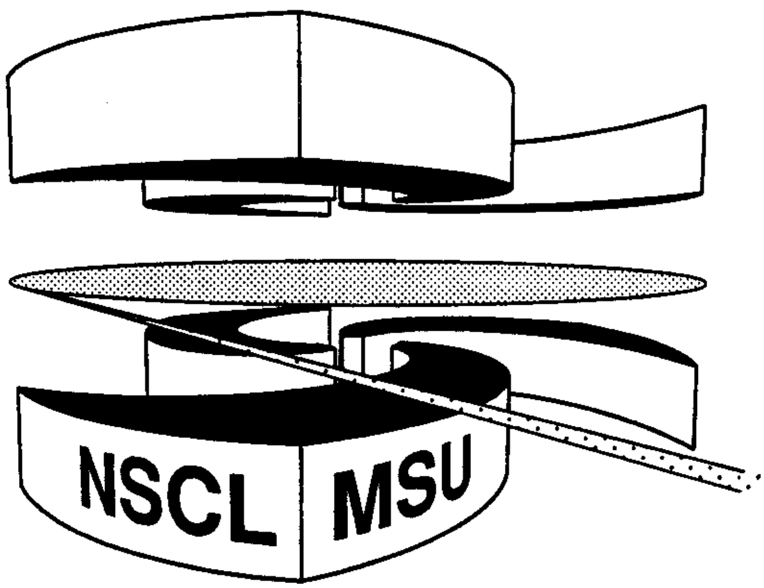
Michigan State University

National Superconducting Cyclotron Laboratory

LINEAR BEAM THEORY

MARTIN BERZ

compiled by WEISHI WAN



MSUCL-814 **DECEMBER 1991**

MSUCL-814
December, 1991

Linear Beam Theory

Martin Berz
compiled by Weishi Wan

December 23, 1991

Abstract

These lectures form a brief introduction into the theory of first order matrix optics or linear optics. Important glass optical and particle optical elements are discussed and used for the design of various simple instruments. The phase space picture of dynamics is discussed. The behaviour of repetitive linear systems (The Courant Snyder Theory) is discussed.

Contents

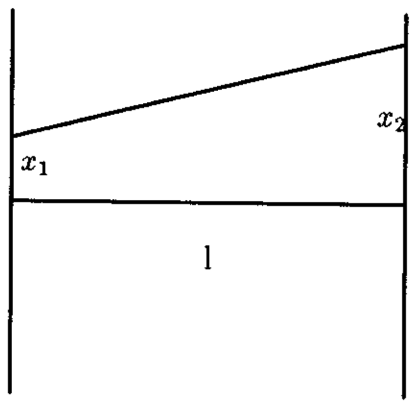


Figure 1: The Drift

1 Matrix Theory of Glass Optics

In this section we will provide an introduction to matrix optics as applied to glass optical elements. It will turn out that the matrix formulation provides a particularly simple and elegant way to describe most of the phenomena in geometric optics.

1.1 The Drift

The simplest part of glass optical elements is a region which doesn't contain any material, the drift. If we denote by x the position of a ray and by m its slope, then the final values x_2 and m_2 after a drift of length l can be connected very simply to the initial values x_1 and m_1:

$$x_2 = x_1 + m_1 \cdot l \tag{1}$$
$$m_2 = m_1 \tag{2}$$

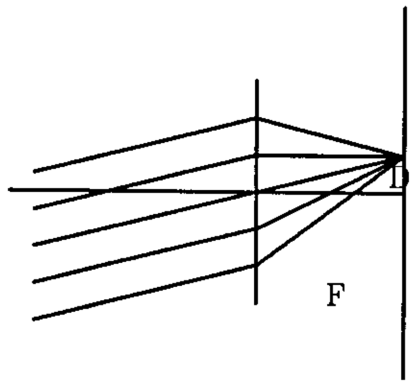


Figure 2: The thin lens

This can obviously be written in a matrix form:

$$\begin{pmatrix} x_2 \\ m_2 \end{pmatrix} = \begin{pmatrix} 1 & l \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ m_1 \end{pmatrix} \tag{3}$$

For the later discussion it will prove important to note that the matrix $\begin{pmatrix} 1 & l \\ 0 & 1 \end{pmatrix}$ depends only on the characteristic properties of the element, which here is the length. On the other hand, the vector $\begin{pmatrix} x_1 \\ m_1 \end{pmatrix}$ depends only on the parameters of the ray. Altogether, a drift performs a linear transformation in x, m space. Note that the determinant of the drift matrix is unity.

