Collider Beam Physics

Frank Zimmermann

Abstract

We review some accelerator physics topics for circular as well as linear colliders, considering both lepton and hadron beams.

*Invited Contribution to Review of Accelerator Science & Technology, Vol. 7 (2014),
Editors Alexander W. Chao and Weiren Chou, World Scientific Pub., Singapore*

*The research leading to these results has received partial funding from the European
Commission under the FP7 Research Infrastructures project EuCARD-2, grant agreement
no.312453.*

Collider Beam Physics

Frank Zimmermann, CERN

Abstract

We review some accelerator physics topics for circular as well as linear colliders, considering both lepton and hadron beams.

Keywords

Colliders, luminosity, beam-beam interaction, beamstrahlung, beam-beam compensation, optics, dynamic aperture, impedance, beam instabilities, luminosity levelling, luminosity tuning, collimation, injection, beam diagnostics, electron cloud, cooling, polarized beams, dust and UFOs

Outline

This report contains 20 sections, roughly organized in the following order:

- 1) collider layouts & types;
- 2) collisions, including integrated and peak luminosity, beam-beam interaction, beamstrahlung, interaction region, long-range collisions. beam-beam compensation, luminosity tuning, and luminosity leveling;
- 3) collective effects and electron clouds;
- 4) dynamic aperture, injection, collimation and beam diagnostics;
- 5) single-beam strategies specific to colliders: cooling schemes and polarization; and
- 6) other features: dust and UFOs.

1. Collider layouts and types

The name “collider” refers to a device which brings into collision two high-energy particle beams moving in opposite direction, possibly at a small nonzero crossing angle. The collision normally happens at an interaction point (IP) located close to the centre of an experimental detector. The main reason for colliding two counter-propagating, usually ultra-relativistic, particle beams is the much higher centre-of-mass energy, E_{cm} , thereby attained, which is increasing linearly with the beam energy (assuming both beams have the same energy),

$$E_{\text{cm}} = 2 E_{\text{beam}}$$

as compared with an increase of the centre-of-mass energy only as the square root of beam energy

$$E_{\text{cm}} = \sqrt{2E_{\text{beam}}m_{\text{target}}c^2}$$

in the case of sending a single beam onto a fixed target (with target particle mass m_{target}). Most colliders constructed so far were electron-positron colliders. Another important branch is hadron colliders, which can be proton-proton, proton-antiproton or heavy-ion colliders. Yet another species are hadron-lepton colliders. For the future also muon and $\gamma\gamma$ colliders have been proposed.

The first hadron collider, the CERN Intersecting Storage Rings (ISR), was an extremely successful machine, from the point of view of accelerator performance. Its peak luminosity of $1.4 \times 10^{32} \text{ cm}^{-2} \text{ s}^{-1}$ was 35 times above the design value, and remained a world record for about two decades. The ISR dc beam current of more than 40 A (single beam current up to 57 A) has not yet been reached by any other collider [1,2]. Many pioneering accelerator studies were performed at the ISR.

The colliding beams can be continuous, as was the case at the ISR, or quasi-continuous as proposed for “superbunch hadron colliders” [3], but in most cases the beams are bunched by a radiofrequency accelerating system.

Almost all colliders so far were circular colliders, in which the same particle bunches are collided again and again over many subsequent revolutions. When colliding protons and antiprotons, or electrons and positrons, at the same energy the two beams can share a common beam pipe as well as the magnetic guiding and focusing system. In the case of colliding equally charged particles, such as protons with protons, two separate beam pipes with opposite magnetic dipole fields are required for guiding the two counter-propagating beams. Two separate beam pipes and magnet systems are also needed in the case of asymmetric e^+e^- factories with different energy for each of the two beams. The almost complete separation of the two beams minimizes the number of parasitic long-range beam collisions, which often limit the dynamic aperture. Namely, the separation avoids most, sometimes all, of the parasitic beam-beam collisions otherwise encountered, and it also allows for an optimized optics tuning and optimization of each of the two rings. The double-ring configuration has been one of the reasons for achieving extremely high luminosities at the asymmetric B factories, PEP-II and KEKB, as well as at the LHC. The advantages are so big that even for future symmetric e^+e^- colliders like the FCC-ee (formerly TLEP) a double-ring scheme is considered. Many-bunch operation in a single ring e^+e^- collider can be achieved with a Pretzel (or helix) scheme based on electrostatic separators, as has been used at LEP-1, at CESR, and in the Tevatron.

In the case of electron-positron colliders, at each revolution, the circulating particles are losing energy in the form of synchrotron radiation. The energy loss per electron per turn, U_0 , increases steeply with beam energy,

$$U_0 = \frac{4\pi r_e}{3(m_e c^2)^3} \frac{E_{\text{beam}}^4}{\rho},$$

where r_e denotes the classical electron radius, m_e the electron mass, c the speed of light, and ρ the bending radius. The strong dependence on the particle mass explains why circular muon colliders have been proposed as a future form of lepton collider and why proton beams can reach much higher collision energies than either electrons/positrons or muons.

To avoid the energy loss (and the associated emittance increase due to “quantum excitation”) from synchrotron radiation, for highest beam energies, e.g. at the TeV scale, one alternative approach for realizing electron-positron collisions is in the form of a linear collider, where each bunch collides with a bunch of the opposing beam a single time. While early proposals of linear colliders considered recovering the energy of the spent beam as an essential ingredient [4,5], in more recent proposals bunches are disposed on a beam dump soon after the collision. This means that to achieve a luminosity comparable to the one provided by a circular collider, where only a small fraction of the beam energy is lost per collision, a linear collider must (and can) have much smaller IP spot sizes. It also implies that a future linear collider must have a much more powerful positron source (e.g. the proposed linear ILC requires about $2 \times 10^{14} e^+/s$ to be compared with $2 \times 10^{12} e^+/s$ in the most extreme parameter sets for the circular SuperKEKB and FCC-ee).

Only one linear collider was in operation so far, namely the Stanford Linear Collider, from about 1989 through 1998. More precisely, the SLC was not completely linear, but accelerated single electron and positron bunches (as well as a second electron bunch used for positron production) in a common single linac, on the same RF pulse, and then deflected these bunches into two separate km-long collider arcs of opposite curvature, to bring them into collision. The SLC positron source – the most powerful source in operation so far – provided about $6 \times 10^{12} e^+/s$. Even so, the luminosity of the pioneering SLC was lower than for probably any circular collider since the time of ADONE (the e^+e^- collider ADONE was built during the second half of the 1960’s, following the first circular e^+e^- colliders AdA in Italy/France; around 1965 also two circular e^+e^- double-ring colliders were in operation: VEP-1 in Russia and CBX in the US). Two future linear e^+e^- colliders are being proposed – the International Linear Collider (ILC), based on superconducting RF powered by klystrons, and the Compact Linear Collider (CLIC), based on warm copper RF cavities powered by decelerating an intense drive beam. A variant of the linear collider is the proposed photon collider, where high-

energy electrons are Compton-back scattered off a high-power laser beam a few mm or a few 100 μm from the focal point so that the resulting Lorentz-boosted gamma quanta collide, instead of the original lepton beams. No linear collider was yet considered for proton beams. For a 100 TeV collider, with accelerating gradients of 100 MV/m as for CLIC, a length of 1000 km would be needed, which is ten times longer than the circumference of the proposed Future Circular Collider (FCC-hh). The latter aims at 100 TeV in the centre of mass using a moderate radiofrequency voltage, well below 100 MV, and SC dipole magnets with a field of 16 T.

2. Peak luminosity

For the head-on collision of short upright Gaussian bunches of population N_b the instantaneous luminosity, L_0 , is given by

$$L_0 = f_{coll} \frac{N_b^2}{4\pi\sigma_x\sigma_y}$$

where f_{coll} is the collision frequency, which, for a circular collider is equal to the product of revolution frequency and number of bunches, $f_{coll}=f_{rev}n_b$, and for a linear collider equal to the product of repetition rate and number of bunches per pulse, $f_{coll}=f_{rep}n_b$.

If the collisions do not occur head on, but with a crossing angle θ_c the luminosity is reduced due to the lower geometric overlap of the colliding bunches. The luminosity also decreases if the bunch length becomes equal to, or larger than, β^* ("hourglass effect"). For the case of electron-positron colliders, with vertical and horizontal rms IP beam sizes fulfilling $\sigma_y \ll \sigma_x$, the luminosity reduction due to both these effects, $R_L \equiv L/L_0$ is expressed as [6]

$$R_L = \sqrt{\frac{2}{\pi}} a e^b K_0(b),$$

where $a \equiv \beta_y^*/(\sqrt{2}\sigma_z)$, and

$$b \equiv a^2 \left[1 + \left(\frac{\sigma_z}{\sigma_x} \tan \frac{\theta_c}{2} \right)^2 \right].$$

In the case of round (e.g. hadron) beams the equivalent expression is [7]

$$R_L = \frac{\cos(\theta_c/2)}{\sqrt{\pi}\sigma_z} \int_{-\infty}^{\infty} \frac{e^{-As^2}}{1 + (s/\beta^*)^2} ds$$

with

$$A \equiv \frac{\sin^2(\theta_c/2)}{\sigma_x^2(1 + (s/\beta^*)^2)} + \frac{\cos^2(\theta_c/2)}{\sigma_z^2}$$

For hadron colliders a common case is $\sigma_z \ll \beta_{x,y}^*$. In this situation one can neglect the hourglass effect; the luminosity loss, only due to the crossing angle, is described by

$$R_\phi = \frac{1}{\sqrt{1 + \phi^2}}$$

where

$$\phi \equiv \frac{\theta_c \sigma_z}{2\sigma_x}$$

is called the Piwinski angle. This geometric luminosity loss is illustrated in Figs. 1 and 2.

Figure 1: Luminosity loss due to crossing angle is related to the effective increase of the transverse beam size, $\sigma \rightarrow \sigma\sqrt{1 + \phi^2}$.

Figure 2: Geometric luminosity loss as a function of the Piwinski angle.

For linear colliders the luminosity is enhanced by the pinch effect, i.e. by the shrinkage of the transverse beam size during the collision under the effect of the strong focusing field of the opposing beam [8]. At the SLC the measured pinch enhancement factor was about 1.4 [9]. For circular colliders the pinch effect can be neglected. However, the additional focusing due to the beam-beam collisions has a (multi-turn) optical effect, referred to as the dynamic beta function, which can also lead to smaller effective beam size at the collision point, and, thereby, to enhanced luminosity.

Especially for hadron colliders, the number of tracks per bunch collision is large, and the detector technology may set limits on the total number of events per crossing (e.g. for calorimetry), as well as on the number of bunch crossings per unit time (or the bunch spacing), on the longitudinal line density of events (for tracking of the primary vertices) and, possibly, also on the number of events per unit time during the collision. These parameters should be optimized in the design, e.g. through a sufficiently long luminous region, as well as through special operational procedures, e.g. luminosity levelling to confine the event pile up.

3. Integrated luminosity

The most important figure of merit of a collider is the useful luminosity it delivers over time, or its integrated luminosity provided to one or several particle-physics detectors.

The integrated annual luminosity can roughly be estimated from the peak luminosity L_{peak} and the total time of the year scheduled for physics operation T_{phys} as

$$L_{int} \approx H L_{peak} T_{phys},$$

where H is an empirical factor ("Hübner factor") which includes the effects of the luminosity decay during periods of data acquisition by the physics experiments, the turnaround time from one physics fill to the next (used for ramp down, injection, acceleration, collimator set up, β^* squeeze), the machine availability, and, for a linear collider, shot-by-shot fluctuations, tuning periods, and limited

tuning precision. For LHC, LEP and SLC, this factor was $H \approx 0.2$, and almost constant from year to year. For PEP-II and KEKB after implementing top-up injection, resulting in almost constant beam current, the Hübner factor increased to $H \approx 0.6 - 0.9$.

For e^+e^- colliders the relevant cross sections decrease with the inverse energy squared, $\propto 1/E_{cm}^2$, which requires the luminosity to increase with the square of the energy. In the case of hadron colliders the situation is more complex, since hadrons are not fundamental particles, but consist of partons (quarks and gluons). At hadron colliders, to produce a new particle at a given energy, it is generally much more efficient to increase the hadron beam energy than to raise the hadron-collider luminosity, since the constituent (parton) luminosities at the energy of interest rapidly increase with higher beam energy. To illustrate this point, and also to elucidate the impact of integrated luminosity on the energy reach of a hadron collider, Fig. 3 compares the reach of a 100-TeV pp collider at two different integrated luminosities (3 and 30 ab^{-1}) with the one of the HL-LHC (3 ab^{-1} at 14 TeV). The different green lines in the picture refer to various types of parton interactions (e.g. valence quark, sea quark, and gluon scattering).

