Nuclear Instruments and Methods in Physics Research A 350 (1994) 45–52
North-Holland

NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH
Section A

A muon collider scenario based on stochastic cooling ☆

Alessandro G. Ruggiero
Department of Advanced Technologies, Brookhaven National Laboratory, Upton, Long Island, New York 11973, USA

Received 8 April 1994

The most severe limitation to the muon production for a large-energy muon collider is the short time allowed for cooling the beam to dimensions small enough to provide reasonably high luminosity. The limitation is caused by the short lifetime of the particles that, for instance, at the energy of 100 GeV is of only 2.2 ms. Moreover, it appears to be desirable to accelerate the beam quickly, with very short bunches of about a millimeter so it can be made immediately available for the final collision.

This paper describes the requirements of single-pass, fast stochastic cooling for very short bunches. Bandwidth, amplifier gain and Schottky power do not seem to be of major concern. Problems do arise with the ultimate low emittance that can be achieved, the value of which is seriously affected by the front-end thermal noise.

Since mixing within the beam bunches is completely absent, methods are required for the regeneration of the beam signal with external and powerful magnetic lenses. The feasibility of these methods are crucial for the development of the muon collider. These methods will be studied in a subsequent report.

1. Introduction

In the quest for the Higgs bosons, a muon collider may be perceived as the experimental device more affordable and more feasible than electron–positron or very large hadron colliders [1–3]. Muons have a mass ten times lighter than protons and are therefore easier to be steered on circular trajectories. On the other hand their mass is a hundred times greater than electrons and their motion is considerably less affected by the synchrotron radiation. Muons are elementary lepton particles, with no internal structure. Like the electrons, they have obvious advantages over the hadron counterpart when they are used as the main projectiles for the production of the Higgs bosons. Moreover, because of their larger mass, they are also better suited than the electrons themselves, due to a considerably larger propagator constant. Unfortunately, muons do not exist in nature and they have to be produced with the only technique we know these days: impinging an intense beam of protons or electrons on a target. Like in the case of production of antiprotons, in order to make the beam of some use for the subsequent collisions, muons also have to be collected and cooled to a sufficiently high intensity and small dimensions before they can be accelerated and injected in the collider proper. To make the situation more complicated there is also the fact that muons are intrinsically unstable particles with a very short lifetime. Accumulation, acceleration and cooling are then to be executed extremely fast if one requires that a large fraction of the particle beam survives to the collision point.

This paper deals with the requirement of betatron stochastic cooling which is to be very effective and fast. The situation being described is altogether different from the usual encountered with coasting beams [8,9]. Now the beam is tightly bunched at a very large frequency. The bunches are very narrow, having a length which is considerably smaller than the wavelength of the bandwidth of available electronic amplifiers. Thus a different method is to be developed based on the correction of the stochastic signal for all particles at the same time in one single-step. The fundamental limitation remains of the ultimate value of the final emittance that can be achieved. The limitation is caused by the thermal noise at the front-end of the amplifier.

We begin by reviewing a possible scenario of a muon collider in Section 2. This scenario assumes that stochastic cooling is done at maximum energy. The performance of the collider luminosity is evaluated in Section 3, where special emphasis is put on the effects of the beam lifetime and on the betatron emittance reduction. The requirements for the stochastic cooling proper are exposed in Section 4, where we underline the particular situation we are facing of very short beam bunches. The analysis of the cooling device itself follows in Section 5. The goal is the determination of an equation which gives the evolution of the betatron emittance with time. This is described essentially by two parameters: the cooling rate and the diffusion rate due to the thermal noise which is by far more important

☆ Work performed under the auspices of the U.S. Department of Energy.

SSDI 0168-9002(94)00522-9

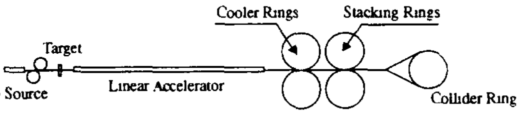

Fig. 1. A conceptual layout of the muon collider.

than any other diffusion process, for instance beam Schottky power. The derivation of the equation for the evolution of the betatron emittance is given in Section 6. A discussion leading to optimization considerations of the cooling process is presented in Section 7. Finally, an application to the muon collider is worked out in Section 8, where the dependence of the performance with a dynamical gain less than optimal and with the bunching frequency is also investigated. It is found that a luminosity of about 10^{24} cm^{-2} s^{-1} can be achieved at the very most. Conclusions are given in the last Section 9.

