

Introduction to Accelerator Theory

E. J. N. Wilson¹

John Adams Institute for Accelerator Science

University of Oxford, United Kingdom

ted.wilson@cern.ch

This paper is the first of two which summarise the basic theory necessary to understand the design of particle accelerators and their limitations. It is intended as preparation for those about to start a career in accelerator physics but also might be considered essential background knowledge for anyone who intends to propose experiments on a future linear collider. Luminosity is of paramount importance to such experimenters and is determined not only by the beam intensity and the optics of the final focus but by the performance of the damping rings which shape the emittance of the beams. Factors which place limits on luminosity are discussed and how they may be reconciled to ensure optimum performance.

The current proposals, ILC and CLIC, stem from two quite different linac technologies – superconducting cavities in the case of ILC and for CLIC: higher frequency room temperature cavities powered from a second drive linac. This leads to quite different parameters for the delivered beam in each case. CLIC offers a more compact accelerator capable of delivering up to 3 TeV in the Centre of Mass. ILC's technology might be considered better established but relies on kilometres of superconducting components within cryostats – a recipe perhaps to be avoided if it does not bring with it an advantage in energy. Furthermore it has yet to be established with LHC data whether the much lower centre of mass energy of ILC will be appropriate for future physics needs.

¹ E. J. N. Wilson

Third Linear Collider Physics School 2009 – LCPS2009
Ambleside, UK
August 17-23, 2009

POS(LCPS2009)006

1. Overview of the history

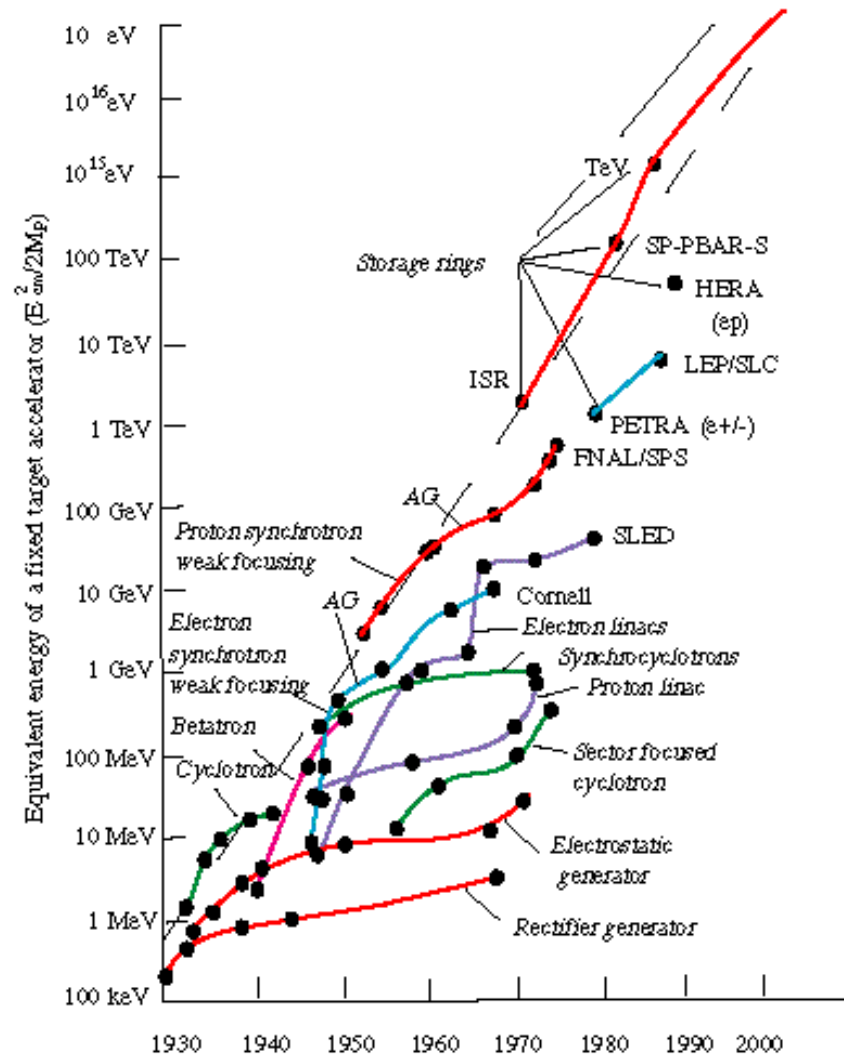


Fig. 1 The development of high energy accelerators labelled with the technological advances which contributed to this progress.

Fig. 1 shows how electron and proton accelerators developed from modest beginnings to become the most powerful tools available today for the study of physics. The motivation to strive for higher energies came from the need to use projectiles whose De Broglie wavelength was short compared with detail in the structure of fundamental particles, hand-in-hand with the understanding of smaller and smaller structures. The result was the discovery of a whole series of ever more massive particles requiring more and more energetic particles to produce them.

As beam instabilities were mastered, fixed target accelerators gave way to storage rings in which two counter rotating beams meet continuously head-on.

2.Linear accelerators



Fig. 2 Inside a typical Alvarez linac showing its drift tubes

In 1927, Wideröe built a three-tube linac model which accelerated sodium ions. It was much later, in the mid-forties, that L. W. Alvarez [1] at the Radiation Laboratory of the University of California started to build the first serious proton linac. By this time suitable high power, high frequency oscillators had become available to meet the needs of war-time radar development. Figure 2 shows an Alvarez linac – a copper-lined cylinder excited by a radio transmitter. As in Wideröe’s linac, particles gained energy from the accelerating potential differences between the ends of the drift tube. The particles are protected from the decelerating phase while inside the metallic drift tubes. The spacing of drift tubes increases as the particle gains energy up to the point that is travelling close to the velocity of light.

In linacs designed for electrons and other highly relativistic particles, the drift tube spacing is uniform since the time of flight between accelerating gaps is just the period of the wave which excites the structure. In fact modern linac structures contain no drift tubes but resemble cylindrical waveguides into which a number of equally spaced diaphragms or “washers” have been inserted. The phase velocity in a simple cylinder without washers would

POS(LCPS2009)006

be larger than c and therefore would not remain in phase with the particle. The effect of the washers is to slow down the wave travelling along the structure to a velocity, slightly below c , matching that of the electrons. In the superconducting linac structure proposed for ILC the washers have been replaced by corrugating the guide as shown in Figure 3



Fig. 3 Superconducting linac structure - TESLA.

2.1 Important parameters of linac structures

Accelerating structures, whether they are simple pill-box cavities or the more sophisticated structures for linear colliders, can be thought of in terms of an equivalent circuit. Take a simple resonant circuit consisting of a capacitor, inductance and resistor in parallel. The accelerating voltage (per meter of structure) is equivalent to the voltage across the capacitor while the energy stored in the fields within the structure, W_s is proportional to the square of this accelerating voltage.

The commonly used figure of merit for an accelerating cavity is its shunt impedance R_s . This is the parameter which relates the accelerating voltage to the power P_d which has to be provided

to balance dissipation in the walls and drive the linac. The voltage V can be defined along a path followed by the beam:

$$V = \int_{\text{path}} |E_z(x, y, z)| d\ell$$

This, of course is at a fixed instant in time. Shunt impedance (with the usual factor 2 consistent with peak and not r.m.s. voltage) is defined:

$$R_s = \frac{V^2}{2P_d}$$

In the world of linear accelerators it is the shunt impedance per unit length and power dissipated per unit length that are usually quoted.

In order to calculate the voltage, dissipated power, and the energy stored in the cavity we use three dimensional relaxation programs to solve the general differential equation for the field:

$$\nabla^2 \mathbf{E} = \mu\sigma \frac{\partial \mathbf{E}}{\partial t}$$

with the boundary conditions at the conducting surface:

$$\mathbf{n} \times \mathbf{E} = 0 \quad \text{and} \quad \nabla \cdot \mathbf{E} = 0$$

where \mathbf{n} is normal to the wall.

To calculate power dissipated, P_d and later W_s , we should first evaluate the linear density (amps per meter) of the current \mathbf{j} along the walls of the selected structure as if the walls were lossless. The losses are then introduced taking into account the finite conductivity σ of the walls. Since for a perfect conductor we have $\mathbf{j} = \mathbf{n} \times \mathbf{H}$, we can write:

$$P_d = \frac{R_{surf}}{2} \int_s |H|^2 ds$$

where s is the inner surface of the structure and $R_{surf} = \sqrt{\pi\omega\mu_0\mu_r/\sigma} = 1/\sigma\delta$ is the surface resistance (for copper $R_{surf} = 2.61 \cdot 10^{-7} \sqrt{\omega}$ ohms).

The other important parameter is the quality factor of a resonator, defined as the ratio of the stored energy, W_s to W_d , the energy dissipated per cycle divided by 2π . The power dissipated is P_d :

$$Q = \frac{W_s}{W_d} = \omega \frac{W_s}{P_d}$$

The stored energy is the integral of

$$W_s = \frac{\epsilon}{2} \int |E|^2 dv \quad \text{or} \quad \frac{\mu}{2} \int |H|^2 dv$$

over the volume of the cavity. The first integral applies at the time the energy is all stored in the E field and the second integral as it oscillates back into the H field. The second term on the right

hand side is a consequence of Maxwell's equations. Modern cavity design programs do these calculations for us. Again we can ask the cavity design program to give us the surface integral of field.

These two important parameters are very different in the superconducting ILC structure and the room temperature CLIC main linac. Not only do they effect the length of structure necessary per TeV but the much higher Q of the ILC demands quite a different time structure for the injected beam.

3.The synchrotron

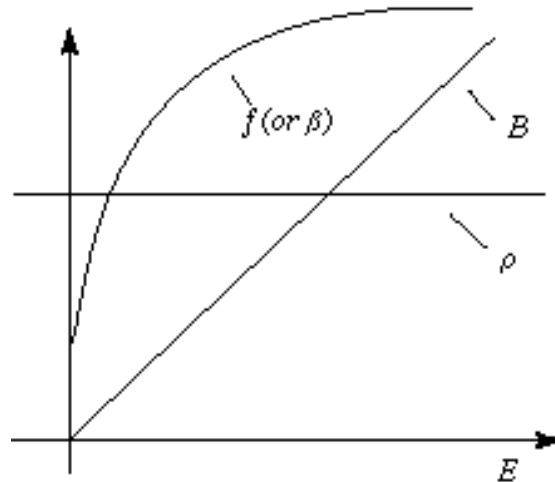


Fig. 4 Field and frequency rise together in a synchrotron.

After the World War II it was realised that cyclotrons, which accelerated a continuous stream of particles from their central region to the edge of their cylindrical poles could not be realistically scaled up to the higher energies were needed to pursue the aims of physics. It was Australian physicist, Mark Oliphant who synthesized three old ideas into a new concept – the synchrotron. The ideas were: accelerating between the gaps of resonators, varying the frequency, and pulsing the magnet.

We see in Fig. 4 how, once a short pulse is injected at low field, the field rises in proportion to the momentum of particles as they are accelerated and this ensures that the radius of the orbit remains constant. Unlike cyclotrons and betatrons, the synchrotron needs no massive poles to support a magnetic field within the beam's circular orbit. The guide field is instead provided by a slender ring of individual magnets. Acceleration is provided by fields within a hollow cylindrical resonator, excited by a radio transmitter. A particle passes through as it completes each turn of the synchrotron receiving another increment in energy at revolution. The fact that the magnet is pulsed and the frequency must be controlled to track the increasing speed of particles is a complication.

4.Phase stability

Although Oliphant was confident that his synchrotron could be made to work, others thought that any slight mistiming of the sine wave of accelerating voltage in the cavity might build up over many turns until particles would gradually get out of step until they were lost? Fortunately

the comforting principle of phase stability, which prevents this happening, was soon to be independently discovered opening the way to the construction of the first synchrotrons.

In order to achieve phase stability, particles orbiting the synchrotron are timed to ride, not on the peak, but on the rising edge of the voltage wave in the accelerating cavity. They receive more energy if they are late and less energy if they are early so that they oscillate about the stable or synchronous phase. For all particles, the time average of their energy gain matches the rising magnetic field. We shall come to discuss this principle of phase stability later. Now we must turn to the methods of constraining the particles in the directions perpendicular to their direction of travel – the transverse phase planes.

