PS/AA/Note 81-2
21 October 1981

SOME ACCELERATOR ASPECTS OF $P\bar{P}$ COLLIDING BEAMS

Rol Johnson

Lectures given at
1981 School of Elementary Particle Physics
Kupari, Yugoslavia
September 28-29, 1981

1. INTRODUCTION

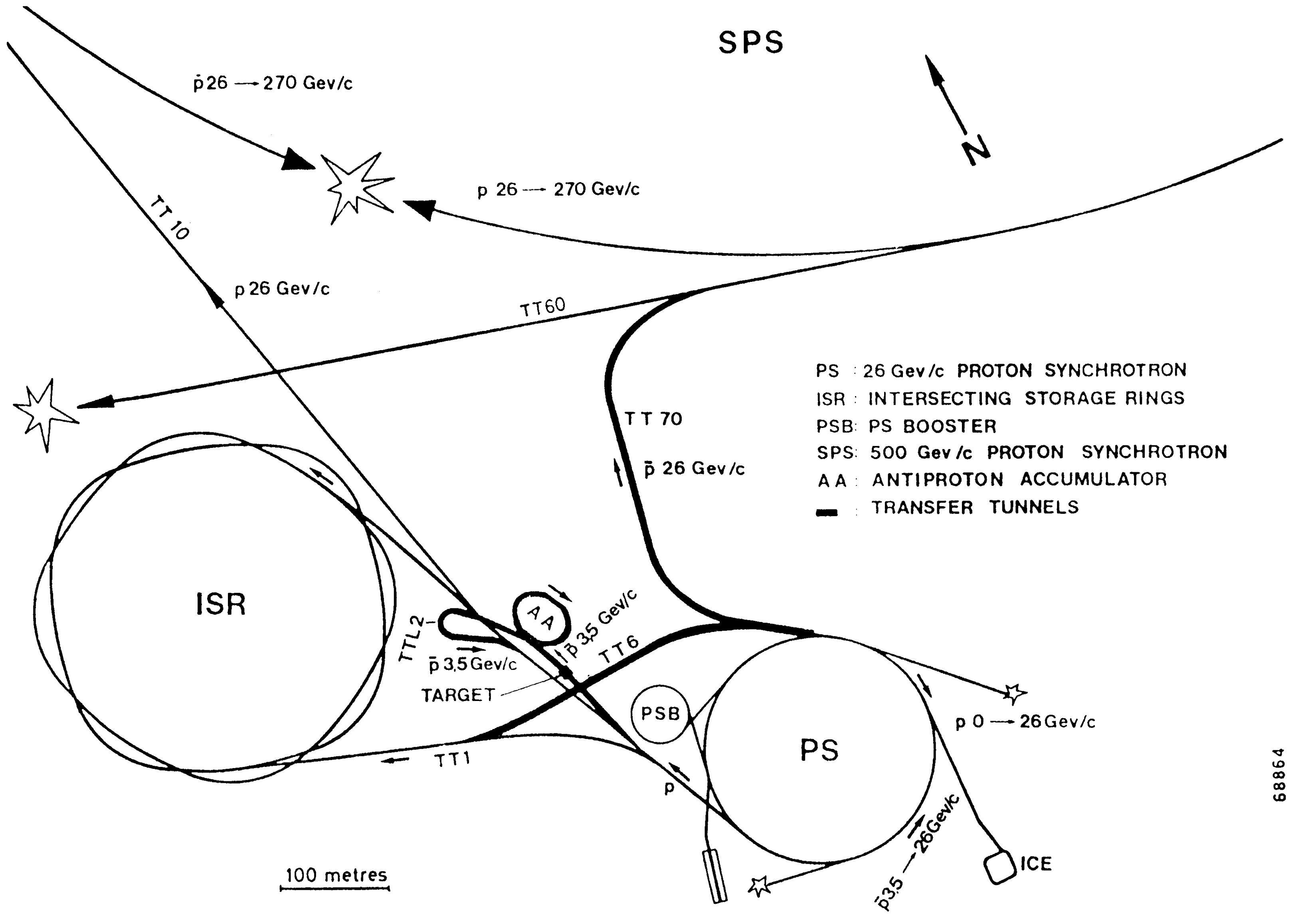
Colliding beams of electrons on positrons or protons on protons has been a fruitful pastime for many years now. The obvious advantage of colliding beams is that the available center-of-mass energy is large at high energy (2E for colliding beams vs. $\sqrt{2m(E + m)}$ for beams on fixed targets). The particular interest in $\bar{p}p$ colliding beams comes from the fact that the particle and antiparticle beams can be stored and made to collide in only one ring of magnets, which means more energy for a certain amount of money spent on magnets. The two cases I will discuss today concern rings of magnets, synchrotrons, which have been or are being built for other purposes so the potential saving of money is even more.

Protons, being heavier than electrons, are less susceptible to limitations due to synchrotron radiation, and for a given amount of money one can build a proton synchrotron of much higher energy than an electron synchrotron. For example, the 1000 GeV proton synchrotron (Tevatron) at Fermilab will cost less than the 70 GeV e^+e^- collider (LEP) at CERN. A further advantage of $\bar{p}p$ collisions over pp collisions is that there are more <u>anti</u>quarks carrying a large fraction of the energy in the antiproton than in the proton. This is of particular importance if you want to see quark-antiquark annihilations. This advantage can be more than an order of magnitude in cross-section for processes where intermediate vector bosons are produced.

Of course $p\bar{p}$ colliders have disadvantages relative to both e^+e^- and pp machines as well. Since protons are made of smaller constituents, quarks, the effective energy you have in constituent collisions has to be degraded by the fact that the energy of a proton or antiproton is shared among its valence and sea quarks and gluons. Detailed calculations typically conclude that a factor of 1/5 in energy is appropriate to compare pp or $\bar{p}p$ machines with e^+e^- colliders. Then again, all those extra quarks and gluons you have in a $\bar{p}p$ or pp collision will tend to mess up your pretty detector displays more than might be expected in the e^+e^- case. So the events may be more difficult to understand and analyze. We don't really know how big a problem this will be and we'll have to wait and see how things turn out.

The disadvantage of $\bar{p}p$ compared to pp mainly follow from the fact that $\bar{p}$'s are hard to get. And having few of them, you have to do some things to make the $\bar{p}p$ interaction rate ($\sim \mathcal{L} \equiv$ luminosity) as large as possible. This is done by squeezing the beams in all three dimensions so the particle density is as large as possible in the interaction region. In the transverse directions the synchrotron structure has to be modified so there is a tight focus on both x and y planes at the center of the detector. This focusing is done with special quadrupoles (called a low β insertion) at the upstream and downstream ends of the detector. In the longitudinal dimension the density is made large by keeping the beams tightly bunched, using the same rf cavities as used to accelerate particles in the synchrotron. With the number of $\bar{p}$'s available and these tricks of making the beam density large, the luminosity of the CERN and Fermilab colliders can be expected to be about 10^{30} $cm^{-2}sec^{-1}$ give or take a factor of ten. (For a $p\bar{p}$ cross-section of 60 mB, ($60 * 10^{-3} * 10^{-24} = 6 \times 10^{-26}$) $* 10^{30}$, this gives 60,000 events/sec, enough to keep most detectors satisfied.)

Since protons are in plentiful supply, the corresponding pp collider doesn't have to work so hard in compressing the beam and one can expect large proton currents in the ring leading to luminosities around 10^{33} $sec^{-1}cm^{-2}$, which more than make up for the factor of 10 due to the lack of antiquarks in the proton compared to antiproton. And since the protons in a pp collider are in different rings, you don't have to worry about the electromagnetic fields around one beam interfering with the particles in the other beam as much as in the case of a $\bar{p}p$ ring where the beams can cross, in principle, all around the ring. This e-m interference of the two beams is called the beam-beam tune shift and we'll talk about that more later, as it is probably the ultimate luminosity limitation of a $\bar{p}p$ collider as it is for e^+e^-.



CERN OVERALL SITE LAYOUT

FIG. 1

At present there are two $\bar{p}p$ projects, one at CERN nearing completion and another at Fermilab in the design stage. What I plan to do in these lectures is to first describe the CERN scheme, with some comments on its present status, and then describe the Fermilab scheme as it is now envisioned. In both cases I will emphasize the accelerator aspects of antiproton accumulation, the most essential part of a $\bar{p}p$ collider.

Then, after having taken a global look at the problem with these two specific examples, tomorrow we can come back to some of the techniques used to acquire antiprotons and try to understand them in more detail. In particular we shall look at stochastic cooling, stochastic stacking, and rf bunch rotation as parts of the antiproton accumulation process. We can try to develop an intuitive feel as well as do some mathematical manipulations to allow us to understand the potentials and limitations of these techniques.

It seems very likely to me that the results from $\bar{p}p$ collisions at the CERN and Fermilab colliders will dominate experimental high energy physics for the next 10 years. The extent of these new and wonderful discoveries, especially in light of competition from single pass colliders (SLAC), LEP (CERN), ISABELLE (BNL), and UNK (Serpukhov), will be in large part determined by our ability to create and manipulate relatively large quantities of antimatter.

