INTRODUCTION TO RF LINEAR ACCELERATORS

Mario Weiss
CERN, Geneva, Switzerland

ABSTRACT

The basic features of RF linear accelerators are described. The concept of the 'loaded cavity', essential for the synchronism wave-particle, is introduced, and formulae describing the action of electromagnetic fields on the beam are given. The treatment of intense beams is mentioned, and various existing linear accelerators are presented as examples.

1 INTRODUCTION

Under the name 'linear accelerators', a great number of devices can be included, having in common the acceleration of particles on a linear path. If we only consider devices which operate with time-varying electric fields, all the d.c. devices are excluded. In fact, we wish to be even more restrictive and only consider devices which operate with sinusoidally varying electromagnetic fields. Such devices are very important and are called RF linear accelerators or RF linacs. These operate either as waveguides or resonators. We shall concentrate our study on them, leaving out devices such as induction linacs which operate in a pulsed manner: a current pulse passes through a magnetic torroid, producing a rapidly changing magnetic field, which by Faraday's law induces a voltage across an accelerating gap.

RF linear accelerators usually operate in the frequency range from 100 MHz up to several GHz. They produce intense beams of very good quality and are applied to scientific research (elementary particle physics), as well as industry or medicine. They can be used at lower energies as injectors into synchrotrons (typically several tens or hundred MeV), or at high energies, e.g. as colliders, accelerating electrons and positrons to 50 GeV, as at the Stanford Linear Accelerator Center (SLAC).

The history of linear accelerators starts with G. Ising, who in 1924 proposed a system of sequentially-pulsed drift tubes, where the particles would be accelerated in gaps between them. In 1928, R. Wideröe suggested the use of RF voltage between successive drift tubes, and successfully tested such a device. He was followed in 1931 by D.H. Sloan and E.O. Lawrence who constructed an improved version of Wideröe's accelerator. It was, however, not until the end of the Second World War that the development of linear accelerators really started. The availability of RF power sources in the MW range and at frequencies of hundreds of MHz, developed for radar purposes, was a great asset. In 1946, L. Alvarez built in Berkeley his famous proton drift tube linac of 32 MeV, operating at 200 MHz. At the same time the development of a 3 GHz electron linear accelerator at SLAC was under way. Since then, many linear accelerator structures have been developed and built, and

the theory and practice have progressed considerably, hand in hand, and are still doing so nowadays.

This paper has been conceived as an introduction to RF linear accelerators. Emphasis has been given to the explanation of the basic principles of operation. It was felt that by understanding these principles, different types of linear accelerators could be treated in a common way.

Sections 2 and 3 are essential for understanding the principles of accelerators with RF electromagnetic fields. For clarity the accelerating structures have been treated as lossless. The treatment of lossy structures is more complicated, but is not basically different. Section 4 divides the RF linear accelerators into two classes: travelling- and standing-wave linacs. Section 5 explains the method by which the action of the electromagnetic field on moving particles is computed. Here one could have gone into more detail, but it was feared that the guiding line of the paper would be interrupted too much. Section 6 exposes the principle of phase stability and derives in a simplified way the equations of the synchrotron motion. It was also felt that is was necessary to mention how intense beams must be handled in accelerators. This is done in Section 7, where this important subject, still in full evolution, is briefly presented. Section 8 reviews some existing linear accelerator structures and completes, where necessary, the theory of their operation.

2 ELECTROMAGNETIC WAVES AND CAVITIES

2.1 General considerations

Electromagnetic waves in free space are of the transverse electromagnetic (TEM) type: both the electric and magnetic field vectors are perpendicular to the direction of propagation of the wave [1]. In bounded media such as waveguides or resonators, where the boundary is a conductor, such a wave type is no longer possible because boundary conditions cannot be satisfied. A good physical insight into the problem is obtained by assuming the boundaries to be perfect conductors where the tangential component of the electric field, E_t, and the normal component of the magnetic field, B_n, must be zero. The second condition can also be stated as

$$\frac{\partial B_t}{\partial n} = 0 \ ,$$

i.e. the normal derivative of the tangential magnetic field at the boundary is zero. In general, boundary conditions are specified either by the value of a field component (Dirichlet condition), or by the derivative of a field component along the normal to the boundary (Neumann condition).

Boundary conditions can be met if the wave has one field component in the direction of propagation. This can be either the electric field (TM wave) or the magnetic field (TE wave). The wave propagation along the waveguide can be explained in terms of its reflection from wall to wall; in this way one gets a field component in the direction of propagation, as will be shown below.

A wave with an oblique incidence on a conducting plane is shown in Fig. 1. The incident and reflected wave combine in such a way as to annul E_t or B_n.

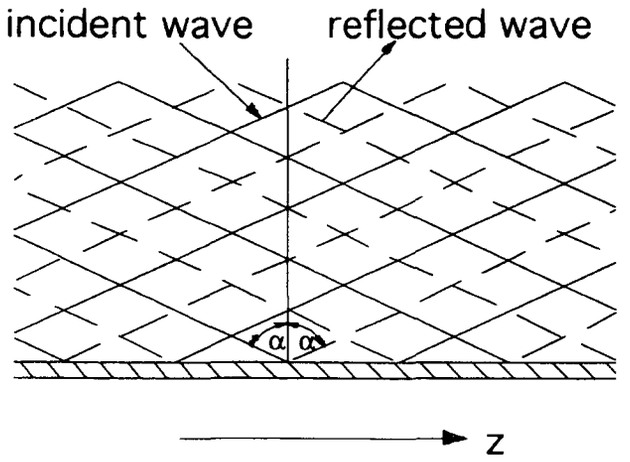

Fig. 1 Oblique wave incidence on a perfectly conducting plane

Observing the incident wave in more detail (Fig. 2), one sees interesting things. In the direction of propagation (z-direction), the distance between two adjacent wave crests, λ_p, is longer than the actual wavelength λ. This means that the wave phenomenon moves in the z-direction with a speed greater than the velocity of light c:

$$v_{ph} = \frac{\lambda_p}{\lambda}\, c = \frac{c}{\sin\alpha} > c\ .$$

The velocity v_{ph} is called the phase velocity. The energy propagates with another velocity, the group velocity v_g, which is according to Fig. 2:

$$v_g = c \sin\alpha\ .$$

In this particular case

$$v_{ph} \cdot v_g = c^2\ .$$

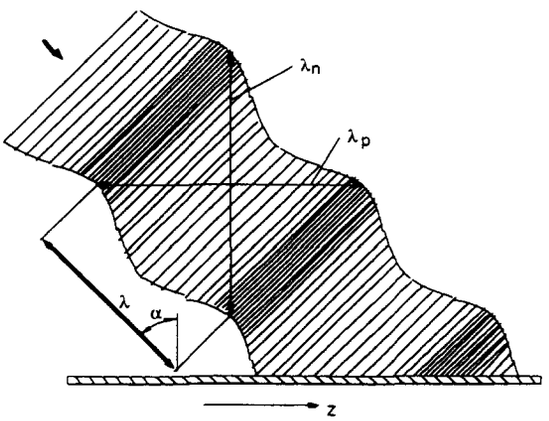

Fig. 2 Oblique incidence of a wave with wavelength λ; λ_p and λ_n are the wavelengths parallel and perpendicular to the boundary, respectively

With one conducting wall as in Fig. 2, the angle of incidence of the wave can have any value. With two walls, as in Fig. 3, one is more limited, and only certain angles are

allowed if boundary conditions are to be satisfied at both walls. For a given wave type (TE or TM) one can therefore have only certain wave modes which present an integer number of half wavelengths between the walls, see Fig. 3. The wave phenomenon in the longitudinal or z-direction is a travelling wave with the wavelength λ_p, and in the transverse or x-direction, between the walls, one has a standing wave with the wavelength λ_n.

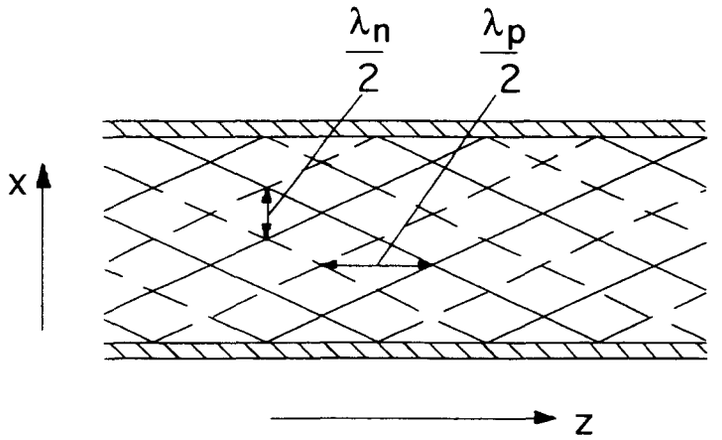

Fig. 3 Waves between two perfectly conducting planes; λ_p is the wavelength of the travelling wave; λ_n is the wavelength of the standing wave

A waveguide of rectangular cross-section has four walls, as shown in Fig. 4. The number of half wavelengths between two parallel planes determine the wave mode, which is written as

$$\mathrm{TE_{mn}} \text{ or } \mathrm{TM_{mn}} ,$$

m and n being the number of half waves in the x and y direction, respectively. In the case of a resonator, a third subscript is added, indicating the number of half wavelengths in the z-direction. These modes, which describe the configuration of electromagnetic fields in a cavity, are called *cavity* modes.

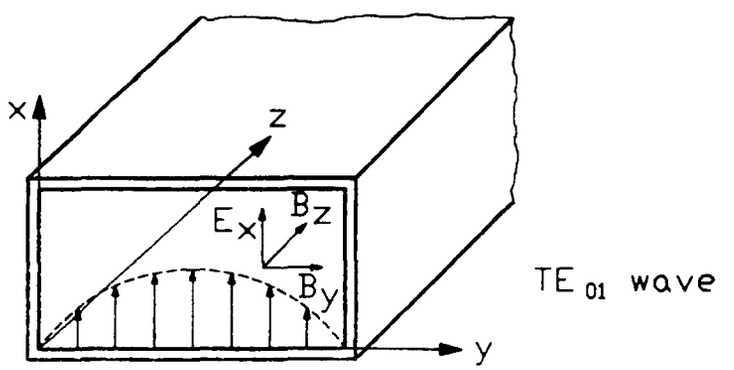

Fig. 4 Rectangular waveguide (schematic)

2.2 Synchronism between wave and particle

If we wish to accelerate charged particles, we need a longitudinal electric field, as the waves of the TM type have (we shall see that 'distorted' TE waves can also be used). In addition, we need to maintain a synchronism between the wave and particle, which means that the wave velocity (phase velocity v_{ph}), and the particle velocity v_p must be equal. This is impossible in the uniform waveguide as we have seen that $v_{ph} > c$.

In order to use electromagnetic waves for the acceleration of particles one must first find means to slow the waves down.

3 WAVE EQUATION AND SLOWING DOWN OF WAVES

3.1 Waves in empty cavities

The process of slowing down of waves is best understood with the help of the wave equation. This is a partial differential equation of the second order, derived from Maxwell's first-order equations. In Cartesian coordinates we have for the electric field vector

$$\left(\underbrace{\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}}_{\text{space}} - \underbrace{\frac{1}{c^2}\frac{\partial^2}{\partial t^2}}_{\text{time}} \right) \vec{E} = 0 \; . \tag{1}$$

The wave equation applies to each field component, but we are mostly interested in the longitudinal field E_z. The solution of the wave equation is usually expressed as the product of functions of one variable:

$$\mathrm{E}_z(x,y,z,t) = \mathrm{X}(x)\ \mathrm{Y}(y)\ \mathrm{Z}(z)\ \mathrm{T}(t) \; . \tag{2}$$

For sinusoidally varying fields, $\mathrm{T}(t) \propto \mathrm{e}^{j\omega t}$ and the propagation in the z-direction is given with $\mathrm{Z}(z) \propto \mathrm{e}^{-jkz}$. The constant k indicates the phase advance of the wave per unit length. $\mathrm{X}(x)$ and $\mathrm{Y}(y)$ are trigonometric functions. We can write:

$$\mathrm{E}_z = \mathrm{F}(x,y)\ \mathrm{e}^{j(\omega t - kz)} \; . \tag{3}$$

The exponent $\omega t - kz$ is typical for travelling waves. In fact, if one moves with the wave crest:

$$\begin{aligned} \omega t - kz &= 0 \\ \frac{z}{t} = v_{ph} &= \frac{\omega}{k} \; , \end{aligned}$$

and we know $v_{ph} > c$ for a uniform waveguide.

Accelerators are usually of a circular cross-section, and the electromagnetic field inside usually has a rotational symmetry. For such a case, the wave equation in cylindrical coordinates has the form:

$$\frac{\partial^2 \mathrm{E}_z}{\partial z^2} + \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial \mathrm{E}_z}{\partial r}\right) - \frac{1}{c^2}\frac{\partial^2 \mathrm{E}_z}{\partial t^2} = 0 \; , \tag{4}$$

and the solution

$$\mathrm{E}_z = \mathrm{Z}(z)\ \mathrm{R}(r)\ \mathrm{T}(t) \; . \tag{5}$$

Expressing Z(z) and T(t) as before, we have an equation for R(r):

$$\frac{d^2R}{dr^2} + \frac{1}{r}\frac{dR}{dr} + \underbrace{\left(\frac{\omega^2}{c^2} - k^2\right)}_{K_r^2} R = 0 \; . \tag{6}$$

This is the Bessel equation of zero order (because we have assumed that there is no azimuthal variation of E_z). The solution, as $K_r^2 > 0$, is given by the Bessel function of the first kind and zero order:

$$R(r) = A \; J_0(K_r r) \; , \tag{7}$$

where A is a constant. At the boundary (cylinder of radius a) we must have

$$J_0(K_r a) = 0 \; . \tag{8}$$

Equation (8) has many roots, each root corresponding to a mode of the TM wave. Each mode is a solution of the wave equation. The first root (mode TM_{01}, the subscript 0 indicating no azimuthal variations) is:

$$\begin{aligned} K_r a &= 2.405 \; , \qquad \text{or} \\ K_r &= \frac{2.405}{a} \; . \end{aligned} \tag{9}$$

To satisfy boundary conditions, the value of K_r is fixed. From (6) we get

$$\frac{\omega^2}{c^2} = K_r^2 + k^2 \; . \tag{10}$$

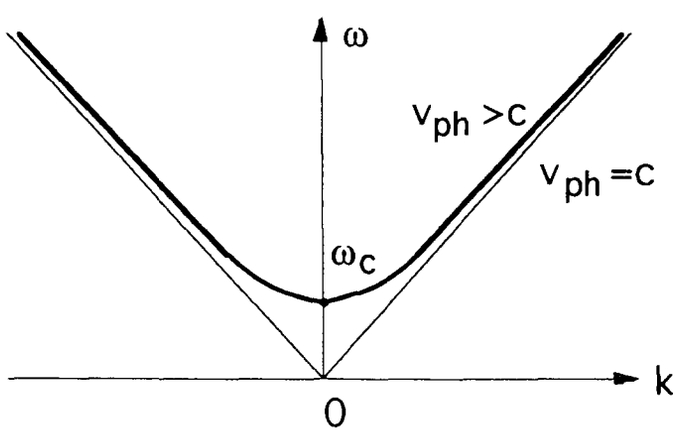

Fig. 5 Dispersion (Brillouin) diagram

This is the dispersion relation. Plotting ω as $f(k)$, one gets the dispersion or Brillouin diagram (see Fig. 5). For $k = 0$, we have the lowest frequency ω_c (cut-off frequency), for which the boundary conditions can still be satisfied. The slope of the radius vector from the origin to a point on the hyperbola gives the phase velocity

$$v_{ph} = \frac{\omega}{k} \; . \tag{11}$$

All points on the hyperbola, being above the asymptote $v_{ph} = c$, have $v_{ph} > c$. The slope of a point on the hyperbola gives the group velocity [2]

$$v_g = \frac{d\omega}{dk} . \tag{12}$$

Note that $v_g = 0$ for the cut-off frequency. The dispersion diagram is symmetric around the origin, which means that waves can travel in both directions, $+z$ and $-z$.

