

FR 9.10 239

LAL/RT 90-15
December 1990

Laboratoire de l'Accélérateur Linéaire

BEAM PHASE SPACE AND EMITTANCE

Jean BUON

*Lecture given at the CERN Accelerator Course
General Accelerator Physics Course
Jülich, 17-28 September 1990*

U.E.R
de
l'Université Paris-Sud



Institut National
de Physique Nucléaire
et
de Physique des Particules

Bâtiment 200 - 91405 ORSAY Cedex

BEAM PHASE SPACE AND EMITTANCE

J. Buon

LAL, 91405 Orsay, France

ABSTRACT

The classical and elementary results for canonical phase space, the Liouville theorem and the beam emittance are reviewed. Then, the importance of phase portraits to obtain a geometrical description of motion is emphasized, with examples in accelerator physics. Finally, a statistical point of view is used to define beam emittance, to study its law of approximate conservation and to treat two particular examples.

1. INTRODUCTION

Beam phase space and beam emittance are topics presented in many textbooks. It is not very necessary to repeat here what can be found elsewhere, particularly in previous CERN Accelerator School courses. It is also assumed that the reader is already familiar with the Hamiltonian formalism of mechanics at least at an elementary level.

To refresh the reader's memory and to guide him, the most important classical and elementary results are reviewed in Section 2. No proofs are given and the reader is referred to the literature. In particular, there are several different proofs of the Liouville theorem of which perhaps the shortest and most direct is that reported by Weiss^[1]. Examples of emittance non-conservation are also only listed, without trying to be exhaustive. After this review, the aim is to develop two particular and important topics : the phase portraits and the statistical concept of emittance.

Section 3 is devoted to the phase portraits : the pictures made by all the possible trajectories in phase space. Their importance lies in the geometrical description of motion that they give. They are particularly useful when the phase space dimensionality is low (2 or 3 at most) and when phase trajectories do not cross. The non-linear pendulum, considered in Section 2 is the simplest example. It is also the case of the uncoupled motion of non-interacting particles submitted to time-independent forces. However, in accelerators particles experience guiding and accelerating forces that vary from place to place. The non-crossing property of the phase trajectories can be restored at the price of increasing the dimensionality by one unit. Fortunately, the Poincaré sections make it possible to reduce the dimensionality later without losing the non-crossing property. The usual turn-by-turn stroboscopic view of the betatron motion in a circular accelerator follows these general lines. It is used here to illustrate the procedure. Finally, some examples of phase portraits in accelerator physics are shown.

In Section 4 a statistical point of view is taken to define and study the beam emittance. It is more satisfactory than the traditional emittance definition using a limiting contour in phase space since the latter suffers from arbitrariness of the contour. Moreover, the statistical definition lends itself more easily to calculations. This definition, initially given by P. Lapostolle^[2], is introduced here in the most natural way, as a measure of the spread in phase space of the points representing beam particles. It generalizes in two dimensions the standard deviation that measures the dispersion of points distributed on a line.

With the statistical definition and in the case of a Hamiltonian motion, the conservation of emittance is not a direct consequence of the Liouville theorem as in the usual definition. It is shown that the theorem applies and the emittance is conserved only if the motion is linear. The conclusion is that emittance is conserved, but only approximately, in accelerators where particle motion is made as linear as possible and where other causes of emittance dilution are minimized.

Finally, the emittance growth by scattering, and the emittance reduction by stochastic cooling, are studied in an elementary way to give two examples of calculation using the statistical definition of emittance.

2. SUMMARY OF CLASSICAL AND ELEMENTARY RESULTS

2.1 The canonical phase space

For many physical systems, their evolution after a time t only depends on their state at that time. This is determined by the values of r coordinates q_i ($i = 1, \dots, r$) and their time derivatives \dot{q}_i , the integer r being the number of degrees of freedom.

In the absence of friction forces, the equations of motion can be written in the form of Hamilton :

$$\begin{cases} \dot{q}_i = + \frac{\partial H}{\partial p_i} \\ \dot{p}_i = - \frac{\partial H}{\partial q_i} \end{cases} \quad i = 1, \dots, r \quad (1)$$

where the Hamiltonian $H(q_i, p_i, t)$ is a function of the coordinates q_i , of the conjugate momenta p_i , and of time t . If H does not explicitly depend on the time t , the Hamiltonian H is the total energy of the system that is invariant.

The first set of these equations allows the time derivatives \dot{q}_i to be expressed in terms of the conjugate momenta p_i , or vice versa. The second set allows the second-order differential equations of the motion to be derived, the so-called equations of Lagrange.

The canonical phase space is the $2r$ -dimensional space with the conjugate coordinates q_i, p_i ($i = 1, \dots, r$). The state of the system at a time t is represented by a point $M(t)$ with coordinates $q_i(t), p_i(t)$ in the canonical phase space. As t increases, the representative points $M(t)$ generate a curve, called a phase trajectory.

When the Hamiltonian is time-independent, there is only one phase trajectory originating from any point M in the phase space, apart from where the derivatives $\frac{\partial H}{\partial q_i}, \frac{\partial H}{\partial p_i}$ vanish. Different trajectories cannot cross each other. The motion along a closed trajectory is periodic. If the phase space is two-dimensional, all trajectories, originating from points inside the region limited by a closed trajectory, are bounded by this closed trajectory.

A beam of N interacting particles is a system of $3N$ degrees of freedom if the internal degrees of freedom (like spin) can be neglected. The $3N$ coordinates q_j are the components of the coordinate vectors \mathbf{r}_j ($j = 1, \dots, N$) of each particle. The canonical phase space is $6N$ -dimensional.

When the N particles are identical and without mutual interaction, only the canonical phase space of one particle can be considered. The latter is 6-dimensional. At a time t , the state of the beam is represented by the set of N points $M_j(t)$, where the point M_j corresponds to the particle j ($j = 1, \dots, N$).

The N particles form a beam only if the N points are clustered in a relatively small volume of the 6-dimensional phase space. When N is large, the beam state can be represented by a phase density $f(\vec{q}, \vec{p}, t)$ such that at time t the number dN of representative points in an infinitesimal volume d^3pd^3q is given by :

$$dN = f(\vec{q}, \vec{p}, t) d^3q d^3p. \quad (2)$$

In general, the phase density $f(\vec{q}, \vec{p}, t)$ explicitly depends on the time t as the density does not stay constant at a fixed point $M(\vec{q}, \vec{p})$.

The phase density can be considered as a probability distribution in the phase space, resulting from the lack of information on the initial state of the beam. Very often the phase density can be approximated either by a parabolic function, by a gaussian, or again by a uniform density in an ellipsoid.

Usually, the longitudinal motion along the beam axis is decoupled from the motion in the plane transverse to the beam axis. In that case, the 6-dimensional phase space can be split into a longitudinal phase space (2-dimensional) and a transverse phase space (4-dimensional).

Finally, if the transverse motion can be decomposed into two independent motions along two orthogonal directions, the transverse phase space can be split into two 2-dimensional phase spaces. This is the ultimate reduction of the phase space dimensionality.

Canonical transformations allow the conjugate variables q_i, p_i to be changed for new conjugate variables Q_i, P_i : the canonical phase space (q_i, p_i) then being mapped into a new phase space (Q_i, P_i) . One usual canonical transformation in beam dynamics is the change of the longitudinal coordinates z, p_z for the phase shift $\delta\varphi$ and the energy increment δE , both relative to a synchronous particle.

