of oscillation derived from the Hamiltonian (22), similar to the expression (19), is

$$\omega_0 = \sqrt{\frac{e}{m_\ell} E \frac{\omega}{v_0} \cos\omega t_0} \quad . \tag{23}$$

The mass that comes into the dynamical problem is the longitudinal mass m_ℓ. Thus we see that the amplitude of oscillation is given by

$$A \sim (v_0/m_\ell E\omega \cos \omega t_0)^{1/4} \quad .$$

We see that as the velocity approaches the velocity of light, and the longitudinal mass becomes infinite, the amplitude of oscillation decreases to zero. This can also be seen directly from the phase space: we have seen that as we approach the velocity of light, the oval paths of representative points become lengthened vertically; thus a given particle must go into an orbit of smaller horizontal width, or smaller amplitude, to keep the area constant. In a similar way, as the accelerating field E increases, the amplitude of oscillation decreases. These relations, as we shall see in the next section, may be used to produce very sharp bunching of the particles around the equilibrium position.

<u>13. Application of Electron Dynamics to Different Types of Accelerators</u>. In the preceding section we have investigated the dynamics of longitudinal bunching of particles in a traveling sinusoidal field, both in the case where the field travels with constant velocity, and when its velocity slowly varies with position. Now that we have acquired this theoretical background, we shall discuss its application to the various types of linear accelerators which have been contemplated. The simplest case is that of the M.I.T. acceler-

ator. In this case we are accelerating electrons, and are injecting them from a Van der Graff generator with two million electron-volts initial energy. The accelerating tube is designed to have a phase velocity exactly that of light. Thus the case is just that of Fig. 9. Electrons are injected with all possible phases. Thus if we draw a horizontal line in Fig. 9, at a height corresponding to two million electron-volts, we shall have a uniform distribution of particles in phase at the injection end. We see from Fig. 9 that something more than half of these phases will lead to orbits which become bound to the wave, and continue indefinitely to gain energy. There will be a particular concentration of electrons in orbits in the neighborhood of that marked by a, b, c in Fig. 9, which is tangent to the horizontal line at height of two million electron-volts. This orbit reaches an asymptotic phase of approximately 90°, and hence eventually will receive a maximum possible energy. Thus with this particular injection energy, corresponding to the accelerating field E which we have chosen, there will be a strong concentration of accelerated electrons in the neighborhood of the maximum possible acceleration. A study of the probable energy spectrum of the emerging electrons has been made by Mr. S. J. Mason, and he finds a rather narrow peak which should contain most of the electrons. We note that if E were smaller, the optimum injection voltage would be correspondingly greater.

A number of the other electron acceleration projects, including the Stanford and G.E. projects, inject their electrons at much lower energies, and use a section of tube with gradually increasing velocity to speed the electrons up to the velocity of light. Some work on such an injection method has been carried on at M.I.T. as well. With the order of magnitude of accelerating field which we are using, and velocities well below the velocity of light, we find from (23) that the angular frequency of oscillation of electrons with respect to the bunch is not greatly less than the angular frequency of the r-f field. Furthermore, the acceleration is so rapid that we are not really justified in our assumptions that the change in relativistic mass over a period of the oscillation is small, or that the parameters involved change slowly with time. Nevertheless, we may assume that our derivation using the adiabatic theorem is not entirely invalid. Suppose we inject with all phases into a tube with varying velocity, so that its phase space is shown approximately by Fig. 11. Some phases will result in electrons which are caught by the wave, other phases in electrons which are lost; it should not be hard to capture at least 50 per cent of the electrons. In fact, if we were to inject at just the velocity of the wave, or at $p' = 0$, we should capture all electrons. Then as the velocity increases, using the adiabatic theorem, we expect the amplitudes of the electron oscillations to decrease, so that the bunches will contract around the phases of the wave which correspond to just enough acceleration to travel with the same velocity as the wave. As we approach the velocity of light, the bunches should narrow down, and in the limit we should approach a situation much like that of Fig. 9, which we have already discussed. The failureof the adiabatic theorem and other approximations, on account of the rapid acceleration, would probably result in a loss of a certain number of the electrons, and a diffuseness in phase of the final electrons which become captured in the wave traveling with the velocity of light, but nevertheless we should expect a large fraction of the injected electrons to come through with high energy. This prediction is in accordance with the experience of those projects which have accelerated electrons in such tubes.

