Fermi National Accelerator Laboratory

FERMILAB-Conf-93/308

Theory, Technology and Technique of Stochastic Cooling

John Marriner

Fermi National Accelerator Laboratory
P.O. Box 500, Batavia, Illinois 60510

October 1993

Submitted to the *Workshop on Beam Cooling and Related Topics*, Montreux, Switzerland, October 3-8, 1993

Disclaimer

THEORY, TECHNOLOGY, AND TECHNIQUE OF STOCHASTIC COOLING

John Marriner
Fermilab* , P.O. Box 500, Batavia, IL 60510

ABSTRACT
The theory and technological implementation of stochastic cooling is described. Theoretical and technological limitations are discussed. Data from existing stochastic cooling systems are shown to illustrate some useful techniques.

1. THEORY

1.1 Introduction

The theory of stochastic cooling has been discussed by a number of authors [1,2,3,4,5,6,7]. I will try to describe the theory and to make the results seem plausible, but I will not give a detailed derivation of the results.

A prototypical transverse stochastic cooling system is shown in Fig. 1. The system senses the particle position at the pickup by measuring the difference in induced current between the two pickup electrodes. The electronic signal is amplified and applied to the kicker electrode when the particle passes between the electrodes. Transverse electric and magnetic fields in the kicker deflect the particle. The angular deflection at the kicker will decrease the amplitude of the betatron oscillations provided that it has the correct sign.

1.2 Schottky Signals

An understanding of Schottky signals is central to the understanding of stochastic cooling. A particle passing a point in the ring with revolution frequency makes a current

$$i(t) = e\sum_{n=-\infty}^{\infty} \delta\left(t+\tau-\frac{n}{f_0}\right)$$
$$= ef_0\left[1+2\sum_{n=1}^{\infty}\cos n\omega_0(t+\tau)\right]. \qquad (1)$$

The current consists of an infinite series of lines at multiples of the revolution frequency. If we consider a beam of particles with random values of τ then the ac current will be nearly zero because of the random phase factor in Eq. (1). The rms current is not zero and is known as the Schottky current. In a beam the revolution frequencies are similar but not identical for all particles. Thus, the Schottky currents organize themselves into bands around the average revolution frequency. The rms current per Schottky band for a beam of N particles is

$$P_t = 2Ne^2 f_0^2. \qquad (2)$$

The betatron motion of a particle in the beam at the observation point is given by

$$x(t) = A\cos(\omega_0 Qt+\phi). \qquad (3)$$

The dipole moment of the beam is the current times the displacement and can be written in terms of sine waves as

$$d(t) = ef_0 A\left[\sum_{n+Q>0}\cos(n+Q)\omega_0(t+\tau) + \sum_{n-Q>0}\cos(n-Q)\omega_0(t+\tau)\right]. \qquad (4)$$

There are 2 Schottky sidebands per harmonic of the revolution frequency. Each band has a rms amplitude

* Operated by the Universities Research Assoc. under contract no. DE-AC02-76CHO3000 with the US Dept. of Energy

$$P_t = \frac{1}{2} N e^2 A^2 f_0^2. \tag{5}$$

Schottky signals are often observed with a pair of strip-line pickups. The beam current is observed when the top and bottom pickups are added. The dipole moment is obtained when the two pickups are subtracted. An example of the observation of the dipole signal is shown in Fig. 2.

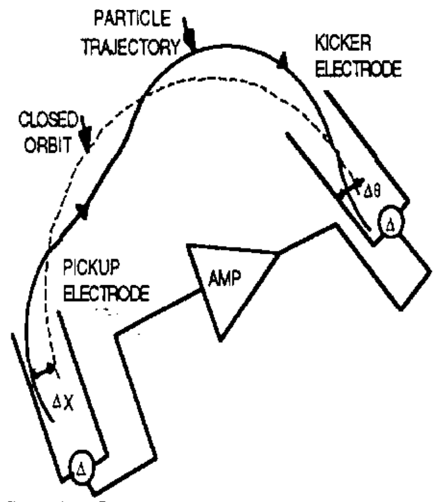

Fig. 1. Schematic of a typical stochastic cooling system. The particle position is sensed at the pickup and a corresponding deflection is applied at the kicker.

