11 Cooling

Many applications of particle accelerators require beam cooling, which refers to a reduction of the beam phase space volume or an increase in the beam density via dissipative forces. In electron and positron storage rings cooling naturally occurs due to synchrotron radiation, and special synchrotron-radiation damping rings for the production of low-emittance beams are an integral part of electron-positron linear colliders. For other types of particles different cooling techniques are available. Electron cooling and stochastic cooling of hadron beams are used to accumulate beams of rare particles (such as antiprotons), to combat emittance growth (e.g., due to scattering on an internal target), or to produce beams of high quality for certain experiments. Laser cooling is employed to cool ion beams down to extremely small temperatures. Here the laser is used to induce transitions between the ion electronic states and the cooling exploits the Dopper frequency shift. Electron beams of unprecedentedly small emittance may be obtained by a different type of laser cooling, where the laser beam acts like a wiggler magnet. Finally, designs of a future muon collider rely on the principle of ionization cooling. Reference [1] gives a brief review of the principal ideas and the history of beam cooling in storage rings; a theoretical dicussion and a few practical examples can be found in [2].

11.1 Damping Rates and Fokker–Planck Equation

In the presence of cooling and in the absence of any excitation by noise, the evolution of a beam distribution function $f(x, x', t)$ for one degree of freedom, here for the horizontal plane, is described by the differential equation [2]

$$\frac{\mathrm{d}f(x, x', t)}{\mathrm{d}t} = \lambda f \,, \tag{11.1}$$

with the solution

$$f(x, x', t) = e^{\lambda t} f_0(x_0, x'_0) \,, \tag{11.2}$$

where the subindex 0 characterizes the initial distribution f_0 or the initial phase-space variables. The latter, x_0 and x'_0, are related to x and x' by the equation of motion including the damping. Note that the phase space

M. G. Minty et al., *Measurement and Control of Charged Particle Beams,*
https://doi.org/10.1007/978-3-662-08581-3_11

density about each particle increases exponentially. Without a cooling force, the system would be Hamiltonian and the local phase-space density conserved ($\mathrm{d}f/\mathrm{d}t = 0$), so that $\lambda = 0$ in this case.

It is common to introduce action-angle variables I and ψ (where I is proportional to the square of the oscillation amplitude) via the relations

$$\frac{x}{\sqrt{\beta}} = \sqrt{2I}\cos\psi \,, \tag{11.3}$$

$$\sqrt{\beta}\left(x' + \alpha\frac{x}{\beta}\right) = -\sqrt{2I}\ \sin\psi \,, \tag{11.4}$$

where α and β are the usual alpha and beta function describing the linear optics (cf. Chap. 1). The angle variable ψ can be identified with the betatron phase, The distribution in the angle ψ is often uniform and random. In these case, the beam distribution function f only depends on the action variable I and, possibly, on the time, i.e., $f(x, x', t) = f(I, t)$. Indeed, in earlier chapters, e.g., Sect. 4.1, we have implicitly taken the randomness of the betatron phase as the definition of a 'matched' beam. For a mismatched or oscillating beam, the initial beam distribution is not uniform in the angle ψ, but any spread of the betatron frequencies results in a phase 'randomization', after a time roughly equal to the inverse of the frequency spread. A spread in the betatron frequency is always present. It arises, e.g., from a nonzero chromaticity and a finite energy spread, or from a dependence of the betatron tune on oscillation amplitude. Our following treatment assumes that this 'randomization time' is much shorter than the cooling time. The phase randomization and the cooling can then be mathematically decoupled, e.g., by averaging the equations describing the time evolution of the action over the betatron phase.

Cooling in the three phase-space dimensions results in an exponential damping of the 3 action invariants:

$$\left\langle \frac{\partial \dot{I}_i}{\partial I_i} \right\rangle = -\lambda_i \,, \tag{11.5}$$

where $i = (x, y, z)$. The angular brackets in (11.5) denote an average over both the angle variables and the azimuthal position around the storage ring, θ, i.e.,

$$\langle ... \rangle = \int_0^{2\pi} \frac{\mathrm{d}\psi}{2\pi}\,\frac{\mathrm{d}\theta}{2\pi}\,(...) \tag{11.6}$$

and the λ_i are the damping rates in the three planes.

Denoting the physical momenta by $p_i = \gamma m c v_i$ (v_i is the velocity for the ith degree of freedom in units of m/s, m the particle mass, and γ the relativistic factor) and considering a 'cooling force' F which changes the particle momenta according to $\dot{p}_k = F_k$, some algebra yields

$$\sum_i \left\langle \frac{\partial \dot{I}_i}{\partial I_i} \right\rangle = \sum_i \left\langle \sum_k \frac{\partial}{\partial I_i} \frac{\partial I_i}{\partial p_k} F_k \right\rangle$$
$$= \sum_i \left\langle \sum_k \frac{\partial}{\partial I_i} \frac{\partial I_i}{\partial p_k} F_k + \frac{\partial}{\partial \psi_i} \frac{\partial \psi_i}{\partial p_k} F_k \right\rangle$$
$$= \left\langle \sum_k \frac{\partial F_k}{\partial p_k} \right\rangle , \tag{11.7}$$

where we have made use of the fact that the average over ψ of any derivative with respect to ψ is zero. The sum of the action damping coefficients is

$$\lambda_x + \lambda_z + \lambda_s = -\left\langle \frac{\partial F_x}{\partial p_x} + \frac{\partial F_z}{\partial p_z} + \frac{\partial F_s}{\partial p_s} \right\rangle = \langle -\boldsymbol{\nabla}_{\boldsymbol{p}} \boldsymbol{F} \rangle , \tag{11.8}$$

independent of any coupling between the three planes of motion.

As an example, cooling due to synchrotron radiation and due to ionization cooling is approximately described by a cooling force that is anti-parallel to the particle velocity $\boldsymbol{v}$ [2],

$$\boldsymbol{F} = -a\boldsymbol{v} , \tag{11.9}$$

where the coefficient a may depend on the particle energy. The cooling is accompanied by a particle energy loss rate W,

$$\frac{\mathrm{d}E}{\mathrm{d}t} = -W = \boldsymbol{F} \cdot \boldsymbol{v} = -av^2 , \tag{11.10}$$

which can be compensated by an rf system. Assuming ultrarelativistic particles ($v \equiv |\boldsymbol{v}| = c$), the cooling force of (11.9) may be rewritten in terms of the energy loss as $\boldsymbol{F} = -\boldsymbol{v}W/c^2$, and direct evaluation then yields:

$$-\boldsymbol{\nabla}_{\boldsymbol{p}} \cdot \boldsymbol{F} = \left(\frac{W}{pc} \right) \left[2 + \frac{\partial \ln W}{\partial \ln p} \right] . \tag{11.11}$$

By inserting this expression into (11.8) the total decrease rate in phase-space volume can be calculated. Equations (11.8) and (11.11) state that the sum of the three damping rates is a constant, only depending on the total rate of energy loss. In the special case of synchrotron radiation, this is known as the 'Robinson theorem'.

One might think it would be possible to produce a beam of nearly zero temperature by cooling for a very long time. However, there is always some noise exciting the beam, which prevents reaching this limit and gives rise to an equilibrium emittance. In the case of synchrotron radiation this noise is due to quantum fluctuations, for ionization cooling it is due to multiple scattering, and in the case of stochastic cooling there is electronic noise in the detector-amplifier chain and Schottky noise arising from the finite number of particles in the beam.

With such noise sources present, the evolution of the distribution function $f(I,t)$ is no longer described by (11.1), but by a Fokker–Planck equation of the form

$$\frac{\partial f(I,t)}{\partial t} = \frac{\partial}{\partial I}\left(-\left\langle\frac{\Delta I}{\Delta t}\right\rangle f(I,t)\right) + \frac{1}{2}\frac{\partial^2}{\partial I^2}\left[\left(\left\langle\frac{(\Delta I)^2}{\Delta t}\right\rangle f(I,t)\right)\right] , \quad (11.12)$$

where now the angular brackets denote an average over the entire beam distribution, including the action variables, and over the noise. For example, if the Fokker–Planck terms $\langle\Delta I\rangle$ and $\langle(\Delta I)^2\rangle$ are linear in I and constant, respectively, the equation reduces to

$$\frac{\partial f}{\partial t} = \frac{\partial}{\partial I}\left(\lambda I f + \frac{D}{2}\frac{\partial f}{\partial I}\right) , \quad (11.13)$$

where $\lambda = \langle\Delta I/\Delta t\rangle/I$, and $D \equiv \langle(\Delta I)^2/\Delta t\rangle$. The beam then asymptotically approaches the distribution, $f_\infty \propto \exp(-I/I_\infty)$, with the equilibrium emittance (for the equality of rms emittance and average action see (1.14) and Ex. 1.1)

$$\epsilon = \langle I\rangle_{t=\infty} = I_\infty = \frac{D}{2\lambda} . \quad (11.14)$$

Using (11.13), this distribution is easily shown to be stationary: $\partial f_\infty/\partial t = 0$.

The cooling of various particles can be coupled, e.g., in stochastic cooling the time resolution may be limited by the amplifier bandwidth, and on each passage through the cooler only the average position of several particles is measured and damped. Under these circumstances, the beam is fully cooled only if the individual particles exchange their positions within the beam, so that on successive turns the measured average position, which is damped, refers to different combinations of particles. This process of particle exchange is called 'mixing'.

11.2 Electron Cooling

Electron cooling was proposed in 1966 by G.I. Budker [3]. The first experiments of electron cooling were performed at the NAP-M storage ring at the INP in Novosibirsk, where a 65-MeV antiproton beam was cooled down to a final momentum spread of 1.4×10^{-6} and to an angular divergence of 12.5 μrad, much smaller than the 3 mrad angular divergence of the 0.3-A 50-keV electron beam. Cooling times of the order of 25 ms were achieved [2].

11.2.1 Basic Description

Electron cooling is based on the heat exchange between a stored hadron beam and an accompanying electron beam via Coulomb collisions. The temperature

of the electron beam is held constant and lower than the temperature of the hadron beam to be cooled. This is easily fulfilled since for equal ion and electron velocities, $v_e \approx v_i$, the temperature of the electron beam is

$$T_e \approx \frac{m_e}{M} T_{\text{ion}} \,, \tag{11.15}$$

where M and m_e denote the ion and electron masses, respectively. Because of their mass ratio, the temperature of the ion beam is much larger than that of the electron beam. The average velocities of the hadron and electron beams should coincide in the cooling interaction region, in order to maximize the Coulomb cross section, which depends on the relative velocity. Viewed in the electron rest frame, moving with the electron beam, the ions are 'stopped' similarly to the slowing down of charged particles traversing matter, because in the Coulomb collisions energy is transferred from the ions to the electrons. The typical layout of an electron cooler and a photo of the electron cooling system at LEAR are depicted in Figs. 11.1 and 11.2, respectively.

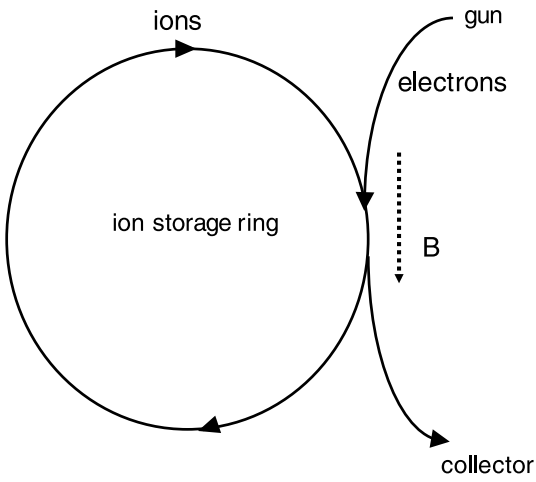

Fig. 11.1. Schematic of electron cooling for an ion storage ring

Transverse and longitudinal temperatures, $T_\perp$ and $T_\parallel$, of the ion beam can be defined by analogy with kinetic gas theory:

$$T_\perp = \frac{M\langle u_\perp^2 \rangle}{k_B} \,, \tag{11.16}$$

and

$$T_\parallel = \frac{M\langle \Delta u_\parallel^2 \rangle}{k_B} \,, \tag{11.17}$$

where M is the ion mass, u the ion velocity, and k_B the Boltzmann constant. The velocity components $\langle u_\perp^2 \rangle^{1/2}$ and $\langle \Delta u_\parallel^2 \rangle^{1/2}$ refer to the transverse

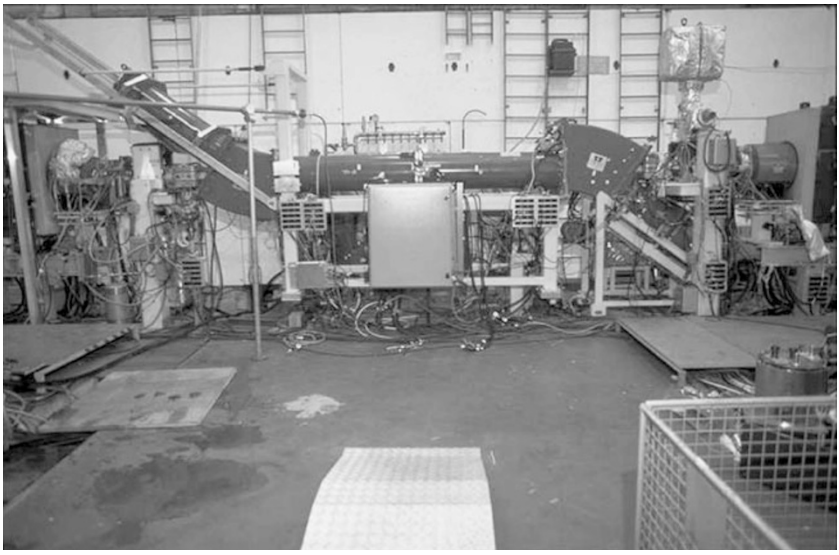

Fig. 11.2. Electron cooling system at LEAR (Courtesy M. Chanel, 1999)

and longitudinal rms velocity spread, repectively. The longitudinal velocity is taken to be the difference from the mean velocity of the ion beam, which is indicated by the prefix Δ. The transverse and longitudinal temperatures are usually not the same. Electron-beam temperatures are defined in the same way.

