

Differences between electron and ion linacs

M. Vretenar

CERN, Geneva, Switzerland

Abstract

The fundamental difference between electron and ion masses is at the origin of the different technologies used for electron and ion linear accelerators. In this paper accelerating structure design, beam dynamics principles and construction technologies for both types of accelerators are reviewed, underlining the main differences and the main common features.

1 Historical background

While the origin of linear accelerators dates back to the pioneering work of Ising, Wideröe, Sloan and Lawrence in the decade between 1924 and 1934 [1], the development of modern linear accelerator technology starts in earnest in the years following the end of World War II, profiting of the powerful generators and of the competences in high frequencies developed for the radars during the war. Starting from such a common technological ground two parallel developments went on around the San Francisco Bay in the years between 1945 and 1955. At Berkeley, the team of L. Alvarez designed and built the first Drift Tube Linac for protons [2], while at Stanford the group of E. Ginzton, W. Hansen and W. Panofsky developed the first disc-loaded linac for electrons [3]. Following these first developments ion and electron linac technologies progressed rather separately, and only in recent years more ambitious requirements for particle acceleration as well as the widening use of superconductivity paved the way for a convergence between the two technologies. The goal of this paper is to underline the basic common principles of ion and electron RF linacs, to point out their main differences, and finally to give an overview of the design of modern linacs.

2 Special relativity, particle velocity and synchronicity

The basic difference between electrons and ions is their rest mass. The mass of an electron at rest is 9.1×10^{-31} kg (corresponding to $511 \text{ keV}/c^2$), while a proton has a mass of 1.6×10^{-27} kg ($938 \text{ MeV}/c^2$), i.e. a factor about 2'000 (exactly, 1'836) difference. Heavy ions used in linear accelerators have even greater masses.

Because of their difference in mass, electrons and ions will react differently to an externally applied electric field. An RF linear accelerator is basically a linear array of small space intervals ('gaps') where a time varying electric field has been generated. The field can be made synchronous with a particle beam in such a way that the particles see in each gap a field in the same longitudinal direction. The change in momentum of a particle inside the beam is determined by the well-known dynamics equation:

$$q[\vec{E} + (\vec{v} \times \vec{B})] = \frac{d}{dt}(m\vec{v})$$

In absence of a magnetic field, the force produced on the particle by the electric field increases its momentum $m\vec{v}$. The velocity however can not increase indefinitely, being limited to the speed of

light c . When the velocity of a particle approaches c , its momentum will keep increasing, the velocity will remain $v \approx c$ while its relativistic mass m will increase, accordingly to the relation

$$m = m_0 / \sqrt{1 - (v/c)^2} = m_0 / \sqrt{1 - \beta^2} = m_0 \gamma$$

where $\beta = v/c$, $\gamma = (1 - \beta^2)^{-1/2}$ and m_0 is the Newtonian mass (mass in the rest frame).

To compute how the velocity increases during acceleration, we have to recall that the kinetic energy W of a particle is equal to its relativistic total energy minus the rest energy:

$$W = mc^2 - m_0c^2 = m_0c^2(1/\sqrt{1 - \beta^2}) - m_0c^2 = m_0c^2(1/\sqrt{1 - \beta^2} - 1)$$

This relation can be used to express the relativistic velocity $\beta = v/c$ as a function of kinetic energy:

$$\beta^2 = 1 - 1 / \left(\frac{W}{m_0c^2} + 1 \right)^2 \quad (1)$$

For small energies ($W \ll m_0c^2$), β follows the classical relation for kinetic energy:

$$\beta^2 = \frac{2W}{m_0c^2} \quad (2)$$

Plotting relation (1) for electrons ($m_0c^2 = 511 \text{ keV}$) and for protons ($m_0c^2 = 938 \text{ MeV}$) we obtain the curves of Fig. 1.

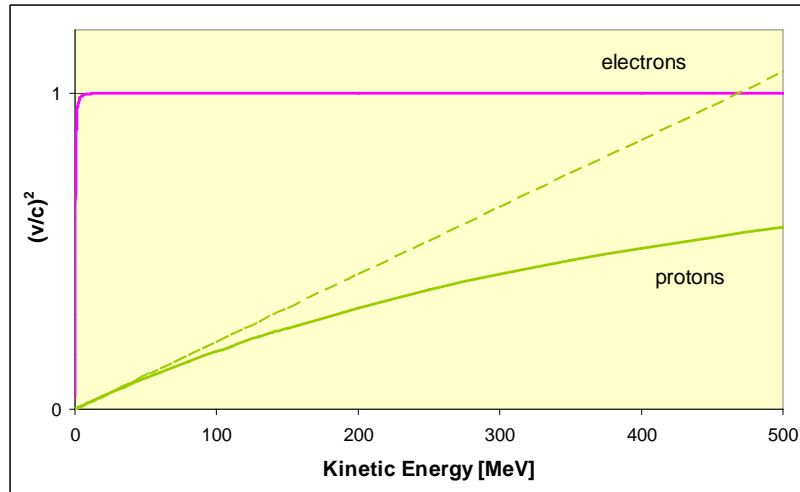


Fig. 1: Relativistic particle velocity (squared) as function of kinetic energy (protons and electrons)

In this plot the horizontal scale shows the beam kinetic energy, but it could easily be interpreted as the distance along a linear accelerator: the energy in a linac increases more or less linearly with the length (typical gradients are 10 to 15 MeV/m for electron linacs and 2 to 10 MeV/m for proton linacs).

The electrons (upper curve) come close to the speed of light already after few MeV of energy, corresponding to about the first meter of acceleration. For the rest of the acceleration process, their

velocity will not change. This already suggests that an electron linac will be made of an injector covering the initial distance, where the velocity rapidly increases to c , followed by identical accelerating structures covering the rest of the accelerator length.

Instead, in a proton linac (lower curve) the particles will increase slowly their velocity, initially following classical mechanics [dotted curve, corresponding to equation (2)] up to an energy of a few MeV. As the energy increases further, the mass starts to increase, accordingly to relativistic mechanics, while the velocity increases more and more slowly. Only from an energy of a few GeV the velocity saturates towards c , well beyond the energy range of most proton linacs. We can already conclude that a proton linac (and a heavy ion linac) will have to adapt to the increase in particle velocity, and will therefore be made of sections of different type.

The need to adapt the accelerating structure to the velocity of the accelerated beam comes from the condition of synchronicity between the particles and the accelerating fields, required to have a maximum transfer of energy to the particles.

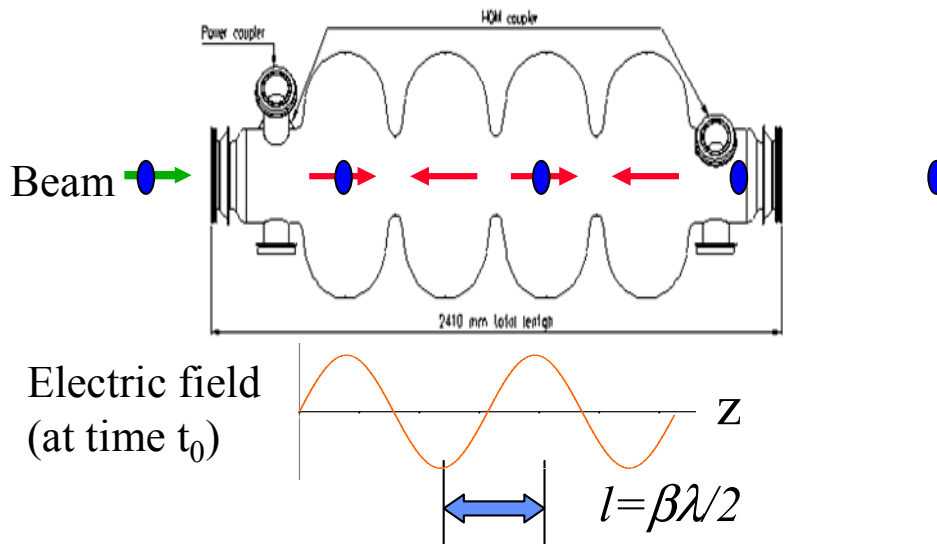


Fig. 2: Synchronism condition for a 4-cell accelerating cavity (LEP-II)