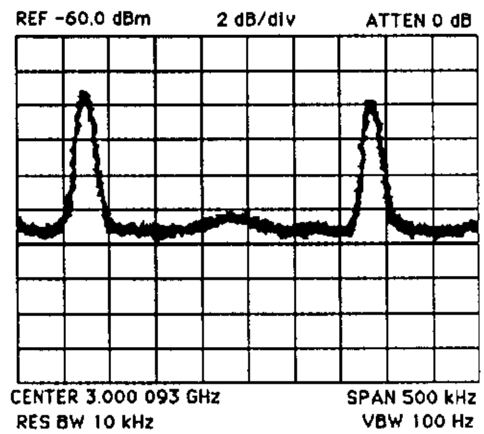

Fig. 2. The Schottky signal from the two betatron sidebands in the FNAL Debuncher ring is shown. The pickup is most sensitive to the dipole moment but has some direct sensitivity to the current which produces a small enhancement at harmonics of the revolution frequency as seen near the center of the trace.

1.2 Action of the Feedback

We consider the action of the feedback system in Fig. 1 on a coasting beam. A more schematic representation is shown in Fig. 3. The dipole moment will be the sum of the dipole moment $d_i(\omega)$ in the absence of the feedback system and a dipole moment $d_f(\omega)$ induced by the excitation applied to the kicker. If this excitation is not too large, then the dipole moment will depend linearly on the deflection.

$$d_f(\omega) = F(\omega)\theta(\omega). \tag{6}$$

We further assume that the angular deflection depends linearly on the dipole moment of the beam:

$$\theta(\omega) = G(\omega)\left[d_i(\omega) + d_f(\omega)\right]. \tag{7}$$

Using the circuit shown in Fig. 3, one can infer that

$$\begin{aligned}\theta(\omega) &= \frac{G(\omega)d_i(\omega)}{1 - F(\omega)G(\omega)} \\ &= H(\omega)d_i(\omega).\end{aligned} \tag{8}$$

Fig. 3 and this formalism are intended to demonstrate that stochastic cooling is formally identical to purely electronic feedback circuits. Thus, one can apply the techniques used in dealing with these more commonplace feedback circuits directly to stochastic cooling systems. Network analyzers can be effectively employed to measure the denominator in Eq. (8). The denominator, in fact, is of considerably more importance than the numerator. Not only does it determine system stability according to the criterion of Nyquist [8], it also provides a measure of the cooling rate relative to the maximum rate that can be obtained.

In order to understand the relationship between the denominator of Eq. (8) and the cooling effect, consider Fig. 4. A segment of a beam is shown in coordinate space. Suppose that a random clumping of particles gives the beam a net dipole moment near the center of the beam. The action of the stochastic cooling system serves to reduce the dipole moment by displacing the centroid of the beam by making the "kink" shown in Fig. 4. This effect of the cooling system is known as "signal suppression." When signal suppression is present, the motion of the beam centroid produces a signal which is opposite in sign to the Schottky signal. If the cooling system is turned off, the Schottky signal will return to its former level.

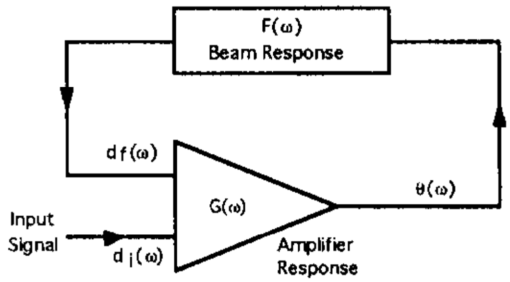

Fig. 3. Equivalent circuit of the beam feedback. The input signal d_i is modified by the addition of the feedback d_f.

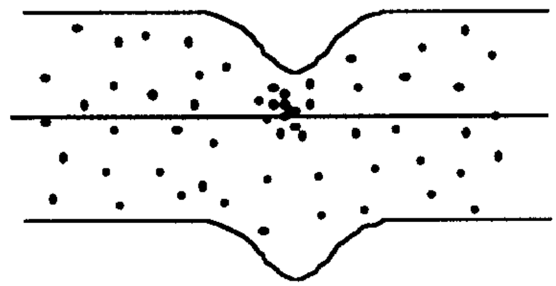

Fig. 4. A cartoon of the behavior of a fluctuation in the dipole moment of the beam under the influence of a stochastic cooling system.

The kink in the beam reduces the amplitudes of the particles that constitute the clump of beam. These particles have slightly different revolution frequencies, so the clump will dissipate as time elapses. After the clump dissipates these particles will continue to have their reduced amplitudes – even if the cooling system is turned off. Note that some particles have their amplitudes increased in this process. However, the net effect is beam cooling since more particles are in the clump of beam than outside it. The length of time that it takes for fluctuations to disappear is defined to be M times the revolution period, where M is known as the mixing factor. Large values of M reduce the maximum cooling rate. The power of the cooling system to resolve two nearby fluctuations is approximately $1/W$, where W is the cooling system bandwidth.