The cooling stops when the temperatures of the electron and ion beam are equal. The velocity of a cooled coasting ion beam (without rf) is equal to that of the electron beam, $v_{\mathrm{ion}} = v_e$. This provides a useful tool for tuning the ion beam energy. For a bunched beam, the rf frequency must be adjusted in order to match the revolution frequency of the ions as determined by the electron beam.

11.2.2 Estimate of the Cooling Rate

The cooling force may be estimated by considering the collision of a single ion with a single electron in a reference frame where the electron is at rest before the collision [2].

To this end, we split the collision into two steps. During the first step, the electron and ion approach each other, and in the second step they are separating again. We assume that during the first part the electron is accelerated by the field of the ion and that it moves in the direction of the impact parameter. The duration of either time step is of the order $\Delta t \approx \rho/u$, where ρ is the impact parameter and u the velocity of the ion. The situation is sketched in Fig. 11.3.

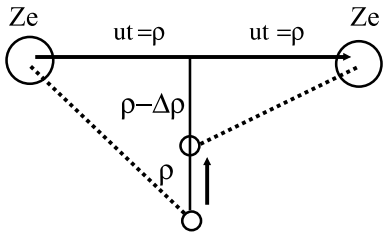

Fig. 11.3. Collision of one ion and one electron during electron cooling [2]

The ion with charge Ze moves from left to right and approaches the electron with an impact parameter ρ. The electron, initially at rest, is accelerated by the ion field. At the end of the first time step, the electron velocity in the direction of the closest approach is

$$\Delta v_e = \frac{Z r_e c^2}{\rho u} \,. \tag{11.18}$$

At this time the electron has moved by about a distance

$$\Delta\rho \approx \Delta v_e \Delta t \approx \frac{Z r_e c^2}{u^2} \,. \tag{11.19}$$

Integrating over the value of the initial impact parameter, the average change of the ion momentum in the direction u is

$$\left\langle \frac{\mathrm{d}p_u}{\mathrm{d}t} \right\rangle = u n_e m_e c^2 \int_{\rho_{\min}}^{\rho_{\max}} \left(\frac{Z r_e c}{\rho u} - \frac{Z r_e c}{(\rho - \Delta\rho) u} \right) 2\pi\rho \ \mathrm{d}\rho \tag{11.20}$$

where n_e denotes the local density of the electron beam, and m_e the electron mass. The limits of integration $\rho_{\max}$ and $\rho_{\min}$ refer to the maximum and minimum impact parameter. Expanding in powers of $\Delta\rho$ and keeping only the leading contribution, one finds

$$\left\langle \frac{\mathrm{d}p_u}{\mathrm{d}t} \right\rangle = \frac{2\pi n_e r_e^2 Z^2 m_e c^4}{u^2} \, L_C \tag{11.21}$$

where we have introduced the Coulomb logarithm $L_C \equiv \ln(\rho_{\max}/\rho_{\min})$. As an upper integration limit $\rho_{\max}$ we may take the Debye shielding length of the electron beam:

$$r_D \approx \left[\frac{k_B T}{4\pi m_e c^2 n_e r_e} \right]^{1/2} \,. \tag{11.22}$$

A lower limit $\rho_{\min}$ can be determined, e.g., from the maximum momentum transfer to the electron (classical head-on collision):

$$\rho_{\min} = \frac{Z r_e c^2}{u^2} \,. \tag{11.23}$$

In numerical estimates, L_C is usually taken to be constant, on the order of 10.

Averaging (11.21) over the electron velocity distribution function f_e results in the cooling force

$$F_{\mathrm{el}} = \left\langle \frac{\mathrm{d}\boldsymbol{p}}{\mathrm{d}t} \right\rangle = 2\pi Z^2 r_e^2 m_e c^4 L_C \int \mathrm{d}^3 v_e f_e(\boldsymbol{v}_e) \frac{\boldsymbol{v} - \boldsymbol{v}_e}{(\boldsymbol{v} - \boldsymbol{v}_e)^3} \,. \tag{11.24}$$

The result of a more precise evaluation of the cooling force starting from the Rutherford cross section agrees within a factor of 2 with (11.24) [4].

The cooling time τ_{el} follows from [5]

$$\frac{1}{\tau_{\mathrm{el}}} = \left| \frac{1}{u} \frac{\mathrm{d}u}{\mathrm{d}t} \right| = \left| \frac{F_{\mathrm{el}}}{Mu} \right| \,. \tag{11.25}$$

In the laboratory frame the cooling time is larger by a factor γ due to time dilation (there is a further factor of γ due to Lorentz contraction if the distribution function f_e is taken to be that in the laboratory frame). In the limit of large ion velocities, the electron velocity may be replaced by a delta function; in the opposite limit an isotropic Gaussian distribution is assumed. The cooling time in the two limits is [5]:

$$\tau = \frac{\gamma^2}{\eta_{\mathrm{el}}} \frac{M}{m_e} \frac{1}{Z^2 r_e^2 c^4} \frac{1}{\rho_L L_C} \begin{cases} \frac{1}{4\pi} u^3 & \text{for } u \gg v_{e,\mathrm{rms}} \\ \frac{3}{2\sqrt{2\pi}} \left(\frac{\frac{3}{2} k_B T_e}{m_e} \right)^{3/2} & \text{for } u < v_{e,\mathrm{rms}} \end{cases} \tag{11.26}$$

where η_{el} is the ratio of the cooling section length to the ring circumference, and ρ_L the electron beam density in the laboratory frame. The equation shows that electron cooling becomes inefficient for high energies, $\gamma \gg 1$, and that the cooling time is short for light ions of high charge. The cooling time of hot beams scales as u^3, while the cooling time of cold beams is independent of the ion velocities and only depends on the electron temperature.

Figure 11.4 shows a schematic of the transverse and longitudinal cooling forces, illustrating the two different cooling regimes incurred for high and low ion velocities.

Example parameters for electron cooling are $k_B T_e \approx 0.2$ eV, $n_L = 3 \times 10^8$ cm^{-3}, $L_C = 10$, $\eta = 0.05$, $\gamma = 1$, and $Z = 1$, which results in a cooling time of 40 s.

In reality, there are two additional effects which considerably help to reduce the cooling times: First, the electron velocity distribution is not Gaussian, but Maxwellian, and due to acceleration in the electron gun, the distribution is compressed in the longitudinal direction. This compression of the longitudinal velocity spread leads to shorter longitudinal cooling times. Second, a longitudinal magnetic field is employed to guide and confine the electron beam. This results in a cyclic motion of the electrons. If the cyclotron period is small compared with a typical ion-electron collision time, the cyclotron motion decreases the effective transverse temperature of the electron beam, and can reduce the cooling times, to values below one tenth of a second [2, 6, 5].

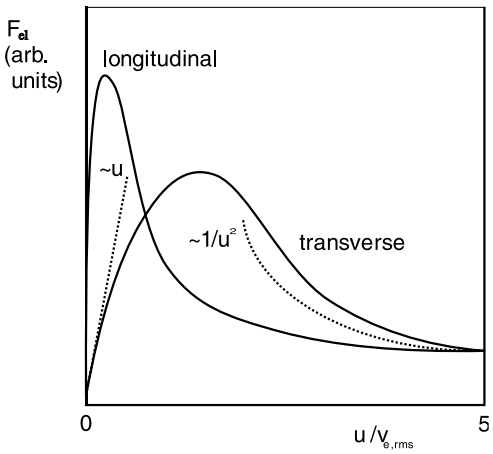

Fig. 11.4. Cooling force $F_{\rm el} = Mu/\tau_{\rm el}$ in a flattened electron beam as a function of ion velocity u in units of the rms electron velocity in the beam frame $v_{e,\rm rms}$. The *dashed curve* corresponds to the asymptotic formulae derived in the text. The difference between the transverse and longitudinal plane is due to the temperature difference, which arises from the longitudinal acceleration. Picture is redrawn from [5]

To relate electron cooling times for different types of particle beams, we note that the cooling rate scales like [1]

$$\frac{1}{\tau} \propto \frac{Z^2}{A} \,, \tag{11.27}$$

where A is the atomic mass of the ion, and Z the atomic number (i.e., the ion charge in units of the electron charge). We thus expect that cooling is faster for highly charged ions. However, these ions can also more easily capture a cooling electrons, thereby changing their charge, and get lost. The rate of recombination due to radiative electron capture scales approximately as [1]

$$\frac{1}{\tau_r} \propto Z^2 \,. \tag{11.28}$$

It is also worth mentioning that for relativistic energies electron cooling becomes less efficient; see, e.g., (11.26), where the cooling time τ increases as γ^2. In addition, higher electron-beam energies would be required in the cooling system.

11.2.3 Optical Functions at the Electron Cooler

If the electron beam temperature is low compared with that of the ion beam, the electron cooling rate varies as

$$\frac{1}{\tau} \propto \frac{1}{u^3} \sim \frac{1}{\theta_{x,y}^3} \,, \tag{11.29}$$

where $\theta_{x,y} = \sqrt{\epsilon/\beta_{x,y}}$ is the transverse rms divergence of the ion beam, and $\beta_{x,y}$ here the lattice beta function at the cooler (we assume that $\alpha_x \approx 0$). One might thus imagine that a large value of $\beta_{x,y}$ would give the best cooling results. However, for a large value of $\beta_{x,y}$ also the beam size is large, and the ions sample the nonlinear space-charge field of the electron beam. This space-charge effect complicates the electron-ion velocity matching. In addition, a large ion beam may only incompletely overlap with the electron beam. For this reason, an intermediate beta function turns out to be optimal, where the ion beam is slightly smaller than the electron beam.

One would also expect that the cooling rate increases in proportion to the electron beam current. In practice, however, for larger current one observes a tendency of saturation. Again, the limit arises from the space-charge force in the electron beam.

Let us take a closer look at the electron space-charge effects, and, in doing so, also explore the effect of a nonzero dispersion function at the electron cooler. Consider a cylindrically symmetric electron beam of radius a with a uniform transverse distribution and with a longitudinal density $\lambda = I/(e\beta c)$, where I is the current, e the electron charge, and c the speed of light. For a uniform charge distribution the space charge force is linear for radial positions $r < a$:

$$E_r = \frac{\lambda e}{2\pi\epsilon_0 a^2} r \ . \tag{11.30}$$

Sufficiently far away from the gun, the electron beam reaches an equilibrium state where the sum of kinetic and potential energy is a constant for all electrons and where the electron energy depends on the radial position as [7]

$$m_e c^2 \gamma(r) = m_e c^2 \gamma(0) + e \int_0^r \mathrm{d}r' \ E_r(r') \tag{11.31}$$

or

$$\gamma(r) = \gamma(0) + \lambda r_e \frac{r^2}{a^2} \ , \tag{11.32}$$

where $\gamma(r)$ is the Lorentz factor characterizing the energy of electrons at radius r, and r_e is the classical electron radius. Since

$$\frac{\Delta v(r)}{v} = \frac{1}{\gamma^2 - 1} \ \frac{\gamma(r) - \gamma(0)}{\gamma} = \frac{1}{\gamma^2 - 1} \ \frac{\lambda r_e r^2}{\gamma a^2} \ , \tag{11.33}$$

the velocity distribution in the electron beam is roughly parabolic as a function of the radial position. For high currents, the increase in the velocity spread of the electron beam degrades the cooling force. The situation is illustrated in Fig. 11.5, which shows the velocity of electron and ion beams as a function of radial position. From the figure, it is evident that a nonzero dispersion at the electron cooler can reduce the average velocity difference between electrons and a beam which is injected off-momentum, thereby improving the performance. In Fig. 11.5, the part of the beam which has already been cooled – the 'stack' – has zero momentum offset.