5. Displacement and divergence

Of course a beam of particles enters the machine as a bundle of trajectories spread about the ideal orbit. At any instant a particle may be displaced horizontally by x and vertically by z from the ideal position and may also have divergence angles horizontally and vertically:

$$x' = dx/ds \quad \text{and} \quad z' = dz/ds .$$

Such mis-steering would cause particles to leave the vacuum pipe were it not for the carefully shaped field which restores them back towards the beam centre so that they oscillate about the ideal orbit. The design of the restoring fields determines the transverse excursions of the beam and the size of the cross section of the magnets and is therefore of crucial importance to the cost of a project.

6. Weak Focussing

Early synchrotrons used a weak focusing. The guide field in the bending magnets forming the circular orbit had a slight negative gradient towards the outside which implied horizontal magnetic fields of increasing strength above and below the median plane of the magnets. In the vertical plane (z, z') these field restored particles towards the median plane. Unfortunately the effect was the opposite in the (x, x') or horizontal plane but provided the gradient was not too strong, sufficient focussing could be provided by the variations in the balance between the bending force and the centripetal acceleration (which became stronger towards the outside radius of the machine). The focussing in both cases resembles that of small sphere rolling down a slightly inclined gutter with constant speed..

Now let us extend this analogy by bending the gutter into a circle rather like the brim of a hat. We provide the necessary instrumentation to measure the displacement of the sphere from the centre of the gutter each time it passes a given mark on the brim and we also have a means to measure its transverse velocity. With the aid of a computer, we might convert this information into the divergence angle shown in Fig. 5:

$$x' = \frac{dx}{ds} = \frac{v_{\perp}}{v_{\parallel}} .$$

Suppose also that we make the brim of a hat out of a slightly different length of gutter than is shown so that the number of oscillations is not an integer. We can plot a point for each arrival of the sphere in a diagram of x' against x which we call a ‘phase space diagram’ of transverse motion. The sphere has a large transverse velocity as it crosses the axis of the gutter and has almost zero transverse velocity as it reaches its maximum displacement.

7. Phase Space Ellipse

The locus of the ‘observations’ will be an ellipse (Fig. 5) and the phase will advance by Q revolutions each time the particle returns. Of course, only the fractional part of Q may be deduced from our observations since we are blind to what happens round the rest of the hat’s brim – a situation we shall find is common in the real life of accelerators.

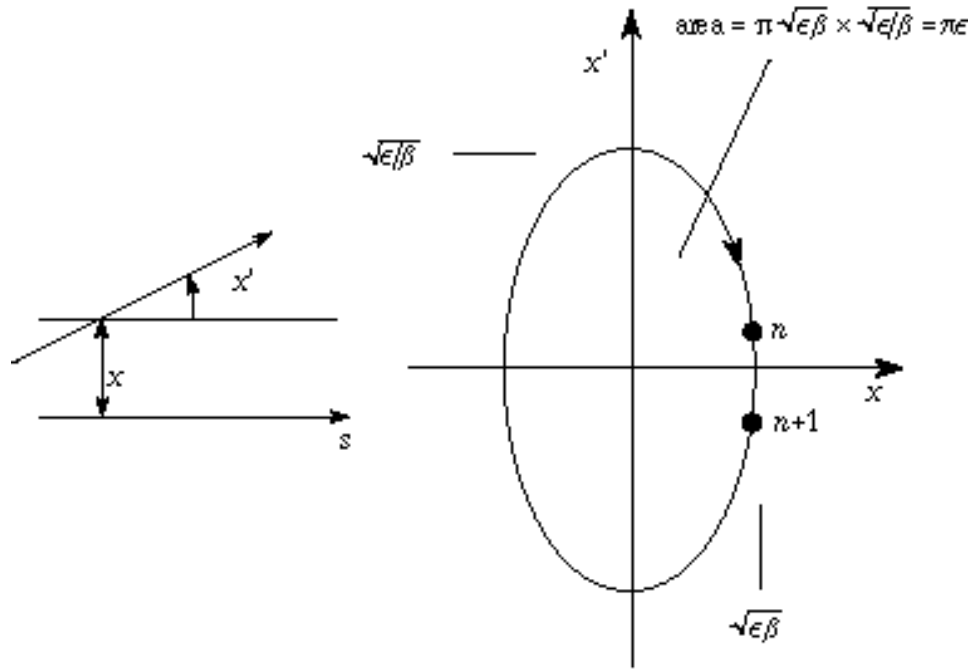


Fig. 5 The elliptical locus of a particle's history in phase space as it circulates in a synchrotron.

We now have to define some of the transverse beam dynamical quantities more rigorously. The area of the ellipse is a measure of how much the particle departs from the ideal trajectory which in the diagram is represented by the origin:

$$\text{Area} = \pi \epsilon \text{ [mm.rad].}$$

In accelerator notation we use ϵ , the product of the semi-axes of the ellipse as a measure of the area called the emittance. The emittance is usually quoted in units of π mm.mradians. The maximum excursion in displacement, the major axis, of the ellipse defined:

$$\hat{x} = \sqrt{\epsilon \beta},$$

hence

$$\hat{x}' = \sqrt{\epsilon / \beta}.$$

The quantity β is a property of the focussing system, not the beam. In the synchrotron it varies around the ring and is the envelope function we have plotted in Fig. 7. The potential well which describes a constant gradient focussing system will vary its width and curvature around the circumference of a modern synchrotron in which focussing alternates in sign. In some way β must follow this variation. Note that the aspect ratio of the ellipse is just β . We will return to these quantities when we have studied more about the alternating gradient focusing systems.

8. Strong focusing

Synchrotrons that relied on weak focussing had rather large aperture magnets. Their apertures were often almost a meter across but a team; including Stan Livingston, Ed Courant and Hartland Snyder, discovered a much more compact focussing system which hinged upon alternating the sign of the gradient around the ring. Alternating gradient, or strong focusing, greatly reduces the beam's excursions and the cross section of the magnet gap can be reduced to become comparable with a hand rather than a whole human body. Its discovery enabled Brookhaven and CERN to build the next generation of proton synchrotrons, AGS and CPS, – yet use beam pipes of only a few centimetres height and width.

The principle is illustrated in Fig. 6 which shows an optical system in which each lens is concave in one plane while convex in the other. It is possible, even with lenses of equal strength, to find a ray which is always on axis at the D lenses in the horizontal plane and therefore only sees the F lenses. The spacing of the lenses would then have to be $2f$. If the ray is also central in the lenses which are vertically defocusing, the same condition will apply simultaneously in the vertical plane. At least one particular particle or trajectory corresponding to this ray will be contained indefinitely.

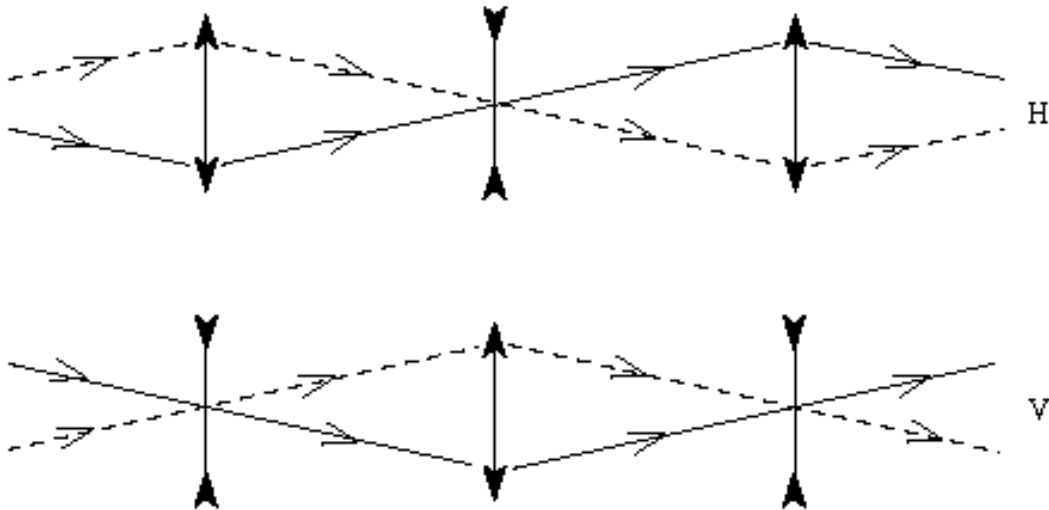


Fig. 6 Optical analogy in which an alternating pattern of lenses.

9.A FODO Lattice

A modern synchrotron or storage ring consists of pure bending magnets and quadrupole magnets or lenses which provide focusing. These are arranged around the ring in a pattern called the lattice. In Figure 7 we show one of about 100 identical “FODO” cells in the ring of the SPS. We have ignored the bending magnets and the curvature they produce and show only the alternating focussing and defocusing lenses – quadrupole magnets – which are the focussing system. While the magnetic term in constant gradient focusing was defocusing, alternating gradient focussing can be strong in both planes.

For obvious reasons this focusing structure is called FODO and in this pattern half of the quadrupoles focus, while the other half, defocus the beam. The envelope of these oscillations follows a function $\beta(s)$ which has waists near each defocusing magnet and has a maximum at the centres of F quadrupoles. Since F quadrupoles in the horizontal plane are D quadrupoles vertically, and vice versa, the two functions $\beta_h(s)$ and $\beta_v(s)$ are one half-cell out

of register in the two transverse planes. The function β has the dimensions of length but the units bear no relation at this stage to physical beam size. The reader should be clear that particles do not follow the $\beta(s)$ curves but oscillate within them in a form of modified sinusoidal motion whose phase advance is described by $\phi(s)$. The phase change per cell in the example shown is close to $\pi/2$ but the rate of phase advance is modulated throughout the cell.

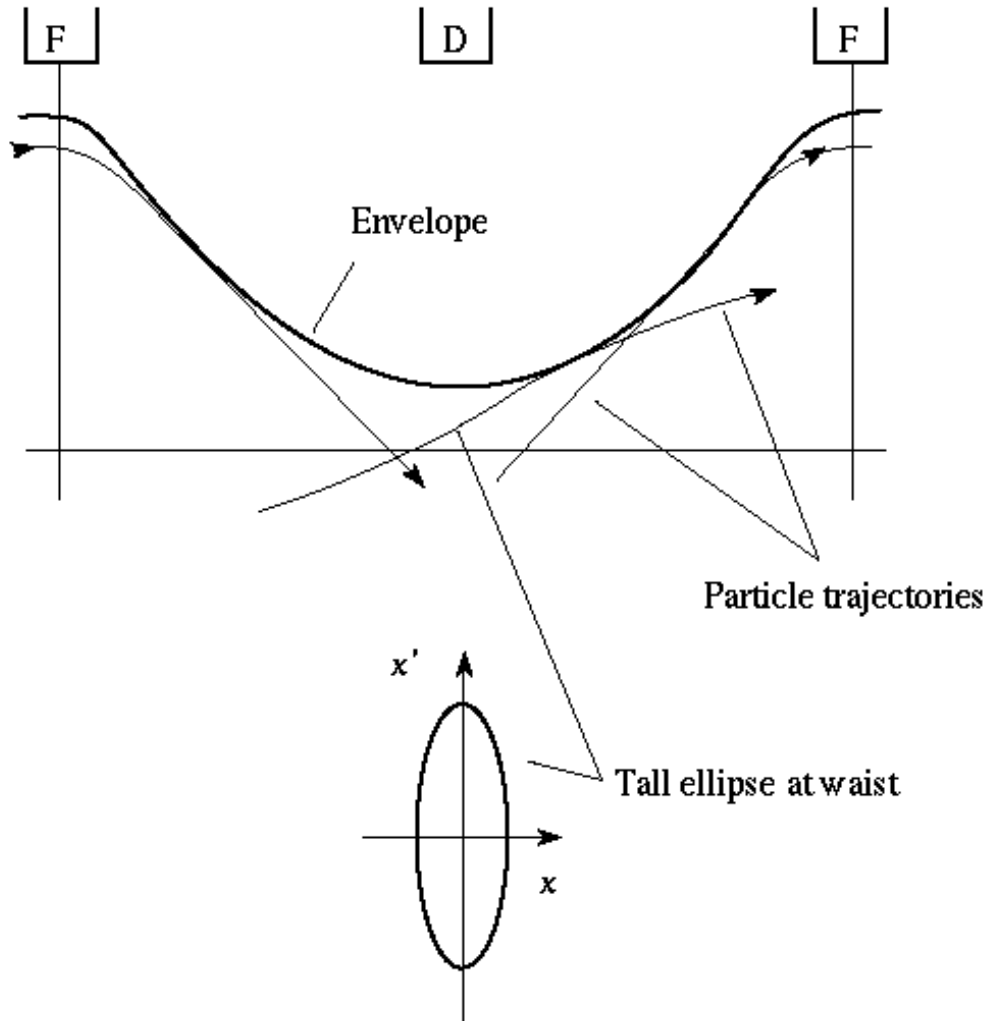


Fig. 7 The paths of particles within a FODO lattice are within the envelope of betatron motion and, like the rays of Fig. 6, are always closer in the D quadrupoles.