2. THE CERN ANTIPROTON-PROTON COLLIDER

Figure 1 shows a schematic of the CERN accelerator complex. The dark lines correspond to things that had to be built for the $\bar{p}p$ project. The general scheme is as follows:

1. The PS accelerates 10^{13} p to 26 GeV every 2.4 sec and directs them to the copper $\bar{p}$ production target.

2. The 3.5 GeV/c antiprotons produced from this target ($\sim$ 2.5E7 $\bar{p}$, in principle) are focused by a pulsed magnetic horn and directed by a system of magnetic quads and dipoles into the outer radius of the ring called the AA. A septum magnet - fast kicker combination injects the $\bar{p}$'s into orbit.
3. $\sim$ 2 KW of longitudinal stochastic cooling is applied to reduce the momentum spread from 1.5% to about 0.2% in about 2 seconds.
4. A normal rf cavity is used to bunch this cooled beam and decelerate it to lower momentum at the same time as shutters which were needed to do the fast longitudinal cooling are pulled out of the middle of the aperture to let this bunched beam pass by.
5. The beam is debunched by removing the rf field, and a stochastic cooling system decelerates the precooled particles into the stack of $\bar{p}$'s which have already been accumulated. This leaves room for the next shot of precooled beam. This step is stochastic stacking, which we'll discuss tomorrow.

The shutters were closed as soon as the bunched, precooled beam had passed them, thus the outer radius of the ring is ready for another shot of $\bar{p}$'s. Steps 1-5 are repeated every 2.4 secs for 24 hours until $\sim$ 6E11 $\bar{p}$'s have been added to the ring, and the SPS is ready for another fill because the stored beams have decayed away due to collisions with gas in the vacuum chamber of the SPS.

The stored $\bar{p}$'s in the AA are extracted by the rf in 6 bunches from the stack, one every 2.4 secs, and sent to the PS. There they are accelerated to 26 GeV, matched to the 200 MHz rf system of the SPS (bunch length $<$ 5 ns), kicked out of the PS, and injected into the SPS. The 6 bunches can be recombined into 3 which circulate oppositely to 3 bunches of protons injected via the other transfer line into the SPS from the PS. The protons and antiprotons are then accelerated to 270 GeV where they stay making life exciting for the groups of experimenters at the two underground experimental areas.

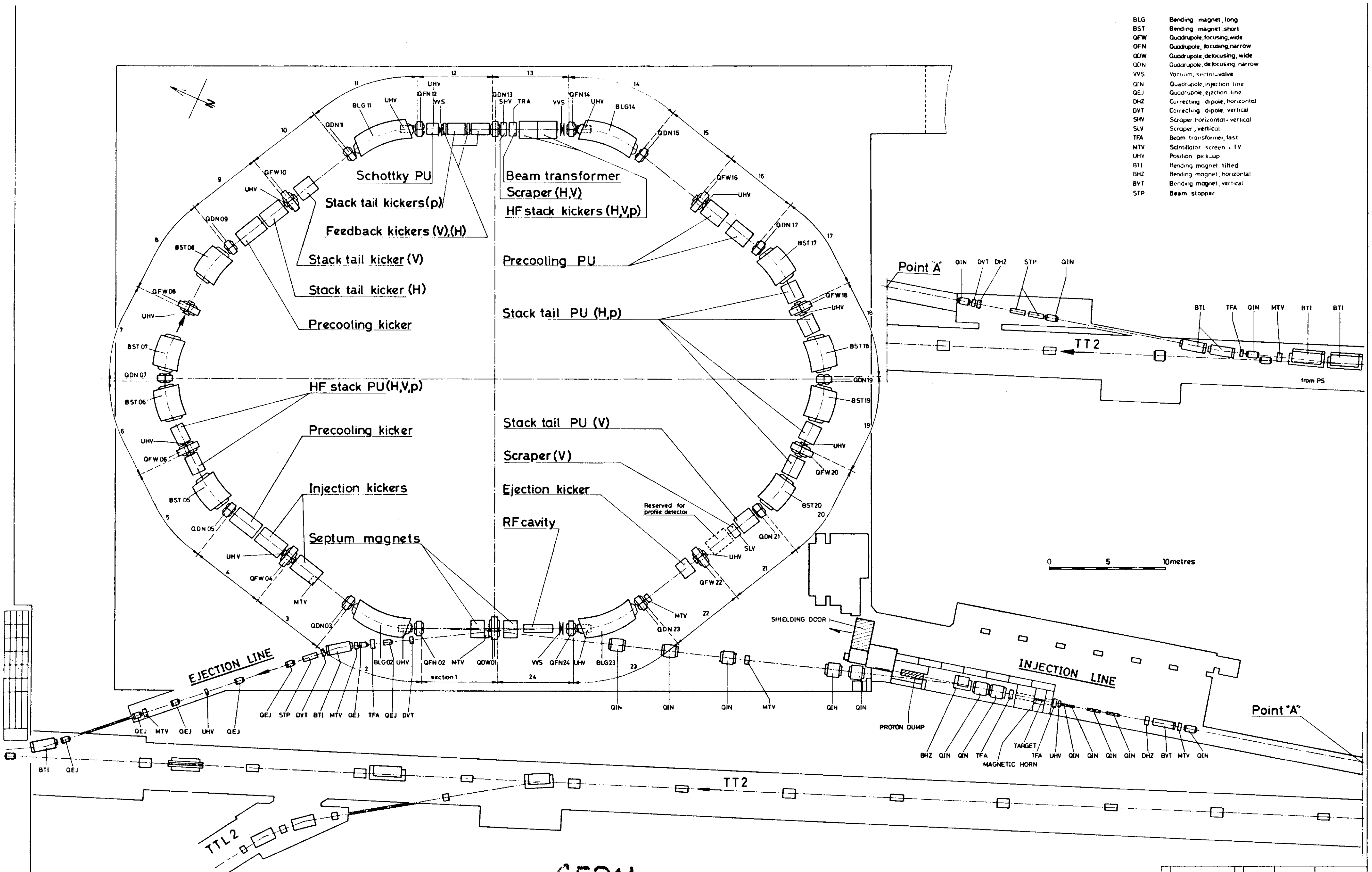
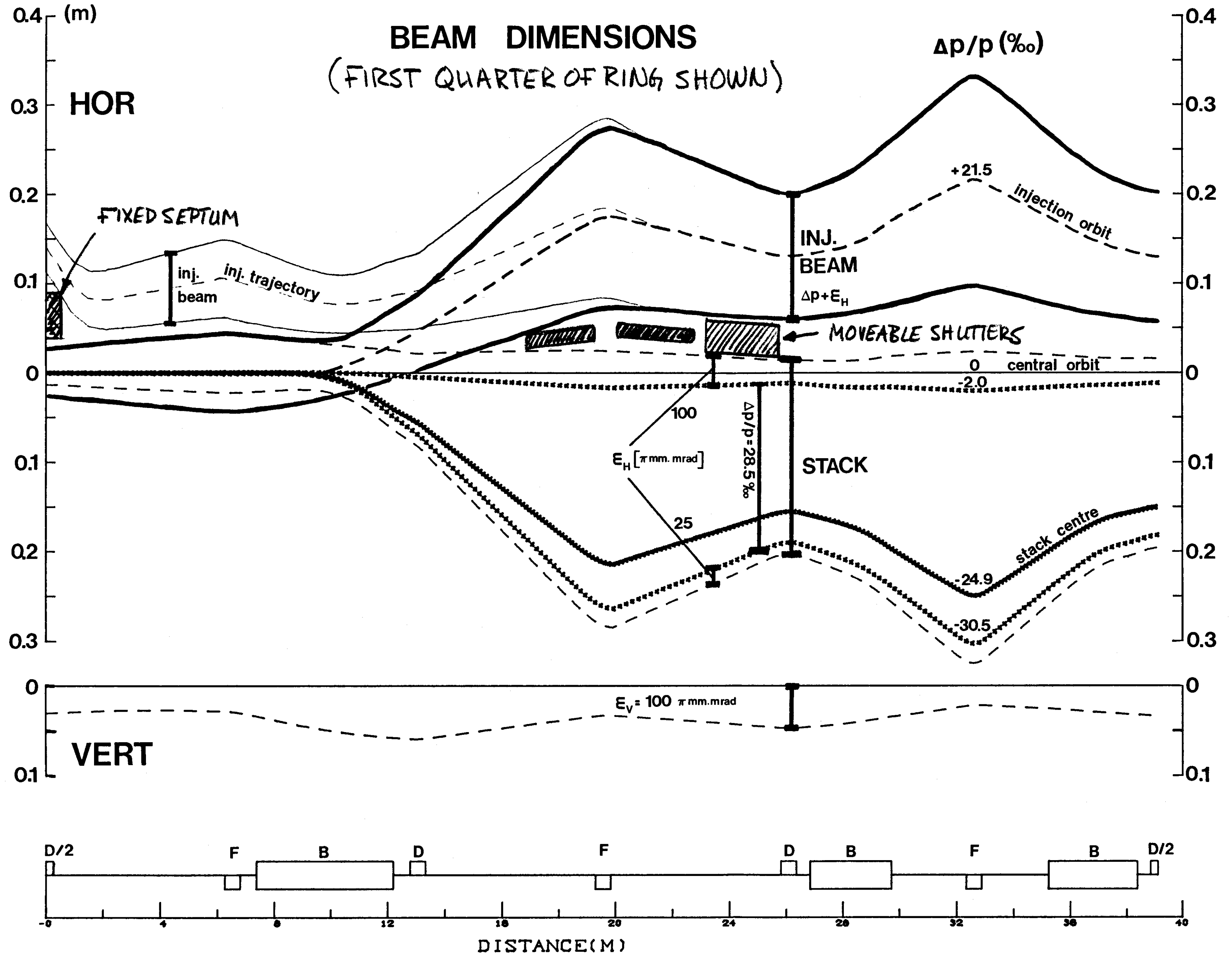