Limits are set by quantum mechanics on the knowledge of two conjugate variables q_i and p_i . According to the Heisenberg Uncertainty Principle, the product of the uncertainties δq_i and δp_i cannot be made smaller than the Planck constant \hbar (divided by 2π) :

$$\delta q_i \delta p_i \geq \hbar c = 2 \times 10^{-10} \frac{\text{MeV}}{c} \times \text{mm}. \quad (3)$$

This limitation can be expressed by saying that the state of a particle is not exactly represented by a point, but by a small uncertainty volume of the order of \hbar^3 in the 6-dimensional phase space.

Nevertheless, the practical accuracy on the knowledge of position and momentum is normally far away from this quantum limit. Classical mechanics can be used to very good approximation to study particle motion, except when there are interactions at the level of atomic size, like the quantum jumps of electrons emitting photons in a magnetic field (synchrotron radiation).

2.2 The Liouville theorem

The evolution of a physical system can also be considered as a mapping of the phase space onto itself, relating the two representative points $M(t)$ and $M(t')$ at two different times. When this evolution is governed by a Hamiltonian, the particular form of the equations of motion implies that there are quantities left invariant by the mapping.

In particular, the Liouville theorem states that volumes in phase space are invariant.

By mapping from time t to time t' , all the points $M(t)$ in a finite volume V are transformed into the points $M(t')$ that define the transformed volume V' . From the Liouville theorem, the measures of the volumes V and V' are the same. This property applies to the $6N$ -dimensional canonical phase space of a beam of N interacting particles, as well as to the 6-dimensional phase space of identical particles without mutual interactions. It is a property of the mapping that results from the Hamiltonian form of the equations of motion and does not depend on the particles themselves. In the 6-dimensional phase space, the invariance of a volume does not depend on how many particles are represented in it. There may be no particles in that volume.

Nevertheless, the hypersurface limiting a finite volume is not invariant in general. Its geometrical form is modified, but in a way such that the enclosed volume is conserved. For instance if the N points, representing a beam in a 2-dimensional phase space, are enclosed in an ellipse at time t , then at another time t' , the N points are enclosed in a new closed curve, transformed from the ellipse and may be with a different form.

Moreover, volumes in a subspace corresponding to a part of the degrees of freedom are not generally invariant. They are only invariant when these degrees of freedom are uncoupled from the other ones. For instance, when the betatron motions are coupled, areas in horizontal and vertical phase spaces are not conserved.

An immediate consequence of the Liouville theorem and of particle conservation, when beam losses are negligible, is the invariance of the phase density $f(\vec{q}, \vec{p}, t)$ in the 6-dimensional phase space. More precisely, the density at the point $M(t)$ is equal to the density at the transformed point $M(t')$ at another time t' :

$$f(\vec{q}(t), \vec{p}(t), t) = f(\vec{q}(t'), \vec{p}(t'), t'). \quad (4)$$

In other words, the phase density stays constant when running along a phase trajectory as a representative point does.

The motion of the representative points in the 6-dimensional phase space is analogous to the flow of an incompressible fluid. In this picture, the phase density is analogous to the mass density of the fluid. However, in general a fluid is homogeneous with a uniform density, while the phase density may not be uniform and often has a gaussian variation in phase space at a given time. On the other hand, there are many examples illustrating the non-invariance of the phase space volumes for non-Hamiltonian systems. For instance, the phase trajectories of a linear damped oscillator are spirals all converging to the origin. Any finite volume shrinks and ultimately vanishes.

2.3 The beam emittance

Neglecting mutual interactions and coupling between the three coordinates of a

particle, one defines the emittance of each degree of freedom ; horizontal, vertical and longitudinal.

The horizontal (vertical) emittance is usually defined considering the $x, x' = \frac{dx}{ds}$ phase space, instead of the x, p_x canonical phase space (s is the curvilinear abscissa along the reference orbit). One is more interested to know about the slope x' than the transverse momentum p_x .

Representative points of particles with the same betatron amplitude, but with different phases, are on the same ellipse at time t . Homothetic ellipses correspond to different amplitudes.

It is usual to define the emittance ϵ_x as the area (optionally divided by π) of the ellipse containing 95 % of all the particles in its interior (see Fig. 1a). The 95 % proportion is somewhat arbitrary. One could equally use 90 % or 99 %.

The emittance is also the product of the two semi-axes of the ellipse. If the ellipse is upright, its axes coincide with the horizontal and vertical axis. The two semi-axes are half the beam size Δx and the beam divergence $\Delta x'$ respectively (see Fig. 1b).

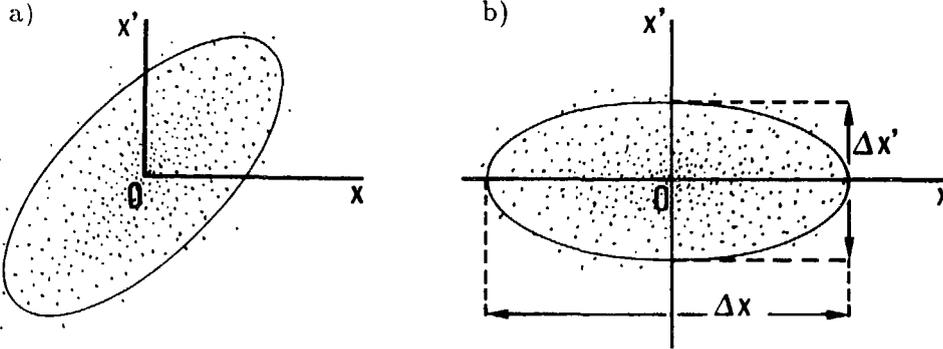


Fig. 1 : A set of points representative of a beam in the 2-dimensional (x, x') phase space

- a) Tilted emittance ellipse.
- b) Upright emittance ellipse.

In a beam transport line with negligible longitudinal magnetic field, the potential vector \vec{A} of the magnetic field is longitudinal. The transverse components of the particle momentum $\vec{p} - e\vec{A}$ are equal to the transverse components of the conjugate momentum \vec{p} . Therefore, the slope x' is proportional to the conjugate momentum p_x :

$$p_x = m\beta\gamma cx' \quad (5)$$

where m is the rest mass of the particles, $\beta = \frac{v}{c}$ and $\gamma = 1/\sqrt{1-\beta^2}$ being the usual relativistic quantities attached to the velocity v of a particle.

In a transport line, or in a ring without acceleration, the energy of a non-radiating

particle is constant, and so is $\beta\gamma$. The Liouville theorem is also valid in the x, x' phase space and the emittance ϵ_x is conserved. Along the line or the ring, the beam passes through focusing or defocusing lenses. The beam size Δx and the beam divergence $\Delta x'$ vary, but the emittance stays constant.

In linear, or circular accelerators, the emittance is not invariant. Instead one defines the so-called normalized emittance :

$$\epsilon_{xN} = \beta\gamma\epsilon_x \quad (6)$$

that refers to the area in the x, p_x canonical phase space. The normalized emittance is conserved during acceleration.

As $\beta\gamma$ increases proportionally to the particle momentum p , the emittance ϵ decreases as $1/p$. It is called "adiabatic damping". An equivalent point of view is to consider that when crossing accelerating cavities, the longitudinal momentum p_y is increased while p_x is not changed. Therefore, the slope $x' = \frac{p_x}{p_y}$ is decreased.