1.2 The Thin Lens

Besides empty space, glass optical devices contain lenses that change the direction of the light ray. We are here primarily interested in the thin lens, which is characterized by the following facts:

1. Positions are not changed, but directions are

2. Any bundle of parallel light is unified in one point a distance f after the lens.

3. A ray lighting the center of the lens goes straight through.

The quantity f that describes the lens is called the focal length. Let us now consider a ray going through the lens; from the picture we read

$$
\begin{aligned}
x_2 &= x_1 && (4)\\
D &= f \cdot m_1 && (5)\\
x_1 + m_2 \cdot f &= D = f \cdot m_1 && (6)
\end{aligned}
$$

From which we infer

$$
\begin{aligned}
x_2 &= x_1 && (7)\\
m_2 &= -\frac{x_1}{f} + m_1 && (8)
\end{aligned}
$$

This relationship can again be written in matrix form:

$$
\begin{pmatrix} x_2 \\ m_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -\frac{1}{f} & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ m_1 \end{pmatrix} \qquad (9)
$$

As in the case of the drift, the matrix $\begin{pmatrix} 1 & 0 \\ -\frac{1}{f} & 1 \end{pmatrix}$ depends only on the lens, whereas the vector $\begin{pmatrix} x_1 \\ m_1 \end{pmatrix}$ depends on the ray.

The simple thin lens we have discussed here represents quite an approximation for several reasons. First of all, any real lens performs a refraction at two different surfaces, so positions do change as one goes through the lens. Furthermore, for most lenses it is not really true that parallel rays all meet at a point a distance f behind the lens. This is connected to the fact that lenses are usually ground with spherical surfaces because anything else is technically difficult. Furthermore, the glass has dispersion and so different colors are affected differently. We note however that Snell's law still allows to determine the true transfer map of a lens in a rather straightforward way. It is important to note, however, that this transfer map will no longer be linear.

1.3 Combinations of Elements

One of the key advantages of the matrix formulation of linear optics is that it is very simple to compute the matrix describing a system that is composed of many pieces. Indeed, if M_1 through M_n are the matrices for the subsystems, then because of the associativity of matrix multiplication, we obtain for the ray after the last subsystem:

$$\begin{pmatrix} x_{n+1} \\ m_{n+1} \end{pmatrix} = M_n \left(\cdots \left(M_1 \begin{pmatrix} x_1 \\ m_1 \end{pmatrix} \right) \cdots \right) \tag{10}$$

$$= (M_n \cdot \cdots \cdot M_1) \begin{pmatrix} x_1 \\ m_1 \end{pmatrix} \tag{11}$$

So we have shown that the matrix of a combined system equals to product of matrices of subsystems. Since especially on computers it is very simple to multiply matrices, this is the method of choice for the basic design of optical systems.

1.4 Two Drifts

In this section, we want to study what happens if the matrices of two drifts are multiplied:

$$\begin{pmatrix} 1 & l_2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & l_1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & l_1 + l_2 \\ 0 & 1 \end{pmatrix} \tag{12}$$

This equals the matrix of a drift of length $l_1 + l_2$ (no surprise).

1.5 Two Lenses

The combination of two lenses is a little more interesting:

$$\begin{pmatrix} 1 & 0 \\ -\frac{1}{f_2} & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -\frac{1}{f_1} & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -\frac{1}{f_1} - \frac{1}{f_2} & 1 \end{pmatrix} \tag{13}$$

So the combination of two lenses provides the same effect as one lens with focus length f, where $1/f = 1/f_1 + 1/f_2$. This is of course a famous law of optics, the derivation of which is trivial in the matrix concept and somewhat more involved in the standard geometric method.

2 Systems with Special Properties

In this section we want to apply the matrix techniques to the study of certain special categories of systems. In particular, we associate certain fundamental properties of systems with properties of the matrix. We begin with the imaging systems.

2.1 Imaging (Point-to-Point, · ·) Systems

Imaging systems are perhaps the most important systems in optics, and they deserve some special thought. Suppose we study the action of a slide projector. At one end of the projector, light is sent through the slide. Suppose the slide shows a tree in the fall with one last green leaf. The image of this tree is to appear on the screen, and the green leaf is to appear at one particular location. This requires that all light going through the green spot on the slide in various directions has to be re-united at one spot on the screen.

