25. ACCELERATOR PHYSICS OF COLLIDERS

Revised August 2003 by K. Desler and D.A. Edwards (DESY).

25.1. Introduction

This article is intended to be a mini-introduction to accelerator physics, with emphasis on colliders. Essential data are summarized in the "Tables of Collider Parameters" (Sec. 26). Luminosity is the quantity of most immediate interest for HEP, and so we begin with its definition and a discussion of the various factors involved. Then we talk about some of the underlying beam dynamics. Finally, we comment on present limitations and possible future directions.

The focus is on colliders because they provide the highest c.m. energy, and therefore, the longest potential discovery reach. All present-day colliders are synchrotrons with the exception of the SLAC Linear Collider. In the pursuit of higher c.m. energy with electrons, synchrotron radiation presents a formidable barrier to energy beyond LEP. The LHC will be the first proton collider in which synchrotron radiation has significant design impact.

25.2. Luminosity

The event rate R in a collider is proportional to the interaction cross section σ_{int}, and the factor of proportionality is called the *luminosity*:

$$R = \mathscr{L} \sigma_{\text{int}} \ . \tag{25.1}$$

If two bunches containing n_1 and n_2 particles collide with frequency f, then the luminosity is approximately

$$\mathscr{L} = f \frac{n_1 n_2}{4\pi \sigma_x \sigma_y}, \tag{25.2}$$

where σ_x and σ_y characterize the Gaussian transverse beam profiles in the horizontal (bend) and vertical directions. Though the initial particle distribution at the source may be far from Gaussian, by the time the beam reaches high energy, the normal form is a very good approximation, thanks to the central limit theorem of probability and diminished importance of space charge effects. The qualifier "approximately" appears because this generic expression requires adaptation to particular cases. Discussion may be found in the article of Furman and Zisman in Sec. 3.1.1 of Ref. 1.

Luminosity is often expressed in units of $\text{cm}^{-2}\text{s}^{-1}$, and tends to be a large number. For example, KEK recently announced that its B factory had reached a peak luminosity in excess of 10^{34} cm^{-2} s^{-1}. The highest luminosity for protons achieved so far is 1.3×10^{32} $\text{cm}^{-2}\text{s}^{-1}$ at the now decommissioned ISR; a goal of the Tevatron run just getting underway as this is written is to challenge that record. The relevant quantity for HEP is the luminosity integrated over time, usually stated in the units normally used for cross sections, such as pb^{-1} or fb^{-1}. B-factory integrated luminosities are moving into the hundreds of fb^{-1} range.

The beam size can be expressed in terms of two quantities, one termed the *transverse emittance*, ϵ, and the other, the *amplitude function*, β. The transverse emittance is a beam quality concept reflecting the process of bunch preparation, extending all the way back to the source for hadrons and, in the case of electrons, mostly dependent on synchrotron radiation. The amplitude function is a beam optics quantity, and is determined by the accelerator magnet configuration.

CITATION: S. Eidelman *et al.*, Physics Letters **B592**, 1 (2004)

The transverse emittance is a measure of the phase space area associated with either of the two transverse degrees of freedom, x and y. These coordinates represent the position of a particle with reference to some ideal design trajectory. Think of x as the "horizontal" displacement (in the bend plane for the case of a synchrotron), and y as the "vertical" displacement. The conjugate coordinates are the transverse momenta, which at constant energy are proportional to the angles of particle motion with respect to the design trajectory, x' and y'. Various conventions are in use to characterize the boundary of phase space. Beam sizes are usually given as the standard deviations characterizing Gaussian beam profiles in the two transverse degrees of freedom. In each degree of freedom, the one-σ contour in displacement and angle is frequently used, and we will follow this choice.

Suppose that at some location in the collider, the phase space boundary appears as an upright ellipse, where the coordinates are the displacement x (using the horizontal plane for instance), and the angle x' with respect to the beam axis. The choice of an elliptical contour will be justified under Beam Dynamics below. If σ and σ' are the ellipse semi-axes in the x and x' directions respectively, then the emittance may be defined by $\epsilon \equiv \pi\sigma\sigma'$. Transverse emittance is often stated in units of mm-mrad.