FIG. 2

CERN
ANTIPROTON ACCUMULATOR GENERAL LAYOUT

FIG. 5


BEAM DIMENSIONS
(FIRST QUARTER OF RING SHOWN)
HOR
(m)
Δp/p (‰)
FIXED SEPTUM
inj. beam
inj. trajectory
+21.5
injection orbit
INJ. BEAM
Δp+ε_H
MOVEABLE SHUTTERS
0
central orbit
-2.0
100
ε_H [π mm. mrad]
Δp/p=28.5‰
STACK
25
stack centre
-24.9
-30.5
ε_V = 100 π mm.mrad
VERT
D/2
F
B
D
F
D
B
F
B
D/2
DISTANCE(M)

Now that's an overview of the process. Lets go back and look at the way the aperture of the AA is utilized to handle these $\bar{p}$'s.

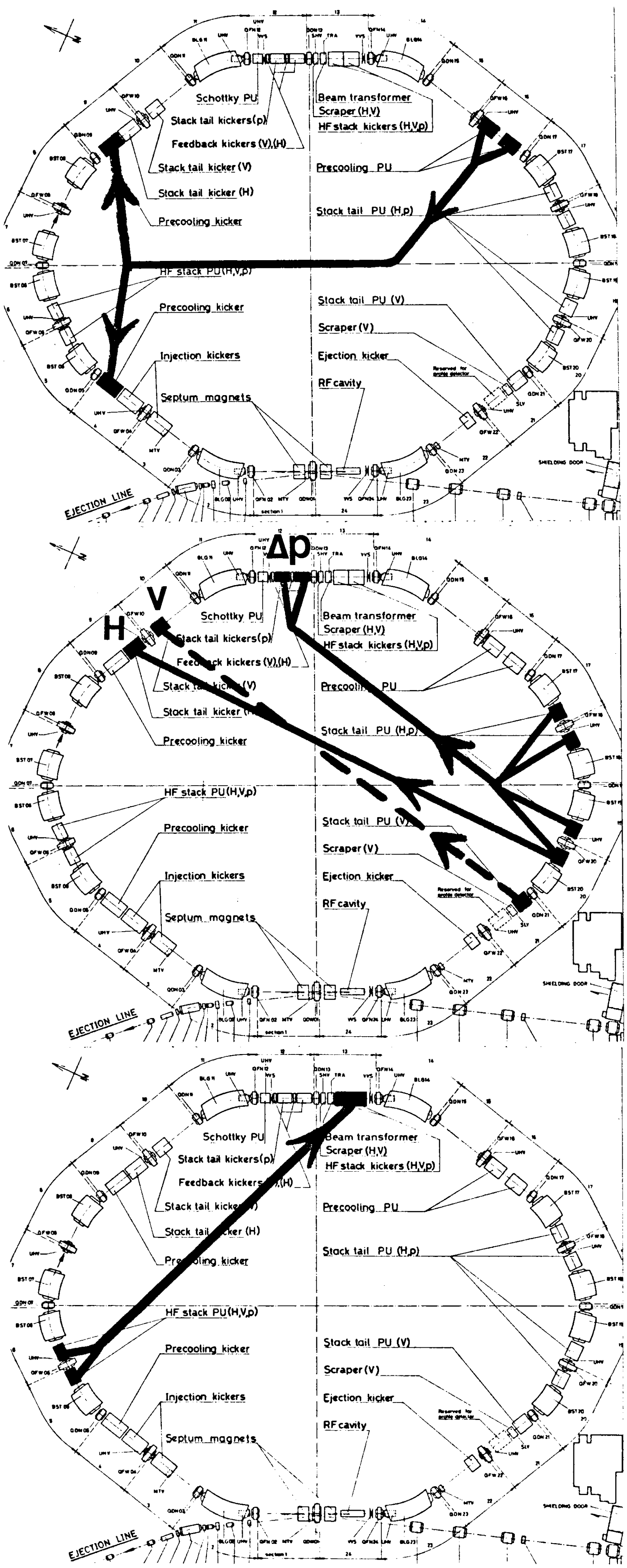
Figure 2 shows a more detailed description of the AA ring.

Figure 3 shows a schematic of the AA ring aperture. The ring has been opened up and straightened and the scales have been modified. The top part shows the view from above. Note that the physical size of the vacuum chamber is close to 3/4 m which in turn corresponds to a momentum spread of $\sim$ 6%. The blocks in the middle are the moveable shutters; the fat ones are the ferrite blocks used to form the single-turn transformers used as momentum precooling pick-ups and kickers. The two thin ones are the sheets of metal which shield the stack region from the fields of the injection kicker magnet which puts the injected $\bar{p}$'s into a circulating orbit. The two blocks at the ends of the diagram are the fixed septa of the two septum magnets.

The injected antiprotons fill the region between the shutters and the outer vacuum chamber wall. Precooling then reduces the injected $\bar{p}$ beam size to that indicated in about 2 seconds. These particles are then trapped in a rf potential well or bucket (carefully) and the frequency is changed to move the particles inwards. The shutters open for just long enough to let the bunched beam pass. The $\bar{p}$ are then released from the rf bucket as close as possible to the $\bar{p}$'s already circulating in the inner half of the ring.

At this point, another shot of $\bar{p}$'s is injected into the ring and the injection cycle starts over again. As you can see, it only remains to merge those particles we just deposited with those already circulating on the inner half of the ring so that the region is clear 2.4 seconds later when we want to deposit the next bunch of precooled antiprotons. If there were particles left in this area and we tried to deposit a fresh bunch in the same place, the rf while depositing the freshly precooled antiprotons would accelerate those left from the previous bunch (by phase displacement) into the now closed shutters.

FIG. 4



STOCHASTIC COOLING
schematic layout

arrows are in the direction from pick-up to feedback kicker

PRECOOLING

Δp only

STACK TAIL COOLING

Δp, hor. and vert. separated as indicated

STACK CORE COOLING

Δp, hor. and vert. combined

So this "merging" is an essential part of the accumulation process. It is done by stochastic cooling and the whole process is called stochastic stacking. We will discuss both stochastic cooling and stacking in more detail tomorrow. The $\bar{p}$'s are effectively decelerated by this cooling until they reach the stack "core" region where the cooling system stops the deceleration and the particle density increases.

So far in the precooling and stochastic stacking we have only been concerned with momentum cooling. And we can conveniently divide these momentum cooling systems conceptually into three distinct systems; (1) precooling, (2) stack tail cooling, and (3) stack core cooling.

In fact there is also a need to cool the transverse or betatron dimensions as well. And as in the stack momentum systems we need to divide transverse cooling onto a system for the less-dense stack tail and another for the stack core. So there are 4 more (Betatron) cooling systems; (4) stack tail horizontal, (5) stack tail vertical, (6) stack core horizontal, and (7) stack core vertical.

Figure 4 shows a schematic of the AA ring with the positions of the P.U.s and K.s of these seven systems. It is fair to point out that these seven stochastic cooling systems are not independent of each other. The stack tail and stack core momentum systems are designed to work together to decelerate and contain the particles in the stochastic stacking process. However, there are also some other interrelationships between the systems that are undesired and, in fact, are some of the problems that are being solved right now.

Here are some of the design parameters of the AA compared with the best obtained results so far:

	Design	Best so far
Number of $\bar{p}$ produced and accepted before precooling per 10^{13} protons on target	2.5×10^7	0.8×10^7
$\bar{p}$ production rate (normalized to incident proton intensity) $\bar{p}$/hour	6×10^{11}/ 24 hours $= 2.5 \times 10^{10}$	$\sim 0.5 \times 10^{10}$
Maximum no. of $\bar{p}$'s stored	10^{12}	1.1×10^{11}
Beam lifetime in SPS for bunches of $\sim 10^{11}$ $\bar{p}$ (hours)	$\gtrsim 24$	~ 10

So in these figures you can see the crux of the accumulation problem, namely that only 1/3 of the number of $\bar{p}$ expected from the production target are getting into the machine. The major part of this, relative to the design figures, we believe is just due to an overly optimistic estimate of the $\bar{p}$ production cross-section (probably a factor of ~ 2 too high).