3.2 Waves in loaded cavities

The waves in an empty cavity analysed so far always have

$$v_{ph} > c .$$

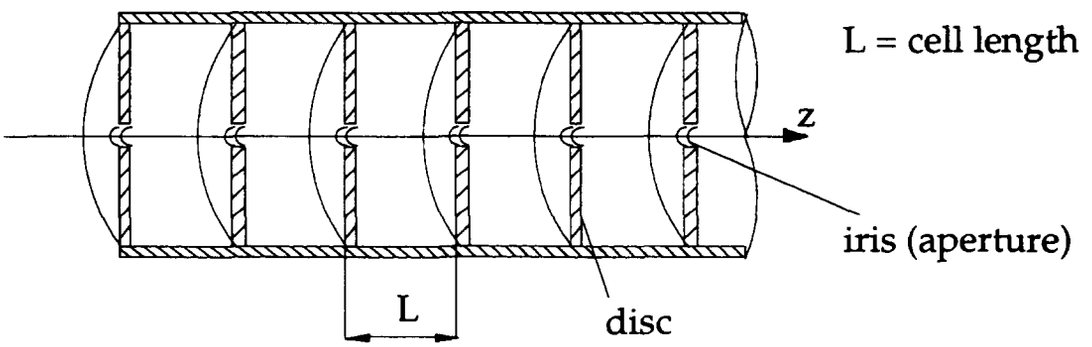

Fig. 6 Disc-loaded cavity (schematic)

In order to slow down the waves, we have to 'load' the cavity by introducing some periodic obstacles into it [3]. Figure 6 shows a disc-loaded cavity with a period (or cell length) L. The study of the solution of the wave equation in such a cavity is based on two essential points:

i) Floquet's theorem for periodic structures, which states that in a given mode of oscillation and at a given frequency, the wave function is multiplied by a constant such as e^{-jk_0L}, when moving from one period to the next (for non-lossless structures this constant is complex, $e^{-\gamma L}$, with $\gamma = \alpha + jk_0$). A simple function which satisfies Floquet's theorem is e^{-jk_0z}.

ii) The complicated boundary conditions cannot be satisfied by a single mode, as was the case with an empty cavity, but by a whole spectrum of so-called 'space harmonics', which is in fact a Fourier series applied to a periodic case.

With the above we write:

$$E_z(r, z, t) = F(r, z)e^{j(\omega t - k_0 z)} \tag{13}$$

$$F(r, z + L) = F(r, z) \tag{14}$$

$$F(r,z) = \sum^{n} a_n(r) e^{-j(2\pi n/L)z} \tag{15}$$

Introducing these expressions in the wave equation for rotationally symmetric waves we get:

$$e^{j\omega t} \sum^{n} e^{-j(k_0+2\pi n/L)z} \left[\frac{d^2 a_n(r)}{dr^2} + \frac{1}{r}\frac{da_n(r)}{dr} + K_r^2 a_n(r) \right] = 0 \ , \tag{16}$$

with

$$K_r^2 = \left(\frac{\omega}{c}\right)^2 - \left[k_0 + \frac{2\pi n}{L}\right]^2 \ . \tag{17}$$

This is the dispersion relation of periodic structures. For each n we have a travelling wave with its own phase velocity

$$v_{ph} = \frac{\omega}{k_0 + 2\pi n/L} = \frac{\omega}{k_n} \ . \tag{18}$$

Comparing (18) with (11), we see that the phase velocity is slowed down and, in principle, we can find an n such that v_{ph} is equal to the velocity of the particles, which is the condition for acceleration.

If we have

$$v_{ph} < c \ ,$$

then $K_r^2 < 0 = -k_r^2$, or $K_r = jk_r$. The solution of the wave equation is now expressed with Bessel functions of imaginary argument, which are called modified Bessel functions. In our case:

$$a_n(r) = A_n I_0(k_r r) \ , \tag{19}$$

with A_n being a constant. A_n usually decreases when $|n|$ increases.

The dispersion diagram of a periodic structure reflects the periodicity of the structure itself. At each obstacle, the iris acts as a scatterer, resulting in a transmitted as well as a reflected wave. When the spacing between irises approaches half the wave length (or a multiple of it) then the transmitted and reflected waves from successive irises interfere strongly and the dispersion curve differs drastically from that of an empty cavity, see Fig. 7.

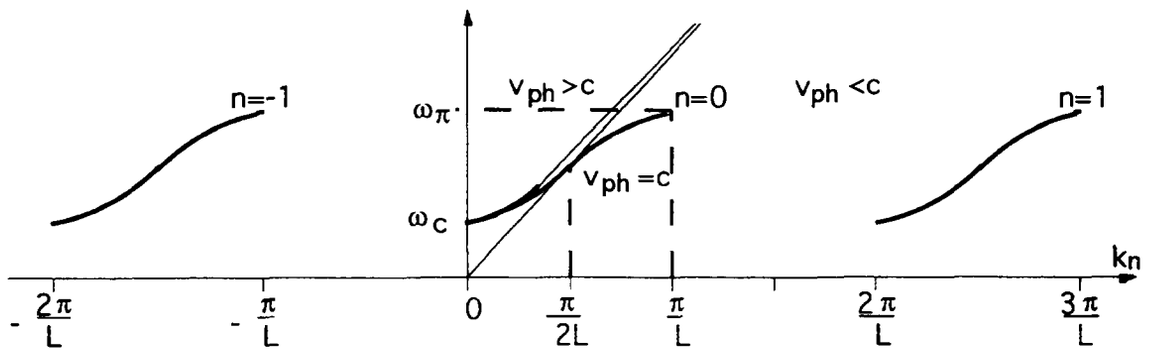

Fig. 7 Dispersion diagram of a periodic structure (waveguide); unloaded structure case included for comparison

The discussion of Fig. 7 is very instructive:

- for a given mode, there is a limited pass band of possible frequencies, $\omega_\pi - \omega_c$; at both ends of the pass band, the group velocity $v_g = d\omega/dk_n$ is zero;
- for a given frequency, one has an infinite series of space harmonics, from $n = -\infty$ to $n = +\infty$. All space harmonics have the same group velocity, but different phase velocities;
- when the group and phase velocity are in the same direction, we speak about forward waves; if the directions are opposite, we speak about backward waves ($n < 0$).

In Fig. 7, the electromagnetic energy propagates in the $+z$ direction ($v_g = d\omega/dk > 0$). At the end of the cavity this energy can either be dissipated into a matched load or be reflected back by a shorting end wall. In the former case, the cavity is in a certain sense a waveguide, and RF accelerators working in this way are called travelling-wave linear accelerators. In the latter case, the energy is reflected back and forth between end walls, as in a resonator, and accelerators working using this principle are called standing-wave linear accelerators. The dispersion diagram of Fig. 8 includes the reflected waves (dotted lines) which are characterized by $v_g < 0$, i.e. the electromagnetic energy flows in the $-z$ direction.

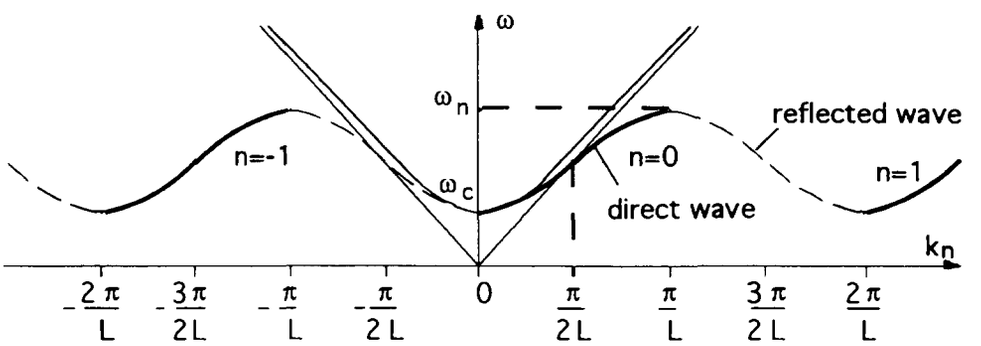

Fig. 8 Dispersion diagram of a periodic structure (resonator)

4 TRAVELLING- AND STANDING-WAVE LINEAR ACCELERATORS

On the basis of the analysis carried out in the preceding Sections, we are now in a position to understand how linear accelerators function. Some important accelerator design parameters still have to be defined [4, 5].

i) Shunt impedance per unit length:

$$Z_s = \frac{E_{zn}^2}{-dP/dz} \tag{20}$$

where E_{zn} is the relevant space harmonic of the axial electric field synchronous with the particle, and $-dP/dz$ is the RF power dissipated in the walls of the accelerator per unit length.

The shunt impedance is a measure of the excellence of a structure to accelerate particles. Note that some authors include a factor of one half in Eq. (20) to be consistent with the standard circuit theory. A high Z_s means an efficient acceleration for a given dissipated RF power. The Z_s is usually expressed in $M\Omega m^{-1}$.

ii) Quality factor:

$$Q = \omega \frac{w}{-dP/dz} \tag{21}$$

where ω is the RF angular frequency, and w is the stored energy per unit length.

iii) Ratio Z_s/Q:

$$\frac{Z_s}{Q} = \frac{E_{zn}^2}{\omega w} \; ; \tag{22}$$

This ratio depends only on the structure geometry and not on the quality of the surface of the cavity walls. It tells us how much accelerating field one has for a given stored energy per unit length.

iv) Frequency: the choice of the frequency of an accelerator depends on many factors. For relativistic particles, such as electrons above 10 MeV, the frequency is usually in the range of several GHz, whilst for protons it is of several hundred MHz. One can also show that

$$Z_s \propto \omega^{1/2} \; ,$$

so higher frequencies are more efficient for acceleration. The breakdown problems due to high electric surface fields, are also diminished at higher frequencies, as indicated by the Kilpatrick law (see Section 8.1). However, as structures at high frequency are smaller, they might not have a sufficient aperture for intense particle beams.

v) Group velocity:

$$v_g = \frac{P}{w} \; ; \tag{23}$$

where P is the power flowing in the structure.

This definition of the group velocity is obvious. Since the stored energy per unit length is proportional to E_{zn}^2, a small v_g, usually 0.01 c to 0.02 c, is preferred for better acceleration efficiency, keeping the accelerator filling time still in reasonable limits, see Eq. (35).

4.1 Travelling-wave accelerators

There is no firm rule with which to decide whether a travelling-wave or a standing-wave accelerator is to be chosen. However, travelling-wave accelerators are usually used when dealing with short beam pulses and when particle velocities approach the velocity of light, as is the case with electrons. One should distinguish between a constant impedance and a constant gradient travelling-wave accelerator [6].

A travelling-wave accelerator operates with the $n = 0$ space harmonic, as the wave amplitude is largest in this case. The structure is usually designed so that the dispersion

curve crosses the $v_{ph} = c$ line for a phase advance per period of $\pi/2$ to $2\pi/3$ (approximately in the middle of the pass band) or

$$k_0 = \frac{\pi}{2L} \div \frac{2\pi}{3L} ,$$

(see Fig. 7). Note that 10 MeV electrons already have a relativistic β_r of 0.999, whilst that of protons is only 0.145. Therefore, the structure with $v_{ph} < c$ is only used for very low electron energies.

Along the accelerator, a certain power is dissipated in the cavity walls and the electric field is attenuated:

$$\frac{dE_{z0}(z)}{dz} = -\alpha(z) E_{z0}(z) , \tag{24}$$

where $\alpha(z)$ is the attenuation constant.

We have correspondingly

$$\frac{dP(z)}{dz} = -2\alpha(z) P(z) . \tag{25}$$

Using the definitions of Q and w, we also have

$$\frac{dP(z)}{dz} = -\frac{\omega\ P(z)}{Q\ v_g(z)} . \tag{26}$$

It follows that

$$\alpha(z) = \frac{\omega}{2Q\ v_g(z)} . \tag{27}$$

A *constant-impedance* accelerator has a uniform structure:

$$\begin{aligned} \alpha(z) &= \text{const} = \alpha \\ E_{z0}(z) &= E_{z0}(0)\ e^{-\alpha z} . \end{aligned}$$

The energy gain of a particle of charge q, sitting on the crest of the wave in a constant-impedance accelerator of length ℓ is

$$\begin{aligned} qV &= q \int_0^{\ell} E_{z0}(0)\ e^{-\alpha z}\ dz = \\ &= q\ E_{z0}(0)\ell\ \frac{1 - e^{-\alpha\ell}}{\alpha\ell} . \end{aligned} \tag{28}$$

This expression is maximum for $\alpha\ell = 1.26$ and

$$qV_{\max} = 0.57q\ E_{z0}(0)\ell . \tag{29}$$

At the end of the accelerator in the optimized case one has:

$$E_{z0}(\ell) = 0.28 E_{z0}(0) , \tag{30}$$

$$P(\ell) = 0.08 P(0) . \tag{31}$$

Less than 10% of the input power goes into the matched load at the end of the structure.