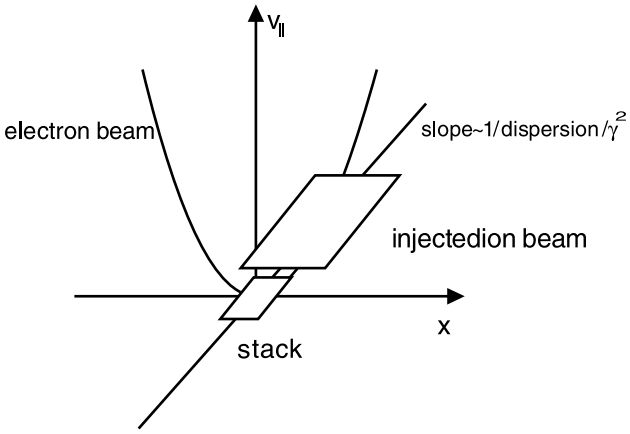

Fig. 11.5. Longitudinal velocity versus horizontal position of the electron and ion beams. Due to space charge the electron velocities lie on a parabola; the ion velocity varies linearly with a slope inversely proportional to the dispersion. Because of betatron oscillations, ions occupy a large area in phase space, as indicated [8, 9, 10] (Courtesy Ch. Carli and M. Chanel, 2002)

The optimum value of the dispersion function scales as [8, 9]

$$D \propto \sqrt{\frac{Ua^2}{I(\Delta p/p)_{\mathrm{rms}}}}\,, \tag{11.34}$$

where U is the accelerating voltage of the electron beam, I the electron current, $(\Delta p/p)_{\mathrm{rms}}$ the rms momentum spread of the ion beam, and a the electron beam size. The positive effect of a nonzero dispersion was confirmed by observations [8, 9].

11.2.4 Outlook

For the cooling of high-energy beams, it has been proposed to store the electron beam in a storage ring, sharing a common straight section with the ion or proton storage ring, where the cooling takes place [11, 12].

The emittance of the electron beam is then maintained by radiation damping. In such scheme, the bucket spacing of the electron storage ring should be an integral multiple of the bucket spacing of the ion storage ring [12]:

$$\frac{C_e}{h_e} = n\frac{C_i}{h_i} \qquad (n \text{ integer})\,, \tag{11.35}$$

where h_e and h_i denote the harmonic numbers for the electron and ion ring, respectively, and C_e and C_i the ring circumferences.

Recently, a novel scheme of high-energy electron cooling was proposed for the Relativistic Heavy Ion Collider (RHIC). The concept includes acceleration in a superconducting recirculating linac, strong solenoidal fields in the

cooling-interaction region, and energy recovery from the electron beam after its passage through the cooling section [13, 14].

11.3 Stochastic Cooling

Excellent rewiews of stochastic cooling are available [1, 5, 15, 16, 17]. Stochastic cooling was conceived in 1968 by van der Meer. Proton beam Schottky noise was first observed at the ISR in 1972, and first experimental demonstration took place in 1975, also at the ISR. In the period 1977–83, cooling tests were performed at CERN, FNAL, Novosibirsk and INS-Tokyo. In the 1977 cooling experiment ICE at CERN the momentum spread of 5×10^7 particles was reduced from 3.5×10^{-3} to 5×10^{-4}. At LEAR, in 1985, the momentum spread of 3×10^9 particles was reduced from 4×10^{-3} to 1.2×10^{-3} in 3 minutes of cooling [1]. At the CERN AA a factor 3×10^8 increase in phase-space density was achieved [1].

11.3.1 Basic Description

Figure 11.6 shows the process of stochastic cooling. A transverse pick up detects the displacement of a particle and feeds a signal related to the measured displacement through an amplifier to a kicker. The kicker applies a deflection which corrects the particle trajectory and reduces its betatron oscillation. The signal pulse induced by an off-axis particle and arriving at the kicker is of length $T_s \sim 1/(2W)$, where W is the bandwidth of the cooling system. The smallest fraction of beam that can be observed, the sample, is

$$N_s = \frac{NT_s}{T_0} = \frac{N}{2WT_0} \,, \tag{11.36}$$

where T_0 is the revolution time, and N the total number of beam particles.

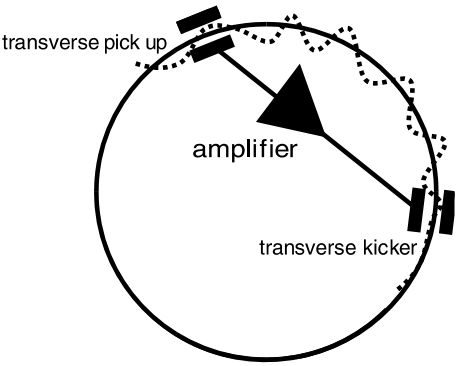

Fig. 11.6. Schematic of stochastic cooling

If the offset of the test particle is x, the applied correction is $-\lambda x$ where λ is related to the sensitivity of the pick up, the strength of the kicker, and the amplification in the cooling loop. The corrected position after the kick is

$$x_c = x - \lambda x - \sum_{\text{sample}'} \lambda x_i \ , \tag{11.37}$$

where the sum with superindex $'$ is over all particles in the sample except for the test particle. This can be rewritten as

$$x_c = x - \lambda N_s \langle x \rangle_s = x - g \langle x \rangle_s \ , \tag{11.38}$$

where $\langle x \rangle_s \equiv \sum_{\text{sample}} x_i / N_s$, and $g \equiv N_s \lambda$ is the fractional correction, also called the *gain*.

If the sample contains only the test particle, and assuming $g = 1$, the cooling time can be roughly estimated as [5]

$$\frac{1}{\tau_x} \equiv -\frac{1}{x}\frac{\mathrm{d}x}{\mathrm{d}t} = \frac{1}{N_s T_0} = 2\frac{W}{N} \ . \tag{11.39}$$

More rigorously, one computes the emittance damping rate

$$\frac{1}{\tau_{x^2}} \equiv \frac{1}{\epsilon_x}\frac{\mathrm{d}\epsilon_x}{\mathrm{d}t} = \frac{W}{N}\left[2g - g^2(M+U)\right] \ , \tag{11.40}$$

where $M \geq 1$ is a 'mixing' term, describing the exchange of particle positions between successive revolutions, and U denotes the noise-to-signal ratio. Simplified one might say that in practice [1]

$$\frac{1}{\tau_{x^2}} \approx \frac{1}{10}\frac{W}{N} \ . \tag{11.41}$$

A typical time constant is $\tau \approx 1$ s for $N \approx 10^7$ and $W \approx 100$ MHz. In (11.40), we have ignored an additional small mixing occurring between the pick up and the kicker.

Comparing (11.41), with the equations for electron cooling we observe that electron cooling works best for cold beams, and stochastic cooling works best for large (hot) beams, where the signal-to-noise ratio is large (U small), and for a small number of particles (small N). Thus stochastic cooling is good for 'halo cleaning', electron cooling for 'core freezing'.

Stochastic cooling for bunched beams has not yet been demonstrated. For this application a much higher bandwidth would be required. In addition, there are large signals at multiples of the revolution frequency, which must be avoided by operating at frequencies well above the $(1/e)$ fall-off frequency of the bunch power spectrum, $f_b \approx \beta c/\sigma_z$. However, at these high frequencies unexpectedly strong coherent signals were observed, which obstruct the observation of Schottky noise and thus the cooling [1, 18].

A promising alternative for bunched beams may be 'optical stochastic cooling', at much higher frequencies and bandwidths [19, 20].

11.3.2 Application: Emittance Growth from a Transverse Damper

It is interesting that the formalism of the stochastic-cooling equations can also be used to estimate the emittance growth induced by the transverse feedback system in a proton storage ring, such as the LHC [21].

We first need to modify (11.40) so as to more accurately include the response of the particle distribution to the cooling ('closing the loop via the beam') and convert the description into the frequency domain. Ignoring the mixing term and only keeping the Schottky noise contribution, the cooling equation becomes [21, 22]

$$\frac{1}{\tau_{x^2}} = \frac{1}{2}\frac{f_0}{N}\sum_{n=-\infty}^{\infty}\left[2\frac{g_n}{1+S_n} - \frac{U_n g_n^2}{(1+S_n)^2}\right] , \tag{11.42}$$

where f_0 is the revolution frequency, N the total number of particles, g_n the 'reduced' feedback gain, S_n the 'feedback via the beam' factor $S_n \approx g_n/(4\delta Q)$ [17], where δQ is the total 'tune spread' (depending on the shape of the distribution approximately equal to 2–3.5 times the rms tune spread), and U_n the ratio of noise and Schottky signal for full mixing. The sum extends over all 'Schottky bands' inside the bandwidth of the system, which, for a bunch-to-bunch damper is $W = 1/(2T_b)$, where T_b is the bunch spacing. In the frequency domain, there are two betatron bands per revolution harmonic and, thus, the total number of beam Schottky bands is $n_b = 2W/f_0$.

The Schottky signal power per band is

$$\left.\frac{\mathrm{d}x^2}{\mathrm{d}f}\right|_{\text{signal}} = \frac{\sigma^2}{Nf_0} , \tag{11.43}$$

where σ denotes the rms beam size. Assume that the amplifier noise is dominated by the quantization of the digital processing, and that the least significant bit of the ADC corresponds to a fraction α of the beam size. Then the amplifier noise is

$$\left.\frac{\mathrm{d}x^2}{\mathrm{d}f}\right|_{\text{noise}} = \frac{\alpha^2\sigma^2}{12W} . \tag{11.44}$$

Dividing this by the Schottky signal power gives

$$U_n = \frac{\mathrm{d}x^2/\mathrm{d}f|_{\text{noise}}}{\mathrm{d}x^2/\mathrm{d}f|_{\text{signal}}} = \frac{\alpha^2}{12}N\frac{f_0}{W} . \tag{11.45}$$

If N is very large, we can neglect the damping term in (11.42) – the first term in the square brackets –, and consider the amplifier-noise component alone. In that case, the above equation yields the emittance growth rate [21]

$$\frac{1}{\tau_{x^2}} = -\frac{4}{3}f_0\alpha^2\delta Q^2 , \tag{11.46}$$

where a negative sign of τ_{x^2} indicates growth. As an example, using LHC parameters, $f_0 = 11$ kHz, $\delta Q \approx 10^{-3}$ (due to beam-beam collisions), and $\alpha = 1/512$ (i.e., a 10 bit effective ADC within $\pm\sigma$), one finds $\tau_{x^2} \approx 50$ hours [21]. A similar treatment can be applied to estimate the impact of other noise perturbations, e.g., ground motion.

Without feedback, the beam becomes unstable, if the imaginary tune shift due to an impedance $\Delta Q_\perp$ exceeds the total tune spread divided by π, i.e., if $\Delta Q_\perp > \delta Q/\pi$ [7]. To obtain a stable beam, the gain g_n of the feedback must be larger than about $4(\pi \Delta Q_\perp - \delta Q)$, where $\Delta Q_\perp$ denotes the imaginary tune shift due to an impedance [7, 21].

11.4 Laser Cooling

11.4.1 Ion Beams

Laser cooling of atoms held in electromagnetic traps is well understood and widely used. In 1981 P. Channel suggested to apply laser cooling also to ions circulating in a storage ring [23].