It is worth noting that, due to adiabatic damping, higher-energy accelerators are able to deliver smaller beams to counteract the decrease of the particle interaction cross-section as the energy increases.

An important challenge in accelerator technology is to preserve beam emittance and even to reduce it. In spite of the Liouville theorem, there are many phenomena that may dilute the emittance. A list, not necessarily exhaustive, of phenomena causing non-conservation of horizontal (vertical) beam emittance^[3] includes :

- coupling between degrees of freedom
- intrabeam scattering
- beam-beam scattering
- scattering on residual gas
- multiple scattering through a thin foil
- electron cooling
- stochastic cooling
- laser cooling
- synchrotron radiation emission
- filamentation due to non-linearities
- wake fields
- space charge effects.

The horizontal ϵ_x and vertical ϵ_z emittances are also related to the beam brightness defined as :

$$B = \frac{I}{\pi^2\epsilon_x\epsilon_z} \quad (7)$$

In most applications either emittances or brightness are the main figures of merit of particle beams.

Similarly, the longitudinal emittance measures the product of the energy spread δE by the phase dispersion $\delta\phi$. In the linear and smooth approximation of the synchrotron motion, the trajectories of representative points of the particles are ellipses in the $\delta E, \delta\phi$

plane. The area (divided by π) of the ellipse containing 95 % of the particles is the longitudinal emittance ϵ_s . As the ellipse is now a phase trajectory, the emittance ϵ_s is also half the action J of this trajectory :

$$J = \frac{1}{2\pi} \oint \delta E d(\delta\phi). \quad (8)$$

One can also say that the conservation of the longitudinal emittance is a consequence of the invariance of the action J . However, the longitudinal Hamiltonian is energy-dependent and varies during acceleration in a synchrotron. Consequently, phase trajectories are not exactly ellipses. Nevertheless, when acceleration is sufficiently slow, ellipses are mapped into ellipses of the same area, according to the adiabatic theorem. The action and the longitudinal emittances are still conserved quantities.

3. PHASE PORTRAITS

The usefulness in considering phase space is not just to represent the state of a dynamical system by a point in such a configuration space, and its evolution with time by a phase trajectory. More important is the topological structure of the trajectories in the phase space. All the possible trajectories form a picture, a “phase portrait”, characteristic of the dynamical system, with typical topological features. For instance, circular trajectories around a centre, as in small oscillations of a pendulum, reveal a state of stable equilibrium (an elliptic point in phase space). On the contrary, diverging trajectories looking like hyperbolae reveal an unstable equilibrium (a hyperbolic, or saddle, point). Many other topological aspects, like separatrix, limit cycle, spirals, islands, attractors, chaotic attractors,..., reveal other dynamical behaviours.

As was recognized by H. Poincaré, a phase portrait gives a complete geometrical description of all the possible behaviours of a dynamical system. Nevertheless, it could be argued that the analytical solution of the equations of motion for linear dynamical systems is enough to obtain a complete description of their behaviour. With that point of view, phase portraits would only be useful for non-linear systems for which analytical solutions are scarcely available. Particle accelerators are made to be linear systems and phase portraits would not be very useful for studying particle dynamics in accelerators. However, a real dynamical system is only nearly linear and its stability and long-term evolution are essentially determined by its non-linear aspects, even if they are weak. The importance of phase portraits is very general, even for particle dynamics in accelerators.

3.1 The non-linear pendulum

The equation of motion of an undamped pendulum is :

$$\frac{d^2\theta}{dt^2} + \omega^2 \sin\theta = 0 \quad (9)$$

with $\omega = \sqrt{\frac{g}{\ell}}$, g being the gravitation intensity and ℓ the pendulum length (Fig. 2).

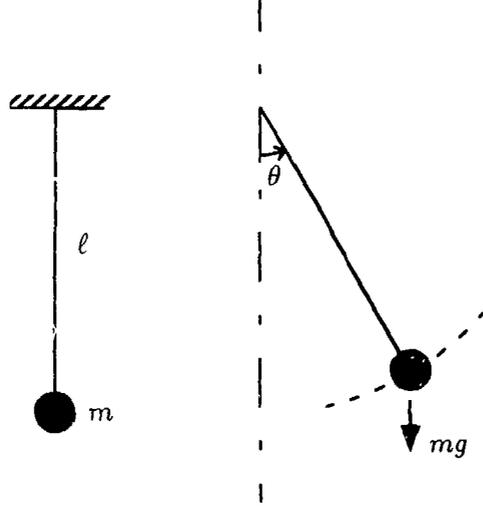


Fig. 2 : A simple and undamped pendulum. The weight of the bob is mg and the length of the rod is ℓ .

Making the approximation of small oscillations, one obtains the linear equation of a harmonic oscillator :

$$\frac{d^2\theta}{dt^2} + \omega^2\theta = 0. \quad (10)$$

The motion of an undamped pendulum gives the classical and simplest example of phase portrait in a two-dimensional phase space. The motion is governed by a Hamiltonian H that is identified with the total energy, sum of the kinetic energy T ($T = \frac{1}{2m\ell^2} p_\theta^2$) and the potential energy V ($V = -mg\ell\cos\theta$) :

$$H(\theta, p_\theta) = T + V \quad (11)$$

where m is the mass of the bob.

The Hamilton equations with the conjugate variables θ, p_θ are :

$$\begin{cases} \frac{d\theta}{dt} = \frac{\partial H}{\partial p_\theta} = \frac{p_\theta}{m\ell^2} \\ \frac{dp_\theta}{dt} = -\frac{\partial H}{\partial \theta} = -mg\ell \sin\theta. \end{cases} \quad (12)$$

These two equations are equivalent to the second-order differential equation :

$$m\ell^2 \frac{d^2\theta}{dt^2} + mg\ell \sin\theta = 0. \quad (13)$$

The canonical phase space is two-dimensional with the conjugate variables θ , p_θ as coordinates. The trajectories in the phase space (Fig. 3a) are the curves of constant energy E :