By suitable choice of strength and spacing of the lenses the envelope function $\beta(s)$ can be made periodic in such a way that it is large at all F quadrupoles and small at all D's. Symmetry will ensure this is true also in the vertical plane. Particles oscillating within this envelope will always tend to be further off axis in F quadrupoles than in D quadrupoles and there will therefore be a net focusing action. We have already seen that β is the aspect ratio of the phase space ellipse. At F quadrupoles the ellipse will be squat and at D quadrupoles it will be tall. In the next chapter we shall define this envelope or betatron amplitude more rigorously and establish how to calculate it for a given lattice of focusing magnets. See also Schmüser

(1987) and Rossbach *et al.* (1992). We will develop the ideas behind this description in the following paragraphs but let us first establish the concepts of the magnets which bend the beam in a circle.

10. Bending magnets

Particle physicists who analyse tracks in detectors should be well acquainted with the concept of magnetic rigidity.

$$(B\rho) = \frac{p}{e} .$$

Strictly we should use the units Newton-second for p and express e in Coulombs to give $(B\rho)$ in Tesla.metres. However, in charged particle dynamics we often talk about the 'momentum' pc which has the dimensions of energy and is expressed in units of GeV. A useful rule of thumb formula based on these units is:

$$B\rho [T.m] = 3.3356pc [GeV]$$

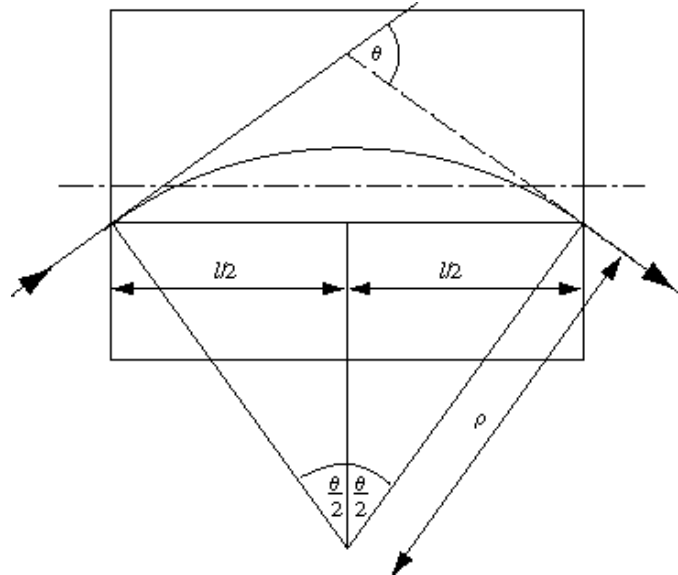


Fig. 8 Geometry of a particle trajectory in a bending magnet of length, l .

Figure 8 shows the trajectory of a particle in a bending magnet or dipole of length, l . Usually the magnet is placed symmetrically about the arc of the particle's path. One may see from the geometry that:

$$\sin(\theta/2) = \frac{l}{2\rho} = \frac{lB}{2(B\rho)} ,$$

and if $\theta \ll \pi/2$

$$\theta \approx \frac{lB}{(B\rho)}$$

11. Quadrupole magnets

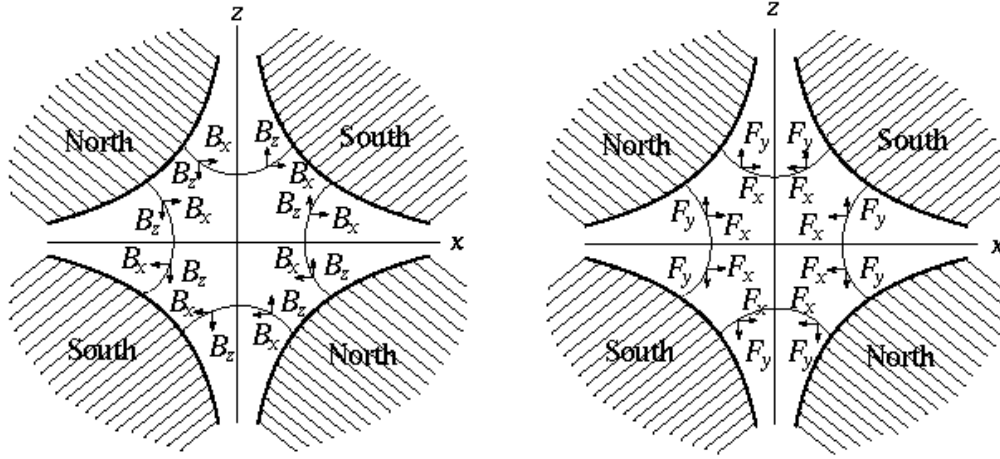


Fig. 9 Components of field and force in a magnetic quadrupole. Positive ions approach the reader on paths parallel to the s axis.

The principal focusing elements in modern synchrotrons are quadrupole magnets. The poles are truncated rectangular hyperbolae and alternate in polarity. Figure 9 shows a particle's view of the fields and forces in the aperture of a quadrupole as it passes through normal to the plane of the paper. The field shape is such that it is zero on the axis of the device but its strength rises linearly with distance from the axis. This can be seen from a superficial examination of Fig. 9 if we remember that the product of field and length of a field line joining the poles is a constant. Symmetry tells us that the field is vertical in the median plane (and purely horizontal in the vertical plane of asymmetry). The field must be downwards on the left of the axis if it is upwards on the right.

This last observation ensures that the horizontal focusing force, $-evB_z$, has an inward direction on both sides and, like the restoring force of a spring, rises linearly with displacement, x . The strength of the quadrupole is characterised by its gradient dB_z/dx normalised with respect to magnetic rigidity:

$$k = \frac{1}{(B\rho)} \frac{dB_z}{dx}.$$

The angular deflection given to a particle passing through a short quadrupole of length, ℓ and strength, k , at a displacement x is therefore:

$$\Delta x' = \theta = \ell B / (B\rho) = \ell B'x / (B\rho) = \ell kx$$

The use of x prime to indicate the divergence angle of a trajectory is defined in Fig. 5. Compare this with a converging lens in optics:

$$\Delta x' = -x / f$$

and we see that the focal length of a horizontally focusing quadrupole is

$$f = -1/(kl)$$

The particular quadrupole shown in Fig. 9 would focus positive particles coming out of the paper or negative particles going into the paper in the horizontal plane. A closer examination reveals that such a quadrupole deflects particles with a vertical displacement away from the axis – vertical displacements are defocused. This can be seen this if Fig. 9 is rotated through 90 degrees.

12.The equation of motion

In the last section we derived an expression for the change in divergence of a particle passing through the quadrupole. The strength of the quadrupole is characterised by its gradient dB_z / dx , normalised with respect to magnetic rigidity:

$$k = \frac{1}{(B\rho)} \frac{dB_z}{dx}$$

If k is negative, the quadrupole is horizontally focusing and vertically defocusing. We first look at the vertical plane. The angular deflection given to a particle passing through a short quadrupole of length ds and strength k at a displacement z is therefore:

$$dz' = -kzds$$

We can deduce from this a differential equation for the motion

$$z'' + k(s)z = 0.$$

This is Hill's Equation, a second order linear equation with a periodic coefficient, $k(s)$ which describes the distribution of focusing strength around the ring. The above form of Hill's equation applies to motion in the vertical plane while in the horizontal plane:

$$x'' + \left[\frac{1}{\rho(s)^2} - k(s) \right] x = 0$$

Here the sign before $k(s)$ is reversed so that the quadrupole defocuses. We include an extra focusing term due to the curvature of the orbit which can be significant in small rings and which is the only form of focusing in a constant gradient synchrotron.

13.Solution of Hill's equation

Hill's equation is reminiscent of simple harmonic motion but has a restoring constant $k(s)$ which varies around the accelerator. In order to arrive at a solution we first assume that $k(s)$ is periodic on the scale of one turn of the ring. The period can also be a smaller unit, the cell, from which the ring is built. The solution, like the differential equation itself, is reminiscent of simple harmonic motion:

$$x = \sqrt{\beta(s)} \mathcal{E} \cos[\phi(s) + \phi_0].$$

In simple harmonic motion the amplitude is a constant but we see that in addition to $\sqrt{\mathcal{E}}$, which can be considered an arbitrary constant, there is another amplitude component, a function $\sqrt{\beta(s)}$. Another difference with harmonic motion is that phase, $\phi(s)$, does not advance linearly

with time and with distance, s , around the ring but is a seemingly arbitrary function. Both these functions of s must have the same periodicity as the lattice and they are linked by the condition:

$$\varphi' = 1 / \beta \quad \text{or} \quad \varphi = \int ds / \beta .$$

This condition is necessary if Hill's Equation is to be satisfied, but for the moment let us just accept it. By simple differentiation we can then find

$$\begin{aligned} x &= \sqrt{\beta(s)\varepsilon} \cos [\phi(s) + \phi_0] \\ x' &= -\sqrt{\varepsilon / \beta(s)} \sin [\phi(s) + \phi_0] + [\beta'(s) / 2] \sqrt{\varepsilon / \beta(s)} \cos [\phi(s) + \phi_0] \end{aligned}$$

If we look at this function where $\beta'(s)$ is zero and hence where the second term in the divergence equation is zero we find an ellipse with semi-axis in the x - direction $\sqrt{\beta\varepsilon}$, and in the x' - direction $\sqrt{\varepsilon / \beta}$ (Fig. 5). Its area is $\pi\varepsilon$, where ε is an invariant of the motion for a single particle or the emittance of a beam of many particles.

14. Liouville's theorem

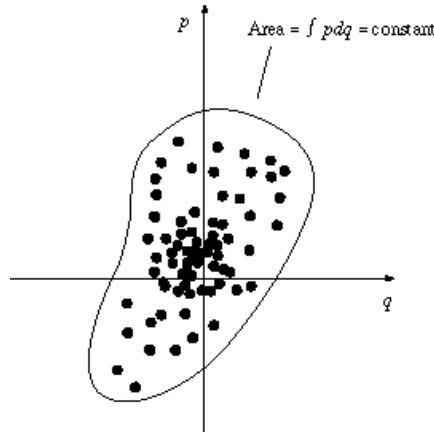


Fig. 10 Liouville's theorem applies to this contour.

Particle beams in an accelerator obey a conservation law of phase space known as Liouville's theorem. To understand this law we must think of a beam of particles as a cloud of points within a closed contour in a transverse phase space diagram (Fig. 10). Liouville's theorem tells us that this area within the contour, $A = \int p dq$, is conserved. The contour is usually, but not always, an ellipse. Earlier, in our discussion of weak focussing, we came across such an elliptical contour – the locus of a particle's motion at a place where the β function is at a maximum or minimum and where the major and minor axes of the upright ellipse are $\sqrt{\varepsilon\beta}$ and $\sqrt{\varepsilon / \beta}$. We could think of this ellipse as the locus of the particle in the beam which has the maximum amplitude of betatron motion and call its area, $\pi\varepsilon$, the emittance. We usually express emittance in units of π mm.milliradians. According to Liouville the emittance area will be conserved as the beam circulates in a synchrotron or as it passes down a transport line or along a linac whatever magnetic focusing or bending operation we do on the beam. Even though the ellipse may appear to have many shapes around or along the accelerator its phase space area will not

change (Fig. 11). At a narrow waist, near a D quadrupole (a) in Fig. 11, its divergence will be large, while in an F quadrupole (d) where the betatron function is maximum; its divergence will be small. The beam is also seen at a broad waist or maximum in the beta function and a place where the beam is diverging.