As to the low lifetime of the high intensity bunches in the SPS, there is no fundamental reason that we know of to prevent reaching the limit of single particle scattering from the residual gas which should be $\gtrsim 200$ hours. It is thought some other mechanism, such as noise in the systems controlling the rf, is causing the beam to blow up. Although such problems are fixable they may take some time to solve.

The $\bar{p}$'s from the AA have been used to provide engineering development work at both the ISR and the SPS. In both cases collisions were seen by experimental groups with more or less success at the rather low luminosity of $\sim 10^{25}$ to 10^{26} cm^{-2}sec^{-1}.

3. THE PLANS FOR THE FERMILAB $\bar{p}p$ COLLIDER

Compared to CERN, the Fermilab $\bar{p}p$ plans have two major advantages. One is that the collisions will take place in the 1000 GeV superconducting Tevatron compared to the 270 GeV SPS. The other is that Fermilab follows CERN and has benefitted greatly from CERN inventions and experience.

Fermilab plans to use the Main Ring at $\sim$ 100 GeV to produce $\bar{p}$'s at $\sim$ 9 GeV/c. The 100 GeV protons will be tightly bunched in time. Since the energy and time are conjugate variables in a synchrotron, this tight time structure can be traded against a wide energy distribution to get effective cooling. That is, a phase space area narrow in time and wide in energy can be rotated to give a narrow energy distribution. This reduction in momentum spread is an effective cooling.

Figure 5 shows the new rings to be added to the Fermilab accelerator complex.

The $\bar{p}$'s are produced and injected into the 1st ring of magnets, the debuncher. Rf cavities in this ring rotate the phase space ellipse. The beam is then transferred from the 1st ring to the 2nd ring, the accumulator, where stochastic cooling is used to merge the injected particles with the antiprotons already accumulated. This accumulator is very much like the CERN AA except that the Debuncher ring at Fermilab replaces the stochastic precooling at the AA.

After sufficient $\bar{p}$'s are accumulated, they are transferred to the Debuncher ring, accelerated to 23 GeV, transferred to the Main Ring, accelerated to 150 GeV, and transferred to the Tevatron. When 3 bunches each of protons and $\bar{p}$'s have been stored at 150 GeV, the p's and $\bar{p}$'s are simultaneously accelerated to 1000 GeV. Collisions should take place in the two experimental areas until something unexpected happens. The

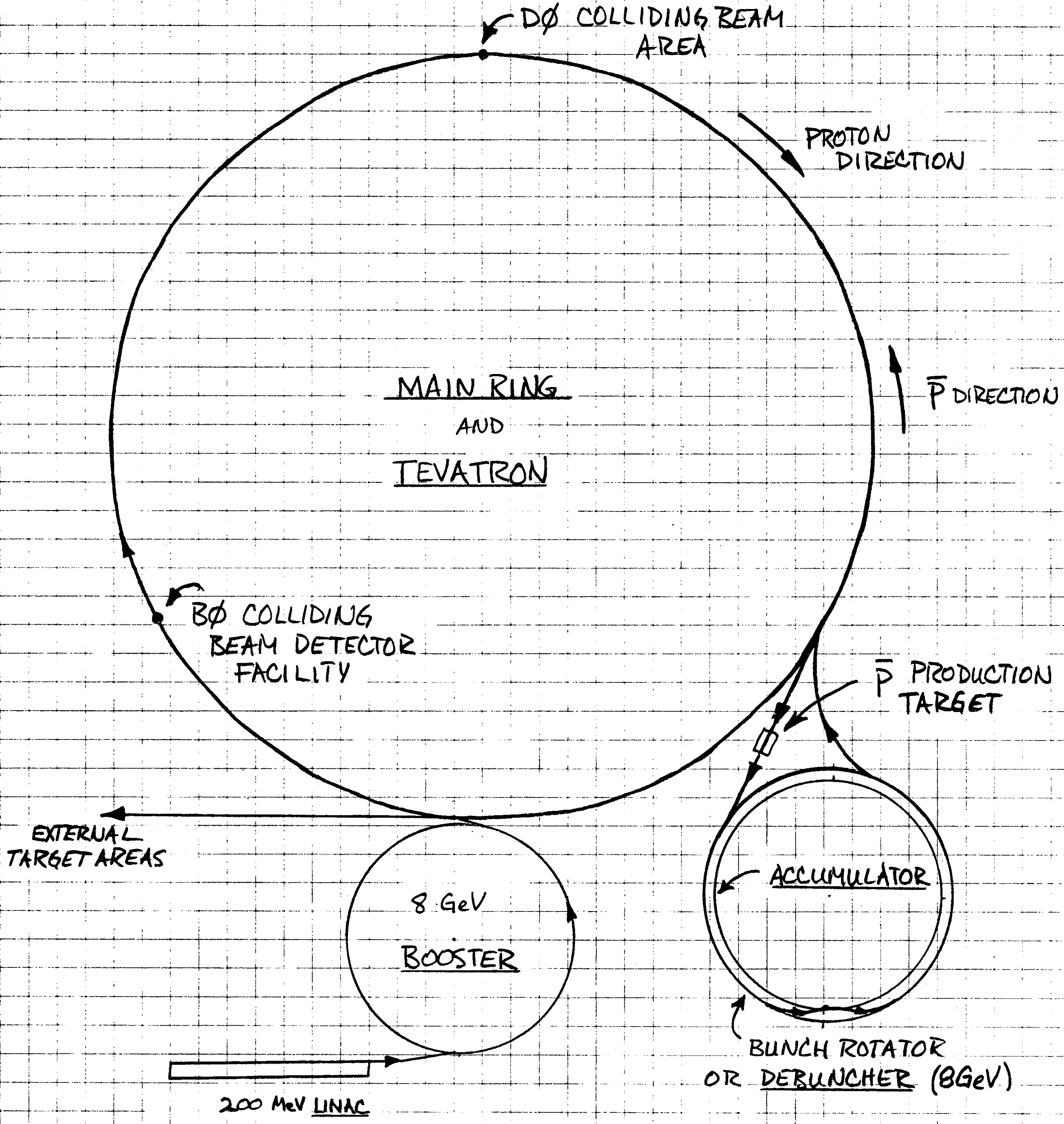


FIG. 5 FERMILAB SITE

good vacuum (cryogenically pumped) and high energy should give very long lifetimes. For very large currents, the beam-beam tune spread will be the lifetime limitation.

Let's go back and look at the $\bar{p}$ accumulation at Fermilab and compare it with CERN's. Obviously 1 ring is different than 2. And bunch rotation isn't used at CERN. As it turns out, the requirements on a synchrotron lattice for bunch rotation are more or less opposite to those for stochastic cooling. For stochastic cooling one needs a strong dependence of revolution frequency on particle momentum or large η. For an rf system in a synchrotron, the size of an rf bucket varies inversely with $\sqrt{\eta}$. Thus to hold the largest momentum spread with the smallest rf gap voltage one needs a small η. Thus the bunch rotation scheme demands 2 separate rings. Which is probably more expensive than one ring and demands more (probably lossy) transfers between rings.

Using 100 GeV protons and gathering a larger momentum spread with the debuncher ring, one expects to have 10 to 20 times more $\bar{p}$'s to accumulate per unit time than at CERN. This larger flux implies more difficulties in the accumulator. The present Fermilab scheme is to use higher frequency stochastic cooling systems to allow higher accumulation rates. Studies are now starting to see how feasible systems of 2-4 GHz and 4-8 GHz will be.

4. COOLING OF $\bar{p}$ BEAMS

The biggest problem to make a $\bar{p}p$ collider is to get enough antimatter to make it all worthwhile. Not only do we have to produce the $\bar{p}$'s but also we must pack them tightly, not only for the subsequent necessity of getting a good luminosity, but just to allow accumulation in a reasonably-sized machine. There are two techniques for this packing, electron cooling and stochastic cooling.

In both cases one reduces the momentum spread, transverse size and angular divergence of an ensemble of particles in a circular accelerator. In fact, the phase space density can be increased, even though the more astute student may note that this seems to violate Liouville's Theorem. But that well-known theorem really is a statement about the conservation of phase-space density of a continuous medium. What we are dealing with is an ensemble of discrete particles; you can check that the formulae we will discuss here contain N, the number of antiprotons, in a way consistent with Liouville's Theorem. As N increases, the particle ensemble acts more and more like a continuous medium and the cooling becomes increasingly difficult.