At either a minimum or maximum of the beam size of a beam circulating in equilibrium, the *amplitude function* β at those points is the aspect ratio σ/σ'. When expressed in terms of σ and β, the transverse emittance becomes

$$\epsilon = \pi\sigma^2/\beta \ . \tag{25.3}$$

Of particular significance is the value of the amplitude function at the interaction point, β^*. To achieve high luminosity, one wants β^* to be as small as possible; how small depends on the capability of the hardware to make a near-focus at the interaction point. For example, in the HERA proton ring, β^* at one of the major detectors is 1 m while elsewhere in the synchrotron, typical values of the amplitude function lie in the range 30–100 m. For e^+e^- colliders, $\beta^*_y \sim 1$ cm.

Eq. (25.2) can now be recast in terms of emittances and amplitude functions as

$$\mathscr{L} = f\frac{n_1 n_2}{4\sqrt{\epsilon_x\,\beta^*_x\,\epsilon_y\,\beta^*_y}} \ . \tag{25.4}$$

Thus, to achieve high luminosity, all one has to do is make high population bunches of low emittance collide at high frequency at locations where the beam optics provides as low values of the amplitude functions as possible.

Depending on the particular facility, there are other ways of stating the expression for the luminosity. In a multibunch collider, the various bunch populations will differ; in a facility such as HERA, the electron and proton bunches may differ in emittance, the variation of the beam size in the neighborhood of the interaction point may be significant, and so on.

25.3. Beam dynamics

A major concern of beam dynamics is stability: conservation of adequate beam properties over a sufficiently long time scale. Several time scales are involved, and the approximations used in writing the equations of motion reflect the time scale under consideration. For example, when, in Sec. 25.3.1 below, we write the equations for transverse stability, no terms associated with phase stability or synchrotron radiation appear; the time scale associated with the last two processes is much longer than that demanded by the need for transverse stability.

25.3.1. *Betatron oscillations*:

Present-day high-energy accelerators employ alternating gradient focussing provided by quadrupole magnetic fields [2]. The equations of motion of a particle undergoing oscillations with respect to the design trajectory are

$$x'' + K_x(s)\,x = 0 \ , \quad y'' + K_y(s)\,y = 0 \ , \tag{25.5}$$

with

$$x' \equiv dx/ds \ , \quad y' \equiv dy/ds \tag{25.6}$$

$$K_x \equiv B'/(B\rho) + \rho^{-2} \ , \quad K_y \equiv -B'/(B\rho) \tag{25.7}$$

$$B' \equiv \partial B_y/\partial x \ . \tag{25.8}$$

The independent variable s is path length along the design trajectory. This motion is called a *betatron* oscillation because it was initially studied in the context of that type of accelerator. The functions K_x and K_y reflect the transverse focussing—primarily due to quadrupole fields except for the radius of curvature, ρ, term in K_x for a synchrotron—so each equation of motion resembles that for a harmonic oscillator, but with spring constants that are a function of position. No terms relating to synchrotron oscillations appear, because their time scale is much longer and, in this approximation, play no role.

These equations have the form of Hill's equation, and so the solution in one plane may be written as

$$x(s) = A\sqrt{\beta(s)}\ \cos(\psi(s) + \delta), \tag{25.9}$$

where A and δ are constants of integration and the phase advances according to $d\psi/ds = 1/\beta$. The dimension of A is the square root of length, reflecting the fact that the oscillation amplitude is modulated by the square root of the amplitude function. In addition to describing the envelope of the oscillation, β also plays the role of an 'instantaneous' $\lambda\!\!\!^{-}$. The wavelength of a betatron oscillation may be some tens of meters, and so typically values of the amplitude function are of the order of meters, rather than on the order of the beam size. The beam optics arrangement generally has some periodicity, and the amplitude function is chosen to reflect that periodicity. As noted above, a small value of the amplitude function is desired at the interaction point, and so the focussing optics is tailored in its neighborhood to provide a suitable β^*.

The number of betatron oscillations per turn in a synchrotron is called the *tune* and is given by

$$\nu = \frac{1}{2\pi}\oint \frac{ds}{\beta} \ . \tag{25.10}$$

Expressing the integration constant A in the solution above in terms of $x,\ x'$ yields the *Courant-Snyder invariant*

$$A^2 = \gamma(s)\, x(s)^2 + 2\alpha(s)\, x(s)\, x'(s) + \beta(s)\, x'(s)^2$$

where

$$\alpha \equiv -\beta'/2, \quad \gamma \equiv \frac{1+\alpha^2}{\beta} \quad . \tag{25.11}$$

(The Courant-Snyder parameters α, β, and γ employ three Greek letters which have other meanings, and the significance at hand must often be recognized from context.) Because β is a function of position in the focussing structure, this ellipse changes orientation and aspect ratio from location to location, but the area πA^2 remains the same.