Laser cooling exploits the Doppler shift in frequency such that the laser beam interacts selectively with ions of a certain energy. The Doppler shifted frequency in the ion rest frame is

$$\omega' = \gamma\omega(1 - \beta\cos\theta) \,, \tag{11.47}$$

where θ is the angle between the ion velocity and the incident laser. We denote by A and B a lower and upper level in the ion electronic state, respectively. Ions with a velocity β so that $\omega = \omega_{AB}$, corresponding to the transition $A \to B$, absorb photons, which are subsequently re-emitted. The emission is isotropic, while the momentum received during absorption is in the direction of the laser. In a single absorption, the ion acquires the recoil velocity:

$$v_r = \frac{\hbar\omega_{AB}}{m_{\text{ion}}c} \,, \tag{11.48}$$

where m_{ion} is the ion mass and $\hbar$ the reduced Planck constant. To avoid isotropic stimulated emission, while yet maintaining a short cooling time, the upper level B of the ion should have a short decay time. The ultimate beam temperature that can be reached is determined either by the energy of a single absorbed photon, or by a balance of cooling and heating due to the randomness in the spontaneous emission recoils,

$$T_{\min} = \frac{7}{20}\frac{\hbar\Gamma}{k_B} \,, \tag{11.49}$$

where Γ is the spontaneous decay rate (inverse lifetime) of the excited ion state. Laser cooling is illustrated schematically in Figs. 11.7 and 11.8.

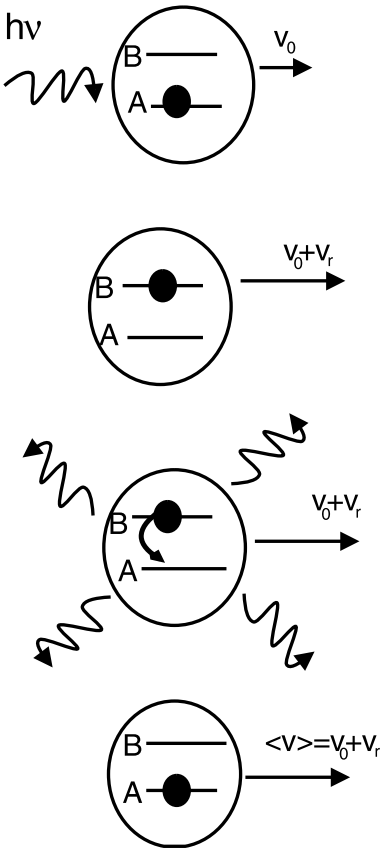

Fig. 11.7. Photon absorption and emission during laser cooling [5]. After each photon absorption the recoil component v_r is added to the initial ion velocity v_0. On the other hand, the emission is isotropic and, thus, on average it does not alter the final ion velocity $\langle v \rangle$

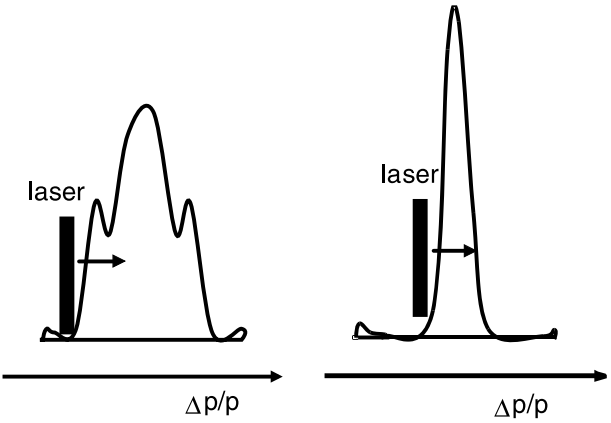

Fig. 11.8. Evolution of ion momentum distribution during laser cooling of a bunched ion beam

As an example [5], consider a 100-keV $^7Li^+$ beam. The ion transition at 548.5 nm is attainable using CW dye lasers. The lifetime of the upper state is 43 ns. The change in energy due to a single absorption is 12 meV. A few mW laser power on a 5-mm spot result in a spontaneous emission of 1.2×10^7 s^{-1}, or about 15 absorptions in an interaction region of 2 m length. This corresponds to a change in energy of 0.2 eV. To cool an ion beam with an energy spread of 1 eV would only require a few revolutions, or a few tenths of microseconds. The ultimate temperature is limited by the recoil momentum acquired in the absorption of a single photon.

Laser cooling requires adequate energy levels and transitions that can be reached by tunable lasers. So far, only 4 ion species fulfill this condition ($^7Li^+$, $^9Be^+$, $^{24}Mg^+$, and $^{166}Er^+$). Laser cooling was demonstrated experimentally in TSR and ASTRID, where energy spreads of less than 10^{-6} were obtained for Li beams [24, 25, 26].

So far laser cooling affects mainly the longitudinal temperature of a beam. However, it is believed that by resonantly coupling the synchrotron and betatron motion, the very fast laser cooling can be extended to the transverse phase space [27]. The coupling between synchrotron motion and horizontal betatron motion may be provided either by a special coupling cavity [28], or, more simply, by momentum dispersion in a regular rf cavity [29]. With such coupling present, the transverse cooling is considerably improved if the tunes are close to a linear resonance:

$$Q_x - Q_s \approx k \,, \tag{11.50}$$

$$Q_x - Q_y \approx l \,. \tag{11.51}$$

where k and l are integers.

11.4.2 Electron Beams

A different type of laser cooling was proposed by Telnov [30] for e^+e^- linear colliders, as a scheme to reduce the transverse emittances and to reach ultimately high luminosities. Collision of an electron beam with a high-power laser beam does not change the beam spot size, nor much the angular divergence. Only the beam energy is decreased, for example, from an inital value E_0 to E. This means that in a laser-cooling stage the two transverse normalized emittances decrease by a factor E/E_0. Telnov estimated that ultimate emittances of $\gamma\epsilon_{x,y} = 2 \times 10^{-7}$ m could be achieved, far better than what can be delivered by conventional damping rings.

More recently, Huang and Ruth studied a laser-electron storage ring (LESR) where radiative laser cooling overcomes the intrabeam scattering effect [31]. The LESR is sketched in Fig. 11.9. It consists of bending magnets, an rf cavity, an injector, and a laser-beam interaction region. A circulating bunch in the ring counterpropagates on each turn through the intense laser pulse. The laser pulse is stored in a high-Q optical resonator, whose path

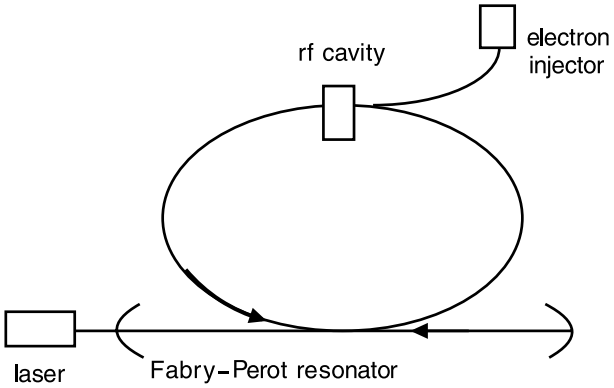

Fig. 11.9. Schematic of a laser-electron storage ring [31]

length is adjusted such that the laser-pulse repetition frequency equals the beam revolution frequency. Thus, a single laser pulse can interact several 10^4 times with the same electron bunch. The LESR can be configured either as a damping ring producing beams with very small transverse emittances, or as a high-intensity X-ray source.

The effect of the laser field is the same as that of a static wiggler with peak field strength [32]

$$B_w = \frac{2}{c}\sqrt{2Z_0 I}\,, \tag{11.52}$$

where I is the laser intensity and Z_R the vacuum impedance (377 Ω). Then the power radiated in the laser field is

$$P_\gamma = \frac{32\pi}{3} r_e^2 \gamma^2 I \tag{11.53}$$

and the energy loss of an electron per turn

$$(\Delta E)_\gamma = \int P_\gamma \frac{\mathrm{d}z}{2c} = \frac{32\pi}{3} r_E^2 \gamma^2 \frac{E_L}{Z_R \lambda_L}\,, \tag{11.54}$$

where Z_R denotes the laser Rayleigh length. The latter characterizes the depth of focus of the laser beam and it is the exact equivalent of a laser-beam 'beta function'. In (11.54), we have assumed that the laser beam is diffraction limited, so that the effective laser emittances are $\epsilon_{L;x,y} \approx \lambda_L/(4\pi)$, in which case its transverse spot area Σ_L at the focal point is given by $\Sigma_L \equiv 2\pi\sigma_{L,x}\sigma_{L,y} = Z_R\lambda_L/2$.

From the energy loss per turn we can compute the longitudinal damping time. It corresponds to a number of turns equal to

$$n_d = \frac{E}{(\Delta E)_\gamma} = \frac{1.6 \times 10^5 \lambda_L[\mu\mathrm{m}]] Z_R[\mathrm{mm}]}{E_L[\mathrm{J}]\ E[\mathrm{MeV}]}\,, \tag{11.55}$$

with E the beam energy.

The transverse emittances are damped at the same rate as the energy spread,

$$I_{x,y}^{\mathrm{RLC}} \equiv -\frac{1}{\epsilon_{x,y}} \left\langle \frac{\mathrm{d}\epsilon_{x,y}}{\mathrm{d}t} \right\rangle = \frac{1}{n_d T_{\mathrm{rev}}} = \frac{\Delta E_\gamma / E}{T_{\mathrm{rev}}} \,, \tag{11.56}$$

where T_{rev} is the revolution time, and RLC stands for 'radiative laser cooling'.

The laser field does not only provide damping, but in the same way as regular synchrotron radiation, it also introduces a quantum excitation. The quantum excitation consists of two parts: a dispersive component, which is dominant in conventional storage rings, and a component due to the finite opening angle of photon emission ($\theta \sim 1/\gamma$). The LESR is designed with zero optical dispersion in the laser-beam interaction region. A small amount of dispersion generated by the wiggler field is negligible compared with the effect of the opening angle, since the wiggle angle is much smaller than $1/\gamma$. This is quite different from the situation in a conventional ring, where the dispersive part is always much larger than the opening-angle contribution. Thus, in a conventional ring the emittance is determined by the dispersion (via the 'curly $\mathcal{H}$'; compare (4.67)), while in the LESR it is defined only by the opening angle.

The number of photons scattered into a frequency interval $\mathrm{d}\omega$ is [33]

$$\frac{\mathrm{d}N_\gamma}{\mathrm{d}\omega} = \frac{1}{\hbar\omega} \frac{\mathrm{d}E_\gamma}{\mathrm{d}\omega} = \frac{3(\Delta E)_\gamma}{\hbar\omega_m^2} \left[1 - 2\left(\frac{\omega}{\omega_m}\right) + 2\left(\frac{\omega}{\omega_m}\right)^2 \right] \,, \tag{11.57}$$

where the energy loss per turn $(\Delta E)_\gamma$ was given above, and $\omega_m = 4\gamma^2\omega_L = 8\pi\gamma^2 c/\lambda_L$ is the maximum photon frequency. The photon frequency ω and the scattering angle θ are related by

$$\omega = \frac{\omega_m}{1 + \gamma^2\theta^2} \,. \tag{11.58}$$

The transverse recoil of the electron is $\delta\psi = \hbar\omega\theta/E$, causing an average change in the normalized transverse emittances of $\Delta\epsilon_{x,y,N} \approx \beta_{x,y}^* \, \delta\psi^2/4$. Here, one factor of 2 is due to the projection onto a transverse plane, the other is due to averaging over the betatron phase.

Integrating over the photon spectrum yields the average emittance excitation per turn

$$\Delta(\epsilon_{x,y,N}) = \frac{\gamma\beta^*}{2} \int_0^{\omega_m} \mathrm{d}\omega \frac{\delta\psi^2}{2} \frac{\mathrm{d}N_\gamma}{\mathrm{d}\omega} = \frac{3}{10} \frac{\lambda_c}{\lambda_L} \frac{(\Delta E)_\gamma}{E} \beta_{x,y}^* \,, \tag{11.59}$$

where $\beta_{x,y}^*$ is the beta function at the laser-electron interaction point, and $\lambda_c = h/(mc) \approx 2.43{\times}10^{-12}$ m the electron Compton wavelength. The average emittance excitation per unit time reads

$$\left\langle \frac{\mathrm{d}\epsilon_{x,y,N}}{\mathrm{d}t} \right\rangle = \frac{(\Delta E)_\gamma}{T_{\mathrm{rev}}} \,. \tag{11.60}$$

As usual, the balance of damping, (11.56), and excitation, (11.60), determines the equilibrium emittance:

$$\epsilon_{x,y,N} = \frac{3}{10}\frac{\lambda_c}{\lambda_L}\beta^*_{x,y} \,. \tag{11.61}$$

According to (11.61), small emittances require a small beta function $\beta^*_{x,y}$. Reducing the value of β^* also facilitates the matching of the electron beam to the laser spot size, thus limiting the required laser-pulse energy.