2. The muon collider

We expose below a possible scenario of a muon collider. A layout of the scheme is shown in Fig. 1. An intense source of either protons or electrons, at the energy of few tens of GeV with an average current of few hundreds of μA, is provided with a conventional fast cycling accelerator [4-8]. In the case of electrons, these are bunched at a large frequency, for instance 3 GHz, and are accelerated in a large-gradient linear accelerator. In the case of protons, after acceleration, the beam is debunched in a stretcher ring and then rebunched also at the frequency of 3 GHz. In either case, the primary beam is made to impinge on a sequence of targets for the production of muon pairs via decay of π mesons. The secondary beam, produced at an energy of about 1 GeV, will be collected with a large production angle yielding a normalized total emittance of about 100 π mm mrad, and with a large momentum bite of few percent. An average intensity of about 50 nA per each component of the pair production is expected, corresponding to a yield of approximately 2×10^{-4} μ-pairs per primary particle. The μ-beams are also tightly bunched at the frequency of 3 GHz so that there are only few particles per bunch (around 100). The lower number of particles per bunch is, as we shall see, a requirement for fast stochastic cooling.

Both types of beam, μ^+ and μ^-, after a preliminary bunch rotation to reduce the momentum spread, are accelerated in a large-gradient linear accelerating structure, operating also at 3 GHz, to the final energy which is in the range of 100 to 1000 GeV. At the end of the acceleration, each beam is transferred to a storage ring where fast stochastic cooling is done to reduce the betatron emittance to the final value. Each beam is then taken to a stacking ring of about the same size, where several cooled beam pulses are stacked side-wise in the momentum phase space. The stacking procedure is carefully done to avoid lengthening of the bunches and increasing of the betatron emittance. With this operation the number of particles per bunch will increase by the number of pulses being stacked. This is required to boost the magnitude of the luminosity of the collider [7]. At the end of stacking, both beams are extracted from their respective stacking ring and transferred to the collider ring proper where they are made to collide and exploited for experimentation. It is to be noticed that all storage rings involved (five, as shown in Fig. 1) operate at the same energy; thus it is expected that they have also about the same size.

Because of the relatively short lifetime of the muon particles, it is obvious that all the operations which have been described above are to be executed very fast.

3. The luminosity performance

What follows is a discussion of the luminosity performance of the muon collider. The average luminosity is given by the following expression

$$L = MN_0^2 f_{\text{bunc}} F\gamma / 4\pi\epsilon_n \beta^*, \quad (1)$$

where $N_0 \sim 100$ is the initial number of particles per bunch at the moment of production, $f_{\text{bunc}} \sim 3$ GHz is the beam bunching frequency during acceleration and stochastic cooling, γ is the energy relativistic factor, $\epsilon_n \sim 25\ \pi$ mm mrad is the initial rms normalized emittance, at the moment of production, and β^* is the focussing amplitude parameter at the interaction point, which for a multiple pass in a collider ring can be as low as 10 cm and for the single-pass mode, where the requirements on the lattice focussing can be relaxed, it is about 1 cm. For an efficient mode of operation, it is important that the bunch length during collision is sufficiently small when compared to β^*. This is obtained with the large bunching frequency. M is the number of beam pulses which are stacked in the momentum phase space of the stacking ring. It is to be noticed that the current $I_\mu = N_0 e f_{\text{bunc}} \sim 50$ nA is a constant equal to the average current of each muon beam at the moment of production. The luminosity expression above shows clearly the advantage of increasing the number of particles per bunch from N_0 to MN_0 with momentum stacking [7]. Finally F is a form factor which includes the losses of particles due to the short lifetime and the emittance reduction due to the stochastic cooling.

We can write the form factor F as the product of many other factors:

$$F = F_{\text{acc}} F_{\text{sto}} F_{\text{stac}}^2 F_{\text{col}}, \quad (2)$$

where F_{acc} is the square of the beam survival fraction after acceleration, F_{sto} reflects the effects on the luminosity of the reduction of the betatron emittance due to stochastic cooling and of the square of the fraction of beam survival after cooling, F_{stac} is the beam survival fraction after

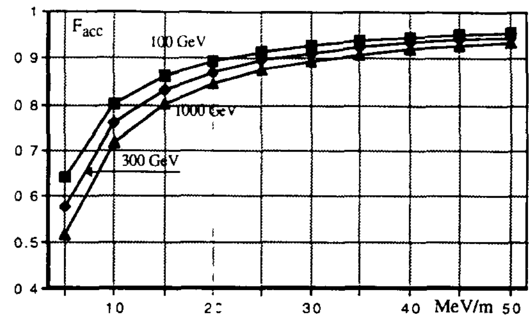

Fig. 2. The acceleration survival factor F_{acc} vs. acceleration gradient.

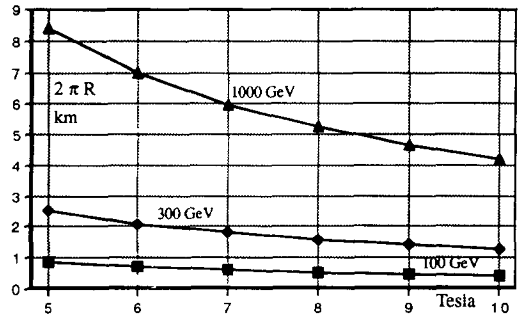

Fig. 4. Circumference of storage ring vs. bending field B for $\eta = 0.5$.

momentum stacking, and F_{col} represents the beam losses during collision.

