IEEE Transactions on Nuclear Science, Vol. NS-30, No. 4, August 1983

STABILITY AND SIGNAL SUPPRESSION OF SCHOTTKY SIGNALS FROM STOCHASTICALLY COOLED BEAMS

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

Introduction

Several thorough treatments of stochastic cooling now exist[1,2] which go far beyond the original concepual picture[3] of "removing fluctuations from the Schott noise". The trend has been to replace this original statistical picture with ensemble evolution equations amenable to continuum treatment. High performance $\overline{P}$ accumulation systems, both designed and contemplated, attempt operation at the limits of system feedback gain, as dictated by instability thresholds and "signal suppession"[SS] degradation. The many admitted as well as implicit assumptions and approximations associated with the current models are most questionable near the high gain limit.

First I will highlight a few of these assumptions by pointing out inconsistancies they lead to. In particular, a fundamental difference between transverse (including "Palmer") cooling and filter cooling is revealed. This critique is not meant to be comprehensive, and thus does not predict new performance limitations for specific systems. Its point is to emphasize the need for further analysis, with more powerful theory.

Finally I point out a candidate for such a refined approach, the "renormalization group"[RG]. Instability thresholds may be viewed as phase stransitions (a fact pointed out, already, by Sacherer[4]). The unique success of the RG approach in describing behavior of systems near phase transitions is well known in other branches of Physics. In this limited format I will merely make this method plausible and sketch its method. The RG is essentially a statistical method, and my arguments in this note argue for a reexamination of a statistical approach to stochastic cooling.

Transverse vs Filter Cooling

This note is concerned with filter momentum cooling, but it will be instructive to compare with the simpler dynamics of transverse cooling systems. For an ideal (e.g. $\alpha_{p,PU-K}=0$ and exact $\pi/2$ PU→K phase advance), but realizable, betatron system with linear transverse PU and K sensitivity, it is easy to understand the phenomena of SS, and instability (heating, in this situation). This is because there is strictly no correlation between the signal (transverse fluctuations) and the particle longitudinal positions. Thus one can actually see, in the simulated dynamics of a few particle beam for $\eta=0$ (global dispersion), that particle betatron phases readjust themselves, canceling their neighbor's amplitudes and giving $V_\beta(t)=0$ after a few turns. Theoretically this is obvious since the dynamics for N particles is linear, in the sense that

$$\begin{pmatrix} X \\ X' \end{pmatrix}_{t+T_0} = M \begin{pmatrix} X \\ X' \end{pmatrix}_t \qquad (1)$$

where $X=(x_i)$ i=1,N is the transverse amplitude at t and M is not (X,X') dependent (but will be (p_i) dependent).

*Operated by Universities Research Association, Inc. under contract with the U.S. Department of Energy.

Such linear behavior is equivalent to that of signals on operational amplifier feedback circuits, which forms the paradigm for the stochastic cooling analysis of Mohl, et al.[1] I wish to contrast the actual behavior of filter cooling to this. First a discussion of Palmer cooling is useful as an intermediate case. Consider the idealized situation of α_p(PU→K) =0; betatron motion uneffected; and PU/K sensitivity independent of η. Two conditions are important (see Fig.1): A. a small momentum spread, σ_p, symmetrical about p_0 (the PU neutral plane); and B. σ_p offset by $\Delta p \gg \sigma_p$ from p_0.

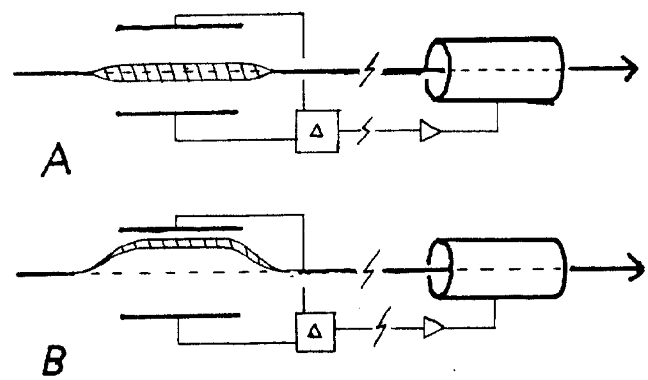

Fig. 1. A normal and B offset Palmer feedback.