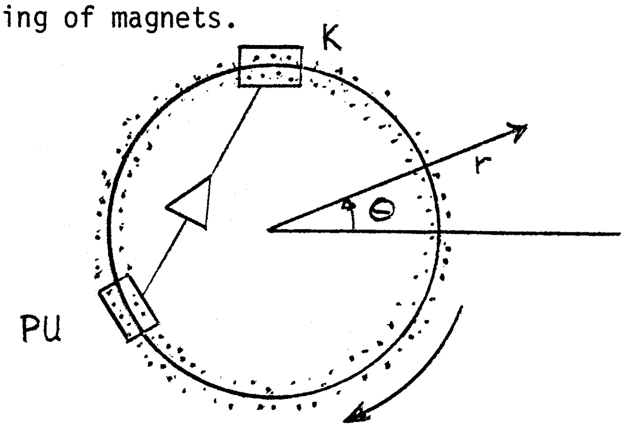
In electron cooling, a high intensity beam of electrons is made to travel parallel to and at the same velocity as the $\bar{p}$'s in a straight section of the $\bar{p}$ accumulation ring. The electrons have the good properties you want the $\bar{p}$'s to have; small angular divergence and small velocity spread. These good qualities rub off on the $\bar{p}$'s, so to speak, in that the ensemble of $\bar{p}$'s and electrons tend to come into thermal equilibrium. You can imagine travelling along with the particles as the few "hot" $\bar{p}$'s collide with the numerous cool electrons, heating the electrons and becoming cooler in the process. And, in fact, here you can see one difficulty with electron cooling. Because if you want to see how fast this cooling takes place in the laboratory, you have to do a Lorentz boost and what you find is that the cooling time has a very strong energy dependence.

But to create lots of $\bar{p}$'s you need incident protons beams of tens of GeV and the maximum $\bar{p}$ production occurs for $\bar{p}$ energies above $\sim$ 3 GeV. And to get good electron cooling times one has to decelerate the $\bar{p}$'s and/or make a cool high intensity electron beam at relatively high energy. The deceleration takes time from the $\bar{p}$ accumulation cycle and such electron beams are technically very difficult. For the low energy $\bar{p}$ beams in LEAR, electron cooling should work very well. For now, we will concentrate on stochastic cooling.

5. STOCHASTIC COOLING

The Basic Idea

Let's first try to develop a mental picture of what happens during stochastic cooling. First consider the case of transverse, or betatron cooling. Suppose that a small number, N, of particles is circulating in a ring of magnets.



Here r is the radial coordinate and θ the azimuthal angle.

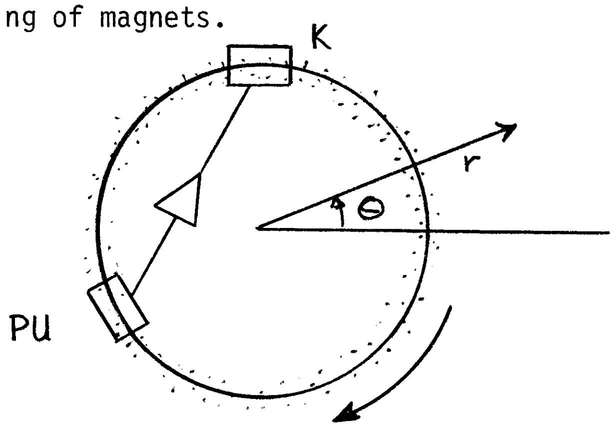
Now suppose a detector, labelled P.U. for pick-up in the drawing, takes a "snapshot" of the distribution within its acceptance. What the detector sees is the center of gravity of the distribution.

But to create lots of $\bar{p}$'s you need incident protons beams of tens of GeV and the maximum $\bar{p}$ production occurs for $\bar{p}$ energies above $\sim$ 3 GeV. And to get good electron cooling times one has to decelerate the $\bar{p}$'s and/or make a cool high intensity electron beam at relatively high energy. The deceleration takes time from the $\bar{p}$ accumulation cycle and such electron beams are technically very difficult. For the low energy $\bar{p}$ beams in LEAR, electron cooling should work very well. For now, we will concentrate on stochastic cooling.

5. STOCHASTIC COOLING

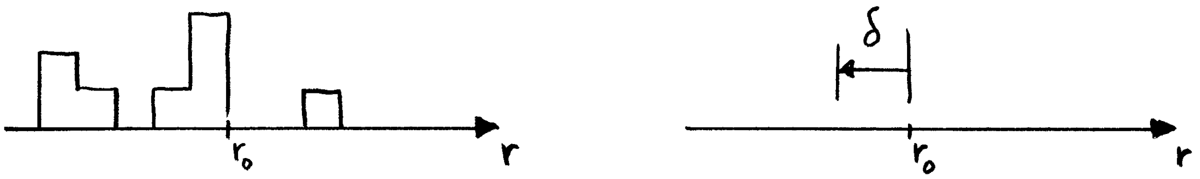
The Basic Idea

Let's first try to develop a mental picture of what happens during stochastic cooling. First consider the case of transverse, or betatron cooling. Suppose that a small number, N, of particles is circulating in a ring of magnets.



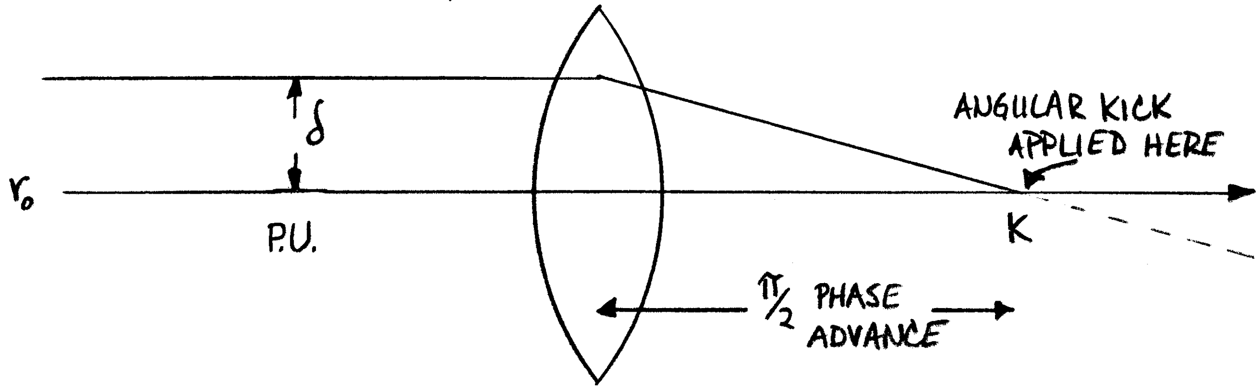
Here r is the radial coordinate and θ the azimuthal angle.

Now suppose a detector, labelled P.U. for pick-up in the drawing, takes a "snapshot" of the distribution within its acceptance. What the detector sees is the center of gravity of the distribution.



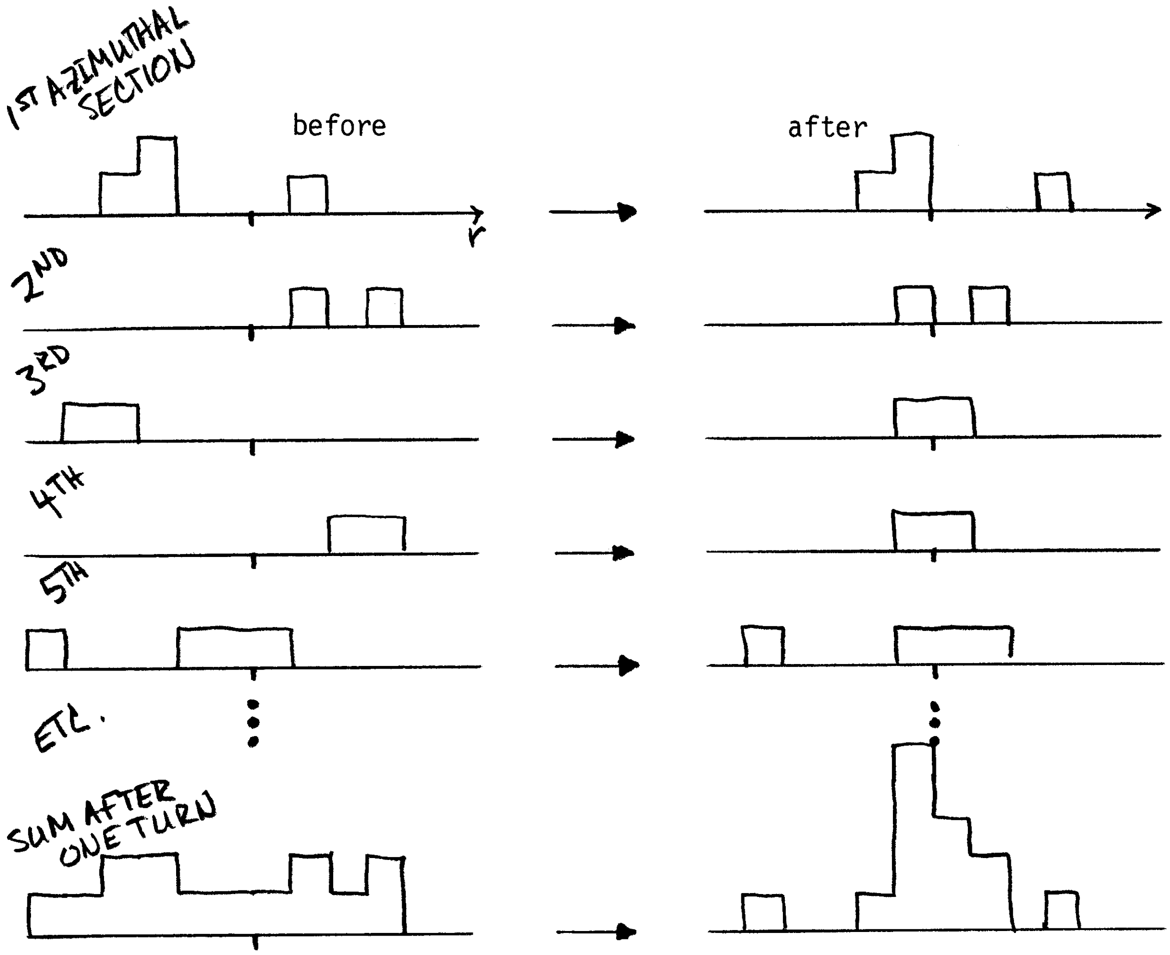
so the distribution → becomes a measure of of displacement

In the case of transverse cooling we can be clever and let the quadrupole magnets in the synchrotron act as a lens to make the correction of the displacement easier.