The overall survival ratio after acceleration can be easily calculated by integrating the instantaneous beam survival over the acceleration cycle, combined with the fact that the lifetime increases linearly with γ. Taking into account the contribution of two beams to the luminosity, we have

$$F_{acc} = (E_{init}/E_{final})^{2E_0/cG\tau_0}, \tag{3}$$

where $E_{init} \sim 1$ GeV is the beam kinetic energy at production, E_{final} is the final energy in the collider, $E_0 = 106$ MeV is the rest energy, $\tau_0 = 2.2$ μs the lifetime of the muon at rest and G is the accelerating gradient in the linear accelerator. The behavior of F_{acc} versus the accelerating gradient is shown in Fig. 2 for various final energies. It is seen that losses are reduced with a larger accelerating gradient and a lower beam energy.

The second factor F_{sto} represents the combined effect of the beam losses during the period of time T_{sto} the beam spends to be cooled and the reduction of the betatron emittance with stochastic cooling

$$F_{sto} = (\epsilon_0/\epsilon_\infty) \exp(-2T_{sto}/\gamma_{final}\tau_0), \tag{4}$$

where $(\epsilon_0/\epsilon_\infty)$ is the ratio of the initial to the final betatron emittance.

The factor F_{stac} represents the survival of the beam integrated over the M pulses being stacked in the stacking ring over a period of time T_{stac}. By denoting with f_0 the circulating frequency, we have $T_{stac} = M/f_0$ and

$$F_{stac} = [1 - \exp(-T_{stac}/\gamma_{final}\tau_0)]/(T_{stac}/\gamma_{final}\tau_0). \tag{5}$$

This quantity is plotted in Fig. 3 versus the parameter $\zeta = T_{stac}/\gamma_{final}\ \tau_0$.

The circumference $2\pi R$ of any of the storage rings can be estimated from the beam energy γ_{final}. Allowing a bending-magnet packing-factor η, denoting with B the banding field in T, and expressing the circumference in meter gives

$$2\pi R = 2.22\gamma_{final}/\eta B \tag{6}$$

from which we derive the revolution frequency

$$f_0 = (135\ \text{MHz})\eta B/\gamma_{final}. \tag{7}$$

Thus is seen that

$$\zeta = 0.0034\ M/\eta B, \tag{8}$$

which does not depend on the beam energy. If we take $B = 6$ T and $\eta = 0.5$ then $\zeta = 0.0011M$. To avoid excessive luminosity losses, it is seen from Fig. 3 that at most we can allow the momentum stacking of $M \sim 900$ pulses. The quantity ζF_{stac}^2 is also plotted in Fig. 3 which shows that at most $MF_{stac}^2 \sim 350\eta B$.

For completeness we display the plots of the circumference $2\pi R$ and of the revolution frequency f_0, respectively

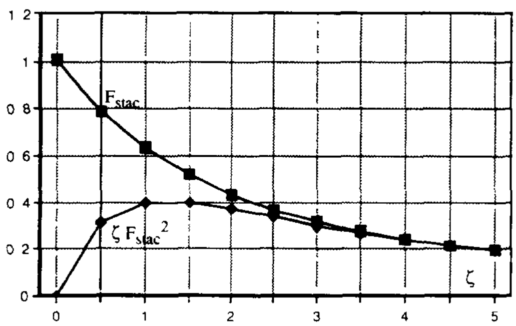

Fig. 3. The survival factors F_{stac} and F_{stac}^2 vs. the parameter.

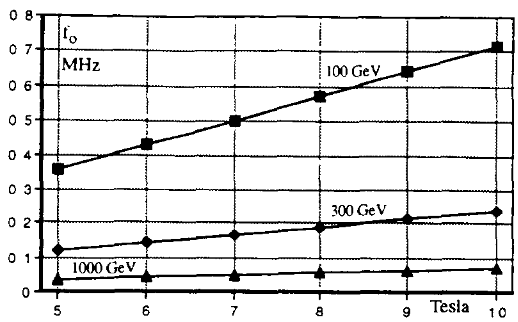

Fig. 5. The revolution frequency vs bending field B ($\eta = 0.5$).

in Figs. 4 and 5, versus the bending field B and for various beam energy values.

Finally, F_{col} represents the loss of average luminosity due to the particle losses during collision which we assume takes a period of time T_{col}. This factor has an expression similar to the one for F_{stac} given by Eq. (5), except that T_{stac} is replaced by $2T_{col}$. It is seen that the best case is given by the single pass mode where after one interaction both beam bunches are immediately disposed. This mode is also more favorable because it does not require a complete collider ring and the final focus may correspond to a lower value of β.