Figure 3: Relative energy reach of a 100-TeV pp collider with an integrated luminosity of 3 ab^{-1} (left) or 30 ab^{-1} (right) compared with the one of the 14-TeV HL-LHC at 3 ab^{-1} . Shown are the parton-parton system masses for an equal number of events, assuming that acceptance and efficiency remain the same. Parton luminosities and associated energy reach were computed using the tool of Ref. [10].

4. Beam-beam interaction

One intrinsic feature of a collider is the collision process. The electric and magnetic fields of the other beam encountered in the collision act like a nonlinear lens, resulting in a transverse force which at small amplitudes (below $1-2\sigma$) increases linearly with the transverse displacement from the centre of the opposing beam and at large amplitudes (say above 3σ) decreases as the inverse of the amplitude. A typical transverse beam-beam deflection is illustrated in Fig. 4.

Figure 4: Normalized single-particle beam-beam deflection angle as a function of normalized displacement from the centre of the opposing beam for the nominal LHC.

For the collision of round Gaussian beam (typical for hadron colliders), of equal charge with horizontal and vertical rms beam size σ , in a “thin-lens approximation” the beam-beam deflection angle is

$$\Delta \vec{r}' = \frac{2r_p N_b}{\gamma} \frac{\vec{r}}{r^2} \left(1 - e^{-\frac{r^2}{2\sigma^2}} \right),$$

where $r \equiv \sqrt{x^2 + y^2}$, $\vec{r} \equiv (x, y)$, $\vec{r}' \equiv (x', y')$, x and y denote the horizontal and vertical distances from the centre of the other beam, while x' and y' are the horizontal and vertical trajectory slopes, γ is the Lorentz factor of the deflected particle, r_p its classical particle radius, and N_b the bunch population of the other beam.

For the case of flat Gaussian beams in (oppositely charged) electron-positron ring colliders, the deflection angle can be expressed analytically in terms of the complex error function w , namely [11]

$$\begin{pmatrix} \Delta x' \\ \Delta y' \end{pmatrix} = -\frac{r_p N_b}{\gamma} \begin{pmatrix} F_x \\ F_y \end{pmatrix},$$

where

$$F_y + iF_x = \sqrt{\frac{2\pi}{\sigma_x^2 - \sigma_y^2}} \left[w \left(\frac{x+iy}{\sqrt{2(\sigma_x^2 - \sigma_y^2)}} \right) - e^{-\frac{x^2}{2\sigma_x^2} - \frac{y^2}{2\sigma_y^2}} w \left(\frac{\frac{\sigma_y}{\sigma_x}x + i\frac{\sigma_x}{\sigma_y}y}{\sqrt{2(\sigma_x^2 - \sigma_y^2)}} \right) \right],$$

and

$$w(z) \equiv e^{-z^2} (1 - \text{erf}(-iz)).$$

At small amplitudes, where the deflection increases approximately linearly with displacement, it resembles the effect of an additional quadrupole (though of equal sign in both planes), whose strength is proportional to the brightness of the opposite beam (where the term “brightness” refers to the ratio of bunch charge and normalized transverse emittance) and characterized by the so-called beam-beam parameter ξ . For a single collision of Gaussian beams without a crossing angle, the latter is defined, in each plane, as

$$\xi_{x,y} \equiv \frac{\beta_{x,y}^* r_p N_b}{2\pi\gamma\sigma_{x,y}(\sigma_x + \sigma_y)}$$

where γ again denotes the relativistic Lorentz factor of the colliding particle, N_b the bunch population of the opposing beam, $\beta_{x,y}^*$ the IP beta function of the particle considered, and, finally, σ_x and σ_y the horizontal and vertical rms beam size of the opposing beam at the collision point. The parameter ξ roughly corresponds to the linear beam-beam (betatron) tune shift experienced by a particle at small amplitude, but, similar to the effect of a quadrupole field error, the exact expression for the real tune shift depends on the working point in tune diagram, as well as on dynamic changes of the beta function (due to the beam-beam focusing). The head-on beam-beam tune shift remains constant if the horizontal and vertical beta functions are varied by the same factor.

The beam-beam interaction increases the tune spread and it also excites betatron resonances of the form

$$mQ_x + nQ_y = p ,$$

with n , m and p integers and $Q_{x,(y)}$ the horizontal (vertical) betatron tune. Overlap of the beam tune spread with resonances up to order 16 has been seen to degrade the performance of hadron colliders. The tune spread, often together with resonances lines, can be visualized in the form of a tune footprint, as is illustrated in Fig 5. The maximum tune shift is reached when relevant resonance lines can no longer be avoided. Figure 6 presents the tune footprint for the nominal LHC, including the effect of parasitic long-range collisions (see later). Based on the experience at earlier hadron colliders the nominal LHC tune footprint had been chosen so as to fit into a square of size $\Delta Q_x \times \Delta Q_y < 0.01 \times 0.01$.

The tune footprint is not a well-defined concept, however. In particular, it depends on the maximum transverse oscillation amplitude considered. For the $SppS$ and early LHC studies, the latter was taken to be 4 sigma, later increased to 6 sigma in order to obtain a better agreement between the footprint criterion and the calculated dynamic aperture. Hence, the footprint criterion should be seen as purely empirical, especially when adding detuning from different sources, e.g. from head-on and long-range collisions.

Figure 5: Schematic tune footprint for head-on collision together with resonance lines.

Figure 6: Nominal LHC beam-beam tune footprint for amplitudes up to 6σ with 4 IPs and nominal bunch intensity $N_b = 1.15 \times 10^{11}$, considering head-on and long-range collisions.

When the relevant tune footprint overlaps with critical resonance lines, poor beam lifetime, emittance growth, or enhanced detector background occur. Such effects typically herald the “beam-beam limit” in hadron colliders. The hadron-collider beam-beam limit may be degraded further by tune or orbit modulation due to magnet power-converter ripple. The experience of the impact of such ripple at the ISR and the $Spp\bar{S}$ [12] provided a strong incentive to reduce the power-converter noise down to record levels of ~ 1 ppm, which represented a considerable effort extending over a decade [13].

In lepton colliders it has been found empirically that increasing the beam current leads to a saturation of the beam-beam tune shift. That is, above a certain beam current the vertical (or horizontal) emittance blows up in such a way that the beam-beam tune shift stays constant, equal to a maximum value. From this (“first”) beam-beam limit onward the luminosity increases only linearly with further raised bunch intensity instead of quadratically. At even higher current a “second” beam-beam limit has been observed, where operation becomes impossible, due to the formation of large beam tails resulting in unacceptable detector background or extremely poor beam lifetime [14].

The beam-beam limit increases (i.e. only lower order resonances remain important) in the case of strong radiation damping, which is especially relevant for circular lepton colliders. LEP observations at different beam energies [15], a simple model [15], and computer simulations [16] (see Fig. 7) all suggest the following dependence on damping time or beam energy

$$\xi_y^{\max}(E) \propto \frac{1}{\tau_s^{0.4}} \propto E^{1.2}$$

Figure 7: Maximum beam-beam tune shift per IP in collisions at 4 IPs as a function of damping decrement – LEP observations, TLEP (FCC-ee) simulations, together with the expected dependence [15,16,17].

Inserting this into the expression for the luminosity yields the following scaling law for the luminosity of a circular collider operating at the beam-beam limit:

$$L_{ring,\xi} \propto \frac{\eta_{ring} P_{wall}}{E^3} \frac{\xi_y}{\beta_y^*} \propto \frac{\eta_{ring} P_{wall}}{E^{1.8}} \frac{1}{\beta_y^*}.$$

Here E is the beam energy, and η_{ring} denotes the efficiency of converting wall-plug power P_{wall} to beam energy (i.e. by compensating for the SR losses), where η_{ring} is defined as $\eta_{ring} \equiv U_0 f_{rev} n_b N_b / P_{wall} = U_0 I_{beam} / (e P_{wall})$, with I_{beam} the beam current and e the electron charge. In the cw operation of a storage ring, the efficiency η_{ring} can be up to 50%.

The above formula shows that in a ring of given size the luminosity decreases as the inverse (1.8)th power of energy. This scaling law ceases to be valid at high energies when, or if, beamstrahlung (synchrotron radiation emitted during the collision; see later), rather than the beam-

beam tune shift, becomes the dominant limitation. In this case the dependence on energy is much stronger,

$$L_{ring,bs} \propto \frac{\eta_{ring} P_{wall}}{E^4} \frac{\sigma_z}{E \sigma_y^*} \propto \frac{\eta_{ring} P_{wall}}{E^4} \frac{1}{\sigma_y^*},$$

assuming $\sigma_z \propto E$.

In circular electron-positron colliders the effect of the beam-beam collision on the optics at small amplitudes (“dynamic beta function” [18]) also changes the dispersion invariant,

$$H_x \equiv (D_x^2 + (\alpha_x D_x + \beta_x D'_x)^2) / \beta_x,$$

around the ring and thereby the horizontal equilibrium emittance (“dynamic emittance”) [19,20]. A clever design and operating strategy profits from these two phenomena to further reduce the spot size at the collision point. The possible equilibria are determined by a self-consistent nonlinear calculation, which, at high intensity, also allows for “flip-flop” states, where one beam blows up and the other stays small [20,21].

The beam-beam collision introduces a large tune spread, which tends to greatly increase the Landau damping [22]. However, under certain circumstances – also depending on the betatron tunes –, coherent beam-beam instabilities can develop [23,24]. Related instabilities – more precisely, a complex interplay between coherent beam-beam interaction, resistive-wall instabilities and the feedback system used to stabilize the latter – have limited the performance of the ISR [13].

For a linear collider the strength of the beam-beam interaction is quantified by the disruption parameter $D_{x,y}$:

$$D_{x,y} \equiv \frac{2r_p N_b \sigma_z}{\gamma \sigma_{x,y} (\sigma_x + \sigma_y)}.$$

For small disruption parameter, the maximum deflection angle of particles during the collision is of order $\theta_{x,y} \cong D_{x,y} \sigma_{x,y} / \sigma_z$.

Comparing the above expressions we can also express the circular-collider beam-beam parameter in terms of the disruption, namely

$$\xi_{x,y} = D_{x,y} \beta_{x,y}^* / (4\pi \sigma_z).$$

The disruption in a linear collider increases the luminosity. When beams of opposite charge collide, the disruption enhances the luminosity. A general approximate expression for this disruption enhancement in the case of “quasi-flat” beams is [25]

$$H_D = H_{D,x}^{1/2} H_{D,y}^{f(R)}$$

with

$$H_{Dx,y} \approx 1 + D_{x,y}^{1/4} \left(\frac{D_{x,y}^3}{1 + D_{x,y}^3} \right) \left\{ \ln \left(\sqrt{D_{x,y}} + 1 \right) + 2 \ln(0.8/A_{x,y}) \right\}$$

and

$$f(R) = \frac{1 + 2R^3}{6R^3}$$

where $A = \sigma_z / \beta_{x,y}^*$.

5. Beamstrahlung

Synchrotron radiation emitted in the strong field of the opposing beam is called “beamstrahlung”.

The typical energy of the beamstrahlung photons is characterized by the parameter Y . This is equal to two thirds the classical critical energy divided by the beam energy E_{beam} [8],

$$Y = \frac{2}{3} \frac{\hbar \omega_c}{E_{beam}} \approx \frac{5}{6} \frac{\gamma r_e^2 N_b}{\alpha \sigma_z (\sigma_x + \sigma_y)},$$

where α denotes the fine-structure constant.

Typical values of Y at the interaction point are $2 \cdot 10^{-3}$ for the SLC, 0.3 for the NLC, and up to 10 for some versions of CLIC. If Y becomes comparable to 1 or larger, a significant portion of beamstrahlung photons convert into real electron-positron pairs in the strong electro-magnetic

fields of the two beams. Unfortunately, linear colliders at multi-TeV energies can hardly avoid operating in this regime.