The problem of accelerating positive ions is much more easily susceptible of analysis, for there, on account of the large rest mass, the acceleration must necessarily be very slow, and our adiabatic theorem and other approximations are true to a high degree of accuracy. Thus suppose we inject at quite slow velocity, in a phase space like Fig. 11. The frequency of oscillation is much smaller than the frequency of the traveling wave, but nevertheless the change of energy of the particle per period of oscillation is much smaller than in the electron case; we find as a matter of fact, by a simple calculation, that the energy picked up by an ion accelerated at the stable position of the wave, per period, is, except for a numerical constant, the same as the kinetic energy change given in (21), though that was calculated for a different problem. This energy is greater for an ion than for an electron, in the ratio of the square root of the masses, on account of the fact that the period of oscillation, and hence the distance traveled in the accelerating field, is greater in this ratio; but this increase of energy forms a smaller fraction of the total energy, in the inverse ratio of the square root of the masses.

It might well be wise, with positive-ion acceleration, to inject with a weak accelerating field E. Then E could be gradually increased, and as we have seen this would result in a contraction of the ion bunches, around the stable positions in the phase space (that is, around the phases where the accelerating field is just enough to speed the particles up enough to keep them in step with the acceleration of the traveling wave). In this way practically all the ions could be bound in the stable bunches. They would accelerate with the wave, having at every point the velocity characteristic of the wave. As the velocity approaches that of light, the bunches will become further narrowed, on account of the increasing longitudinal mass. Finally, at extremely high energies, it should be possible to join onto a section of guide whose velocity was equal to that of light, with a phase space as shown in Fig. 9, and with the ions in small enough bunches so that practically all ions would have maximum acceleration in this guide. This suggests a definite advantage in starting the particles at less than the velocity of light: it introduces a bunching effect, which allows the bunches to be so phased that practically all particles receive maximum acceleration, instead of the situation described earlier for the M.I.T. accelerator, where the electrons are fed without bunching into a section with $v_o = c$. It is possible for this reason that an eventually perfected electron accelerator may wish to use such a bunching feed. Nevertheless the difficulties of defocussing which would be introduced by such a feed, and which will be described in the next section, are so great that it has seemed wise to avoid them in the M.I.T. accelerator by using the Van de Graaff generator as injector.

<u>14. The Dynamics of Transverse Motion and Focussing</u>. In the preceding sections we have studied the dynamics of longitudinal motion of particles along the axis of the accelerator, and have found that particles moving with less than the velocity of light have a tendency to form stable bunches surrounding stable points in the field, points where the force on a particle is just enough to accelerate it to keep step with the field, and in which the accelerating field is increasing with time. The frequency of oscillation of a particle in

such a bunch around its position of equilibrium reduces to zero as the wave and particle approach the velocity of light, and in the limit of waves traveling with the speed of light, the bunching phenomenon disappears; but a bunch already formed at a lower velocity will persist as the particles are accelerated to the speed of light. Now we consider the transverse motions of the particle, and consequent focussing or defocussing effects; and we make the disconcerting discovery that the phase of the stable bunch is inherently unstable with respect to transverse motion, so that it is inherently defocussing.

We might have forseen this result from general principles. For if the wave is traveling more slowly than the velocity of light, we can make a Lorentz transformation and transform it to rest. It then forms a static solution of Maxwell's equations. The stable particle similarly transforms to rest, and the oscillatory particles to particles executing oscillations about this equilibrium position. This Lorentz transformation would have accomplished the same thing for us that our Newtonian transformation to moving axes did, but it is a little more complicated to carry out, and for that reason we have used the Newtonian transformation in our discussion of longitudinal motion. Once we get our transformed static problem, however, we meet Earnshaw's theorem, which states that an electrostatic potential cannot have a maximum or minimum in empty space, but at most a saddle point. If the field leads to stable equilibrium for motion in one direction, the equilibrium must be unstable in directions at right angles to this. Since we have longitudinal stability, or stable bunches, we cannot have stable focussing. Or conversely if we choose the phase to give stable focussing, there is bunching instability: this corresponds to the position of unstable equilibrium which we have discussed in connection with the longitudinal motion. On the other hand, as we approach the velocity of light, the bunching stability disappears, and is replaced by a neutral equilibrium. We saw this in our earlier treatment by the vanishing of the frequency of oscillation in this limit. If we carry out the Lorentz transformation, it shows itself in the wavelength of the radio-frequency field becoming infinite as we go to infinite velocity, so that the accelerating force becomes independent of position, and in any finite neighborhood we have a constant field. Correspondingly the defocussing effect vanishes at the velocity of light, and we have a neutral situation as regards focussing or defocussing.