As noted above, the transverse emittance is a measure of the area in x, x' (or y, y') phase space occupied by an ensemble of particles. The definition used in Eq. (25.3) is the area that encloses 39% of a Gaussian beam.

For electron synchrotrons, the equilibrium emittance results from the balance between synchrotron radiation damping and excitation from quantum fluctuations in the radiation rate. The equilibrium is reached in a time which is small compared with the storage time.

For present-day hadron synchrotrons, synchrotron radiation does not play a similar role in determining the transverse emittance. Rather, the emittance during storage reflects the source properties and the abuse suffered by the particles throughout acceleration and storage. Nevertheless, it is useful to argue as follows: Though x' and x can serve as canonically conjugate variables at constant energy, this definition of the emittance would not be an adiabatic invariant when the energy changes during the acceleration cycle. However, $\gamma(v/c)x'$, where here γ is the Lorentz factor, is proportional to the transverse momentum, and so qualifies as a variable conjugate to x. So often one sees a normalized emittance defined according to

$$\epsilon_N = \gamma \frac{v}{c}\, \epsilon, \tag{25.12}$$

which is an approximate adiabatic invariant, *e.g.* during acceleration.

25.3.2. ***Phase stability***: The particles in a circular collider also undergo synchrotron oscillations. This is usually referred to as motion in the *longitudinal* degree-of-freedom because particles arrive at a particular position along the accelerator earlier or later than an ideal reference particle. This circumstance results in a finite bunch length, which is related to an energy spread.

For dynamical variables in longitudinal phase space, let us take ΔE and Δt, where these are the energy and time differences from that of the ideal particle. A positive Δt means a particle is behind the ideal particle. The equation of motion is the same as that for a physical pendulum, and therefore is nonlinear. But for small oscillations, it reduces to a simple harmonic oscillator:

$$\frac{d^2 \Delta t}{dn^2} = -(2\pi\nu_s)^2 \Delta t \tag{25.13}$$

where the independent variable n is the turn number, and ν_s is the number of synchrotron oscillations per turn, analogous to the betatron oscillation tune defined earlier. Implicit in this equation is the approximation that n is a continuous variable. This approximation is valid provided $\nu_s \ll 1$, which is usually well satisfied in practice.

In the high-energy limit, where $v/c \approx 1$,

$$\nu_s = \left[\frac{h\eta\, eV\ \cos\phi_s}{2\pi E}\right]^{1/2} \quad . \tag{25.14}$$

There are four as yet undefined quantities in this expression: the harmonic number h, the slip factor η, the maximum energy eV gain per turn from the acceleration system, and the synchronous phase ϕ_s. The frequency of the RF system is normally a relatively high multiple, h, of the orbit frequency. The slip factor relates the fractional change in the orbit period τ to changes in energy according to

$$\frac{\Delta\tau}{\tau} = \eta\frac{\Delta E}{E} \quad . \tag{25.15}$$

At sufficiently high energy, the slip factor just reflects the relationship between path length and energy, since the speed is a constant; η is positive for all the synchrotrons in the "Tables of Collider Parameters" (Sec. 26).

The synchronous phase is a measure of how far up on the RF wave the average particle must ride in order to maintain constant energy to counteract synchrotron radiation. That is, $\sin\phi_s$ is the ratio of the energy loss per turn to the maximum energy per turn that can be provided by the acceleration system. For hadron colliders built to date, $\sin\phi_s$ is effectively zero. This is not the case for electron storage rings; for example, the electron ring of HERA runs at a synchronous phase of 45°.