In case A, $V_p(t)$ ("P" for Palmer) is exactly equal to the energy fluctuation record of the beam as a function of time (at PU). That is, if $\sigma\to0$, $V_p(t)\to0$: configuration A has zero sensitivity to longitudinal density fluctuations. In case B we have the opposite: as $\sigma_p\to0$, the signal does not diminish; $V_p(t)$ is eventually a record of density fluctuations. Henceforth A will be assumed (of course, a configuration B will eventually cool to A). Notice that B is relevant for momentum stack systems. They will be subject to contamination by such density fluctuation information, and are therefore (as will be made clear below), more akin to filter cooling systems. If $\eta\to0$ a configuration A must rapidly evolve to an exactly suppressed $V_p(t)\equiv0$, in analogy to the $V_\beta(t)$ discussion. If $\eta\to\infty$ there will still be normal cooling (and heating, past some gain threshold) but no suppression because energy-density cross correlations are absent.

Of course, for momentum cooling one is interested only in the energy fluctuations, so that one seeks a filter, for filter cooling, which would ideally convert a sum PU signal, $V_\Sigma(t)$, consisting entirely of density fluctuations, into $V_p(t)$. But this is impossible because $V_\Sigma(t+t')$ cannot be predicted from a knowledge of $V_p(t)$ (with $t<\tau_{cooling}$) alone. That is, a given energy fluctuation record can evolve into infinitely many different density fluctuation records. Now, in general, the filter cooling signal, $V_F(\omega)=V_\Sigma(\omega)\cdot F(\omega)$, where F is the filter response, may be expressed as,

$$V_F(t) = \int d\omega\, e^{i\omega t} V_\Sigma(\omega)\, F(\omega) = \int dt'\, V_\Sigma(t')\, F(t'-t) \qquad (2)$$

A unique correspondence to $V_p(t)$ can result only if $F(t'-t) = \delta(t'-t)$, in which case $V_F(t) = V_\Sigma(t)$.

There are elementary, practical consequences of this distinction. For instance, a Palmer system, with fixed bandwidth [W], may be constructed with $\eta \to 0$ to obtain optimum performance (no suppression). A similar optimum limit does not occur for filter cooling since correlations, as (2) indicates, are necessary to form a signal containing energy information. $V_F(t)$ cannot equal $V_p(t)$; therefore $V_F(t)$ inevitably carries a contamination of density fluctuation noise.

Even this rather obvious property is not manifest in existing treatments. I will now make reference directly to expressions in Ref.2 [BL...], which marks great progress in placing stochastic cooling on a firm theoretical basis. Expressions BL99a and BL99b for the $\varepsilon_{\pm 1}$ factors, appropriate to filter and Palmer cooling respectively, can clearly be made identical for arbitrary $G_1(\omega)$. Especially in the limit $\eta \to \infty$, it is necessary, quite independently of the theory of BLsect.4A,B&5, to introduce some constraint that $G_1(\omega) \to 0$. Ultimately the lack of derived restriction on $G_1(\omega)$ may be traced back to BL29:

$$\dot{P}_i = \sum_j G(q_i, q_j, P_j) \tag{3}$$

which is a sufficiently general form for the dynamics of a Palmer system, but represents an enormous simplification for filter cooling. In the sequel (BL30-50) it is assumed implicitly that (3) is local. That is,

$$\dot{P}_i(t) = \sum_j G\left(q_i(t), q_j(t), P_j(t)\right) \tag{4a}$$

As we have seen, nonlocality is an essential feature of filter cooling. Strictly one needs a from

$$\dot{P}_i(t) = \sum_j G\left(\int q_i(t')\, F(t'-t), \ldots\right) \tag{4b}$$

and a crude local approximation would be:

$$\begin{aligned}\dot{P}_i(t) &\approx \sum_j G\left(q_i(t), q_j(t), P_j(t), \kappa[P_i - P_j]\right) \\ &= \sum_j \tilde{G}\left(q_i(t), q_j(t), P_i(t), P_j(t)\right)\end{aligned} \tag{4c}$$

which is still different enough from BL29 to be incompatible with their development.

The kinetic equation approach (BL26-34) is certainly an adequate point of departure. Notice that $D(q,p)$, the full phase space distribution function, is a field on the phase space. In the filter coooing case one faces a nontrivially[5] non-local and non-linear field theory. Reducing the dynamics, eqs (4), to an effective form (3) can only result from a self consistent chain. An effective instantaneous $G(q_i, q_j, p_j)$ depends on the past history of the dynamics, the correlations and the mixing rate. That is , it is a function of the $\varepsilon_{\pm 1}$'s, and the $\varepsilon_{\pm 1}$'s are themselves determined by the $G(i,j)$'s.