To correct the position of the C.O.G. of the measured distribution we must have the time of signal transit through the P.U., Amplifier, and Kicker match the time for the particles to go between P.U. and K.

Now imagine that we correct the positions of all the particles in the ring, operating on one azimuthal section at a time.



So after 1 turn the total width of the distribution has been reduced (or cooled). But. If we were to repeat this process on the same sections of azimuth for the next turn, the distributions wouldn't change!

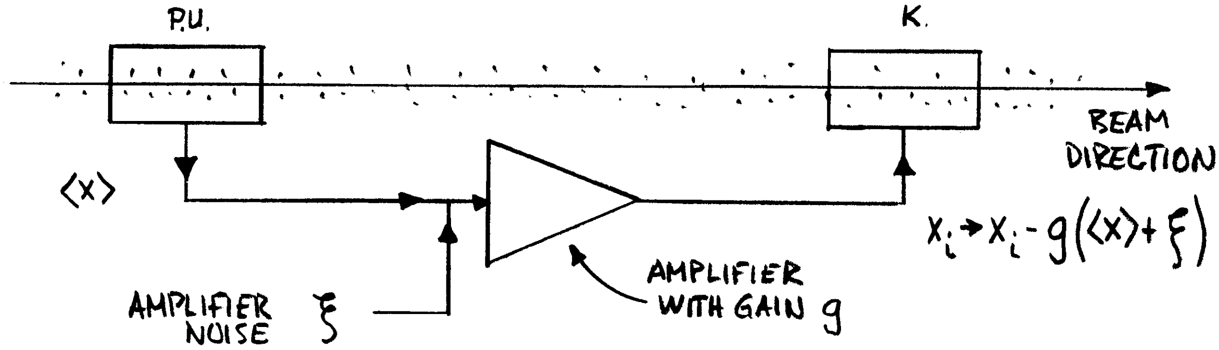
We could do this trick again if we could have a different sample of particles in each of our azimuthal sections. Luckily, this mixing of the distributions happens naturally because the revolution time depends on the particle momentum. Having a distribution of momentum, some particles advance into another azimuthal section, others fall back. And we have new distributions within the P.U. acceptance to kick around.

There you have it -- we can now cool (or reduce the widths of the distributions) on each successive revolution until

1) The momentum distribution is so small that mixing becomes very slow (this can happen if there is simultaneous momentumg cooling). Or ...

2) The detector signals measuring the displacements of the distributions become so small that the noise signal from the amplifiers is overwhelming. Or ...
3) The intrabeam scattering forces cancel the cooling.

Let's try to come up with a numerical estimate of cooling rates. Consider the P.U., amplifier, K. system below.



The P.U. measures the center of gravity of the particles, $\langle x\rangle$. After a single pass through the system, each particle, i, in the sample will have its position changed by $g(\langle x\rangle + \xi)$ or

$$x_i \rightarrow x_i - g(\langle x\rangle + \xi)$$

The mean square of the distribution becomes

$$\langle(x-g(\langle x\rangle +\xi))^2\rangle = \langle x^2 + g^2(\langle x\rangle + \xi)^2 - 2gx(\langle x\rangle +\xi)\rangle$$

$$= \langle x^2 + g^2(\langle x\rangle^2 + \xi^2 + 2\langle x\rangle\xi) - 2gx\langle x\rangle - 2gx\xi\rangle$$

$$= \langle x^2\rangle + g^2\langle x\rangle^2 + g^2\xi^2 + 2g^2\langle x\rangle\xi - 2g\langle x\rangle^2 - 2g\langle x\rangle\xi$$

$$= \langle x^2\rangle - 2g\langle x\rangle^2 + g^2\langle x\rangle^2 + g^2\xi^2 + (2g^2\xi - 2g\xi)\langle x\rangle$$

The last term, proportional to $\langle x\rangle$, can be ignored because $\langle x\rangle \rightarrow 0$, averaged over many samples. The change in the mean square is

$$\langle x^2\rangle - \langle(x - g(\langle x\rangle + \xi))^2\rangle$$

$$= 2g\langle x\rangle^2 - g^2(\langle x\rangle^2 + \xi^2)$$

Now you have to remember something you probably learned in your first physics lab. Namely that the error on the measurement of the center of a distribution is equal to the width of the distribution divided by the square root of the number of elements in the distribution.

$$\langle x\rangle = \frac{x_{rms}}{\sqrt{n}} \quad \text{or} \quad \langle x\rangle^2 = \frac{\langle x^2\rangle}{n}$$

Substituting this into the previous expression for the change in the mean square and dividing by $\langle x^2\rangle$ to get the fractional change of the mean square

$$\frac{dA}{A} \begin{matrix}\text{per}\\ \text{turn}\end{matrix} = \frac{2g}{n} - g^2\left(\frac{1}{n} + \frac{\xi^2}{\langle x^2\rangle}\right)$$

$$= \frac{1}{n}\left(2g - g^2\left(1 + \frac{\xi^2}{\langle x\rangle^2}\right)\right)$$

The time for a particle to go around the ring = 1/f, the time constant of the amplifier = 1/2W, and the total number of particles in the ring = N. Thus the number of particles in the azimuthal sample is

$$n = \frac{Nf}{2W}$$

and the cooling rate of the mean square amplitude becomes

$$\left(\frac{dA}{A}\ \text{turn}^{-1}\right)\left(f\ \text{turns sec}^{-1}\right) = \frac{2W}{N}\left(2g - g^2(1 + \rho)\right)$$

This is for cooling the mean square amplitude in one plane only. Considering rms values, ($\sigma^2 = \sigma_o^2\, e^{-2t/\tau} \rightarrow \sigma = \sigma_o e^{-t/\tau}$), we get only 1/2 this rate. The rate is halved again because the measured displacement

is a projection of the particle's phase space position ($<\sin^2> = \frac{1}{2}$).
The final expression for the cooling rate is

$$\frac{1}{\tau} = \frac{W}{2N} (2g - g^2 (1 + \rho))$$

where $\rho = \frac{\text{amplifier noise power}}{\text{signal power}}$.

Note that this expression can be made positive even with noise > 1 by choosing a smaller gain!

There are some complications to this expression. One is the trivial correction for not having an odd multiple of 90° phase advance between the P.U. and kicker. So multiply the term in g by the sine of the actual phase advance. A second correction is for the mixing between the K. and P.U. being incomplete. This is added as a term $M \gtrsim 1$ in place of the 1 in the g^2 term. M depends mainly on the machine lattice parameter

$$\eta = \frac{df/f}{dp/p} = \frac{1}{\gamma^2} - \frac{1}{\gamma_t^2}$$

Another correction is for mixing between P.U. and Kicker. A fourth correction concerns the feedback from the K. to the P.U. via the beam. These effects as well as complications due to the relatively complicated response functions of the P.U., K., amplifiers and filters are best handled in the frequency domain. The best place to start for the interested student is with Frank Sacherer's articles.

The technique of stochastic cooling was originally invented by Simon Van der Meer in 1968. It was first successfully demonstrated in the CERN ISR where the vertical beam size was cooled to improve the luminosity.

It might be noted that stochastic cooling really involves the coherent displacement of a distribution of particles. The stochastic or random part is the mixing of the particles in a sample. And as said above there is good mixing (between K and P.U.) and bad mixing (between P.U. and K.). The challenge to the machine designers is to invent a system of dipoles and quads, a lattice, which has perfect mixing for one part of the circumference and no mixing in the rest of the ring.