4. Requirements on stochastic cooling

We can express the actual average luminosity in terms of the ideal value L_0 without stochastic cooling, without momentum stacking and for an infinitely long muon lifetime

$$L = MFL_0, \tag{9}$$

where

$$L_0 = N_0^2 f_{\text{bunc}} \gamma / 4\pi\epsilon_n \beta^*. \tag{10}$$

With the values of the parameters given in the previous two sections and for the single-pass mode, which is the most favorable, we have $L_0 = 1.0 \times 10^{15}\gamma$ cm^{-2} s^{-1}. Correspondingly, to achieve a luminosity of about $1.0 \times 10^{27}\gamma$ cm^{-2} s^{-1}, which is required for the high energy physics experimental exploitation of the collider, we need that the overall enhancement factor $MF \sim 1 \times 10^{12}$. Since at most $MF_{stac}^2 \sim 1000$, even assuming $F_{acc}F_{col} \sim 1$, we need $F_{sto} \sim 1 \times 10^9$, which is a very large requirement for the betatron stochastic cooling. The requirement is independent of the beam energy; a normalized emittance of $25 \times 10^{-9}\pi$ mm mrad is required at all energies. Assuming a cooling period T_{sto} short with respect to the muon lifetime $\gamma_{final}\,\tau_0$, one requires a reduction of the betatron emittance by nine orders of magnitude (!). Thus the fundamental question concerns the ultimate emittance that can be realistically achieved at the end of cooling.

There are major differences between stochastic cooling for the case of bunched beams we are investigating here and the usual approach for coasting beams encountered, for instance, during production and accumulation of antiproton beams [9,10]. The muon bunches have a length considerably smaller than the shortest wavelength in the frequency bandwidth of the system. It is indeed a good approximation to assume that the beam bunches have no longitudinal extension, and that all the particles are distributed on a disk with a center slightly displaced from the axis of the pickups. The beam current signal is therefore highly organized and coherent. The transverse beam position, on the other hand, has a very stochastic behavior. Because of the low number of particles, there is a random fluctuation of the beam centroid that can be related statistically to the overall transverse beam size. For the same reasons, the internal motion can be completely ignored and no mixing occurs between the detection of the beam signal at the pickups and the application of the deflection at the kickers. Moreover, the conventional analysis in the frequency domain [9,10] would be hardly applicable, since the Schottky bands remain well separated from each other. Actually from the current point of view, it is improper to refer to the beam Schottky signal. As pointed out already, the stochastic behavior appears only in the transverse displacement of the beam centroid.

The overall system, between pickups and kickers, including the power amplifier, has a bandwidth W large enough to detect and correct displacement of individual bunches. If we assume a bandwidth W extending over an octave, where the frequency at the upper end is twice than the frequency at the lower end, an optimum is given by choosing the bandwidth equal to an integer m times $2/3$ of the bunching frequency. For instance, if the beam is bunched at 3 GHz, the bandwidth could be 2 GHz ($m = 1$) extending from 2 to 4 GHz. A larger bandwidth, for instance from 4 to 8 GHz ($m = 2$) or from 6 to 12 GHz ($m = 3$), is of course also possible. In this mode of operation, it is possible to process the pickup signals to allow complete rejection of a beam bunch signal on another bunch.

The lack of mixing causes a serious limitation on the effectiveness of stochastic cooling. Once the initial beam displacement has been corrected, there is no more signal from the beam that can be used. Thus everything is done in a single step with a relatively small reduction of the beam size. Between steps, the signal from the beam has to be regenerated, for instance by rearranging the particle mutual position with the aid of powerful magnetic lenses. We shall assume below that this is indeed the case. We shall investigate in a separate report the amount of particle rearrangement required and how this can be realized in practice.

Because of the low number of particles per bunch, after amplification, the beam Schottky power is not expected to be excessive. In order to obtain a very fast cooling, one requires to optimize the overall gain to correct the instantaneous beam displacement immediately in one single step. This may require a very large electronic gain. A more serious problem is associated with the thermal noise at the front-end of the amplifier, which will set a limitation of the final beam transverse dimension. The design of the cooling device is to be optimized to reduce this limitation.