As an example, we can consider the 4-cell superconducting accelerating cavity of Fig. 2, which can be used in a high-energy electron or proton linac. The cavity is operated in the so-called π -mode, meaning that the fields have a phase difference of 180° between subsequent cells. The arrows in Fig. 2 represent the orientation of the electric field on the axis at a certain time, accordingly to the sinusoidal field distribution shown in the lower part of the figure. The curve and the arrows represent the field at a given time; the fields will oscillate in time at the frequency of the cavity, and after half of the RF period ($T/2$) the direction of the field in the gaps will be reversed. The particles enter the cavity already grouped in bunches (the “bunching” is performed in the initial section of a linac), spaced exactly by an RF period T . The maximum energy gain in a multi-cell structure corresponds to the condition that a particle inside the bunch travels from one gap to the next in a time equal to $T/2$. In this case the particle will see an accelerating field in all the gaps. For a particle travelling at the relativistic velocity β , the time to cross one cell is $\tau = l/\beta c$. The condition $\tau = T/2$ gives the required cell length for a π -mode structure:

$$l = \frac{\beta c}{2} T = \frac{\beta \lambda}{2}$$

From this synchronism condition, it is clear that an electron linac, where the particles have $\beta=1$ apart from a short initial section, will be made of a sequence of identical accelerating structures, with cells of the same length $l=\lambda/2$. The case of the proton linac is totally different, because β increases slowly by orders of magnitude, and the cell length has to follow a precise β profile. For example, in the CERN proton linac (Linac2) the value of β and of the basic cell length increase by about a factor 200 between the injection at 90 keV and the end of the linac at 50 MeV. In order to keep the synchronism without using too short (mechanically difficult) or too long (inefficient) accelerating cells, a proton or ion linac usually includes different types of structures, operating on different modes and sometimes at multiple frequencies.

3 Accelerating structures

The accelerating structures for electrons and for ions are usually different, but can be derived from the same fundamental RF element, the cylindrical waveguide. This is a simple metallic pipe, inside which we can excite, by means of an appropriate coupler, an RF wave of given frequency and wavelength (Fig. 3).

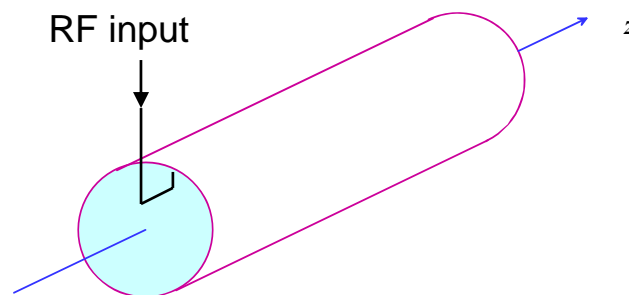


Fig. 3: The cylindrical waveguide

In order to use this structure for the acceleration of particles, we need to excite a longitudinal electric field on the axis, which can then transfer energy to the beam. From basic RF theory [4] we know that in such a waveguide can be excited well defined “modes” (electromagnetic field configurations), at frequencies beyond a “cut-off” frequency characteristic of each mode. Modes with longitudinal electric field on axis are in the family of “TM-modes” (Transverse-Magnetic). The main candidate to be used for particle acceleration is the simplest TM mode, the TM₀₁, whose field configuration (i.e. a snapshot taken at a certain time t) is shown in Fig. 4. The mode will propagate in the waveguide, meaning that the field pattern of Fig. 4 will travel in the z direction, and the position of field maxima and minima will continuously change with time. In this respect, this travelling wave mode is different from the standing wave mode of Fig. 2, which will be analysed later in this section.

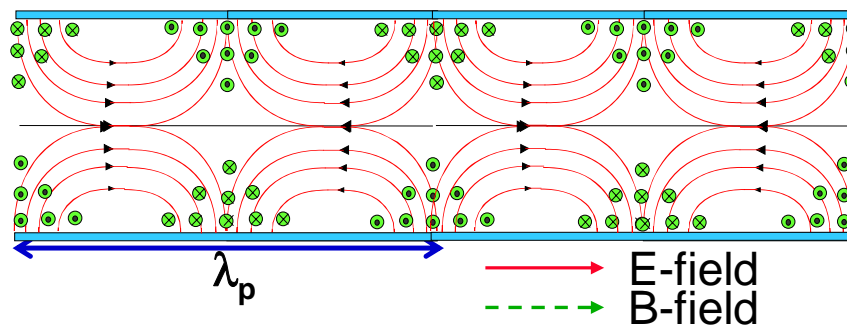


Fig. 4: TM₀₁ mode in a cylindrical waveguide

The wavelength in the direction of propagation λ_p (distance between two subsequent maxima or minima of the field) depends on the dimensions of the waveguide and on the excitation frequency. This can be understood remembering that the propagation within the waveguide is the result of multiple reflections of the wave on the metallic walls of the cylindrical pipe. High frequencies will have short wavelengths, and at the limit of wavelengths much smaller than the dimensions of the waveguide their propagation will not be influenced by the waveguide (the wave does not “see” the waveguide). On the contrary, free space wavelengths comparable to the size of the waveguide will be influenced by the guide and at the limit of very long wavelengths (or small frequency) the wave will not be able to propagate. This behaviour is summarised by the so-called “dispersion curve” of the waveguide, shown in Fig. 5, which can be calculated solving the wave equations in the bounded medium.

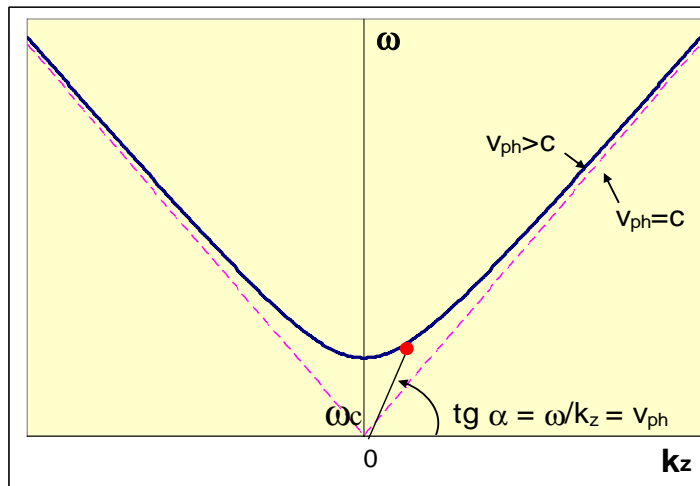


Fig. 5: Dispersion curve of a generic waveguide

The dispersion curve completely characterizes the behaviour of the waveguide. It gives the basic relation between the frequency of the propagating wave and its propagation constant k_z (also called wave number) along z , $k_z = 2\pi/\lambda_p$, which is in turn related to the propagation wavelength, and allows to calculate the propagation velocity. In this context, $\varphi = k_z z$ is the difference in phase of the wave between two points at distance z along the direction of propagation. The straight lines correspond to the free space case, $\omega = 2\pi c/\lambda$, and at high frequencies the dispersion curve tends to the free space curve. Waves with frequencies below cut-off ($\omega < \omega_c$) cannot propagate, with ω_c depending only on the transverse dimensions of the waveguide. The phase velocity of a propagating wave at frequency ω , defined as the apparent velocity of a maximum (or minimum) of the wave in the z direction, is:

$$v_{ph} = \lambda_p/T = \omega/k_z$$

From the curve of Fig. 5, we can see that phase velocity at ω is the tangent of the angle between the corresponding point on the curve and the k -axis. Above cut-off, phase velocity decreases from $v_{ph} = \infty$ at cut-off towards $v_{ph} = c$, in the limit case of $\omega \rightarrow \infty$ (free space). Since the phase velocity is always higher than the speed of light we have to conclude that we can never achieve synchronism between the wave and a particle beam. This means that we cannot accelerate particles in a standard cylindrical waveguide. It is important to observe that the phase velocity is only an “apparent” quantity, and there is no RF power or any information traveling at this velocity. This is why phase velocity can be higher than c , without violating the relativity principle. Real quantities, like the energy or the information carried by variations of the wave will travel at the “group velocity” $v_g = d\omega/dk_z$,

represented by the tangent to the dispersion curve of Fig. 5. It is easy to see that group velocity will range from $v_g = 0$ at cut-off to $v_g = c$ in the limit of very high frequencies.

In order to use the waveguide for particle acceleration we have to “slow down” the wave, i.e. to find a geometrical modification which can bring for some frequencies the phase velocity down to c or below. Considering that wave propagation occurs via multiple reflections, the most obvious way to produce a “slow-wave” structure is by adding to the cylindrical waveguide some discs, as in Fig. 6. It is first of all important to notice that we have now a “periodic” structure, with period (cell length) l , and very likely the waves most affected by the discs will have propagation wavelength equal to multiples of l : $\lambda_p \sim l$.