Now if one has a synchrotron oscillation with amplitudes $\widehat{\Delta t}$ and $\widehat{\Delta E}$,

$$\Delta t = \widehat{\Delta t}\ \sin(2\pi\nu_s n)\ , \qquad \Delta E = \widehat{\Delta E}\ \cos(2\pi\nu_s n), \tag{25.16}$$

then the amplitudes are related according to

$$\widehat{\Delta E} = \frac{2\pi\nu_s E}{\eta\tau}\widehat{\Delta t} \quad . \tag{25.17}$$

The longitudinal emittance ϵ_ℓ may be defined as the phase space area bounded by particles with amplitudes $\widehat{\Delta t}$ and $\widehat{\Delta E}$. In general, the longitudinal emittance for a given amplitude is found by numerical integration. For $\sin\phi_s = 0$, an analytical expression is:

$$\epsilon_\ell = \left[\frac{2\pi^3 EeVh}{\tau^2\eta}\right]^{1/2} (\widehat{\Delta t})^2. \tag{25.18}$$

Again, a Gaussian is a reasonable representation of the longitudinal profile of a well-behaved beam bunch; if $\sigma_{\Delta t}$ is the standard deviation of the time distribution, then the bunch length can be characterized by

$$\ell = c\,\sigma_{\Delta t} \quad . \tag{25.19}$$

In the electron case, the longitudinal emittance is determined by the synchrotron radiation process just, as in the transverse degrees of freedom. For the hadron case, the history of acceleration plays a role, and because energy and time are conjugate coordinates, the longitudinal emittance is a quasi-invariant.

For HEP, bunch length is a significant quantity, because if the bunch length becomes larger than β^*, the luminosity is adversely affected. This is because β grows parabolically as one proceeds away from the IP, and so the beam size increases, thus lowering the contribution to the luminosity from such locations.

25.3.3. ***Synchrotron radiation*** [3]: A relativistic particle undergoing centripetal acceleration radiates at a rate given by the Larmor formula multiplied by the 4th power of the Lorentz factor:

$$P = \frac{1}{6\pi\epsilon_0}\frac{e^2a^2}{c^3}\gamma^4. \tag{25.20}$$

Here, $a = v^2/\rho$ is the centripetal acceleration of a particle with speed v undergoing deflection with radius of curvature ρ. In a synchrotron that has a constant radius of curvature within bending magnets, the energy lost due to synchrotron radiation per turn is the above multiplied by the time spent in bending magnets, $2\pi\rho/v$. Expressed in familiar units, this result may be written

$$W = 8.85 \times 10^{-5} E^4/\rho \ \text{ MeV per turn} \tag{25.21}$$

for electrons at sufficiently high energy that $v \approx c$. The energy E is in GeV and ρ is in kilometers. The radiation has a broad energy spectrum which falls off rapidly above the *critical energy*, $E_c = (3c/2\rho)\hbar\gamma^3$. Typically, E_c is in the hard x-ray region.

The characteristic time for synchrotron radiation processes is the time during which the energy must be replenished by the acceleration system. If f_0 is the orbit frequency, then the characteristic time is given by

$$\tau_0 = \frac{E}{f_0 W} \quad . \tag{25.22}$$

Oscillations in each of the three degrees of freedom either damp or antidamp depending on the design of the accelerator. For a simple separated-function alternating gradient synchrotron, all three modes damp. The damping time constants are related by Robinson's Theorem [4], which, expressed in terms of τ_0, is

$$\frac{1}{\tau_x} + \frac{1}{\tau_y} + \frac{1}{\tau_s} = 2\frac{1}{\tau_0} \quad . \tag{25.23}$$

Even though all three modes may damp, the emittances do not tend toward zero. Statistical fluctuations in the radiation rate excite synchrotron oscillations and radial betatron oscillations. Thus there is an equilibrium emittance at which the damping and excitation are in balance. The vertical emittance is non-zero due to horizontal-vertical coupling.

Polarization can develop from an initially unpolarized beam as a result of synchrotron radiation. A small fraction $\approx E_c/E$ of the radiated power flips the electron spin. Because the lower energy state is that in which the particle magnetic moment points in the same direction as the magnetic bend field, the transition rate toward this alignment is larger than the rate toward the reverse orientation. An equilibrium polarization of 92% is predicted, and despite a variety of depolarizing processes, polarization above 80% has been observed at a number of facilities.

The radiation rate for protons, is of, course down by a factor of the fourth power of the mass ratio, and is given by

$$W = 7.8 \times 10^{-3} E^4/\rho \;\; \text{keV per turn} \tag{25.24}$$

where E is now in TeV and ρ in km. For the LHC, synchrotron radiation presents a significant load to the cryogenic system, and impacts magnet design due to gas desorption and secondary electron emission from the wall of the cold beam tube. The critical energy for the LHC is 44 eV.