5. Analysis of the cooling device

Consider a very narrow bunch made of N particles all with the same electric charge. The bunch is periodically traversing a beam position pickup made of two parallel striplines each of length l and separated by a distance d. The striplines are shorted at one end and terminated at the upstream end to their characteristic impedance R_p. During the occurrence of a traversal assumed at the instant $t = 0$, the bunch current can be represented as a pulse of zero duration, proportional to the average displacement $\bar{x}$, that is

$$I_p(t) = Ne(\bar{x}/d)\delta(t). \quad (10)$$

This current leaves at the upstream end of one pickup a voltage signal given by

$$V_p(t) = Ne(\bar{x}/2d)R_p[\delta(t) - \delta(t - 2l/c)], \quad (11)$$

that is a voltage pulse occurring simultaneously to the current pulse, followed by another at the delay of $2l/c$. Since the beam bunch duration is considerably shorter than the delay between the two voltage pulses, only the first pulse is relevant to our analysis and we shall ignore the second one which we assume can be disposed properly without disruptions to the subsequent bunches. The voltage signal is then filtered by the bandwidth of the system, mostly caused by the power amplifier, and properly amplified by the linear gain A. The resulting voltage is an oscillating and decaying signal of which only the front-end is of relevance here since it constitutes the part that is to be applied in phase with the bunch at the location of the kickers. The amount of the properly correspondent voltage is simply

$$V_A = Ne(\bar{x}/2d)R_p AW. \quad (12)$$

In the case there is only one pickup and one kicker, this is also the voltage that would appear across the kicker assuming an ideal impedance matching all along the transfer of the signal. To include the case of more pickups and kickers, we modify Eq. (12) as follows:

$$V_k = Ne(\bar{x}/2d)AW\sqrt{R_p R_k n_p/n_k}, \quad (13)$$

where $n_{p(k)}$ is the number of pickups (kickers) and R_k is the characteristic impedance of kickers. This expression gives the voltage across one single kicker. We are assuming here that both the pickups and the kickers are closely packed and that they extend over a length of the storage ring where the beam position $\bar{x}$ and the lattice functions do not vary appreciably. Moreover kickers have exactly the same geometrical configuration and size of the pickups.

Assuming small deflection angles and full correlation among the kicks, the total deflection angle each particle will be subject to at the traversal of the kickers is

$$\theta_S = eV_k l n_k/\beta^2 Ed, \quad (14)$$

where E is the particle total energy and βc is the velocity. Statistically, over many revolutions, the following relation holds between the average beam displacement $\bar{x}$ and the rms beam size σ:

$$\sigma^2 = Nx^2. \quad (15)$$

We prefer writing the expression for the deflection θ_S as follows:

$$\delta_S = g_0\sigma/d, \quad (16)$$

where

$$g_0 = \frac{lA}{2\beta^2 Ed}\sqrt{Ne^4W^2 n_k n_p R_k R_p}. \quad (17)$$

Another signal is induced to the kickers which is random and independent of the particle position. The finite temperature of the terminating resistors of the loop and of the preamplifiers creates at the input to the preamplifier a signal of power

$$P_T = k_B(T_A + T_R)W, \quad (18)$$

where $k_B = 8.6171 \times 10^{-5}$ eV/K is the Boltzmann constant, T_A is the equivalent temperature of the amplifier and T_R is the temperature of the resistor. Proceeding in the same way as for the beam Schottky signal, we calculate the total deflection angle due to the thermal noise:

$$\theta_T = \frac{elA}{\beta^2 Ed}\sqrt{n_k R_k P_T}. \quad (19)$$

To assess the effectiveness of stochastic cooling a useful parameter is the ratio of Schottky to the thermal power

$$S = \theta_S^2/\theta_T^2 = Ne^2W^2 n_p R_p \sigma^2/4d^2 P_T. \quad (20)$$

It is seen that this ratio increases linearly with the bandwidth. This result is different from the one it was derived for coasting beams [10], in which case the ratio is independent of the bandwidth. The difference is due to the fact that now the current distribution of the beam bunch is highly organized and does not exhibit a stochastic behavior.

6. The equation for the evolution of the beam emittance

Both Schottky and thermal power kicks apply simultaneously when a particle is crossing the location of the kickers. Let us calculate the effect of both kicks to the beam emittance which we can define as follows

$$\epsilon = \sum_i \left(\gamma x_i^2 + 2\alpha x_i x_i' + \beta x_i'^2\right)/N. \quad (21)$$

Here α, β and γ are the lattice Twiss parameters.

At the kickers each particle receives the same kick, that is $x_i' \to x_i' + \theta_S + \theta_T$. The corresponding emittance change is

$$\Delta\epsilon = 2(\alpha\bar{x} + \beta\bar{x}')_k(\theta_S + \theta_T) + \beta_k(\theta_S + \theta_T)^2, \quad (22)$$

where the subscript k(p) denotes that the corresponding quantity is evaluated at location of the kickers (pickups). $\bar{x}$ and $\bar{x}'$ are the average values of the particle positions and angle. We are interested in the expectation value of $\Delta\epsilon$

over many revolutions. In this case, since there is no correlation between particle position and thermal noise, the previous equation reduces to

$$\langle \Delta\epsilon \rangle = 2(\alpha\bar{x} + \beta\bar{x}')_{k}\theta_{S} + \beta_{k}\theta_{S}^{2} + \beta_{k}\theta_{T}^{2}. \tag{23}$$