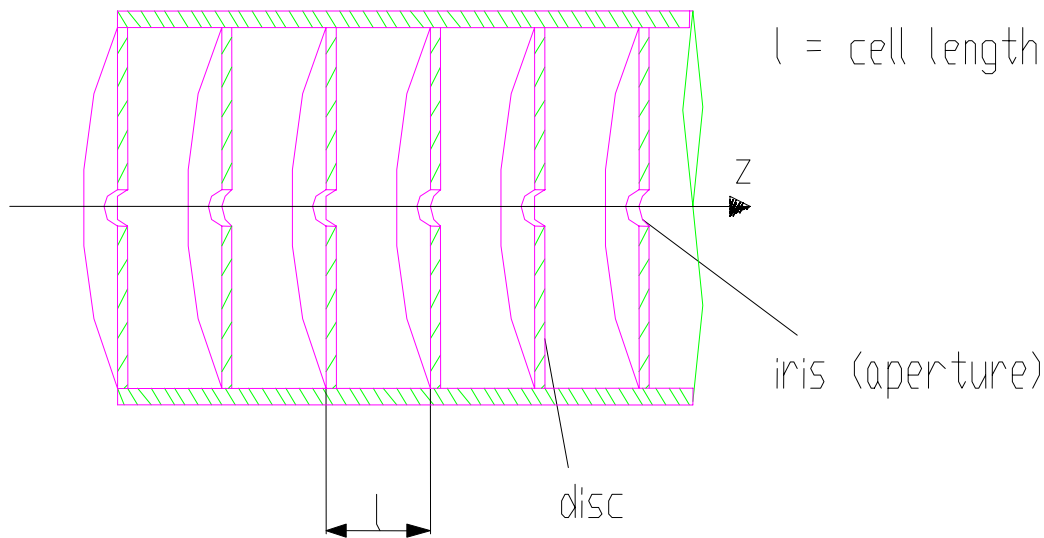


Fig. 6: Disc-loaded waveguide

On the contrary, waves with $\lambda_p = \infty$ or $\lambda_p = 0$ will not “see” the discs: for $\lambda_p = \infty$ (cut-off) the discs are a small perturbation, while for $\lambda_p = 0$ the wave is confined between the cylinder walls and does not interact with the discs. Therefore, we can already observe that the dispersion curve of the disc-loaded system will be identical to that of the standard cylindrical waveguide at the two limits $k \sim 0$ and $k \rightarrow \infty$. Instead, when the distance between discs is equal to half the propagation wavelength ($k_z = 2\pi/2l$), we can expect that the wave is confined between two discs, and the cell behaves like a small resonator. What actually happens for $\lambda_p = l/2$ ($k_z = \pi/l$) is that two frequencies are possible, corresponding to the two solutions shown on the left side of Fig. 7. In one case, the electric field is heavily perturbed by the discs, and the frequency will be lower than for the standard waveguide. In the other case, the magnetic field is perturbed, and the frequency is higher. For continuity, we can now draw the dispersion curve of the disc-loaded waveguide (right side of Fig. 7). The dispersion curve is split in two branches, separated by a “stop-band”. Observing the lower branch of the dispersion curve, it is clear that at a certain frequency it will cross the “free space” curve corresponding to $v = c$: all frequencies above this one will present a phase velocity $v < c$. The conclusion is that a disc-loaded waveguide presents a range of frequencies with phase velocity $v \leq c$, and can be used for particle acceleration. The range of possible frequencies depends on the diameter of the waveguide and on the distance between discs.

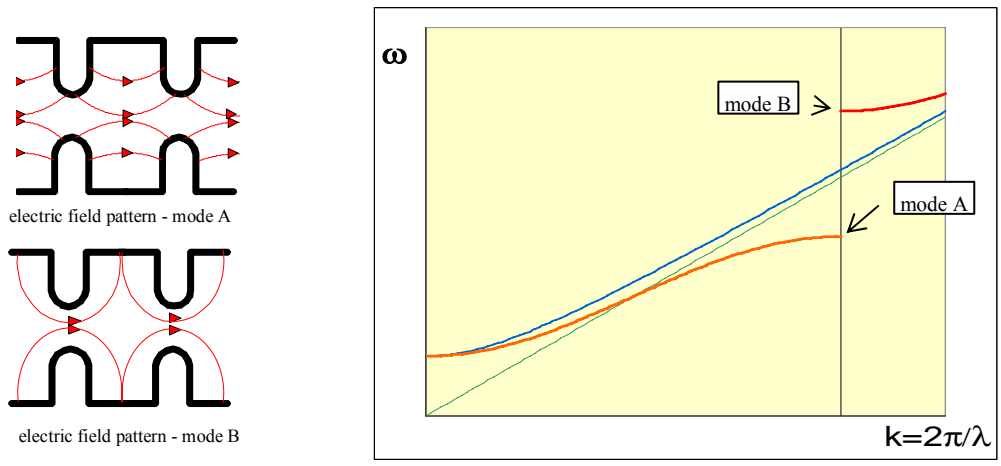


Fig. 7: Electric field distribution of $k = \pi/l$ modes and dispersion curve for a disc-loaded waveguide

The travelling-wave linac generally used for electrons at $\beta = 1$ (Fig. 8) is a disc-loaded structure designed to have phase velocity $v = c$ for a given operating frequency. It is equipped with an input and output coupler, to inject the RF wave coming from a power source, and to extract at the end of the structure the remaining RF power. Electrons at $\beta = 1$ entering the structure when the electric field is at its maximum (i.e. the beam needs to be “bunched” and to be injected at the correct phase) will travel along the linac with the same velocity as the wave, seeing the maximum accelerating electric field all along the structure. The result will be an increase of their kinetic energy. The beam energy comes from the RF wave: part of the RF energy of the wave will be dissipated in the structure walls, usually made out of copper, part will be absorbed by the beam, and the rest will be extracted via the output coupler, to be absorbed in a matched load at the end of the structure. Usually, the length of a standard linac structure is such that about 30% of the input power goes to the load.

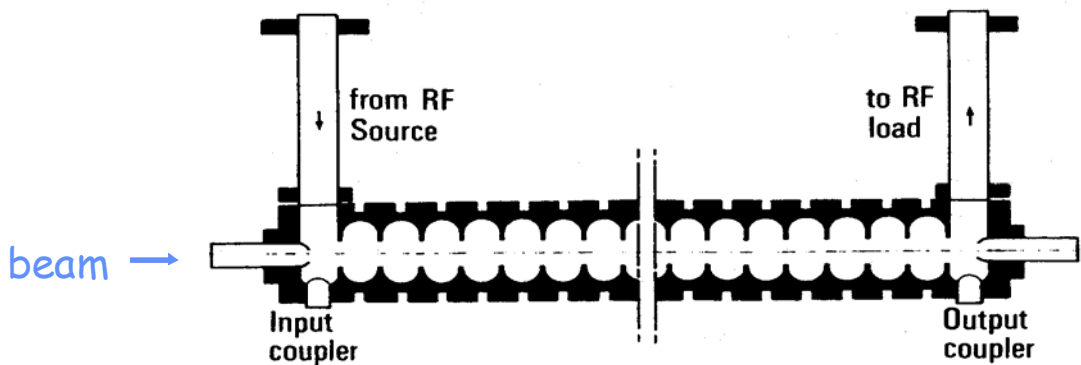


Fig. 8: Travelling-wave linac structure

Travelling-wave structures are designed for a fixed phase velocity, and can not be used for ions at $\beta < 1$ and increasing velocity. Even if the structure could be designed for a phase velocity corresponding to the initial β of the particles, when gaining energy the beam velocity would increase and the synchronism with the wave would be lost. However, up to a certain limit a travelling-wave structure designed for $\beta = 1$ can accept an electron beam with $\beta < 1$ and rapidly increasing velocity, as we will see in the following.

For protons and heavy ions, instead of matching the phase velocity of a travelling wave to the particle velocity, we can realize the synchronism between moving particles and field by keeping the phase of the wave constant, and changing its space distribution accordingly to the synchronism condition. This is obtained in standing-wave structures, which can be basically described as travelling-wave structures closed at the two ends with metallic walls (Fig. 9). We have already seen an example of standing-wave structure in the case of the superconducting cavity of Fig. 2.