1.3 Betatron Cooling

In order to quantitatively connect the signal suppression phenomenon and the cooling rate, consider the equations of motion. The emittance is a constant of motion in the absence of the cooling system. Under the action of the cooling system the change in amplitude of the i^{th} particle is:

$$\begin{aligned}\Delta\left(A_i^2\right) &= \frac{\beta}{2}\left[\left(\theta_i+\Delta\theta\right)^2-\theta_i^2\right] \\ &= \beta\theta_i\Delta\theta+\frac{\beta}{2}\Delta\theta^2, \end{aligned} \tag{9}$$

where β is the beta function at the kicker, θ_i. is the angle of the i^{th} particle at the kicker, and $\Delta\theta$ is the angular deflection at the kicker. The deflection at the kicker depends linearly on the signals the various particles created at the pickup. Therefore one can write the deflection as a sum over particles

$$\Delta\theta = \sum_{j=1}^{N}\Delta\theta_j, \tag{10}$$

where $\Delta\theta_i$ is contribution of the motion of the i^{th} particle to the total deflection. We can use Eqs. (9) and (10) to compute the change in emittance as follows:

$$\Delta\varepsilon = \frac{1}{N}\sum_{i=1}^{N}\Delta\left(A_i^2\right) = \frac{\beta}{2N}\sum_{i=1}^{N}\left(2\theta_i\Delta\theta_i + N\Delta\theta_i^2\right). \tag{11}$$

In writing Eq. (11) it has been assumed that the particle motions are random, so that the $i \neq j$ terms do not *on the average* contribute to $\Delta\varepsilon$ and the double sum collapses to a single sum.

An evaluation of $\Delta\theta$ and a more detailed account of some other significant points are given in reference 7. The result is that the cooling rate is given by

$$\frac{1}{\varepsilon}\frac{d\varepsilon}{dt}=\frac{1}{4\pi N}\sum_{i=1}^{N}\omega_{0i}\sum_{n=-\infty}^{\infty}\left\{-2\,\mathrm{Re}\left[jH((n+Q)\omega_{0i})e^{j\mu_k}e^{-j(n+Q)\omega_{0i}t_{ki}}\right]\right.$$
$$\left.+\left[M((n+Q)\omega_{0i})+U((n+Q)\omega_{0i})\right]\left|H((n+Q)\omega_{0i})\right|^2\right\}, \tag{12}$$

where Q is the betatron tune, μ_k is the phase advance between pickup and kicker, ω_{0i} is the angular revolution frequency, and t_{ki} is the transit time delay between pickup and kicker. The first term on the right hand side of Eq. (12) is the cooling term. The cooling rate is maximized when $\mu_k=\pi/2$ and $H(\omega)$ is real except for a phase factor $e^{j(n+Q)\omega_{0i}t_{ki}}$. The phase factor corresponds to a delay equal to the particle pickup to kicker delay.

The mixing factor M is the ratio of the Schottky power density of particles with the same revolution frequency as the i^{th} particle to the average Schottky power. Since no particles have *exactly* the same revolution frequency as the i^{th} particle, the calculation of M must be understood as a limit. If the cooling is applied for a time T, the number of particles that have a revolution frequency that is indistinguishable from the i^{th} particle is proportional to $1/T$. The growth in amplitude squared from the remaining particles, however, is proportional to T^2. Thus the net effect on the i^{th} particle is a growth in amplitude squared that is linear with time.

The factor $U(\omega)$ is the ratio of thermal noise power to the average Schottky power. It enters into the equations in exactly the same way as the Schottky power. The only difference between the Schottky noise and thermal noise is that the system designer has some control over the value of the thermal noise. In fact, in low density beams the major design challenge may be to achieve an acceptable signal to noise ratio.

Since the cooling term in Eq. (12) is proportional to $H(\omega)$ and the heating term is proportional to $|H(\omega)|^2$, there is a value of H that achieves the highest cooling rate. This value is known as the optimum gain. Gain values which exceed the optimum gain will result in a reduced cooling rate.

Equation (12) is the exact expression for cooling in the frequency domain. The cooling of particles with different frequencies will vary because of the differences in the phase factors in the cooling term and the mixing factor M. The cooling system is usually phased for the center of the beam distribution – where M usually has a peak. The heating term decreases as the frequency departs from the center, but the phase error of the cooling term increases so the cooling rate doesn't vary too much. It is conventional to estimate the cooling rate by using the cooling value at the peak value of M. If H is at the optimum gain, the formula for the cooling rate becomes:

$$\frac{1}{\varepsilon}\frac{d\varepsilon}{dt}=\frac{W}{N(\overline{M}+\overline{U})}. \tag{13}$$

where $W=f_{max}-f_{min}$ is the cooling system bandwidth and $\overline{M}$ and $\overline{U}$ are the peak values of M and U.