The betatron emittance is also defined as

$$\epsilon = (\sigma^2/\beta)_{p,k}. \tag{24}$$

At the same time it can be proven that

$$\langle (\alpha\bar{x} + \beta\bar{x}')_{k}\bar{x}_{p} \rangle = -\langle \bar{x}_{p}^{2} \rangle \sqrt{(\beta_{k}/\beta_{p})} \sin\psi_{pk}, \tag{25}$$

where ψ_{pk} is the betatron phase advance between pickup and kicker. Manipulating some of the previous equations gives

$$\langle \Delta\epsilon \rangle = -(2g \sin\psi_{pk} - Ng^2)\epsilon + \beta_{k}\theta_{T}^{2} \tag{26}$$

with the dynamical gain

$$g = g_0\sqrt{\beta_{k}\beta_{p}/Nd^2}. \tag{27}$$

Finally, the beam emittance evolution is described by the following equation

$$\mathrm{d}\epsilon/\mathrm{d}t = -\lambda\epsilon + D, \tag{28}$$

where the cooling rate

$$\lambda = n_{S}f_0(2g \sin\psi_{pk} - Ng^2) \tag{29}$$

and the diffusion coefficient

$$D = n_{S}t_0\beta_{k}\theta_{T}^{2}. \tag{30}$$

with f_0 the revolution frequency. In deriving these equations we have assumed a total of n_s identical cooling systems in the storage ring. In the following we shall assume that the distance between pickups and kickers is adjusted so that $\sin\psi_{pk} = 1$.

7. Optimization of the cooling performance

An optimum cooling rate is obtained by setting $g = 1/N$ and is

$$\lambda_{opt} = n_{S}f_0/N, \tag{31}$$

which corresponds to correcting the instantaneous beam bunch displacement in one single step. At the same time we can also derive the required amplifier gain

$$A = \frac{2\beta^2Ed^2/lN}{\sqrt{e^4W^2n_{k}n_{p}R_{k}R_{p}\beta_{k}\beta_{p}}}, \tag{32}$$

which decreases linearly with the bandwidth and the number of particles in the bunch.

The equilibrium value of the emittance for this case is

$$\epsilon_{\infty} = D/\lambda_{opt} = N\beta_{k}\theta_{T}^{2}. \tag{33}$$

By combining some of the equations we obtain for the equilibrium emittance

$$\epsilon_{\infty} = \frac{4d^2P_{T}}{Ne^2W^2n_{p}\beta_{p}R_{p}}, \tag{34}$$

which, similarly to the amplifier gain A, also exhibits the same dependence with bandwidth W and number of particles N. It can be seen that the equilibrium emittance corresponds to the situation where the ratio of the Schottky to thermal power $S = 1$.

It is to be noticed that an optimum bandwidth is related to the bunching frequency f_{bunc} by the relation

$$W = 2mf_{bunc}/3, \tag{35}$$

where m is a positive integer. Also the quantity

$$I_{\mu} = Nef_{bunc} \tag{36}$$

is the average muon current essentially equal to the one produced at the target. Thus when these relationships are taken into account we have for the amplifier gain

$$A = \frac{3\beta^2Ed^2/lI_{\mu}m}{\sqrt{e^2n_{k}n_{p}R_{k}R_{p}\beta_{k}\beta_{p}}} \tag{37}$$

and for the equilibrium emittance

$$\epsilon_{\infty} = \frac{6d^2P_{T}/W}{mn_{p}\beta_{p}eI_{\mu}R_{p}}. \tag{38}$$

Both of these expressions show the same dependence on the bandwidth factor m and on the average beam current. Noticing that the thermal power P_T is proportional to W, it is seen that both A and ϵ_{∞} do not depend explicitly on how the beam is bunched. On the other hand the cooling rate λ depends very strongly with the number N of particles per bunch.

8. An application of the optimal system

As an application of these expressions we take the following values:

$d = 1$ cm,
$\beta_{p} = \beta_{p} = 200$ m,
$n_{p} = n_{k} = 1024$,
$R_{p} = R_{k} = 100\ \Omega$.

We chose $m = 3$, that is a bunching frequency $f_{bunc} = 3$ GHz, corresponding to a bandwidth $W = 6$ GHz ranging between 6 and 12 GHz. The length of the pickups is adjusted to match the bandwidth according to $l = 6$ cm/W (GHz) so that for $m = 3$ it is $l = 1$ cm. We assume also a bending field $B = 6$ T and a packing factor in the storage ring $\eta = 0.5$. Finally we set the temperature of the amplifier and resistor $T_A = T_R = 1$ K which is very likely an unrealistic value, for which moreover we cannot really