Power Spectrum Sum Rules

Another characteristic constraint on $G_1(\omega)$ which is missing from non-self-consistent theories involves the integrated power spectrum:

$$E(W,T) \equiv \int_0^W P(\omega)\, d\omega \tag{5}$$

where $P(\omega) = |V_\Sigma(\omega)|^2/Z_0$, and T is a sufficiently long observation time (but $\ll \tau_{cooling}$). In order to examine the entire Schottky spectrum imagine an ideal Σ PU with ∞ bandwidth (but excluding zero frequency). However, the electronics (filter) following it has a sharp cutoff at W. Therefore $V_F(\omega)$, the influence on the beam, contains no frequency components > W. Suppose W_* is a fequency sufficiently high that each particle's passage is resolved. Then $E(W_*,T) \simeq NW_* T/T_0$ for N particles circulating with period T_0. If the loop is open $E(W,T) \simeq NW(T/T_0 \pm 1)$ for any W. Now, a characteristic of the existent treatments (BL64, and particularly explicit for the beam response presented by Van der Meer[6]) is that $\varepsilon(\omega > W_*) = 1$ (at least below instability thresholds). Thus $E(W_*,T) - E(W,T)$ does not change when the loop is closed, and therefore $E(W,T)$ cannot either.

In most treatments it is convenient to make a further assumption: that the distributions are q_i independent and dynamics depends only on differences $q_i - q_j$ (BL39-42). This simplification is usually referred to as a model of the "non-overlapping" Schottky bands situation. This leads to independent suppression over each band, thus allowing the tighter constraint,

$$\int_{BAND} P(\omega)\, d\omega \geq NT/T_0 \tag{6}$$

Physically, constraints (5) and (6) represent the fact that N point charges circulate past the Σ PU each turn, and that W limits the scale of time correlations possible in the rearrangement of their azimuthal spacings. The entire problem may be studied in terms of PU arrival time difference correlations by converting (5) or (6) via the Wiener-Kintchine theorem.

For Palmer momentum cooling (case A above) consider the spectrum from an auxiliary Σ PU observing the beam (but not feeding back). The dynamics are particularly simple:

$$\dot{P}_i \approx \sum_j H(q_i - q_j)\, \tilde{G}(P_j) \tag{7}$$

That is, the G(j)'s of (3) factorize, which leads to a band by band constraint (6). Factorization of G is evidently a powerful assumption. If the cooling term is neglected in the kinetic evolution equation for the single particle distribution function, $f_1(q,p)$ (BL33), an initially factorized f_1 evidently remains so in time. If f_1 is exactly factorizable the $V_\Sigma(\omega)$ signal, seeing only density fluctuation, will not change. The $\eta \to 0$ limit of Palmer cooling was an obvious example of this.

The one good piece of data immediately available to me relevant to (6) is in Fig 26 of Ref.1. A careful analysis of the filter cooled Schottky band pictured gives a closed to open loop trace area ratio of $\simeq 0.9$. Whether such a marginal suppression contradicts (6) can depend on several other considerations: 1. For filter cooling the "non overlapping" bands approximation is not exact ((5) is the proper statement). 2. The PU's or K's may have some transverse sensitivity bias. 3. Other non ideal realities of the experiment become significant, such as $\alpha_{p,PU-K} = 0$, or amplifier noise (which, after all, determines the shape of the asymptotic band illustrated!).

This same figure from Ref.1 illustrates another feature implicit in the above sum rule for filter cooling. As I pointed out for Palmer cooling, $V_p(t)$ is suppressed. Indeed the kicker signal must experience a net suppression when the loop is closed on any [well adjusted] cooling system. In other words, a system which cools all particles presents net negative feedback to energy fluctuations which constitute some optimal portion of the kicker signal content. But for filter cooling, V_Σ and V_F (kicker signal) are proportional. Thus a normally behaved notch filter circuit will induce a closed loop distortion of V_Σ that must have the form of a peaking at the notch (see Fig. 2).