1.4 Momentum Cooling

Momentum cooling is similar to betatron cooling except that a technique is required to develop a signal that is proportional to the momentum of the particle. The simplest technique is to place a difference pickup in a region of non-zero dispersion. A pickup that measures the dipole moment of the particles in a region of dispersion will produce a Schottky voltage proportional to the momentum fluctuations in the beam. This voltage can be made to accelerate or decelerate the beam in a kicker and therefore cool the momentum spread. This cooling method is sometimes called the Palmer method. A second technique is to use a notch filter. The filter response changes sign depending on whether the particle revolution frequency is above or below the desired frequency. This method of cooling [9] is sometimes called the Thorndahl method.

The two methods have disadvantages and advantages. The pickup placed in the region of dispersion will have poor signal to noise ratio (nominally 0) at the center of the pickup. The notch filter method avoids the low signal to noise ratio by filtering the noise as well as the signal, and the signal to noise ratio follows the particle density throughout the notch. The filter method can only be used if the revolution frequency versus momentum relationship is unique (non-overlapping Schottky bands). The filter introduces undesirable phase characteristics that reduce the cooling rate. Thus, the filter method is used in situations where the signal to noise ratio is critical; otherwise the pickup in dispersion is used.

Momentum cooling is conventionally described by a Fokker-Planck equation. The beam is described by a distribution $N(E,t)$, which is the number of particles having energy greater than E at time t. The density Ψ is given by

$$\Psi(E,t) = \frac{\partial N(E,t)}{\partial E}, \tag{14}$$

and the flux is given by

$$\Phi(E,t) = \frac{\partial N(E,t)}{\partial t}. \tag{15}$$

The flux is related to the beam distribution by the Fokker-Planck equation

$$\Phi = F\Psi + \left(D_0 + D_1 + D_2\Psi\right)\frac{\partial\Psi}{\partial E}. \tag{16}$$

The term proportional to $F = F(E)$ is the cooling term. $F(E)$ is proportional to the cooling system gain G. The term D_1 is a heating term that is proportional to thermal noise in the system. The term proportional to D_2 is the Schottky heating term. The Schottky heating is proportional to $\Psi(E,t)$ the number of particles at that energy (or equivalently that revolution frequency). Both D_1 and D_2 are proportional to G^2. The term D_0 is used to describe external heating mechanisms such as intrabeam scattering that are independent of cooling system gain.

One can describe momentum cooling in terms of moments similar to the equation describing betatron cooling. However, moments tend to be less useful because of the non-linear nature of Eq. (16): as the cooling proceeds the density increases and the Schottky heating term increases. Similarly, one can write a Fokker-Planck equation to describe betatron cooling. However, most betatron cooling applications are fairly well described by gaussian distributions where the variation in the rms beam size completely describes the beam evolution. Momentum cooling, however, has been applied to antiproton momentum stacking, working with non-gaussian distributions. Eq. (16) can be solved analytically for a number of situations where $\Phi(E,t) = \Phi_0$ =constant. A solution for antiproton stacking is discussed by van der Meer [10].

1.5 Good and Bad Mixing

Perhaps the most fundamental limitation in achieving effective cooling comes from limitations in the mixing. The fastest cooling is achieved if the Schottky signal is completely randomized between successive passes through the pickup. The randomization is complete in the limit that $M = 1$. However, the Schottky signal should not change between pickup and kicker (M should be large). The randomization between successive passes through the pickup is sometimes called the "good mixing" and the randomization between pickup and kicker is sometimes called the "bad mixing". A reasonable rule of thumb is that M=3 for an optimized betatron cooling system.

Antiproton sources CERN AA and FNAL Accumulator were engineered to have a specific value of mixing by using appropriate magnetic lattice design techniques. The lattice parameter $\eta = 1/\gamma_t^2 - 1/\gamma^2$ specifies the spread in revolution periods. At a particular frequency in the cooling band the time spread in revolution frequencies leads to phase errors. As the frequency is increased, the phase error increases. Thus, attempts to upgrade the antiproton source cooling systems are constrained by the limitation imposed by the "bad mixing." At other accelerators it is a matter of luck if the lattice parameter η is appropriate for effective cooling at some frequency.