Table 1
Stochastic cooling performance

Beam energy, GeV	100	300	1000
$2\pi R$, m	700	2100	7000
n_s	8	24	80
$1/\lambda$, ms	0.030	0.030	0.030
A	1×10^9	3×10^9	1×10^{10}
$\epsilon_n = \gamma\epsilon_\infty$, π mm mrad	32	96	320
L_0, cm^{-2} s^{-1}	1×10^{18}	3×10^{18}	1×10^{19}
MF	1000	300	100
L, cm^{-2} s^{-1}	1×10^{21}	1×10^{21}	1×10^{21}
Ng_{max}	0.0068	0.0023	0.0007
f_{max}	40	120	400
L_{max}, cm^{-2} s^{-1}	4×10^{22}	1.2×10^{23}	4×10^{23}

foresee the behavior of the thermal noise at the front end. The summary of the results of our calculations are shown in Table 1, where we have also assumed the optimum gain $g = 1/N$.

To be observed is the increase of the circumference of the storage ring with the beam energy, and that we let the number n_s of cooling systems vary proportionally. As a consequence, the cooling time $1/\lambda$ is constant with energy, whereas the amplifier gain A and the equilibrium emittance ϵ_∞ increase linearly with energy. As one can see, even at the very low temperature of 1 K, thermal noise dominates over the beam signal, and the equilibrium emittance is just about comparable to the initial beam emittance at the energy of 100 GeV. For larger energies there is actually stochastic heating accompanied by an increase of the beam emittance.

Since for the optimum gain the cooling time is 0.03 ms which is considerably shorter than the beam lifetime, it is reasonable to lower the amplifier gain. The results are shown in Figs. 6–8 where the cooling rate, the amplifier gain and the equilibrium emittance are plotted versus the dynamical gain $g/g_{opt} = Ng$.

It is seen that, as the gain g is lowered, both the amplifier gain and the equilibrium emittance reduce also, but on the other hand, unfortunately, the cooling time

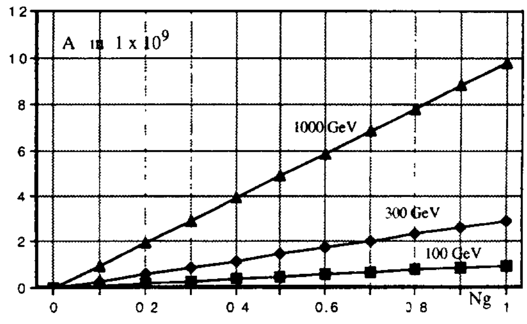

Fig. 7. Amplifier gain vs. dynamical gain Ng.

increases. If the increase is too large then the particle losses would also be too large.

The luminosity depends on the gain g only through the F_{sto} factor. We can chose an expression of the luminosity which shows the explicit dependence on g as follows:

$$L = f(Ng)L_{opt}, \quad (39)$$

where L_{opt} is the luminosity obtained with the parameters shown in Table 1, that is for the optimum gain $g = 1/N$ and

$$f(u)[(2-u)/u]\exp[-k/(2u-u^2)], \quad (40)$$

with

$$k = 2/(\gamma\tau_0\lambda_{opt}) \ll 1. \quad (41)$$

The first factor of Eq. (41) represents the reduction of the betatron emittance and the exponential factor represents the beam loss. It can be seen that the faction $f(Ng)$ has a maximum for $Ng_{max} \sim k/2$ where it takes the value $f_{max} \sim 1/k$. Thus at most the luminosity can be increased by $f_{max} \sim 0.04\gamma$, that is an increase which is proportional with the beam energy as required. The values of Ng_{max} and of f_{max} with the corresponding increased luminosity are also shown at the bottom part of Table 1. The luminosity figures are still well below the desired values.

The only other parameter that can be varied is the bunching frequency f_{bunc}. For the largest realistic bandwidth, this is equivalent to vary the frequency integer

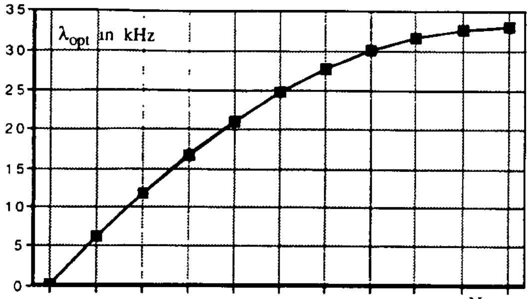

Fig. 6. Cooling rate vs. dynamical gain Ng.