Besides Y , there is a second parameter of interest, namely the number of beamstrahlung photons emitted per electron, N_γ . For Gaussian colliding beams with low disruption, N_γ is approximately [8]

$$N_\gamma \approx \frac{10\pi}{2} \frac{\alpha\sigma_z}{\gamma\lambda_e} Y \frac{1}{(1 + Y^{2/3})^{1/2}} \approx \frac{2\alpha r_e}{\sigma_x + \sigma_y}$$

For a high-energy e^+e^- storage ring the beamstrahlung has two main effects:

(1) Increased energy spread and bunch lengthening due to the additional synchrotron radiation with large quantum fluctuation [26], which is described by the self-consistent formula [27]

$$\sigma_{\delta,tot} = \left(\frac{1}{2} \sigma_{\delta,SR}^2 + \left(\frac{1}{4} \sigma_{\delta,SR}^4 + \frac{1}{4} \frac{\tau_z n_{IP}}{T_{rev}} \sigma_{\delta,BS}^2 \frac{\sigma_{\delta,SR}^2}{\sigma_{z,SR}^2} \right)^{1/2} \right)^{1/2}$$

with [21]

$$\begin{aligned} \sigma_{\delta,BS} &\approx \delta_{BS} \left(0.333 + \frac{4.583}{N_\gamma} \right)^{\frac{1}{2}} \\ \delta_{BS} &\approx 0.86 \frac{r_e^3 \gamma N_b^2}{\sigma_z \sigma_x^2} \\ N_\gamma &\simeq 2.1 \frac{\alpha r_e N_b}{\sigma_x}. \end{aligned}$$

(2) Reduction of the beam lifetime due to the emission of photons of so large an energy that the emitting electron, or positron, falls outside of the ring momentum acceptance and is lost over subsequent turns. The momentum acceptance δ_{acc} is not only the static acceptance from the RF system (bucket size), but foremost given by the (smaller) off-momentum dynamic aperture. The latter is determined by the off-momentum optical design of the interaction region. One approximate analytical formula for the beamstrahlung lifetime is [28,29]

$$\begin{aligned} \tau_{BS} &\approx \frac{1}{n_{IP} f_{rev}} \frac{4\sqrt{\pi}}{3} \sqrt{\frac{\delta_{acc}}{\alpha r_e}} \\ &\exp \left(\frac{2}{3} \frac{\delta_{acc} \alpha}{r_e \gamma^2} \frac{\gamma \sigma_x \sigma_z}{\sqrt{2} r_e N_b} \right) \frac{\sqrt{2}}{\sqrt{\pi} \sigma_z \gamma^2} \left(\frac{\gamma \sigma_x \sigma_z}{\sqrt{2} r_e N_b} \right)^{3/2}. \end{aligned}$$

For linear colliders the main effect of the (much stronger) beamstrahlung is the degradation of the luminosity spectrum, which is closely related to the number of beamstrahlung photons emitted per electron, N_γ . The fraction of the luminosity at the nominal energy depends exponentially on N_γ as [30]

$$\frac{\Delta L}{L} \approx \frac{1}{N_\gamma^2} (1 - e^{-N_\gamma})^2$$

As a consequence, linear collider designs aim at parameters which ensure $N_\gamma \leq 1$. For higher energy this is ever more difficult to fulfil.

Introducing N_γ the luminosity scaling for a linear collider becomes

$$L \propto \frac{\eta_{linac} P_{wall}}{E} \frac{N_\gamma}{\sigma_y^*},$$

which, at constant wall plug power and vertical IP spot size, decreases inversely with energy. The wall-plug power to beam power conversion efficiency η_{linac} is typically less than 20% (17% for the ILC). The lower efficiency than for storage rings is due to the pulsed mode of operation, i.e. a large part of the RF pulse is dumped onto a load after the beam passage; in addition linear-collider klystrons are operated far from the working point with maximum efficiency, used at LEP, in order to provide margin for feedbacks.

Note that by reducing the bunch length σ_z , one can reach a parameter regime where Y is large and the spot size small, but where one can still ensure that $N_\gamma \leq 1$, thanks to the quantum correction

term — the last fraction in the central expression for N_γ above. This is sometimes referred to as the quantum suppression of beamstrahlung. It arises, roughly speaking, from the fact that the electrons cannot radiate photons of energy higher than the beam energy. In this extreme quantum regime the classical spectrum of synchrotron-radiation photon energies is modified [8]. Exploiting the quantum suppression of photon emission by intentionally operating in a regime of large γ has been suggested as option for a far future high energy linear collider [31,32].

6. Interaction region

The interaction region (IR) where both beams are squeezed and brought into collision is one of the most complicated parts of a collider, together with the matching sections between the IR and the arcs, which need to satisfy a large number of constraints. Either the IR or the matching section may limit the collider performance. The LHC IR is sketched in Fig. 8.

Figure 8: Interaction region of the LHC, including final-focusing quadrupole triplets and pairs of separation dipoles, leading the beams into two separate magnetic arc channels.

The dispersion at the collision point should normally be zero, since a non-zero value increases the IP beam size and also excites synchro-betatron resonances, by introducing a nonlinear coupling between the transverse and longitudinal oscillations. The natural horizontal dispersion from the arcs can be cancelled by a so-called “dispersion suppressor,” an accelerator-optical design element consisting of dipole and quadrupole magnets, at the entrance of the straight section in which the IP is located.

The final focusing quadrupoles introduce a large chromaticity which must be corrected by sextupoles in dispersive regions placed either locally (“compact final focus” [33]), semi-locally (“modular final focus” [34]) or in the collider arcs (for which the ATS scheme is a particularly powerful example [35]). Residual higher-order chromatic aberrations are important, and so is the field quality of the triplet quadrupoles and the separation dipoles, which are located in the regions with the largest beta functions of the collider. Often the multipole errors of the magnets at large β determine the dynamic aperture (i.e. the stable region in phase space) at collision energy as well as – e.g. for the LHC with common final quadrupoles for both beams –, the maximum acceptable crossing angle. In addition to the imperfect magnetic fields, there are unavoidable aberrations, due to magnet edges and kinematic terms which can as well contribute to a limited dynamic aperture, e.g. in the case of SuperKEKB or FCC-ee. The vertical tune shifts scale as [36]

$$\Delta Q_y \approx \frac{l^*}{(2\pi)} \frac{J_y}{\beta_y^{*2}} \quad \text{or} \quad \Delta Q_y \approx \frac{|k|l^{*3}}{(12\pi)} \frac{J_y}{\beta_y^{*2}}$$

for the kinematic and fringe effect respectively, where l^* denotes the free length from the IP, β_y^* the vertical beta function, J_y the vertical action variable, and k the quadrupole strength in units of m^{-2} . For low energy colliders the effect of the detector solenoid also requires careful consideration and dedicated compensation schemes (antisolenoids, or cancellation based on sets of skew quadrupoles).

On top of all these effects, in the IR both beams are present with the possibility of parasitic long-range collisions, which often are a dominant aberration.

7. Long-range collisions

When colliding beams consisting of trains of closely spaced bunches a number of “parasitic” collisions are unavoidably encountered at either side of each main IP.

The total beam-beam tune shift is the sum of the individual beam-beam tune shifts induced at each bunch collision around the ring.

In order to keep the total tune shift low parasitic head-on collisions at places other than the main interaction points inside the particle physics detectors should be avoided. The unwanted parasitic encounters are converted to long-range collisions by colliding the beams at a crossing angle or, sometimes for e^+e^- collisions, by installing electrostatic separators before and after the IP.

Figure 9 illustrates the collision scheme with crossing angle and the long-range forces incurring for trains of closely spaced short bunches in the vicinity of the primary IP. The importance of the long-range collisions depends on the normalized separation and, to a lesser extent, on the bunch charge.

Figure 9: Schematic of long-range collisions experienced by closely spaced short bunches on either side of the primary head-on interaction point. Bunches at the end of a bunch train do not encounter the full number of long-range interactions, and are called PACMAN bunches.

To be effective the crossing angle θ_c should be several times the rms beam divergence at the IP, σ'^* , or $n_\theta \equiv \theta_c/\sigma'^* \gg 1$. On either side of the main IP, through the entrance of the first quadrupole or dipole magnet, the normalized separation of the two colliding beams stays constant and the betatron phase almost does not vary. Therefore, the effects of the various long-range encounters occurring in this region add up linearly. The tune shift generated by n_{LR} long-range collisions around the main IP with head-on collision tune shift ΔQ_{HO} is approximately given by [37,38]

$$\Delta Q_{LR} = \mp 2n_{LR} \frac{\Delta Q_{HO}}{n_\theta^2},$$

where the upper sign refers to the plane of crossing and the lower sign to the orthogonal plane. The crossing angle does not only reduce the geometric overlap of the colliding bunches, but it also reduces the beam-beam tune shift induced by the wanted central collision. For round beams with alternating planes of crossing in two IPs the reduction of the beam-beam tune shift is approximately given by the same factor R_ϕ as for the luminosity [39] (see above).

In the LHC, for example, with 25 ns bunch spacing, there is a parasitic long range collision about every 3.75 m, and 16 such collisions between the IP and the entrance of the separation dipole. With two sides of 4 IPs, this yields a total of about 128 long-range collisions. Half of these are associated with the high-luminosity insertions of ATLAS and CMS, where the nominal long-range separation is only 9.5σ , significantly contributing to tune shift and orbit effects.

Bunches at the head or end of a train experience a smaller number of long-range collisions, and, as a consequence, exhibit a slightly different orbit and tune. These bunches are referred to as

‘PACMAN’ bunches [37,40], in allusion to the classical computer game, since their lifetime will likely be worse than for the nominal bunches. However, the opposite situation may also occur, and the lifetime of these non-nominal bunches might be longer. This latter phenomenon was frequently observed at the Fermilab Tevatron, where it was called the ‘scallop’ effect [41].

In first approximation the dynamic aperture caused by the long-range collisions is equal to the distance from the other beam at the parasitic encounters. In second approximation the dynamic aperture is smaller than this separation, the distance from the latter growing with the square root of the beam intensity [42]. The larger the crossing angle, the larger is the dynamic aperture due to the long-range encounters. In the case of the LHC, the dynamic aperture is also affected by multipole errors of the high gradient quadrupole magnets in the interaction regions, getting larger for larger angles, and, accordingly, there is a finite value of the crossing angle which maximizes the dynamic aperture due to the combined effect of IR multipole errors and long-range collisions [43].

8. Beam-beam compensation

Various proposals and attempts have been made for compensating the harmful effect of head-on or long-range collisions. Such beam-beam compensation schemes include “four beam collisions,” where the forces and fields from two co-propagating e^+ and e^- beams exactly cancel. Such four-beam collisions were tested at the DCI in e^+e^- collisions at 0.8 GeV [44,45] however without yielding any clear performance improvement, due to coherent beam-beam instabilities. They were also contemplated for linear colliders [46,47].

A somewhat similar scheme is the auxiliary collision with an “electron lens” [48], i.e. a pulsed low-energy electron beam with a current of the order of 1 A (or higher), which collides with one of the two main beams. The low-energy electron beam is guided and stabilized by a strong solenoid field. The collision with the electron beam can cancel part of the head-on tune spread (and possibly of the resonance excitation) induced in the same main beam by primary collisions with a positively charged (proton) beam. A first set of electron lenses was implemented and operated at the Tevatron [49]. Instead of compensating head-on beam-beam effects, the Tevatron electron lenses were successfully used as pulsed quadrupoles to equalize the tunes along a bunch train. Electron lenses are presently being installed and tested in RHIC [50] and are also proposed as an additional tool for the HL-LHC.

Other compensation schemes are based on optimizing the phase advance or the planes of crossing between successive interaction points, or on the use of octupoles [51,52] which can add to, or reduce, the linear tune shift with action induced by the beam-beam interaction.

Finally, for the LHC a compensation of long-range collisions based on current-carrying wires was proposed [53]. At a transverse distance r , the wires generate the same transverse force of shape $1/r$, as the field of the opposing beam at the parasitic long-range encounters [53]. In order to correct all non-linear effects the correction must be local. For this reason, there needs to be at least one wire compensator, on one side of each primary interaction point (IP) for either beam, in a region where the two beams are already physically separated, but otherwise as close as possible to the common region where the long-range encounters occur. The proposed LHC compensator is located upstream of the separation dipole D1, on both sides of IP1 and IP5, at a location where the horizontal and vertical beta functions are equal. Figure 10 illustrates how one wire cancels the effect of all 16 long-range encounters occurring on one side of the IP. The betatron phase difference between the long-range (LR) wire compensator and the average LR collision is 2.6° (ideally it should be zero).

Figure 10: Illustration of the LHC long-range beam-beam compensation principle [45].