Possible solutions to the mixing problem are special lattices that have no mixing between pickup and kicker and yet have large mixing between kicker and pickup. Such a lattice was considered at least briefly as a possible design for the CERN Antiproton Collector, but I do not know of any accelerator built on this principle. Another possibility is to design a special pickup placed in a region of dispersion to cancel the transit time differences of the particles in the beam. The basic idea is that both the transit time difference and the position at the pickup are proportional to the momentum offset. If the signal at the pickup can be made to have a time delay that is proportional to the horizontal position of the particle, it may be possible to have the time delay difference in the pickup exactly compensate the pickup to kicker transit time difference.

2. TECHNOLOGY

2.1 Bandwidth considerations

Stochastic cooling systems have been built with bandwidths ranging from 100 MHz to 4 GHz. The highest frequency systems operate in the 4-8 GHz band. Systems typically are chosen to operate in a band where the upper frequency is twice the lower frequency (known as an octave bandwidth). It becomes extremely difficult for a variety of technological reasons to maintain uniform gain amplitude and phase over bands that exceed one octave.

From a technological point of view it appears possible to extend present stochastic cooling technology to frequencies in the 10's of GHz. Aside from the good mixing/bad mixing problem mentioned above some of the technical difficulties in extending the frequency range of stochastic cooling systems are discussed below.

2.2 Pickups and Kickers

Pickups are devices that convert the mechanical beam energy into electrical energy. Kickers perform work on the beam with the applied electrical energy. The same structure can act as either a pickup or a kicker. The reciprocity theorem states that any device that converts the mechanical energy of the beam into electrical energy can be used to convert electrical energy into mechanical energy with the same coupling. A more precise statement of this theorem and other theoretical and practical pickup considerations can be found in the article by Lambertson [11]. The major practical differences between pickup and kicker design is that pickups may be cooled to cryogenic temperatures to reduce the thermal noise level while kickers may be required to dissipate significant rf power.

2.3 Stripline Pickup Response Model

Although other types of pickups are possible [12,13], the most common pickup is the strip-line type indicated in Figs. 5 and 6. These figures are also intended to suggest an electrical model of the beam current and stripline. The beam current is given by:

$$i(z,t) = i_b e^{j(kz-\omega t)} \tag{17}$$

where $v = \omega / k$ is the beam velocity. A current equal in magnitude but opposite in sign (the image current) flows on the walls of the beam chamber. A fraction f of the image current flows onto the strip line at one end and off at the other end. The pickup response can be modeled as a transmission line with the beam acting as a current source at the two ends. The solution of these equations shows that the output voltage is given by

$$V_p = f Z_p i_b \sin k\ell \tag{18}$$

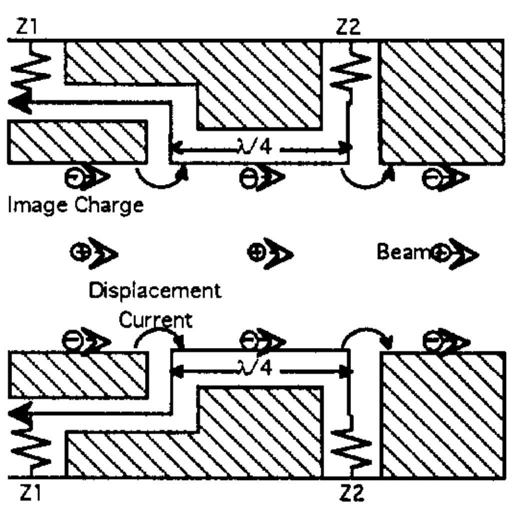

Fig. 5. Schematic representation of a quarter-wave strip-line electrode interacting with a high energy charged particle (side view).

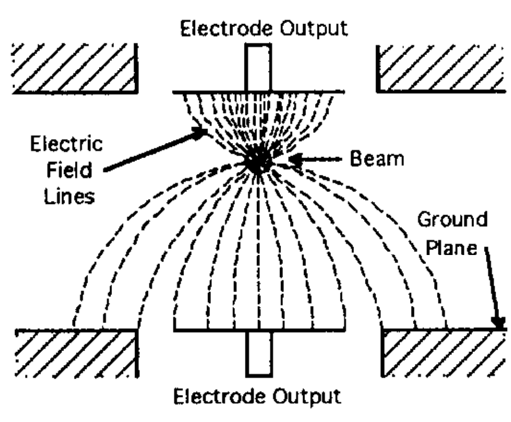

Fig. 6. The electric field lines from a single particle in the vicinity of a pickup electrode. More field lines terminate on the upper plate because the particle is closer to it than to the lower plate (cross-section).