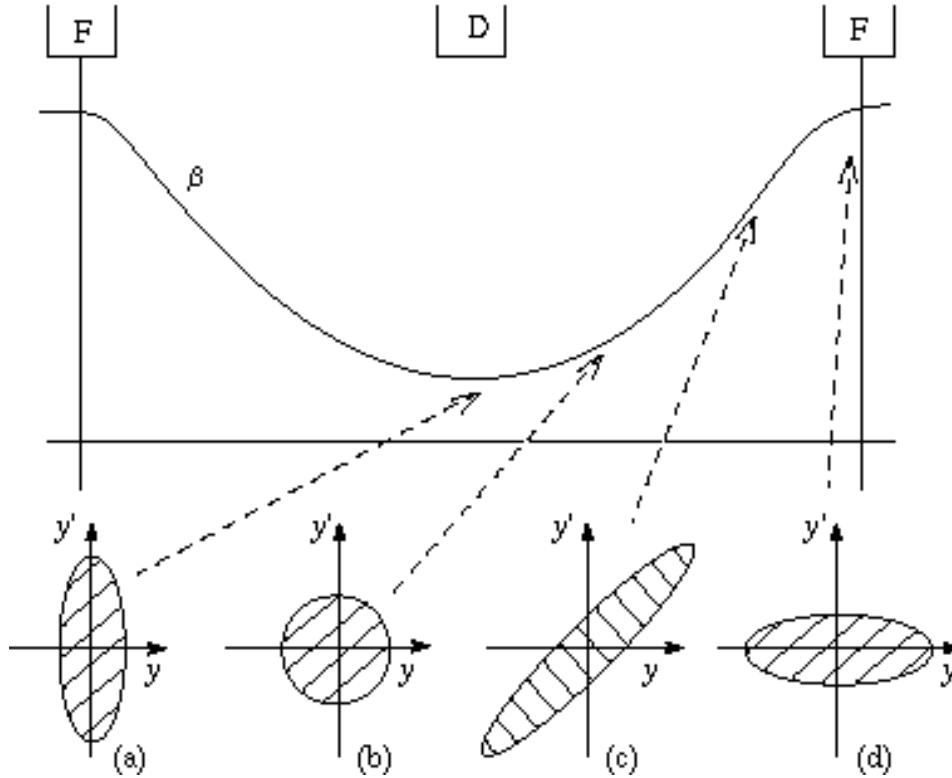


Fig. 11 How the conserved phase space appears at different points in a FODO cell.

The invariance of the area of this ellipse as we move to different points in the ring is an alternative statement of Liouville's theorem.

One word of caution – the strict version of Liouville's theorem says that:

'In the vicinity of a particle, the particle density in phase space is constant if the particles move in an external magnetic field or in a general field in which the forces do not depend upon velocity.'

and although this implies limited conservation of phase space area, it rules out the application of Liouville's theorem to situations in which space charge forces within the beam play a role or when particles emit synchrotron light – a velocity dependent effect.

However, with these precautions in mind, we may reliably apply Liouville to proton beams (which do not normally emit synchrotron light) and to electrons travelling for a few turns in a synchrotron. This is usually too short a time for electrons to emit enough synchrotron light energy to affect their transverse motion.

Louville's theorem only applies to the coordinates of Hamilton's canonical phase space (q, p) which are related to but not identical to accelerator coordinates. The quantity to be conserved as protons are accelerated is

$$\int pdq = m_0 c(\beta\gamma) \int y' dy = p_0 \int y' dy$$

where p_0 is the momentum in the direction of motion of the particle. Accelerator physicists often call this the invariant or 'normalised' emittance:

$$\varepsilon^* = \beta\gamma\varepsilon [\pi \cdot \text{mm.mrad}] .$$

As acceleration proceeds in a proton synchrotron the normalised emittance is conserved and we therefore expect the beam dimensions to shrink as, $1/p^2$ a phenomenon called 'adiabatic damping'.

The same idea applies to electrons in linac but in a damping ring the emittance will be determined by the equilibrium between the statistical inflation of the beam due to the recoil of quantum emission and the damping mechanism due to the continuous energy loss by radiating electrons as they circulate.

15. Longitudinal Motion

Earlier we tried to explain how phase stability restored errant particles in a linac, and also in a synchrotron. Suppose the particles are well below the velocity of light. A particle, *B*, (Fig. 12) which arrives late, receives an extra energy increment which will cause it to speed up and overtake the synchronous particle, *A*. In so doing, its energy defect, ΔE , grows and provided the amplitude is not too large, its trajectory will follow an ellipse in phase space. This describes this motion up and down the r.f. wave and may remind some readers of the representation of a simple harmonic oscillator, or pendulum, when plotted in a phase space diagram of velocity versus displacement. The trajectory is closed and over many turns the average deviation from the synchronous energy is zero. This *phase stability* depends upon the fact that δE is positive when $\phi - \phi_s$ is small and positive. (Le Duff 1992, Montague 1977).

Even if the oscillation is so large that the particle reaches the non-linear part of the r.f. wave and over the top of the wave, it will still be restored and oscillate about the stable phase provided it does not reach and pass the point where it receives less voltage than the synchronous particle. On this non-linear part of the curve the motion is no longer an ellipse but is distorted into a fish-shape but its trajectory is still closed and stable. However, if a particle, *C*, oscillates with such large amplitude that it falls below the synchronous voltage; an increase in ϕ will cause a negative ΔE which causes ϕ to move further away from the synchronous particle. This particle is clearly unstable and will be continuously decelerated. There is a particle which, starting at $\phi = \pi - \phi_s$, would trace out a limiting fish-shaped trajectory which is the boundary or *separatrix* between stable and unstable motion. The region within this separatrix is called the r.f. bucket and is shown in the lower half of Fig. 12.

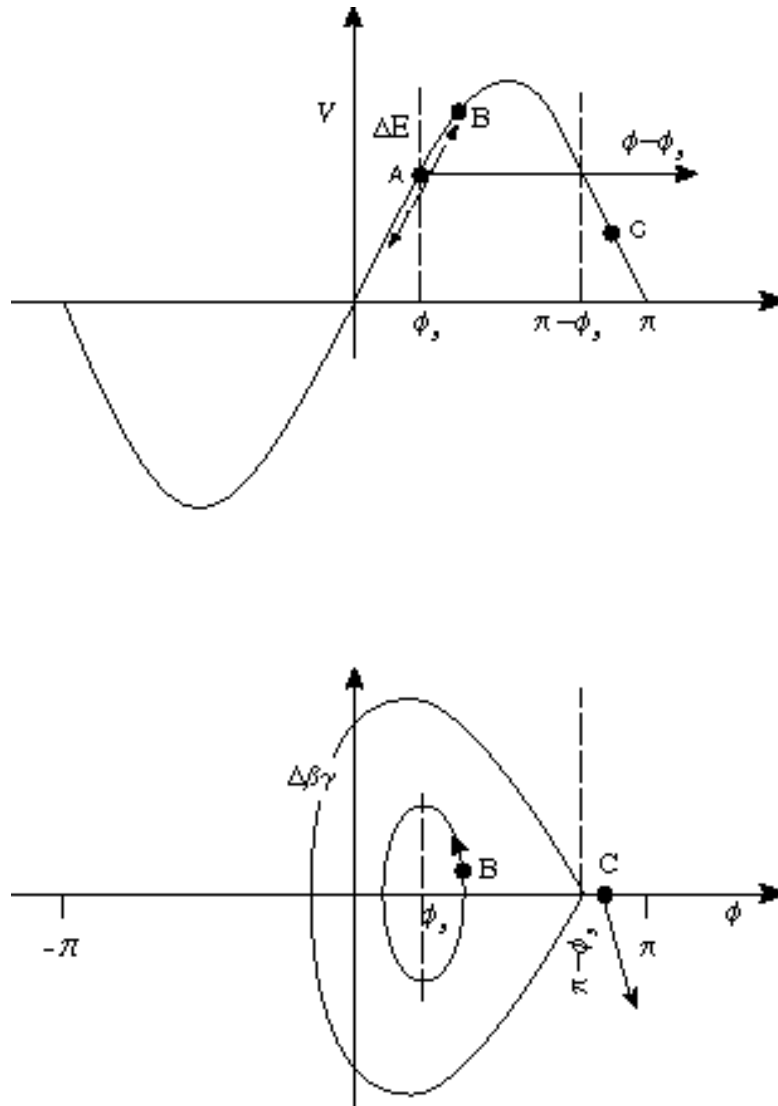


Fig. 12 The limiting trajectory for a particle in a ‘moving’ or accelerating bucket when the stable phase is not zero.

16. Phase stability for relativistic particles

The principle of phase stability works equally well for a linac as for a synchrotron. To understand this one should think of each repetitive cell of the linac as a turn of the synchrotron. However, phase stability seems to depend on the momentum (or energy) dependence of the revolution frequency. Why therefore does it work in an electron linac or synchrotron where the revolution frequency is almost that which one might calculate from the velocity of light and changes only imperceptibly with momentum? In an electron linac we need not worry for the phase of errant electrons and this, once established, cannot change from cell to cell. In a synchrotron however, the revolution time depends on the orbit length which will be larger for higher momentum particles. Thus electrons in a synchrotron, if given extra energy, will actually take longer to return to the cavity. A moments consideration should convince the reader that phase stability can be restored by accelerating on the descending side of the rf wave. To

understand this quantitatively we must know how to calculate the orbit length for an off momentum particle.

17. Closed orbit

The bending field of a synchrotron is matched to some ideal (synchronous) momentum p_0 . A particle of this momentum and of zero betatron amplitude will pass down the centre of each quadrupole, be bent by exactly 2π by the bending magnets in one turn of the ring and remain synchronous with the r.f. frequency. Its path is called the central (or synchronous) momentum closed orbit. In Fig. 7 this ideal orbit was the horizontal axis. We see particles executing betatron oscillations about it but these oscillations do not replicate every turn. In contrast the synchronous orbit closes on itself so that x and x' remain zero.

18. Orbit of a low momentum particle

We now take a bird's eye view and look at a closed orbit which differs in the horizontal plane due to the particle having a momentum defect.

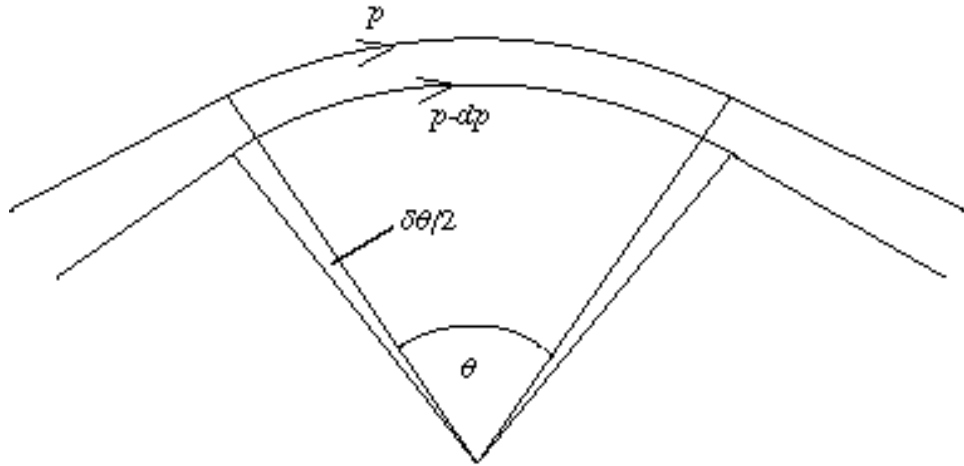


Fig. 13 Orbits in a bending magnet.

Figure 13 shows a particle with a lower momentum $\Delta p/p < 0$ and which therefore is consistently bent horizontally more in each dipole of a FODO lattice. We might argue that the total deflection in a turn of the ring, being more than 2π would cause it to spiral in. But there is a closed orbit for this lower momentum in which the extra bending forces are compensated by extra focusing forces as the orbit is displaced inwards in the F quadrupoles and less so in the D's. This is shown in Fig. 14. We may describe the shape of this new closed orbit for a particle of unit $\Delta p/p$ by a *dispersion* function $D(s)$.