Sometimes the change of E_{z0} along the accelerator is inconvenient and one prefers to have it constant. Such a structure is called a *constant gradient* accelerator and to achieve $E_{z0} = \text{const}$, the structure dimensions have to be suitably varied along the accelerator. It is enough, for example, to increase progressively the size of the iris. It is assumed here that these variations practically do not affect the Z_s.

From

$$E_{z0} = \text{const}$$

and

$$Z_s = \frac{E_{z0}^2}{-dP(z)/dz} = \text{const} \ ,$$

it follows that the variation of P along the accelerator is also constant:

$$-\frac{dP(z)}{dz} = \text{const} = \frac{d}{dz}\left(P(0) + \frac{P(\ell) - P(0)}{\ell} z\right) .$$

With

$$P(\ell) = P(0)e^{-2\int_0^\ell \alpha(z)dz} = P(0)e^{-2\tau} \ ,$$

one can write

$$\frac{dP(z)}{dz} = -P(0)\frac{1 - e^{-2\tau}}{\ell} = \text{const} \ , \tag{32}$$

where the total attenuation along the accelerators is

$$\tau = \int_0^\ell \alpha(z)dz \ . \tag{33}$$

The group velocity in a constant gradient accelerator is not constant, but decreases linearly as $P(z)$:

$$v_g(z) = -\frac{\omega P(z)}{Q dP(z)/dz} \ \propto P(z) \ , \tag{34}$$

because $dP(z)/dz = \text{const}$.

The time needed for the electromagnetic energy to fill the accelerator, prior to the acceleration of particles, is called the filling time t_F:

$$t_F = \int_0^\ell \frac{dz}{v_g(z)} \ . \tag{35}$$

Figure 9 shows schematically a travelling-wave accelerator, with its input and output couplers. These couplers are in fact wave guides, matched to the characteristic impedance of the accelerator, by which the input RF power is fed into the cavity, and the output RF power brought to the matched load.

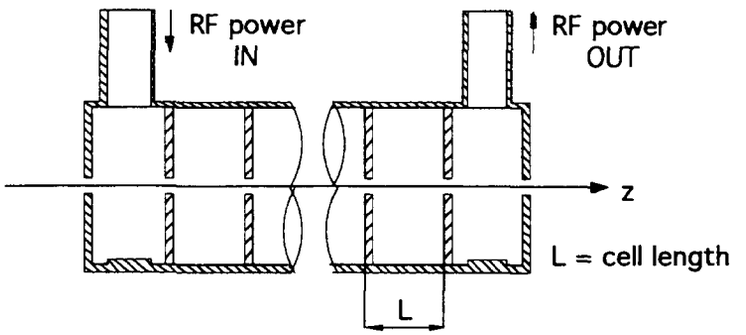

Fig. 9 Travelling-wave linear accelerator with input and output couplers (schematic)

4.2 Standing-wave accelerators

A direct and a reflected sinusoidally varying wave, travelling with the same velocity but in opposite directions, combine to create a standing-wave pattern. If the amplitudes of the direct and reflected wave are A_0 and B_0, respectively, the standing-wave pattern has maxima $A_0 + B_0$ and minima $A_0 - B_0$, distant from each other by

$$d = \frac{\pi}{2k_0} ,$$

with $k_0 = \omega/v_{ph}$. The average amplitude of the standing-wave pattern is A_0 [6], hence the same as the direct travelling wave.

Such a standing-wave pattern is not useful as the reflected wave only dissipates power travelling backwards, but does not contribute to the acceleration of particles.

Standing-wave accelerators, however, use both the direct and the reflected wave to accelerate particles. How this comes about can be understood by inspecting Fig. 8: at the points where the direct (solid line) and reflected (dotted line) space harmonics join, they have the *same* phase velocity, and if this velocity is synchronous with the particle, both harmonics contribute to the acceleration.

From Fig. 8 one understands that standing-wave linear accelerators operate either at the lowest or at the highest frequency of the pass band, where

$$k_n L = N\pi, \qquad N = 0, \pm 1 \ . \tag{36}$$

In other words, we say that the *structure* mode in standing-wave accelerators is either 0 or π. In the first case, the fields in all the cells are in phase; in the second, the fields in adjacent cells are in phase opposition.

Apart from the structure mode, we also have the *cavity* mode, as already mentioned in Section 2.1. For each wave type, TM or TE, the modes are specified by three subscripts (two for travelling waves) m, n, and p, which in the case of a cylindrical cavity indicate the following:

- m is the number of full period field variations azimuthally;
- n is the number of zeros of the axial field radially ($0 < r \leq a$);
- p is the number of half-period field variations longitudinally (only for standing waves).

For standing waves, the number of half-period variations along the cavity of length ℓ must be an integer:

$$k_n \ell = p\pi, \qquad p = 0, 1, 2, \ldots . \tag{37}$$

This condition is a result of the necessity to annul the tangential electric field on end walls; for rotationally symmetric fields one has:

$$\mathrm{div}\vec{E} = \frac{1}{r}\frac{\partial}{\partial r}\left(rE_r\right) + \frac{\partial E_z}{\partial z} = 0 \ , \tag{38}$$

giving

$$E_r(r) = -\frac{1}{r}\int_0^r \frac{\partial E_z}{\partial z}\, r'\, dr' \ . \tag{39}$$

To have $E_r = 0$, the field E_z must have an extremum at the end walls, and therefore it can only have amplitude distributions along the cavity (relative values) as shown in Fig. 10. Obviously, the only interesting case is $p = 0$, which is used in standing wave accelerators. Other cases, $p \neq 0$, are present to satisfy boundary conditions if there are errors in the structure dimensions along the cavity.

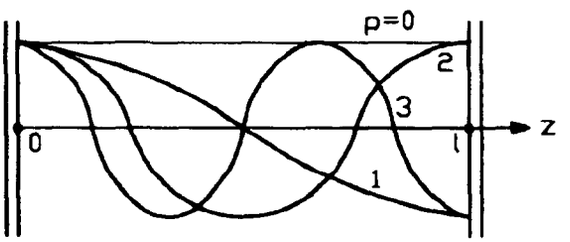

Fig. 10 Standing-wave amplitude distribution (relative values) in a cavity, for different cavity modes, $p = 0, 1, 2, \ldots$.

There exists yet another mode which is important: the so-called *beam mode* [7]. The beam mode is defined by the RF phase change during the time the beam travels over a structure period. It differs from the structure mode $k_n L$ and is given by

$$k_n L + 2h\pi, \quad h = 0, 1, 2 \ . \tag{40}$$

For example, for a zero structure mode, the beam mode is usually $2\pi(h = 1)$. The value $4\pi(h = 2)$ can be used for slower particles, but the accelerating field amplitude is reduced, as was already mentioned before in connection with space harmonics having $n > 1$, see Section 3.2. To get a better idea of the operation of standing-wave accelerators, a cavity operating in the π structure mode is shown in Fig. 11. The fields in adjacent cells are in counterphase. At a given moment (as indicated in Fig. 11), the particles to be accelerated are in odd-numbered cells, the even-numbered ones are empty. Half an RF cycle later, the particles will be accelerated in even-numbered cells and the odd-numbered ones will be empty.

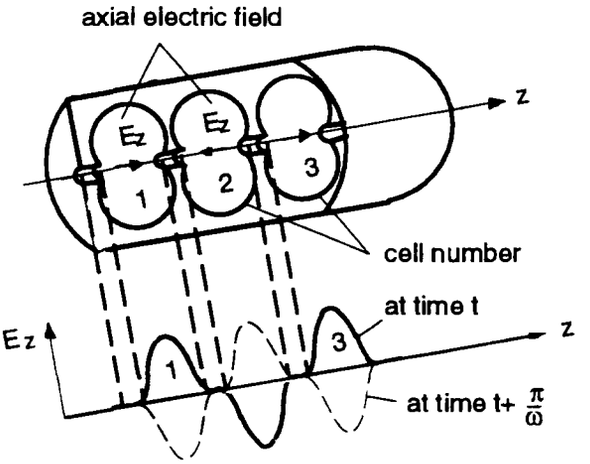

Fig. 11 Standing-wave accelerator operating in the π structure mode (schematic)

In contrast to the π structure mode, the zero structure mode has the electromagnetic fields in adjacent cells in phase. The walls, separating the cells, could therefore be left out without affecting the electromagnetic field distribution in the cavity [8]. An example of an accelerator working in the zero mode is the drift tube linear accelerator, represented schematically in Fig. 12. This accelerator, very frequently used for the acceleration of protons and light ions, was invented by L. Alvarez, as mentioned in Section 1. It operates with an electromagnetic field configuration TM_{010} (always defined with respect to the empty cavity), and accelerates particles in the gaps between drift tubes. When the longitudinal electric field E_z reverses sign, the particles are screened inside the drift tubes. As the velocity of particles increases, the cell length becomes correspondingly longer in order to keep the synchronism between the arrival of particles and the accelerating field in the gap.

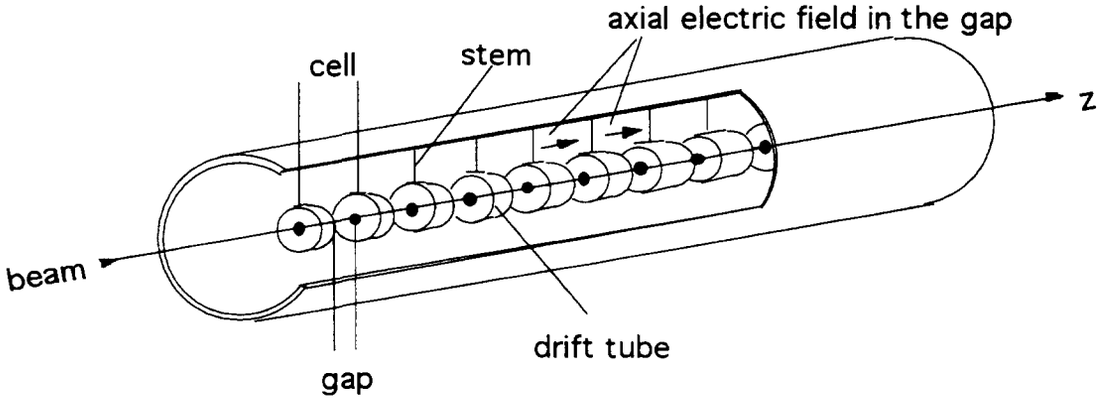

Fig. 12 Alvarez linear accelerator operating in the 0 structure mode (schematic)

It should be noted that the electromagnetic energy builds up *progressively* in time in a standing-wave accelerator. The filling time of such a cavity lasts until the input power just covers the dissipation losses and is given by

$$t_F \propto \frac{Q}{\omega} ,$$

with Q being the quality factor of the cavity, and ω the RF angular frequency. When the beam is injected, additional RF power is fed into the cavity to compensate for the beam loading of the cavity, i.e. for the power taken by the beam.

5 ELECTROMAGNETIC FORCES AND THEIR ACTION

5.1 Method of computation

So far we have tried to explain the mechanism of the action of electromagnetic fields in a cavity on moving charged particles. Now we shall analyse the electromagnetic forces in more detail, and show how one can compute their action [9, 10].

As an example we shall take an Alvarez linear accelerator, which operates in the TM_{010} cavity mode, and thus has field components E_z, E_r, and B_ϑ. Figure 13 presents schematically an Alvarez cell, containing two half drift tubes and a gap. The centre of the coordinate system is placed at mid gap, and the E_z field is symmetric around the origin.

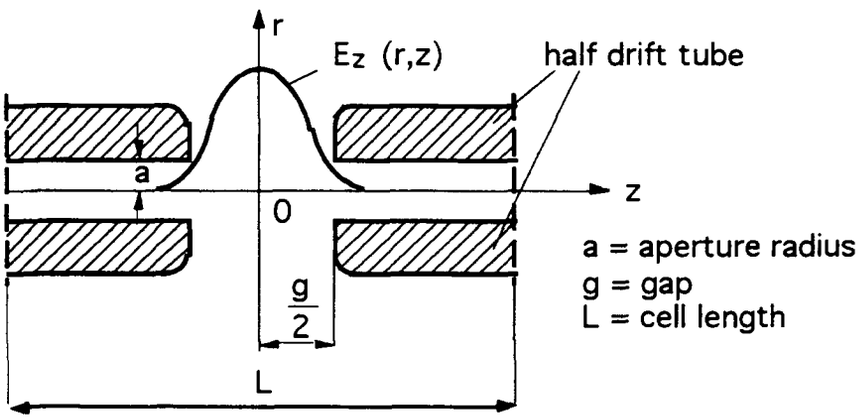

Fig. 13 Cell of an Alvarez linear accelerator (schematic)

In Section 3.2 we expressed the E_z field of a periodic structure by a Fourier series, see Eqs. (13), (14), and (15). In the same way, but in order to be more general, we express this field now by a Fourier integral

$$E_z(r,z) = \int_{-\infty}^{\infty} A_k \; I_0(k_r r) \cos kz \; dk \; , \tag{41}$$

where cos kz has replaced the exponential function because E_z is an even function of z and

$$k_r = \left[k^2 - \left(\frac{\omega}{c}\right)^2\right]^{1/2} .$$

By inversion of the Fourier integral one gets

$$A_k \; I_0(k_r r) = \frac{1}{2\pi} \int_{-\infty}^{\infty} E_z(r,z) \cos \; kz \; dz \; . \tag{42}$$

To determine the constant A_k, we shall assume that at the drift tube bore radius a, the E_z field has a constant value in the gap, and zero within the drift tube:

$$E_z(a,z) = \text{const} = E \; , \qquad -\frac{g}{2} \le z \le \frac{g}{2} \; . \tag{43}$$

Solving the integral we get

$$A_k = \frac{Eg}{2\pi} \frac{\sin\ kg/2}{kg/2} \frac{1}{I_0(k_r a)} \ , \tag{44}$$

and for the field

$$E_z(r,z) = \frac{Eg}{2\pi} \int_{-\infty}^{\infty} \frac{\sin kg/2}{kg/2} \frac{I_0(k_r r)}{I_0(k_r a)} \cos kz \ dk \ . \tag{45}$$

Let us now compute the gain in energy of a particle q, when passing through the i-th cell of an Alvarez accelerator:

$$W_i - W_{i-1} = q \int_{-L_i/2}^{L_i/2} E_z(r,z,t)dz = q \int_{-L_i/2}^{L_i/2} E_z(r,z) \cos\ (\omega t + \varphi)dz \ , \tag{46}$$

L_i is the cell length and φ is the RF phase (counted from the crest of the wave) when the particle crosses the mid gap ($t = 0$). At a given time t, the particle will be at the position z given by

$$z = v_p t \ , \tag{47}$$

where v_p is the velocity of the particle, considered here as constant. Substituting $t = z/v_p$ in the integral one gets

$$W_i - W_{i-1} = q \int_{-L_i/2}^{L_i/2} E_z(r,z) \left[\cos \frac{\omega z}{v_p} \cos\varphi - \sin \frac{\omega z}{v_p} \sin\varphi \right] dz \ . \tag{48}$$

We recognize that the first term under the integral gives the inverse Fourier integral multiplied by $\cos\varphi$, and the second term is zero, as the even function E_z is multiplied by the odd function $\sin \omega z/v_p$. As the inverse Fourier integral has already been computed (note that $k = \omega/v_p$), we have

$$W_i - W_{i-1} = q\ Eg \frac{\sin\ kg/2}{kg/2} \frac{I_0(k_r r)}{I_0(k_r a)} \cos\varphi \ . \tag{49}$$

The result is interesting: $q\ Eg \cos\varphi$ would be the energy gain due to an alternating voltage Eg traversed at a phase φ from the peak; the factor

$$\frac{\sin\ kg/2}{kg/2} \ ,$$

indicates the reduction in the energy gain due to the finite time the particle travels across the gap; the factor

$$\frac{1}{I_0(k_0 a)} \ ,$$

is the additional reduction due to the fact that the field E_z on the axis penetrates into the drift tube aperture and is weaker than E, the field at the drift tube bore radius. Of course

$$\int_{-L/2}^{L/2} E_z(0,z)dz = Eg \ .$$

The field off axis is always stronger than the field on axis [factor $I_0(k_r r)$].

