Beam cooling

H. Danared
Manne Siegbahn Laboratory, Stockholm, Sweden

Abstract
Beam cooling is the technique of reducing the momentum spread and increasing the phase-space density of stored particle beams. This paper gives an introduction to beam cooling and Liouville's theorem, and then it describes the three methods of active beam cooling that have been proven to work so far, namely electron cooling, stochastic cooling, and laser cooling. Ionization cooling is also mentioned briefly.

1 Introduction

Particles moving in an accelerator never have exactly the same velocities. In the system of reference moving with the average velocity of the particles, the velocity spread can be seen as a thermal motion, and beam cooling deals with the reduction of this velocity spread. Since beam size and velocity spread are coupled in a focusing system, cooling also means that the beam size shrinks, and thus the emittance is reduced and the phase-space density increases.

Depending on the type of accelerator and the application, the velocity spread can have several sources. It could come from the particle source, from the way that injection is made, from the acceleration process, from intra-beam scattering, from scattering in an internal target, etc. Reasons why one wants to cool beams include accumulation or stacking of rare particles, increase of luminosity in colliders, emittance control during deceleration, improved precision in measurements, and many others.

Although there are many different heating mechanisms, and many reasons to have cold beams, only a few methods for beam cooling exist. Three different methods have been demonstrated: stochastic cooling, electron cooling, and laser cooling. Of these, only the first two are used routinely. A fourth cooling method, ionization or muon cooling, will be tested within the next few years, but it is hardly applicable to 'small accelerators' which is the topic of this accelerator school.

We should add that we will only discuss active cooling in this text. Electrons have the nice property of cooling themselves on timescales much less than a second by emitting synchrotron radiation, which is a topic that will not be treated here. Furthermore, laser cooling and electron cooling, as well as other cooling methods, are in frequent use in traps, but no reference will be made to cooling in traps although, for instance, laser cooling in traps can be quite similar to laser cooling of stored ion beams.

The different methods for beam cooling are presented below, where the order and the amount of detail reflects their usefulness for small accelerators. First, however, we need to take a quick look into phase space, and a mention of Liouville's theorem is more or less obligatory in an introductory text on beam cooling.

2 Liouville's theorem

We have already said that to cool a beam is to increase its phase-space density. Can this not be done by inserting some specially constructed magnet into the ring, or by applying some electrical field with

a time dependence calculated once and for all by some clever person? The answer is no, and this follows from Liouville's theorem that one can look up in textbooks on classical mechanics.

Liouville's theorem is a theorem about the density of points in phase space (Fig. 1). It says that, in a dynamical system under the influence of conservative forces, the density of points, measured along the trajectory of a given point, does not change in time. The consequence is, as it seems, that no increase in phase-space density and no beam cooling can take place.

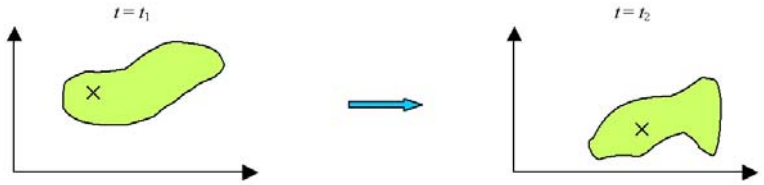

Fig. 1: Liouville's theorem says that the density of points in phase space, measured along the trajectory of a given point (like the one marked by an x) does not change in time

There are actually two different ways of formulating of the theorem [1] which apply to two different kinds of phase space. One is the so-called Γ space which has $6N$ dimensions for a beam with N particles, each particle having three space coordinates (x, y, and z) and three momentum coordinates (p_x, p_y and p_z). The whole beam with all its particles is thus represented by a single point in this space, and one gets a distribution of points in phase space by considering all possible beams that can be injected into the machine. Then there is the, perhaps more familiar, μ space, which has only six dimensions, and where each particle in a given beam is represented by a point. The distribution of points in μ space thus represents the particle density in one particular beam. Strictly speaking, the theorem as it is written above only applies to Γ space. In μ space it is in addition required that the particles do not interact. The theorem is still is a good approximation also in μ space, however, as long as hard collisions are not important. If the beam then is like a smooth fluid, self-forces cannot be distinguished from external forces.

One can conclude that beam cooling is a question about how Liouville's theorem can be circumvented using methods that violate the assumptions behind the derivation of the theorem.

3 Electron cooling

3.1 Introduction

With electron cooling, a particle beam stored in a synchrotron ring (or similar) is coupled to an external heat sink through a beam of cold electrons. This clearly violates Liouville's theorem unless the heat sink and the temperature of the rest of the world is included in the equations, and this is, of course, not interesting for the purpose of beam cooling.

Heat is transferred from the hot, stored particles to the cold electrons when the two beams, which must have the same average velocity, are merged in a section of the ring. Since the electrons are much lighter than the particles to be cooled, the required electron velocity can be reached by acceleration through a simple voltage gap in all electron coolers built so far. For this reason, it is easier to produce cold electrons than to have a circulating particle beam that is cold from the beginning.

The principle of electron cooling was suggested by Budker [2] in 1966, and the method was proven to work at the NAP-M storage ring in Novosibirsk in 1974. Since then, electron cooling has been used in many storage rings for many different kinds of particles including protons, light and heavy atomic ions up to uranium, molecular ions as well as antiprotons. For simplicity, we will refer to all these particles as ions in this text.

Figure 2 shows, as an example, a cross-section of the CRYRING electron cooler. The main components are magnets in the shape of solenoids and toroids (bent solenoids), the vacuum system, an electron gun and a collector. The magnets produce a field running from the electron gun sitting at one end of the device to the collector at the other end. The magnetic field guides the electrons, preventing the electron beam from blowing up under the influence of its own space charge. The electrons thus follow the field lines, making a spiralling cyclotron motion around them. The ions, which have a much higher mass, are not much influenced by the magnetic field, and are only slightly deflected in the toroids. This deflection is compensated by correction dipoles outside of the cooler.

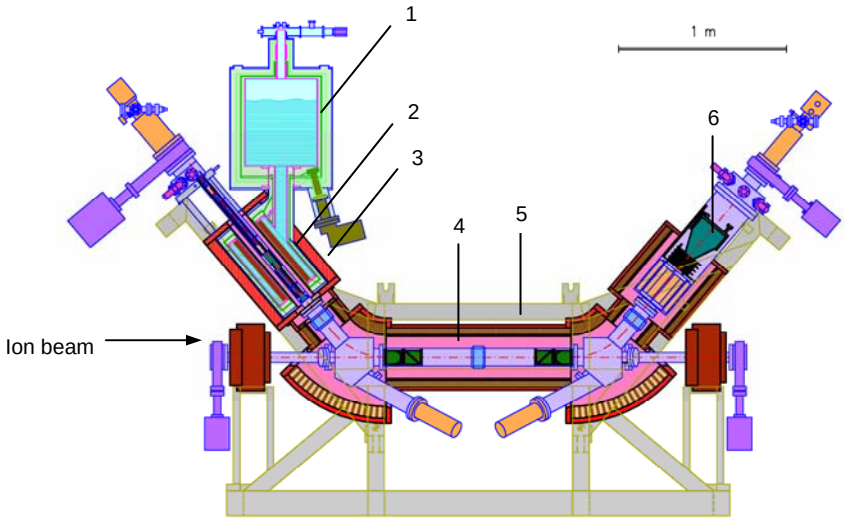

Fig. 2: The CRYRING electron cooler in cross-section with 1) liquid-helium reservoir, 2) electron gun, 3) superconducting solenoid, 4) interaction region, 5) normal-conducting magnets, 6) electron collector

3.2 Electron beam

In the electron gun, the electrons are emitted thermally from a cathode heated to T = 900–1000ºC, implying that the electrons have a Maxwellian velocity distribution with an energy spread kT in the order of 100 meV, k being Boltzmann's constant. Electrons are extracted from the cathode by an electrical field produced by an anode, like in a vacuum tube. The electron gun is operating in a space-charge-limited mode, which makes the current determined by the cathode-to-anode voltage and not by the cathode temperature. Having passed the anode, the electrons are accelerated toward earth potential through further electrodes or an acceleration tube that can be quite long if the voltage needed to reach the same velocity as the ions is high. In front of the collector, the electrons are decelerated again in order to reduce the power dissipation in the collector.