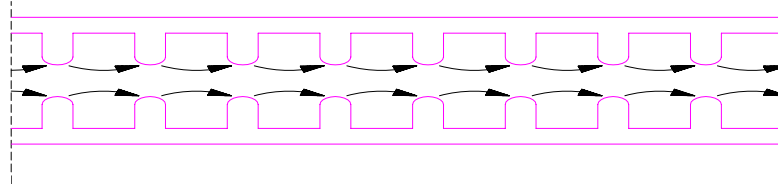


Fig. 9: Standing-wave linac structure (with an example of electric field distribution)

A wave injected into this structure will be reflected by the metallic walls at the two ends, and the resulting field distribution will be the result of the superposition of two waves travelling in opposite directions. For each of these two waves, the dispersion curve of the loaded waveguide (Fig. 7) is still valid, but now not all the frequencies on the curve will be possible. For most of the excitation frequencies, the phase difference of the two superimposed waves, which depends only on the distance between the two end walls, will continuously change, and the sum of these waves will be zero. Only for frequencies where the phase of the waves at each reflection are identical (or 180° apart) the interference will be constructive and the two waves will build up a field inside the structure. The condition for the establishment of a “resonant mode” is that the length of the structure L contains an integer number of half-wavelengths: $L = n\lambda_p/2$. In the dispersion curve the allowed “modes” will then correspond to

$$k_z = 2\pi/\lambda_p = \pi n/L \quad (n = 0, 1, \dots)$$

where n is the mode index. As an example, Fig. 10 shows the dispersion curve of a 7-cell disc-loaded structure with its 7 allowed modes. Remembering that the main branch of the dispersion curve (Fig. 7) extends from $k_z = 0$ to $k_z = \pi/l$, we can already observe that the number of permitted modes will be equal to the number of periods (cells) in the structure: the highest mode corresponds to $\pi N/L = \pi/l$ or $N = L/l$. The longitudinal electric field resulting from the superposition of the two travelling waves will have a maximum or a minimum on the end walls (the only case where transverse components of the field are zero on the metallic boundary). The sum of two waves travelling in opposite direction along the z axis with propagation constant k_z will have amplitude proportional to:

$$e^{jk_z z} + e^{-jk_z z} = 2 \cos k_z z = 2 \cos \frac{\pi n}{L} z = 2 \cos \frac{\pi n}{Nl} z \quad (n = 0, 1, \dots, N)$$

i.e. a standing wave with period Nl/n . The sinusoidal solutions are shown on the right side of Fig. 10, together with the corresponding electric field distribution for some of the modes. The phase difference of the mode amplitude between adjacent cells $\Delta\phi = \pi n/N$ ranges from 0 ($n = 0$) to π ($n = N$) and is used to identify the mode.

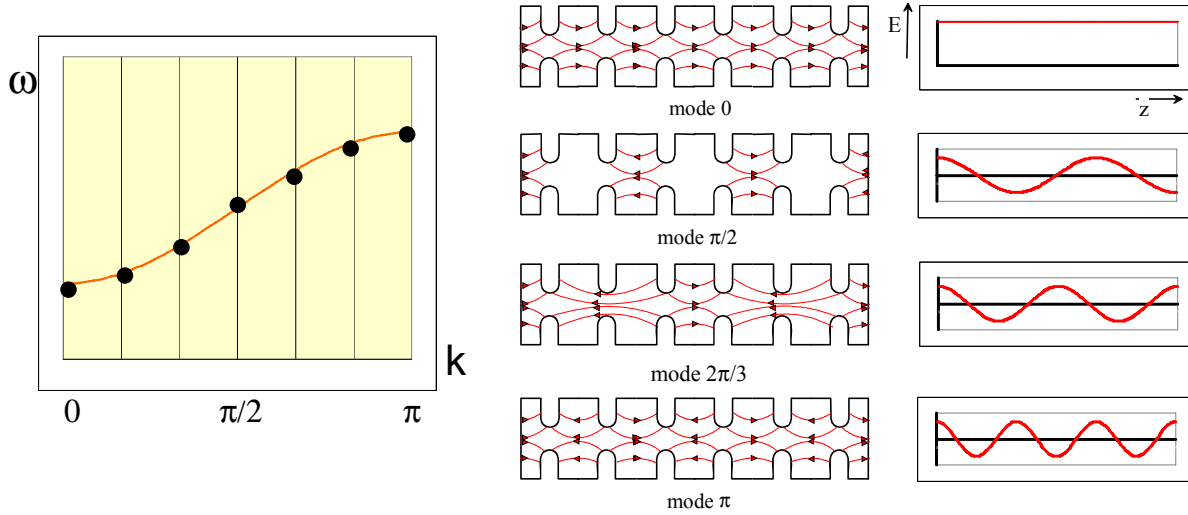


Fig. 10: Dispersion curve and field patterns for standing-wave linac structure

For particle acceleration, the most commonly used modes are the 0, $\pi/2$ or π . In the case of the π -mode, the electric field has a maximum at the centre of each cell, and changes sign from one cell to the next. The condition for acceleration has been already calculated in Section 2: cell length must be $l = \beta\lambda/2$. The conclusion is that a standing wave structure can be matched to any possible value of β , simply changing the cell length. Standing-wave linacs are ideal structures for the acceleration of protons and heavy ions, and for the initial acceleration of an electron beam. Inside a single accelerating structure the cell length can be easily increased proportionally to the increase in β , allowing to maintain the synchronism.

In practice, we have to make two more remarks on the use of standing wave structures for ion linacs. First of all, not only the π -mode can be used for acceleration. The 0-mode can be used as well, with synchronism condition $l = \beta\lambda$, and some particular structures even make use of the $\pi/2$ mode. Secondly, in a standing-wave linac the particle see the electric field only in the centre of the cells. It is therefore important to concentrate the field in this region, by adding “noses” to the discs separating the cells. But reducing the dimension of the central aperture the wave propagation would be more difficult, hindering the operation of the structure. To avoid this problem, “slots” can be opened in the cell walls, to restore the correct propagation. Figure 11 shows a 3D view of a standard cell for a π -mode standing-wave linac for protons. In a standing-wave linac, the RF power is introduced via an input coupler usually placed in the middle, as can be seen in the comparison table of travelling and standing wave structure given in Fig. 12.

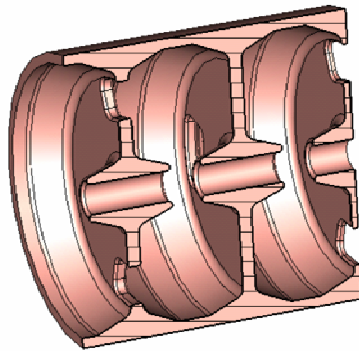


Fig. 11: An example of standing-wave linac structure

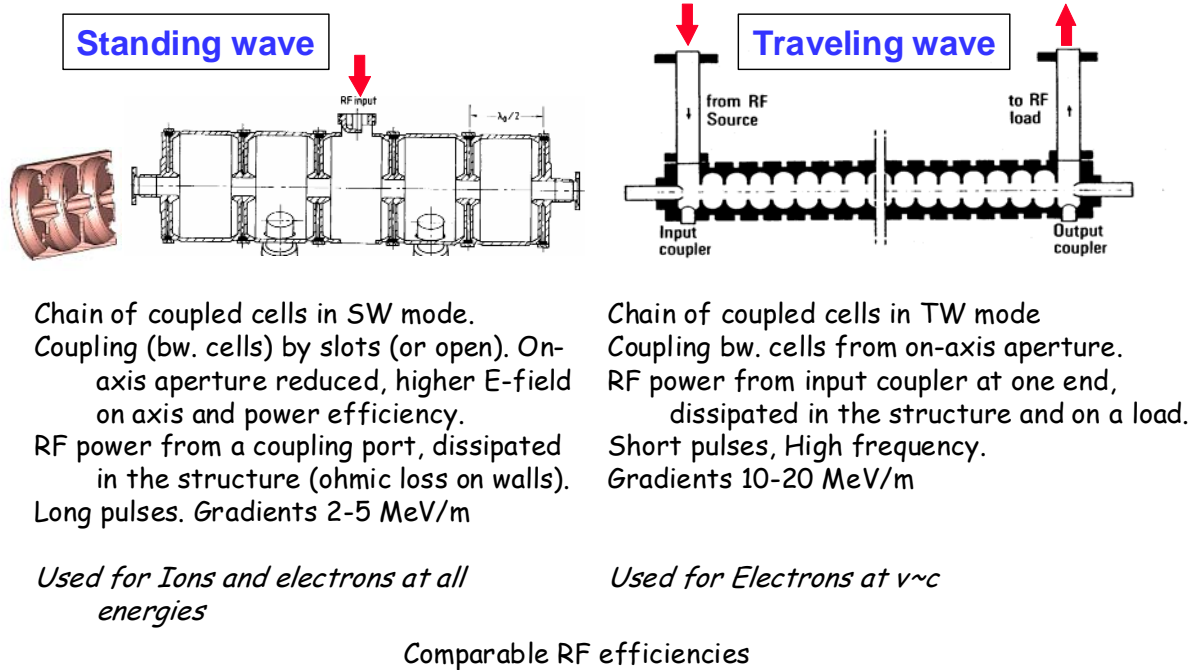


Fig. 12: Summary of properties of standing-wave and traveling-wave linac structure

An interesting example of a 0-mode structure for protons at relatively low energy (from about 3 to 90 MeV) is the Drift Tube Linac (DTL, Fig. 13). In this structure, the aperture on the beam axis has been greatly reduced in order to increase the efficiency, and thanks to the particular configuration of the fields the walls between cells have been completely removed, to maximise coupling between cells as well as power flow inside the structure and to minimise wall losses. What remains inside the structure are “drift tubes”, which can be made large enough to contain a focusing quadrupole, for the transverse beam focalization that is particularly important at low energy. The drift tubes are kept in position by supports (“stems”) from the top of the structure. Tuning plungers allow keeping the structure on resonance in presence of fluctuations of temperature or pressure, while “post-couplers” provide a stabilization of the fields against mechanical errors. Figure 14 shows the electric and magnetic field densities inside a DTL structure.