Substituting Eg by E_0L (E_0 is the average longitudinal field on axis in a cell), we get finally

$$W_i - W_{i-1} = q\ E_0L\ T(k,r)\cos\varphi\ , \tag{50}$$

with

$$T(k,r) = \frac{\sin\ kg/2}{kg/2}\ \frac{I_0(k_r r)}{I_0(k_r a)} = T(k)I_0(k_r r)\ . \tag{51}$$

$T(k,r)$ is the so-called transit time factor. In deriving the formula for the energy gain, we have seen that out of the whole spectrum of waves, only the wave having the same velocity as the particle ($k = \omega/v_p$) has given a contribution. This is in agreement with our earlier considerations, and we can regard the transit time factor as the relative amplitude of the space harmonic synchronous with the particle. The value of the transit time on the axis is $T(k)$, as $I_0(0) = 1$. In standing wave accelerators, the $T(k)$ appears with the average longitudinal field E_0 in the expression for the shunt impedance

$$Z_{\text{seff}} = \frac{(E_0T)^2}{-dP/dz}\ , \tag{52}$$

which is called the effective shunt impedance. The factor T is the average transit time factor on the axis in the accelerator.

In our analysis we have kept the particle velocity constant in a cell, which is obviously only an approximation. However, the energy gain formula stays sufficiently accurate if one takes for v_p its value in the mid gap. How one computes this value is beyond the scope of this paper.

Knowing E_z, one can compute the other electromagnetic field components: E_r with Eq. (39) and B_ϑ via

$$\text{rot}\ \vec{B} = \frac{1}{c^2}\ \frac{\partial \vec{E}}{\partial t}\ , \tag{53}$$

giving

$$\frac{\partial B_\vartheta}{\partial z} = -\frac{1}{c^2}\ \frac{\partial E_r}{\partial t}\ . \tag{54}$$

One also has to make use of a property of Bessel functions:

$$\int\ x\ I_0(x)dx = xI_1(x)\ . \tag{55}$$

The formulae for all the TM_{010} field components are written below:

$$E_z(r,z,t) = \frac{E_0L}{2\pi}\int_{-\infty}^{\infty} T(k)I_0(k_r r)\cos kz\cos(\omega t+\varphi)dk \tag{56}$$

$$E_r(r,z,t) = \frac{E_0L}{2\pi}\int_{-\infty}^{\infty} T(k)\ \frac{k}{k_r}I_1(k_r r)\sin kz\cos(\omega t+\varphi)dk \tag{57}$$

$$B_\vartheta(r,z,t) = \frac{E_0L}{2\pi}\ \frac{\omega}{c^2}\int_{-\infty}^{\infty} T(k)\ \frac{1}{k_r}\ I_1(k_r r)\cos kz\sin(\omega t+\varphi)dk \tag{58}$$

The components E_r and B_ϑ produce a radial force on the moving particle. The change of the radial momentum p_r of the particle in a cell is

$$p_{r,i} - p_{r,i-1} = q\int_{-L_i/2}^{L_i/2}\left[E_r\left(r,z\right)\cos\left(\frac{\omega z}{v_p}+\varphi\right) - v_pB_\vartheta\left(r,z\right)\sin\left(\frac{\omega z}{v_p}+\varphi\right)\right]\frac{dz}{v_p}\ . \tag{59}$$

The computation proceeds in an analogous way as with $W_i - W_{i-1}$.

5.2 Actual computations

The method of computation presented in the preceding paragraph has given us a physical insight into the problem. It was shown that in standing-wave accelerators the action of the electromagnetic field on a particle is not computed point by point, but given by means of integral expressions. Such integrals are usually computed in programs which analyse the structure of linear accelerators.

The sequence of computer programs used in the design of standing-wave accelerators is the following:

i) A structure program, such as SUPERFISH [11], URMEL [12] or MAFIA [13] (in earlier days CLAS [14]) solves the wave equation in a cell of the structure and determines the cell dimensions corresponding to a given RF frequency. All important structure parameters such as Z_{seff}, Q, dP/dz etc. are computed, as well as various integrals of the electromagnetic field analogous to (42). These integrals are called transit-time coefficients and are used for beam dynamics calculations, in particular to determine the exact energy and phase of the synchronous particle at mid gap [10, 15, 16]. One of them, which we already know, is called the transit time factor and is computed with

$$T(k) = \frac{\int_{-L/2}^{L/2} E_z(0,z) \cos\ \omega z/v_p\ dz}{\int_{-L/2}^{L/2} E_z(0,z) dz} \ . \tag{60}$$

ii) Having computed a certain number of structure cells corresponding to different particle velocities, an accelerator is generated using a program such as GENLIN [16] or PARI [17]. Such a program interpolates between precalculated structure cells to determine all the cells in a certain energy (velocity) range. Then, with the help of transit-time coefficients, exact energies and phases of the synchronous particle in all the cells are computed, and the accelerator is thus progressively determined.

iii) An accelerator has to accept and accelerate particles which lie in a certain range around the synchronous particle. This range is called the acceptance of the accelerator, and the beam has to be matched to it. The calculation of the acceptance, beam matching parameters and, if required, an external transverse focusing is done with programs like ADAPT [18], where the influence of beam space charge forces is also included.

iv) As a final step, the accelerator designed with the above procedure is tested by simulating a beam through it. Simulation programs, such as PARMILA [19] represent the beam with up to several thousands of macro-particles (their number is limited by the required computer time) and follow them through the accelerator. The number of particles at the output and their distribution is a measure of the quality of the designed accelerator.

A travelling-wave linear accelerator structure is designed in a similar way. One can also use SUPERFISH to determine the structure, although this program computes only standing-wave fields. With a 'trick' it is possible to obtain a travelling-wave by conveniently adding two standing waves [20]. In fact, only one standing wave is computed by

SUPERFISH, but the result is then multiplied by phase constants to give two standing waves, which, when added, result in the desired travelling wave. All structure parameters for the travelling-wave accelerator can be deduced from the standing-wave ones. In particular, the group velocity can be computed by

$$v_g = \frac{P}{w} = \frac{1/2 \int E_r H_\vartheta ds}{w} , \tag{61}$$

where the numerator represents the average power flow expressed by the Poynting vector, and the denominator indicates the stored energy. Note that the fields in the numerator are travelling-wave fields, and the stored energy in the denominator is half the energy stored in the standing-wave case (no reflected waves).

The travelling-wave structures are usually made for a constant velocity ($v_{ph} = c$), and hence programs like GENLIN are not needed. The beam simulation for low energies, where space charge is still important, is done by programs such as PARMELA [21]. At higher energies programs such as TRANSPORT [22] are used.

6 PHASE STABILITY AND ACCEPTANCE OF AN ACCELERATOR

The particles in a beam are specified in the six-dimensional phase space by two sets of conjugate coordinates, p and q. The coordinates are conjugate if they satisfy the conditions

$$\frac{\partial q}{\partial t} = \frac{\partial H}{\partial p} \tag{62}$$

and

$$\frac{\partial p}{\partial t} = -\frac{\partial H}{\partial q} , \tag{63}$$

where $H = H(p, q, t)$ is a function called the Hamiltonian. The coordinates of the particles of a beam occupy a volume in the phase space which is constant. This is Liouville's theorem [23]. The six-dimensional phase space is usually analysed through its two-dimensional projections, of which we distinguish two transverse, and one longitudinal. The area occupied by the particles in each phase plane is called the beam emittance. The area accepted by the accelerator is called the acceptance. In the longitudinal phase plane the principle of phase stability, discovered in 1945 by McMillan and Veksler, is of essential importance for the acceleration of particles by sinusoidally varying RF fields.

6.1 Phase stability

The principle of phase stability states that particles with phases and energies in the neighbourhood of the synchronous one are also accelerated, and in doing so, they oscillate in energy and phase around the synchronous particle. These oscillations are called synchrotron oscillations.

The synchrotron oscillations can easily be understood if one observes what happens in a gap of the Alvarez linear accelerator. From Fig. 14a one can see that a particle arriving at the mid gap with a phase φ different from the synchronous phase φ_s, gains an energy different from the nominal one. The energy difference at the end of the Alvarez cell is

$$\Delta W = qE_0TL \, (\cos\varphi - \cos\varphi_s) ,$$

or its variation with z

$$\frac{d}{dz}\Delta W = qE_0T\ (\cos\varphi - \cos\varphi_s)\ . \tag{64}$$

(Note that one usually uses z as the independent variable and not t).

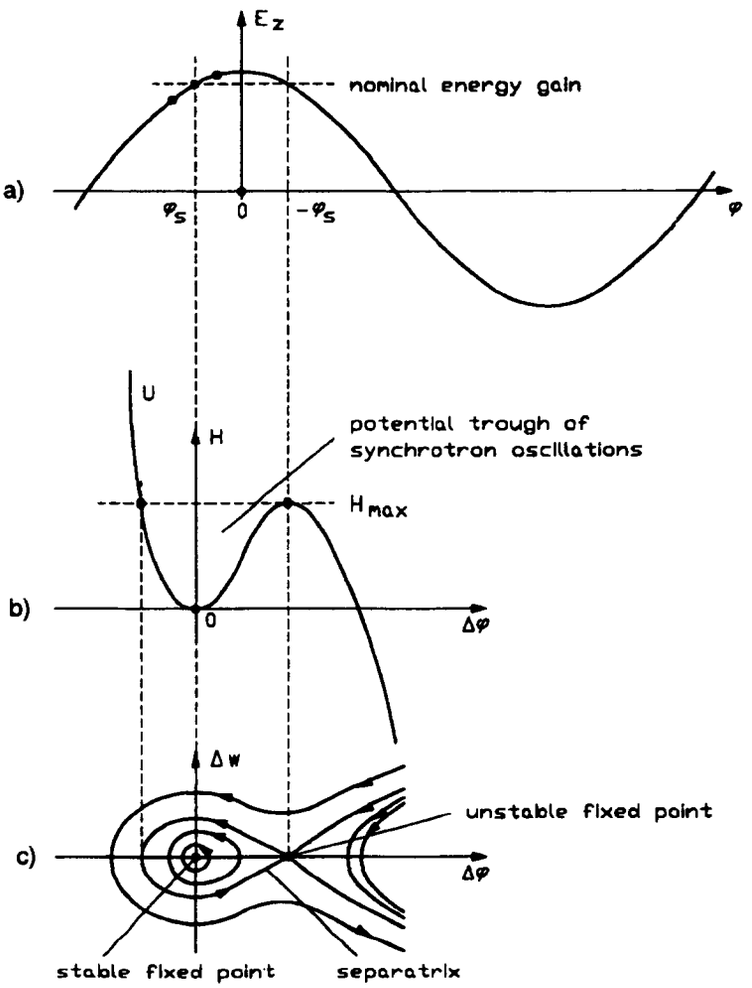

Fig. 14 a) Acceleration in an RF gap; b) potential trough of synchrotron oscillations; c) stable region bounded by the separatrix

Corresponding to the difference in energy, there is also a difference in phase

$$\Delta\varphi = \varphi - \varphi_s\ ,$$

or its variation with z

$$\frac{d}{dz}\ (\varphi - \varphi_s) = \omega\left(\frac{dt}{dz} - \frac{dt_s}{dz}\right) = \frac{\omega}{c}\left(\frac{1}{\beta} - \frac{1}{\beta_s}\right) \cong -\frac{\omega}{\beta_s c}\ \frac{\Delta\beta}{\beta_s}\ ;$$

$\Delta\beta/\beta_s$ can be expressed via $\Delta W/W_s$ and one obtains

$$\frac{d}{dz}\ \Delta\varphi = -\frac{\omega}{2c\beta_s\gamma_s^2}\ \frac{\Delta W}{W_s}\ . \tag{65}$$

The set of first-order differential equations, (64) and (65), describes the synchrotron oscillations. If the frequency is constant, the variables ΔW and $\Delta\varphi$ are canonically conjugate, and they describe the longitudinal phase plane.

Figure 14b shows the potential trough of synchrotron oscillations. The trough is asymmetric as the set of equations is not linear. To obtain a solution, one assumes adiabaticity ($\beta_s, \gamma_s \cong$ const), and transformes the first-order set into a second-order differential equation:

$$\frac{d^2}{dz^2}\,\Delta\varphi + K\Big[\cos\,(\varphi_s + \Delta\varphi) - \cos\varphi_s\Big] = 0\ , \tag{66}$$

with

$$K = \frac{\omega q E_0 T}{2c\beta_s\gamma_s^2\,W_s}\ .$$

The equation (66) is transformed in the usual way by considering the second term as $\partial U/\partial\Delta\varphi$ and multiplying all by $d\Delta\varphi/dz$

$$\frac{d}{dz}\left[\frac{1}{2}\left(\frac{d\Delta\varphi}{dz}\right)^2 + U(\Delta\varphi)\right] = 0\ . \tag{67}$$

The expression in the bracket is constant and it represents the total energy of oscillation, usually denoted by H. All particles with $H \leq H_{\max}$ will remain in the potential trough (see Fig. 14b). The potential function $U(\Delta\varphi)$ has two points where the first derivative (with respect to $\Delta\varphi$) is zero: one is a minimum (stable point), and the other a maximum (unstable point). At both points, the velocity in the phase plane is zero, and therefore they are called fixed points.