The electron current density from a space-charge limited gun is given by Child's law,

$$j = 2.33\times10^{-6}U^{3/2}/d^2\,, \tag{1}$$

where U is the voltage between cathode and anode, d is the distance between cathode and anode, and the numerical factor is just a combination of fundamental constants. Multiplying with the cathode area, one sees that the current is proportional to $U^{3/2}$ and that the constant of proportionality, known as the perveance P, is determined only by the geometry of the gun. In particular, the perveance is unchanged if the gun is scaled up or down, as long as its relative proportions are maintained. A non-relativistic uniform cylindrical electron beam propagating in vacuum has a theoretical maximum perveance of 32 μA/V$^{3/2}$. Above that limit, the electron beam will be repelled and reflected by its own space charge. The gun in an electron cooler has a much lower perveance, typically around 2 μA/V$^{3/2}$ or less.

Note that the electron current in principle is independent of the beam energy, since the cathode and the anode, as well as the collector, normally are connected to a high-voltage platform which is floating with respect to earth potential. Acceleration from platform to earth potential happens after the electrons have come out of the gun. The electron current, which depends on the voltage between cathode and anode, is thus independent from the electron energy which is given by the platform voltage.

In a well-designed gun, the electron temperature remains more or less equal to the cathode temperature transversally, but the longitudinal temperature decreases when it is measured in the frame of reference moving with the electron average velocity. If the average electron energy is E_0 (given by the cathode potential times the elementary charge), and a particular electron is faster that the average one by an amount Δv such that its energy is $E_0 + \Delta E$ in the laboratory frame, its energy in the moving frame is, non-relativistically,

$$E = \frac{1}{2} m_\mathrm{e} (\Delta v)^2 = \frac{1}{2} m_\mathrm{e} \left[\left(2(E_0 + \Delta E)/m_\mathrm{e}\right)^{1/2} - \left(2E_0/m_\mathrm{e}\right)^{1/2} \right]^2 \approx \frac{(\Delta E)^2}{4E_0} . \tag{2}$$

It is thus seen that the energy spread in the moving frame is lower than that in the laboratory frame by a factor $\Delta E/(4E_0)$.

Although temperature is a scalar quantity in thermodynamics, it is customary to define transverse and longitudinal beam temperatures. The anisotropic velocity distribution of the electron beam can then be written

$$f(\mathbf{v}_\mathrm{e}) = \frac{m_\mathrm{e}}{2\pi k T_\perp} \left(\frac{m_\mathrm{e}}{2\pi k T_\parallel} \right)^{1/2} \exp\left(-\frac{m_\mathrm{e} v_{\mathrm{e}\perp}^2}{2kT_\perp} - \frac{m_\mathrm{e} v_{\mathrm{e}\parallel}^2}{2kT_\parallel} \right), \tag{3}$$

where

$$kT_\perp \approx kT_\mathrm{cath} \quad \text{and} \quad kT_\parallel \approx \frac{(kT_\mathrm{cath})^2}{4E_0} \tag{4}$$

if all of the energy spread comes from the cathode temperature.

In reality, the lower limit of the longitudinal electron temperature is usually set by scattering processes within the electron beam rather than the cathode temperature, but $kT_\parallel$ can reach down to about 0.1 MeV. Also the transverse temperature can be reduced below the cathode temperature, as discussed below. On the other hand, imperfections in the magnetic field can give rise to transverse velocity components and an increased effective transverse temperature. It is therefore important that the quality of the magnetic field is high, with a very straight field in the interaction region. This is particularly important at high energies where a small misalignment between the ion and electron beams results in large relative velocities.

3.3 Theory

In order to understand the physics of electron cooling and to estimate cooling times, one needs to look in more detail at the Coulomb interaction between ions and electrons. This can be done by studying binary collisions between the two particles. Interestingly, expressions for cooling times can be taken from Spitzer's 1940 paper on the stability of star clusters [3] since the Coulomb interaction between charged particles has the same $1/r^2$ dependence as the gravitational force between massive bodies. We will, however, instead start from the Rutherford cross-section

$$\sigma(\theta) = \frac{Z^2 q^2}{4(4\pi\varepsilon_0)^2 \mu^2 u^4} \frac{1}{\sin^4(\theta/2)}. \tag{5}$$

Here, Z is the ion charge state, q the elementary charge, μ is the reduced mass $m_e m_i/(m_e + m_i)$ and u is the relative velocity. As usual in scattering problems, we have separated the relative motion from the motion of the centre of mass, which is the reason for the reduced mass appearing in (5), and we first concentrate on the relative motion.

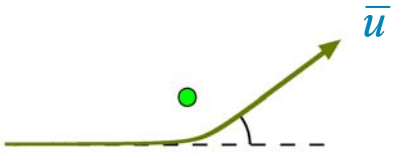

Fig. 3: Rutherford scattering between two charged particles with relative velocity u and scattering angle θ. The z axis is along the incoming velocity vector.

To begin with, we want to calculate the change in velocity due to a scattering event. From Fig. 3 it is seen that the change in longitudinal velocity is

$$\Delta u_z = u(1 - \cos\theta) = 2u \sin^2(\theta/2). \tag{6}$$

We do not need to worry about the transverse velocity change, since its average value is zero for symmetry reasons. Using the definition of the cross-section, we find the expectation value of the velocity change by integrating over all scattering angles:

$$\langle \Delta u_z \rangle = \int_\Omega \Delta u_z \sigma(\theta) u \, \mathrm{d}\Omega = 2\pi \int_{\theta_{\min}}^{\theta_{\max}} \Delta u_z \sigma(\theta) u \sin\theta \, \mathrm{d}\theta. \tag{7}$$

This integral in principle just consists of sine functions, so it is simple to evaluate except for the fact that the Coulomb force has an infinite range, making both the cross-section and the integral diverge when θ approaches zero. In a real electron beam, the ion charge becomes screened by the electrons as they rearrange around the ion. This so-called Debye screening limits the range of the force. Mathematically, we can thus solve the problem, at least superficially, by introducing the cut-offs $\theta_{\min}$ and $\theta_{\max}$ in the scattering angle, and we then find

$$\langle \Delta u_z \rangle = -4\pi \left(\frac{Zq^2}{4\pi\varepsilon_0} \right)^2 \frac{L_C}{\mu^2 u^2}, \tag{8}$$

where

$$L_C = \ln \frac{\sin(\theta_{\max}/2)}{\sin(\theta_{\min}/2)} \approx \ln \frac{\sin(\theta_{\max})}{\sin(\theta_{\min})} \approx \ln \frac{b_{\max}}{b_{\min}} \tag{9}$$

is called the Coulomb logarithm, and $b_{\min}$ and $b_{\max}$ are the cut-off values for the impact parameter (the perpendicular distance between the incoming trajectory and the scattering centre). The question is then instead how to calculate the Coulomb logarithm, but the answer to that is outside the scope of this text. We just state that, as a first approximation, it typically has a numerical value of about 10.

We are more interested in the change of absolute ion velocity than in the change of relative velocity, and they are related via the velocity of the centre of mass through $v_i = v_{cm} + u\, m_e/(m_i + m_e)$. Also, we use the fact that ions are much heavier than electrons, so the reduced mass can be approximated by the electron mass, and in addition we multiply the whole expression with the ion

mass in order to convert the velocity change to a force. As a result, we can write the average force acting on an ion as

$$\mathbf{F}_\mathrm{i}(\mathbf{v}_\mathrm{i}) = -4\pi\left(\frac{Zq^2}{4\pi\varepsilon_0}\right)^2 \frac{L_\mathrm{C}}{m_\mathrm{e}} \frac{\mathbf{v}_\mathrm{i}-\mathbf{v}_\mathrm{e}}{\left|\mathbf{v}_\mathrm{i}-\mathbf{v}_\mathrm{e}\right|^3}. \tag{10}$$

Finally, the force has to be integrated over the (flattened) Maxwellian velocity distribution of the electrons according to equation (3), giving

$$\mathbf{F}_\mathrm{i}(\mathbf{v}_\mathrm{i}) = -4\pi\left(\frac{Zq^2}{4\pi\varepsilon_0}\right)^2 \frac{n_\mathrm{e} L_\mathrm{C}}{m_\mathrm{e}} \int f(\mathbf{v}_\mathrm{e}) \frac{\mathbf{v}_\mathrm{i}-\mathbf{v}_\mathrm{e}}{\left|\mathbf{v}_\mathrm{i}-\mathbf{v}_\mathrm{e}\right|^3} \mathrm{d}^3 v_\mathrm{e} \tag{11}$$

if n_e is the electron density and the Coulomb logarithm is treated as a constant and taken outside the integral. The ion and electron velocities are measured in the moving frame, which is seen from the way $f(v_\mathrm{e})$ is defined in (3).