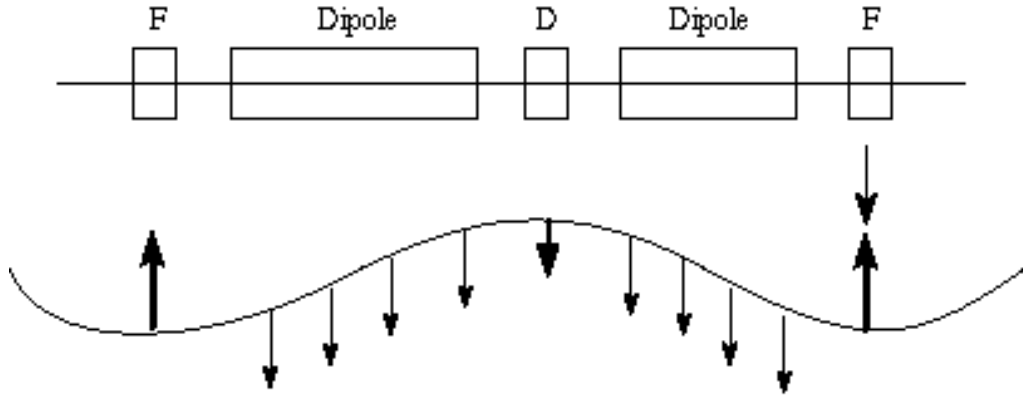


Fig. 14 The extra inward force given to a low momentum particle by the dipoles is balanced by the focusing the quadrupoles and defines a dispersion function.

This clearly means the beam will be wider if it has momentum spread. In Fig. 5.7 we see how the effect of dispersion for off momentum orbits combines with betatron motion. The betatron motion of each of the three kinds of particles $\Delta p/p < 0, = 0$ and > 0 is confined to an ellipse in physical (x, z) space. The ellipses for each momentum are separated by a distance $D(S)\Delta p/p$. Hence the minimum semi-aperture required for the beam will be:

$$a_v = \sqrt{\beta_v \epsilon_v}, a_H = \sqrt{\beta_H \epsilon_H} + \left| D(s) \frac{\Delta p}{p} \right|$$

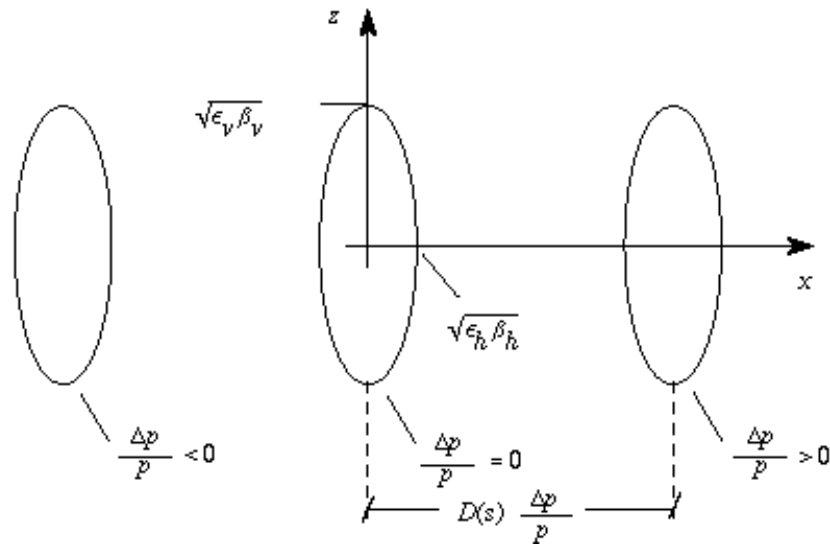


Fig. 15 The beam cross sections in real space for beams of three different momenta at a point where the dispersion function is large.

19. Synchrotron Motion

It is not hard to deduce from special relativity that the momentum may be written

$$p = m_0 c (\beta \gamma) .$$

We shall therefore use $\Delta(\beta\gamma)$ as the momentum co-ordinate in longitudinal phase space. The other co-ordinate is the particle's arrival phase, ϕ , with respect to the zero crossing of the r.f. voltage at the cavity. Let us consider the simplest case of a small oscillation in a stationary bucket, $\phi_s = 0$ when the particle is not being accelerated.

A particle with a small phase error will describe an ellipse in phase space which one may write parametrically as

$$\begin{aligned}\Delta(\beta\gamma) &= \Delta(\beta\gamma)_{\max} \sin 2\pi f_s t \\ \phi &= \phi_{\max} \cos 2\pi f_s t\end{aligned}$$

where f_s is the frequency of execution of these oscillations in phase which we call the synchrotron frequency.

In order to find the differential equation behind this motion we must first remember that the angular frequency $2\pi f$ of an oscillator is nothing other than the rate of change of phase, $\dot{\phi}$ or to be exact $-\dot{\phi}$. (The negative sign stems from the fact that ϕ is a phase lag.) For example a 5 Hz frequency changes phase at a rate of 10π radians per second. We may therefore relate the rate of change in arrival phase to the difference in revolution frequency of the particle, compared to that of the synchronous particle.

$$\dot{\phi} = -2\pi h [f(\Delta\beta\gamma) - f(0)] = -2\pi h \Delta f .$$

Note that we have multiplied by, h , the harmonic number of the r.f. since ϕ is the phase angle of the r.f. swing while $f(\Delta\beta\gamma)$ refers to the revolution frequency. Here we can use the definition of the slip factor η

$$\eta = \left(\frac{df}{f} \right) / \left(\frac{dp}{p} \right)$$

and then simply use some standard relativistic relations to end up with Δf as a function of ΔE , the energy defect with respect to the synchronous particle:

$$\Delta f = \eta f \frac{\Delta p}{p} = \eta f \frac{\Delta(\beta\gamma)}{(\beta\gamma)} = \frac{\eta f}{\beta^2} \frac{\Delta\gamma}{\gamma} = \frac{\eta f}{E_0 \beta^2 \gamma} \Delta E$$

where E_0 is the particle's total energy.

Having made the substitution, we differentiate once more to obtain a second order differential equation which we hope to resemble a simple oscillator.

$$\ddot{\phi} = - \frac{2\pi h \eta f}{E_0 \beta^2 \gamma} (\Delta \dot{E})$$

We recall from the arguments of phase stability that the extra energy given per turn to a particle whose arrival phase is ϕ will be

$$\Delta E = e V_0 (\sin \phi - \sin \phi_s)$$

and that the rate of change of energy will therefore be this times, f , the revolution frequency we can write

$$\ddot{\phi} = - \frac{2\pi e V_0 \eta h f^2}{E_0 \beta^2 \gamma} (\sin \phi - \sin \phi_s)$$

This is actually a fundamental and exact description of the motion provided the parameters should change slowly (the adiabatic assumption). Nowadays we simply integrate to find its solution numerically but to see an analytic solution for small amplitudes we set $\phi_s = 0$ and $\phi = \sin \phi$

$$\ddot{\phi} + \frac{2\pi e V_0 \eta h f^2}{E_0 \beta^2 \gamma} \phi = 0 .$$

Clearly the frequency of these synchrotron oscillations in longitudinal phase space is

$$f_s = \sqrt{\frac{|\eta| h e V_0}{2\pi E_0 \beta^2 \gamma}} f .$$

or writing $f_{rf} = hf$ we could equally express

$$f_s = \sqrt{\frac{|\eta| h e V_0}{2\pi E_0 \beta^2 \gamma h}} f_{rf}$$

20. Gradient errors

We now have covered the ideal behaviour of particles in an ideal machine as they are focussed by the lattice while their phase is stabilised in the longitudinal plane. We now turn to the effect of field imperfections. Quadrupoles also have errors and understanding the effect of these gradient errors is a useful preparation for the study of non-linear errors. We represent a ring of magnets as a circle in Fig. 6.11 and the matrix for one turn starting at A as:

This is a transport matrix which operates on the displacement, divergence vector (y, y') takes it through one turn of the machine. the parameters α, β, γ are called Twiss parameters and are all functions of the envelope parameter β and its derivative. The trace of this matrix Now consider a small gradient error which afflicts a quadrupole in the lattice between B and A. The unperturbed matrix for this quadrupole is m_0

$$m_0 = \begin{pmatrix} 1 & 0 \\ -k_0(s)ds & 1 \end{pmatrix} ,$$

and when perturbed the quadrupole matrix is

$$m = \begin{pmatrix} 1 & 0 \\ -[k_0(s) + \delta k(s)]ds & 1 \end{pmatrix} .$$

The unperturbed transfer matrix for the whole machine includes m_0 . To find the perturbed transfer matrix we make a turn, back-track through the small unperturbed quadrupole (m_0^{-1}) and then proceed through the perturbed quadrupole (m). Translating this into matrix algebra we have,

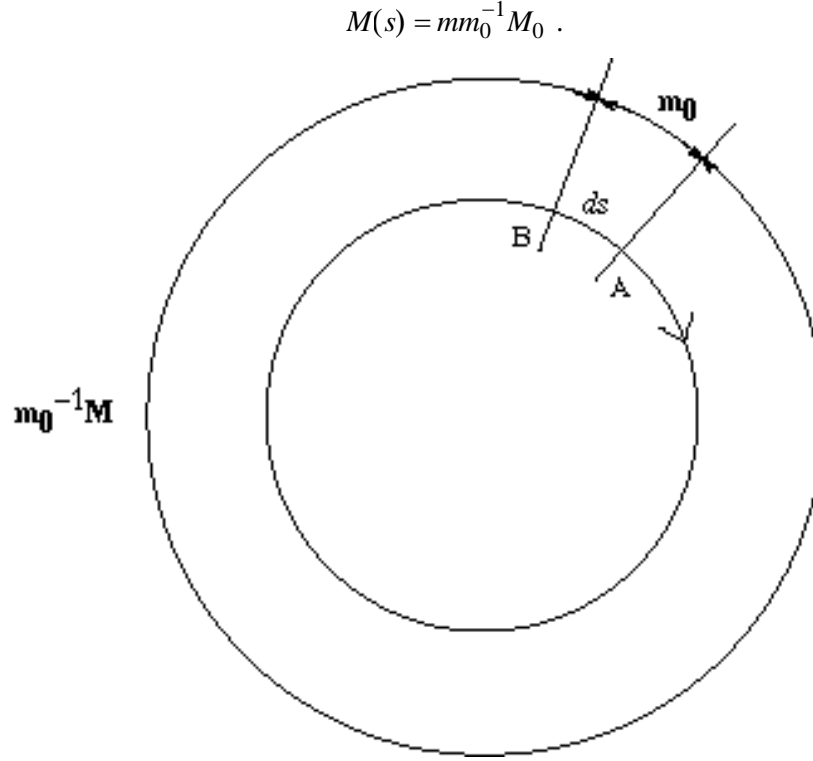


Fig. 16 Matrix representation of a small quadrupole, m_0 , subject to an error which is a component of the matrix for the whole ring, M .

Now

$$mm_0^{-1} = \begin{pmatrix} 1 & 0 \\ -\delta k(s)ds & 1 \end{pmatrix}$$

So

$$M = \begin{pmatrix} \cos \phi_0 + \alpha_0 \sin \phi_0 & \beta_0 \sin \phi_0 \\ -\delta k(s)ds(\cos \phi_0 + \alpha_0 \sin \phi_0) - \gamma \sin \phi_0 & -\delta k(s)ds\beta_0 \sin \phi_0 + \cos \phi_0 - \alpha_0 \sin \phi_0 \end{pmatrix}$$

Now $TrM = 2\cos \phi$. So the change in $\cos \phi$ is

$$\Delta(\cos \phi) = -\Delta\phi \sin \phi_0 = \frac{\sin \phi_0}{2} \beta_0(s) \delta k(s) ds$$

$$2\pi\Delta Q = \Delta\phi = \frac{\beta(s)\delta k(s)ds}{2}$$

Since betatron phase is not involved in this equation we can just integrate around the ring to obtain

$$\Delta Q = \frac{1}{4\pi} \int \beta(s) \delta k(s) ds .$$

This is a very useful result and is worth memorising. It is perhaps surprising that the change is independent of the phase of the perturbation.

21. The working diagram

This is simply a diagram with Q_h and Q_v as its axes. The beam can be plotted on it as a point but because there is a certain Q -spread among particles of different momenta we had better give the point a finite radius ΔQ (Fig. 6.12).

On the diagram we see a mesh of lines which mark danger zones for the particles. We have hinted above that if Q in either the vertical or the horizontal plane is a simple vulgar fraction, then

$$nQ = p ,$$

where n and p are integers and $n < 5$, a resonance takes over and walks the proton out of the beam. In general this is true when

$$\ell Q_H + m Q_V = p$$

where $|\ell| + |m|$ is the order of the resonance and p is the azimuthal frequency which drives it.

This equation just defines a set of lines in the Q diagram for each order of resonance and for each value of the integer p . Fig. 17 shows these lines for the SPS.