These results, which could be proved by relativistic methods, can also be demonstrated by elementary and straightforward calculation, which is considerably simpler. Let us write down the expressionsfor the electric and magnetic fields in our traveling wave which resonates with the particle, in our ordinary fixed frame of reference. We can easily show that the components are

$$E_z = E \sin \omega(t - z/v_o) \left[J_o(w)\right]$$

$$E_r = E \cos \omega(t - z/v_o) \frac{(-jJ_1(w))}{\sqrt{1 - v_o^2/c^2}}$$

$$B_\Theta = E \cos \omega(t - z/v_o) \left\{ \frac{v_o}{c^2} \frac{(-jJ_1(w))}{\sqrt{1 - v_o^2/c^2}} \right\}$$

where

$$w = \frac{j\omega r}{v_o} \sqrt{1 - v_o^2/c^2} \ .$$

Now the r component of force on a particle of charge e, moving along the z axis with a velocity v_o (we neglect in the defocussing problem the slight difference between the velocity of the particle and the velocity of the wave resulting from the oscillation) is the r component of $e(E + v \times B)$, or is $e(E_r - v_o B_\Theta)$. This is

$$F_r = eE \cos \omega(t - z/v_o) \left\{ \sqrt{1 - v_o^2/c^2} \quad (-jJ_1(w)) \right\}.$$

In other words, the magnetic force is $-v_o^2/c^2$ times the electric force, so that in the limit as the velocity approaches the velocity of light any transverse force reduces to zero, and focussing and defocussing effects disappear.

We may now use these results to discuss the transverse force F_r for small displacements of the particle from the axis. The Bessel function $J_o(w)$ approaches 1 for small values of its argument, and $J_1(w)$ approaches $w/2$. Thus for small r we have

$$F_z = eE \sin \omega(t - z/v_o)$$

$$F_r = \frac{\omega}{2v_o} \left(1 - \frac{v_o^2}{c^2}\right) eE \cos \omega(t - z/v_o) \qquad (24)$$

We are interested in F_r particularly at the phase of the wave where we can have stable bunching. This is the region where F_z is positive and increasing with time; that is where $\sin\omega(t - z/v_o)$ is positive, and its time derivative $\omega\cos\omega(t - z/v_o)$ is also positive. Thus we see that under these circumstances F_r is a positive force, proportional to r, tending to drive particles away from the axis, or to defocus them.

The motion of a particle under the action of a force repelling it from a point proportional to the distance is an exponential function of the time. If we set up the equation of motion of a particle in the force F_r above, remembering that we are now dealing with transverse motion and so must use the transverse mass $m_o/\sqrt{1 - v_o^2/c^2}$, we find that this combines with the factor $1 - v_o^2/c^2$ to give the longitudinal mass, and that we find the motion to be given by

$$r = r_o e^{t/T}, \quad \frac{1}{T} = \sqrt{\frac{eE\omega}{2m_o v_o} (1 - v_o^2/c^2)^{3/2} (\cos \omega t_o)}, \qquad (25)$$

where $t_0 = t - z/v_0$. We find this reciprocal time to be $1/\sqrt{2}$ times the angular frequency of oscillation about the stable position, given in (23). Hence we conclude that the time required for a given value of r to increase in the ratio e is of the order of magnitude of a period of the oscillation about the stable position. For electrons in the M.I.T. accelerator, we have no problem, because the period of oscillation is essentially infinite; straightforward study of the transverse motion shows that electrons injected at one side of the axis, with two million electron-volts energy, will spread to only about twice their initial distance from the axis in reaching any arbitrarily high energy. In electron acceleration from a low velocity, the problem still is not very serious. With the high accelerations that we have here, we have already seen that it requires only a relatively small number of oscillations in the bunch for the electrons to pick up enough energy to reach substantially the velocity of light: thus, though they may multiply their initial r by e a number of times, the resulting factor need not be unreasonably large, and need not result in an impossible spreading of the beam, if we start with a very sharp and well-focussed beam at injection. Furthermore, it is easy to focus electrons, in this initial region of r-f defocussing, by the action of a longitudinal magnetic field, as we shall discuss shortly.