The integral has to be evaluated numerically except for some special cases. Using $kT_\perp = 100$ meV and $kT_\parallel = 0.1$ meV, the result is shown in Fig. 4 for the longitudinal and transverse force components as functions of longitudinal and transverse ion velocity, respectively. It is first seen that the forces and the velocities have opposite signs, which of course is a requirement for cooling. For small ion velocities, the force is linear in velocity. This results in an exponential decrease in beam temperature characterized by a time constant τ. Another way to say the same thing is that the cooling rate $1/\tau$ is constant in time. For larger velocities, the force decreases with v_i^{-2} and the cooling thus becomes slower. The force components have maxima near the thermal rms velocities which are

$$v_{\mathrm{e}\parallel,\mathrm{rms}} = (kT_\parallel / m_\mathrm{e})^{1/2} \quad \text{and} \quad v_{\mathrm{e}\perp,\mathrm{rms}} = (2kT_\perp / m_\mathrm{e})^{1/2}. \tag{12}$$

This is seen more clearly in Fig. 5.

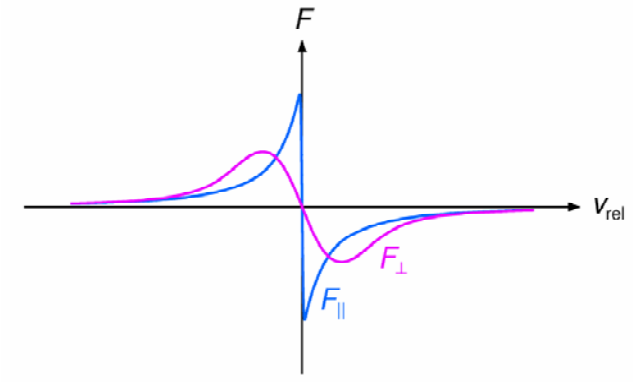

Fig. 4: Cooling force components according to the binary-collision model described in the text, although with a somewhat more elaborate Coulomb logarithm. The longitudinal electron temperature is 0.1 meV, and the transverse temperature is 100 meV. The longitudinal force is drawn as a function of a purely longitudinal relative velocity and the transverse force as a function of a purely transverse relative velocity.

From Fig. 5, it looks as if all ions will cool down to zero relative velocity and zero temperature. This is because we have integrated over the electron-velocity distribution to get an average force. In reality, individual ion–electron collisions lead to diffusion. Taking this into account, or using common sense, one finds that the ions will cool down to thermal equilibrium with the electrons, at best.

Normally, there are also heating mechanisms present, like intra-beam scattering, and then the equilibrium ion temperature will be higher than the electron temperature.

We already discussed the reduction in longitudinal electron temperature that follows automatically from the acceleration of the electrons. Using a magnetic expansion of the electron beam, also the transverse temperature can be reduced. This is achieved by using a strong magnetic field at the electron gun and a weaker field in the region where ions and electrons interact. In addition one has to make sure that the transition between the two fields is adiabatic, or slow, with respect to the cyclotron motion that the electrons perform around the magnetic field lines. If this adiabatic condition is fulfilled, one can show (see, for example, Ref. [4]) that the transverse energy (defined as $W_\perp = m_e v_{e\perp}^2/2$) divided by the longitudinal magnetic field $B_\parallel$ is an invariant. With a magnetic field which is 100 times stronger at the electron gun it should thus be possible to reduce the transverse electron temperature from approximately 100 meV to 1 meV. In practice, temperatures down to 2–3 meV have been observed. The electron-beam radius is proportional to the square root of the magnetic field, so if the beam radius in the interaction region is limited by the size of the beam pipe, the electron gun must be made smaller when an expanded electron beam is used. Note, however, that this does not change the electron current that one can reach, since the perveance of a space-charge limited electron gun is unchanged as long as all its dimensions are scaled with the same factor.

Although cold electrons in principle cool better than hot ones, the magnetic field in the cooler also has a positive influence on the cooling process. For low relative velocities between ions and electrons, i.e., for ion beams that are already relatively cold, the time during which ions and electrons interact can be long with respect to the cyclotron period of the electrons (the time it takes for an electron to make one turn in its spiral motion around a field line). The transverse electron velocity averaged over many cyclotron periods is thus very low, and, as a result, the ion interacts with an electron which has a very low effective transverse temperature. The simplest way to treat the 'magnetization' of the cooling is to insert this low effective electron temperature in the expressions for the cooling force, and then the result is the same as lowering the electron temperature through adiabatic expansion. This is illustrated in Fig. 5 where the lower, dashed curves are the same as in Fig. 4 but on a logarithmic scale. The upper curves differ only in that the transverse temperature (effective or real) has been lowered from 100 meV to 1 meV. The force is calculated for singly charged ions and is normalized to an electron density of 1×10^{14} m^{-3}.

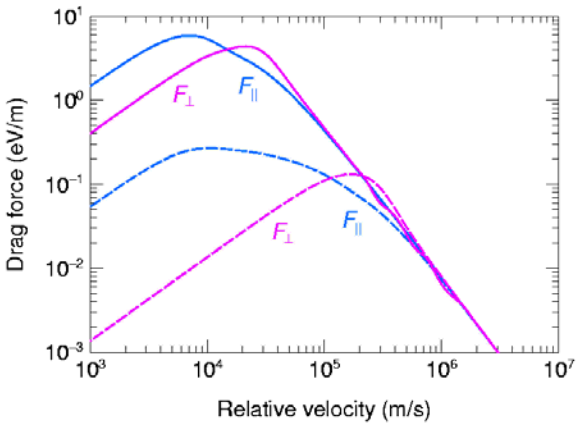

Fig. 5: Cooling force components according to the binary-collision model described in the text, although with a somewhat more elaborate Coulomb logarithm. The longitudinal electron temperature is 0.1 meV, and the transverse (effective or real) temperature is 100 meV for dashed curves and 1 meV for full-drawn curves. The longitudinal force is drawn as a function of a purely longitudinal relative velocity and the transverse force as a function of a purely transverse relative velocity. The force is normalized to $n_e = 1 \times 10^{14}$ m^{-3}.

As predicted above, the maximum of the longitudinal force in Fig. 5 is relatively close to the longitudinal thermal r.m.s. velocity $v_{e\parallel,\mathrm{rms}} = 4.2\times10^3$. m/s defined in equation (12). Similarly, the transverse force has maxima near $v_{e\perp,\mathrm{rms}} = 1.9\times10^4$ and 1.9×10^5] m/s for transverse temperatures 1 and 100 meV, respectively.

The cooling rate (in one dimension) is obtained from the force through

$$\frac{1}{\tau} = -\frac{1}{v_\mathrm{i}}\frac{\mathrm{d}v_\mathrm{i}}{\mathrm{d}t} = -\frac{F(v_\mathrm{i})}{m_\mathrm{i}v_\mathrm{i}}, \tag{13}$$

but one must take into account that the ions in a storage ring are not interacting with the electrons all the time, and one must therefore multiply the cooling rate with the ratio η_c between cooler length and ring circumference in order to get to the value one measures in practice. This is the cooling rate for the velocity. The cooling rate for the emittance, which is proportional to the transverse velocity squared, is twice as big.

For large relative velocities, where the force is proportional to v_i^{-2}, the cooling rate changes as the beam gets colder, and it is more interesting to talk about the total time it takes to cool the beam than about the rate. Using $F/m_\mathrm{i} = \mathrm{d}v_\mathrm{i}/\mathrm{d}t = \mathrm{const.}\times v_\mathrm{i}^{-2}$ and integrating, one finds that the cooling time becomes proportional to v_i^3 in this case. Clearly, electron cooling becomes less efficient when the beam is very hot.