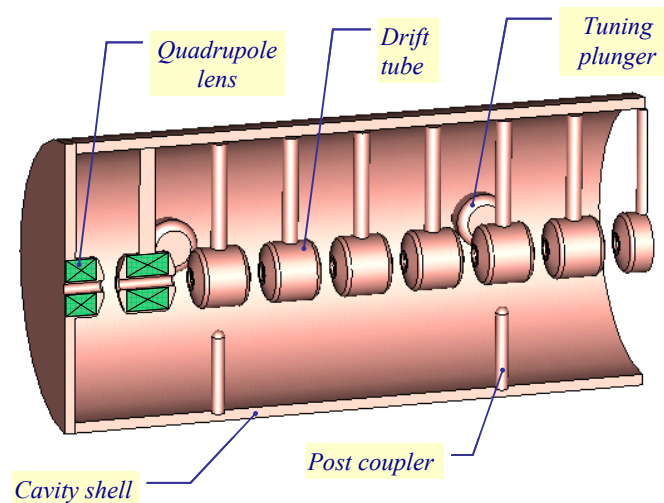


Fig. 13: Drift tube linac structure

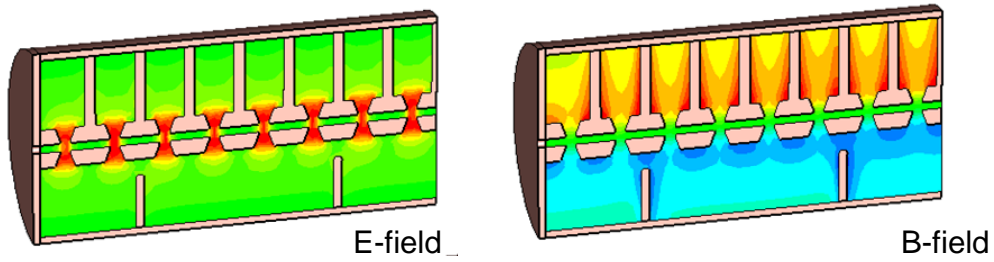


Fig. 14: Electric and magnetic field distribution inside a drift tube linac structure

4 Beam dynamics

4.1 Longitudinal

We have seen that, in order to achieve the maximum acceleration, bunches of particles must be synchronous with the accelerating wave. This means that they have to be injected in the linac on a well defined phase with respect to the accelerating sinusoidal field, and then they have to maintain this phase during the acceleration process. However, usual linac beams are made of a large number of particles with a certain spread in phase and in energy. If the injection phase corresponds to the crest of the wave ($\varphi = 0^\circ$ in the linac definition) for maximum acceleration, particles having slightly higher or lower phases will gain less energy. They will slowly lose synchronicity until they are lost. The important principle of phase stability, strictly valid only for particles not yet relativistic, provides a means to keep the particles longitudinally bunched during acceleration. If the injected beam is not centred on the crest of the wave but on a slightly lower phase, a “synchronous phase” φ_s whose typical values are between -20° and -30° , particles that are not on the central phase will oscillate around the synchronous phase during the acceleration process. The resulting longitudinal motion is confined, and the oscillation is represented by an elliptical motion of each particle in the longitudinal phase plane, the plane $(\Delta\varphi, \Delta W)$ of phase and of energy difference with respect to the synchronous particle. The relation between the synchronous phase in an accelerating sinusoidal field and the longitudinal phase plane is represented in Fig. 15.

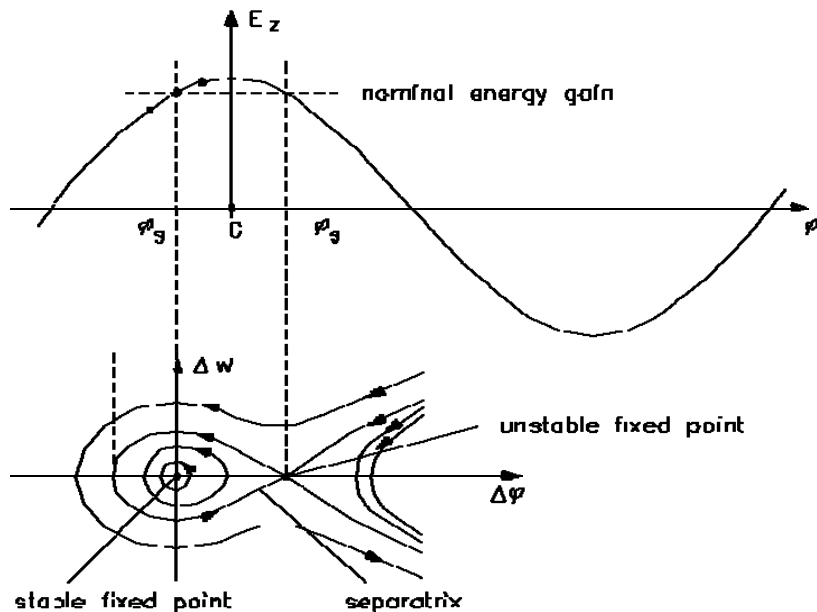


Fig. 15: Longitudinal motion of an ion beam

It is interesting to observe that the frequency of longitudinal oscillations, i.e. the number of oscillations in the longitudinal phase plane per unit time depends on the velocity of the beam. A simple approximate formula for the frequency of small oscillations ω can be found for example in [5]:

$$\omega_l^2 = \omega_0^2 \frac{qE_0 T \sin(-\varphi_s) \lambda}{2\pi mc^2 \beta \gamma^3}$$

Here ω_0 and λ are the RF frequency and wavelength, $E_0 T$ is the effective accelerating gradient and φ is the synchronous phase. The oscillation frequency is proportional to $1/\beta\gamma^3$: when the beam becomes relativistic, the oscillation frequency decreases rapidly. At the limit of $\beta\gamma^3 \gg 1$, the oscillations will stop and the beam is practically “frozen” in phase and in energy with respect to the synchronous particle. For example, in a proton linac $1/\beta\gamma^3$ and correspondingly ω can decrease by 2 or 3 orders of magnitude from the beginning of the acceleration to the high energy section.

Another important relativistic effect for ion beams is “phase damping”, the shortening of bunch length in the longitudinal plane. It can be understood considering that, as the beam becomes more relativistic, its length in z seen by an external observer will contract due to relativity. A precise relativistic calculation shows that the phase damping is proportional to $1/(\beta\gamma)^{3/4}$:

$$\Delta\varphi = \frac{\text{const}}{(\beta\gamma)^{3/4}}$$

When a beam becomes relativistic, not only its longitudinal oscillations slow down, but also the bunch will compact around the center particle.

In an electron linac instead, after the initial section $\beta\gamma^3 \rightarrow \infty$ and phase stability does not apply. There is no movement in the longitudinal phase plane, and during acceleration, the beam remains at the injection phase. The phase length $\Delta\varphi$ of an electron bunch tends to zero, and at the relativistic limit an electron beam will be compacted around the central particle. As a consequence, acceleration of a relativistic electron beam can take place on the crest of the wave. The beam will remain bunched, and acceleration will be the highest.