For positive ions, however, the defocussing is extremely serious, and may very possibly prove a fatal objection to the use of linear accelerators for positive ions. We have seen that it will require a great many periods of oscillation for a positive ion to pick up enough energy to reach approximately the velocity of light, and in this time the radius of the beam will be multiplied by e, a great many times, or will spread by an entirely inadmissible amount. There is, it is true, one focussing feature which we have passed over by our method of analysis, and which may conceivably make the result not quite so disastrous as it seems at first sight. We may understand this feature by comparison with the theory of the cyclotron. There the ions in the accelerating region between the dees find themselves in a field as in Fig. 12. As they enter the region, they are focussed, and as they leave they are defocussed. The defocussing can be less than, or

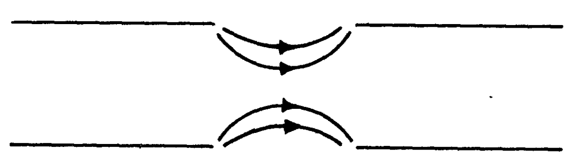

Figure 12. Lines of force in a gap.

greater than, the focussing for two reasons. First, the field may be increasing with time as the ions go through the region. Then the field will be stronger as they leave than as they enter, and the net effect will be defocussing. This is the effect we have so far been considering. But the second reason is that as the ions go through the gap they will be speeded up. Thus they will be going faster when they go through the defocussing field than when they go through the focussing field, and this will make the focussing more effective than the defocussing. We have missed this effect in our analysis by neglecting second-order effects of change of particle velocity during the acceleration.

As a result of this effect, there is a small range of phase when there can be both bunching and focussing. This is the range where the accelerating field has almost reached its maximum, so that cos ωt_0 has almost reduced to zero. The angular frequency of oscillation in the bunches given in (23) is still real, though small, so that there is still slight stability in the bunches. At the same time, the quantity $1/T$ of (25) has also almost reduced to zero, so that the defocussing on account of the outward component of r-f field is very small. This may be small enough so that it is counteracted by the other effect which we have just been describing, with a result that there is net focussing. This effect is discussed in the TRE reports, and it is there shown that there is a phase region of a few degrees where this situation holds. Under some circumstances, with the field concentrated in very narrow gaps, the favorable phase can be considerable, and it seems likely that the success of the early linear accelerators resulted from using this favorable region of phase. It seems not impossible that the final success of positive-ion linear accelerators may depend on exploiting this focussing feature.

The other possible method which has been suggested for counteracting the defocussing in a positive-ion accelerator is the use of thin metallic foils for grids, as proposed by the Berkeley group. If we could place a grid over the exit to an accelerating gap, as shown in Fig. 13, it is clear that we should preserve the focussing, and

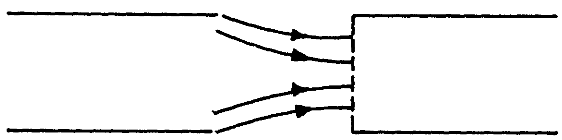

Figure 13. Lines of force in a gap with focussing grid or foil.

completely eliminate the defocussing. This does not contradict our analysis which has led to an inherent defocussing. That analysis was based on the resolution of the field along the axis into Fourier components, based on the assumption that it satisfied Maxwell's equations in that region, with no charge density. On the other hand, if there are grids, they can carry charge, and the potential can effectively obey Poisson's rather than Laplace's equation. This allows a real minimum of potential energy, and allows stable bunching and stable focussing at the same time. The problem is an experimental one, of realizing the grids in practice. The Berkeley project proposes to use thin metallic foils, so thin that the ions can pass through them with relatively little scattering. If they can be made mechanically strong enough, and at the same time with little enough scattering, there is no apparent electrical reason why they should not work.

For electron focussing with velocities less than the velocity of light, we have already mentioned the possibility of a longitudinal magnetic field, which would have the effect of bending the orbits into helices. Let us consider the effect of such a magnetic field. We can solve the problem completely by setting up the Lagrangian equations for a particle, with the forces F_z and F_r already considered, and in addition the forces resulting from a constant magnetic induction B along the z axis. The effect on B on the radial motion proves to be to add an effective potential energy term proportional to r^2, or an effective restoring force proportional to r. If the magnitude of this term is great enough to balance the repulsive force F_r, which is also proportional to r, there will be stable orbits, the stable helical paths of the electrons. Rather than carrying through this whole analysis, we adopt a simple discussion which will lead to a value of the critical magnetic field for focussing.