$$H(\theta, p_\theta) = E. \quad (14)$$

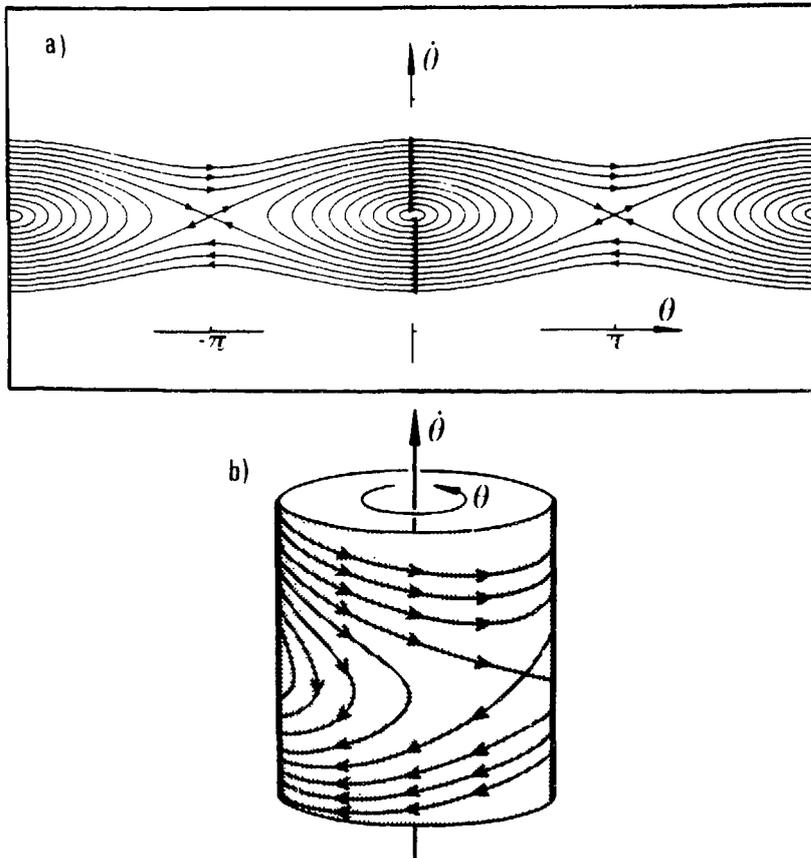


Fig. 3 : Phase trajectories of the non-linear pendulum :
 a) on a plane,
 b) on a cylinder.

The direction of the arrows on these curves are readily obtained by observing that θ increases when p_θ is positive.

For small-amplitude oscillations, the phase trajectories are ellipses centred on the origin 0. It is the point of stable equilibrium with the pendulum hanging vertically without swinging.

The second equilibrium position is where the pendulum is balanced vertically on end. It is an unstable equilibrium state for which $\theta = 180^\circ$ and $\frac{d\theta}{dt} = 0$. It appears as a saddle point in phase space, associated with the crossing of a particular phase trajectory, named the separatrix. This is the curve corresponding to the energy of the saddle point. If the pendulum is launched with this energy, it rotates towards the unstable equilibrium

position. Its velocity decreases when approaching it such that it would take an infinite time to reach the position.

The wavy lines above and below the separatrix correspond to the whirling motion of the pendulum.

As the pendulum motion is periodic, with a cyclic coordinate θ (the state of the pendulum is unchanged by adding 2π), one can also use a cylindrical phase space as illustrated in Fig. 3b.

3.2 The non-crossing trajectory property

The phase portrait is especially clear when phase trajectories do not cross, as illustrated by the example of the pendulum. It is an advantage to have the "non-crossing" condition.

To also include the case of a separatrix crossing itself at a saddle point, one only requires that there is no crossing in finite time. A sufficient condition is that the Hamiltonian of the system should not explicitly depend on time. Therefore, the Hamilton equations show that the tangent to the phase trajectory at any point is uniquely determined (if not an equilibrium position). From this one can infer the uniqueness of the phase trajectory originating from that point.

The topological consequences of this non-crossing property are best illustrated in the case of a two-dimensional phase space. If a phase trajectory is a closed loop, the corresponding motion is periodic. This trajectory divides the plane into two regions, the interior and the exterior. Any trajectory, starting from inside, stays inside for ever and is then bounded.

When the Hamiltonian is time-dependent, one can nevertheless satisfy the non-crossing property at the price of increasing the phase space dimensionality. It is exactly what happens for betatron motion and it can be taken as an example of the procedure.

In the linear approximation, the betatron motion along one direction x is governed by an equation of Hill's type :

$$\frac{d^2 x}{dt^2} + \omega^2(t)x = 0. \quad (15)$$

The angular frequency ω depends on time t as the focusing strength varies along the ring. The Hamiltonian H is :

$$H = \frac{p_x^2}{2m} + \frac{1}{2} m \omega^2(t) x^2. \quad (16)$$

To the canonical variables x and p_x , one adds a third variable $s = ct$ that can be interpreted as the abscissa of a particle with velocity c along its reference orbit. The system of the two Hamilton equations :

$$\begin{aligned} \frac{dx}{dt} &= \frac{p_x}{m} \\ \frac{dp_x}{dt} &= -m\omega^2(s)x \end{aligned} \quad (17)$$

becomes a system of three equations, with the third equation :

$$\frac{ds}{dt} = c. \quad (18)$$

The phase space is now three-dimensional with the coordinates x, p_x and s . There is only one phase trajectory originating from any point (x, p_x, s) . Figure 4 shows an example of a betatron phase trajectory in the (x, p_x, s) phase space.

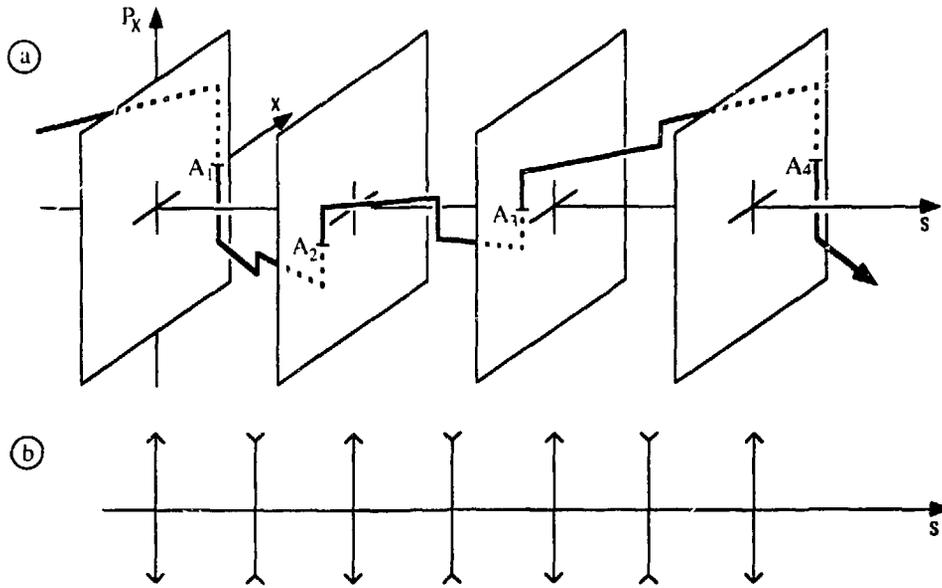


Fig. 4 : A betatron phase trajectory in the (x, p_x, s) phase space (a), corresponding to a periodic FODO transport line in the thin lens approximation (b). The betatron phase advance is 120° per period.

3.3 The Poincaré section

The phase portrait gives a geometrical and global view of motion. However, this view becomes difficult to perceive when the phase space dimensionality is larger than 2 or 3. Nevertheless, the topological structure can still be revealed by only looking at the intersection of the phase trajectories with a hypersurface in phase space. Such a hypersurface is called a Poincaré section.

Again, one can illustrate the procedure with the example of the betatron motion, here in the periodic lattice of a ring. The angular frequency $\omega(t)$ is periodic with the period $T = L/c$ (L is the lattice period).

In the three-dimensional (x, p_x, s) phase space one looks at the intersection of any phase trajectory with the planes x, p_x whenever s is a multiple of the lattice

period L (see Fig. 5). These intersections give a series of x, p_x dots. It can be considered as a stroboscopic technique where a snapshot of the phase trajectory is taken at each period L .