The bracket of (67) can be written as

$$\frac{\omega}{2c\beta_s\gamma_s^2 W_s}\left\{\frac{\omega}{4c\beta_s\gamma_s^2 W_s}\,\Big(\Delta W\Big)^2 + q\,E_0T\Big[\sin\Big(\varphi_s + \Delta\varphi\Big) - \Delta\varphi\cos\varphi_s - \sin\varphi_s\Big]\right\} = H\ , \tag{68}$$

where one has used (65) and expressed $U(\Delta\varphi)$ in order for it to be zero for $\Delta\varphi = 0$. For each H, one has a trajectory in the $\Delta\varphi, \Delta W$ phase plane (see Fig. 14c). For $H_{\max}$ [equal to $U(-2\varphi_s)$], we get the separatrix, dividing the stable from the unstable region:

$$\frac{\omega}{4c\beta_s\gamma_s^2\,W_s}\,\Big(\Delta W\Big)^2 + q\,E_0T\Big[\sin(\varphi_s + \Delta\varphi) + \sin\varphi_s - (2\varphi_s + \Delta\varphi)\cos\varphi_s\Big] = 0\ . \tag{69}$$

The separatrix is shown in Fig. 14c. The maximum allowed ΔW is found for $\Delta\varphi = 0$:

$$\Delta W_{\max} = \pm 2\left[\frac{2q\,E_0T\,c\beta_s\gamma_s^2 W_s(\varphi_s\cos\varphi_s - \sin\varphi_s)}{\omega}\right]^{1/2}\ . \tag{70}$$

The area inside the separatrix is called the bucket. Note that the oscillations are stable only for $\varphi_s < 0$.

To analyse the motion around the stable phase, one can linearize (64) by $\cos\varphi - \cos\varphi_s \cong -\sin\varphi_s\Delta\varphi$, and thus obtain a linear set of equations of the form

$$\frac{d}{dz}\Delta\varphi = -b^2(z)\Delta W\ , \tag{71}$$

$$\frac{d}{dz}\Delta W = a^2(z)\Delta\varphi \ , \tag{72}$$

with $a^2(z)$ and $b^2(z)$ being slowly varying functions of z. Eliminating dz and integrating, one obtains the equation of an ellipse

$$\frac{(\Delta\varphi)^2}{a^2} + \frac{(\Delta W)^2}{b^2} = \text{const} \ ,$$

(see Fig. 14c).

The linearized system (71), (72) can be solved without the assumption of adiabaticity. The second-order differential equation is

$$\frac{d}{dz}\left(\beta_s^3\gamma_s^3 \ \frac{d}{dz} \ \Delta\varphi\right) - \frac{\omega q \ E_0 T \sin\varphi_s}{mc^3}\Delta\varphi = 0 \ , \tag{73}$$

m being the rest mass of the particle. Multiplying (73) by $\beta_s^3\gamma_s^3$ and substituting

$$\beta_s^3\gamma_s^3 \ \frac{d}{dz} = \frac{d}{du} \ ,$$

one gets an equation of the form

$$\frac{d^2}{du^2} \ \Delta\varphi + K(u)\Delta\varphi \ , \tag{74}$$

where $K(u)$ is a slowly varying function of u. This equation is solved by the usual BKW (Brillouin, Kramer, Wenzel) method, which assumes solutions of the form

$$w(u)e^{j\psi(u)} \ .$$

Transforming the solution back to the independent variable z, we have

$$\Delta\varphi(z) = C \ \widehat{\Delta\varphi}(z) \sin\left(\int_0^z \Omega_s(z)dz' + \varphi_0\right) \ ,$$

with C and φ_0 being integration constants and

$$\widehat{\Delta\varphi}(z) \ = \ \left[\frac{-mc^3}{\omega q E_0 T \sin\varphi_s \beta_s^3\gamma_s^3}\right]^{1/4} \ , \tag{75}$$

$$\Omega_s(z) \ = \ \left[\frac{-\omega q E_0 T \sin\varphi_s}{mc^3\beta_s^3\gamma_s^3}\right]^{1/2} \ . \tag{76}$$

Equations (75) and (76) give the amplitude and frequency of small phase oscillations. In the course of acceleration, $\widehat{\Delta\varphi}$ becomes smaller

$$\widehat{\Delta\varphi} \ \propto \ (\beta_s\gamma_s)^{-3/4} \ . \tag{77}$$

This is called phase damping. If the area in the phase plane is conserved, the energy spread in the beam is increased correspondingly

$$\widehat{\Delta W} \ \propto \ (\beta_s\gamma_s)^{3/4} \ . \tag{78}$$

Figure 15 is a more complete picture of the $\Delta\varphi, \Delta W$ phase plane, also showing unstable trajectories outside the separatrix. Note that the stable synchronous phase φ_s is always counted from the crest of the wave. The plot of Fig. 15 is valid for particle velocities v_p smaller than c.

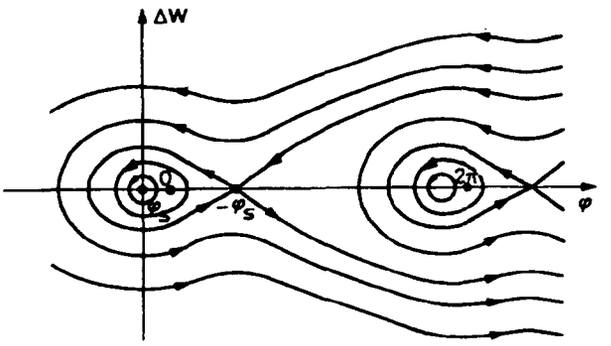

Fig. 15 Trajectories in the longitudinal phase plane (φ_s is counted from the crest of the wave in ion linacs)

What happens when v_p approaches c? It is obvious that the trajectories in the phase plane will change considerably. Because the velocity of light is approached with electrons, we shall consider the longitudinal phase plane in electron linear accelerators somewhat differently from the above.

Electron linear accelerators are usually travelling-wave accelerators, and it is common practice to count the phase from the zero crossing of the wave, and not from its crest. The first-order differential equations, analogous to (64) and (65), are written in the form [24]

$$\frac{d\gamma}{dz} = -\frac{eE_0}{mc^2}\sin\varphi \ , \tag{79}$$

$$\frac{d\varphi}{dz} = \frac{\omega}{c}\left(\frac{1}{\beta_{ph}} - \frac{1}{\beta}\right) = \frac{\omega}{c}\left(\frac{1}{\beta_{ph}} - \frac{\gamma}{(\gamma^2-1)^{1/2}}\right) \ , \tag{80}$$

with γmc^2 the total energy of the electron, e the electron charge, $\beta_{ph}c$ the wave velocity, E_0 the relevant space harmonic, and βc the electron velocity.

Eliminating dz one gets

$$\frac{\omega}{c}\left[\frac{1}{\beta_{ph}} - \frac{\gamma}{(\gamma^2-1)^{1/2}}\right] d\gamma = -\frac{eE_0}{mc^2}\sin\varphi \ d\varphi \ , \tag{81}$$

and after integration

$$\cos\varphi = \frac{\omega\, mc}{eE_0}\left(\frac{\gamma}{\beta_{ph}} - \beta\gamma\right) + C \ . \tag{82}$$

The extremum of phase φ is obtained with (81) putting

$$\frac{d\varphi}{d\gamma} = 0 \ ,$$

which gives

$$\frac{1}{\beta_{ph}} = \frac{\gamma}{(\gamma^2-1)^{1/2}} = \frac{1}{\beta} \ .$$

The constant C in (82) is:

$$C = \cos\varphi_{\max} - \frac{\omega\, mc}{eE}\,\frac{(1-\beta_{ph}^2)^{1/2}}{\beta_{ph}}\ ,$$

and finally

$$\cos\varphi - \cos\varphi_{\max} = \frac{\omega\, mc}{eE_0\,\beta_{ph}}\left[\gamma(1-\beta\beta_{ph}) - (1-\beta_{ph}^2)^{1/2}\right]\ . \tag{83}$$

Analysing Eq. (83), we assume at first that $v_{ph} < c$. For a given $\varphi_{\max}$, Eq. (83) represents a trajectory in the longitudinal phase plane. A plot of Eq. (83), with $v_{ph} < c$, is given in Fig. 16, where the relativistic factors of the right-hand side of the equation have been replaced by the kinetic energy of the electron $W = (\gamma - 1)\ mc^2$. The plot is centred on $\varphi = 0$, which has been assumed as the stable phase [compare Eq. (79) with Eq. (64)]. The reason why we have taken $\varphi_s = 0$ will become clear when analysing the case $v_{ph} = c$.

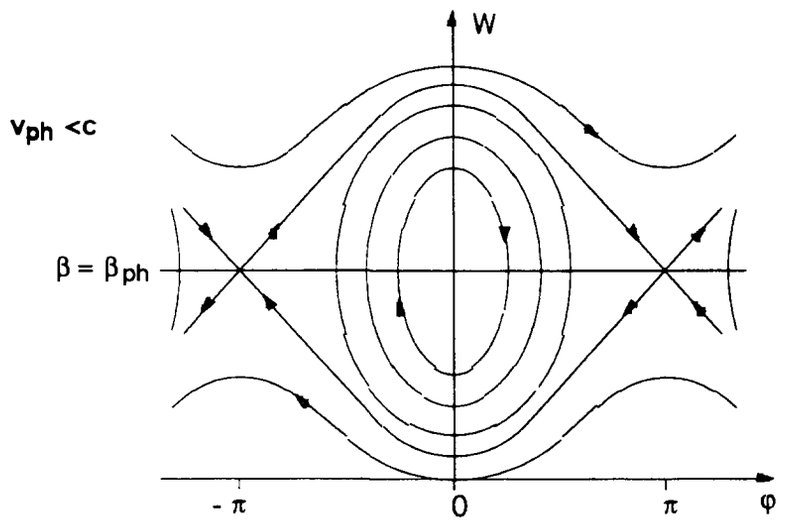

Fig. 16 Trajectories in the longitudinal phase plane (φ_s is counted from the zero of the wave in electron linacs)

Electrons rapidly gain highly relativistic velocities and therefore most electron linacs operate with $v_{ph} = c$. In this case $\beta_{ph} = 1$, and the particle trajectories cannot be closed curves any longer as (80) is always negative. This situation is shown in Fig. 17. Equation (83) becomes

$$\cos\varphi - \cos\varphi_{\max} = \frac{\omega\, mc}{eE_0}\gamma(1-\beta)\ . \tag{84}$$

Electrons at first lag behind the wave, but as they are accelerated and approach the velocity of light, they also approach asymptotically a phase of the wave, $\varphi_{\max}$. The best asymptotic phase corresponds to the crest of the wave: $\varphi_{\max} = -\pi/2$. If the electrons are injected at $\varphi = 0$, the amplitude of the accelerating space harmonic must be

$$E_0 = \frac{\omega\, mc}{e}\,\gamma(1-\beta)\ , \tag{85}$$

in order that the electrons remain trapped by the crest of the wave (see Fig. 17). To have a reasonable energy spread at the output of the accelerator the electrons should have a small phase spread around $\varphi_{max} = -\pi/2$. Electrons having too small an energy at $\varphi = 0$ are not trapped at all, because they will drift past the asymptotic phase (see Fig. 17).

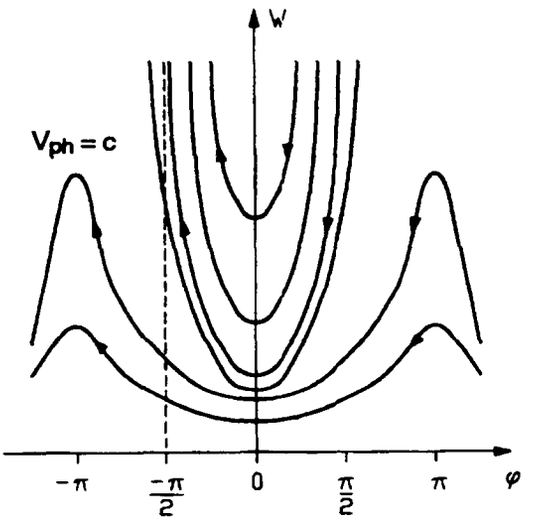

Fig. 17 Trajectories of ultra-relativistic particles in a linac with $v_{ph} = c$

6.2 Acceptance of an accelerator

It is obvious that for a good capture and acceleration the beam must be matched to the accelerator acceptance. Usually the beam emittances in the three phase planes are separately matched to the corresponding acceptances. In the longitudinal phase plane, the beam has to be bunched beforehand in order to have the required energy and phase spread at the input of the accelerator. Transversally, the beam emittances must be matched to an alternating-gradient focusing system. In fact, all linear accelerators (except the RFQ, as will be seen later) require an external focusing system. The focusing system is required:

i) to handle the finite beam emittance;

ii) to compensate for RF defocusing;

iii) to compensate for space charge forces in intense beams.

The RF defocusing can best be understood in a frame moving with the wave: one has thus an electrostatic case with

$$\mathrm{div}\vec{E} = \frac{\partial E_x}{\partial x} + \frac{\partial E_y}{\partial y} + \frac{\partial E_z}{\partial z} = 0 \ ,$$

and requiring $\partial E_z/\partial z < 0$ (longitudinal focusing for phase stability), the other terms being necessarily positive (defocusing).

The space-charge forces are important at lower energies. At higher energies, the magnetic self field focuses and compensates partly for the defocusing electric self field ($\propto 1 - \beta^2$). At $v_p = c$ there are no transverse self forces in the beam.

The acceptance of a linear accelerator in all phase planes is computed as if it had a periodic focusing system. Note that ion linear accelerators, where the cell length increases, are periodic if one considers t and not z as the independent variable. For intense beams, space-charge forces must somehow be included when quoting the acceptances. There are computer programs which compute the beam matching parameters and the accelerator acceptances in the presence of beam self forces.

7 HANDLING OF INTENSE BEAMS

In Section 6.1 we saw that with linear equations of motion the particle trajectories are ellipses. In the longitudinal plane the linear region is limited to the neighbourhood of the synchronous phase. In the transverse planes the external focusing forces are essentially linear.

The beam emittances in the phase planes are also specified by ellipses. To match a beam to the accelerator, the beam emittances must have the same shape as the trajectories and lie inside them.