In the absence of the force F_r, a particle in the presence of a magnetic field B will execute circular motion in the plane at right angles to the field. In a familiar way, we may equate the force driving the particle toward the axis, when it moveswith angular velocity $\dot{\Theta}$ in a circle of radius r, to its transverse mass times its centripetal acceleration. That is, we have $eBr\dot{\Theta} = m_t\dot{\Theta}^2r$, from which at once $\dot{\Theta} = eB/m_t$, the Larmor angular frequency. On the other hand, if we have an additional force kr tending to increase r (our force F_r), the equation will become $eBr\dot{\Theta} - kr = m_t\dot{\Theta}^2r$. This gives a quadratic equation for $\dot{\Theta}$, whose solutions are

$$\dot{\Theta} = \frac{eB}{2m_t} \pm \sqrt{\left(\frac{eB}{2m_t}\right)^2 - \frac{k}{m_t}} \quad .$$

This will represent stable rotation in a circle, provided $\dot{\Theta}$ is real. This gives a limiting case

$$\frac{eB}{m_t} = \sqrt{\frac{4k}{m_t}} \quad .$$

If we take k from our formula (24) for F_r, and use (23), we have

$$\frac{eB}{m_t} = \sqrt{2}\,\omega_0 \ . \qquad (26)$$

In other words, to counteract the defocussing, we must use a longitudinal magnetic field strong enough so that the resulting Larmor frequency is a least $\sqrt{2}$ times the frequency of oscillation in the bunches. Any additional magnetic field will result in positive stability.

We now may examine the magnitudes of the required magnetic fields. For electrons, we have seen that the frequency ω_0 may approach the frequency ω of the r-f field for quite high accelerations, although in practical cases it is likely to be considerably less. Thus B will be less than that required to produce a Larmor frequency equal to the r-f frequency. With the r-f frequencies contemplated, this leads to a magnetic field of a few hundred or at most a few thousand gauss, which therefore is in the range that can be conveniently set up in practice, particularly since it needs to be used only in the part of the accelerator where the electrons have appreciably less velocity than the velocity of light. With positive ions, on the other hand, the required magnetic fields become quite impossible. The frequencies ω_0, it is true, are less for positive ions than for electrons in the ratio of the square root of the mass of the electron divided by the mass of the ion, but the additional factor of the transverse mass in (26) results in the magnetic field itself being greater for a positive ion than for an electron in the ratio of the square root of the ratio of mass of electron to mass of ion. Thus instead of perhaps a thousand gauss, we should require magnetic fields of perhaps forty or fifty thousand gauss for overcoming the defocussing effect with positive ions. Furthermore, this would have to be applied to the whole length of the accelerator in which the velocity of the ions was appreciably less than the velocity of light. While magnetic fields of this magnitude can of course be obtained in limited regions, the problem of creating such a field over the whole of a very large accelerator would be prohibitively difficult, and would certainly make the positive-ion linear accelerator a more elaborate and expensive machinethan the proton synchrotron.

There is one additional problem of electron focussing which we should consider. Even though there is no defocussing force acting on an electron traveling with the velocity of light, still the injected beam of electrons in such an accelerator as the M.I.T. one will have a certain spread of directions of the electrons. We may well ask, will the injected beam have to have such a small spread that all the incident electrons are aimed to hit the final iris aperture at the far end of the tube? This would correspond to a very small solid angle, and it would be almost impossible to concentrate an incident beam to such an extent. Fortunately this is not necessary. To see this point, we may think about the problem in a very simple way. An incident electron whose direction is not quite along the axis will have two components to its momentum: one along the axis, p_z, and the other along the radius, p_r. In the limiting case we are considering, there will be no radial force, and p_r will remain constant during the subsequent motion of the electron, whereas the longitudinal momentum p_z will increase proportionally to the time, on account of the constant accelerating force in that direction. As a result, the resultant momentum vector, and the velocity vector which is parallel to it, will turn more and more along the z axis as the particle picks up momentum. The problem is like that of a falling body, falling along the z axis, subject to a constant accelerating force in that direction, but started with a component of momentum at right angles. The falling body would fall in a parabola, and its horizontal motion in the time of fall would equal its initial velocity in that direction, times the time of fall. Here, however, on account of the relativistic nature of the motion, we may not infer from the constancy of the transverse momentum that there is also a constant transverse velocity; as a matter of fact, we shall find that the transverse velocity decreases along the path, so that the net transverse motion is much less than we should find in the non-relativistic case.