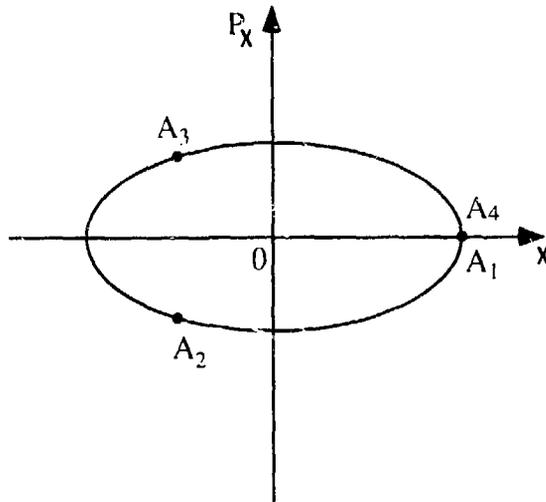


Fig. 5 : Intersection points A_1, A_2, A_3, A_4 of the phase trajectory shown on Fig. 4. A_1 and A_4 coincide as the tune is $1/3$. The ellipse passing through the points is also shown.

The series of x, p_x dots is a mapping of the (x, p_x) plane into itself, called a Poincaré mapping.

According to the Floquet theorem, this Poincaré mapping is a linear mapping. The dots are located on a common ellipse, assuming that the motion is stable. If the phase trajectories have same period as the lattice, any dot is mapped into itself and the series is reduced to a single dot. It corresponds to an integer betatron tune. If the phase trajectories are closed after n periods, the dot series reduces to n points. The betatron tune is a rational π/n .

Due to the periodicity of the lattice, the Poincaré mapping is invariant by a translation of L . It means that the Poincaré mapping is time-independent and satisfies the non-crossing property. The ellipses in the (x, p_x) plane do not cross and one has the phase portrait of a linear oscillator.

Similarly, the finite difference equations of the synchrotron motion between RF cavities also define a Poincaré mapping.

3.4 Examples of phase portraits in accelerator physics

Three examples taken from previous CAS lectures will be briefly mentioned, the reader being referred to the CAS proceedings mentioned for further details.

Resonant extraction uses the action of a sextupole non-linearity near a third-integer resonance^[4,5]. The time-dependence of the Hamiltonian is suppressed by the use of a

rotating reference frame. Figure 6 shows the phase portrait in the J, φ canonical phase space.

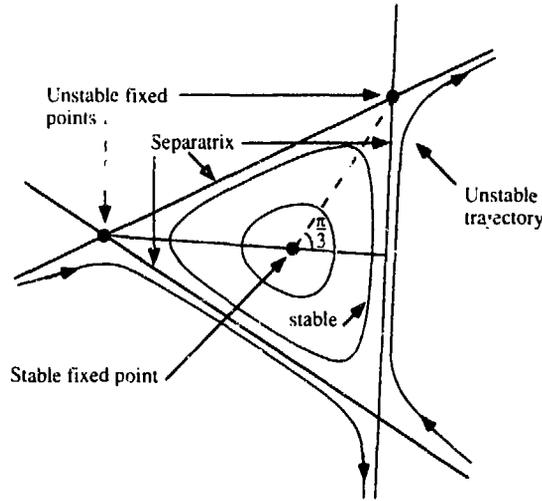


Fig. 6 : Phase portrait of a resonant extraction^[4,5]

The phase portrait (Fig. 7) of the synchrotron motion depends on the RF phase ϕ_s of the synchronous particle^[6,7,8]. The phase space domain inside the separatrix is the RF bucket.

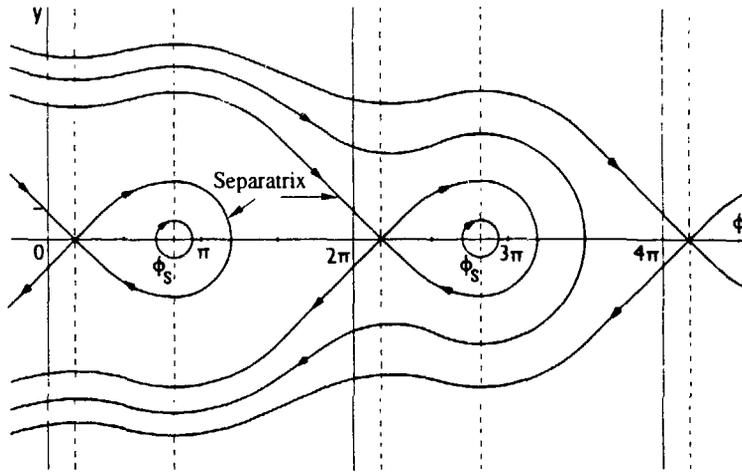


Fig. 7 : Phase portrait of synchrotron motion^[6,7,8] ($\frac{\pi}{2} < \phi_s < \pi$)

The betatron motion, driven by the non-linear beam-beam force, is commonly studied by the technique of the Poincaré sections and mapping. Two-dimensional Poincaré

sections show peculiar phase portraits^[9], with eventually stochastic domains (Fig. 8).

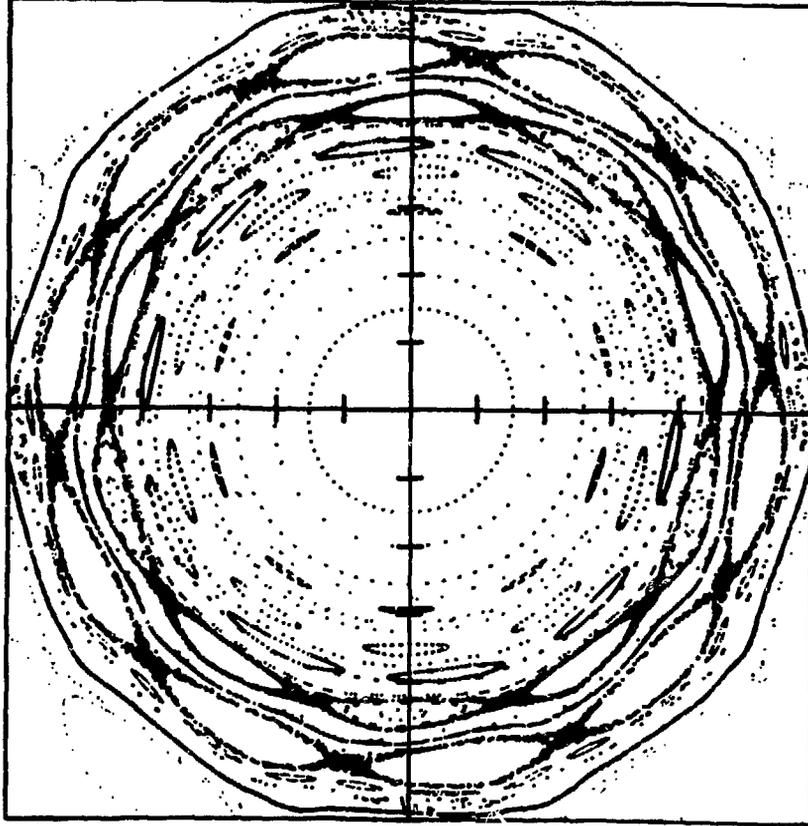


Fig. 8 : Phase portrait of betatron motion perturbed by a beam-beam force^[9]

4. BEAM EMITTANCE : A STATISTICAL POINT OF VIEW

Considering a beam as a statistical set of points in a two-dimensional phase space, the beam emittance is a measure of their spread in that plane (see Fig. 1). The beam emittance is similar to the standard deviation σ that measures the dispersion of a set of points on a line :

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2} \quad (19)$$

where \bar{x} is the average of the abscissa x_i of each point. Similarly, the beam emittance should be a "root mean square" quantity, i.e. a function of the second-order moments of the point distribution.