Somehow, by careful adjustment of the quadrupoles in the lattice and by keeping the Q -spread (chromaticity) small, we must coax the beam up to full energy without hitting the lines. To make things more difficult, each line has a finite width, proportional to the strength of the imperfection which drives it. In some cases we must compensate the imperfections with correction multipoles to reduce this width.

We will discuss resonances and their correction in more detail later but before this, a word about chromaticity.

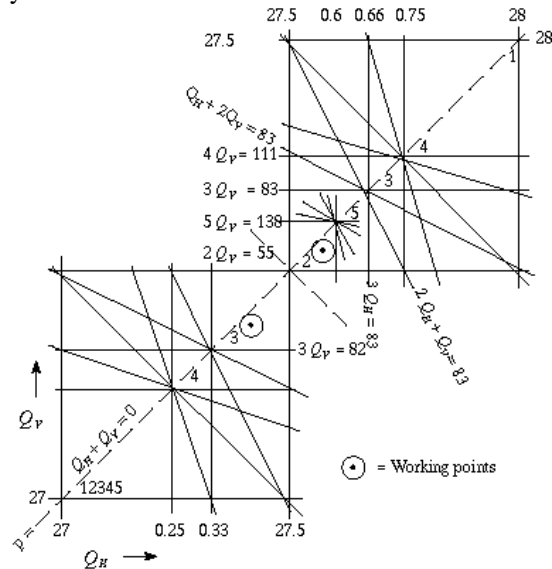


Fig. 17 SPS working diamond.

22. Multipole fields

Before we come to discuss the non-linear terms in the dynamics, we need to describe the field errors which drive them. The magnetic *vector potential* of a magnet with $2n$ poles in Cartesian co-ordinates is:

$$\mathbf{A} = \sum_n \mathbf{A}_n f_n(x, z)$$

where f_n is a homogeneous function in x and z of order n . If the magnet is long we can ignore end fields and assume the vector potential is entirely parallel to the magnet axis, A_y .

Table 1 Cartesian solutions of magnetic vector potential

Multipole	n	Regular f_n	Skew f_n
Quadrupole	2	$x^2 - z^2$	$2xz$
Sextupole	3	$x^3 - 3xz^2$	$3x^2z - z^3$
Octupole	4	$x^4 - 6x^2z^2 + z^4$	$4x^3z - 4xz^3$
Decapole	5	$x^5 - 10x^3z^2 + 5xz^4$	$5x^4z - 10x^2z^3 + z^5$

Table 1 gives $f_n(x, z)$ for low-order multipoles, both regular and skew. Figure 18 shows the distinction between regular and skew multipoles. We can obtain the function for other multipoles from the binomial expansion of

$$f_n(x, z) = (x + iz)^n$$

The real terms correspond to regular multipoles, the imaginary ones to skew multipoles.

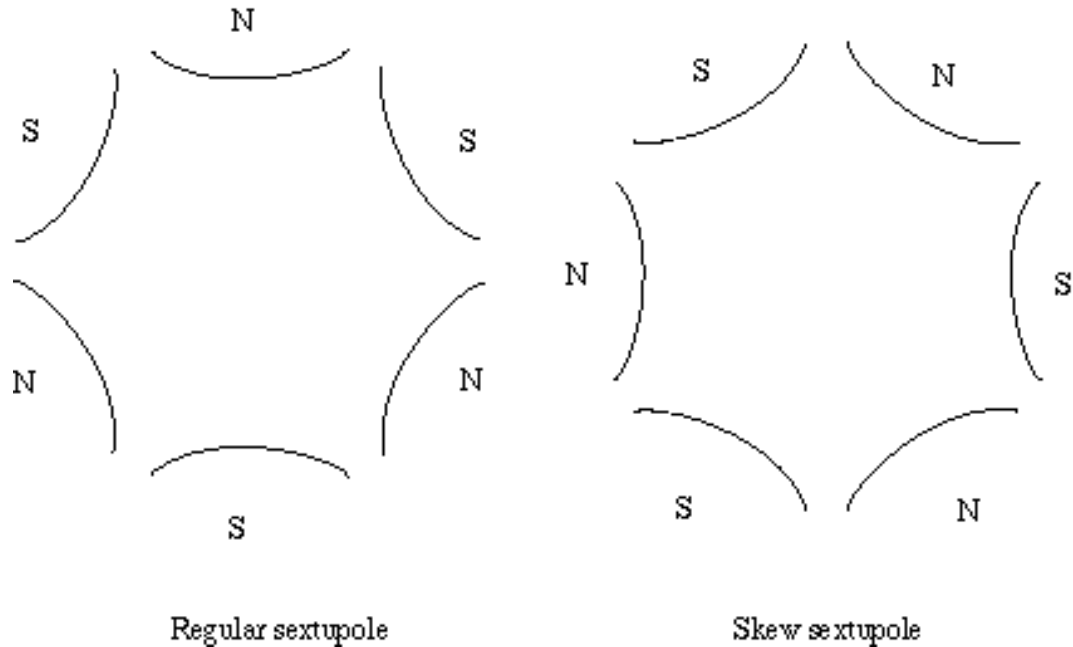


Fig. 18 Pole configurations for a regular sextupole and a skew sextupole.

Of course we are interested in the field and for a long regular magnet with a purely paraxial potential.

$$B_z(z=0) = -\frac{\partial A_x}{\partial x} = -\sum_{n=1}^{\infty} n A_n x^{(n-1)} = -\sum_{n=1}^{\infty} \frac{1}{(n-1)!} \left(\frac{d^{(n-1)} B_z}{dx^{(n-1)}} \right)_0 x^{n-1}.$$

We could equally well express the field as a Taylor expansion and by equating these expressions find:

$$A_n = \frac{1}{n!} \left(\frac{d^{(n-1)} B_z}{dx^{(n-1)}} \right)_0.$$

A more modern convention (in Europe) is to speak of multipole coefficients, b_n , for normal components and a_n for skew components, where R_n is some reference radius (10 mm for the LHC), B_1 is the magnitude of the nominal dipole field, B_y , and $Z = x + iz$.

$$B + iB_x = B_1 \sum_n (b_n + ia_n) (Z/R_r)^{n-1}.$$

The suffix, $n = 1$ for the dipole, 2 for the quadrupole and 3 for sextupole etc. Note that US notation starts with $n = 0$ for the dipole and different laboratories use other reference radii. In spite of the possibilities for confusion this has the advantage that the coefficients are a measure of the tolerated fraction of field error inside the reference circle where beam is supposed to be stable. It is therefore much easier to compare the designs of the different machines.

23. Resonant conditions

As we have said earlier, we must steer Q well away from certain fractional values which can cause the motion to resonate and result in loss of the beam. To understand how some Q values are dangerous, let us return to the case of closed orbit distortion. Imagine a perturbation due to a small dipole:

$$\Delta y' = \frac{\Delta(B\ell)}{B\rho}$$

If Q is an integer value this will become infinite after a number of turns. What happens physically is that the beam receives a kick at the same phase on every turn and just spirals outwards. An error in gradient can have the same effect if the Q value is close to one of the lines

$$\left. \begin{array}{l} 2Q_h = p, \quad 2Q_v = p \\ Q_h = Q_v = p, \quad Q_h + Q_v = p \end{array} \right\}$$

where p is an integer.

At this stage in the description of transverse dynamics we can only hint at the explanation for this. Particles spiral outwards in phase space if the perturbation has the same effect on each turn and this can only build up in this way if the particle returns to the same point in phase space on each turn ($Q = p$). The perturbation from a dipole is independent of the transverse displacement but a quadrupole error has field proportional to x and if a particle makes half turns in phase space it will see alternately positive and negative kicks in divergence but both will reinforce the growth. This can happen if Q is a half integer. One may extend this argument to understand why sextupole errors which have a quadratic x dependence excite third-order resonances as they are called near the lines.

$$\left. \begin{aligned} 3Q_h &= p \\ 2Q_h + Q_v &= p \\ Q_h + 2Q_v &= p \\ 3Q_v &= p \\ 2Q_h - Q_v &= p \\ Q_h - 2Q_v &= p \end{aligned} \right\}$$

24. Chromaticity

The steering of Q depends on careful regulation of quadrupole and dipole power supplies. In fact, much of the setting up time of a large circular accelerator is devoted to tuning Q to be constant as the fields and energy rise. Once beam has been accelerated the problem becomes one of reducing all effects which produce a spread in Q among the particles in the beam. The limit to this is usually reached when beam intensity is high enough to cause space-charge focusing effects, whose strength varies with the local beam density and which modulate Q as particles move up and down the bunch. Before reaching this limit one must correct the tune spread due to momentum – the chromaticity. This is exactly equivalent to the chromatic aberration in a lens. It is defined as a quantity Q'

$$\Delta Q = Q' \frac{\Delta p}{p}.$$

The chromaticity arises because the focusing strength of a quadrupole has $(B\rho)$ in the denominator and is therefore inversely proportional to momentum:

$$k = \frac{1}{(B\rho)} \frac{dB_z}{dx}.$$

A small spread in momentum in the beam, $\pm\Delta p/p$, causes a spread in focusing strength:

$$\frac{\Delta k}{k} = \mp \frac{\Delta p}{p}.$$

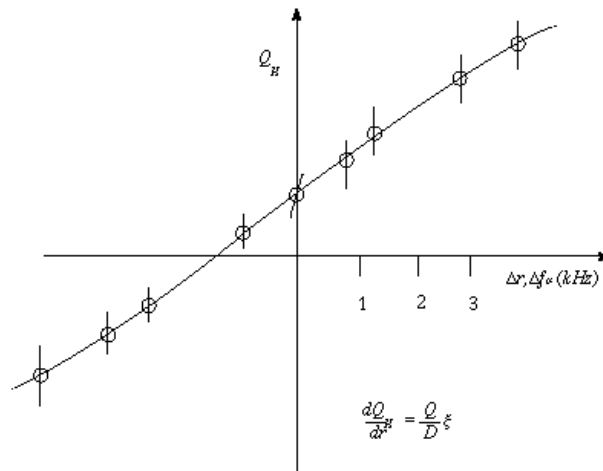


Fig. 19 Measurement of variation of Q with mean radius made by changing the r.f. frequency.

An equation we derived earlier in the section on gradient errors

$$\Delta Q = \frac{1}{4\pi} \int \beta(s) \delta k(s) ds .$$

enables us to calculate Q' rather quickly:

$$\Delta Q = \frac{1}{4\pi} \int \beta(s) \delta k(s) ds = \left[\frac{-1}{4\pi} \int \beta(s) k(s) ds \right] \frac{\Delta p}{p} .$$

The chromaticity Q' is just the quantity in square brackets. To be clear, this is called the natural chromaticity. For most alternating gradient machines, its value is about $-1.3Q$. Of course there are two Q values relating to horizontal and vertical oscillations and therefore two chromaticities.

One way to correct chromaticity is to introduce some focusing which gets stronger for the high momentum orbits near the outside of the vacuum chamber – a quadrupole whose gradient increases with radial position is needed. A sextupole whose field is:

$$B_z = \frac{B''}{2} x^2$$

in a place where there is dispersion it will introduce a normalised focusing correction:

$$\Delta k = \frac{B'' D}{(B\rho)} \frac{\Delta p}{p} .$$

We use an earlier expression for the effect of this Δk on Q and obtain

$$\Delta Q = \left[\frac{1}{4\pi} \int \frac{B''(s)\beta(s)D(s)ds}{(B\rho)} \right] \frac{\Delta p}{p} .$$

To correct chromaticity we have to make the quantity in the square bracket balance the chromaticity. There are of course two chromaticities, one affecting Q_H , the other Q_V and we must therefore arrange for the sextupoles to cancel both. For this we use a trick which is common and will crop up again in other contexts. Sextupoles near F-quadrupoles where β_x is large affect mainly the horizontal Q , while those near D-quadrupoles where β_z is large, influence Q_v .

25. Luminosity and performance of colliders

Many of the factors which limit performance are common to both linear and circular colliders. Luminosity, the final focus design, the beam-beam interaction and even the intensity limitations due to instabilities are calculated with the same techniques. We now review each of these limits in turn. Luminosity is the primary figure of merit for any collider and it is worth examining it precisely. Take a probe particle in one beam which sees the oncoming beam as a cloud of N particles, each a disc of a cross-section, σ (Fig. 20). The probability of the rare process occurring is just the fraction of the beam's area that is occluded:

$$P = \frac{N}{A} \sigma = \ell \sigma .$$

In this trivial case of a single encounter of one particle passing through a beam, $\ell = N/A$, is a sort of luminosity for this one encounter. It is independent of the cross-section

under study and only depends on the beam geometry. We may think of it as a probability of producing an event normalised to unit interaction cross section.