25.3.4. ***Beam-beam tune shift***: In a bunch-bunch collision, the particles of one bunch see the other bunch as a nonlinear lens. Therefore, the focussing properties of the ring are changed in a way that depends on the transverse oscillation amplitude. Hence, there is a spread in the frequency of betatron oscillations.

There is an extensive literature on the subject of how large this tune spread can be. In practice, the limiting value is hard to predict. It is consistently larger for electrons because of the beneficial effects of damping from synchrotron radiation.

In order that contributions to the total tune spread arise only at the detector locations, the beams in a multibunch collider are kept apart elsewhere in the collider by a variety of techniques. For equal energy particles of opposite charge circulating in the same vacuum chamber, electrostatic separators may be used assisted by a crossing angle if appropriate. For particles of equal energy and of the same charge, a crossing angle is needed not only for tune spread reasons, but also to steer the particles into two separate beam pipes. In HERA, because of the large ratio of proton to electron energy, separation can be achieved by bending magnets.

25.3.5. ***Luminosity lifetime***: In electron synchrotrons, the luminosity degrades during the store primarily due to particles leaving the phase stable region in longitudinal phase space, as a result of quantum fluctuations in the radiation rate and bremsstrahlung. For hadron colliders, the luminosity deteriorates due to emittance dilution resulting from a variety of processes. In practice, stores are intentionally terminated when the luminosity drops to the point where a refill will improve the integrated luminosity.

25.4. Status and prospects

Present facilities represent a balance among current technology, the desires of High Energy Physics, and public support. For a half century, beam optics has exploited the invention of alternating gradient focussing. This principle is employed in all colliders both linear and circular. Superconducting technology has grown dramatically in importance during the last two decades. Superconducting magnets are vital to the Tevatron, HERA, and to the future LHC. Superconducting accelerating structures are necessary to CESR, LEP, HERA, Jefferson Laboratory, the spallation neutron source, and other facilities requiring high-gradient long pulse length RF systems. Present room temperature accelerating structures produce very short pulses, but with gradients well in excess of the superconducting variety [1].

At present, the next facilities will include the LHC, and possibly an electron linear collider. The LHC is an approved project that will represent a major step forward in superconducting magnet technology. No linear collider project has been approved as yet, and the conventional and superconducting approaches compete for prominence.

In addition to the possibilities of the preceding paragraph, there are other synchrotron-based collider studies underway. Despite formidable R&D challenges, a muon-muon collider may become feasible. Proponents of a very large hadron collider at higher energy than the cancelled SSC project are exploring low-cost magnets and tunnels for a facility on the 100 TeV c.m. energy scale.

The approach to collider design sketched here—guidance and focussing provided by external magnetic fields, and acceleration produced by RF resonators—has led to ever larger and more costly facilities with increase of c.m. energy. Support for new HEP facilities has diminished as proposals have climbed into the multi-billion dollar range.

There is no shortage of ideas for departure from the current design paradigm. Wakefield accelerators, plasma-laser combinations, and related investigations may, if successful, deliver gradients far higher than any realized today in existing HEP facilities. However, staging and energy efficiency are major hurdles. These approaches are exceedingly challenging technologically, and require a strong R&D program if they are to succeed.

Other important references include Ref. [5–7], which are not cited in the text above.

References:

1. *Handbook of Accelerator Physics and Engineering*, eds. A.W. Chao and M. Tigner, World Scientific Publishing Co. (Singapore, 2nd printing, 2002).
2. E.D. Courant and H.S. Snyder, Ann. Phys. **3**, 1 (1958). This is the classic article on the alternating gradient synchrotron.
3. M. Sands, *The Physics of Electron Storage Rings—An Introduction*, SLAC Publication SLAC 121, UC-28 (ACC), 1970.
4. K. Robinson, Phys. Rev. **111**, 373 (1958).
5. D.A. Edwards and M.J. Syphers, *An Introduction to the Physics of High Energy Accelerators*, John Wiley & Sons, 1993. This is an elementary textbook on accelerator

physics. The next two references are more advanced, and are cited here for readers who may wish to pursue beam physics in greater depth.

6. A. Wu Chao, *Physics of Collective Beam Instabilities in High Energy Accelerators*, John Wiley & Sons, 1993.
7. M. Reiser, *Theory and Design of Charged Particle Beams*, John Wiley & Sons, 1994.