Longitudinally, the energy spread increases due to the energy fluctuation of the emitted photons:

$$\left\langle \frac{\mathrm{d}(\sigma_E)^2}{\mathrm{d}t} \right\rangle = \frac{1}{T_{\mathrm{rev}}} \int_0^{\omega_m} \mathrm{d}\omega \; (\hbar\omega)^2 \; \frac{\mathrm{d}N_\gamma}{\mathrm{d}\omega} = \frac{7}{10} \; \frac{\hbar\omega_m \; (\Delta E)_\gamma}{T_{\mathrm{rev}}} \,. \tag{11.62}$$

As in a normal storage ring, the longitudinal damping occurs at a rate

$$\frac{1}{\sigma_E^2} \left\langle \frac{\mathrm{d}(\sigma_E)^2}{\mathrm{d}t} \right\rangle = -2\frac{\Delta E_\gamma / E}{T_{\mathrm{rev}}} \equiv -\Gamma_s^{\mathrm{RLC}} \,. \tag{11.63}$$

Equating the excitation and damping terms yields the equilibrium energy spread [31]

$$\sigma_\delta \equiv \frac{\sigma_E}{E} = \sqrt{\frac{7}{5}\frac{\lambda_c}{\lambda_L}\gamma} \,, \tag{11.64}$$

which tends to be much larger than in a conventional storage ring.

The increased energy spread widens the beam size in the arcs of the laser-electron storage ring, where the dispersion function is large. Thereby it both reduces the emittance growth rate due to intrabeam scattering and it keeps the incoherent space-charge tune shift at an acceptable value (for the above parameters, a bunch population of 10^{10}, an average beta function of 0.1 m, and 6 mm rms bunch length, the tune shift is about 0.01) [31]. However, the large energy spread demands a good chromatic correction, and a high-frequency rf system in order to maintain a short bunch length.

The depletion of the laser pulse due to its interaction with the electron beam is negligible. Neither does the laser-pulse energy significantly decrease over several damping times, provided the mirror reflectivity in the optical resonator is sufficiently high (i.e., 99.99% or better).

This scheme has not yet been demonstrated in practice, but several projects have been proposed and proof-of-principle experiments are under way [34, 35].

11.5 Thermal Noise and Crystalline Beams

Laser or electron cooling produce extremely cold beams. These beams have unusual noise spectra [36]. Suppose the azimuthal density of a stored unbunched proton beam is described by a Fourier expansion as

$$\rho(\theta, t) = \sum_{n=-\infty}^{\infty} \frac{A_n(t)}{2\pi} \exp(in\theta) \tag{11.65}$$

and

$$A_n(t) = \sum_{i=1}^{N} e^{-in\theta_i(t)} , \tag{11.66}$$

where i counts the particles and N is the total number of particles in the beam. In an ordinary beam, where the fluctuations arise from so-called Schottky or 'shot' noise, we have $\langle |A_n|^2 \rangle = N$.

Interaction of the particles via the external environment (characterized by the longitudinal impedance) suppresses the density fluctuation at the nth revolution harmonic as [36]

$$\left\langle |A_n|^2 \right\rangle = \frac{N}{1 + N/N_{\text{th}}} , \tag{11.67}$$

where the threshold number N_{th} follows from equating the longitudinal coherent frequency shift for the nth revolution harmonic $\Delta\Omega_n$,

$$(\Delta\Omega)_n^2 = n^2 \frac{N r_p mc 4\pi \omega_r (\mathrm{d}\omega_{\text{rev}}/\mathrm{d}p)}{Z_0 C} \left(\frac{Z_n}{n}\right) , \tag{11.68}$$

where p is the momentum, ω_r the angular revolution frequency, Z_0 the vacuum impedance, C the circumference and r_p the classical particle radius, to the spread in the revolution frequency ($n\ \delta\omega_r$) near $n\omega_r$. This yields [36]:

$$N_{\text{th}} = \frac{C Z_0\ (\delta\omega_{\text{r}})^2}{4\pi r_p mc \omega_r (\mathrm{d}\omega_r/\mathrm{d}p)} \left(\frac{n}{Z_n}\right) . \tag{11.69}$$

When the beam is cooled, N_{th} becomes smaller than N. Under these conditions the noise power of the beam no longer depends on the number of particles. Instead it is a direct measure of the beam temperature:

$$\left\langle |A_n|^2 \right\rangle \approx N_{\text{th}} \propto (\delta\omega_r)^2 \left(\frac{n}{Z_n}\right) . \tag{11.70}$$

The impedance Z_n/n can be determined from the observed shift in coherent frequency as a function of beam current. The remarkable suppression of the noise spectrum for a cold beam was first observed with an electron-cooled proton beam at the NAP-M storage ring in Novosibirsk [36].

The fast cooling techniques open up the exciting possibility to generate a new state of matter: a crystalline beam. Crystalline beams were proposed by Dikanski and Pestrikov [37], motivated by the observation at NAP-M [36]. Theoretical studies of crystal beams were first performed by Schiffer and Rahman [38, 39], and later by Wei, Li, Sessler, Okamoto, and others [40, 41, 42]. A crystalline beam is an ordered state, where the particles forming the beam

'lock' into fixed positions so that the repelling intra-particle Coulomb forces just balance the external focusing force. Crystalline beams might provide a route to obtaining ultra-high luminosity in colliders.

The generation and possible maintenance of the ordered state was investigated with molecular dynamics (MD) methods [38, 39] starting from a Hamiltonian describing the external focusing and the inter-particle forces in the beam frame. For example, and without derivation, in a combined-function cyclotron magnet, this Hamiltonian is [40]

$$H = \frac{1}{2}(P_x^2 + P_y^2 + P_z^2) - \gamma x P_z + \frac{1}{2}(1-n)x^2 + \frac{1}{2}ny^2 + V_c(x,y,z)\ , \quad (11.71)$$

with $n \equiv -\partial B_y/\partial x\ \rho/B_0$ measuring the strength of the quadrupole field, ρ the bending radius associated with the dipole field B_0, and the inter-particle potential

$$V_c = \sum_j [(x_j - x)^2 + (y_j - y)^2 + (z_j - z)^2]^{-1/2}\ , \quad (11.72)$$

where the summation is over all other particles. In the above Hamiltonian all dimensions were scaled by the characteristic distance $\xi = r_p\rho^2/\beta^2\gamma^2$, time is measured in units of $\rho/(\beta\gamma c)$, and energy in units of $\beta^2\gamma^2 Z^2 e^2/\xi$.

The beam-frame is an accelerated frame of reference, and the above Hamiltonian includes, so to speak, the relativistic generalization of centrifugal and Coriolis forces [40]. The effect of shear, given by the term $\gamma x P_z$, can render the Hamiltonian unbounded. This and the time dependent focusing in an alternating gradient focusing lattice may heat and melt the crystal.

Studying the circumstances under which the crystal is stable, one finds that two conditions have to be fulfilled in order to maintain the crystalline state [41]:

1. the storage-ring must be alternating focusing and the beam energy must be below the transition energy, and
2. the ring lattice periodicity should be larger than 2 times the maximum betatron tune.

The first condition arises from the requirement of stable kinematic motion. The second condition ensures that these is no linear resonance between crystal phonon modes and the machine lattice periodicity [42].

Although the crystalline ground state will show a periodic variation of its shape as the beam travels around the storage ring, at low temperatures very little heat is absorbed by the crystal and the crystal can remain stable for a very long time.

When the ion density is very low, the crystalline ground state is a 1-dimensional chain stretching around the ring. As the ion density increases, the 1-dimensional crystal changes into a 2-dimensional crystalline structure. This transition from the 1-dimensional to a 2-dimensional configuration occurs when the nearest-neigbor distance Δ_z (in the scaled units) obeys the equality [40]

$$\min(Q_y^2, Q_x^2 - \gamma^2) = \frac{4.2}{\Delta_z^3} \,. \tag{11.73}$$

The 2-dimensional structure extends into the transverse plane of weaker focusing. At still larger densities, a 3-dimensional crystal should be formed.

One-dimensional crystal beams have been observed in the ESR and SIS rings at the GSI Darmstadt, where they are generated by electron cooling [43].

11.6 Beam Echoes

An echo is a coherent oscillation which grows out of a quiet beam with some delay after the application of two independent pulse excitations. Echoes can occur in unbunched and in bunched beams, both transversely and longitudinally. The shape and magnitude of the echo signal contains information on diffusion processes in the beam and on the beam temperature (e.g., on the energy spread). Echoes may thus become a useful diagnostics tool for beam cooling.

We first give a simple illustration how an echo signal can arise. We next calculate the echo signal in the transverse plane induced by the successive application of a dipole kick and a quadrupole kick, closely following the pioneering work by Stupakov [44]. Then, we discuss experimental results, addressing both longitudinal echoes in unbunched beams and a different type of transverse echo, which is induced by two dipole kicks.

11.6.1 Illustration

The successive application of a dipole kick (at time $t = 0$) and a quadrupole kick (at a later time $t = \tau$) can generate an echo signal (at time $t \approx 2\tau$), as illustrated schematically in Fig. 11.10.

The first picture shows the dipole kick, which deflects two beam particles to different radial positions in phase space. After the kick, the two particles execute betatron oscillations, which are represented as circular movements about the phase-space origin. If the betatron tune depends on the radial position in phase space (i.e., on the amplitude of the oscillation), the two particles rotate at different angular velocities. This difference in angular velocity is indicated by the different lengths of the dashed arrows. As we shall see, the nonzero tune shift with amplitude is essential for producing the echo signal.

Next, after a certain time interval, a quadrupole kick is applied (the right picture). In this example, at the moment of the kick the particle with the larger betatron amplitude has no transverse offset ($x = 0$), and hence its motion is not affected by the quadrupole kick. On the other hand, the amplitude of the second particle is changed by the kick in such a way that its betatron

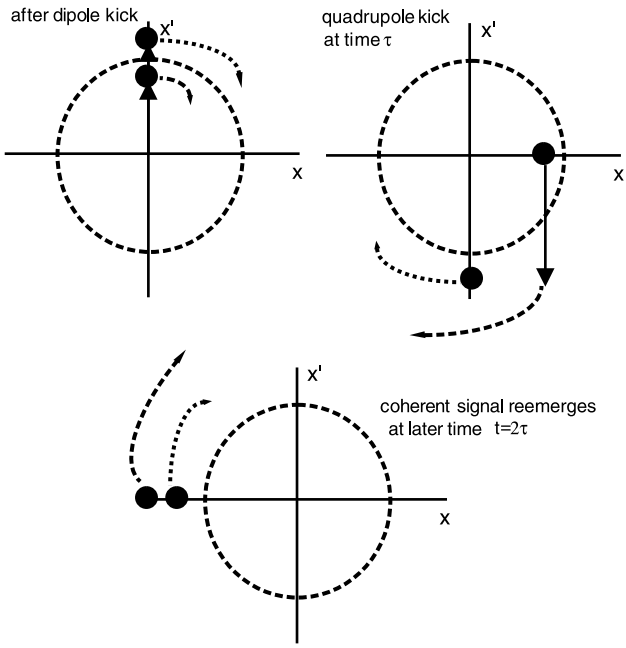

Fig. 11.10. Two-particle model of signal recoherence after applying first a dipole kick and then a quadrupole kick [45]. For the echo generation it is essential that the betatron tune depends on the oscillation amplitude

amplitude increases and now exceeds that of the first particle. Therefore, the quadrupole kick inverts the difference in betatron frequency between the two particles, i.e., the previously more slowly oscillating particle becomes faster and vice versa. After a further time interval, comparable to the time between the two kicks, the particles are again in phase (the last picture). This represents the 'echo'.

By considering the motion of additional particles in phase space, one could also illustrate that, at the moment of the echo, the betatron motion does not 'recohere' for all the particles in the beam, but only for a certain subset.