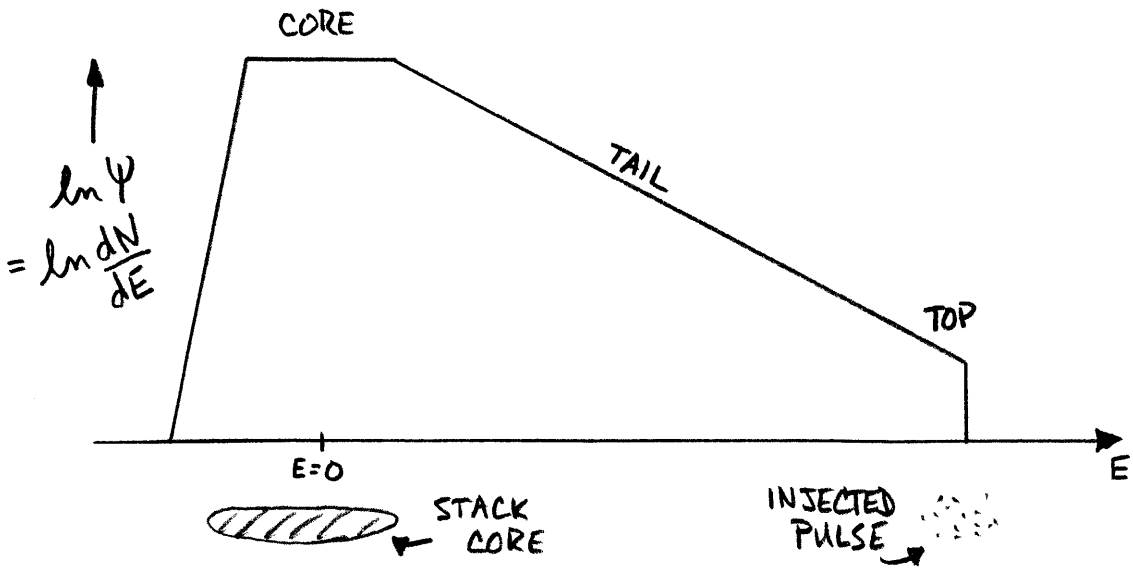
The example discussed was for radial cooling which is the same as for vertical cooling. For momentum cooling the "zero value" of the distributions can be determined with a resonant device (e.g. a shorted co-axial cable or notch-filter). In fact the cable can be attached such that the effective amplifier gain is stronger as the center of gravity of the sample momentum distribution is farther from that corresponding to the revolution frequency given by the cable. This clever technique invented by Lars Thorndahl has the advantage that the amplifier power is used efficiently where it is needed. Another technique, invented by Bob Palmer, uses a radial position pick-up and the correlation between radial position and momentum to do momentum cooling.

6. STOCHASTIC STACKING

Now in a real system for accumulating $\bar{p}$'s, the performance limitation, if it's not due to the meager $\bar{p}$ supply, is in how fast one can accept $\bar{p}$'s and absorb them to make room for new $\bar{p}$'s. Van der Meer calls this stochastic stacking. Namely, by a system of filters and position sensitive pick-ups, you can vary the gain profile of a stochastic cooling network across the aperture of a $\bar{p}$ accumulation synchrotron such that you can quickly cool, or accelerate, low density particles away from the region needed to accept new particles. The gain near the high density of already-accepted $\bar{p}$'s must be low to prevent instabilities of the feedback system. As we will show below, this required low gain follows as well from the stacking rate optimization.

In most statistical mechanics problems one can often get the right answer by taking a macroscopic view (i.e. using Thermodynamic concepts). The same is true for the problem of the density distribution of $\bar{p}$'s in an accumulation synchrotron. The actual feedback loops and filters for the CERN AA are quite complicated in detail. However, one can do quite nicely by considering the particle density as a function of energy. (In a synchrotron you could as well use radius or momentum or revolution frequency in place of energy. All four variables are simply related.)

Following Van der Meer, we can try to derive the optimum stack profile



The flux, $\phi = \frac{dN}{dt}$, passing a certain value of E is function of Ψ and $\frac{d\Psi}{dE}$.

$$\phi = - D \frac{d\Psi}{dE} + F\Psi \qquad (1)$$

If you use the contiuity condition $\frac{\partial\Psi}{\partial t} + \frac{\partial\phi}{\partial E} = 0$, the equation becomes

$$\frac{\partial\Psi}{\partial t} = \frac{\partial}{\partial E} D \frac{d\Psi}{dE} - \frac{\partial}{\partial E} F\Psi,$$

the usual form of the Fokker-Planck equation.

The first term on the right corresponds to the noise effect, a diffusion toward lower density, and varies like g^2. The second term on the right is the cooling effect with F = dE/dt depending on E and varying like g. For cooling, i.e. an increase in particle density, this term in F must push particles towards a region of decreasing F so they pile up.

If the energy loss per turn = V,(proportional to the gain), then

$$F = \frac{-V}{T} ; \quad \text{T the revolution period}$$

$$\text{and} \quad D = AV^2\Psi \qquad (2)$$

$$\text{With} \quad A = \frac{\beta p \Lambda}{4T^3W^2|\eta|}$$

β = particle velocity /c

p = momentum in eV/c

W = bandwidth (H_z) of feedback system between f_{min} and f_{max}

Λ = $\ell n\ (f_{max}/f_{min})$

T = revolution time = $\frac{1}{f}$

$\eta = \frac{-pdT}{Tdp} = \frac{1}{\gamma^2} - \frac{1}{\gamma_t^2}$

This is not meant to be a derivation including these parameters. The functional relationship of the components of A can be found in Sacherer's lectures.

(1) and (2) →

$$\frac{d\Psi}{dE} = -\frac{\phi_o}{AV^2\Psi} - \frac{1}{AVT} \qquad (3)$$

at each value of E we want to chose V to make $d\Psi/dE$ as negative as possible.

Differentiating (3) and setting = 0 gives

$$V = -\frac{2\phi_o T}{\Psi}$$

or
$$\frac{d\Psi}{dE} = \frac{\Psi}{4A\phi_o T^2} = \frac{\Psi}{E_d}$$

The solution is

$$\Psi = \Psi_1 e^{(E_1-E)/E_d}$$

$$V = -\frac{2\phi_o T}{\Psi_1} e^{-(E_1-E)/E_d}$$

with
$$E_d = -4A\phi_o T^2 = -\frac{\beta p\Lambda\phi_o}{TW^2|\eta|} ,$$

the inverse of the exponential slope of the distribution.

$$E_d = \frac{-(E_o - E_1)}{\ell n\Psi_o - \ell n\Psi_1} ,$$

and Ψ_1 and E_1 are the initial density and energy at the top of the stack. (E = 0 is the center of the core.)

Let's write the expression in terms of ϕ_o and discuss the limitations and possibilities to allow the maximum flux of $\bar{p}$'s to be accumulated

$$\phi_o = \frac{E_d TW^2|\eta|}{\beta p\Lambda}$$

E_d is the slope $-(E_0 - E_1)/(\ell n\Psi_0 - \ell n\Psi_1)$. To make it larger you can increase the numerator by making the synchrotron have a larger momentum acceptance or decrease the denominator by having a smaller maximum (core) density.

W^2 is the bandwidth of the rf system. To increase this is to encounter technological problems of higher frequencies, i.e. tighter tolerances on all dimensions. But you can clearly win big as this term goes as the square.

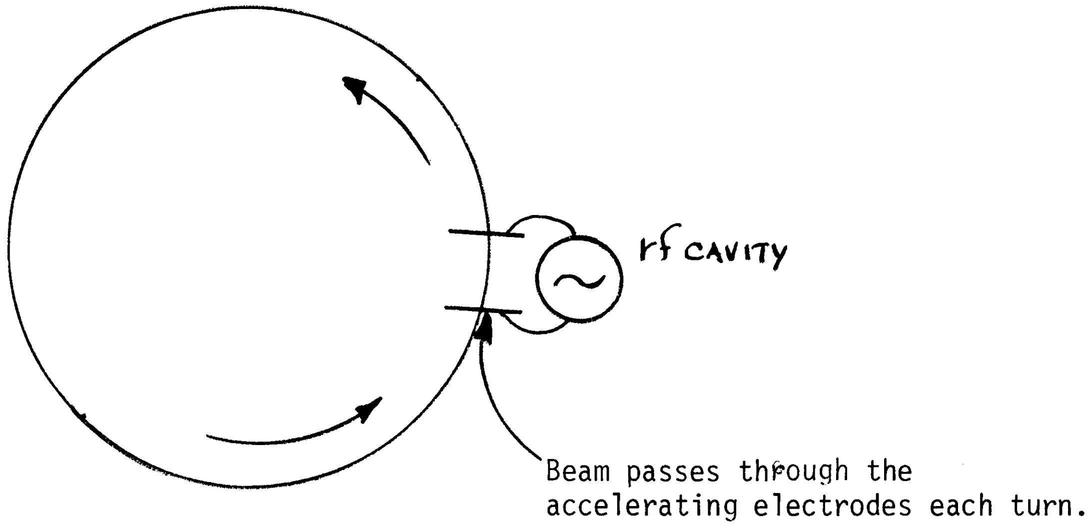
$|\eta|$ is the mixing parameter and depends on how you design the lattice (i.e. pattern of quadrupoles).