For cooling at high energies it is also necessary to take relativistic effects into account. The electron density is lower in the moving frame, which the cooling force refers to, than in the laboratory frame by a relativistic γ factor because the electron beam is Lorentz-contracted in the lab frame. Also, time passes slower in the laboratory frame due to time dilation, introducing another γ factor. This causes the cooling rates to decrease and cooling times to increase by a factor γ^2 when they are measured in the laboratory frame, assuming a given relative velocity in the moving frame and an electron density that is measured in the laboratory frame. In addition, one may want to find cooling rates or cooling times for quantities measured in the laboratory frame rather than the moving frame, such as the beam emittance. Thus, for instance, the total emittance cooling time, in the range of transverse velocities where the force scales with v_i^{-2}, is proportional to γ^5 for a given initial (unnormalized) emittance and an electron density measured in the laboratory. We see that electron cooling of relativistic beams is difficult because cooling times become long, and it is also technically difficult to produce cold electron beams of high energy.

3.4 Experiment

Figure 6 shows, as an example, Schottky spectra of an uncooled and a cooled coasting deuteron beam of 23 MeV per nucleon at CRYRING. The Schottky spectrum of a coasting beam without any collective motion reflects the momentum distribution of the beam (see, e.g., Ref. [9]). Such is the case for the uncooled beam in the figure, where the relative momentum spread $\Delta p/p$ has a σ of approximately 6×10^{-4}. After about a second of cooling, the momentum distribution has shrunk by a large amount, but there has also developed some collective motion which is characteristic of cold beams, resulting in the double-peak spectrum seen in the figure. The discussion of such effects is outside the scope of this text, so we just state that the true relative momentum spread is smaller than indicated by the width of the double peak and in the order of 1×10^{-5}.

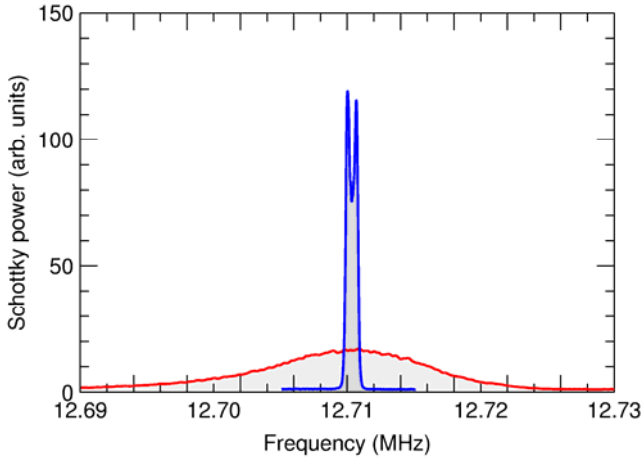

Fig. 6: Schottky spectra of uncooled and cooled 23 MeV/u deuteron beams at CRYRING

In Fig. 7, a sequence of transverse beam profiles from CRYRING is shown. Here, the beam was H^- at 3 MeV and the time interval between the frames was 300 ms. It is seen how the core of the beam cools faster than the tails, in accordance with $1/v^2$ decrease of the cooling force for large relative velocities.

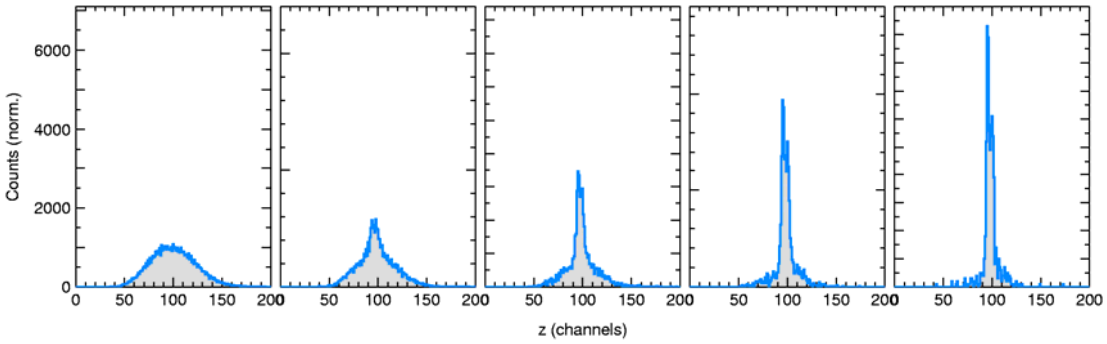

Fig. 7: Transverse profiles of a 3 MeV H^- beam at CRYRING. On the horizontal scale, 1 mm corresponds to 5 channels

It is a rather simple exercise to compare the observed transverse cooling rate with the theory that we have presented: From the figure we can estimate that the cooling time (reduction of transverse velocity or beam radius by a factor e) is, say, 0.8 s in the outer parts of the beam. Knowing the beam energy and that the vertical beta function at the position of the beam profile monitor is 5 m, we can calculate that a beam with a measured radius of 1 cm has transverse velocities up to 5×10^4 m/s at the profile monitor and 7×10^4 m/s in the cooler where the beta function is only 2.8 m. Using Eq. (13), we then obtain a cooling force of 1×10^{-3} eV/m. In order to compare this value with the theory in Fig. 5, we need to normalize to an electron density of 1×10^{14} m^{-3} and divide by the fraction of the ring occupied by the cooler (the quantity η_c defined below equation (13) which is 0.017 in our case). The electron current was 18 mA at the measurement and the electron-beam radius was 20 mm, so we find that the electron density was 3.8×10^{12} m^{-3}. The final result is thus a normalized transverse force in the order of 1 eV/m at 7×10^4 m/s. This is quite close to the curve in Fig. 5 for 1 meV transverse electron temperature, which is also the theoretical temperature since the beam expansion factor (ratio between magnetic field in the gun and in the interaction region) was 100 at this measurement. This quick calculation of course contains several simplifications. For instance, we have assumed that the cooling force is linear, and we have neglected that the transverse velocity of an ion varies from its maximum

value down to zero due to its betatron motion. The latter effect increases the expected cooling time since no transverse velocity means no transverse cooling. The comparison with theory also neglects the influence of the magnetic field on the cooling force.

Many electron coolers are used not only for cooling but also as electron targets for recombination experiments where cross-sections and reaction rates for recombination between stored ions and the cooler electrons are measured. The ions are normally first cooled so that ions and electrons get the same average velocity v_{cool}. The electron velocity is then detuned from its value at cooling by an amount v_{d}, giving a relative velocity which is the sum of this detuning velocity and the thermal electron velocity (assuming that the ions have a negligible velocity spread because they are heavier than the electrons but of similar temperature due to the cooling). The recombination cross-section or rate is then measured as a function of this detuning velocity v_{d}, or the corresponding energy $E_{\mathrm{d}} = m_{\mathrm{e}} v_{\mathrm{d}}^2/2$. Here, it is clear that thermal electron velocities of the same magnitude as the detuning velocity will blur recombination spectra, or, inversely, that colder electrons improve the spectral resolution.

Recombination spectra can be used to measure the electron temperature. Such measurements can be very accurate if the recombination spectra has narrow lines. These will then essentially be convoluted with the Maxwell distribution (3), and by fitting one can extract both the longitudinal and the transverse electron temperatures. This is discussed in more detail in Ref. [5]. Figure 8 shows an example from dielectronic recombination of C^{3+} ions. This spectrum was taken with an adiabatic expansion of a factor 10, such that one would expect a transverse electron temperature of around 10 meV, which indeed is very close to the measured value of 9.4 meV.