Figure 11: The first (left) and the second prototype wire compensator (right) installed in the CERN SPS in 2002 and 2004, respectively.

Photographs of wire-compensator prototypes installed in the SPS are shown in Fig. 11. A different set of wire compensators were built and tested at RHIC [54]. Wire compensators, installed outside the vacuum chamber, successfully reduced the effect of parasitic long-range collisions and improved the beam lifetime in the e^+e^- collider DAFNE [52].

The possibility of compensating the long-range beam-beam effects using multipole correctors has been explored in simulations [55].

9. Collision schemes

Quite a number of collision schemes were proposed, and partly realized, to boost collider luminosities (also see [56]).

Some of these aim at maximizing the achievable head-on beam-beam tune shift, e.g. round beam collisions preserving an additional integral of motion [57] implemented at VEPP-2000, crab crossing [58, 59] as used at KEKB, “crab waist collision” [60] demonstrated at DAFNE, and collisions with a negative momentum compaction factor [61] (also tested at DAFNE).

Other schemes reduce the beam-beam tune shift in order to increase the beam current or to lower the emittance, e.g. collisions with a large Piwinski angle [62] – including “superbunch” schemes in hadron colliders [3,39]; “nanobeam schemes” proposed for the former SuperB (together with crab waist collisions [60]) and adopted for SuperKEKB, in which the short length of the beam-beam overlap enables the implementation of an extremely low β_y^* [63]; colliding hadron bunches with longitudinally flat profile [64,65]; or compensating the beam-beam tune shift with electron(-beam) lenses [48] as implemented at the Tevatron [49] and at RHIC [50], and proposed for the HL-LHC.

The main purpose of a third class of collisions schemes is mitigating the effect of long-range collisions: alternating the planes of crossing in different IPs [37] as for the LHC, and dedicated beam-beam wire compensation [53].

A fourth group of approaches was conceived to maximize the geometric overlap of the colliding bunches, e.g. the crab crossing planned at the (HL-)LHC, crab waist collisions [60,66], and longitudinal strong rf focusing [67]. For linear colliders, the “traveling focus” scheme has been proposed, where each longitudinal slice of the bunch has its minimum size at the moment when it encounters the centre of the opposing beam [68] (the traveling focus can be realized with a correlated energy spread plus non-zero chromaticity or by means of RF quadrupoles). A fifth class optimizes the quality of the experimental data for a particle physics experiments, e.g. collisions at much higher beam energy with extremely large crossing angle [69], luminosity levelling schemes such as the novel “crab kissing” concept [70] (see below), or a reduction in the effective centre-of-mass energy spread at the collision point. The latter is achieved by introducing a nonzero vertical dispersion at the IP, such that the product of dispersion times the relative rms energy spread, $D_y^* \sigma_\epsilon$, is larger than the vertical betatron beam size, $\sqrt{\beta_y^* \epsilon_y}$, and such that higher energy particles of one beam preferentially collide with lower-energy particles of the other beam, e.g. with dispersion of opposite sign [71,72]. The relative rms spread in c.m. collision energies W is given by

$$\frac{\sigma_W}{W} = \sqrt{\frac{\epsilon_y}{2 \left(\frac{D_y^{*2}}{\beta_y^*} + \frac{\epsilon_y}{\sigma_\epsilon^2} \right)}}$$

where σ_ϵ the relative rms energy spread of the beam. This scheme, known as “mono-chromatization” [71,72], has been proposed for particle factories performing high-precision measurements of narrow resonance states.

10. Luminosity tuning

It is evident that to achieve maximum performance the two counter-propagating beams should be optimally centred transversely and longitudinally. Transversely this can be achieved by maximizing the luminosity (or zeroing the beam-beam centre-of-mass deflection angle detected by beam-position monitors (BPMs)) when varying the horizontal and vertical separations in small steps using orbit correctors. This is also called a van-der-Meer scan. If the beam-beam separation drifts away from zero quickly the centering can be controlled by an IP-orbit feedback system using data from a luminosity monitor and/or measured deflection angles as input signals. The relative RF phase for the two beams is varied to ensure that their centres collide with each other in the case of a crossing angle, and at the desired place inside the physics detector. The linear optics at the collision points is adjusted and controlled by varying quadrupole strengths or remotely moving sextupole magnets horizontally in order to adjust the waist position of either beam so as to coincide with the collision point. Skew quadrupole magnets or vertical sextupole movements are optimized to minimize the betatron coupling, which is especially important for the collision of flat lepton beams. Other quadrupoles, at upstream locations with nonzero design dispersion, are optimized to zero any spurious dispersion at the IP. Also the next order of IP optics correction is relevant for luminosity performance. Sextupole settings are adjusted to minimize the chromaticity and to maximize the energy bandwidth, and, perhaps, also to correct the second-order dispersion at the IP. Skew sextupoles at location with nonzero dispersion can control the chromatic coupling, the cancellation of which was of critical importance for KEKB. For the purpose of the aforementioned IP optics corrections, sets of magnet strengths or of magnet positions (if magnets are mounted on remotely controlled movers) are varied together so as to modify only a single aberration at the IP. These combinations of excitation elements are known as orthogonal tuning knobs. If the orthogonality is not perfect the various tuning steps may need to be alternately iterated. If no action is taken, in the course of time the optics of an accelerator degrades, for example due to thermal effects

changing the mechanical position of the accelerator beam pipe or affecting the BPM readings (which will lead to a new actual orbit if a feedback system attempts to correct the perceived changes), magnet-current drifts, slow RF phase variation, or due to ground motion. The degraded optics leads to sub-optimal performance. Therefore, most of the optics corrections have to be redone typically at least once per year, or once per day (for the ISR), or in the case of lepton facilities sometimes every 8-hour shift (KEKB) or even every few hours (SLC). The tuning of optical parameters, especially those at the collision point, is necessary in order to obtain either the maximum beam-beam tune shift in a storage ring [whose performance is also extremely sensitive to the working point in tune diagram] or the smallest spot sizes (and thereby highest luminosity) for a linear collider. Figure 12 sketches the process of iterative tuning with intermittent spot size degradation. One important exception is the LHC, whose optics has proven remarkably stable and reproducible, even from one year to the next, without requiring any further correction in the following year. During the LHC commissioning, a segment-by-segment local IR optics correction has been applied together with global corrections, and IP parameter scans are found not to be necessary [73].

Figure 12: Schematic of tuning effect and spot-size increase between tunings. (This conceptual picture goes back to J. Irwin.)

Luminosity imbalance between different detectors is another important consideration for circular colliders operating with more than one experiment (LEP, LHC). In 2012, the LHC luminosity imbalance between the ATLAS and CMS experiments was reduced to less than 3% thanks to the applied optics corrections [73]. Overall, the LHC optics control has set an impressive new record for hadron colliders. While in circular colliders, which have a well defined and easily measurable emittance (including possible blow up from the beam-beam interaction) and steady-state optics, the collision point spot size often agrees with the expectation, e.g. at the LHC, KEBB or LEP2 (though, it is reported, not at the ISR [13], with extremely high beam current and a working point close to the integer resonance), this has not always been the case in single-pulse linear machines. For example, at low current, $N_b \sim 5 \times 10^9$, the 1994 SLC final focus routinely delivered the expected vertical rms spot size of 420 nm [74], whereas at high current ($N_b \sim 3.5 \times 10^{10}$ particles per bunch) the vertical spot size increased to values of 600-900 nm [74]. The much larger beam size was attributed, in parts, to deteriorating upstream emittances, to reduced orbit stability, to final-focus wake fields, and to insufficient tuning accuracy.

Figure 13: Incremental IP corrections of vertical waist, vertical dispersion and skew coupling in the North and South final focus during the 1996 SLC run. Shown dotted is the average resolution of an aberration scan as quoted by the SLC control system [75].

Figure 13 illustrates the frequency and magnitude of various standard IP optics corrections performed in regular few-hour intervals at the SLC from May to July 1996: vertical waist location, vertical IP dispersion, and horizontal-vertical coupling [75].

The KEKB circular collider also required regular tuning, e.g. scans and corrections of the primary aberrations during every 8-hour shift, in order to operate at the maximum beam-beam tune shift and, hence, maximum luminosity. As an example, Fig. 14 illustrates the efficiency of IP coupling correction at KEKB. The local coupling at the IP was measured with nearby turn-by-turn BPMs. The correction was accomplished through vertical orbit changes at various sextupole magnets. Figure 15 shows the global progress in vertical beam size tuning at the linear-collider final-focus test facility ATF2. After a few years of tuning (only running for a fraction of each calendar year, and slowed down by an earthquake) and after several hardware improvements, world record small beam sizes below 50 nm were obtained, not far from the design value of 37 nm, albeit so far only at low bunch charge [77].

Figure 14: Phase-space correlations before and after optics correction at KEKB [76]. The units of y and p_y should be mm and mrad, respectively.

Figure 16 illustrates the recovery of a <60 nm beam size within 32 hour from the start of operation after a few-week shutdown in 2014.

Figure 17 shows a measurement of the beam size degradation as the ATF-2 bunch intensity is increased towards the design value. Namely, a large decrease in the measured modulation depth of the laser-interferometer (“IPBSM”) is observed as a function of bunch charge. In first approximation, the modulation depth M is related to the beam size σ_y as [78]

$$\sigma_y \approx \frac{d}{2\pi} \sqrt{2 \ln \left(\frac{|\cos \theta|}{M} \right)},$$

where θ denotes the full crossing angle between the two laser beams, and $d = \lambda_L / (2 \sin (\theta/2))$ the spacing of the fringes.

Figure 15: Tuning of vertical beam-size at the ATF-2 as a function of time over 5 years. The reduction of the vertical beam size at the end of 2012 was achieved after lowering the bunch charge by an order of magnitude, from a few 10^9 to a few 10^8 electrons per bunch. The design rms beam size at nominal intensity is 37 nm [77].

Figure 16: Vertical rms beam size versus time in hours counting from the start of operation after 3 weeks of shutdown during some days in April 2014. The different colours refer to different crossing angles (or ranges of beam-size sensitivity) of the laser interferometer beam-size monitor used for tuning [77].

Figure 17: ATF-2 IPBSM modulation as a function of bunch population, indicating a significant beam-size increase for higher bunch changes. The two measurements were done with a laser-beam crossing angle θ of 174 degrees (left) and 30 degrees (right) [77].

The measured intensity dependence of the ATF2 vertical beam size has been fitted as

$$\sigma_y^2(q) = \sigma_y^2(0) + w^2 q^2$$

with $w \sim 100$ nm/nC [77], which means that a beam size of 50 nm at 0.1 nC increases to 110 nm at 1 nC. Interestingly, the SLC observations from two decades earlier [74] would be described by an extremely similar value of w , despite more than an order of magnitude difference in beam energy, bunch charge and nominal beam size.

11. Luminosity levelling

Limiting the event pile up (the number of collisions per bunch crossing) while maximizing the integrated luminosity is a paramount design consideration for the LHC and for the next generation of energy-frontier proton colliders. The LHC experiments were designed to cope with about 20 events per crossing, but in 2012 they have managed to handle up to twice that many; see Fig. 18.

Figure 18: Recorded luminosity versus the mean number of interactions per crossing at the CMS detector during the 2012 LHC run.

The LHC luminosity upgrade, HL-LHC, will be designed for high “virtual” peak luminosity, with the underlying concept that the actual instantaneous luminosity will be reduced intentionally to a level acceptable for the experiments.

The “effective beam lifetime,” τ_{eff} , is defined, and computed, by the following equations,

$$\frac{dN_{\text{tot}}}{dt} = -\frac{N_{\text{tot}}}{\tau_{\text{eff}}} = -\sigma_{\text{tot}} n_{\text{IP}} L_{\text{lev}},$$

yielding

$$\tau_{\text{eff}} = \frac{N_{\text{tot}}}{n_{\text{IP}} \sigma_{\text{tot}} L_{\text{lev}}},$$

where n_{IP} denotes the total number of high-luminosity interaction points (IPs), with, e.g. $n_{\text{IP}}=2$ for the LHC, and σ_{tot} the total cross section. For a given levelled luminosity, L_{lev} , the effective beam lifetime scales with the total beam intensity, N_{tot} . It is evident that, to obtain a decent beam lifetime at a given target luminosity, a minimum beam intensity will be required.

The general luminosity formula is

$$L = \frac{f_{\text{rev}} n_b N_b^2}{4\pi\beta^* \varepsilon} F(\phi_{\text{piv}}, \Delta x, \dots),$$

where F denotes the geometric reduction factor from crossing angle, beam-beam offset and hourglass effect.