If one deals with intense beams, space-charge forces must be taken into account. The charge density distribution in the beam is usually non-uniform, giving rise to non-linear self forces. The question now is the following: how can a non-uniform beam be matched to a structure with predominently linear forces?

It can be done. Sacherer [25] has demonstrated that the evolution of the root-mean-square (r.m.s) beam size essentially depends on the linearized part of the self forces. This is valid for beams of any charge distribution, provided they are of the ellipsoidal type, i.e. the isodensity curves are concentric ellipses. Hence, the r.m.s beam sizes and an emittance defined with r.m.s values can be matched. All beams having the same intensity, and the same r.m.s values behave in the same way. Therefore, once we know the r.m.s values of a real beam it can be replaced by an 'equivalent' one having the same r.m.s sizes, but a uniform density distribution.

The r.m.s beam sizes are computed via the second momenta of a normalized density distribution function f

$$\tilde{x} = (\overline{x^2})^{1/2} = \int \dots \int (x - \bar{x})^2 \; f(x, x', y, y', z, z') dx \; dy \; dz \; dx' \; dy' \; dz' \; , \tag{86}$$

where $\bar{x}$ is the first moment (average value), usually zero. The same is valid for other phase space coordinates. Note that in Eq. (86) the longitudinal phase plane is described with z and z', replacing $\Delta\varphi$ and ΔW. The important r.m.s envelope equation is

$$\tilde{x}'' + K(s)\tilde{x} - \frac{\varepsilon_{\text{rms}}}{\tilde{x}^3} - \frac{\overline{xF_s}}{\tilde{x}} = 0 \; , \tag{87}$$

where primes indicate the derivative with respect to the independant variable s (which indicates the position on the longitudinal axis) and

$$\varepsilon_{\text{rms}} = \left[\overline{x^2} \; \overline{x'^2} - (\overline{xx'})^2 \right]^{1/2} \; ,$$

is the r.m.s emittance ellipse and F_s the non-linear self force. Sacherer has shown that

$$\frac{\overline{xF_s}}{\tilde{x}^2} \; x = K_{sc}(s) x \; , \tag{88}$$

with $K_{sc}(s)$ being the linearized self-force constant. The r.m.s emittance was assumed constant. More recent studies allow for a change in $\varepsilon_{\rm rms}$ and describe mechanisms which influence it [26]. Emittance increases can be detected by multiparticle simulation programs. Between the r.m.s beam size and the overall one there is a fixed ratio, which depends on the distribution. For example, for a uniform and parabolic n-dimensional distribution this ratio is

$$\frac{\tilde{x}}{\hat{x}} = \left(\frac{1}{n+2}\right)^{1/2} \quad \text{and} \quad \frac{\tilde{x}}{\hat{x}} = \left(\frac{1}{n+4}\right)^{1/2} , \tag{89}$$

respectively, and for the emittances

$$\frac{\varepsilon_{\rm rms}}{E} = \frac{1}{n+2} \quad \text{and} \quad \frac{\varepsilon_{\rm rms}}{E} = \frac{1}{n+4} . \tag{90}$$

If the distribution changes, so do the ratios. Therefore, when designing an accelerator with r.m.s beam values, one must allow sufficient aperture for overall beam sizes.

To assess the importance of space charge in an accelerator, one uses the smooth beam-envelope equation. The smooth envelope is the average r.m.s beam size in a period, where the wiggle caused by the alternating gradient focusing is left out. In a transverse phase plane, the smooth envelope equation is

$$r'' + \overline{K} r - \frac{\varepsilon_{\rm rms}^2}{r^3} - \frac{k_{sc} I}{r} = 0 , \tag{91}$$

with $\overline{K}$ the average external focusing, k_{sc} the space charge proportionality factor, and I the beam current. For a matched beam

$$r'' = 0 . \tag{92}$$

In the absence of space charge, $I = 0$, one has with (91) and (92)

$$\overline{K} = \frac{\varepsilon_{\rm rms}^2}{r_0^4} = \overline{\Omega}_{\beta 0}^2 , \tag{93}$$

$\overline{\Omega}_{\beta 0}$ being the average betatron frequency without space charge. With space charge one has

$$\overline{K} - \frac{k_{sc} I}{r^2} = \frac{\varepsilon_{\rm rms}^2}{r^4} , \tag{94}$$

or

$$\overline{\Omega}_{\beta 0}^2 (1 - \mu) = \overline{\Omega}_{\beta}^2 , \tag{95}$$

where

$$\mu = \frac{k_{sc} I}{\overline{\Omega}_{\beta 0}^2 r^2}$$

is the space-charge factor, which must always be < 1, in order to have focusing. From (93) and (94) we see that

$$\overline{\Omega}_{\beta 0}^2 \propto \frac{1}{r^4} ,$$

while the space charge-force is only $\propto 1/r^2$. Therefore, μ is smaller for smaller radii, a result which might seem contradictory at first glance.

8 VARIOUS STRUCTURES OF LINEAR ACCELERATORS

In the analysis of the operation of linear accelerators, we have shown schematically some structures which have helped us to understand the concept of a 'loaded' cavity and how it functions. Now we shall present a few typical examples of existing accelerators in order to check our understanding and, eventually, extend it.

8.1 Superconducting LEP cavity

At the Large Electron-Positron storage ring (LEP) at CERN superconducting cavities are now being progressively installed. Thirty-two of these will be equivalent, as far as beam acceleration is concerned, to the 128 normally conducting cavities installed originally. Their power consumption (RF + cryogenics) is, however, six times smaller [27]. The cavities in a collider have to operate in a standing-wave mode as electrons and positrons travel in opposite directions. The frequency chosen for LEP cavities is 350 MHz.

Figure 18 shows schematically the four-cell superconducting LEP cavity in its cryostat. It is a disc-loaded cavity, operating in the TM_{010} cavity mode, while the structure and beam mode is π. One sees that the forms are smooth (compare Fig. 11), typical for superconducting cavities. In fact, smooth forms guarantee the best superconducting surface quality and permit their fabrication by spinning (half cell) and electron welding. The material is niobium or niobium sputtered on copper. Figure 19 shows a photo of the LEP Nb cavity.

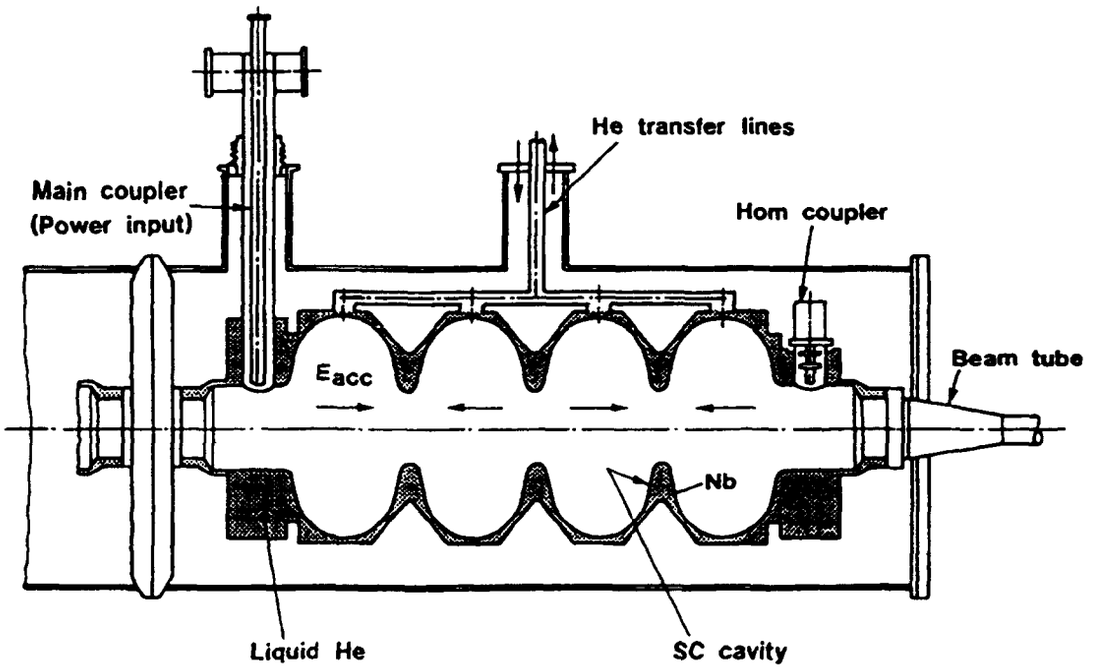

Fig. 18 Cross-section of the superconducting LEP cavity

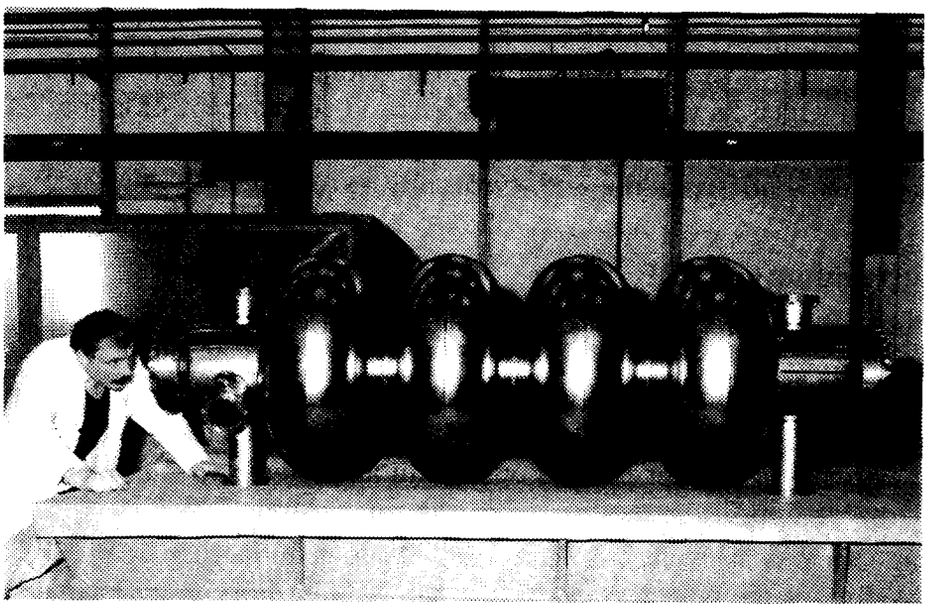

Fig. 19 LEP Nb cavity

With smooth forms, the ratio of the maximum electric field on the surface, E_s, to the average accelerating field on axis, E_0, is low:

$$\frac{E_s}{E_0} \cong 2 \ , \quad \text{with } E_0 \cong 6\text{MV/m} \ .$$

In normally conducting cavities, as in Fig. 11, this ratio is 5 to 6. The form in Fig. 11 is chosen in order to increase the transit-time factor $T(k)$; in superconducting cavities $T(k)$ is lower but, as dissipation losses are negligible, one can increase E_0. The quality factor Q_0, being about 5×10^9, is five orders of magnitude bigger than with normally conducting cavities. In the presence of beam (beam loading), the Q drops to about 10^6.

A cavity designer has to avoid too large surface fields, which could produce sparking and breakdowns. A good guideline is the Kilpatrick formula [28], which gives the critical conditions for breakdown. Kilpatrick's formula is usually brought into a form valid for most RF accelerators, where the critical field E_k is linked to the RF frequency:

$$f = 1.64 \ E_k^2 \ e^{-8.5/E_k} \ , \tag{96}$$

where f is in MHz and E_k in MV/m. The formula (96) is pessimistic for clean surfaces and short pulses, so $E_s \cong 2E_k$ is usually tolerated.

Coming back to Fig. 18, we see that the RF power is fed into the cavity by a coaxial line; the end of the line being open, the field in the cavity is electrically (capacitively) coupled to the electromagntic field in the line. At the other end of the cavity, there is a higher-order mode (HOM) coupler. In a cavity with a very high Q, various disturbing modes can be excited by the beam itself (a short bunch is like a 'delta' function containing all frequencies), and have to be damped. A HOM coupler, conveniently designed and placed, couples out the unwanted modes and dissipates their energy, leaving the fundamental mode undisturbed (the coupler appears as a short circuit for the operating mode).

8.2 CERN Linac 2

For the experimental programme with protons, the main CERN activity before the advent of LEP, Linac 2 is used to inject particles into the booster synchrotron, the first circular accelerator in the chain of CERN's synchrotrons for hadrons.

Linac 2 is a 200 MHz (exactly 202.56 MHz) Alvarez drift tube linear accelerator, accelerating protons from 0.75 up to 50 MeV. It is a standing-wave accelerator with a cavity mode TM_{010}, structure mode 0, and beam mode 2π. It is approximately 33 m long and composed of three cavities, called tanks, which are practically contiguous, see Fig. 20. The tanks accelerate to 10, 30, and 50 MeV, respectively.

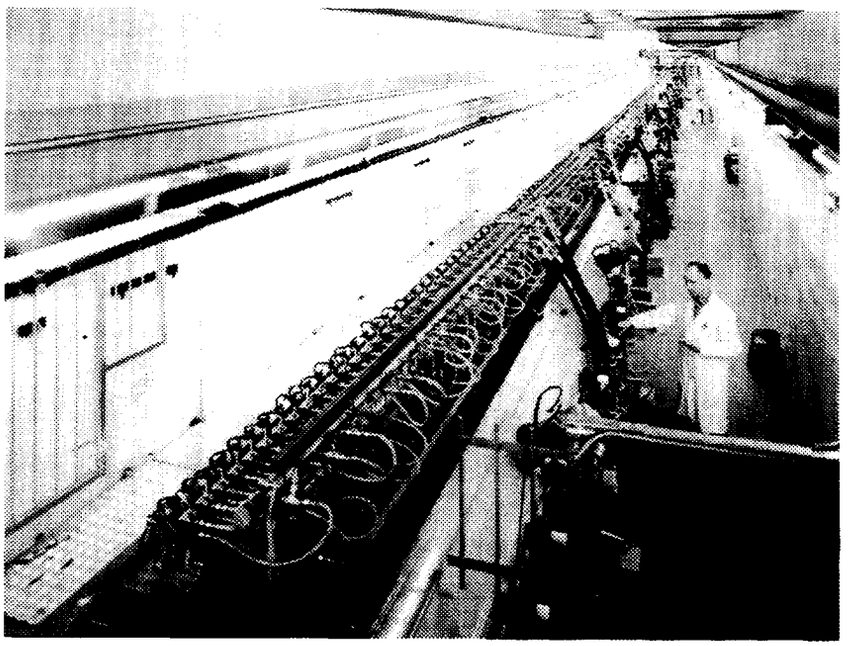

Fig. 20 CERN proton linear accelerator (Linac 2)

On the top of the tanks are girders, which hold the drift tubes, see Fig. 21. Each drift tube houses a magnetic quadrupole. As can be seen, the drift-tube length increases to correspond to the increased proton velocity along the accelerator. The cell length (see Fig. 13), increases in the same way, and so does the energy-gain-per-cell of the particle, according to formula (50).