11.6.2 Calculation of Transverse Echo

We now calculate the response of the beam centroid to the dipole and quadrupole kick and derive an analytical expression for the echo response. To describe the transverse motion of particles in a storage ring, we here employ the normalized coordinates

$$\hat{y} = \frac{y}{\sqrt{\beta}} \quad \text{and} \quad \hat{p} = \frac{1}{Q\omega_r}\frac{\mathrm{d}y}{\mathrm{d}t}\,, \tag{11.74}$$

where β is the beta function, Ω_r the angular revolution frequency, and Q the tune. The beam dynamics can be studied using the distribution func-

tion $\rho(\hat{p},\hat{y},t)$ which is normalized so that $\int \rho(\hat{p},\hat{y},t)\ \mathrm{d}\hat{p}\,\mathrm{d}\hat{y} = 1$. The initial distribution is assumed to be Gaussian,

$$\rho(\hat{p},\hat{y},0) = \frac{1}{2\pi I_0}\exp\left(-\frac{\hat{p}^2+\hat{y}^2}{2I_0}\right)\ , \tag{11.75}$$

where I_0 is a constant equal to the rms beam emittance. It is customary to introduce action-angle coordinates (I,ϕ) via

$$\hat{y} = \sqrt{2I}\cos\phi\ , \tag{11.76}$$

$$\hat{p} = -\sqrt{2I}\sin\phi\ . \tag{11.77}$$

The initial distribution function then assumes the form

$$\rho_0(I,\phi) = \rho(I,\phi,0) = \frac{1}{2\pi I_0}\exp\left(-\frac{I}{I_0}\right) \tag{11.78}$$

and the transformation corresponding to free betatron oscillations conserves the action J:

$$I(t) = I(0)\ , \tag{11.79}$$

$$\phi(t) = \phi(0) + Q\omega_r t\ . \tag{11.80}$$

In the original coordinates this oscillation reads

$$\hat{p}(t) = \hat{y}(0)\ \cos(Q\omega_r t) + \hat{p}(0)\ \sin(Q\omega_r t)\ , \tag{11.81}$$

$$\hat{y}(t) = -\hat{y}(0)\ \sin(Q\omega_r t) + \hat{p}(0)\ \cos(Q\omega_r t)\ . \tag{11.82}$$

We assume that the tune Q depends on the amplitude of the oscillation as

$$Q(I) = Q_0 + \Delta Q\frac{I}{I_0}\ , \tag{11.83}$$

where ΔQ has the meaning of a tune shift with amplitude, which is crucial for the echo effect.

From the distribution function $\rho(I,\phi,t)$ we can calculate the evolution of the averaged (centroid) displacement, by means of a simple integration:

$$\langle\hat{y}\rangle = \int_{-\infty}^{\infty}\mathrm{d}\hat{p}\int_{-\infty}^{\infty}\hat{y}\rho(\hat{p},\hat{y},t)\,\mathrm{d}\hat{y} = \sqrt{2}\int_0^{\infty}\sqrt{I}\ \mathrm{d}I\ \int_0^{2\pi}\cos\phi\ \rho(I,\phi,t)\ \mathrm{d}\phi\ . \tag{11.84}$$

Our strategy is to compute $\rho(I,\phi,t)$ after applying the two transverse excitations, and then to obtain the echo signal in the motion of the beam centroid from (11.84). The evolution of the distribution function is governed by the Vlasov equation:

$$\frac{\partial\rho}{\partial t} + \frac{\partial\rho}{\partial\hat{y}}\frac{\mathrm{d}\hat{y}}{\mathrm{d}t} + \frac{\partial\rho}{\partial\hat{p}}\frac{\mathrm{d}\hat{p}}{\mathrm{d}t} = 0\ . \tag{11.85}$$

Alternatively and equivalently, the distribution function at time t can be obtained from that at time 0 by simply expressing the coordinates $\hat{p}(t)$ and $\hat{y}(t)$, or the corresponding action-angle variables, in terms of those at time 0. In other words, the Hamiltonian mapping induces the following transformation of the distribution function:

$$\rho(\hat{p}(0), \hat{y}(0), 0) \to \rho(\hat{p}(t), \hat{y}(t), t) = \rho(\hat{p}(\hat{p}(t), \hat{y}(t), 0), \hat{y}(\hat{p}(t), \hat{y}(t), 0), 0) \ ,$$

which links the distributions at times 0 and t. We will use this second method for computing $\rho(\hat{p}, \hat{y}, t)$.

Suppose that at time $t = 0$, the beam is displaced from the closed orbit by a transverse dipole kick of size $\Delta\hat{p} = \epsilon$. This dipole kick gives rise to the new distribution function:

$$\rho_1(\hat{p}, \hat{y}) = \rho_0(\hat{p} - \epsilon, \hat{y}) \ , \tag{11.86}$$

where ρ_0 is the initial distribution function at time $t = 0$, which we assume to be Gaussian. Assuming that the kick ϵ is small, we can expand the above equation to first order in ϵ:

$$\rho_1(\hat{p}, \hat{y}) \approx \rho_0(\hat{p}, \hat{y}) - \epsilon \frac{\partial \rho_0}{\partial \hat{p}} = \rho_0(I) + \epsilon\sqrt{2I} \sin\phi \ \frac{\mathrm{d}\rho(I)}{\mathrm{d}I} \ . \tag{11.87}$$

The first kick is followed by a free betatron oscillation over a time τ. This changes the distribution function as

$$\rho_2(I, \phi, \tau) = \rho_1(I, \phi - Q\omega_r\tau) \ . \tag{11.88}$$

Inserting the previous expression for ρ_1 we find

$$\rho_2 = \rho_0(I) + \epsilon\sqrt{2I} \sin(\phi - Q(I)\omega_r\tau) \frac{\mathrm{d}\rho_0(I)}{\mathrm{d}I} \ . \tag{11.89}$$

Using (11.83), (11.84), and (11.89), and performing the integration, we calculate the centroid motion after the dipole kick:

$$\langle\hat{y}\rangle = \epsilon \left[\frac{1 - \Delta Q^2\omega_r^2\tau^2}{(1 + \Delta Q^2\omega_r^2\tau^2)^2} \sin(Q_0\omega_r\tau) + \frac{2\Delta Q \ \omega_r\tau}{(1 + \Delta Q^2\omega_r^2\tau^2)^2} \ (\cos Q_0\omega_r\tau) \right] \ .$$

It is illustrated in Fig. 11.11 and clearly shows the decoherence of the signal. For large τ, the average displacement $\langle\hat{y}\rangle$ decreases as τ^{-2}.

At a later time, $t = \tau$ we apply a quadrupole kick of strength q:

$$\hat{p}_{\mathrm{new}} = \hat{p}_{\mathrm{old}} + \Delta\hat{p}_{\mathrm{quad}} = \hat{p}_{\mathrm{old}} - q\hat{y} \ . \tag{11.90}$$

The new distribution function is

$$\rho_3(\hat{p}, \hat{y}) = \rho_2(\hat{p} - \Delta\hat{p}_{\mathrm{quad}}, \hat{y}) \ . \tag{11.91}$$

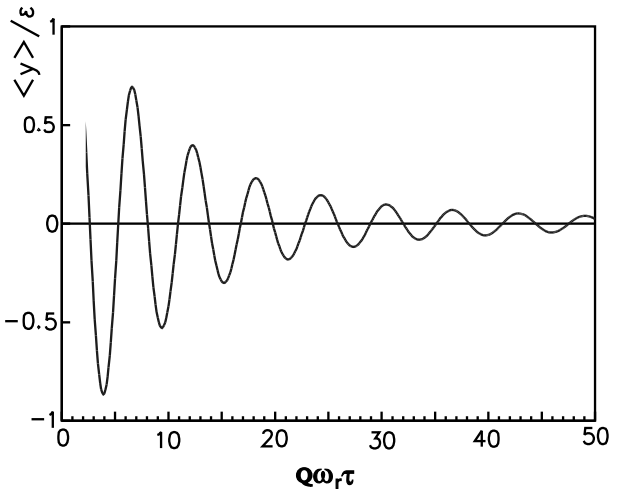

Fig. 11.11. Average displacement of the beam as a function of time following a dipole kick, for a tune spread $\Delta Q \approx 10\%$ [44]

We again perform a Taylor expansion, also assuming that the quadrupole kick is small, or, more precisely, that

$$qQ\omega_r\tau \ll 1 \; . \tag{11.92}$$

Inserting all the terms from above we have

$$\rho_3(\hat{p},\hat{y}) \approx \rho_0(I) + \epsilon\sqrt{2I}\sin(\phi - Q(I)\omega_r\tau)\;\frac{\mathrm{d}\rho_0(I)}{\mathrm{d}I} + q\hat{y}\frac{\partial}{\partial\hat{p}}\left[\rho_0(I) + \epsilon\sqrt{2I}\sin(\phi - Q(I)\omega_r\tau)\frac{\mathrm{d}\rho_0(I)}{\mathrm{d}I}\right] \; . \tag{11.93}$$

The echo effect is contained in the last term on the right-hand side of this equation [44]. Using the relation

$$\frac{\partial}{\partial\hat{p}} = -\sqrt{2I}\sin\phi\frac{\partial}{\partial I} - \frac{1}{\sqrt{2I}}\cos\phi\frac{\partial}{\partial\phi} \; , \tag{11.94}$$

the largest term that contributes to the echo comes from the derivative of $\sin(\phi - Q(I)\omega_r\tau)$ with respect to I. Denoting this term by ρ_3^{echo} one has

$$\rho_3^{\mathrm{echo}} \approx 2\epsilon q\Delta Q\;\omega_r\tau\sin(\phi)\cos(\phi - Q(I)\omega_r\tau)\frac{I}{I_0}\frac{\mathrm{d}\rho_0(I)}{\mathrm{d}I} \; . \tag{11.95}$$

Following the quadrupole kick, there is another free betatron oscillation of duration s (here the variable s is in units of time), with

$$\rho_4(I,\phi) = \rho_3^{\mathrm{echo}}(I,\phi - Q\omega_r s) \; . \tag{11.96}$$

Putting this into (11.84) and integrating, we finally obtain the equation for the echo response:

$$\langle \hat{y}^{\text{echo}} \rangle \approx q\epsilon \Delta Q \omega_r \tau \left[\frac{A(A^2-3)}{(1+A^2)^3} \cos(Q_0 \omega_r (\tau - s)) + \frac{3A^2-1}{(1+A^2)^3} \sin(Q_0 \omega_r (\tau - s)) \right] , \tag{11.97}$$

where $A \equiv \Delta Q \ \omega_r(\tau - s)$. The echo is illustrated in Fig. 11.12 as a function of the time difference $(\tau - s)$, for a tune spread of 10% ($\Delta Q/Q_0 = 0.1$). The peak of the echo signal is proportional to the strengths of the two kicks. It does not depend on the time interval between them. However, the time of the echo occurrence around $s = \tau$ (or $t = 2\tau$) of course does.

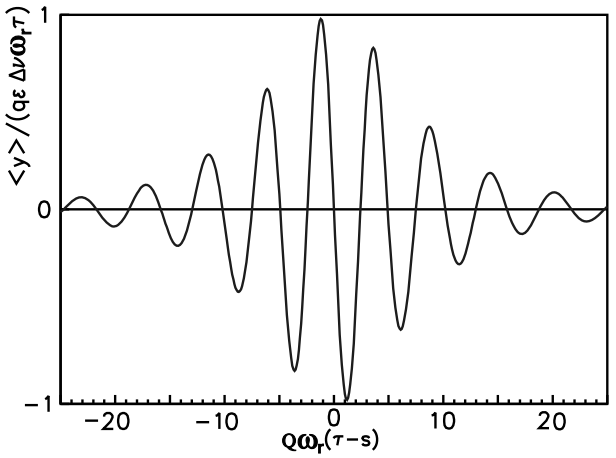

Fig. 11.12. Echo signal of the beam after a second (quadrupole) kick was applied [44]

11.6.3 Measurements of Longitudinal Echoes

Experimental results were first obtained for longitudinal echoes in unbunched beams. Such experiments were performed at the Fermilab Accumulator [46] and at the CERN SPS [47]. In these studies, two rf kicks were applied at frequencies f_{kick1} and f_{kick2}. The response was observed at the difference frequency [46]

$$f_{\text{echo}} = f_{\text{kick2}} - f_{\text{kick1}} \ . \tag{11.98}$$

For example if $h_{\text{kick1}} = 10$ and $h_{\text{kick2}} = 9$ (h is the harmonic number), the response occurred near the fundamental frequency $h_{\text{echo}} = 1$. Generalizing

the previous discussion, now the time of the echo, counted from the first kick, is

$$t_{\text{echo}} = \frac{f_{\text{kick2}}}{f_{\text{kick2}} - f_{\text{kick1}}}\tau \ , \tag{11.99}$$

where τ as before denotes the time separation between the two kicks.

The presence of diffusion destroys the reversibility of the decoherence. Diffusion thus reduces the response of the echo signal, especially for echoes at large times t_{echo}. The amplitude of the echo is of the form [46]

$$I_{\text{echo}} \propto J_1(\delta k_1 \tau) \exp(-D k_2 t_{\text{echo}}^3) \ , \tag{11.100}$$

where δ is proportional to the kick strength, k_2 is a constant which depends on the two kick harmonics (and on the echo harmonic), D is the diffusion rate (or collision rate), and t_{echo} the time interval from the first kick to the center of the echo. The decorrelation due to diffusion results in an exponential decay of the echo signal as t^3. By comparing the echo responses for different sets of harmonics, the contributions from the Bessel function and from the diffusion can be distinguished. In the Tevatron Accumulator, a diffusion rate of $d \approx 3 \times 10^{-4}$ Hz was measured [46].