An important feature of electron linacs is the capture condition of an electron beam not yet relativistic into an accelerating structure designed for $\beta_c = 1$. The beam is not “synchronous”, and the phase of a particle injected with velocity $\beta = \beta_0$ at phase φ_0 into a travelling wave structure will change with β accordingly to the relation [5, page 191]:

$$\sin\varphi(\beta) = \sin\varphi_0 + \frac{2\pi mc^2}{\lambda_g qE_0} \left[\sqrt{\frac{1-\beta_0}{1+\beta_0}} - \sqrt{\frac{1-\beta}{1+\beta}} \right]$$

At the relativistic limit ($\beta = 1$) the phase is increased by:

$$\Delta(\sin\varphi) = \frac{2\pi mc^2}{\lambda_g qE_0} \left[\sqrt{\frac{1-\beta_0}{1+\beta_0}} \right]$$

This expression shows that a non-relativistic electron bunch injected into a $\beta = 1$ structure will increase its phase, converging asymptotically to a well defined phase shift depending on the injection energy and on the accelerating field. In the limit case of injection phase -90° (the minimum to ensure

acceleration), the condition for having the beam converging to the crest (0° , maximum acceleration) corresponds to taking $\Delta(\sin\phi) = 1$ in the previous relation and is called the capture condition:

$$\frac{2\pi mc^2}{\lambda_g qE_0} \left[\frac{1-\beta_0}{1+\beta_0} \right] \leq 1$$

For a given gradient E_0 this expression gives the minimum β_0 that can be accepted by the structure, or for a given β_0 at injection it gives the minimum gradient to be applied in order to capture the beam. For example, taking $E_0 = 8$ MV/m, a 3 GHz structure has a minimum capture energy of 150 keV. Figure 16 shows how the phase of the injected bunch changes during acceleration, combining the phase shift towards the crest of the wave and the decrease of phase spread (phase damping).

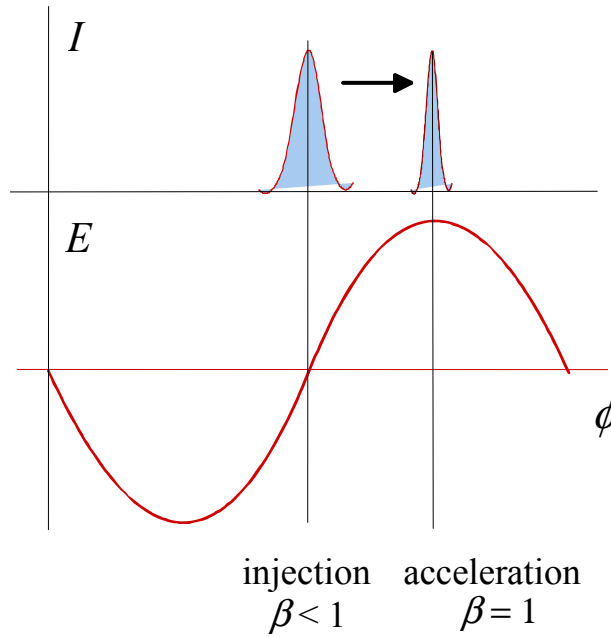


Fig. 16: Longitudinal motion of an electron beam

4.2 Transverse

Transversally in a linac, the beam will be subject to an external focusing force, provided by an array of quadrupoles or solenoids. This force has to counteract the defocusing forces that either develop inside the particle beam or come from the interaction with the accelerating field. The main defocusing contributions come from:

Space charge forces

Representing the Coulomb repulsion inside the bunch between particles of the same sign. In the case of high intensity linacs at low energy, space charge forces are one of the main design concerns. However, at relativistic velocity the space charge repulsion starts to be compensated by the attraction due to the magnetic field generated by the beam, and finally disappears at the limit $v = c$. It is quite simple to calculate the electric and magnetic field active on a particle at distance r from the axis of an infinitely long cylindrical bunch with density $n(r)$ travelling at velocity v (Fig. 17):

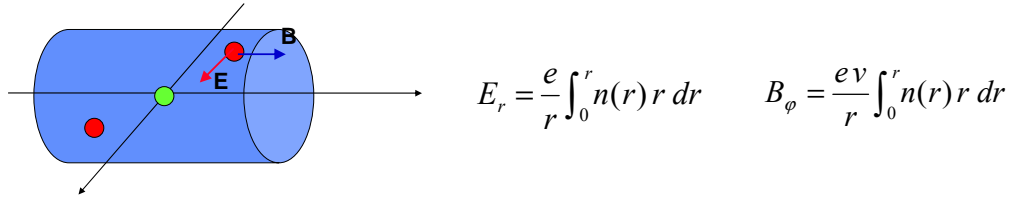


Fig. 17: Forces acting on a particle inside an infinitely long bunch

The overall force acting on the particle is directed radially, and has intensity:

$$F = e(E_r - vB_\phi) = eE_r(1 - \frac{v^2}{c^2}) = eE_r(1 - \beta^2) = \frac{eE_r}{\gamma^2}.$$

The overall space charge force is then proportional to $1/\gamma^2$ and will disappear for $\gamma \rightarrow \infty$.

RF defocusing forces

Which represent the transverse defocusing experienced by a particle crossing an accelerating gap on an RF phase that is longitudinally focusing. We have seen in the previous section that for longitudinal stability the beam will cross the gap while the field is increasing ($\varphi_s < 0$). Figure 18 shows a schematic configuration of the electric field in an accelerating gap. In correspondence to the entry and exit openings for the beam, the electric field has a transverse component, focusing at the entrance to the gap and defocusing at the exit, proportional to the distance from the axis. Because the field is increasing, the defocusing effect will be stronger than the focusing one, and the net result will be a defocusing force, proportional to the time required by the beam to cross the gap. Again, this effect is proportional to $1/\gamma^2$, and will disappear at high beam velocity.

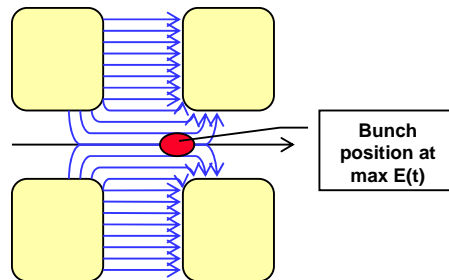


Fig. 18: Electric field line configuration around a gap and position of the bunch at maximum field

The conclusion is that the beam must be transversally in equilibrium between an external focusing force and some internal defocusing forces. This equilibrium will result in an oscillation in time and in space of the beam parameters, with a frequency that will depend on the difference between focusing and defocusing forces. Instead of defining the frequency of oscillation with respect to time, it is convenient to characterise the oscillation in terms of the phase advance per focusing period of the oscillation σ_i , the focusing structure being periodical. Alternatively, we can use the phase advance per unit length k_i . Clearly, if L is the period length of the focusing structure, $k_i = \sigma_i/L$. In an ion linac, σ_i should always be $< 90^\circ$, to avoid instabilities, but should always be higher than some $10\text{--}20^\circ$, in order to avoid that the amplitude of the oscillations becomes too high and the beam size becomes too large.

We can find an approximate relation for the phase advance as function of focusing and defocusing forces in a simple case. First of all, one has to limit the analysis to beam oscillations in a simple FODO quadrupole lattice (focusing-drift-defocusing-drift) under smooth focusing approximation, i.e. averaging the localised effect of the focusing elements. Then, adding together the focusing and RF defocusing contributions to phase advance as derived for example in [5, 7.103] and subtracting the space charge term as approximately calculated in the case of a uniform three-dimensional ellipsoidal bunch [5, 9.51] we obtain for the phase advance per unit length:

$$k_t^2 = \left(\frac{\sigma_t}{N\beta\lambda} \right)^2 \cong \left(\frac{qGl}{2mc\beta\gamma} \right)^2 - \frac{\pi q E_0 T \sin(-\varphi)}{mc^2 \lambda \beta^3 \gamma^3} - \frac{3qI\lambda(1-f)}{8\pi\epsilon_0 r_0^3 mc^3 \beta^2 \gamma^3}.$$

Here $N\beta\lambda$ is the length of the focusing period in units of $\beta\lambda$. The first term on the right side of the equation is the focusing term: Gl is the quadrupole strength, expressed as product of gradient G and length l of the quadrupole. The second term is the RF defocusing: E_0T is the effective accelerating gradient, λ the RF wavelength and φ the synchronous phase. For $\varphi < 0$, corresponding to longitudinal stability in the linac definition, $\sin(-\varphi)$ is positive and this term is negative, i.e. defocusing. The third term is the approximate space charge contribution: I is the beam current, f is the ellipsoid form factor ($0 < f < 1$) and r_0 is the average beam radius. The other parameters in the equation define the particle and medium properties (charge q , mass m , relativistic parameters β and γ free space permittivity ϵ_0). It must be noticed that this simple equation shows, although in an approximate simplified case, how the beam evolution in a linear accelerator depends on the delicate equilibrium between external focusing and internal defocusing forces. Even if real cases can be solved only numerically, the parametric dependence given by this equation remains valid, and allows to determine how the dynamics will change with the particle β .