Within the statistical point of view, the emittance is not related to any contour limiting the area occupied by the points. The usual definition of emittance as the area

of a limiting contour suffers not only from the arbitrariness of the chosen contour, but also from its distortion when time is running. It may become so twisted and bent that its area is no more representative of the spread of the particles. Higher-order moments of the point distribution could also be introduced to parametrize other aspects, like skewness, but the emittance is the main parameter of the distribution.

4.1 The statistical definition

A distribution of x, x' points in phase space can be translated and rotated without changing its spread. One is free to choose an origin and two coordinate axes. It is natural to take the origin at the barycentre of the points, such that the averages \bar{x} and \bar{x}' of their coordinates x_i, x'_i are null. It is also convenient to orientate the axes to minimize (maximize) the sum of their squared distances to these axes. If θ is the polar angle of such an axis, the distance d_i of a point x_i, x'_i to this axis is (see Fig. 9) :

$$d_i = |x'_i \cos\theta - x_i \sin\theta|. \quad (20)$$

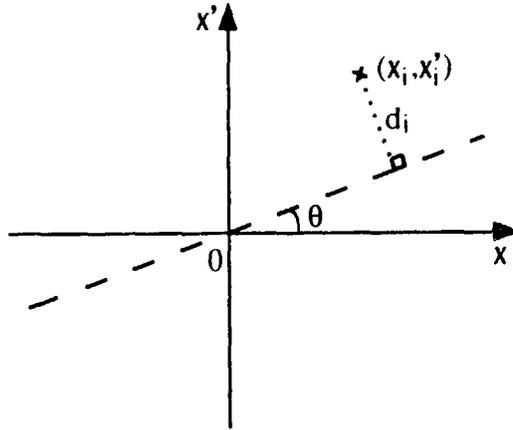


Fig. 9 : The distance d_i to an axis with polar angle θ

The mean square distance $\overline{d^2} = \frac{1}{N} \sum_{i=1}^N d_i^2$ is minimized with respect to the angle θ . The solution of $\frac{\partial \overline{d^2}}{\partial \theta} = 0$ gives :

$$\tan 2\theta = \frac{2 \overline{xx'}}{\overline{x^2} - \overline{x'^2}} \quad (21)$$

where :

$$\overline{x^2} = \frac{1}{N} \sum_{i=1}^N x_i^2, \quad \overline{x'^2} = \frac{1}{N} \sum_{i=1}^N x_i'^2 \quad \text{and} \quad \overline{xx'} = \frac{1}{N} \sum_{i=1}^N x_i x_i' \quad (22)$$

are the central second-order moments of the point distribution. It can be seen that the quantity to minimize is not the same as the one used to find the regression line. In general, the axis will not coincide with the latter. In fact, two solutions θ and $\theta + \frac{\pi}{2}$ giving two orthogonal axes have been found, one with a minimum value $(\overline{d^2})_1$, the other with a maximum value $(\overline{d^2})_2$ of $\overline{d^2}$:

$$(\overline{d^2})_{1,2} = \frac{1}{2} \left(\overline{x^2} + \overline{x'^2} \pm \frac{2 \overline{xx'}}{\sin 2\theta} \right). \quad (23)$$

$(\overline{d^2})_1$ and $(\overline{d^2})_2$ measure the dispersion of the points in the direction perpendicular to each axis. It is then natural to define the emittance ϵ , i.e. the spread of the distribution, as :

$$\epsilon = 2\sqrt{(\overline{d^2})_1} \cdot 2\sqrt{(\overline{d^2})_2}. \quad (24)$$

The expression so obtained :

$$\epsilon = 4\sqrt{\overline{x^2} \cdot \overline{x'^2} - (\overline{xx'})^2}. \quad (25)$$

was first given by P. Lapostolle^[2], and is sometimes called either "effective emittance" or "r.m.s. emittance". The factor 4 in front of the square root is optional. The expression in the square root (in fact always positive) is also the determinant of the covariance matrix :

$$\epsilon = 4\sqrt{\det \begin{pmatrix} \overline{x^2} & \overline{xx'} \\ \overline{xx'} & \overline{x'^2} \end{pmatrix}}. \quad (26)$$

With the use of the standard deviations σ_x , $\sigma_{x'}$ and the correlation coefficient r :

$$r = \frac{\overline{xx'}}{\sqrt{\overline{x^2} \overline{x'^2}}}, \quad (27)$$

the expression of the emittance can also be written as :

$$\epsilon = 4 \sigma_x \sigma_{x'} \sqrt{1 - r^2} \quad (28)$$

and the latter expression is still valid when the mean values \bar{x} and $\bar{x'}$ do not vanish.

Now, a straightforward algebraic manipulation produces :

$$\epsilon = \frac{2\sqrt{2}}{N} \sqrt{\sum_{i=1}^N \sum_{j=1}^N (x_i x'_j - x_j x'_i)^2}. \quad (29)$$

The expression in the square root is just the sum of the squared areas A_{ij}^2 of all the

triangles formed by any couple of points M_i , M_j and the origin 0 (ignoring a factor 2) :

$$\epsilon = \frac{4}{N} \sqrt{2 \sum_{i=1}^N \sum_{j=1}^N A_{ij}^2}. \quad (30)$$

(Caution : each M_i , M_j couple enters twice in the double sum.)

Thus the emittance can be considered as a statistical area : the r.m.s. area of the triangles. One could have taken this definition a priori as well. Here again, this statistical area is not the area of a limiting contour, but the measure of the spread of the points around their barycentre.

This statistical definition of the emittance leads to some differences with the usual definition with a contour. For instance, both definitions give a null emittance to a distribution on a straight line in the phase space. However, if the line has a curved shape, the usual emittance is still null (using a non-elliptic contour in this case), while the statistical one is not (Fig. 10).

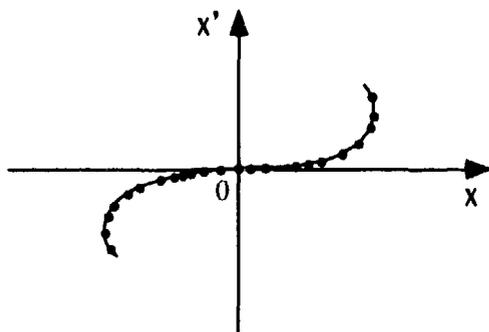


Fig. 10 : Distribution of points on a curved line

This example also shows that the emittance is not unambiguously related to the entropy. The latter is defined as the logarithm of the area occupied by the points in the two-dimensional phase space. Again in this example, the area is null for a curved line, while the statistical emittance is not.

Care is also needed with the tail of a point distribution when considering the statistical emittance ϵ , as well as the standard deviation. Points at very large distance from the barycentre give large contribution to ϵ . If there is a long but small tail, the statistical emittance overestimates the spread of the distribution.

In the case of a long tail, one can also extend to two dimensions the usual alternative to measure the spread of a one-dimensional distribution. Instead of the standard deviation, the spread is better represented by the full width at half maximum (FWHM) of the distribution. Figure 11a shows a one-dimensional gaussian distribution with a standard deviation 1 on which is superimposed a 10 % gaussian tail with a standard deviation 10. The tail increases the overall standard deviation by 29 %, while the FWHM

is only 7 % larger. Similarly, the emittance can be represented as the area limited by the contour at half-maximum. Figure 11b shows a bi-dimensional gaussian distribution with standard deviations $\sigma_x = 2$ and $\sigma_{x'} = 1$. A 10 % gaussian tail, with standard deviations $\sigma_x = \sigma_{x'} = 10$, is superimposed.