T/p is almost a constant and β doesn't change very much.

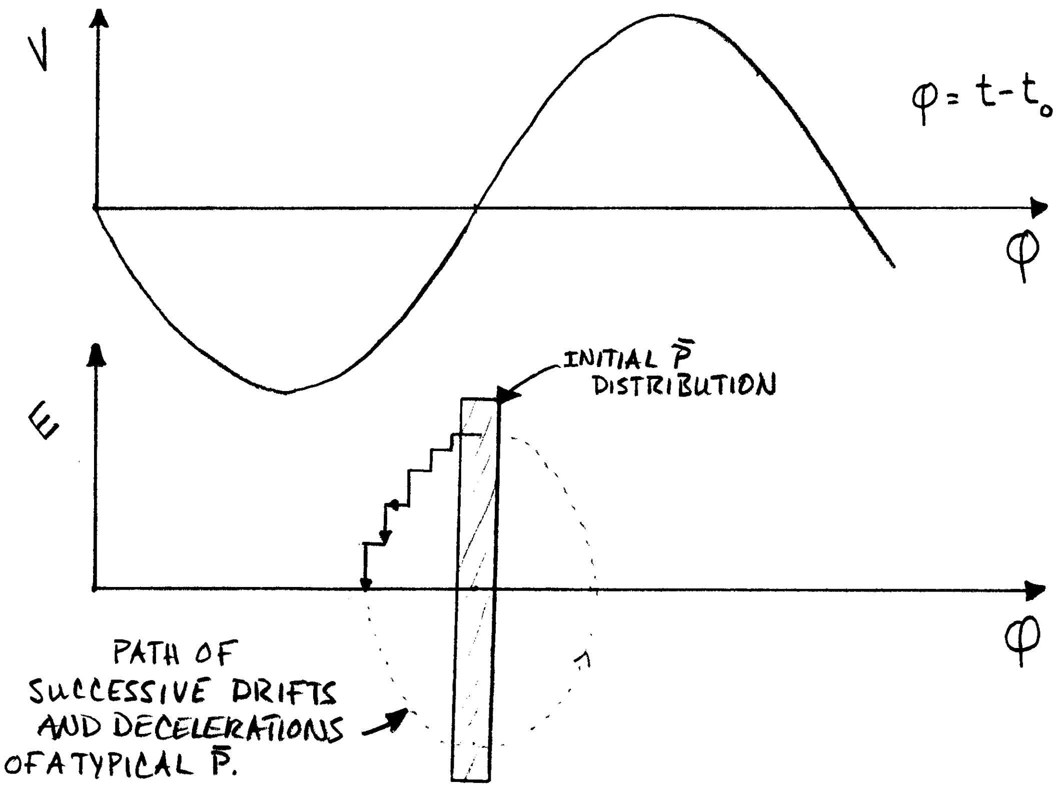
Λ is the log of the ratio of the upper and lower frequencies in the amplifier bandwidth. There is some chance of improving this factor but not much. Amplifiers generally come with constant gain over octaves, i.e., f high/f low = 2/1. Also, the pick-ups are usually tuned to give good response at a particular frequency and lose effective gain as the frequency approaches the edges of the bandwidth. When this effect is taken into account, Λ becomes a less interesting parameter. The expression for A is for the case of constant gain over the bandwidth of the system. For optimum cooling the gain should increase with frequency. Then the lower frequencies contribute less and the extension of the bandwidth to more than an octave helps very little.

7. RF BUNCH ROTATION

Let's consider the system and process of trading momentum or energy spread for the tight bunch structure of the $\bar{p}$ beam. This scheme, envisioned for the Fermilab $\bar{p}$ source, calls for the tight bunching of the Main Ring beam at $\sim$ 100 GeV/c to produce $\sim$ 9 GeV/c $\bar{p}$'s which have the same time structure. The $\bar{p}$'s are injected into a ring with a large momentum acceptance and an RF cavity "rotates" the particles in energy-rf phase space. This is how it works:



The rf cavity produces a sine wave of High Voltage at the revolution frequency of the central energy particle. The rf-phase ϕ is then the time relative to the zero crossing of the rf sine wave.



A particle with higher energy takes longer to go around the ring (above transition, the change in bending radius for higher momentum is a bigger effect than the change in velocity for the ultrarelativistic $\bar{p}$'s). It arrives later than the synchronous particle at the rf cavity and sees a negative voltage and is decelerated by an amount eV sin ϕ. On successive turns, the alternate drift and ΔE increments cause the particle to trace an elliptical path on the E-ϕ diagram. After the particles have travelled 90° around the ellipse and the energy spread is at its minimum the particles are transferred to the other ring.

If the energy spread is small and the oscillations are near the zero crossing of the rf, the sine wave looks like a linear restoring force and the motion is that of a simple harmonic oscillator. This sort of longitudinal oscillation was first seen in synchrotrons and is called synchrotron motion. The number of oscillations per turn is the synchrotron tune:

$$Q_s = \nu_s = \frac{f_s}{f} = \sqrt{\frac{hV\eta}{2\pi E\beta^2}}$$

And the phase space area contained by the rf voltage

$$A = \frac{16\ \beta}{2\pi f} \sqrt{\frac{VE}{2\pi|\eta|h}} \quad \text{(eV-seconds)}$$

This maximum area contained by the rf HV is called the bucket.

The area is measured in eV-s and is an invariant even when the bunches are transferred to another machine.
The Energy is the natural variable as the energy is the thing incremented at each pass of the rf cavity.
β and γ are the usual relativistic variables.
η is the change in rotation frequency vs. momentum.

$$\eta = \frac{df/f}{dp/p} = \frac{1}{\gamma} - \frac{1}{\gamma_t} \ .$$

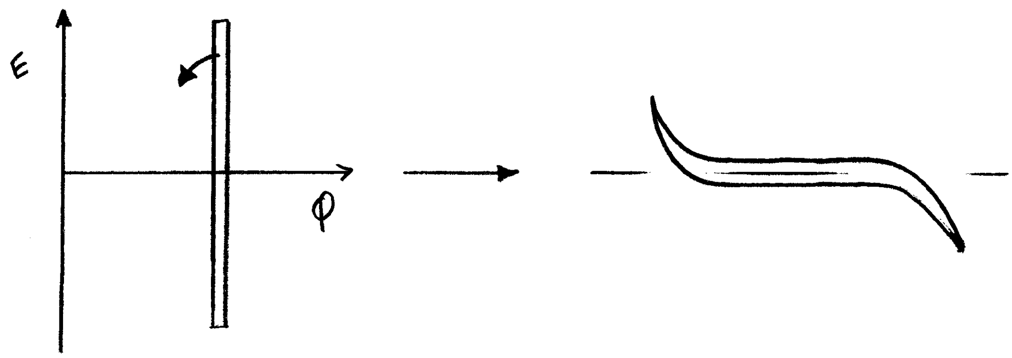
f is the revolution frequency and h the harmonic number = f_{rf}/f.

Here you can see something interesting. As η gets smaller and smaller, the area you can contain with a certain rf voltage increases.

Now η is a consequence of the placements and strengths of the quadrupoles in the lattice. So for bunch rotation you need a lattice with small η to be able to get a large phase space area with a certain rf voltage. But for stochastic cooling you need good mixing or large η. And as mentioned before, this is why you need two rings for Fermilab.

Another interesting comment can be made about the time it takes to do a 90° synchrotron phase oscillation, $1/4f_s$, which also depends on η. If η varies across the aperture, some particles get rotated less and some more than the 90° you would like. The effective cooling is thereby diminished. As it turns out it is difficult to design a large momentum acceptance ring with a constant value of across the aperture. For the AA, for example, η changes from -.078 to -0.111 over its 6% momentum bite. It is a non-trivial challenge to design a machine with the proper virtues.

That the synchrotron frequency is a constant for all particles assumes the linear restoring force. Sine we have a sine-like restoring force, the particles at larger and larger momentum amplitude rotate more and more slowly. The rotated phase space area is correspondingly distorted and the effective reduction in energy spread diminished, i.e.



8. CONCLUSIONS

We have looked at some of the aspects of $\bar{p}p$ collisions, especially $\bar{p}$ production, with as little of the jargon of the trade as I could manage. There are more topics which are interesting and are more or less standard accelerator techniques. The beam-beam tune shift, for example, has been studied and measured for years now. Unfortunately, even though the conceptual picture for this phenomenon is not too complicated, the actual calculational solution is so difficult (and controversial) that there is not much to say at the level of this course.

For those of you interested in learning more about accelerator physics, I recommend Ted Wilson's "Proton Synchrotron Accelerator Theory", CERN 77-7 and references therein, as a good place to start. For more on stochastic cooling there is D. Möhl, G. Petrucci, L. Thorndahl and S. Van der Meer "Physics and Technique of Stochastic Cooling", CERN/PS/AA 79-23, or Physics Reports, 58, N2, 76, 1980, and Frank Sacherer's "Stochastic Cooling Theory", CERN-ISR-TH/78-11 (1978).