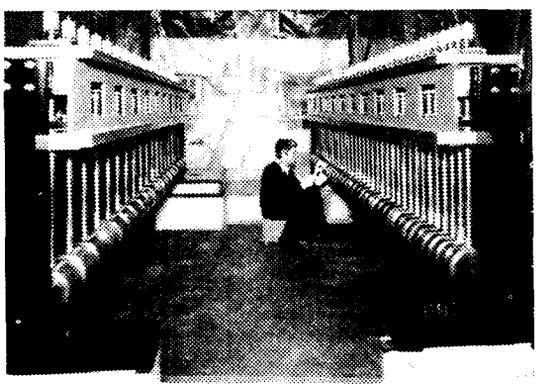

Fig. 21 Girders with drift tubes of Linac 2

The average field on the axis, E_0, is about 2 MV/m, except at the input to tank 1, where it is lower. One of the reasons for this is to keep the ratio of the energy gain to the kinetic energy of the particle small. This 'quasi adiabaticity' is necessary to maintain the beam matched. The effective shunt impedance, Z_{eff}, is about 40 MΩ/m, from mid tank 1 onwards. At lower energies, Z_{eff} drops to about half the above value. The synchronous phase, φ_s, varies from $-35°$ at the input of tank 1 to $-25°$ at the output of tank 1, and stays constant thereafter. With $\varphi_s = -25°$, the acceleration is more efficient, but the bucket size is reduced; however, due to phase damping [see formula (77)], the beam remains well inside.

Linac 2 accelerates 150 mA of protons in pulses of about 100 μs, and with a repetition rate of up to 2/s. The RF power fed into the cavity has to cover not only the dissipation losses in the structure, but must also compensate for the 'beam loading', which is very important in this case. In the 0 structure mode the energy cannot propagate along the cavity ($v_g = 0$), so the electromagnetic field configuration must get somewhat distorted in order to make $v_g \neq 0$. The distortion results in a droop of the electromagnetic field from the input coupler onwards, and in a phase shift [29], $\propto (\ell/\lambda)^2$, where ℓ and λ are the cavity length and the RF wavelength, respectively. This 'power-flow phase shift' is often disagreeable and one tries to avoid it. It has been avoided in Linac 2 and such structures, which are called stabilized structures, will be analysed in the next paragraph.

8.3 LAMPF side-coupled cavity linear accelerator

The standing-wave accelerators operate at 0 or π structure mode, where Z_{eff} is the largest. The region around the $\pi/2$ mode, where v_g is the largest, is not interesting as there the direct and reflected wave do not combine (see Fig. 8).

A cavity is composed of many cells and each cell can be regarded as an oscillator or resonator. All the resonators are coupled together, for example, via the aperture on the beam axis (electric coupling), and they form a passband as shown in Fig. 22 (curve a).

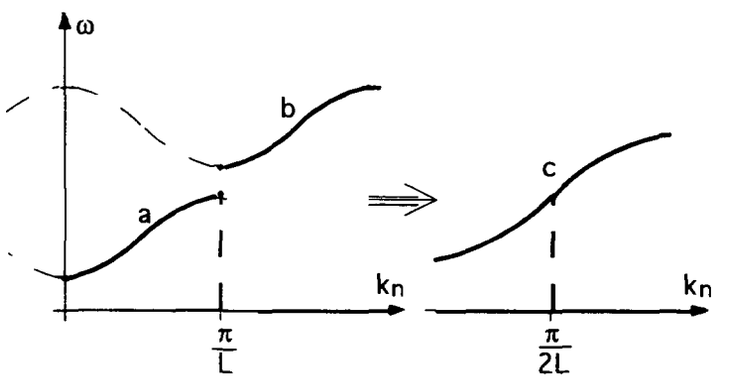

Fig. 22 Confluence of two passbands

One can imagine another chain of resonators, magnetically coupled and forming another passband, as shown in Fig. 22, curve b. If both chains are present in an accelerator (biperiodic chain), and if they are made to meet at $k_n = \pi/L$ (confluence), they will influence each other in such a way as to form a dispersion diagram according to curve c in Fig. 22. At the point of confluence, where v_g was originally zero, one now has a finite v_g.

Using the above principle [30], an accelerator was already constructed at the Los Alamos National Laboratory (LANL) in the late sixties. It is the 800 MHz 'Side-Coupled Cavity' linear accelerator (SCC) of the Los Alamos Meson Physics Facility (LAMPF). It accelerates protons from 100 MeV (an Alvarez accelerator becomes less efficient at higher energies), to the final energy of 800 MeV. Figure 23 shows schematically a part of the SCC Linac. One distinguishes 'accelerating cells' on the beam axis and 'coupling cells' (second passband) on the side. As far as the *beam* is concerned, the cavity operates in π mode, and the direct and reflected waves combine. For the propagation of *electromagnetic energy*, the biperiodic cavity operates in $\pi/2$ mode, and has a finite v_g. Biperiodic cavities are, in addition, less sensitive to constructional errors.

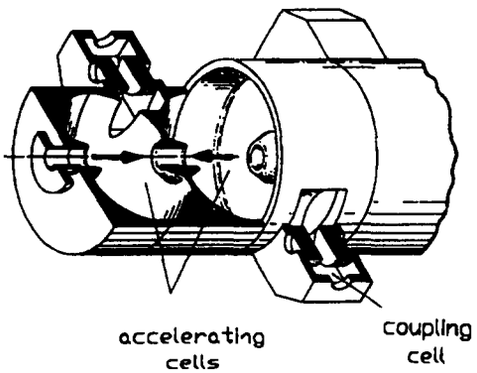

Fig. 23 Side-coupled linear accelerator at LANL

Biperiodic chains can also be made to have a confluence at 0 mode. This is the case with CERN Linac 2, where the second chain is formed by resonant coupling posts (see Fig. 24). The coupling posts face the drift tubes and are placed in the horizontal plane, perpendicular to the drift tube stems. The length of the post (inductance), and the distance to the drift tube (capacitance), are chosen to make the post resonant at the operating frequency. If there is no need for the RF power to flow, the posts are not excited. When RF power has to flow the posts are excited, but their fields are screened by the drift tubes and are not felt by the beam. The RF power is fed into the Alvarez cavity by a coaxial line terminated with a coupling loop, as shown schematically in Fig. 24. The magnetic field in the loop couples to the magnetic field in the cavity.

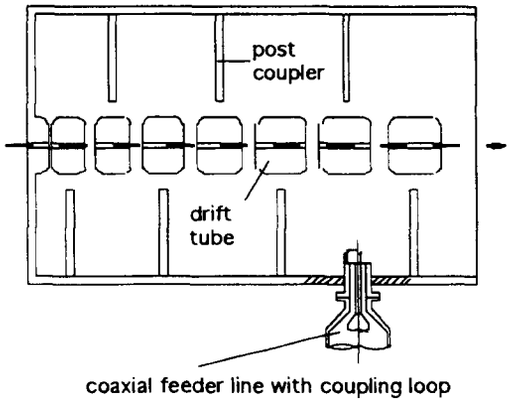

Fig. 24 Stabilized cavity with post couplers

8.4 Interdigital or IH structure at GSI

All the accelerators presented so far operate with electromagnetic fields of the TM type. Now we shall describe an accelerator operating with a TE type field, nominally TE_{110} [31]. Such a field type and mode in an empty cavity is shown in Fig. 25a. The transverse electric field in the horizontal plane must somehow be deviated into the longitudinal direction. Figure 26 shows how this is done in an interdigital or IH structure. The photo presents the open first tank of the HLI (Hoch Ladungs Injektor) linear accelerator for heavy ions of the Gesellschaft für Schwerionenforschung (GSI) at Darmstadt. The transverse electric field, short-circuited by the stems, exists only in the gaps between drift tubes. If we form a closed-line integral of the electric field, starting in the median horizontal plane and coming back on the lower half cylinder of the tank, we always enclose the same flux of the magnetic field. This means that the voltage between the drift tubes is always the same, irrespective of the cell length. This is a characteristic of the TE fields which makes them efficient at low particle velocities (Z_{eff} in the range of 200 MΩ/m, for a relativistic factor in the range $0.02 < \beta < 0.1$).

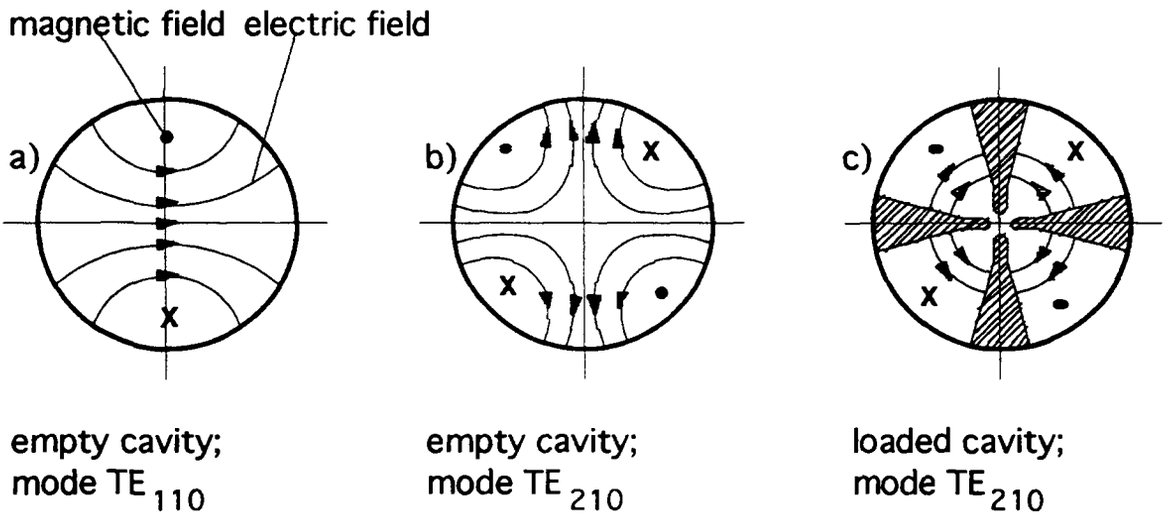

Fig. 25 Some TE field configurations

The HLI accelerator operates with a structure and beam mode of π. The drift tubes are small (less capacity, less stored energy, smaller dissipation), and do not contain quadrupoles. For focusing, there are special cylinders, which house quadrupole triplets (see Fig. 26). The first tank of HLI operates at 108 MHZ and accelerates particles from 0.3 to 1.4 MeV per atomic mass unit (MeV/u). A similar accelerator is being constructed for the CERN Lead Ion Project [32].

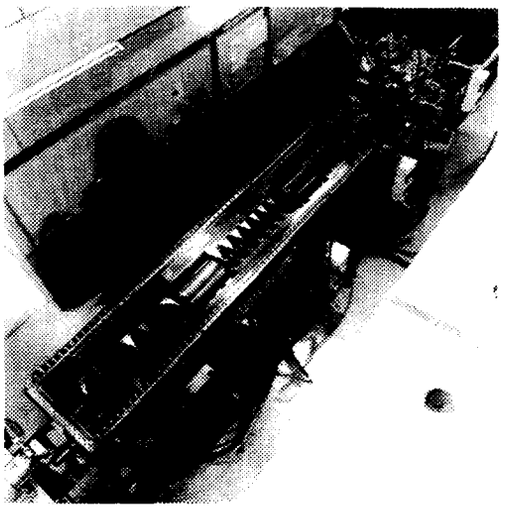

Fig. 26 First tank of the GSI IH linear accelerator

8.5 Radio frequency quadrupole (RFQ) linear accelerator

As a last example of a linear accelerator, we treat the very particular Radio Frequency Quadrupole (RFQ) accelerator which does not need external focusing because it performs all the required actions on the beam with the RF fields. The RFQ accepts a continuous beam, bunches it, focuses it, and accelerates it. A very good conventional bunching scheme places about 80% of the particles into an Alvarez bucket. An RFQ has a particle trapping efficiency of nearly 100%.

The RFQ was invented in the late sixties, in Russia, by I. M. Kapchinskiy [33] and V. Teplyakov. The LANL [34] has made important contributions to its design from 1980 onwards.

The operation of the RFQ can best be understood by considering a long electric quadrupole with a sinusoidally varying voltage on its electrodes. Particles moving along this quadrupole, and staying in it during several periods of the RF voltage, will experience an alternating gradient focusing action. If the electrodes, instead of being smooth, are modulated in the longitudinal direction, a certain part of the transverse field will also be deviated into the longitudinal direction. The modulation is made such that it is displaced by half a period between the horizontal and vertical electrodes (see Fig. 27). Evidently the length of the period is linked to the velocity of the particles (synchronism), and the RFQ is an accelerator of the TE type.

It is convenient to analyse separately the beam dynamics and electrodynamics properties of the RFQ [35].

8.5.1 Beam dynamics

The electrodes of the RFQ form a well-defined boundary along the beam axis. Due to the symmetry, the magnetic field on the axis is zero, and nearly zero in the close neighbourhood. Therefore, in this region we can replace the wave equation by the much simpler Laplace equation in cylindrical coordinates

$$\frac{1}{r}\,\frac{\partial}{\partial r}\left(r\,\frac{\partial U}{\partial r}\right)+\frac{1}{r^2}\,\frac{\partial^2 U}{\partial \vartheta^2}+\frac{\partial^2 U}{\partial z^2}=0\,, \tag{97}$$

with $U(r, \vartheta, z)$ being the electric field potential. The potential has, in addition, to be multiplied by a time factor, such as $\sin(\omega t + \varphi)$, to take care of its a.c. character. A general solution of (97) has the form [36]

$$U(r,\vartheta,z) = \frac{V}{2}\Big[\sum^{n} A_{on}\, r^{2n} \cos 2n\vartheta + \sum^{n}\sum^{\ell} A_{\ell n}\, I_{2n}\,(\ell k r)\, \cos 2n\vartheta \cos \ell k z\Big] , \qquad (98)$$

where

$\ell + n = 2p + 1, \quad p = 0, 1, 2, ...$
$\pm V/2$ is the electrode potential with respect to the axis
$I_{2n}(x)$ is the modified Bessel function of order $2n$
$k = 2\pi/\beta\lambda$, β is the relativistic factor and λ the wavelength.