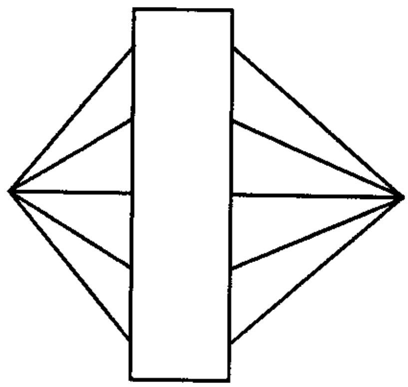


Figure 3: Point-to-point or imaging system

This means that the final position of a ray is independent of its initial angle and only depends on the initial position. In terms of transfer matrices

$$M = \begin{pmatrix} (x \mid x) & (x \mid m) \\ (m \mid x) & (m \mid m) \end{pmatrix} \tag{14}$$

this means that the element $(x \mid m)$ has to vanish. Obviously the element (x, x) also has an important interpretation: it is the magnification.

2.2 Drift-Lens-Drift (DLD) Systems

Let us sandwich a lens between two drifts:

$$\begin{pmatrix} 1 & l_2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -\frac{1}{f} & 1 \end{pmatrix} \begin{pmatrix} 1 & l_1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & l_2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & l_1 \\ -\frac{1}{f} & 1 - \frac{l_1}{f} \end{pmatrix} \tag{15}$$

$$= \begin{pmatrix} 1 - \frac{l_2}{f} & l_1 + l_2 - \frac{l_1 l_2}{f} \\ -\frac{1}{f} & 1 - \frac{l_1}{f} \end{pmatrix} \tag{16}$$

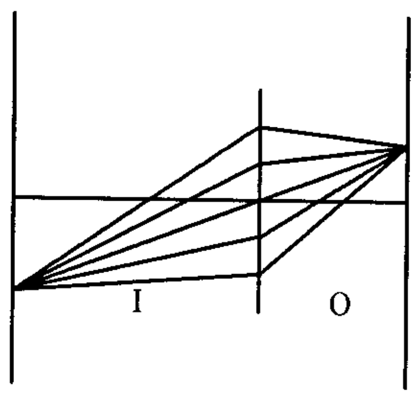


Figure 4: Illustration of a DLD system

If such a system is supposed to be imaging, we have to satisfy $l_1 + l_2 - l_1 l_2 / f = 0$. That is

$$\frac{1}{l_1} + \frac{1}{l_2} = \frac{1}{f} \tag{17}$$

In this case, the magnification is $M = 1 - l_2/f = -l_2/l_1$.

This principle is used in several different devices. In the slide projector, l_1 is very small and l_2 is very large, providing a large magnification. Probably the most important imaging system is the eye. Here the situation is just the opposite: l_1 is large and l_2 is small, allowing for large things to be mapped on the small retina of the eye.

2.3 Combination of Two Imaging Systems

It is interesting to study the combination of two imaging systems:

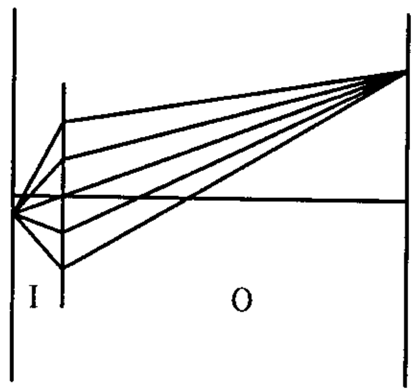


Figure 5: The Slide Projector

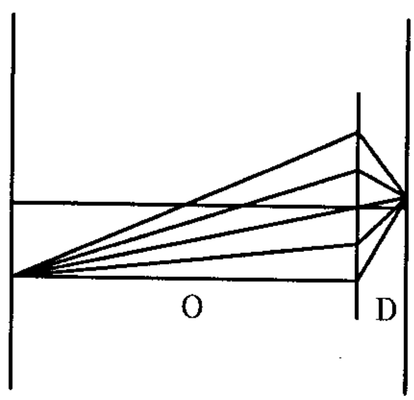


Figure 6: Imaging of The Eye

$$\begin{pmatrix} (x \mid x)_2 & 0 \\ (m \mid x)_2 & (m \mid m)_2 \end{pmatrix} \cdot \begin{pmatrix} (x \mid x)_1 & 0 \\ (m \mid x)_1 & (m \mid m)_1 \end{pmatrix} \tag{18}$$

$$= \begin{pmatrix} (x \mid x)_2(x \mid x)_2 & 0 \\ (m \mid x)_2(x \mid x)_1 + (m \mid m)_2(m \mid x)_1 & (m \mid m)_2(m \mid m)_1 \end{pmatrix} \tag{19}$$