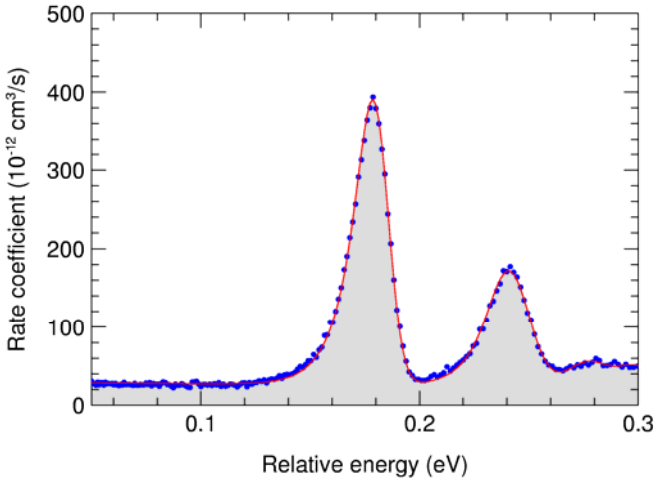

Fig. 8: Spectrum of dielectronic recombination of C^{3+} ions in CRYRING [6]. By fitting delta functions convoluted with the Maxwell distribution of equation (3) to the peaks, electron temperatures of 9.4 meV and 0.08 meV, transversely and longitudinally, respectively, were found.

Recombination can also be a problem since it can reduce the beam lifetime. This is particularly serious in the case of very highly charged ions that have large recombination cross-sections. Lifetimes can then be reduced by orders of magnitude.

Electron cooling has been implemented at more than a dozen storage rings since the early seventies, some of which have been taken out of operation since then, and 10 coolers were in use at the time of this accelerator school. These were at AD at CERN, ASTRID in Århus, CELSIUS in Uppsala, COSY in Jülich, CRYRING in Stockholm, the electrostatic ring in Tokyo, ESR and SIS in Darmstadt, HIMAC in Chiba, and TSR in Heidelberg. A few more low-energy coolers are under

development, i.e., coolers for electron energies below a few hundred keV. In addition, an electron cooler for 8.9 GeV/*c* antiprotons, requiring electrons of 4.4 MeV kinetic energy, is being commissioned at Fermilab, and there is a development programme at Brookhaven for cooling of 100 GeV/u gold ions using 54 MeV electrons from an energy recovery linac. In spite of technical difficulties and long cooling times, there is thus a trend in electron cooling in going toward higher energies.

4 Stochastic cooling

4.1 Introduction

It is quite clear that one can achieve cooling if one could have a detector that measures the orbit of each individual particle in the beam and some kicking device that can address individual particles and give them the desired momentum one by one. This resembles a 'Maxwell's demon', and such creatures are not included in the derivation of Liouville's theorem.

For reasonably intense beams, it is not possible to build a detector with a bandwidth high enough to resolve individual particles. Simon van der Meer realized, however, that the method works even if the particle orbits are not corrected one at a time, but that one can instead work with ensembles of particles if their average orbits are corrected repeatedly. This was written down in 1972 [7], and cooling was demonstrated at the Intersecting Storage Rings (ISR) at CERN a few years later. The idea earned van der Meer a Nobel Prize in 1984.

The principle is illustrated in Fig. 9 in the case of transverse cooling. A transverse pickup detects the centre of gravity of a sample of the beam. This centre of gravity results from the transverse positions that the particles of the sample happen to have at the instant when the sample is taken, and if the particles move independently, the pickup signal can be seen as a result of random fluctuations in the beam. We thus have to expect that the centre of gravity is a little bit off the nominal orbit (dashed line), and a signal is then induced in the pickup, amplified, and sent to a transverse kicker. Since the signal makes a straight shortcut it can catch up with the beam, and if the kicker is positioned an odd number of quarter betatron wavelengths downstream of the pickup where the offset in position has been transformed to an angular offset, the kicker can damp the betatron oscillation of the sample. In order for the cooling to proceed beyond the first turn, the samples that have been corrected once need to mix before they reach the pickup again, so that new fluctuations can be detected and corrected.

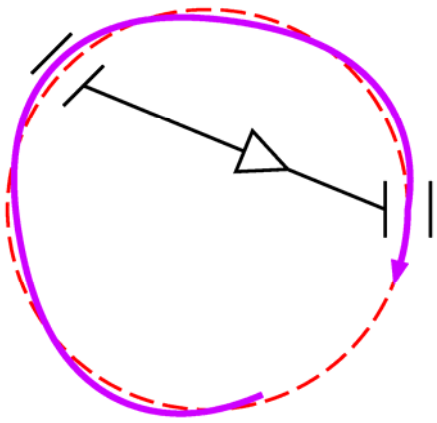

Fig. 9: Principle of transverse stochastic cooling: a transverse pickup detects that a sample of the beam is offset from the nominal orbit, the pickup signal is amplified and fed into a kicker that aligns the sample with the nominal orbit

4.2 Cooling time

Let us now try to estimate how long time it takes to cool a beam stochastically. We consider the case of transverse cooling (betatron cooling), and begin with a transverse pickup that has a bandwidth W. It samples the beam with a time resolution $T_s = 1/(2W)$ according to the Nyquist theorem. If the beam has a total of N particles, the number of particles in the sample is given by $N_s/N = T_s/T$, where T is the revolution time of the particles, or

$$N_s = \frac{N}{2WT}. \tag{14}$$

With x_i being the position of the individual particles we can write the centre of gravity of the sample as

$$\langle x \rangle = \frac{1}{N_s} \sum_{i=1}^{N_s} x_i. \tag{15}$$

We then assume, for the time being, that the kicker makes a full correction, such that it effectively displaces the sample by an amount $\Delta x = -\langle x \rangle$. (As we shall see further down, this is not necessarily optimal.) In order to quickly reach an approximate expression for the cooling time we make the trick of separating a test particle, which is given an index t, from the rest of the sample, such that

$$\langle x \rangle = \frac{1}{N_s} \sum_i x_i = \frac{x_t}{N_s} + \frac{1}{N_s} \sum_{i \neq t} x_i \equiv \frac{x_t}{N_s} + \langle x \rangle^*. \tag{16}$$

The test particle is displaced by the same amount as all other particles in the sample, so we have

$$\Delta x_t = -\langle x \rangle = -\frac{x_t}{N_s} - \langle x \rangle^*. \tag{17}$$

Clearly, the displacement of the test particle depends on its own position x_t but also on the position of the other particles through $\langle x \rangle^*$. We now take the, seemingly drastic, step of neglecting the influence of the other particles and just write

$$\Delta x_t = -\frac{x_t}{N_s}, \tag{18}$$

which gives a cooling rate per turn (skipping the minus sign) which is

$$\frac{\Delta x_t}{x_t} = \frac{1}{N_s}. \tag{19}$$

Multiplying with the revolution frequency and using the expression (14) we had for N_s, we end up with the cooling rate per second

$$\frac{1}{\tau} = \frac{1}{TN_s} = \frac{2W}{N}. \tag{20}$$

Surprisingly, this simple derivation overestimates the actual, optimal cooling time by only a factor 2.

A more rigorous derivation (see, for example, Ref. [8]) gives an expression

$$\frac{1}{\tau} = \frac{W}{N}\left[2g - g^2\left(M + U\right)\right]. \tag{21}$$

Note that this is the cooling rate for the amplitude x (or beam radius σ_x if the beam is Gaussian). The cooling rate for the emittance is twice that of the amplitude since the emittance is proportional to the amplitude squared. The expression contains the gain g, the mixing factor M and the noise-to-signal ratio U.

The gain g is proportional to the electronic gain, and it is defined as the ratio $-\Delta x/\langle x\rangle$ which was set equal to 1 in the simple analysis above.

The mixing factor M is related to the fact, also mentioned above, that the samples that have been corrected in the kicker need to mix again for new fluctuations to appear and be corrected. More precisely, M is equal to the number of turns it takes for the samples to get randomized after a correction. This is equal to the ratio of the sample time T_s to the spread in revolution time among the beam particles which we write as ΔT. Using the standard relation $\Delta T/T = \eta \Delta p/p,$ where η is the frequency-slip factor, we see that the mixing factor can be written as

$$M = \frac{T_S}{\Delta T} = \frac{1}{2WT\eta\,\Delta p/p}. \tag{22}$$

Note that efficient mixing means small M, but that M always has to be ≥ 1. As a side remark, one can quite easily see that M is related to the amount of separation between the Schottky bands, such that $M = 1$ means that the bands overlap completely. (Schottky spectra are discussed in, e.g., Ref. [9]) What we have discussed now is the so-called good, or desirable mixing. There is also a bad, or unwanted mixing which occurs between the pickup and the kicker and which we neglect here.