A few further points should be mentioned. Exactly at the center of the echo the measured signal is zero. We have seen the same behavior in our above analysis for the transverse echo (compare Fig. 11.12). It is related to the fact that the echo signal is proportional to the slope of the distribution function, which is zero at the center of Gaussian a bunch. The separation of the two peaks, on either side of this zero, is inversely proportional to the energy spread within the bunch as [46]

$$\Delta t_{\text{peak}} = \frac{\beta^2}{h_{\text{echo}} \pi f_{\text{rev}} |\eta| \frac{\sigma_E}{E}} \ , \tag{11.101}$$

where η is the slippage factor, f_{rev} the revolution frequency, and β the velocity divided by the speed of light ($\beta = v/c$). This equation was confirmed experimentally.

If the distribution function is not Gaussian, the shape of the echo response changes. The echo signal thus permits a reconstruction of the actual beam distribution. Care has to be taken, as the echo shape may also be modified by longitudinal wake fields.

Another interesting observation is that for sufficiently large energy spread the notch at the center of the echo signal disappears. A possible explanation is the contribution from higher-order momentum compaction (or slippage) to the spread in revolution frequencies:

$$\frac{\Delta f_{\text{rev}}}{f_{\text{rev}}} = -\frac{\eta}{\beta^2}\frac{\Delta E}{E_0} = -\frac{1}{\beta^2}\frac{\Delta E}{E_0}\left(\eta_0 + \eta_1 \frac{1}{\beta^2}\frac{\Delta E}{E_0} + \dots\right) \ . \tag{11.102}$$

For a larger energy spread, the nonlinear contributions destroy the linear correlation between particle energy and phase.

In 2000, the first longitudinal echoes for a bunched beam were observed at HERA [48]. In the HERA experiments, first an rf phase kick was applied, which was then followed by an rf amplitude kick. This excitation pattern is the exact longitudinal analogue to the combination of a dipole kick and a quadrupole kick in the transverse plane, which we have analysed in Sect. 11.6.2.

11.6.4 Measurements of Transverse Echoes

Recently, it was discovered by F. Ruggiero that a sequence of two dipole kicks of largely different amplitude can result in a transverse echo [49]. Since all rings are equipped with (injection) kicker magnets, this scheme overcomes the difficulties that had been associated with transverse echo measurements, e.g., the assumed necessity of a quadrupole exciter.

As in Sect. 2.7.4, we assume that the betatron tune changes quadratically with amplitude

$$Q = Q_0 - \mu a^2 \,, \tag{11.103}$$

where a denotes the oscillation amplitude in units of σ and μ characterizes the strength of the nonlinear detuning. We denote the turn number by N and the magnitude of a dipole kick in units of $\sigma_{x'}$ by $Z = \beta \Delta x'/\sigma_x$ (for simplicity, we here assume that $\alpha_x = 0$ at the kicker).

During filamentation following a large kick Z_1, the beam distribution in betatron phase space will evolve into a spiral-like shape with an increasing number of closely spaced filaments. After N_t turns the distance between two adjacent filaments (i.e., occupied circular regions of phase space) is $\Delta a \approx 1/(2\pi\mu Z_1 N_t)$.

Numerical simulations suggest that for a maximum echo signal the amplitude of the second kick, Z_2, also normalized to $\sigma_{x'}$, should be chosen as half this distance, or

$$Z_{2,\mathrm{opt}} \approx \frac{N_{t,1/2}}{N_t Z_1} \,. \tag{11.104}$$

where we have introduced the number of turns, $N_{t,1/2}$, after which the initial signal amplitude has decreased by a factor of two (compare the discussion of filamentation in Chap. 2)

$$N_{t,1/2} \approx \frac{1}{4\pi\mu} \,. \tag{11.105}$$

Figure 11.13 presents a simulation result of the beam centroid response to two subsequent kicks of strength 5σ and 0.25σ, respectively. In the simulation a clear echo is observed if, as here, the first kick is several times the rms divergence and the second kick is so small that the displacement in phase space caused by the latter corresponds to roughly half the distance between filaments. Note also that the normalised rms beam size after applying a kick of strength Z and subsequent filamentation is given by

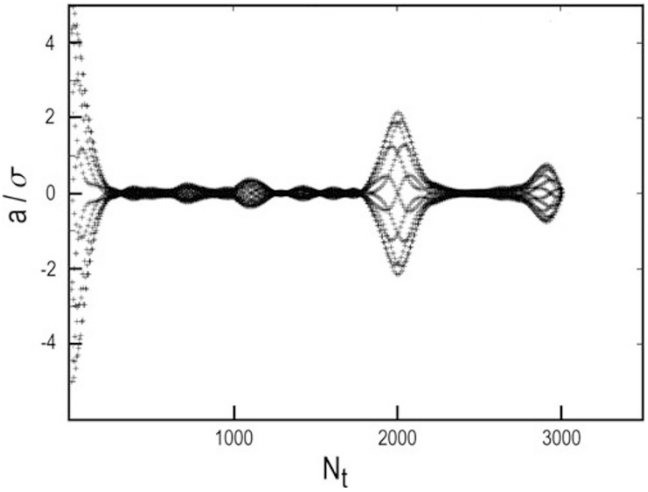

Fig. 11.13. Simulated centroid position in units of σ as a function of turn number for a normalized detuning $\mu = -2 \times 10^{-4}$ [49]. A first kick of 5σ applied at turn 0 is followed by second kick of strength 0.25σ at turn 1000. A clear echo signal is seen around turn 2000, and a second echo at turn 3000

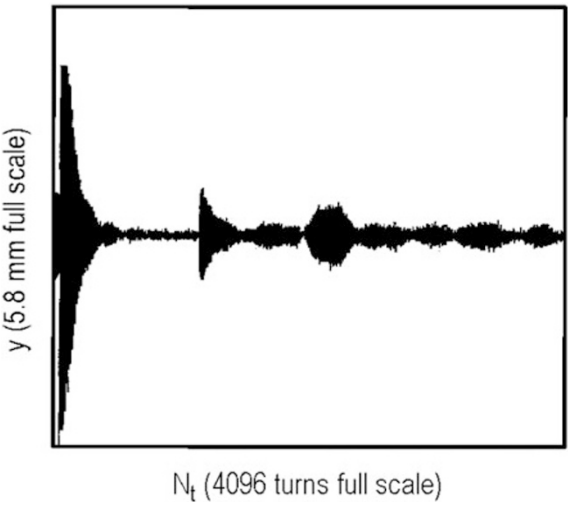

Fig. 11.14. Measured echo signal over 4096 turns [49]. The vertical centroid position is shown as a function of turn number. A first kick at 12884 ms (turn 0) of strength 5 kV, about 0.9σ, is followed at 12910 ms (turn 1128) by a second kick of strength 1 kV, approximately 0.2σ. An echo signal is visible around turn 2000. Octupole magnets were used to adjust the detuning with amplitude, μ, while the chromaticity was corrected by sextupoles. The bunch population was about 8×10^9 protons

$$\left\langle a^2(Z)\right\rangle_\phi = \frac{1}{(2\pi)^2}\int_0^{2\pi} \mathrm{d}\phi \int_0^{2\pi} \mathrm{d}\phi_0 \int_0^{\infty} \mathrm{d}a\; a^3 \sin(\phi+\phi_0)^2$$
$$\times \exp\left[-(a^2+Z^2-2aZ\cos\phi_0)/2\right] = \left(1+\frac{Z^2}{2}\right) ,$$

where ϕ_0 is the inital betatron phase of a particle prior to the kick, and ϕ is the phase advance over many turns, which becomes a random variable due to filamentation.

Figure 11.14 shows an echo signal measured during an exploratory experiment at the SPS [49]. In the vertical plane we first see the strong kick. Then the centroid motion is damped due to the decoherence induced by the nonlinearities. Next a second kick is applied after roughly 26 ms. And finally the echo signal appears about 52 ms (2256 turns) after the first kick.

A simulation was performed for parameters approximating those of the experiment, i.e., for a first kick of 0.9σ followed by a second kick of 0.2σ, separated by about 1000 turns. The result shown in Fig. 11.15 resembles the measurement in Fig. 11.14.

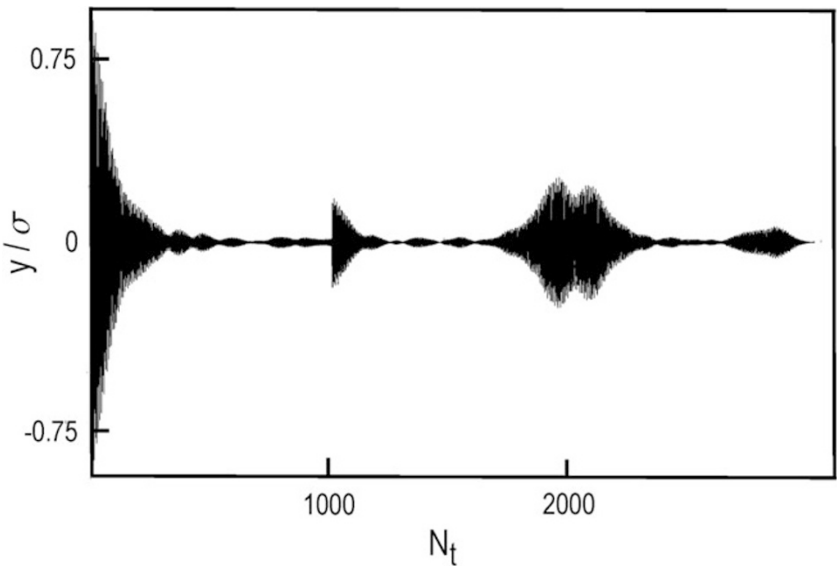

Fig. 11.15. Simulated centroid position in units of σ as a function of turn number for conditions similar to those in the previous figure [49]. A first kick of 0.9σ applied at turn 0 is followed by second kick of strength 0.2σ at turn 1000. Fig. 11.14 The simulation assumed a vertical tune of 0.5785, and a detuning with amplitude equal to $\mathrm{d}Q/\mathrm{d}I = 4000\ \mathrm{m}^{-1}$ corresponding to a normalized detuning $\mu = (-1)/(4\pi N_{t,1/2}) \approx -8\times10^{-4}$, with $N_{t,1/2} \approx 100$

11.7 Ionization Cooling

The successful operation of a future muon collider requires a reduction of the 6-dimensional beam phase space by about a factor of 10^{-6} [50]. The approach proposed to achieving this reduction is ionization cooling. Ionization cooling is similar to electron cooling, but the electron beam is replaced by a solid or liquid.

In ionization cooling the muon beam is passed through some material, in which the muons lose energy, experiencing an average force opposite to their momentum, as in (11.9). The average energy loss is described by the Bethe–Bloch formula

$$-\frac{\mathrm{d}E_\mu}{\mathrm{d}s} = 4\pi N_A r_e^2 m_e c^2 \rho \frac{Z}{A}\frac{1}{\beta^2}\left[\ln\left(\frac{2m_e c^2\beta^2\gamma^2}{I_{\mathrm{ion}}}\right) - \beta^2 - \frac{\delta(\gamma)}{2}\right] , \quad (11.106)$$

where N_A is Avogadro's number, the product $4\pi N_A r_e^2 m_e c^2$ equals 0.3071 MeV cm^2 g^{-1}, ρ is the material density, A and Z are mass number and atomic number, respectively, I_{ion} the average ionization energy, and $\delta(\gamma)$ in this expression represents a *density effect* (shielding by the atomic electrons), which at high energies approaches the value $2\ln\gamma$. The energy loss per length for Beryllium is shown in Fig. 11.16 as a function of the momentum of the incident muon. The muons lose kinetic energy in the direction of their motion. Only the longitudinal momentum is restored by subsequent rf sections, resulting in a transverse emittance reduction. This cooling effect is similar to the radiation damping arising from the energy loss due to synchrotron radiation in an electron storage ring. The ionization-cooling process must be repeated many times to achieve a significant emittance reduction. Figure 11.17 illustrates the concept of transverse ionization cooling.