For the luminosity with levelling we can write

$$L_{\text{lev}} = f_{\text{lev}}(t) L_{\text{max}}(t)$$

where f_{lev} designates a time-dependent levelling factor, $f_{\text{lev}} \leq 1$, which characterizes the amount of “levelling detuning” with respect to the unlevelled maximum luminosity that would be possible at this moment in time.

The “virtual peak luminosity” is defined as

$$\hat{L} \equiv L_{\text{max}}(0) = \frac{f_{\text{rev}} n_b N_b^2(0)}{4\pi\beta^*(0)\varepsilon} F(\phi_{\text{piv}, \text{min}}(0)) = \frac{L_{\text{lev}}}{f_{\text{lev}}(0)}$$

i.e. as the levelled luminosity divided by the initial value of the levelling detuning factor.

Various levelling schemes are already being used at the LHC. These, plus new ones, are also being considered for the HL-LHC:

(1) Varying the beam-beam offset Δx (successfully applied during LHC operation in 2010), which gives rise to

$$L_{\text{lev}} = \hat{L} \exp\left(-\left(\frac{\Delta x}{2\sigma^*}\right)^2\right)$$

Figure 19 shows the normal luminosity decay in ATLAS and CMS together with a levelled constant luminosity in the LHCb experiment. The levelling was accomplished by a feedback system which varies and controls the beam-beam separation at LHCb.

Figure 19: Luminosity decay in ATLAS and CMS and levelled luminosity at LHCb – typical example from LHC run 1 (Courtesy C. Gaspar).

(2) Varying the Piwinski angle ϕ_{piw} , that is the bunch length σ_z , the (effective) crossing angle θ_c , or the crab voltage V_{crab} , and thereby changing the luminosity as

$$L_{lev} \approx \hat{L} \frac{1}{\sqrt{1 + \phi_{piw}^2}}.$$

The effective Piwinski angle can be changed, for example, by adjusting the crab-cavity voltage: Figure 20 sketches the “crab crossing” configuration for luminosity leveling in the crossing plane.

Figure 20: “Crab-crossing” configuration in the crossing plane [70].

(3) Varying the IP beta function β^* e.g. at constant ϕ_{piw} , leading to (for round beams):

$$L_{lev} \approx \hat{L} \frac{\hat{\beta}^*}{\beta_{lev}^*}$$

(4) “Crab kissing” – a newly proposed scheme [70] based on additional crab cavities deflecting the two beams in the non-crossing plane in anti-phase, as is illustrated in Fig. 21. This particular scheme has the further advantage of not reducing the longitudinal extent of the luminous region during the levelling process, a feature much appreciated by the particle-physics detectors.

For a given levelled luminosity, the maximum levelling time in units of τ_{eff} is a function of the virtual peak luminosity according to

$$\frac{t_{lev}}{\tau_{eff}} = 1 - \sqrt{\frac{L_{lev}}{\hat{L}_{peak}}}$$

Figure 21: Illustration of “crab kissing” configuration for luminosity levelling in the plane orthogonal to the crossing plane [70].

Figure 22: HL-LHC luminosity evolution as a function of time without (red), and with levelling at a pile up of 140 for a bunch spacing of 25 ns (solid blue) and 50 ns (dashed blue), assuming the optimum levelling and decay times for a turnaround of $t_{ta} \sim 5$ h [79].

Figure 22 compares the planned luminosity levelling at the HL-LHC, with two different bunch spacings, and the luminosity decay without levelling [79]. Without levelling the ideal average luminosity is

$$L_{ave} = \hat{L} \frac{\tau_{eff}}{(\tau_{eff}^{1/2} + T_{ta}^{1/2})^2}$$

where T_{ta} denotes the turnaround time from the end of physics data taking to the start of the next data taking, including the times needed for ramp down, injection, acceleration, squeeze and

collimation set up ($T_{ta} \sim 2\text{-}3$ hours in case of the LHC). With levelling the time-averaged luminosity is [79]

$$L_{ave} = L_{lev} \frac{\left(1 - \sqrt{\frac{L_{lev}}{\hat{L}}}\right) \tau_{eff} + \frac{t_{dec} \tau_{eff}}{t_{dec} + \tau_{eff}}}{t_{dec} + \left(1 - \sqrt{\frac{L_{lev}}{\hat{L}}}\right) \tau_{eff} + T_{ta}}$$

where t_{dec} denotes the optimum decay time after the end of the levelling ($k \equiv \hat{L}/L_{lev}$):

$$t_{dec} = \tau_{eff} \frac{1}{2 - \frac{1}{\sqrt{k}}} \left(-1 + \frac{1}{\sqrt{k}} + \sqrt{\left(1 - \frac{1}{\sqrt{k}}\right)^2 + \left(2 - \frac{1}{\sqrt{k}}\right) \frac{t_{ta}}{\tau_{eff}}} \right)$$

For the example of Fig. 22 the levelling time t_{lev} amounts to 7.8 h or 17.2 h, and the optimum decay time t_{dec} to 3.7 h and 2.9 h, for a bunch spacing of 25 ns and 50 ns, respectively.

12. Collective instabilities

The interaction of the beam with its environment through self-generated electromagnetic fields can result in single-bunch and multi-bunch instabilities. The strength of this interaction can be described using the concept of impedance, which is the Fourier transform of the wake field.

Often a broadband resonator impedance is chosen as a helpful approximation to represent the sum of many small individual (geometric) impedance contributions, due to changes in beam-pipe cross section, or to generate an ad-hoc model which matches some beam observations. The broadband resonator impedance has the form

$$Z_{\parallel} = \frac{R_{\parallel}}{1 + iQ \left(\frac{\omega}{\omega_r} - \frac{\omega_r}{\omega} \right)}, \text{ or}$$

$$Z_{\perp} = \frac{R_{\perp}}{1 + iQ \left(\frac{\omega}{\omega_r} - \frac{\omega_r}{\omega} \right)}$$

for “monopole” and dipole modes in the transverse and longitudinal plane respectively, where $R_{\parallel}$ denotes the longitudinal resonator impedance in units of Ω , and $R_{\perp}$ the transverse impedance in units of Ω/m (which would be related to a longitudinal dipole mode impedance $R_{\parallel,1}$ [in units of Ω/m^2] through $R_{\perp} = R_{\parallel,1} c/\omega$), while Q is the quality factor, often – somewhat arbitrarily – chosen equal to 1.

The broad-band impedance is mainly relevant for single-bunch effects, while narrow resonator impedances, with a high quality factor Q , for which the beam-induced wake field does not much decay between successive bunch passages, can be responsible for multi-bunch instabilities as well as for local beam-induced heating of components.

Among primary sources of impedance in a collider figure the resistive wall of the vacuum chamber, the higher-order modes in the radiofrequency accelerating cavities, and the resistive or geometric impedance of collimators. For the arc vacuum chamber in superconducting hadron storage rings, like the LHC or the FCC-hh, the anomalous skin effect and magneto-resistance may need to be taken into account when computing the resistive wall impedance [80].

For non-metallic collimators, such as those in the LHC, where the skin depth is not small compared with the collimator gap, the classical formula of the transverse thick-wall resistive-wall impedance per unit length L ,

$$\frac{Z_{\perp}}{L} = \frac{c}{\omega} \frac{1 - \text{sgn}(\omega)i}{\pi b^3 \delta_{\text{skin}} \sigma}$$

with b the chamber radius, σ the conductivity and δ_{skin} the skin depth at frequency ω , needs to be modified. Under certain approximations (applicable for the LHC collimators) it can be expressed as [81,82]

$$\frac{Z_{\perp}}{L} = \frac{2}{cb^2} \frac{1 - \text{sgn}(\omega)i}{i + \text{sgn}(\omega)(1 + b/\delta_{\text{skin}})}.$$

Transverse coherent tune shifts and rise times of multi-bunch head-tail modes driven by the accelerator impedance, including the effect of chromaticity Q' , are determined by a formula going back to Sacherer [83],

$$\Delta Q_{m,l,Q'} = -i \frac{n_b N_b e \beta}{4\pi T_0 E} Z_{\text{eff}}(m, l, Q'),$$

where β denotes the beta function at the location of the impedance, and Z_{eff} the effective impedance for the l th multi bunch mode, and the m th head-tail mode (e.g. $m=0$ rigid bunch motion, $m=1$ one node along the bunch etc.), defined as

$$Z_{\text{eff}}(m, l, Q') \equiv \frac{\sum_p Z_{\perp}(\omega_{\beta} + (Mp+l)\omega_0 + m\omega_s) h_m(\omega_{\beta} + (Mp+l)\omega_0 + m\omega_s - \omega_{\xi})}{\sum_p h_m(\omega_{\beta} + (Mp+l)\omega_0 + m\omega_s - \omega_{\xi})}$$

with ω_0 the angular revolution frequency,

$$\omega_{\xi} \equiv \frac{Q' \omega_0}{\eta}$$

the chromatic frequency shift, and $\eta = \alpha_c - 1/\gamma^2$ the slippage factor. For a Gaussian model, the weight functions are

$$h_m(\omega) = \left(\frac{\omega \sigma_z}{c} \right)^{|2m|} e^{-\omega^2 \sigma_z^2 / c^2}.$$

The latter satisfy the following relation:

$$\sum_p h_m(\omega_{\beta} + (Mp+l)\omega_0 + m\omega_s - \omega_{\xi}) \approx \frac{(2m-1)!!}{2^m}.$$

Once the complex coherent tune shifts $\Delta Q_{x,y} \equiv \Delta Q_{(x,y)m,l,Q'}$ are known, stability or instability of the beam follows from the dispersion relation (here formulated for the special case of rigid dipole-mode oscillations, $m=0$) [84]

$$-\frac{1}{\Delta Q_{x,y}} = \iint_0^{\infty} \frac{J_{x,y} df / dJ_{x,y}}{Q - q_i(J_x, J_y)} dJ_x dJ_y,$$

where f is the distribution function in action space, $J_{x,y}$ are the two linear action variables, and $q_i(J_x, J_y)$ describes the detuning with amplitude, which can be due to dedicated Landau-damping octupole magnets or due to the beam-beam interaction, or due to the combined effect. For simplicity, here we have neglected the longitudinal dimension. The values of coherent tune shifts, $\Delta Q_{x,y}$, traced for different real values of Q (i.e. for actual Q values which correspond to neither growth nor decay) define the stability border in the complex tune-shift plane. This border has to be compared with the complex tune shifts calculated from the effective impedance using the Sacherer equation above.

Concerning Landau damping, in a collider the beam-beam collision normally is the largest source of nonlinearity and the largest source of betatron tune spread. Beams tend to be extremely resistant to impedance-driven classical single-beam instabilities while they are subjected to head-on collisions. The design of the collider together with the desired beam parameters establishes an impedance budget. The impedance of each beam-pipe component should be carefully examined in order to assure it stays within the allocated budget. Since the days of the CERN Intersecting Storage Rings (ISR) in the late 1960's, it is customary to construct an impedance model [85,86]. To construct this model, each beam-pipe component installed, or to be installed, in an accelerator is analysed in one or several ways: (1) by numerical computation using standard software codes, e.g., HFSS,

MAFIA, CST Microwave Studio, GdfidL, ECHO3D, Omega3P, (2) by laboratory “wire” measurements, (3) analytically, or (4) via beam experiments [85,86]. Normally at least two of these four possible methods are applied to the same component, for the purpose of validation. Alternatively, once the accelerator has been built, one can use beam measurements to fit a few key parameters in a simplified impedance description, e.g., by a broadband resonator model with $Q \approx 1$ as for the CERN SPS [87]. The approach is illustrated in Fig. 23.

Figure 23: Multi-pronged approach to constructing and validating the accelerator impedance (Courtesy E. Metral).

Two other collective effects to be considered in the collider design and operation are (1) intrabeam scattering, leading to emittance growth, and (2) Touschek scattering, which can limit the beam lifetime, especially if the momentum acceptance is limited, e.g. by the off-momentum dynamic aperture.