We immediately see that at low velocities ($\beta \ll 1$, $\gamma \sim 1$) the defocusing terms are dominant. In order to keep the beam focused with a large enough phase advance per unit length one has to increase the integrated gradient Gl and/or decrease the length of the focusing period $N\beta\lambda$, i.e. minimise the distance between focusing elements. This is for example the case of the Radio Frequency Quadrupole (RFQ), the structure of choice for low energy ion beams (from $\beta \approx 0.01$ to $\beta \approx 0.1$). The RFQ provides a high focusing gradient by means of an electrostatic quadrupole field, with short cells at focusing period $\beta\lambda$. At higher energy, standard electromagnetic quadrupoles have a sufficiently high gradient and a structure alternating accelerating gaps and quadrupoles can be used. The typical structure for protons at >3 MeV is the Drift Tube Linac (DTL, Fig. 14), which has $2\beta\lambda$ focusing period when focusing and defocusing quadrupoles alternate inside the drift tubes. Going further up in energy, the defocusing terms ($\propto 1/\beta^3\gamma^3$, $1/\beta^2\gamma^3$) decrease much faster than the focusing term ($\propto 1/\beta\gamma$). The focusing period can be increased, reducing the number of quadrupoles and simplifying the construction of the linac. At energies beyond about 50 MeV, modern proton linacs usually adopt multi-cell structures operating in π -mode spaced by focusing quadrupoles, like the ones of Fig. 12 left (normal conducting) and Fig. 2 (superconducting). For example, in the CERN Linac4 design (90 keV to 160 MeV beam energy) the focusing period increases from $\beta\lambda$ in the RFQ to $15\beta\lambda$ in the last π -mode accelerating structure.

Heavy ions differ from protons for the fact that usually ion currents are small and the space charge term can be neglected. Immediately after the RFQ, the focusing period can be some 5–10 $\beta\lambda$.

In the case of electron linacs transverse dynamics becomes simpler. The two defocusing terms disappear, and phase advance is zero independently from the focusing term. The consequence is that the beam does not oscillate in the transverse plane. However, some external focusing is usually required to control the beam emittance and to stabilise the beam against instabilities like wakefields and beam breakup. The length of the focusing period is not very important in this case, and usually

focusing elements are spaced by several meters. Moreover, as focusing elements are often used solenoids placed around the accelerating structure, because of the low gradients required for the focalisation, which make effective even solenoids with large apertures.

5 Linac technologies

5.1 Particle sources

One of the few aspects of linac technology where electron and ion linacs drastically differ is in the technology of the source of particles. The same fundamental problem, the extraction of charged particles from matter, is solved in completely different ways in the case of light electrons to be extracted from the external atomic layers and in the case of heavier particles to be extracted from the nuclei.

Electron sources have to give energy to the free electrons inside a metal to overcome the potential barrier at the boundary. The energy can be provided by heat, via the standard thermoionic effect, or by an intense localized laser pulse.

An ion source instead has first of all to separate the electrons from the nuclei creating a plasma. Then the plasma conditions have to be optimized in terms of heating, confinement and loss mechanisms in order to produce the desired ion type. The final step in the process is to remove ions from the plasma via an aperture and a strong electric field. The sources differ for the process used to produce the plasma (electrical discharge, RF fields, electron cyclotron resonance excitation, etc.) and for the type and shape of the extraction field.

5.2 Low-energy injectors

After the source, both electron and ion linacs include a low-energy injector, which in modern linacs tends to be a quite sophisticated and critical section, especially when high currents are required. This is the only section in a linac where the requirements for electrons and ions are quite similar. The injector has to basically to perform three tasks:

- Compact the beam particles into bunches at the required accelerating frequency (both electron and ion sources deliver a continuous beam pulse).
- Transport over a distance of a few meters a low-energy beam, usually of high current and subject to strong space charge defocusing forces.
- Accelerate smoothly the beam from the extraction energy of the source (usually 10 to 100 keV) up to the input energy of the main accelerating structure (usually between 1 and 5 MeV), in an energy range where beam velocity considerably increases.

The first function, called bunching, is achieved by dedicated RF gaps (“buncher” cavities) operated on the longitudinally focusing phase ($\varphi_s = -90^\circ$), followed by a drift distance. The bunchers can be operated at different frequencies to maximise the beam transmission of the injector. An alternative for ion beams is to apply a distributed adiabatic bunching inside a special accelerator containing many bunching and accelerating cells, the Radio Frequency Quadrupole (RFQ). RFQ’s are nowadays commonly used for ion linacs, where the source current is intrinsically limited and beam transmission through the pre-injector should be as close as possible to 100%.

Focusing in an electron pre-injector is usually provided by an array of solenoids placed around the beam line. For ion beams subject to severe space charge forces, a stronger focusing is needed together with the shortest possible length of the focusing period. Again, the RFQ structure can provide

a electric focusing quadrupole field, which contrarily to magnetic focusing is independent from beam velocity and more effective at low energy, distributed all along the structure length.

Acceleration can be easily provided by the bunching gaps, some of which can be phased towards the accelerating synchronous phase. Again, this is easily realised in the RFQ, by simply modulating the length of the bunching and accelerating cells. In high-current ion linacs the injector is completed by one or more bunching cavities between the RFQ and the main linac, needed to preserve the beam bunching during the transport to the next accelerating structure.

As examples of pre-injectors, Fig. 19 shows two schemes from CERN. The CTF3 electron injector includes as many as five bunching cavities following the electron gun. Then, bunching is completed in a travelling-wave structure that performs the first acceleration as well. Focusing solenoids surround the injector. In the RFQ2 high-current proton injector instead bunching acceleration and focusing are performed in a 1.8 m long RFQ at 200 MHz. Two focusing solenoids after the source prepare the beam for injection into the RFQ and two bunchers after the RFQ preserve the bunching before the injection into the following DTL.

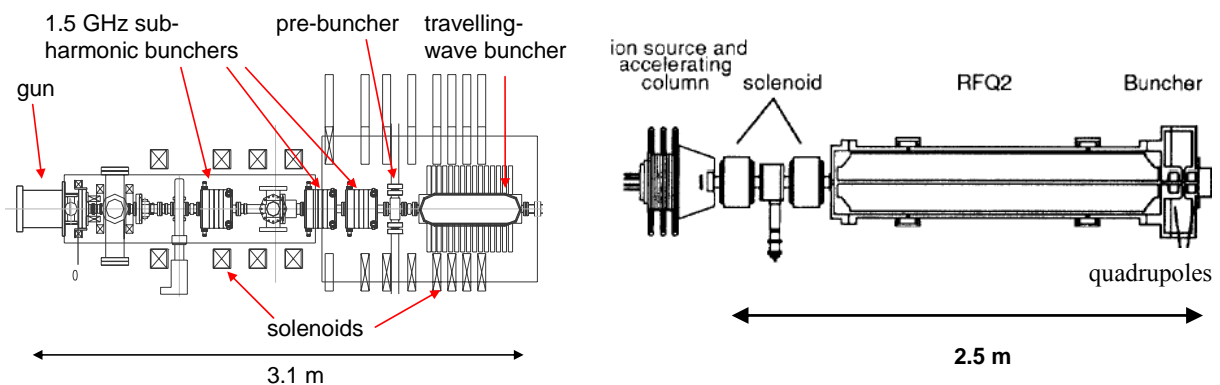


Fig 19: The CTF3 electron injector (left) and the RFQ2 proton injector (right) at CERN

5.3 Accelerating structure technology

We have seen that the standard accelerating structures used in linear accelerators are disc-loaded travelling wave structures for electrons and coupled cell standing wave structures, which can be used for all type of particles. In the project of a linear accelerator the first problem for the designer is the choice of the frequency, which has to take into account many factors coming from beam dynamics, Radio Frequency and mechanical considerations. The important parameters of a linear accelerator scale differently with the frequency. First of all, accelerator dimensions and cell length are proportional to the wavelength λ . Other important factors scaling with the wavelength are the machining tolerances, which can have a considerable impact on the cost of an accelerator. In the beam dynamics section, we have seen how the RF defocusing scales with $1/\lambda$. Other important parameters are the RF power efficiency, usually called the shunt impedance, which in first approximation scales as \sqrt{f} , and the maximum electric field achievable before breakdown, which accordingly to an empirical law scales as well with \sqrt{f} .