We may handle the problem in an elementary way. The energy of the particle will be eEz, where z is measured from an assumed zero at which the energy would be zero, if the accelerating force were uniform the whole way. Thus the momentum is eEz/c, in the relativistic range. This momentum is p_z; p_r is constant. Now the actual path of the particle will be parallel to the momentum; thus we have $dr/dz = p_r/p_z = p_r c/eEz$. This shows us at once that, since $dz = c\,dt$ in the region where the velocity equals the velocity of light, the radial velocity decreases inversely proportionally to z. Integrating the equation above, we have

$$r_2 - r_1 = \frac{p_r c}{eE} \ln z_2/z_1 \quad .$$

This gives us the change in radius from z_1, where the particle is injected, to z_2, where it emerges. We can express it in more convenient form by using the angle ϕ_o between the incident direction of the electron, and the axis of the tube. Then we have ϕ_o equal to the ratio of p_r to the value of p_z when $z = z_1$. This leads to

$$r_2 - r_1 = \phi_o z_1 \ln \frac{z_2}{z_1} \quad .$$

When we put numbers into the equation above, we find very small values for the spreading, even for a very long accelerator. Thus let us assume that we inject at two million electron-volts, and that this corresponds to a length $z_1 = 60$ cm. If we accelerate to two billion electron-volts, we have $z_2/z_1 = 10^3$. Let us assume that $\phi_o = 10^{-3}$, corresponding to a collimation through a 1-mm aperture at a meter's distance, a collimation which should be possible with a well-adjusted Van de Graaff beam. Then we find $r_2 - r_1 = 4$ mm. Even with an accelerator ten times as long, the displacement is only increased by a factor of 4/3, to 5.3 mm. Thus we may expect that a beam can be collimated well enough so that it will not spread to be too large for the size of the iris holes, for as long an accelerator as we are likely to contemplate.

References

1. J. Schwinger, "Electron Radiation in High Energy Accelerators", Phys. Rev., 70,797 (1946).

   N. H. Frank, "Stability of Orbits in the Synchrotron", Phys. Rev., 70, 117 (1946).

2. E. M. McMillan, "Resonance Acceleration of Charged Particles", Phys. Rev., 70, 800 (1946).

3. E. O. Lawrence and D. H. Sloan, "Production of High-Speed Canal Rays without the Use of High Voltages", Proc. Nat. Acad. Sci., 17, 64 (1931).

   D. H. Sloan and E. O. Lawrence, Phys. Rev., 38, 2021 (1931).

   D. H. Sloan and W. M. Coates, Phys. Rev., 46, 539 (1934).

4. J. B. Fisk, H. D. Hagstrum, and P. L. Hartman, "The Magnetron as a Generator of Centimeter Waves", Bell Sys. Tech. J., 25, 167 (1946).

5. L. W. Alvarez, "The Design of a Proton Linear Accelerator", Phys. Rev., 70, 799 (1946).

6. J. Halpern, E. Everhart, R. A. Rapuano, and J. C. Slater, "Microwave Linear Accelerators", Phys. Rev. 69, 688 (1946).

   J. C. Slater, "The Design of Linear Accelerators", Phys. Rev., 70, 799 (1946).

   Quarterly Progress Reports of the Research Laboratory of Electronics, M.I.T.

7. Reports of the Linear Accelerator Project, Stanford University.

8. R. O. Haxby, E. S. Akeley, A. Ginzbarg, R. N. Smith, H. W. Welch, and R. M. Whaley, "Preliminary Studies on the Purdue Microwave Linear Accelerator", Phys. Rev. 70, 797 (1946).

9. H. L. Schultz, R. Beringer, C. L. Clarke, J. A. Lockwood, R. L. McCarthy, C. G. Montgomery, P. J. Rice, and W. W. Watson, "Cavity Accelerator for Electrons", Phys. Rev., 72, 346 (1947).

10. T.R.E. Reports on Linear Accelerators.

11. J. C. Slater, "Microwave Electronics", Rev. Mod. Phys., 18, 441 (1946).

12. J. E. Evans, R. C. Fletcher, and F. F. Rieke", R-F Phasing of Pulsed Magnetrons", Rad. Lab. Report No. 1051, Feb. 6, 1946.