In practice the probe beam has as many particles as its opposing target and such encounters occur as often as the many bunches in the circulating beams in a storage ring meet each other. The luminosity for two such equal beams colliding is:

$$L = \frac{N^2 f_b}{A}$$

where: N is the number of particles per bunch, f_b , is the bunch frequency, and A , the transverse beam area at the crossing.

A luminosity of $10^{33} \text{ cm}^{-2} \text{ s}^{-1}$ will produce one event per second from a process for which the cross section is 10^{-33} cm^{-2} .

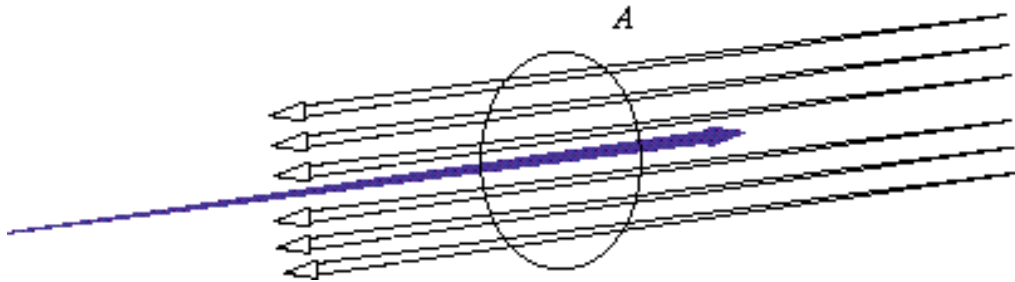
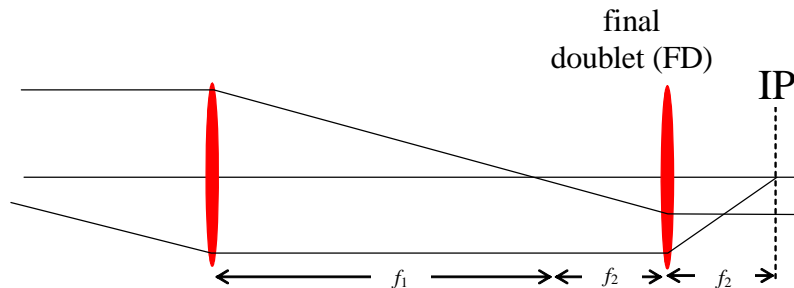


Fig. 20 A probe particle encounters a target – a beam of particles with cross-sectional area A travelling in the opposite sense.

26. designing the final focus optics to maximise luminosity

The cross sectional area of the beam, A , is proportional to $\sqrt{\epsilon_h \beta_h \epsilon_v \beta_v}$, while emittances in the horizontal and vertical planes in the case of a lepton collider are determined by the damping ring which is designed to make them as small as possible. The betatron amplitudes β_h, β_v , can be made small by using the magneto-optical equivalent of a telescope. Figure 21 shows such a telescope.



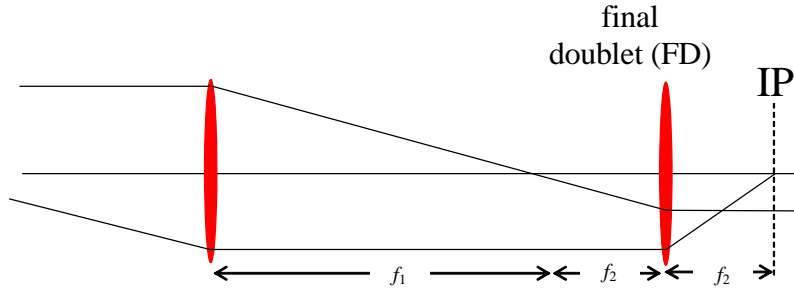


Fig. 21 A telescope to produce a low beta at the interaction point. Each lens consists of at least two quadrupoles of opposite sign which form a doublet.

There is a limit to how low a beta may be useful in practice. The beam shape at the interaction point resembles an hour-glass. The extent of the region over which beta is small is roughly equal to the value beta in meters. If the beta is smaller than the bunch length, σ_z colliding bunches many of the particles collide on either side of the waist where the beam is significantly larger than at the waist. We see this in Figure 22

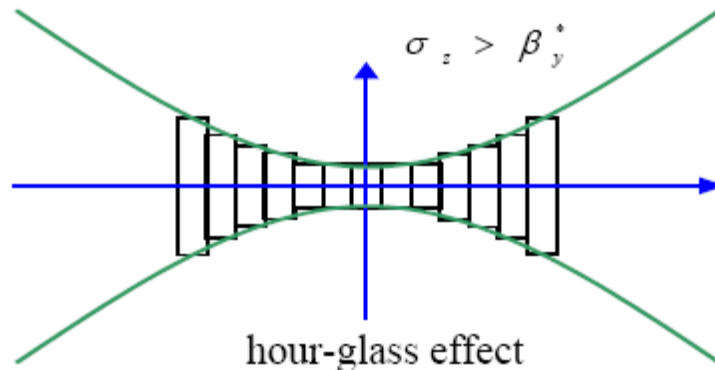


Fig. 22 The rectangles show where particles at different distances from the centre of the bunches collide. The rectangle at the extreme right will be where a particle near the beginning of the bunch coming from the left collides with the particle near the end of the bunch coming from the right.

In adapting the beam from the end of the main linac to the interaction point the lattice designer must suppress any dispersion there is in the beam leaving the end of the linac so that all momenta collide in the same horizontal line. Another momentum dependent property to be corrected is the chromaticity. We see in Figure 23 how this was done for the SLAC Linear Collider.

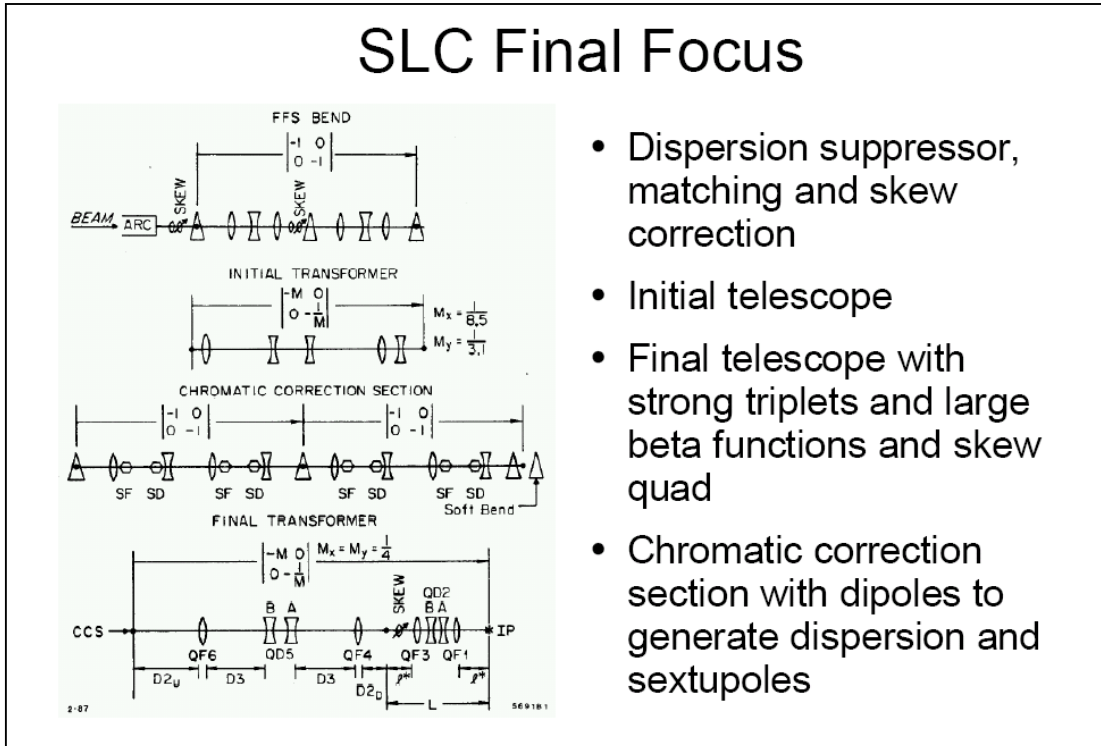


Fig. 23 The description of the final focus optics for the SLAC Linear Collider taken from their design description.

27. Intensity limitations which restrict luminosity

Obviously the number of colliding particles (and the frequency of the encounter) is crucial. In either a circular or linear collider, one of the fundamental limits to luminosity will be the number of particles per colliding bunch. This will usually be limited by a collective instability. Let us look at the general properties of such instabilities. They are caused by bunches of particles leaving wake fields behind as they pass through cavities. In the case of a linear collider the wakes will cause currents to flow in the walls of the accelerating structure and these currents will set up accelerating, decelerating or deflecting fields which act back on the beam. The particles in the bunch will be disturbed and if this disturbance is of the right phase such that it produces even bigger wake field, the instability will grow exponentially with positive feedback.

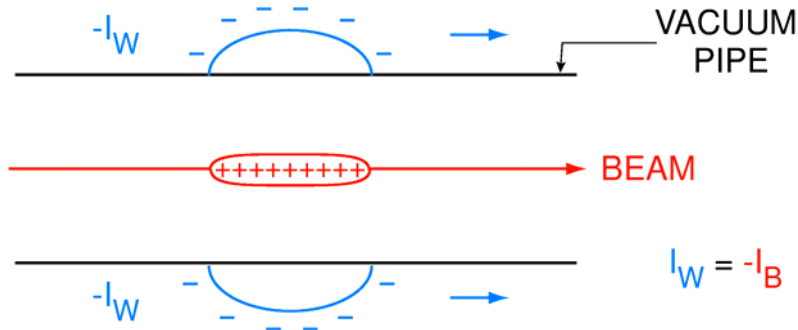


Fig. 24 A very simple example of how image charges and fields are set up in a conducting beam pipe by a passing bunch.

In Figure 24 we see a very simple example of how image charges and fields are set up in a conducting beam pipe by a passing bunch. In this case the image currents are in phase with the beam current, any field they produce acts back on the beam in phase and cannot excite longitudinal motion. However, the beam may pass through a cavity, exciting field in the cavity at its resonant frequency. Such fields will act back on the beam, and if delayed in phase can act on the beam to excite longitudinal oscillations. The strength of this reaction on the beam can be characterised by an impedance which is just the ratio of the voltage experienced by the beam divided by its own current. The instability of the beam is then like the response of a resonant LC circuit to an oscillating wave at its input. The impedance can be resistive, capacitive or inductive depending on the phase of the reactive fields on the beam. The impedance has to be calculated from a three dimensional solution of the waves excited in the pipe or cavity by the beam's passage

We can plot the impedance as a point in an Argand diagram (Figure 25).

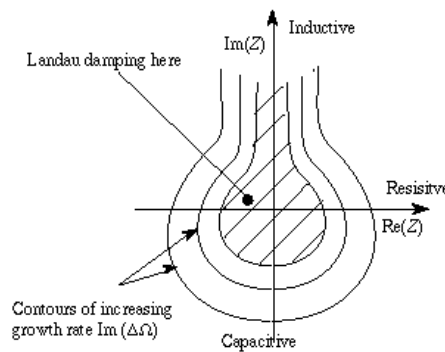


Fig. 25 Instability diagram with damping.

POS(LCPS2009)006

The contours show lines of constant growth rate for the instability – the larger the impedance, further they are from the origin and the stronger the growth rate. The shaded region in the centre is where the growth rate is too slow to be sustained in the face of a competing damping mechanism – Landau damping.

28. Landau damping

The above treatment and the diagram make an important assumption: that all the particles in the beam oscillate with the same frequency Ω_0 . However, if this condition is relaxed to allow for a spread in natural oscillation frequency due to, say, momentum spread in the beam, instability is suppressed for small values of β . Any collective motion becomes confused because of the frequency spread on a time scale which is shorter than the time it takes to develop instability. (Hereward 1976). Remember that we rely upon the motion of the centre of charge of all the particles to excite the cavity. Figure 26 illustrates how two particles of different revolution frequency get out of step.