The noise-to-signal ratio U depends both on the beam and pickup (for the signal strength) and on the electronics (for the noise). We will come back to this quantity further down.

It is seen (by checking when the derivative with respect to g becomes zero) that the highest cooling rate is obtained when $1/g = M + U$. The cooling rate is thus at best

$$\frac{1}{\tau} = \frac{W}{N}\frac{1}{M+U}. \tag{23}$$

Since M and U in general change during the cooling process, $1/\tau$ is just an instantaneous rate, and optimal cooling requires that the gain is adjusted as the beam gets colder.

Longitudinal stochastic cooling or momentum cooling can be made according to two different basic principles illustrated in Figs. 10 and 11. One technique, Palmer cooling, uses the correlation between momentum and position in regions of high dispersion. A transverse pickup of the same type as is used for transverse cooling connected to a longitudinal kicker or rf gap then can provide cooling. If the dispersion is positive, the phase of the amplifier of course has to be adjusted such that a displacement outward from the centre of the ring in the pickup results in a decelerating field in the kicker and vice versa. The derivation of the cooling time is very similar to the transverse case, and the resulting expression looks the same.

The other technique uses notch filters to separate fast particles from slow ones. What is needed is a filter sitting between a longitudinal pickup and a longitudinal kicker. The filter is periodic, the period is equal to the desired revolution frequency ω_0, and the output signal changes sign when ω goes through $n\omega_0$ as illustrated in Fig. 11. The reason for having a filter with many such notches is that one then can add the signal from many Schottky bands and thus gain signal strength. It is seen that cooling will occur if the overall phase is adjusted such that when the pickup detects a sample with too high an average frequency, the field in the kicker decelerates the beam, and vice versa.

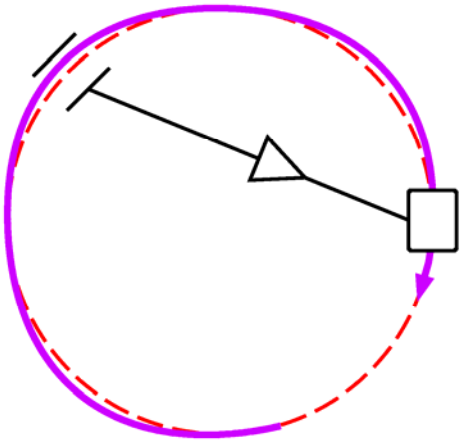

Fig. 10: Principle of longitudinal stochastic cooling using dispersion to correlate beam momentum with position. A transverse pickup can detect if a sample has a momentum which is too low or too high, and a longitudinal kicker can correct the momentum error.

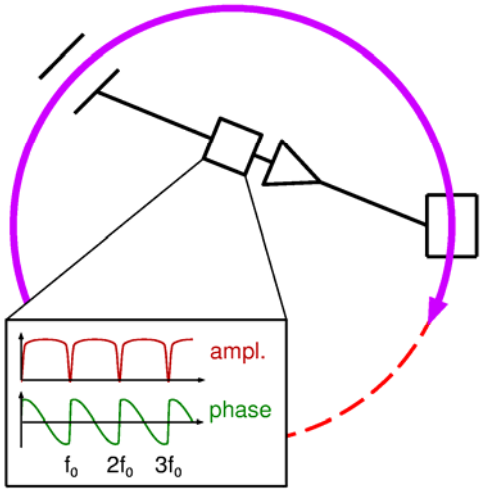

Fig. 11: Principle of longitudinal stochastic cooling using a longitudinal pickup and notch filters to detect samples that have a too low or too high momentum. A longitudinal kicker corrects the momentum error.

4.3 Noise and signals

We are not yet able to calculate the cooling time in seconds, because we have not analysed what kind of signal-to-noise ratio one can expect. Such an analysis is more a question of diagnostics (see, for example, Ref. [10]) than beam cooling, and the impatient reader can skip immediately to the following section for the conclusion. Nevertheless, we can illustrate the principle here by considering the simplest possible detector for a non-relativistic beam in a small accelerator, i.e., an electrostatic pickup. Such a pickup is not sufficiently sensitive to be used for stochastic cooling in reality, but it allows us to calculate a signal-to-noise ratio from first principles, and we will be able to draw conclusions of general validity.

The noise, to begin with, from a good charge-sensitive amplifier with GaAs FETs has a power density (reduced to the input of the amplifier) in the order of $s_{\text{noise}} = 1$ nV^2/Hz, independent of

frequency. This noise is the voltage noise from the FETs themselves. (The actual measured power, in watts, is obtained after multiplying with the amplification factor and the resistance over which the voltage is measured, such as 50 ohm, and integrating over frequency.)

As for the signal from a capacitive pickup, the beam induces a voltage on the pickup plates, and the voltage is proportional to the current passing through the pickup – the Schottky current. If we first look at the longitudinal current, we can let each passing particle be represented by a delta function in time. Multiplying with the ion charge Zq, summing over all turns made by the particle, from plus to minus infinity, and also summing over all N particles, we find that the current is [11]

$$I_{\parallel}(t) = Zq\sum_{a=1}^{N}\sum_{n=-\infty}^{\infty}\delta\left(t-\frac{\theta_{0,a}}{\omega_a}-\frac{2\pi n}{\omega_a}\right)\approx\frac{Zq\omega_0}{2\pi}\sum_{a=1}^{N}\sum_{n=-\infty}^{\infty}\exp\left(jn(\omega_a t-\theta_{0,a})\right). \tag{24}$$

Here, in addition, ω_a is the revolution frequency of particle a and $\theta_{0,a}$ is its position (phase) relative to the pickup at time zero. We assume that all particles have approximately the same revolution frequency ω_0 and we have used the fact that a sum of delta functions can be rewritten as a sum of exponentials. As expected, the unit for this current is amperes.

An electrostatic pickup is simple in that the voltage induced on the pickup electrodes can be written just as $U = Q/C$, where Q is the charge on the electrodes and C is the pickup's capacity to earth (typically in the order of 100 pF). The charge from a single particle is approximately the current it represents multiplied by the time it spends inside the pickup, so for a longitudinal pickup the voltage signal is

$$U(t) = \frac{Q(t)}{C} = \frac{I_{\parallel}(t)l}{vC}, \tag{25}$$

where l is the length of the pickup and v the particle velocity. (This applies as long as one looks at frequencies $U(t)$ that are low compared to the inverse of the time it takes for the particles to pass through the pickup, and this is one of the reasons why capacitive pickups are not used for stochastic cooling of relativistic beams.)

If we want to look at transverse cooling, or longitudinal cooling using the Palmer method, we need transverse pickups rather than longitudinal ones. A transverse electrostatic pickup detects a signal which is proportional to the current above but also to the transverse displacement of the particle relative to the centre of the pickup. Having transverse cooling in mind, we therefore define the transverse Schottky current as

$$\begin{aligned} I_{\perp}(t) &= Zq\sum_{a=1}^{N}A_a\cos(Q_a\omega_a t-\phi_a)\sum_{n=-\infty}^{\infty}\delta\left(t-\frac{\theta_{0,a}}{\omega_a}-\frac{2\pi n}{\omega_a}\right) \\ &\approx\frac{Zq\omega_0}{2\pi}\sum_{a=1}^{N}A_a\cos(Q_a\omega_a t-\phi_a)\sum_{n=-\infty}^{\infty}\exp\left(jn(\omega_a t-\theta_{0,a})\right) \end{aligned}. \tag{26}$$

Here, A_a is the amplitude of the betatron oscillations of particle a, Q_a its tune and φ_a the phase of its betatron oscillation. The unit for the transverse current is amperes times metres.

The pickup now measures a difference signal between opposite pickup plates, and we can assume that, if the distance between the plates is $2d$ and a specific particle has a displacement x, it induces a charge difference between the plates that is x/d times the summed charge. Thus, we have

$$U(t) = \frac{\Delta Q(t)}{C} = \frac{I_{\perp}(t)l}{dvC} \tag{27}$$

for the transverse pickup signal.