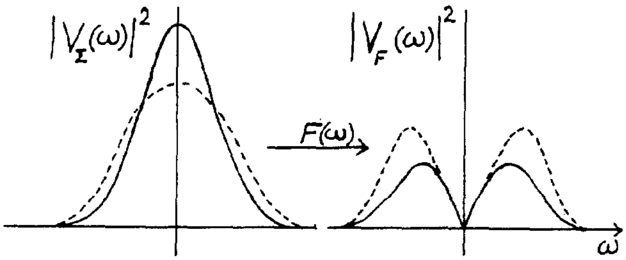

Fig. 2. Sum PU and Kicker signals (one band) with loop opened(-----) and closed(-----). Each Sum PU curve has same area.

The Renormalization Group

The RG method[7,8] provides a procedure for treating fundamentally difficult problems in statistical mechanics. Because of their difficulty and the RG's generality, one does not expect to present compact formulae (or even possibly complicated expressions!) representing a "solution" to the problems. In fact the real work of applying RG theory to a specific problem is in the numerical analysis, the theory serving only as a sort of inspiration for how to proceed, what sorts of behavior to watch for, etc. To gain a flavor of this context refer to the reviews of Wilson[7,9].

Also, as many approximations and assumptions as the existing treatments require, are necessary to arrive at a tractable numerical procedure. Despite these disadvantages the RG can provide deeper insight into high gain cooling limits because: 1. It retains an essential self-consistency in the analysis that cannot be retained by other means; and 2. A consistent operational procedure is established to evaluate the relevancy or irrelevancy of various assumptions.

A classical application of the RG is to the spin ordering phenomena near the Curie point, a static problem (i.e. temperature equilibrium). The initial conditions are lattice spacing, the local spin-spin Hamiltoniam, etc. and the parameter is temperature. The RG procedure starts with a basic, microscopic Hamiltonian, then examines the crystal [e.g. correlations] on a systematically larger scale. The original local spin Hamiltonian is transformed consistently at each scale change until a macroscopic $\mathcal{H}$effective is produced. In this way one bridges the gap of $\rightarrow\infty$ scale between the microscopic $\mathcal{H}$ and the long range correlation effects known to dominate near phase transitions.

Strictly cooling is non-static (and the RG approach can be extended to nonequilibrium problems[8]). However, an analogy to the spin problem is possible for quasi-static intervals of time $T \ll \tau_{cooling}$ and well past any loop closing transients. Analogous initial conditions are mean particle density, T_0, $G(q_i,q_j,p_j)$, etc. and the analogous parameter is mean kick per turn. Notice that time replaces lattice spacing as the scale variable.

The RG approach relies on scale invariance and locality of interaction. The concept that a frequency domain signal can be used to describe cooling is equivalent to scale invariance in time since $V(\omega)$ is an average over T. This also indicates the necessary feature of long range correlation importance near critical points (instability). A beam instability threshold is characterized by a growing coherent density mode. For this to show up in the Fourier spectrum it must persist over the whole time scale T. Locality of interaction is approximately met in practice since $N/T_0W \ll N$. However $1 \ll N/T_0W$, which is a difficulty corresponding to spin problems including n^{th} nearest neighbors ($n \gg 1$). This is only a complexity in establishing the first step "microscopic" (i.e. one turn) interaction G_{eff}. T_0 is the minimum microscopic scale, at least for filter cooling, due to the inherent filter correlation time ($\simeq T_0$).

The kinetic equation approach to simplifying the full statistical problem is to integrate over all N-1 [phase] space degrees of freedom. The inconsistency of doing this with the basic many particle nature of the interaction leads one to work backwards through the BBGKY hierarchy, which must be truncated in some ad hoc manner (BL35-36). The RG approach starts with a microscopic interaction, then self-consistently integrates out time correlations (frequencies).

References

1. D. Mohl, et.al., Phys. Reports, 58 ,73(1980).

2. J. Bisognano and C. Leeman, in Physics of High Energy Accelerators, R. A. Carrigan editor, AIP Conf. Proc. 87, p583(1982).

3. D. Mohl, CERN/PS/DL 78-25, unpublished (1978).

4. F. Sacherer, CERN-ISR-TH/78-11, unpublished(1982).

5. A trivial example is just the shift in time incorporated into the Vlasov theory approach (see Ref. 2 equation 56 and Ref. 6 section 3).

6. S. Van der Meer, CERN/PS/AA 80-4, unpublished (1980).

7. K. Wilson and J. Kogut, Phys. Reports, 12C, 78 (1974).

8. S. K. Ma, Modern Theory of Critical Phenomena W. A. Benjamin (1976).

9. K. Wilson, Rev. Mod. Phys. , 47 ,773 (1975).