13. Vacuum and electron cloud

Hadron and lepton colliders require beam pipes with an ultrahigh vacuum to provide an acceptable beam lifetime and good background conditions in the experimental detectors. Advanced pumping schemes developed for storage-ring colliders include various types of “non-evaporable getter” (NEG) as well as “cryo-pumping,” where a (slightly) warmer beam screen with pumping slots is surrounded by a cold mass of 1.9 K, and/or cryosorber material at 4.5 K. In the vacuum chambers of (positively charged) proton and positron beams, such as in the LHC or the e^+ damping rings of proposed linear colliders, an electron cloud can build up due to photoemission from synchrotron radiation and/or due to a beam-induced multipactoring process. Namely, seed electrons created by ionization of the residual gas at injection, or photo-electrons liberated from the wall by synchrotron radiation, are accelerated in the electric field of the positively charged bunches. When these electrons then hit the opposite wall, they generate secondary electrons which can in turn be accelerated by the field of the following bunch if they are slow enough to remain in the vacuum chamber until the next bunch arrives. This mechanism is illustrated in Fig. 24 for the LHC, with a bunch spacing of 25 ns.

Figure 24: Illustration of the beam induced multipacting process leading to the build-up of an electron cloud in the LHC beam pipe. The horizontal axis is time, indicating the bunch spacing. Blue curvy lines represent synchrotron radiation photons hitting the chamber wall at the moment of a bunch arrival. (Courtesy F. Ruggiero).

Generally speaking, the shorter the spacing the more easily an electron cloud is produced. Only for very short spacing (the value depending on the beam pipe aperture and bunch charge) when each electron interacts with several passing bunches the build-up is suppressed: in the limit of a coasting beam electrons move in a static electric potential and can no longer gain energy from the beam.

If present, the electron cloud has potential implications for beam stability, emittance growth, and, for a superconducting machine like the LHC, heat load on the cold beam screen in the arcs. Electron cloud effects have been actively investigated at CERN since 1997, by analytic estimates, simulations, and experimental tests [88-90]. Mitigation measures against electron cloud include solenoid windings with a field of 30-50 G in otherwise field-free regions [91], and coating of the beam pipe with a material of low secondary emission yield, e.g. TiN [92], NEG, or amorphous carbon [93]. The electron cloud can ultimately cure itself since the secondary emission yield of the chamber surface decreases as a result of the continuing bombardment with electrons of sufficiently high energy (e.g. above about 30 eV). The variation of the maximum secondary emission yield, $\delta_{\max}$, with electron dose measured in the laboratory for the colaminated copper surface of the LHC beam screen (at perpendicular incidence of the primary electrons) is illustrated in Fig. 25.

Figure 25: Variation of the secondary emission yield of copper with the incident electron dose for different energies of the impinging electrons at normal incidence on colaminated Cu of the LHC beam screen. The squares represent the $\delta_{\max}$ after an additional dose of $10^{-2} \text{ C mm}^{-2}$ at 200 eV [94] (Courtesy R. Cimino).

The LHC strategy to overcome electron-cloud related limitations relies on this process of surface conditioning (intentionally enhanced in dedicated “scrubbing runs”).

Electron-cloud effects and a reduction of the secondary emission yield with time have indeed been observed in the first years of LHC operation, where operation with 50-ns spacing has become possible thanks to rapid surface improvement (lowering of the secondary emission yield) in scrubbing' runs with 25-ns beams [95,96,97]. Figure 26 presents the LHC surface conditioning observed for the cold arcs in 2011. The final value of $\delta_{\max}$ achieved by the scrubbing, using 25-ns beam, was well below the threshold for multipacting with 50-ns beam, which allowed electron-cloud free operation of the LHC at 50-ns spacing during all of 2011 and 2012.

Figure 26: LHC surface conditioning in the cold arcs during 2011 – beam intensity (top) and estimated evolution of the maximum secondary emission yield (bottom) as inferred by benchmarking of simulations with cryogenic heat-load measurements (Courtesy G. Iadarola and G. Rumolo).

Figure 27: SPS “doublet” beam, with alternately 20-ns and 5-ns bunch spacing, may be used for “scrubbing” the LHC, in order to allow for cloud-free operation at 25 ns bunch spacing (Courtesy: G. Iadarola and G. Rumolo).

In order to safely operate the LHC at 25 ns bunch spacing and to push the value of $\delta_{\max}$ below the threshold for 25-ns spacing a more “powerful” scrubbing beam is needed, Figure 27 shows one such beam, consisting of more bunches with alternating 20 and 5 ns spacing, which can be produced in the LHC injector complex.

14. Dynamic aperture

The dynamic aperture of a storage ring refers to the stable region in phase space, within which beam particles survive over relevant time scales, e.g. for a few radiation damping times (ms) in the case of $e^{\pm}$ storage rings, or for the duration of a physics run (hours, days) in the case of hadron colliders.

In collision the dynamic aperture is often limited by the optics in the interaction region with large beta function and significant local chromaticities as well as by head-on and long-range beam-beam effects. For superconducting (SC) hadron storage rings, the dynamic aperture at injection energy has also been a great concern, since the block structure of their coils unavoidably gives rise to higher-order field harmonics on top of the nominal field, and since, in addition, large time-dependent

persistent current field errors are induced in the SC dipole and quadrupole magnets during an acceleration cycle. A larger magnet aperture leads to lower field errors at the expense of increased magnet costs.

The actual LHC dynamic aperture at injection was measured with the help of kicked beams [98]. The measured and simulated dynamic aperture values agree well. It was also shown that through a minimization of readily observed nonlinear parameters it is possible to increase the LHC dynamic aperture [98].

The following scaling law for the turn dependence of the dynamic aperture well parameterizes the results of numerical simulations and is motivated by fundamental theorems for dynamical systems [99]:

$$D(N) = D_{\infty} \left(1 + \frac{b}{(\log N)^{\kappa}} \right).$$

Here D_{∞} represents the asymptotic value of the amplitude of the stability domain, while b and κ are additional parameters, which characterize the variation of the simulated dynamic aperture $D(N)$ with the number of turns, N , considered. The above scaling law for the dynamic aperture can be translated into a scaling law for the intensity evolution in hadron storage rings [100]. Experimental data from the SPS and from the Tevatron are in remarkable agreement with the predicted beam-current evolution [100], providing a means for non-invasive dynamic-aperture measurements [101,102].

15. Injection

In collider complexes, techniques of stacking, accumulation or damping of charged particle beams are often employed, for producing bunches of sufficiently high intensity and low emittance. Examples are the antiproton accumulation with stochastic cooling for $p\bar{p}$ colliders; the heavy-ion accumulation with electron cooling, e.g. for the heavy-ion programme at the LHC; the stacking of proton beams from an H linac by charge-exchange injection into an accumulation ring through a striping foil (where the change of the particle charge overcomes the Liouville theorem on the conservation of phase-space density); as well the positron and electron damping rings for linear colliders, where radiation damping reduces the beam emittance.

For a storage ring collider, magnet cycles for injection and acceleration as well as varying beam currents (leading to variation in the beam pipe temperature) degrade the efficiency of the IP tuning. Latest lepton ring colliders, like PEP-II, KEKB, SuperKEKB as well as the proposed FCC-ee and CepC employ top-up injection, where the collider-ring magnets do not need to cycle, and the beam currents are held almost constant, being continually replenished from a full-energy injector. This “top-up injection” requires an additional high-energy, fast cycling or pulsed, injector(s), as well as an appropriate injection scheme. Longitudinal injection appears particularly well suited for top-up, with regard to minimizing injection-induced detector background (with zero or low dispersion across the IR) and in view of the faster radiation damping in the longitudinal plane.

16. Collimation

Historically collimators have been used to minimize the particle-physics detector background, which is caused by particle losses around the interaction region. In hadron colliders, dedicated collimators, located elsewhere around the ring, can remove particles in the transverse beam tails, which otherwise could be lost near the IR. In lepton colliders synchrotron radiation from the arcs can cause detector background, which is minimized by sophisticated masking schemes, permitting only multiply reflected photons to hit critical portions of the detector. By contrast, for the LHC, the primary functions of its more than 100 collimators is the minimization of beam losses in the cold superconducting arcs so as to avoid magnet quenches [103], and machine protection by protecting critical components (such as the final quadrupoles) against sudden beam impact in case of a critical failure, e.g. the accidental pre-firing of an extraction kicker. Since the LHC primary and secondary

collimators are the first objects impacted by an unstable beam they need to be extremely robust, even at the expense of higher impedance. The present collimator jaws are made from carbon-fibre reinforced carbon. For the HL-LHC even stronger collimator materials are being developed and explored. Alternative, more advanced collimation schemes, considered for the longer-term future, include the use of short bent crystals to impart larger deflection angles on tail particles (together with much-reduced energy deposition from an incident proton beam), and the use of hollow electron-beam lenses, generating a strong nonlinear electro-magnetic field on their outside to deflect protons at large amplitude.

17. Beam diagnostics

The development of colliders has been tightly linked to the progress of beam instrumentation and accelerator control systems. Non-destructive diagnostics are one specific feature of colliders. Circular lepton colliders and the LHC are profiting from synchrotron-light monitors. Beam-position monitors with turn-by-turn readout allow for fast optics verifications. The discovery of the “Schottky noise” as a diagnostics for hadron beams [104,105], and associated measurements of beam-transfer functions (revealing the beam distribution) have been real breakthroughs [106,107]. An example measurement is shown in Fig. 28. At the ISR, “the beam stability diagram, formerly confined to theoretical consideration, could be displayed online,” with provisions to smooth it by appropriate control of higher-order chromatic terms generated by pole-face winding [108]. The Schottky noise of a coasting beam introduces a tiny beam energy loss due to the resistive wall impedance, which could also be measured in the ISR [109]. Simultaneous injection of protons and alpha particles were used as well at the ISR, to test special relativity [108].

Figure 28: Example of a vertical beam transfer function measured at the ISR – amplitude (top) and phase response (bottom). The natural Schottky spectrum is shown for comparison (top right) [106].

18. Cooling schemes

While the radiation damping times of high energy circular e^+e^- colliders often are of the order of milliseconds, this is not yet the case for hadron colliders. The longitudinal (amplitude) damping time of the LHC is about 1 day, the damping time of the FCC-hh 100-TeV (c.m.) 100-km collider around 1 hour.

In the past, certain beams, especially antiproton beams and ion beams, were accumulated and conditioned, in the respective injector complex, by stochastic cooling and electron cooling. The beams of linear colliders are cooled with the help of synchrotron radiation in low-energy damping rings. In circular lepton colliders the horizontal equilibrium emittance is a function of the optics alone (including the optical effects of the beam-beam collision), but it can be modified in either direction with the help of wiggler magnets in dispersive or non-dispersive regions.

The radiation damping is known to lead to higher permissible beam-beam tune shifts, and it counteracts, or recovers, any unwanted emittance dilution. For this reason various cooling techniques have been proposed for hadron colliders. The only effective cooling method so far successfully implemented for colliding high-energy hadron bunches is the 3-D stochastic cooling of heavy-ion beams at RHIC [110,111]. Figure 29 compares the luminosity evolution during a RHIC store for various configurations of the stochastic cooling. With full 3D cooling, the luminosity increases 5-fold over the first 1-2 hours of the store thanks to the cooling of all three beam emittances.

Figure 29: Effect of different cooling configurations on the luminosity in 2012 uranium-uranium collisions at RHIC. No cooling is shown in red, longitudinal only in brown, longitudinal and vertical in green, and all 3 systems in blue. The initial luminosity for these stores varies slightly due to other improvements in RHIC and the injectors [111].

Other proposed schemes for future hadron facilities are optical stochastic cooling [112], coherent electron cooling based on the FEL principle [113], and coherent electron cooling based on the microbunching instability [114]. On the other hand, in future hadron colliders, too strong a radiation damping increases not only the luminosity, but also the event pile up, and the beam-beam tune shift. It is expected that the FCC-hh will require a continuous controlled blow up by transverse and longitudinal noise excitation to preserve desired bunch dimensions during a physics run [96].

19. Polarized beams

Some colliders operate with polarized beams. In electron-positron ring colliders the beams can be self-polarized due to the Sokolov-Ternov effect [115]. In flat rings without solenoids the polarization is perpendicular to the machine plane (i.e. normally vertical) and it reaches a maximum value of 92.4%. Vertical bends, field errors, misalignments etc. lead to depolarization and to equilibrium polarization values lower than the maximum [116].