As is to be expected, the total system is again imaging, and the magnification is just the product of the individual magnifications.

$$M_{total} = M_1 \cdot M_2 \tag{20}$$

2.4 Parallel-to-Point ($\| \cdot$) Systems

As we saw above, the human eye observing a nearby object is one of the prime examples of an imaging system. But what happens if the eye looks at things farther and farther away, in particular at the stars, a pastime of the human race and scientists for eternity? The length of the first drift l_1 becomes larger and larger, and for all practical purposes the light coming from one star reaches the eye as a parallel bundle. So what the eye is to interpret now is the angle under which the light comes in, and hence the position on the retina should depend only on the initial angle, but not on the initial position at which the light strikes the eye.

This requires that $(x|x) = 0$. If we look at the eye as a DLD system, this requires $1 - l_2/f = 0 \Leftrightarrow l_2 = f$, l_1 arbitrary. Thus the retina has to be exactly at the focal length; almost as important is that the distance to the object is arbitrary since we cannot change our distance to the stars significantly.

2.5 Point-to-Parallel ($\cdot \|$) Systems

Another important class of systems is the point to parallel systems. In point to parallel systems, the final slope does depend only on the initial position, but not on the initial slope. So we have $(m|m) = 0$.

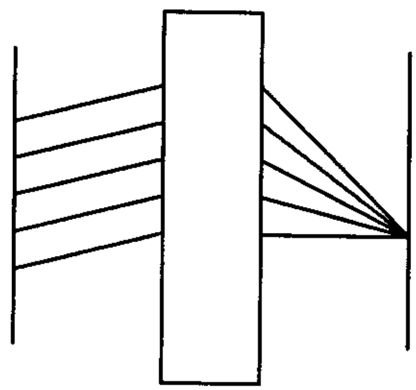


Figure 7: Parallel-to-point system

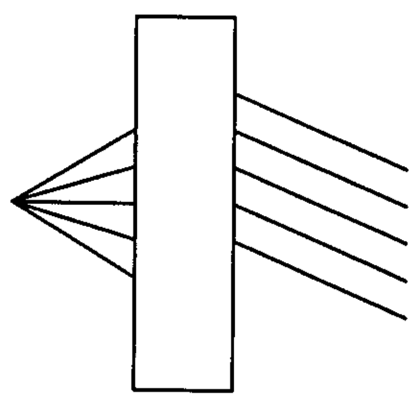


Figure 8: A Point-to-parallel system

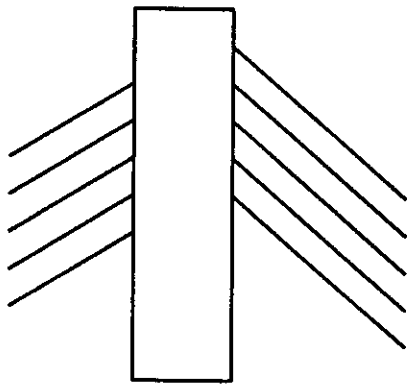


Figure 9: Parallel-to-parallel system

From the transfer matrices, it follows rather directly that the combination of a point to parallel and a parallel to point system forms a point to point system. Using the relaxed eye as the parallel to point system, we can thus build a microscope by putting a suitable point to parallel system in front of the eye. It is interesting to see how the lengths in a point to parallel system have to be chosen; we obtain $(m|m) = 0 \Leftrightarrow 1 - l_1/f = 0 \Leftrightarrow l_1 = f$, l_2 arbitrary. The first part is as expected; the latter part is helpful because it allows the eye to move with respect to the microscope.

2.6 Parallel-to-Parallel ($\| \; \|$) Systems

The final important system is the parallel to parallel system. By putting it between the eye and the stars, a magnification of angles can be achieved. This is the principle of the telescope.

The system has to be such that the final slope depends on the initial slope, but not on the initial x, which requires $(m|x) = 0$. If we try to achieve this with a DLD system, then we have to satisfy $-1/f = 0$, which is impossible. This entails that a telescope has to contain at least two lenses.