Combining equations (26) and (27), we have the voltage from the pickup, but we need to express the signal in terms of a power density in order to compare it with the noise power density. The power density $s_{\text{sign}}(\omega)$ is defined as the Fourier transform of the autocorrelation function of $U(t)$ where the latter is defined as

$$R(\tau) = \int U(t)U^*(t-\tau)\mathrm{d}t\,. \tag{28}$$

This integral is quite simple to perform if we assume that the particles move independently, so that cross terms in (28) with different a become zero. Also, signals from different Schottky bands are independent, so cross terms with different n also become zero. What remains after a Fourier transformation is then

$$s_{\text{sign}}(\omega) = \frac{l^2}{d^2v^2C^2}\frac{(Zq)^2\omega_0^2}{8\pi}\sum_{a=1}^{N}A_a^2\sum_{n=-\infty}^{\infty}\left(\delta[\omega-(n+Q_a)\omega_a]+\delta[\omega-(n-Q_a)\omega_a]\right). \tag{29}$$

This is the transverse Schottky signal as seen on, e.g., a spectrum analyzer, and it has the form of sidebands separated by $\pm Q_a\omega_a$ from the harmonics of the revolution frequency. If we assume that all particles have the same tune Q and that betatron amplitudes A_a and revolution frequencies ω_a are uncorrelated, we can approximate the sum $\sum...$ over a with an integral $N\int\int...F_A(A')f_0(\omega')\mathrm{d}A'\mathrm{d}\omega'$, where $f_0(\omega')$ is the distribution of revolution frequencies and $f_A(A')$ is the distribution of betatron oscillation amplitudes. This gives

$$s_{\text{sign}}(\omega) = \frac{l^2}{d^2v^2C^2}\frac{N(Zq)^2\omega_0^2}{8\pi}\langle A^2\rangle\sum_{n=-\infty}^{\infty}\left[\frac{1}{n+Q}f_0\left(\frac{\omega}{n+Q}\right)+\frac{1}{n-Q}f_0\left(\frac{\omega}{n-Q}\right)\right], \tag{30}$$

where $\langle A^2\rangle$ is the r.m.s. value of the oscillation amplitudes. Integrating over a sideband and multiplying by two because of the two sidebands at each n gives

$$S_{\text{sign}} = \frac{\pi N(Zq)^2\lambda^2\langle A^2\rangle}{d^2C^2}\,. \tag{31}$$

We have here introduced λ as the ratio between pickup length and ring circumference.

The noise power integrated over the same frequency interval between two harmonics of the revolution frequency is

$$S_{\text{noise}} = s_{\text{noise}}\,\omega_0/(2\pi)\,, \tag{32}$$

so we can now write down the noise to signal ratio U required in equation (23). It is

$$U = \frac{S_{\text{noise}}}{S_{\text{sign}}} = \frac{s_{\text{noise}}\omega_0 d^2C^2}{2\pi^2N(Zq)^2\lambda^2\langle A^2\rangle}. \tag{33}$$

The cooling rate is then

$$\tau = \frac{N}{W}(M+U) = \frac{N\omega_0}{4\pi W^2\eta\Delta p/p} + \frac{s_{\text{noise}}\omega_0 d^2C^2}{2\pi^2W(Zq)^2\lambda^2\langle A^2\rangle}. \tag{34}$$

Note that the mixing factor M cannot be smaller than one, so the first term cannot be smaller than N/W. With other types of pickups which are used in real stochastic-cooling systems, the exact appearance of the second term will be different. Some general conclusions can nevertheless be drawn from expression (34).

4.4 Conclusions

First, one can see that the cooling time is proportional to the particle number N, as in the simple theory we started with, as long as N is not too small. This is in contrast to electron cooling where the cooling time is independent of the beam intensity. For small N, however, the cooling time becomes independent of particle number also with stochastic cooling since noise then becomes important, and the second term in expression (34) dominates.

Secondly, hot beams cool faster stochastically than cold beams. Cold beams give smaller signals (in our example because the betatron amplitude $\langle A^2 \rangle$ becomes smaller), and noise again starts to become a limiting factor. This is opposite to electron cooling where hot beams cool slower because of the $1/v^2$ decrease of the cooling force. Also, when $\Delta p/p$ is small, mixing can become less efficient (M increases).

It is also seen that a high bandwidth is desirable. This is difficult to achieve for slow beams in a small accelerator. Clearly, a fast particle induces a shorter pulse than a slow particle when it goes through a pickup. Not only does it spend a shorter time passing the pickup, but if it is relativistic, the charge distribution becomes Lorentz contracted along the direction of motion. Once more, this is different from electron cooling, where we have seen that relativistic beams cool more slowly, and, in addition, technology becomes much more complicated for high energies.

A similarity, finally, between stochastic cooling and electron cooling is that highly charged ions cool better in both cases: in electron cooling because the Coulomb interaction gets stronger and in stochastic cooling because the signal strength increases.

As a numerical example of the cooling time at a rather low energy we can take a 10 MeV, 100μA proton beam in a synchrotron similar to CRYRING. Using $N = 7 \times 10^8$, $\omega_0 = 5$ MHz, $W = 300$ MHz, $\eta = 1$, $\Delta p/p = 1 \times 10^{-3}$, $s_{\text{noise}} = 1 \times 10^{-18}$ V^2/Hz, $d^2/\langle A^2 \rangle = 25$, $C = 100$ pF and $\lambda = 2 \times 10^{-3}$, our expression gives a cooling time of 2000 seconds, completely dominated by the second term in equation (34). This beam can be electron cooled in a second, so it is clear that stochastic cooling is not an optimal choice for such a slow beam even if the noise can be reduced by cooling the amplifiers and the pickup sensitivity could be increased.

On the other hand, stochastic cooling at the antiproton accumulators at CERN (now decommissioned and replaced by the AD) or Fermilab is made at GeV energies. Here, M is close to 1, and the signal is strong when large antiproton currents have been accumulated, so that the cooling time differs from N/W with a factor less than 10. Also, the bandwidth is high – up to 8 GHz. Also at the two other rings, COSY at Jülich and ESR at GSI, where stochastic cooling is performed today, bandwidths go up to 2 or 3 GHz. In these cases, electron cooling is slow and/or difficult because of hot beams and high energies.

5 Laser cooling

In laser cooling, a velocity dependent force is obtained when ions in a storage ring are excited by laser photons and thus absorb the momentum of the photons. Due to the Doppler shift, only ions within a certain velocity range are affected by the photons. The process is illustrated in Fig. 12: A laser beam is shining along the direction of motion of the ions (or opposite to it), and if the ion has the right velocity, so that the Doppler-shifted laser frequency matches a line in the atomic spectrum, photons can be absorbed. A photon with a (Doppler-shifted) frequency ν carries a momentum $h\nu/c$, which is gained by the ion at the absorption. At a later time, the photon is re-emitted, but this emission is isotropic, so the ion gets no average momentum change due to the emission. Now, the ion has returned to its ground state, but its net momentum has increased by $h\nu/c$ on the average, and it is ready to absorb a new photon. This is thus a way to accelerate ions with a velocity-dependent force.

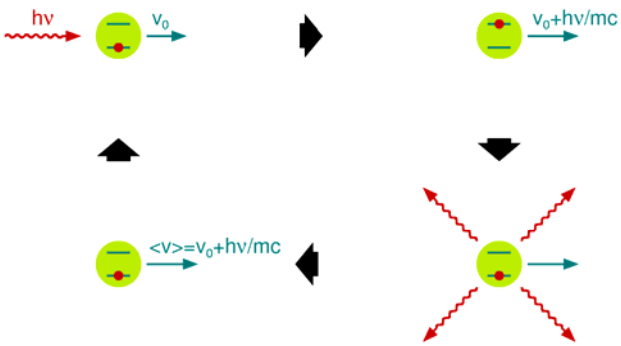

Fig. 12: Principle of laser cooling. An ion is accelerated by absorbing photons and the momentum they carry from one direction only. The re-emission is isotropic and gives no acceleration on the average.

For this process to work efficiently, several conditions have to be fulfilled. First, the atomic transition has to have a (Doppler-shifted) wavelength that can be reached with existing cw lasers of reasonable power. Secondly, it must be a closed transition, meaning that the ion has to decay back to its original state. If instead the excited ion would decay to a third metastable state with another energy which is not matching the laser frequency, and the ion would be trapped in this state, the cooling would stop. Also, it must be a transition with a large oscillator strength, such that the absorption coefficient is large. Then, the decay back to the initial state is also fast, and many absorption–emission cycles can take place in a given time, resulting in a large momentum transfer per unit time. For these reasons, laser cooling has only been applied to a small number of ion species, namely $^{7}Li^{+}$, $^{9}Be^{+}$, $^{24}Mg^{+}$ and, recently, C^{3+}.