The spin motion in the particle rest frame is governed by the Thomas-BMT equation, which in the absence of electrical fields takes the form

$$\frac{d\vec{S}}{dt} = \vec{\Omega} \times \vec{S} = -\frac{e}{\gamma m} \left[(1 + G\gamma) \vec{B}_\perp + (1 + G) \vec{B}_\parallel \right] \times \vec{S}.$$

This equation indicates how, with the help of “spin rotators,” consisting either of strong solenoids, or, especially at high energies, of sets of vertical and horizontal dipoles (which sometimes

are mechanically moved together with changes in the field excitation so as to follow the varying beam orbit), the polarization can be arranged to be longitudinal at the IP. In (lepton) storage rings, a nonzero beam polarization also allows for precise energy determination using a technique called resonant depolarization. The spin depolarization in a ring is especially strong at spin orbit resonances defined by

$$Q_{spin} = n + mQ_x + oQ_y + pQ_s$$

with n , m , o and p denoting integers, $Q_{spin} = G\gamma$ the spin tune, and G (or a) the anomalous magnetic moment (for protons $G=1.7928474$, for electrons $G=0.00116$). Integer spin resonances of the form $Q_{spin} = n$ are particularly harmful and should be avoided (as well as their synchrotron sidebands) by an appropriate choice of beam energy. For electrons the primary integer spin resonances are spaced 440.6 MeV apart in beam energy. Most other resonances can be suppressed by “strong synchro-beta spin matching” – a technique involving deterministic changes in the closed orbit and in quadrupole strengths –, which achieved 70% longitudinal polarization at 27.5 GeV in HERA, and 57% vertical polarization at 46 GeV in LEP. At 60 GeV LEP still had a polarization of 7%. See Fig. 30. For even higher energies the energy spread, increasing with the square of the beam energy, became so large that strong overlap with synchro-spin resonances could no longer be avoided and the polarization dropped to zero. According to the LEP experience, the polarization disappears completely for an rms beam energy spread which is larger than about 1/10th of the distance between linear spin resonances (i.e. larger than 1/10th of 440 MeV).

Figure 30: Measured maximum polarization in LEP for different beam energies. The data is compared to the higher-order (solid curve) and the linear (dashed curve) prediction, assuming equal residual imperfections after correction [117].

For hadron colliders polarized beams must be generated at the source and accelerated across many spin resonances. For a single resonance crossing the initial and final polarization levels are related by the Froissart-Stora formula

$$P_f = P_i \left(2e^{-\frac{\pi|\varepsilon|^2}{\alpha}} - 1 \right),$$

where ε denotes the strength of the resonance and α the speed of the resonance crossing. Based on this formula techniques for preserving polarization levels when crossing a resonance include fast tune jumps (for crossing so-called imperfection resonances, caused by magnetic field errors and magnet misalignments) and excitation of a large oscillation using an “AC dipole” [118] (for strong intrinsic resonances, excited by the horizontal magnetic fields which provide the vertical focusing; inducing a full spin flip). However, the most effective method for crossing multiple resonances, e.g. during acceleration, is the use of Siberian snakes [119]. The snake may consist of a group of dipole magnets with alternating horizontal and vertical dipole field such that it rotates the spin vector by 180°. One or a group of snakes can be arranged so that the spin tune equals ½. This breaks the coherent build-up of perturbations on the spin vector and avoids all spin resonances. By means of

Siberian snakes RHIC has achieved average proton beam polarization levels of 58% at 100 GeV and 52% at 255 GeV [120].

Also linear colliders aim at operation with polarized beams from a polarized source. Polarization levels of 80-90% are attainable from strained GaAs photocathode materials in DC guns [121]. Polarized RF guns are also under development. In linear colliders a fraction of the polarization is lost during the collision due to two different processes. The first process is spin precession in the magnetic field of the opposing beam. The corresponding depolarization formula is [8]

$$\Delta P \approx 0.006(N_\gamma/U_0)^2$$

The second process is spin-flip radiation. Its magnitude can be estimated by [8]

$$\Delta P = 2N_\gamma U_{f0},$$

where the function $U_{f0} = U_{f0}(Y)$ is depicted in [8]. Note that the depolarization for both processes mainly depends on N_γ . For either process the more relevant luminosity-weighted effective depolarization is [122]

$$[\Delta P] \approx 0.27 \Delta P$$

For example, at CLIC about 6.2% effective polarization is lost in the collision, where spin precession and spin flip radiation contribute by similar amounts (i.e. causing a loss of about 3% each) [123].

20. Dust or “UFOs”

Unidentified falling objects (UFOs) or dust particles interacting with the beam have affected the performance of several colliders. A prominent example is HERA, whose electron storage ring suffered from repeated dust events, which led to sustained long periods of short beam lifetime [124]. At HERA dust particles were also seen, with the help of local beam-loss monitors, to move around the ring [125]. Calculations and benchmarking with observations indicated that these particles most likely consisted of silica and were of about micron size [124]. The most probable source of the dust particles were distributed ion pumps located inside the dipole and quadrupole magnets. The dust problem was “solved” first by switching operation to positron beams, and finally by replacing the distributed ion pumps by (voltage-free) NEG pumps. The electron rings of KEKB and PEP-II also suffered from occasional dust events, which were primarily causing background spikes in the detectors. Once a dust particle falls into the electron beam it is rapidly ionized, attracted to the transverse centre of the beam, and there heated by the passing beam particles. In HERA the beam density was so low that the temperature of the silica did not reach the melting point. It is thought that in KEKB and PEP-II any dust particles trapped by the beam quickly melted and then exploded [126]. “UFOs” are a major concern for the LHC [126]. Differently from all previous colliders suffering from dust, in the LHC such events are seen with a positively charged beam. The dust, once falling into the beam, is rapidly ionized and repelled within a few revolution periods. However, the local beam loss experienced during so short a time interval is sometimes sufficient to trigger the LHC beam abort system, which is designed to protect against local quenches of SC magnets. At the LHC the generation of the UFOs could be related to the presence of an electron cloud, as both show conditioning behaviour and both occur more strongly, or more frequently, at the 25-ns bunch spacing (see Fig. 31) [127].

- [15] R. Assmann, K. Cornelis, "The Beam-Beam Interaction in the Presence of Strong Radiation Damping," Proc. EPAC 2000 Vienna (2000) p. 1187
- [16] S. White and N. Mounet, "Beam-Beam Studies for TLEP (and Update on TMCI)," 6th TLEP workshop, CERN (2013).
- [17] J. Wenninger, Lepton Collider Overview," Future Circular Collider Kick-Off Meeting, Geneva, 12-15 February 2014
- [18] M. A. Furman, "Beam-beam tune shift and dynamical beta function in PEP-II," LBL-34957, EPAC 94, London (1994).
- [19] K. Hirata and F. Ruggiero, "Treatment of Radiation in Electron Storage Rings," LEP Note 611 (1988).
- [20] A.V. Otboyev and E.A. Perevedentsev, "Self-Consistent Beta Functions and Emittances of Round Colliding Beams," PRST-AB 2, 104401 (1999).
- [21] F. Zimmermann, "Beam sizes in collision and flip-flop states at KEKB," in Proc. Workshop on Beam-beam effects in circular colliders. Batavia, USA, June 25-27, 2001 (2001).
- [22] X. Buffat et al., "Head-On and Long Range Beam-Beam Interactions in the LHC: Effective Tune Spread and Beam Stability Due to Landau Damping," Proc. IPAC 2013 Shanghai
- [23] N. Dikansky, D. Pestrikov, "Nonlinear Coherent Beam-Beam Oscillations in the Rigid Bunch Model," Proc. EPAC 1990 Nice (1990) 1642-1643
- [24] Y.I. Alexahin, "On the Landau Damping and Decoherence of Transverse Dipole Oscillations in Colliding Beams," Particle Accelerators, Volume 59 (1999) p. 43. Y. Alexahin et al., "Coherent Beam-Beam Effects in the LHC," Proc. HEACC2001 Tsukuba (2001).
- [25] P. Chen, "Disruption Effects from the Collision of Quasi-Flat Beams," IEEE PAC93 Washington,
- [26] K. Yokoya, "Quantum Correction to Beamstrahlung due to the Finite Number of Photons," Nucl. Instrum. Meth. A251 (1986) 1.
- [27] K. Ohmi, F. Zimmermann, "FCC-ee/CepC Beam-Beam Simulations with Beamstrahlung", Proc. IPAC'14 Dresden.
- [28] V. Telnov, "Restriction on the Energy and Luminosity of e^+e^- Storage Rings due to Beamstrahlung," Phys. Rev. Letters 110 (2013) 114801
- [29] A. Bogomyagkov, E. Levichev, D. Shatilov, "Beam-beam effects investigation and parameters optimization for a circular e^+e^- collider at very high energies," Phys. Rev. ST – Accel. Beams 17, 041004 (2014).
- [30] P. Chen, "Differential Luminosity under Multi-Photon Beamstrahlung," Physical Review D46, 1186 (1992).
- [31] T. Himel, J. Siegrist, "Quantum Effects in Linear Colliding Scaling Laws," AIP Conf.Proc. 130 (1985) 602-612
- [32] M. Xie, "Quantum Suppression of Beamstrahlung for Future Linear Colliders," Proc. EPAC98 Stockholm (1998).
- [33] P. Raimondi, A. Seryi, "Novel Final Focus Design for Future Linear Colliders," Phys. Rev. Lett. 86, 3779 (2001).
- [34] K. Oide, "Design of Optics for the Final Focus Test Beam at SLAC," PAC89 Chicago (1989) 1319
- [35] S. Fartoukh, "Achromatic telescopic squeezing scheme and application to the LHC and its luminosity upgrade," Phys. Rev. ST Accel. Beams 16, 111002 (2013)
- [36] K. Ohmi, H. Koiso, "Dynamic Aperture Limit caused by IR Nonlinearity in Extremely Low-beta B Factories," Proc. IPAC'10 Kyoto (2010) p. 1548-1550
- [37] D. Neuffer, S. Peggs, "Beam-Beam Tune Shifts and Spreads in the SSC: Head-on, Long Range and PACMAN Conditions," SSC-63 (1986).
- [38] F. Zimmermann, "Beam Dynamics Challenges for Future Circular Colliders," Proc. EPAC'04 Lucerne (2004), p. 83
- [39] F. Ruggiero, F. Zimmermann, "Luminosity Optimization Near the Beam-Beam Limit by Increasing Bunch Length or Crossing Angle," PRSTAB 5, 061001 (2002).

- [40] W. Herr, "Tune Shifts and Spreads due to Short and Long Range Beam-Beam Interactions in the LHC," CERN-SL/90-06-AP, CERN-LHC-NOTE-119 (1990).
- [41] V. Shiltsev, private communication (2002).
- [42] J. Irwin, "Diffusive Losses from SSC Particle Bunches due to Long-Range Beam-Beam Interactions," SSC-233 (1989).
- [43] W. Chou, D. Ritson, "Dynamic aperture studies during collisions in the LHC," PAC97 Vancouver (1997) p.113-115
- [44] J. E. Augustin et al., Vol.2, Proc. 7th Int. Conf. HEA (1969) 113
- [45] Orsay Storage Ring Group, PAC 79, San Francisco
- [46] V.E. Balakin and N.A. Solyak, Proc. XIIIth Int. Conf. on High Energy Accelerators, Novosibirsk (1986) 151.
- [47] D. Whittum, R. Siemann, Proc. PAC 97 Vancouver
- [48] E. Tsyganov et al., SSCL-519, (1993); Phys. Part. Nuclei 27 (1996) 279
- [49] W. Shiltsev et al., NJP 10 (2008) 043042
- [50] D.M. Gassner et al., "Status of the RHIC Head-on Beam-beam Compensation Project," Proc. PAC 2011 New York (2011)
- [51] S. Myers, "Increase of the Beam-beam Limit by Compensation of the Beam-beam Tune Dependence on Amplitude," CERN-LEP-Note-400 (1982). A.B. Temnykh, Proc. 3rd Adv. ICFA Beam Dynamics Workshop, Novosibirsk (1989).
- [52] C. Milardi et al. "DAFNE Lifetime Optimization with Compensating Wires and Octupoles," Proc. CARE-HHH-APD IR'07 Workshop, Frascati, 7-9 November 2007, CERN-2008-006, arXiv:0803.1544
- [53] J.-P. Koutchouk, "Correction of the Long-Range Beam-Beam Effect in LHC using Electromagnetic Lenses," Proc. PAC 2001 Chicago (2001) p. 1681.
- [54] H.J. Kim et al., "Long-Range Beam-Beam Compensation in RHIC," Proc. IPAC2010 Kyoto (2010)
- [55] J. Shi, L. Jin, O. Kheawpum, "Multipole compensation of long-range beam-beam interactions," Proc. 29th ICFA Advanced Beam Dynamics Workshop on Beam Halo Dynamics, Diagnostics & Collimation (HALO 03), AIP Conf.Proc. 693 (2004) 265-268.
- [56] M. Zobov, F. Zimmermann, "Collision Schemes for Ring Colliders," in Handbook of Accelerator Physics and Engineering: 2nd Edition (edited by A.W. Chao, K.H. Mess, M. Tigner, F. Zimmermann), World Scientific, Singapore (2012)
- [57] V.V. Danilov et al., EPAC 96, 1149
- [58] R.B. Palmer, SLAC-PUB-4707, 1988
- [59] K. Oide, K. Yokoya, Phys. Rev. A40 (1989) 315
- [60] P. Raimondi, in 2nd SuperB Workshop (2006); physics/0702033 (2007)
- [61] M. Zobov et al., physics/0607036 (2006)
- [62] K. Hirata, Phys. Rev. Lett. 74 (1995) 2228
- [63] Y. Ohnishi et al., "Accelerator Design at SuperKEKB," PTEP 2013 (2013) 03A011
- [64] F. Ruggiero et al., KEK Proc. 2002-30, p. 131;
- [65] C. Bhat, F. Zimmermann, IPAC 11, 1879
- [66] Y. Ohnishi, ICFA BDN 48 (2009) 252
- [67] A. Gallo et al., physics/0404020 (2004)
- [68] V. Balakin, Travelling Focus, LC-91, 1991
- [69] P. Raimondi, DAFNE upgrade workshop Alghero Sardinia (2003)
- [70] S. Fartoukh, "Pile up management at the High Luminosity LHC and introduction to the crab-kissing concept," CERN-ACC-2014-0076, submitted to Phys. Rev. ST – Accel. Beams (2014).
- [71] A. Renieri, Frascati preprint INF-75/6(R), 1975
- [72] J.M. Jowett, "Feasibility of a Monochromator Scheme in LEP," LEP Note 544 (1985); also J.M. Jowett et al., "The Tau-Charm Factory," Proc. HEACC92 Hamburg (1992) 439-441.
- [73] R. Tomás et al., "Record low β beating in the LHC," Phys. Rev. ST Accel. Beams 15, 091001 (2012).