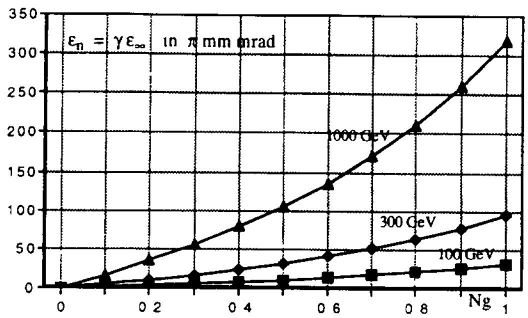

Fig. 8. Equilibrium emittance vs. dynamical gain Ng.

parameter m. If we denote with m_{opt} the value of m used previously, corresponding to $f_{bunc} = 3$ GHz, we can introduce the ratio $\mu = m/m_{opt}$. The expression for the luminosity can then be modified as follows:

$$L = f(Ng, \mu) L_{opt}, \quad (42)$$

where

$$f(u, \mu) = \mu[(2-u)/u] \exp[-\mu k/(2u-u^2)]. \quad (43)$$

The maximum of this function is the one found before and does not depend on the ratio u. That is, the optimum of the luminosity does not depend on the choice of the bunching mode number m, once optimization with respect to the dynamical gain g has been carried out.

9. Conclusions

We have determined that it is indeed feasible that the luminosity of a muon collider scales linearly with the beam energy, as it is required by physics argumentations. Unfortunately, even with the stretching of our imagination, it is seen from our results shown in Table 1 that at the very most only a luminosity of $10^{20}\gamma$ cm^{-2} s^{-1} can be obtained. This is seven orders of magnitude below what it is actually required.

The most important limitation is the effect of thermal noise to the ultimate emittance that can be achieved. This is to be coupled with the requirement on the cooling rate which is to be large compared to the inverse of the beam lifetime. To achieve very fast cooling, a large linear electronic gain is needed, which has also the effect to amplify to a larger level the front-end noise. Moreover a large cooling rate can be obtained only with a few number of particles per bunch. Even by postulating the feasibility of momentum stacking, it is rather difficult to accumulate more than 10^5 particles per bunch.

The comparison of the performance of a muon collider with respect to a proton–antiproton collider is in order. The methodology is essentially the same: both types of particle are produced from a target both need cooling to reduce their dimensions and both are to be compressed longitudinally in bunches. But the antiproton particles have an infinitely long lifetime and it is thus possible to accumulate 10^{10-11} particles per bunch after a long session of stochastic cooling. Moreover it is more convenient for stable particles to operate the collider in a storage mode with multiple passes of the same beams at the collision point.

Our estimates of the performance of stochastic cooling are based on the simple scenario of production, acceleration, cooling and collision we have proposed here. Other scenarios may be possible and we believe that an optimum configuration has still to be searched and is highly desirable. But we also believe that stochastic cooling has to be an integral part of the scheme. Then the following features are to be investigated in more details.

We have seen first that thermal noise at the front-end of the amplifier plays a crucial limiting role to the final beam emittance. We have also seen that this can be better measured with the ratio of Schottky to thermal power, given by Eq. (20). The problem is that as cooling proceeds the beam power reduces whereas the noise signal remains constant. An invention would be highly desirable where the level of noise signal can also be reduced accordingly, as for instance done in momentum cooling with notch filters [9].

The other issue relates to the complete absence of mixing of the particle motion. It is also important to demonstrate that there are ways to regenerate the beam Schottky signal. Several methods have in the meantime been proposed, like the introduction of sextupole to generate a coupling of the transverse motion with the particle momentum error, skew quadrupoles to introduce mixing between the two transverse plane of oscillations, and nonlinear magnetic lenses which cause a dependence of the betatron motion with the amplitude of the motion itself.

These two issues are of paramount importance and are to be investigated carefully if we want to keep the option of a muon collider for the search of the Higgs boson still of some interest.

References

[1] P. Chen and K.T. McDonald, Proc. Workshop on Advanced Accelerator Concepts, AIP Conf. Proc. 279 (1992) 853.

[2] Proc. Mini-Workshop on $\mu^+\mu^-$ colliders: Particle Physics and Design, sponsored by UCLA Center for Advanced Accelerators, Napa Valley, California, Dec. 9–11, 1992.

[3] Proc. Muon Collider Workshop, LA-UR-93-866, LAMPF Auditorium, Feb. 22, 1993.

[4] A.N Skrinsky, Proc. 20th Int. Conf. on High Energy Physics, AIP Conf. Proc. 68 (1980) 1056.

[5] D. Neuffer, Particle Accel. 14 (1983) 75.

[6] R. Noble, Proc. Workshop on Advanced Accelerator Concepts, AIP, Conf. Proc. 279 (1992) 949.

[7] W.A. Barletta and A.M. Sessler, Characteristics of a high energy $\mu^+\mu^-$ collider based on electro-production of muons, LBL-33613 (1993).

[8] A.G. Ruggiero, Proc. Workshop on Advanced Accelerator Concepts, AIP Conf. Proc. 279 (1992) 958.

[9] S. van der Meer, Physics of Particle Accelerators, AIP Conf. Proc. 153, Volume 2 (1987) 1628.

[10] B. Autin, J. Marriner, A. Ruggiero and K. Takayama, IEEE Trans. Nucl. Sci. NS-30 (4) (1983) 2593.