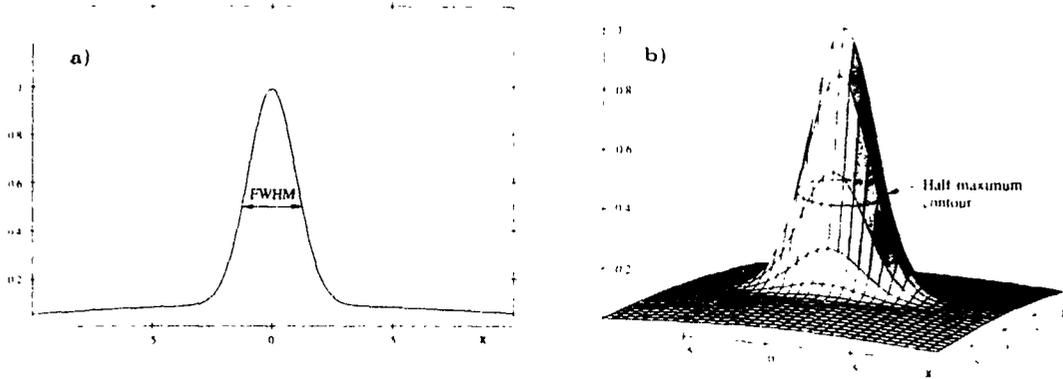


Fig. 11 : a) Full width at half-maximum of a one-dimensional distribution, b) Contour at half-maximum of a two-dimensional distribution.

- Exercises :
- 1) Show that the emittance of three points at the tips of an equilateral triangle, inscribed in a circle of radius R is : $\epsilon = 2R^2$.
 - 2) Show that the emittance of a uniform distribution in a circle of radius R is $\epsilon = R^2$.
 - 3) A distribution of points has a gaussian density :

$$f(x, x') = \frac{1}{2\pi\sigma_x\sigma_{x'}} e^{-\frac{x^2}{2\sigma_x^2} - \frac{x'^2}{2\sigma_{x'}^2}}.$$

Show that the probability p that a point lies outside the ellipse :

$$\frac{x^2}{2\sigma_x^2} + \frac{x'^2}{2\sigma_{x'}^2} = c$$

is $p = e^{-c}$. Show that the emittance ϵ_0 , traditionally defined as the area (divided by π) of the ellipse containing 95 % of the points, is $\epsilon_0 = 6 \sigma_x \sigma_{x'}$ and that the ratio of the r.m.s. emittance ϵ and of the traditional emittance ϵ_0 is $\frac{\epsilon}{\epsilon_0} = \frac{2}{3}$.

4.2 Emittance conservation

Consider now the time evolution of the beam emittance ϵ . The particle motion is assumed to be Hamiltonian, not significantly coupled and in general non-linear :

$$H(x, x') = \frac{x'^2}{2} + f(x). \quad (31)$$

The Hamilton equations of motion, with respect to the independent variable s , are :

$$\begin{aligned}\frac{dx}{ds} &= x' \\ \frac{dx'}{ds} &= -\frac{\partial H}{\partial x}.\end{aligned}\tag{32}$$

The equations of motion of the barycentre (\bar{x}, \bar{x}') are :

$$\begin{aligned}\frac{d\bar{x}}{ds} &= \bar{x}' \\ \frac{d\bar{x}'}{ds} &= -\frac{\partial \bar{H}}{\partial \bar{x}}.\end{aligned}\tag{33}$$

If the motion is linear, $\frac{\partial \bar{H}}{\partial \bar{x}}$ is proportional to \bar{x} . Then the barycentre stays at rest ($\frac{d\bar{x}}{ds} = \frac{d\bar{x}'}{ds} = 0$) if it coincides with the origin ($\bar{x} = \bar{x}' = 0$) at some time.

However, if the motion is non-linear, the time derivative of \bar{x}' involves higher-order moments of the coordinate x , and generally does not vanish. The barycentre moves in the phase space.

Moreover, the R.H.S. of the second equation of motion $\frac{\partial \bar{H}}{\partial \bar{x}}$ is not in general equal to $\frac{\partial H}{\partial x}(\bar{x})$. It means that in the non-linear case, the barycentre does not generally move according to the same law as the representative points. Even if this was so, the straight lines joining a couple of points M_i, M_j and the barycentre O are not mapped into straight lines, and the triangle OM_iM_j is not mapped into another triangle, except in the linear case. Therefore, the Liouville theorem does not imply that the area A_{ij} of the triangles OM_iM_j is conserved. The emittance here defined as a sum of areas A_{ij} is not conserved except when the motion is linear. That is in contrast with the conservation of the traditionally defined emittance due to the Liouville theorem.

To obtain this result analytically, one calculates the derivative $\frac{d\epsilon}{ds}$ of the emittance. Here, one cannot in general assume $\bar{x} = \bar{x}' = 0$. One must replace \bar{x}^2, \bar{x}'^2 and $\bar{x}\bar{x}'$ by the central second-order moments μ_{20}, μ_{02} and μ_{11} respectively in the expression of the emittance :

$$\epsilon = 4 \sqrt{\mu_{20} \mu_{02} - \mu_{11}^2}\tag{34}$$

where :

$$\begin{aligned}\mu_{20} &= \sigma_x^2 = \overline{(x - \bar{x})^2} = \overline{x^2} - \bar{x}^2 \\ \mu_{02} &= \sigma_{x'}^2 = \overline{(x' - \bar{x}')^2} = \overline{x'^2} - \bar{x}'^2 \\ \mu_{11} &= r\sigma_x\sigma_{x'} = \overline{(x - \bar{x})(x' - \bar{x}')} = \overline{xx'} - \bar{x} \cdot \bar{x}'.\end{aligned}\tag{35}$$

A straightforward calculation gives the derivatives of the moments :

$$\begin{aligned}\frac{d\mu_{20}}{ds} &= 2\mu_{11} \\ \frac{d\mu_{02}}{ds} &= -2\mu_{01H} \\ \frac{d\mu_{11}}{ds} &= \mu_{02} - \mu_{10H}\end{aligned}\tag{36}$$

where :

$$\begin{aligned}\mu_{10H} &= x \frac{\overline{\partial H}}{\partial x} - \bar{x} \cdot \frac{\overline{\partial H}}{\partial x} \\ \mu_{01H} &= x' \frac{\overline{\partial H}}{\partial x} - \bar{x}' \cdot \frac{\overline{\partial H}}{\partial x}.\end{aligned}\tag{37}$$

With these relations, the derivative $\frac{d\epsilon}{ds}$ of the emittance is :

$$\frac{d\epsilon}{ds} = \frac{16}{\epsilon} (\mu_{11} \mu_{10H} - \mu_{20} \mu_{01H}).\tag{38}$$

This equation has been derived in a different way by I. Hofmann^[10].