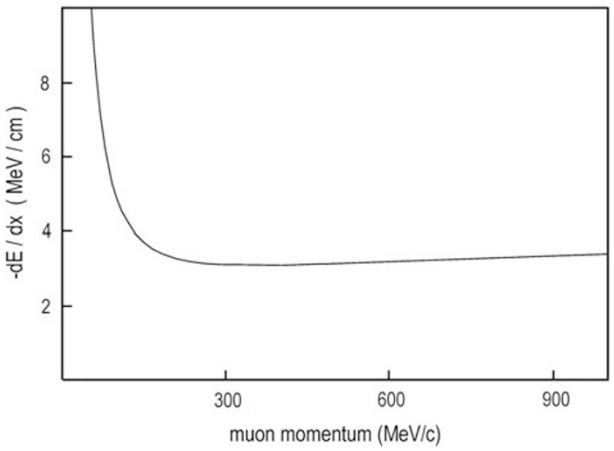

Fig. 11.16. Average muon energy loss per length in beryllium [51]

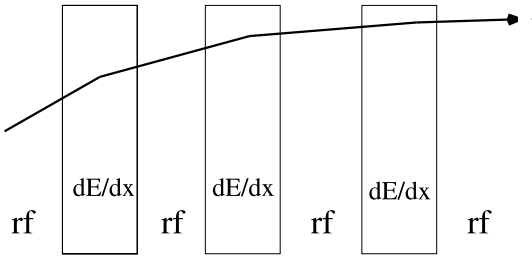

Fig. 11.17. Schematic of ionization cooling in the transverse phase space using a series of low-Z energy absorbers and reacceleration [52]

The equation describing the transverse cooling is [53]

$$\frac{\mathrm{d}\epsilon_N}{\mathrm{d}s} = -\frac{1}{\beta^2}\frac{\mathrm{d}E_\mu}{\mathrm{d}s}\epsilon_N + \frac{1}{\beta^3}\frac{\beta_\perp}{2}\frac{(14\mathrm{MeV})^2}{E_\mu m_\mu c^2 L_R}\,, \tag{11.107}$$

where ϵ_N is the normalized transverse emittance, E_μ the total muon energy, $\beta_\perp$ the beta function at the absorbing material, $\mathrm{d}E_\mu/\mathrm{d}s$ the energy loss per unit length, and L_R the radiation length. The first term in this equation describes the cooling, and the second the heating term due to multiple scattering. The heating is minimized if $\beta_\perp$ is small, and L_R large (low Z material).

If no further action is taken, the energy spread σ_E evolves according to

$$\frac{\mathrm{d}\sigma_E^2}{\mathrm{d}s} = -2\frac{\mathrm{d}\left(\frac{\mathrm{d}E_\mu}{\mathrm{d}s}\right)}{\mathrm{d}E_\mu}\langle\sigma_E^2\rangle + \frac{\mathrm{d}\left\langle(\Delta E_\mu)^2_{\mathrm{straggling}}\right\rangle}{\mathrm{d}s}\,, \tag{11.108}$$

where the first term is the cooling (or heating) due to the average energy loss and the second is the "energy-straggling" term given by [54]

$$\frac{\mathrm{d}\left\langle(\Delta E_\mu)^2_{\mathrm{straggling}}\right\rangle}{\mathrm{d}s} \approx 2\pi(r_e m_e c^2)^2 N_A \frac{Z}{A}\rho\gamma^2\,, \tag{11.109}$$

where N_A is Avogadro's number and ρ the density.

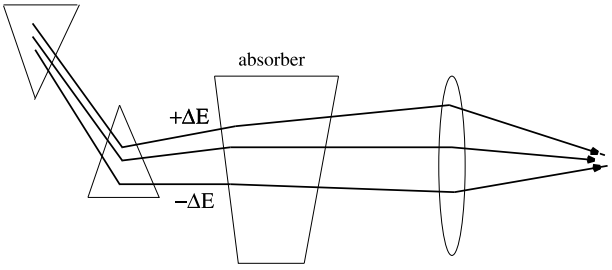

Fig. 11.18. Schematic of ionization cooling in the longitudinal phase space using a wedge [52]

The energy spread can be reduced by a transverse variation in the absorber thickness at a location with dispersion, as shown in Fig. 11.18. The use of such wedges reduces the longitudinal emittance, and it increases the transverse emittance. In other words, the longitudinal cooling is based on emittance exchange with the transverse plane.

Experiments to demonstrate the feasibility of ionization cooling are under study [52].

11.8 Comparison of Cooling Techniques

Table 11.1 compares the different cooling methods. Synchrotron radiation is most suitable for electrons and positrons, ionization cooling can be used for muons, laser cooling for ions, and, possibly, laser cooling of a different kind for electrons and positrons. Stochastic cooling and electron cooling are rather universal, and complementary. Stochastic cooling functions best for a hot beam and the beam halo. Electron cooling times decrease with decreasing beam temperature and electron cooling tends to damp the beam core. Both electron and stochastic cooling are by now well established and used at various storage and accumulation rings. Because of their strong complementarity often these two cooling schemes are employed together in the same ring.

Table 11.1. Comparison of cooling techniques, modified from [5]; N is the number of particles in the beam.

Technique	Stoch.	Electron	Synchr. rad.	Laser (ions)	Laser ($e^{\pm}$)	Ioniz.
Species	All	Ions	e^+, e^-	Some ions	e^+, e^-	Muons
favored beam velocity	high	medium $0.01 < \beta < 0.1$	very high $\gamma > 100$	any $\gamma < 5$	very high $\gamma > 100$	medium
beam intensity	low	any	any	any	any	any
cooling time	$N \cdot 10^{-8}$ s	1–10^{-2} s	$\sim 10^{-3}$ s	$\sim 10^{-4}$–10^{-5} s	$\sim 10^{-2}$–10^{-5} s	$< 10^{-7}$ s
favored temp.	high	low	any	low	any	any

It is predicted that stochastic cooling could be improved by orders of magnitude and possibly extended to bunched beams, if it is applied at optical frequencies. Like optical stochastic cooling, also the laser cooling of electron or positron beams is still waiting for experimental verification. Laser cooling of ion beams in a synchrotron has already been demonstrated, however. This scheme is presently applicable to 4 types of ions, Mg^+, Li^+, Be^+ and Er^+, for which impressively small momentum spreads of less than 10^{-6} have been achieved. Laser cooling has opened the path to a new regime of low temperatures and to the generation of *crystalline beams*. The cooling times for laser cooling are of the order of 10s or 100s of microseconds. They are surpassed only by the even shorter time scale projected for ionization cooling.

Exercises

11.1 Longitudinal Damping Rate with Beam Cooling

Consider two particles which interact simultaneously with the cooling system [2]. Let the cooling act on the momentum variable only. The equations of motion then read

$$\frac{\mathrm{d}p_1}{\mathrm{d}t} = -\lambda(p_1 + p_2) \tag{11.110}$$

and

$$\frac{\mathrm{d}p_2}{\mathrm{d}t} = -\lambda(p_1 + p_2)\ . \tag{11.111}$$

Calculate the damping rate of the centroid motion and the momentum spread.

11.2 Temperature of a Cooled Beam

For each plane of motion a beam temperature can be defined by analogy with kinetic gas theory:

$$\frac{\langle p_x^2\rangle}{2m} = \frac{1}{2}k_B T_x\ , \qquad \frac{\langle p_y^2\rangle}{2m} = \frac{1}{2}k_B T_y\ , \qquad \frac{\langle \Delta p_\parallel^2\rangle}{2m} = \frac{1}{2}k_B T_\parallel\ , \tag{11.112}$$

where all quantities refer to the beam rest frame.

a) Show that [55] $T_x = \frac{mc^2}{k_B}\beta\ \gamma\frac{\epsilon_{x,N}}{\beta_x}$ and $T_\parallel = \frac{mc^2}{k_B}\beta^2\sigma_p^2$ where $\epsilon_{x,N} \equiv (\gamma\beta)\epsilon_{x,y}$ is the normalized emittance (assumed to be equal in horizontal and vertical plane), $\sigma_p = (\Delta p/p)_{\mathrm{rms}}$ the rms momentum spread in the laboratory frame, and β_x the horizontal beta function (so the temperature is position dependent).

b) Calculate the horizontal and longitudinal temperature for the beam from a proton linac at injection into a cyclotron, with $\epsilon_{x,N} = 0.5$ mm mrad, $\beta\gamma \approx 0.7$, $\beta_x = 10$ m, and $\Delta p/p \approx 10^{-3}$. Compare this with the transverse and longitudinal temperatures of an electron, which is generated at the cathode with $k_B T^c = 0.1$ eV in all directions and then accelerated by a voltage $U_0 = 100$ kV.

c) What is the transverse Debye shielding length of the electron beam at this temperature? Assume a typical electron-beam density of 3×10^8 cm^{-3} in the laboratory frame.

d) For a longitudinal solenoidal guide field of strength 500 Gauss, calculate the electron cyclotron period and compare it with a typical impact time of $\sim r_D/u_\perp$, where $u_\perp$ is the relative transverse velocity (since the electron beam temperature is much smaller, this is determined by the temperature of the proton beam).

11.3 Recombination of Ion Beams during Electron Cooling

Assume an electron cooler for protons provides a cooling time of 10 ms, with a recombination time of 10^5 s. Suppose the same cooling system is used for a beam of fully stripped lead ions ($A = 207$, $Z = 82$). What is the fraction of lead ions that would be lost by recombination during one cooling time?

11.4 Electron-Beam Energy for Electron Cooling

What would be the electron-beam energy required to cool the 7-TeV LHC proton beam?

11.5 Derivation of the Debye Length

Derive the formula for the Debye length, (11.22), by calculating the electron density distribution in the potential of the ion charge and assuming the electrons are in thermal equilibrium. Make appropriate approximations.

11.6 Interaction Probabilities with Electron Cooling

Compare the minimum ion-electron interaction time $\Delta t = \rho_{\min}/u$ (in the beam frame), with the time of traversal through a 10 m cooling section. Estimate the ion velocity u assuming a normalized emittance 10 μm and a 5 m beta function. Can the two times become equal?

11.7 Beam Temperature with Ion-Beam Laser Cooling

Consider laser cooling for 100 keV Mg^+ ions ($A = 24$). Suppose the laser operates at a wavelength of 280 nm, equal to a short-lived transition with a natural linewidth of 46 MHz. (a) Which relative ion velocity $\Delta\beta/\beta$ corresponds to the laser tuning range of 20 GHz? (b) What is the ultimate temperature one might hope to achieve?

11.8 Damping Times with Electron-Beam Laser Cooling

Calculate n_d for the parameters $E_L \approx 1$ J, $\lambda_L \approx 1$ μm, $Z_R \approx 1$ mm, and $E \approx 100$ MeV. What is the equivalent damping time for an average ring radius of 1 m, assuming that electron and laser beams collide on each turn?

11.9 Equilibrium Emittances with Electron-Beam Laser Cooling

As an example, consider a ring with $E = 100$ MeV and $\beta^*_{x,y} = 1$ cm, and a laser with wavelength $\lambda_L = 1$ μm. Calculate the equilibrium emittance and relative energy spread.

11.10 Damping Rates and Equilibrium Emittances with Ionization Cooling

The emittance evolution in an ionization cooling system is described by the equation

$$\frac{\mathrm{d}\epsilon_N}{\mathrm{d}s} = -\frac{1}{\beta^2}\frac{dE_\mu}{ds}\frac{\epsilon_N}{E_\mu} + \frac{1}{\beta^3}\frac{\beta_\perp}{2}\frac{(14\mathrm{MeV})^2}{E_\mu m_\mu c^2 L_R}\,, \tag{11.113}$$

where ϵ_N is the normalized transverse emittance, $\beta_\perp$ the beta function at the absorbing material, $\mathrm{d}E_\mu/\mathrm{d}s$ the energy loss per unit length, and L_R the radiation length. Consider a muon beam with an initial normalized emittance of 0.01 m-rad and a kinetic energy E_k of 150 MeV. The muon mass is about 105.7 MeV. Assume that the beta function at the absorber is 10 cm, and that the minimum energy loss per length, $\mathrm{d}E_\mu/\mathrm{d}s$ is 0.29 MeV/cm.

(a) Calculate the average cooling rate λ (in units of m^{-1}) and the emittance reduction in a 10-m long cooling section containing 320 cm of liquid H_2 (radiation length L_R equal to 890 cm).

(b) Ignoring the second (heating) term, how many such stages and which total length would be required to damp the transverse emittance by a factor 10? In reality the complete cooling system might have a length of 500 m. Which fraction of muons is left after traversing this distance at 150 MeV? Note that the muon lifetime at rest is 2.2 μs.

(c) What is the minimum normalized emittance that can be achieved in such a cooling system of arbitrary length?

(d) Can you derive (11.113)? Note that the projected angular distribution due to multiple scattering is approximately Gaussian with an rms width after distance s equal to

$$\theta \approx \frac{14\ \mathrm{MeV}}{\beta cp}\sqrt{\frac{s}{L_R}}\,. \tag{11.114}$$