The scheme described above can be used to change the ion velocity, but for real cooling a few more ingredients are needed. Normally, the laser light has a bandwidth that is much smaller than the width of the Doppler-broadened atomic transition, so something must be done for the laser to interact with all ions. Also, it is not enough just to accelerate the ions, but in order to reach an equilibrium situation one needs to add a counteracting force, such that one can have the total force F equal to zero at some velocity where also $\mathrm{d}F/\mathrm{d}v$ is negative. Then, the ions will be pulled toward this velocity, and one can have a cold equilibrium also in the presence of heating processes like intra-beam scattering. A few different methods to tackle these issues have been used in experiments with laser cooling.

One way to interact with all ions is to sweep the laser frequency over the momentum distribution of the ion beam, as illustrated in Fig. 13. This requires a tunable laser and, of course, adds to the complexity of the setup. The counteracting force is produced by a fixed-frequency laser shining from the opposite direction.

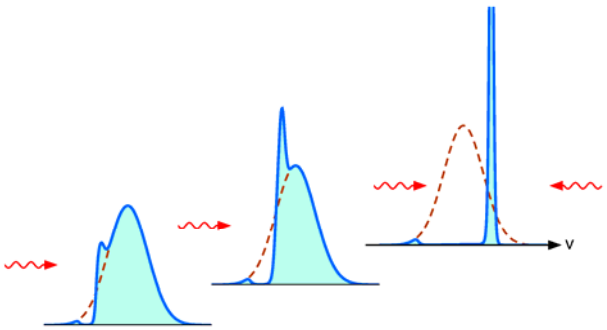

Fig. 13: The laser frequency width is usually smaller than the Doppler width of the beam, so sweeping the laser frequency across the Doppler profile will reduce the momentum spread of the beam. A counteracting laser (or another counteracting force) is needed to get to a cold equilibrium.

An alternative to a tunable laser is to use a fixed-frequency laser but instead sweep the ion velocity using, for example, an induction accelerator. With this approach, one automatically has forces from two directions simultaneously, giving the possibility for a cold equilibrium without a second laser. An induction accelerator is like a transformer where the ion beam constitutes the secondary winding, and the beam is thus accelerated as long as the current in the primary winding and the magnetic field in the iron core increase. The acceleration, and the cooling, can thus proceed until the core saturates.

A third alternative is to laser cool a bunched ion beam and use the rf as the counteracting force. Then the ions will perform synchrotron oscillations, with the hottest ions having the largest amplitude in phase and velocity. The initial laser frequency has to be set to match the largest velocity deviation, and as the beam cools, the laser frequency is scanned to match ions more and more in the centre of the bucket where they have the smallest velocity deviation.

Laser cooling is quite strong. An example, taken from Ref [12], is a $^7Li^+$ beam of 100 keV energy. The cooling transition between the metastable 1s2s ^{3}S state and the more highly excited 1s2p ^{3}P state has a wavelength of 5485 Å. When a 100 keV Li^+ ion absorbs the momentum of a photon of this wavelength, it gains 12 meV of energy. The cooling rate depends on the laser power, but if the laser power is increased too much, stimulated emission starts to dominate over spontaneous emission. Stimulated emission does not give cooling, since the emitted photon then is sent out in the same direction as the incoming one had, and not isotropically as in Fig. 12. The lifetime of the upper ^{3}P state is 43 ns, so if the laser power is chosen such that spontaneous emission is equal to stimulated emission, 1.2×10^7 cycles of absorption and spontaneous emission can take place per second. During the time it takes for an ion to pass the, say, 2 m long section where the ion and laser beams overlap, about 15 absorption–emission cycles will take place, changing the ion energy by 180 meV. For any reasonable energy spread in the beam, it is clear that the cooling time is equal to quite a small number of revolution periods. Also, the final beam temperature is quite low, given by the recoil of 12 meV from the emission of a single photon.

Laser cooling is not in routine use in any storage ring at present. The reason is mainly that only a small number of ion species have suitable optical transitions. A second reason is that transverse cooling is difficult – direct transverse laser cooling has not yet been demonstrated although several schemes have been suggested. Like with electron cooling, however, there are proposals for laser cooling of highly relativistic ions. Using the ensuing large Doppler shifts and a laser shining in the direction opposite to the beam, cooling can be performed using, e.g., the $n\mathrm{S}_{1/2} - n\mathrm{P}_{1/2}$ or $n\mathrm{S}_{1/2} - n\mathrm{P}_{3/2}$ transitions in lithium-like, sodium-like, etc., heavy ions. An advantage in laser cooling of relativistic ions, with high γ factors, is that the ions experience a stronger cooling force since the momentum of the absorbed photons increases with a factor γ. Also, the re-emitted photons are no longer emitted isotropically in the laboratory frame, but preferentially in the forward direction, which improves the cooling.

Laser cooling experiments have been performed at TSR in Heidelberg, ASTRID in Århus and ESR in Darmstadt.

6 Ionization cooling

Ionization cooling is also referred to as muon cooling, because it has been proposed as a method to cool muon beams for neutrino factories or muon colliders. Since these are certainly not small accelerators, we mention this method briefly, just for completeness. The principle is most straightforward for transverse cooling, and it is illustrated in Fig. 14.

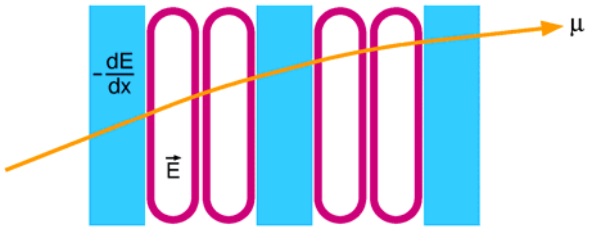

Fig. 14: In ionization cooling, particles (muons) lose momentum along their direction of motion through ionization of matter (square blocks in the figure). Longitudinal momentum is restored in acceleration cavities (ovals in the figure), leading to a net reduction of transverse but not longitudinal momentum.

The muon beam, with a typical momentum of 100 MeV/*c*, traverses a set of energy absorbers and acceleration cavities. In the absorbers, which should be of a low-*Z* material to reduce the transverse scattering that accompanies the energy loss, both longitudinal and transverse momentum is reduced. In the acceleration cavities, only longitudinal momentum is restored, and the result is a net reduction of transverse momentum. Ionization cooling is not suited for electrons which are too light and produce bremsstrahlung, nor is it suitable for ions which would make nuclear reactions, but for muons it has the advantage of being, potentially, very fast, which is important since the muon's lifetime is just 2.2 μs in the rest frame.

A setup to test muon cooling, MICE (Muon Ionization Cooling Experiment), is at present being installed at Rutherford Appleton Laboratories, and a proof that this cooling method works may come during 2006 or 2007.

References

[1] See, e.g., J. D. Lawson, *The Physics of Charged-Particle Beams*, 2nd ed. (Clarendon, Oxford, 1988).

[2] G. I. Budker, Proc. Int. Symp. on Electron and Positron Storage Rings, Saclay, 1996, eds. H. Zyngier and E. Crémieu-Alcan, p. II-1-1.

[3] L. Spitzer, Jr., Mon. Not. R. Astron. Soc. **100** (1940) 396.

[4] H. Goldstein, *Classical Mechanics*, 2nd ed. (Addison-Wesley, Reading, MA, 1980).

[5] H. Danared, Phys. Scr. **T59** (1995) 121.

[6] S. Mannervik *et al*., Phys. Rev. Lett. **81** (1998) 313.

[7] S. van der Meer, CERN/ISR-PO/72-31 (1972).

[8] D. Möhl, CERN Accelerator School, CERN 95-06 (1995), p. 587.

[9] D. Boussard, *ibid.*, p. 749.

[10] See, e.g., contribution by U. Raich to these proceedings.

[11] S. Chattopadhyay, CERN 84-11 (1984).

[12] S. P. Møller, CERN Accelerator School, CERN 94-01 (1994), p. 601.