The general solution contains all the harmonics in infinite series, but an RFQ is usually sufficiently well described with only a few harmonics. To illustrate this, we consider the lowest-order solution, which contains only two terms, one out of each infinite series:

$$U(r,\vartheta,z) = \frac{V}{2}\Big[A_{01}\, r^2 \cos 2\vartheta + A_{10} I_0(kr) \cos kz\Big] . \qquad (99)$$

The first term describes the potential of an electric quadrupole (focusing term); the second, containing $\cos kz$, is linked with the acceleration. Imposing the condition that the potential is constant on the electrode (see Fig. 27)

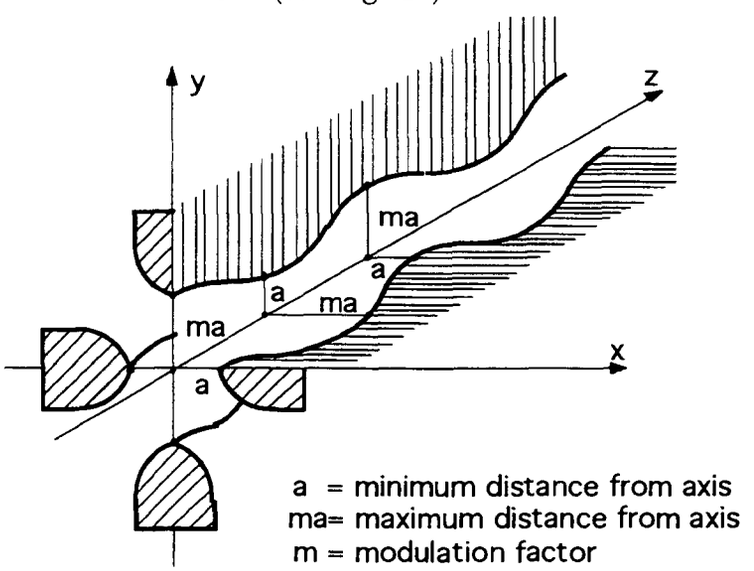

Fig. 27 Modulated vanes of an RFQ (schematic)

$$U(a,0,0) = U\left(ma, 0, \frac{\beta\lambda}{2}\right) = \frac{V}{2} ,$$

one gets

$$A_{10} = \frac{m^2 - 1}{m^2 I_0(ka) + I_0(mka)} , \qquad (100)$$

$$A_{01} = \frac{1}{a^2}\Big[1 - A_{10} I_0(ka)\Big] = \frac{\chi}{a^2} . \qquad (101)$$

The constants A_{01} and A_{10} are expressed by geometric parameters, a, the minimum aperture radius, and m, the modulation factor. By increasing m one gets more acceleration; by decreasing a one gets more focusing. From (101)

$$1 - A_{10} I_0(ka) = \chi \ ,$$

we get, after multiplying by V:

$$\chi V + A_{10} I_0(ka) V = V \ . \tag{102}$$

Equation (102) tells us that the interelectrode voltage V is composed of a part required for focusing (χV), and a part required for acceleration $[A_{10} I_0(ka) V]$. As the electrode modulation can be produced in a very precise manner (e.g. by a milling machine), we can obtain a precise and controlled action on the beam. The RFQ can handle very intense beams, where one tries to keep the beam dimensions constant. In this case, the self forces are constant, and so are the RF forces.

The field components in cylindrical coordinates, derived from (99), are

$$\begin{aligned} E_r &= -\frac{\partial U}{\partial r} = -\frac{V}{2}[2A_{01}\, r\, \cos 2\vartheta + k A_{10} I_1(kr) \cos kz] \\ E_\vartheta &= -\frac{1}{r}\frac{\partial U}{\partial \vartheta} = V A_{01} r \sin 2\vartheta \\ E_z &= -\frac{\partial U}{\partial z} = \frac{V}{2} k A_{10} I_0(kr) \sin kz \ . \end{aligned}$$

The RFQ, being a TE type accelerator, is efficient at low β (from a few keV per atomic mass unit to a few MeV).

As already mentioned, the electrodes have to be shaped so as to produce fields required by the beam dynamics. From (99), one obtains for the electrode surface

$$S(r, \vartheta, z) = A_{01} r^2 \cos 2\vartheta + A_{10} I_0(kr) \cos kz = \pm 1 \ . \tag{103}$$

8.5.2 Electrodynamics

For the beam dynamics it was sufficient to consider a small region around the axis. For the electrodynamics one has to consider the whole cavity (resonator), and solve the wave equation. There are several types of RFQ cavities, and we shall consider the 'four-vane' type introduced by the LANL [34].

The RFQ operates with the TE_{210} field configuration, shown schematically, in Fig. 25b (empty cavity) and Fig. 25c (cavity loaded with vanes). The vane modulation, essential for acceleration, has little effect on the cavity as a whole. This can be understood by Fig. 28, which shows the inside of the CERN RFQ1 accelerator, operating at 202.56 MHz, and accelerating protons from 50 keV to 520 keV. One can see that only the vane tips are modulated. The modulation corresponds to a structure and beam mode of π. To get an idea of the RFQ dimensions, Fig. 29 shows the cavity of the CERN RFQ2 accelerator, which is now the injector into Linac 2. The length of the RFQ2 is about 1.75 m and it nominally accelerates 200 mA of protons, from 90 to 750 keV. With the

advent of the RFQ2, one could eliminate the high d.c. voltage installation (750 kV), and the bunching and focusing system that existed before at the input to Linac 2.

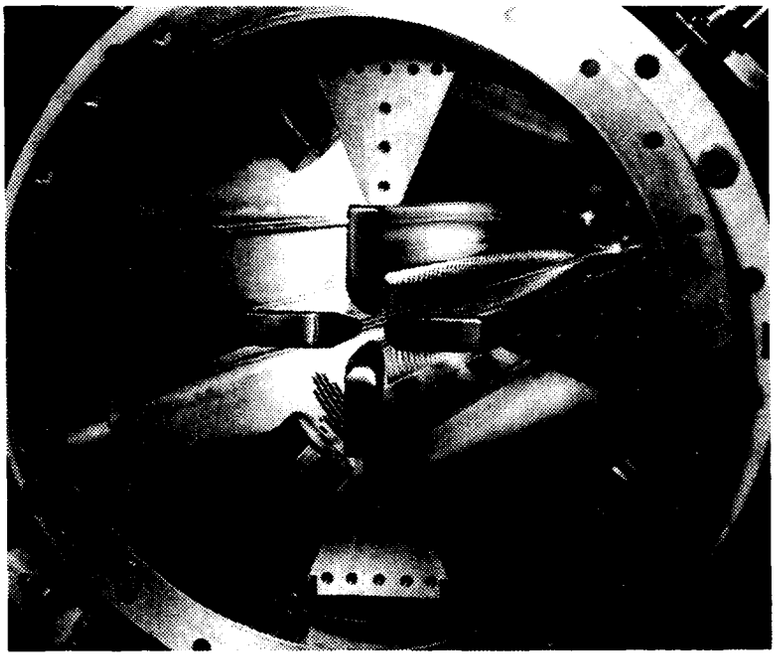

Fig. 28 Inside view of the CERN RFQ1

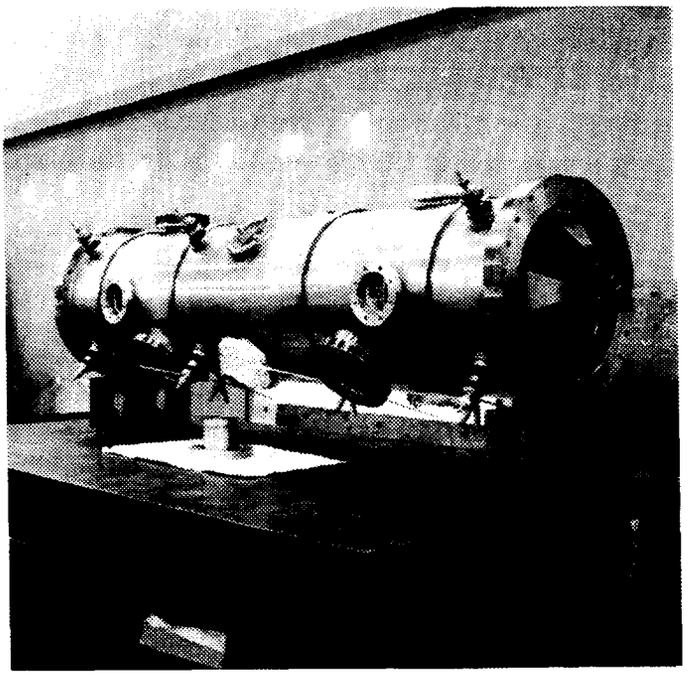

Fig. 29 Cavity of the CERN RFQ2

9 CONCLUSION

The basic concepts underlying how RF linear accelerators function have been presented. This paper, as the title indicates, is only an introduction, where the subjects have been treated to facilitate further study. The beam dynamics equations in cavities have been dealt with in the correct way, but with some simplifications, as have the theories of coupled oscillators and intense beams. By including design principles and examples of existing linear accelerators in this paper, it is hoped that enough material is presented in order to understand what is involved when dealing with linacs.

ACKNOWLEDGEMENTS

This paper has profited from discussions with G. Dôme, T.P. Wangler and in particular P.M. Lapostolle. D.J. Warner has read and commented on the manuscript. To all of them go my thanks.

References

[1] C. Ramo, J.R. Whinnery and T. van Duzer, Fields and waves in communication electronics (John Wiley, New York, 1965).

[2] J.C. Slater, Microwave electronics (D. van Nostrand, New York, 1951), p. 395.

[3] Ibid., pp. 169–177.

[4] G.A. Loew and R.B. Neal, Accelerating structures, *in* Linear Accelerators, by P.M. Lapostolle and A.L. Septier (North Holland, Amsterdam 1970), pp. 47–52.

[5] G. Dôme, Review and survey of accelerating structures, *in* Linear Accelerators (see Ref. [4]).

[6] G.A. Loew and R.B. Neal, Accelerating Structures, *in* Linear Accelerators (see Ref. [4]), pp. 52–65.

[7] P. Lapostolle, Introduction à la théorie des accélérateurs linéaires, CERN 87–09 (1987), pp. 43–44.

[8] J. Le Duff, Dynamics and acceleration in linear structures, Proc. CERN Accelerator School, Gif-sur-Yvette, 1984, CERN 85–19 (1985), p. 146.

[9] P.M. Lapostolle, Introduction à l'étude des accélérateurs linéaires, CERN 66–20 (1966).

[10] P.M. Lapostolle, Equations de la dynamique des particules dans un accélérateur linéaire à protons, Int. Rep. CERN AR/Int. SG/65–11.

[11] K. Halbach and R.F. Holsinger, SUPERFISH-A computer program, Part. Accel. **7** (4), pp. 213–222 (1976).

[12] T. Weiland, On the computation of resonant modes in cylindrically symmetric cavities, Nucl. Instrum. Methods **216**, pp. 329–248 (1983).

[13] R. Klatt et al., MAFIA-A three dimensional electromagnetic CAD system, Proc. Linear Accelerator Conference, 1986, SLAC report 303, pp. 276–278.

[14] M. Martini and D.J. Warner, Numerical calculations of linear accelerator cavities, CERN 68–11 (1968).

[15] B. Schnizer, General properties of fields and beam dynamics in a linac gap, CERN 69–3 (1969).

[16] M. Promé, Effets de la charge d'espace dans les accélérateurs linéaires à protons, Thèse, Orsay No. 761 (1971).

[17] K. Crandall, Documentation for PARMULT, Memo AT–1: 92–151, LANL, Los Alamos.

[18] B. Bru, Calcul de la focalisation quadripolaire d'un linac et de l'adaptation du faisceau en présence de charge d'espace, Int. Rep. MPS/LIN 70–10 (1970).

[19] G.P. Boicourt and J. Merson, PARMILA Users and Reference Manual, LA–UR–90–127, LANL, Los Alamos.

[20] G.A. Loew, R.H. Miller and R.A. Early, Computer calculations of travelling wave periodic structure properties, IEEE Trans. Nucl. Sci., **NS–26**, pp. 3701–3704 (1979).

[21] K. Crandall, PARMELA, Computer code available from AT–6, H829, LANL, Los Alamos.

[22] K.L. Brown et al., Transport - A computer program for designing charged particle beam transport systems, CERN 80–04.

[23] M. Weiss, A short demonstration of Liouville's Theorem, Proc. CERN Accelerator School, Aarhus, Denmark 1986, CERN 87–10 (1987).

[24] R.H. Helm and R. Miller, Particle dynamics, *in* Linear Accelerators (see Ref. [4]), pp. 115–120.

[25] F. Sacherer, RMS envelope equation with space charge, Int. Rep. CERN/SI/Int. DL/70–12.

[26] T.P. Wangler, K.R. Crandall, R.S. Mills and M. Reiser, Field energy and r.m.s. emittance in intense particle beams, Workshop on High Brightness, San Diego, AIP Conf. Proc. **139** (1986), p. 133.

[27] Ph. Bernard, CERN, private communication.

[28] W.D. Kilpatrick, Criterion for vacuum sparking designed to include both RF and d.c., Rev. Sci. Instrum., **28** (1957).

[29] H.G. Hereward, Some examples of energy flow in the Alvarez structure, Int. Rep. CERN MPS/DL Int. 65–1 (1965).

[30] E.A. Knapp, High energy structures, *in* Linear Accelerators (see Ref. [4]), pp. 601–616.

[31] U. Ratzinger et al., The Upgraded Munich Linear Heavy Ion Postaccelerator, Nucl. Instrum. Methods **A263** pp. 261–270 (1988).

[32] D.J. Warner (ed.), CERN Heavy-Ion Facility Design Report, CERN 93–01 (1993).

[33] I.M. Kapchinskiy, Strahldynamik in den Sektionen mit Räumlich Uniformer Fokussierung, GSI–tr–14/76 (Translation from Russian).

[34] R.H. Stokes et al., The radio frequency quadrupole, Proc. of 11th Int. Conf. on High Energy Accelerators, Geneva, 1980 (Birkhauser, Basle, 1980).

[35] M. Weiss, Radio frequency quadrupole, CERN Accelerator School, Aarhus, Denmark, 1986, CERN 87–10 (1987), pp. 196–230.

[36] C. Biscari, Computer programs and methods for the design of high intensity RFQs, Int. Rep. CERN/PS 85–67 (Li).