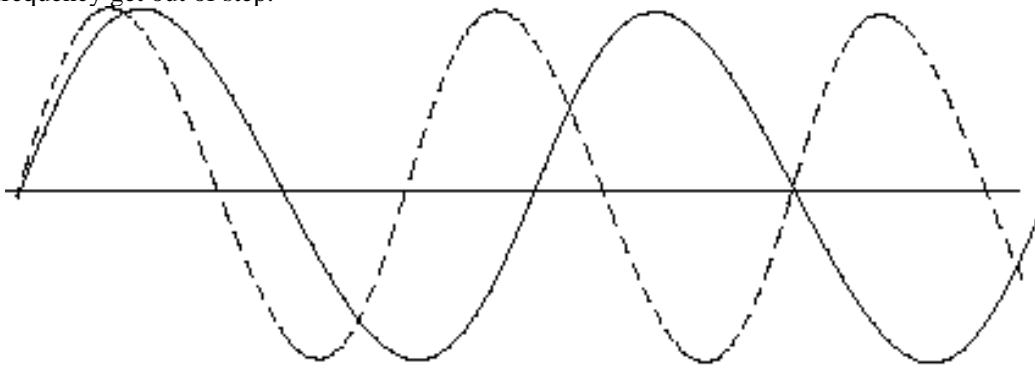


Fig. 26 Decoherence of two particles.

We can think of the shaded area in Fig. 9.9 as a stable region where Landau damping is stronger than the instability. Thus, provided the impedance is inductive, a modest resistive component is tolerable but beyond a certain X value a threshold in stability is passed and growth at a rate characterised by β occurs.

There are a multitude of kinds of instability. Above we have described a longitudinal mode but there are many frequencies at which resonance can occur and fields may act back on a single bunch or on a train of bunches. In the latter case the number of modes is comparable to the number of bunches in the train. Resonant objects around the beam can react at higher harmonics of their fundamental frequency and transverse instabilities can be setup by deflecting modes in cavities. Linear colliders have modes which are analogous to those in circular machine and others that are peculiar to linacs. Inevitably when all have been considered and the impedances which drive them minimised, there is an irreducible handful of modes which can limit colliding bunch intensity and therefore luminosity.

29. Beam-Beam Effect

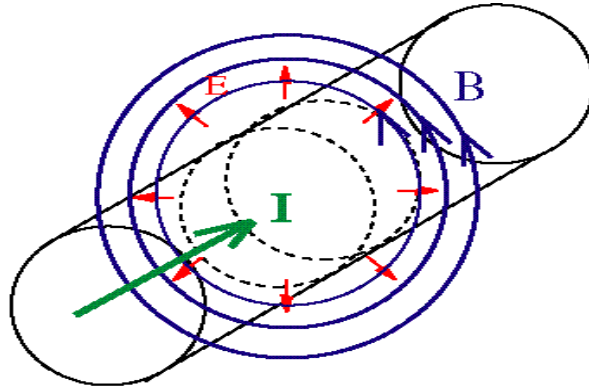


Fig. 27 The electromagnetic fields that surround a beam of particles

Figure 27 shows the electric and magnetic fields produced by a beam approaching oncoming particles in a linear collider. These fields produce a linear focussing force which helps pinch both electron and positron beams to enhance luminosity. This is equivalent to the linear beam-beam tune shift in a circular collider.

The fields shown in Figure 27 tend to have a linear focusing effect on an oncoming beam of particles with the opposite sign. This focussing (the so-called enhancement factor, H) tends to pinch both beams down to a narrower focus which increases the luminosity. However it is also a non linear force which generates a “disruption” of the beam characterised by the quantity D below::

$$D = \frac{r_e N \sigma_z}{\gamma \sigma_x^2}$$

the disruption factor has to be kept within limits but appears in the expression for the enhancement, H.

$$H = 1.6 + 0.43 [\ln(D/A)]^2,$$

This effect, which is unity at zero current, improves the luminosity, and this is expressed by:

$$L = H \frac{f N^2}{4\pi \sigma_x \sigma_y}$$

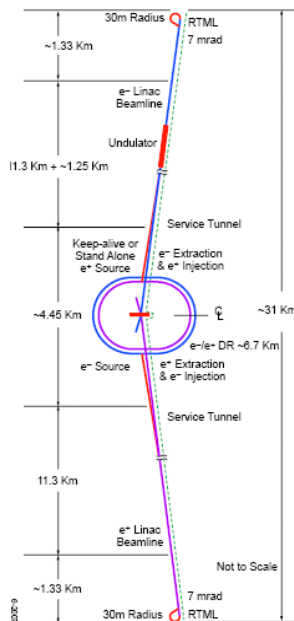
Another intensity dependent effect which limits luminosity is the energy spread in the interacting particles as they lose energy to synchrotron radiation, bent by the in the electromagnetic field of the oncoming beam .

$$\frac{\Delta E}{E} = \frac{8 r_e^3 \gamma N^2}{21 \pi^{1/2} \sigma_z \sigma_x \sigma_y} H \Delta .$$

Limits on this “beamstrahlung” are determined by the tolerable uncertainty accepted by the experimenters.

30.Linear colliders and beam power.

There is an essential difference between linear and circular colliders. Particles, once accelerated to high energy in a circular collider may be re-used almost indefinitely producing a new encounter each time they circulate, and the beam power – the product of the number of particles, the charge they carry and the voltage to which they are accelerated – is not wasted. On the other hand in a linear collider a fresh batch of particles must be accelerated for each encounter. The beam power must be continuously renewed via the accelerating cavities which form the linac and which must ultimately be powered from the electricity supply to the site. Other restrictions come from the fact that the frequency of encounter can be only as fast as the linac can cycle. Not only must the efficiency of the r.f. system which converts power from the wall plug into beam power



be high, but the voltage gradients in the accelerating cavities must be exceptional if the linacs

Figure 28 the ILC layout

are to be only 20 km or so in length. Higher frequency linac structures help, since the energy stored in their fields is smaller and they can be run at higher accelerating

POS(LCPS2009)006

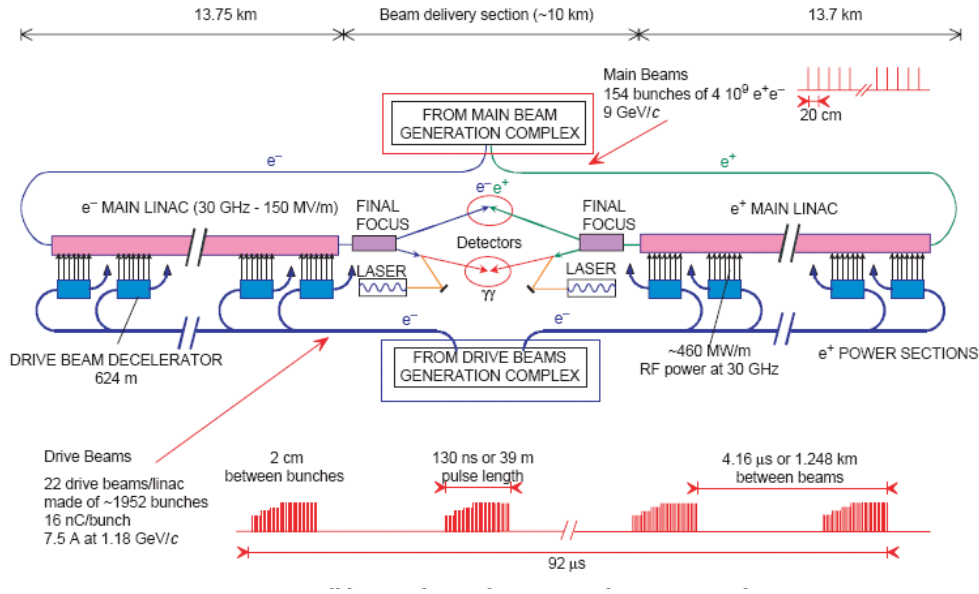


Fig. 29 Layout of CLIC at 3TeV centre of mass energy.

field so that the total length of the linac may be shorter. For the highest frequencies it will also be necessary to develop r.f. power sources at frequencies far beyond the bands where power tubes for normal telecommunication use are available. CERN, as part of a world-wide collaboration, is studying the design of a linear collider of at least 2×1000 GeV, which may well be the minimum required for a significant step into the future. (Figure 29)

It is not sufficient just to design an accelerating structure with maximum MeV per metre. The beam should sweep up as much as possible of the electrical energy stored in the cavity or the residual energy will be wasted on warming the copper walls of the cavity. If the cavity energy stored per metre of cavity is W' and if N electrons of charge, e , are accelerated by a field E when passing through the cavity, the sweeping efficiency is:

$$\eta = \frac{eNE}{W'}$$

The r.f. power needed is the power of the beam divided by the sweeping efficiency:

$$P_{RF} = \frac{P_{beam}}{\eta} = \frac{NfeU}{\eta}$$

Here U is the overall energy gain of the particles and f is the repetition rate of the device. We can imagine that small high-frequency cavities are likely to be a good idea because the number of Joules, W , needed to fill the interior of the cavity is minimised. In addition there are the arguments for high frequency, based on reducing the overall length of the device. If we try to sweep too much energy from the cavity, the sagging of energy loss between the head and the tail of the bunch will be large but this may be compensated to some extent by splitting the bunch into a train of bunches and filling the cavity with energy as they pass.

The parameters f and N in the formula for power appear again in the expression for the luminosity:

$$L = H \frac{fN^2}{4\pi\sigma_x\sigma_y} .$$

Here H is the enhancement of the luminosity due to the electromagnetic forces in the collision, an effect, which is unity at low intensity and is beyond the scope of this description. Comparing the equations for efficiency and power it looks as if we are free to reach any luminosity by increasing N and reducing f in proportion to keep a constant r.f. power. However, there is a limit to this, because N^2 appears not only in the equation for L but in an equation that describes the energy spread in the electrons as they radiate photons meeting the bunch of positrons coming the other way. This beamstrahlung – as it is called – is caused by the acceleration due to the electromagnetic field of the oncoming bunch. If the fractional energy spread is greater than 20% to 30%, it will be impossible to define the energy of the incoming particles with the precision needed to interpret the event.

31. Conclusion

We have reviewed the basic theory of particle accelerators. More is to be found in the References [1-5] below. Our aim has been to understand enough about the workings of linear colliders as well as the damping rings which are an essential feature of both CLIC and ILC. The difference between these two great proposals stems mainly from the choice of low frequency superconducting cavities for ILC, and for CLIC: room temperature cavities at a higher frequency and driven, not by klystrons, but by a separate low energy drive beam linac. The design paths of the two machines have diverged from there leaving CLIC with a possible energy in the centre of mass three times that of ILC.

The two ideas started off with R&D programs matched to an earlier approval date for ILC than for CLIC. This is rapidly being overtaken by events and it may be that when LHC results inform us of the desirable energy for a linear collider CLIC will be the more attractive option. As I write we are all still wondering how long the delay in start up and the subsequent delays in operation will be for LHC as a result of superconductivity and whatever happens, one should carefully re-evaluate the wisdom of extensive use of superconducting element in ILC. The Tevatron, HERA and latterly, LHC were forced to exploit this technology to reach their centre of mass goals. The opposite seems to be the case for the linear collider choice. Both machines depend on advanced technology but in the case of ILC, the lower energy option, the technology will be wrapped up in tens of kilometres of cryostats which would be time consuming to warm up and cool down to adjust or modify hardware.

References

- [1] <http://preprints.cern.ch/yellowrep/2000/2000-008/p1.pdf> “An Introduction to Particle Accelerators” – E. Wilson – OUP <http://ukcatalogue.oup.com/product/9780198508298.do>
- [2] “Engines of Discovery” A. Sessler and E. Wilson - WSP
<http://www.worldscibooks.com/physics/6272.html>
- [3] Alvarez, L.W. (1946). The design of a proton linear accelerator, Phys. Rev. 70 799B.
- [4] Schmüser, P. (1987). Basic course on accelerator optics, Proceedings of the 1986 CERN Accelerator School, CERN 87-10
- [5] Rossbach, J. and Schmüser, P. (1992). Basic course on accelerator optics. Proceedings of the 1986 CERN Accelerator School, Jyväskylä, Finland, CERN 87-1
<http://doc.cern.ch/yellowrep/2005/2005-012/p55.pdf>

POS(LCPS2009)006