Table 1: Scaling with frequency of some basic linear accelerator parameters

Parameter	Scaling
RF defocusing	$\sim f$
Cell length ($\sim\beta\lambda$)	$\sim 1/f$
Peak electric field	$\sim\sqrt{f}$
RF power efficiency (shunt impedance)	$\sim\sqrt{f}$
Acc. structure dimensions	$\sim 1/f$
Machining tolerances	$\sim 1/f$

From a rapid analysis of these parameters, it is clear that high frequencies are economically convenient. At higher accelerating frequency the linac is shorter, makes use of less RF power and can reach higher peak and average accelerating field. However, limitations to the frequency come from the availability and price of high frequency RF power sources and from the mechanical construction costs. Higher frequency means smaller dimensions but also tighter and more expensive machining tolerances. Additionally, in ion linacs the beam dynamics sets a upper limit to the frequency in the injector section, where RF defocusing is strong.

Electron linacs usually adopt higher frequencies than ion linacs, in the range between 3 and 30 GHz, because they are not limited by RF defocusing and because disc-loaded structures have larger dimensions and lower machining tolerances than standing-wave structures at the same frequency. The standard commercial frequency for electron linacs is 3 GHz.

Proton linacs need to start with relatively low frequencies in the RFQ and injector section, ranging in modern linacs from 200 to 400 MHz, in order to reduce the effect of RF defocusing and to make possible the construction of the extremely short first RFQ cells ($\beta\lambda$ long). After the injector and the first accelerating structure (usually a DTL), it is sometimes convenient to double the frequency. Standard medium-high energy sections of modern proton linacs operate at frequencies between 400 and 800 MHz, which usually result from a compromise between focusing requirements, cost and dimensions.

Heavy ion linacs tend to use even lower frequencies (30–200 MHz). The choice of the initial frequency is dominated by the low particle velocity in the first sections. For example, the first cell of the CERN lead ion RFQ (100 MHz, Pb27+ ions at 25 keV/u input energy) is only 7 mm long ($\beta\lambda$). Lower ion charge to mass ratios require even lower frequencies. However, in heavy ion linacs as well it is convenient to double (or even triple) the basic frequency as soon as the ion velocity has reached a sufficiently high value.

As for the construction technology, the structures that we have considered have to carry large RF currents with minimum power loss, and are usually made out of copper. Alternatively, structure with low duty cycle, where the high thermal conductivity of copper is not required, can be made of copper plated steel. The thin copper layer usually covers few penetration depths of the RF current.

Schematically, a linac accelerating structure is an assembly of precisely machined copper or copper plated parts, each making one cell or part of a cell. Joining together these parts to form an adequate RF and vacuum envelope is an art by itself. Furnace brazing is commonly used for small high frequency cells made out of copper, while combinations of welding and gasket mounting are used for larger parts at lower frequency.

5.4 RF technologies

The Radio Frequency power source can be an important cost driver in a linac project, and has to be carefully selected. In general, linacs have the advantage of being pulsed machines, with duty cycles usually below 10%. On the other hand, the beam passes only once through an accelerating section, and in order to reduce the linac length and cost often high accelerating gradients are required, which translate into high values of the peak power required from the RF system. The result is that linac RF systems have the particularity of requiring high peak powers, but relatively low average powers as compared to other RF systems. Klystrons are commonly used as RF power sources for high frequencies (≥ 300 MHz). They are the typical RF source for electron linacs, and are more and more used as power source for proton linac projects. Low frequency proton and ion linacs use instead triode or tetrode based amplifiers as main RF power source, which can be effectively used up to about 400 MHz.



Fig. 20: A 3 GHz klystron amplifier for the LPI electron linac (left) and a 200 MHz triode-based amplifier for the Linac3 ion linac (right), both used at CERN

5.5 Modern trends

Unquestionably the most important development of recent years in the field of linear accelerators is the remarkable progress in superconducting cavity technology, which is leading to a wider application of superconductivity to linac projects. This development went in parallel with a general increase in the accelerating gradients, for both normal conducting and superconducting structures, and of the operating frequencies.

While superconducting accelerating structures were once limited to heavy ion linacs, where short superconducting structures are particularly suited for accelerating different types of ions in CW mode, and to electron linacs operating in CW mode, nowadays high-gradient multi-cell standing wave structures like the one in Fig. 2 are widely used in linacs for both electrons and protons. The vigorous TESLA R&D programme at DESY [6] has defined construction and preparation technologies for superconducting cavities that allow achieving high accelerating gradients. In modern linac designs, safe design values for the gradient of $\beta=1$ cavities are now in the order of 25 MV/m, a value that scales down to some 19 MV/m at $\beta=0.65$ [7]. Moreover, the improvement of RF control and cavity stiffening techniques now allow to effectively operating in pulsed mode a superconducting RF system even in presence of the large cavity detuning from the forces generated by the accelerating field. Whereas the TESLA development at 1.2 GHz is directed to electron linear colliders, a lot of parallel work went on for multi-cell cavities for protons, usually at lower frequency (700–800 MHz), with considerable synergies between electron and proton developments. Superconducting cavities are now

used for the high-energy part of the new SNS 1 GeV linac at Oak Ridge and are proposed for almost all high-energy linac projects worldwide.

In parallel, the technology of Normal Conducting linacs has also evolved towards higher gradients, and in particular a lot of effort went into better understanding of breakdown phenomena [8]. While proton linac nowadays commonly start from frequencies around 400 MHz and some projects even go in the GHz range for the high energy section, for electron linacs are now exploited frequencies that few years ago were considered almost unreachable, as for the 30 GHz acceleration frequency in the CTF Test facility [9].

Acknowledgements

The author is grateful to F. Gerigk, A. Lombardi, S. Minaev, L. Rinolfi, N. Pichoff, C. Rossi and R. Scrivens for their help in preparing the lecture and for kindly providing some of the pictures.

References

- [1] See for example J.P. Blewett, The history of Linear Accelerators, in P. Lapostolle, A. Septier, Linear Accelerators (North Holland, 1970).
- [2] L.W. Alvarez et al., Rev. Sci. Instr. **26**, 111, (1955).
- [3] E.L. Ginzton, W.W. Hansen, W.R. Kennedy, Rev. Sci. Instr. **19**, 89 (1948); M. Chodorow et al., Rev. Sci. Instr. **26**, 134 (1955).
- [4] See for example S. Ramo, J.R. Whinnery, T. Van Duzer, Fields and Waves in Communication Electronics, Wiley, New York (1994).
- [5] T. Wangler, Principles of RF Linear Accelerators, Wiley, New York (1998).
- [6] TESLA Technical Design Report: the superconducting electron-positron linear collider with an integrated x-ray laser laboratory, DESY-2001-011 (2001).
- [7] F. Gerigk (ed.), The SPL Conceptual Design Report 2, CERN (in preparation).
- [8] W. Wuensch, High Gradient Breakdown in Normal-Conducting RF Cavities, EPAC2002 Proceedings, Paris, France, 2002.
- [9] G. Geschonke, A. Ghigo (eds.), CTF3 Design Report, CERN-PS-2002-008-RF (2002).

Bibliography

- P. Lapostolle, A. Septier (editors), Linear Accelerators (Amsterdam, North Holland, 1970).
- T. Wangler, Principles of RF Linear Accelerators (Wiley, New York, 1998).
- I.M. Kapchinskii, Theory of resonance linear accelerators (Harwood, Chur, 1985).
- M. Puglisi, The Linear Accelerator, in E. Persico, E. Ferrari, S.E. Segré, Principles of Particle Accelerators (W.A. Benjamin, New York, 1968).
- P. Lapostolle, Proton Linear Accelerators: A theoretical and Historical Introduction, LA-11601-MS, 1989.
- P. Lapostolle, M. Weiss, Formulae and Procedures useful for the Design of Linear Accelerators, CERN-PS-2000-001 (DR), 2000.

P. Lapostolle, R. Jameson, Linear Accelerators, in Encyclopaedia of Applied Physics (VCH Publishers, New York, 1991).

J. Clendenin, L. Rinolfi, K. Takata, D.J. Warner (eds.), Compendium of Scientific Linacs, CERN/PS 96-32 (DI), 1996.

S. Turner (ed.), CAS School: Cyclotrons, Linacs and their applications, CERN 96-02 (1996).

M. Weiss, Introduction to RF Linear Accelerators, in CAS School: Fifth General Accelerator Physics Course, CERN-94-01 (1994), p. 913.

N. Pichoff, Introduction to RF Linear Accelerators, in CAS School: Basic Course on General Accelerator Physics, CERN-2005-04 (2005).