In the case of a linear motion :

$$\frac{\partial H}{\partial x} = ax, \quad \frac{\overline{\partial H}}{\partial x} = a\bar{x},\tag{39}$$

one has :

$$\mu_{10H} = a\mu_{20} \quad \text{and} \quad \mu_{01H} = a\mu_{11}.$$

The derivative $\frac{d\epsilon}{ds}$ vanishes and the emittance is conserved.

When the motion is non-linear, $\frac{d\epsilon}{ds}$ cannot vanish for all s , and the emittance is not conserved. In practical situations, one generally observes an increase of the emittance through the process of filamentation. However, non-linearities are harmful to the stability of motion. They are maintained at the smallest possible level in particle accelerators. Moreover, other causes of emittance dilution (coupling, scattering, radiation,...) are also minimized.

It can be concluded that in accelerators the beam emittance is nearly conserved, but not exactly.

Exercise : Show that, assuming a single even multipole term in the Hamiltonian and neglecting the coupling with the other transverse coordinate z :

$$H = \frac{x'^2}{2} + a_{2n} \frac{x^{2n}}{2n}, \quad (n > 1)$$

the barycentre stays at rest if it coincides with the origin 0 at an initial time and if the point distribution is symmetrical with respect to the origin. Prove without any calculation that the emittance is not conserved.

4.3 Emittance growth by scattering

The statistical definition of beam emittance allows the emittance growth by scattering to be calculated very simply.

Consider a beam passing through a thin foil where particles are scattered at random. Considering the horizontal direction x only, when a particle is scattered, its position x is not changed, but its horizontal slope is changed from x' to $x' + \delta x'$. The change $\delta x'$ is independent of x and x' and averages to zero. The change of $\overline{x'^2}$ and $\overline{xx'}$ are :

$$\begin{aligned}\Delta \overline{x'^2} &= \overline{(x' + \delta x')^2} - \overline{x'^2} = \overline{\delta x'^2} \\ \Delta \overline{xx'} &= \overline{x(x' + \delta x')} - \overline{xx'} = 0.\end{aligned}\tag{40}$$

Finally, according to the statistical definition, the change of the emittance is simply given by :

$$\Delta \epsilon^2 = 16 \overline{x^2 \delta p^2}.\tag{41}$$

A similar use of the statistical definition for studying emittance growth by radiation fluctuations can be found in a paper by M. Sands^[11].

4.4 Stochastic cooling

The statistical definition of beam emittance is particularly useful in understanding the principle of stochastic cooling which aims to reduce the statistical spread of particles in phase space. It is based on a non-Hamiltonian process and one cannot claim any violation of the Liouville theorem in spite of an apparent decrease of a phase space area.

The principle of stochastic cooling can be described as follows^[12] : A pick-up electrode makes it possible to select a small sample of n particles among the N particles of a beam ($n \ll N$), and to measure their mean displacement \bar{x} , i.e. the x -barycentre of the sample. The latter may be finite, even if the x -barycentre of the beam is vanishing. The signal on the electrode is proportional to \bar{x} and is used to apply a correction signal to another electrode such that the x -barycentre of the sample is reduced, eventually down to zero if the correction is complete. This procedure is repeated many times on different samples of particles from the same beam, so reducing the beam emittance.

In order to understand the mechanism of stochastic cooling, the simplest way is to consider samples of only one particle ($n = 1$) whose position x_0 becomes zero after correction. This correction modifies the second-order momenta of the N particles :

$$\overline{x^2} \rightarrow \overline{x^2} - \frac{x_0^2}{N}, \quad \overline{x'^2} \rightarrow \overline{x'^2} \quad \text{and} \quad \overline{xx'} \rightarrow \overline{xx'} - \frac{x_0 x_0'}{N}.\tag{42}$$

The change of the emittance is :

$$\Delta\epsilon^2 = 16 \left[\left(\overline{x^2} - \frac{x_0^2}{N} \right) \overline{x'^2} - \left(\overline{xx'} - \frac{x_0 x_0'}{N} \right)^2 - \overline{x^2} \cdot \overline{x'^2} + \overline{xx'^2} \right] \quad (43)$$

$$\Delta\epsilon^2 = -\frac{16}{N} \left[x_0^2 \overline{x'^2} - 2 x_0 x_0' \overline{xx'} + \frac{x_0^2 x_0'^2}{N} \right].$$

As the correction procedure is repeated on many samples, the final change of the emittance is expected to be given by the preceding expression averaged over all the possible x_0, p_0 values. One has just to replace x_0^2 and $x_0 p_0$ by their respective averaged values $\overline{x^2}$ and \overline{xp} . Finally, neglecting the $\frac{1}{N}$ term in the bracket, one obtains:

$$\Delta\epsilon^2 = -\frac{16}{N} \left[\overline{x^2} \cdot \overline{x'^2} - 2 \overline{xx'^2} \right] \quad (44)$$

or :

$$\Delta\epsilon^2 = -\frac{16}{N} \overline{x^2} \cdot \overline{x'^2} (1 - 2r^2). \quad (45)$$

The emittance is reduced only if $r < \sqrt{\frac{1}{2}}$. It is easily seen that the emittance would be increased for larger values of r , in particular when x and p are linearly dependent ($r = 1$). The correction is also most efficient when x and p are not correlated ($r = 0$). In fact, the correction is incomplete and x_0 is only replaced by $(1 - g)x_0$ where g is the gain ($0 < g < 1$). The final expected change $\Delta\epsilon^2$ becomes (the proof is left as an exercise) :

$$\Delta\epsilon^2 = -\frac{16}{N} \overline{x^2} \cdot \overline{x'^2} g(2 - g - 2r^2). \quad (46)$$

In practice, the gain is small and the emittance is reduced, except if r is very close to 1.

REFERENCES

- [1] M. Weiss, CERN Accelerator School Proceedings (Editor : S. Turner)
CERN 87-10 p 162-163 (1987)
- [2] P.M. Lapostolle, IEEE Trans. Nucl. Sci. NS-18, No 3, 1101 (1971)
- [3] A.H. Sørensen, CERN Accelerator School Proceedings (Editor : S. Turner)
CERN 89-05 p 18-36 (1989)
- [4] G. H. Rees, CERN Accelerator School Proceedings (Editors : P. Bryant, S. Turner)
CERN 85-19; vol. II p 346 (1984)
- [5] E. N. Wilson, CERN Accelerator School Proceedings (Editor : S. Turner)
CERN 87-03 ; vol. I p 41 (1987)
- [6] J. Le Duff, CERN Accelerator School Proceedings (Editor : P. Bryant, S. Turner)
CERN 85-19 ; vol. I p 125 (1984)

- [7] G. Dôme, CERN Accelerator School Proceedings (Editor : S. Turner)
CERN 87-03 ; vol. I p 110 (1987)
- [8] E. Ciapala, CERN Accelerator School Proceedings (Editor : P. Bryant, S. Turner)
CERN 85-19 ; vol. I p 195 (1984)
- [9] L. R. Evans & J. Gareyte, CERN Accelerator School Proceedings
(Editor : S. Turner) CERN 87-03 ; vol. I p 159 (1987)
- [10] I. Hofmann, CERN Accelerator School Proceedings (Editor : S. Turner)
CERN 87-03 ; vol. I p 337-338 (1987)
- [11] M. Sands, SLAC Report, SLAC/AP-47 (1985).
- [12] S.P. Møller, CERN Accelerator School Proceedings (Editor : S. Turner)
CERN 89-05 p 1-17 (1989), and
D. Möhl, these proceedings.