- [74] F. Zimmermann et al., "Performance of the 1994 / 1995 SLC Final Focus system," Proc. PAC95 Dallas (1995) 656-658
- [75] P. Emma, L.J. Hendrickson, P. Raimondi, F. Zimmermann, "Limitations of interaction point spot size tuning at the SLC," Proc. PAC'97 Vancouver (1997) 452-454
- [76] K. Ohmi et al., "Optics Measurement at the Interaction Point using Nearby Position Monitors in KEKB," Proc. IPAC'14 Kyoto, p. 1539.
- [77] K. Kubo, "Towards International Linear Collider: Experiments at ATF2," Proc. IPAC14 Dresden (2014); also G. White et al., "Experimental validation of a novel compact focusing scheme for future energy-frontier linear lepton colliders," Phys. Rev. Lett. 112, 034802 (2014).
- [78] P. Tenenbaum, T. Shintake, "Measurement of Small Electron-Beam Spots," Annu. Rev. Nucl. Part. Sci. 1999. 49, (1999) p. 125
- [79] O. Brüning, F. Zimmermann, "Parameter Space for the LHC High-Luminosity Upgrade," Proc. IPAC'12 New Orleans (2012).
- [80] W. Chou, F. Ruggiero, "Anomalous Skin Effect and Resistive Wall Heating," CERN LHC-Project-Note-2 (1995).
- [81] B. Zotter, "New Results on the Impedance of Resistive Metal Walls of Finite Thickness," CERN-AB-2005-043.
- [82] A. Koschik, "Simulation of Transverse Multi-Bunch Instabilities of Proton Beams in LHC," PhD Thesis TU Graz, CERN-THESIS-2004-059
- [83] F. Sacherer, "Methods for Computing Bunched Beam Instabilities," CERN Report CERN/SI-BR/72-5 (1972).
- [84] J.S. Berg, F. Ruggiero, "Landau damping with Two-Dimensional Betatron Tune Spread," CERN-SL-AP-96-071-AP (1996).
- [85] B. Spataro et al., "RF Coupling Impedance Measurements versus Simulations," Proceedings of 1st CARE-HHH-APD workshop "HHH-2004", CERN-2005-006 (2005).
- [86] L. Palumbo, "Impedances," Francesco Ruggiero Memorial Symposium," CERN, 3 October 2007.
- [87] D. Boussard, J. Gareyte, "Measurements of the SPS Coupling Impedance," CERN Improvement Report no. 181 (1980).
- [88] F. Zimmermann "A Simulation Study of Electron-Cloud Instability and Beam-Induced Multipacting in the LHC," CERN-LHC-Project-Report-95} (1997).
- [89] O. Gröbner, "Beam Induced Multipacting," PAC 97, Vancouver (1997). J.S. Berg, "Energy Gain in an Electron Cloud during the Passage of a Bunch," CERN-LHC-Project-Note-97 (1997). G. Stupakov, "Photoelectrons and Multipacting in the LHC: Electron Cloud Build-up," CERN-LHC-Project-Report-141 (1997). O. Brüning, "Simulations for the Beam-Induced Electron Cloud in the LHC beam screen with Magnetic Field and Image Charges," CERN-LHC-Project-Report-158 (1997). F. Caspers et al., "Multipacting tests with a Resonant Coaxial Setup," CERN-LHC-Project-Note-110} (1997). M. Furman, "The Electron Cloud Effect in the Arcs of the LHC," CERN-LHC-Project-Report-180 (1998). V. Baglin et al., "Beam-Induced Electron Cloud in the LHC and Possible Remedies," EPAC 1998, Stockholm. F. Zimmermann, "Electron Cloud Simulations: An Update," Proc. Chamonix 2001, CERN-SL-2001-003 DI (2001). See also web page on Electron Cloud in the LHC at <http://wwwslap.cern.ch/collective/electron-cloud>
- [90] F. Zimmermann, "Electron Clouds – Operational Limitations and Simulations," Proc. Chamonix 2003, CERN-AB-2003-008-ADM (2003).
- [91] H. Fukuma et al., "Status of Solenoid System to Suppress the Electron Cloud Effects at the KEKB," AIP Conference Proceedings 642 (2002) p. 357
- [92] K. Kennedy et al., "TiN Coating of the PEP-II Low-Energy Ring Aluminum Arc Vacuum Chambers," PAC97 Vancouver (1997)
- [93] Y. Vallgren et al., "Amorphous carbon coatings for the mitigation of electron cloud in the CERN Super Proton Synchrotron," Phys. Rev. ST Accel. Beam 071001 (2011)
- [94] R. Cimino et al., "Nature of the Decrease of the Secondary-Emission Yield by Electron Bombardment and Its Energy Dependence," Phys. Rev. Lett. 109 (2012) 0604801.

- [95] O. Dominguez et al., "First Electron-Cloud Studies at the Large Hadron Collider," Phys. Rev. ST Accel. Beams 16, 011003 (2013).
- [96] O. Dominguez, "Electron Cloud Studies for the LHC and Future Proton Colliders," PhD thesis EPFL Lausanne (2013).
- [97] G. Rumolo, G. Iadarola, "Analysis of the Electron Cloud Observations with 25 ns Bunch Spacing at the LHC," Proc. IPAC'14 Dresden (2014) p. 1410.
- [98] E. H. Maclean et al., "Measurement of Nonlinear Observables in the Large Hadron Collider using Kicked Beams," Phys. Rev. ST Accel. Beams 17, 081002 (2014).
- [99] M. Giovannozzi, W. Scandale, and E. Todesco, "Prediction of Long Term Stability in Large Hadron Colliders," Part. Accel. 56 (1997) 195-225.
- [100] M. Giovannozzi, "Proposed Scaling Law for Intensity Evolution in Hadron Storage Rings Based on Dynamic Aperture Variation with Time," Phys. Rev. ST Accel. Beams 15, 024001 (2012).
- [101] M. Giovannozzi et al., "First Experimental Observations from the LHC Dynamic Aperture Experiment," Proc. IPAC12 New Orleans, p. 1362.
- [102] M. Giovannozzi et al., "Experimental Observations from the LHC Dynamic Aperture Machine Development Study in 2012," Proc. IPAC13, p. 2606.
- [103] R. Bruce et al., "Simulations and Measurements of Beam Loss Patterns at the CERN Large Hadron Collider," Phys. Rev. ST Accel. Beams 17, 081004 (2014).
- [104] L. Thorndahl, W. Schnell, "Q Measurements from Vertical Schottky Scans," ISR Performance Report, 9 November 1973.
- [105] J. Borer et al., "Non-Destructive Diagnostics for Coasting Beams with Schottky Noise," Proc. IXth HEACC, SLAC (1974).
- [106] J. Borer et al., "Runs 981 and 982, 28.9.78 and 12.10.78, The use of the beam transfer function obtained from noise excitation to measure the influence of the transverse wall impedance, active feedback and resonances," ISR Performance Note 1978
- [107] A. Hofmann, J.-P. Koutchouk "Run 1142, 18.7.80, The beam-beam transfer function," ISR Performance Note 1981.
- [108] J.-P. Koutchouk, private communication (2014).
- [109] A. Hofmann, T. Risselada, "Measuring the ISR Impedance at Very High Frequencies by Observing the Energy Loss of a Coasting Beam," Proc. PAC83, IEEE Trans.Nucl.Sci. 30 (1983) 2400-2402.
- [110] M. Blaskiewicz, J.M. Brennan, K. Mernick, "Three-Dimensional Stochastic Cooling in the Relativistic Heavy Ion Collider," Phys.Rev.Lett. 105 (2010) 094801
- [111] J.M. Brennan, M. Blaskiewicz, K. Mernick, "Stochastic Cooling in RHIC," Proc. IPAC2012 New Orleans (2012).
- [112] A.A. Mikhailichenko, M.S. Zolotarev, "Optical Stochastic Cooling," Phys.Rev.Lett. 71 (1993) 4146-4149
- [113] V.L. Litvinenko, Y.S. Derbenev, "Coherent Electron Cooling," Phys. Rev. Lett. 102 (2009) 114801
- [114] D. Ratner, "Microbunched Electron Cooling for High-Energy Hadron Beams," Phys. Rev. Lett. 111 (2013) 8, 084802
- [115] A.A. Sokolov and I.M. Ternov, Sov. Phys. Dokl. 8 (1964) 1203.
- [116] D.P. Barber, Summary on electron/positron polarization in rings," Snowmass 2001 Working Group M5.
- [117] R. Assmann et al., Observation of Radiative Spin Polarization at 60.6 GeV," Proc. PAC 1999 New York
- [118] M. Bai et al., "Overcoming Intrinsic Spin Resonances with an rf Dipole," Physical Review Letters 80 (1998) 4673-4676.
- [119] Ya.S. Derbenev and A.M. Kondratenko, "Polarization Kinematics of Particles in Storage Rings," Sov. Phys. JETP 37 (1973) 968-973; Zh.Eksp.Teor.Fiz.64 (1973) 1918-1929
- [120] A. Anders et al., "RHIC Polarized Proton Operation in Run 12," Proc. IPAC 2012 New Orleans (2012).

- [121] H. Tang, et al., "The SLAC Polarized Electron Source," International Workshop on Polarized Beams and Polarized Gas Targets, Cologne, Germany, SLAC-PUB-6918 (1995); T. Nakanishi, et al., "Highly Polarized Electron Source Development in Japan," APAC'98 Tsukuba (1998).
- [122] K. Yokoya and P. Chen, "Depolarization due to Beam-Beam Interaction in Electron-Positron Linear Colliders," 8th International Symposium on High Energy Spin Physics, Minneapolis (1988).
- [123] R.W. Assmann, F. Zimmermann, "Polarization issues at CLIC," APS / DPF / DPB Summer Study on the Future of Particle Physics, Snowmass, 30 June - 21 July 2001.
- [124] F. Zimmermann, "Trapped Dust in HERA and DORIS," DESY-HERA-93-08 (1993).
- [125] D.R.C. Kelly, "Dust in Accelerator Vacuum systems," Proc. PAC97 Vancouver (1997)
- [126] F. Zimmermann et al., "Trapped Macroparticles in Electron Storage Rings," Proc. PAC1995 Dallas (1995)
- [127] T. Baer et al., "UFOs in the LHC: Observations, Studies and Extrapolations," IPAC2012 New Orleans (2012), p. 3936; update at CMAC #7, 14 March 2013.