3 Electromagnetic Elements with Straight Axis

In the previous sections, we have concentrated on glass optical systems where light rays are affected by refractive surfaces. Our main goal are particle optical systems, where the trajectories of charged particles are affected by suitable electromagnetic fields. As a first step in this direction, let us first review the equations of motion in electromagnetic fields.

3.1 The Equations of Motion in Electromagnetic Fields

For arbitrary electromagnetic fields, the equations of motion of a charged particle are

$$\dot{\vec{r}} = \vec{v} \tag{21}$$

$$\dot{\vec{p}} = e\left(\vec{E} + \vec{v} \times \vec{B}\right) \tag{22}$$

Since these equations contain both the momentum and the velocity, they are not very convenient for relativistic particles. In this case we have

$$\vec{p} = m\vec{v} = \frac{m_0}{\sqrt{1 - \frac{v^2}{c^2}}}\vec{v} \tag{23}$$

$$\Rightarrow \vec{v} = \frac{\vec{p}}{m_0}\left(1 + \frac{\vec{p}^2}{m_0^2 c^2}\right)^{-\frac{1}{2}} \tag{24}$$

So we can eliminate the velocity from the equations:

$$\dot{\vec{r}} = \frac{\vec{p}}{m_0}\left(1 + \frac{\vec{p}^2}{m_0^2 c^2}\right)^{-\frac{1}{2}} \tag{25}$$

$$\dot{\vec{p}} = e\left(\vec{E} + \frac{\vec{p}}{m_0}\left(1+\frac{\vec{p}^2}{m_0^2}\right)^{-\frac{1}{2}} \times \vec{B}\right) \tag{26}$$

This is of crucial importance for the motion in accelerating cavities. However, for focusing elements the effect is very minor; in magnetic systems, there are no changes in velocities, and since electric fields mainly act perpendicular to the particle trajectory, the effect is not very important there either. So we make the assumption that the mass changes adiabaticlly if at all, which leads to

$$\ddot{\vec{r}} = \frac{e}{m}\left(\vec{E} + \dot{\vec{r}} \times \vec{B}\right) \tag{27}$$

In the following, we are interested in what happens to a beam that moves in z direction. This means that the main velocity components of the beam is in z-direction, and because of $v_z = \sqrt{v^2 - v_x^2 - v_y^2}$, to first order we have $v_z = v = const$ if there are no accelerating fields. It turns out that this approximation is enough to obtain the first order part of the transfer map. In this case the equations can be transformed to z as the independent variable, and dots (derivatives with respect to t) can be replaced by primes (derivatives with respect to z). Using

$$\vec{r}'' = \frac{d^2\vec{r}}{dz^2} = \frac{\ddot{\vec{r}}}{v_z^2} \tag{28}$$

$$\vec{v}' = \left(\frac{v_x}{v_z}, \frac{v_y}{v_z}, 1\right), \tag{29}$$

we obtain the equations of the particle trajectory

$$x'' = \frac{e}{mv_z^2}E_x + \frac{e}{mv_z}\left(y'B_z - B_y\right) \tag{30}$$

$$y'' = \frac{e}{mv_z^2}E_y + \frac{e}{mv_z}\left(B_x - x'B_z\right) \tag{31}$$

$$\tag{32}$$

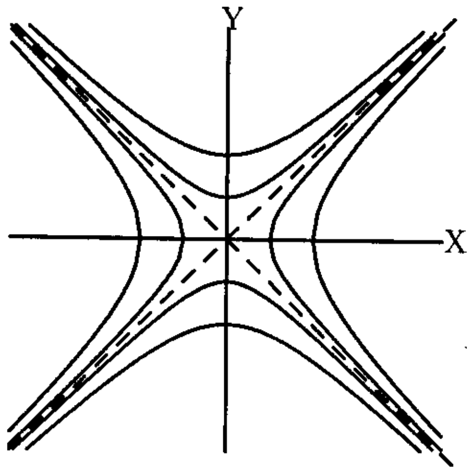


Figure 10: Equipotential lines of electric quadrupole

In the following sections, we will study the various types of motion for several important cases.

3.2 Electric Case, $E_z = 0$

We begin with elements in which there are no fields in z direction. Since we are interested in cases in which the motion is linear, we have to have linear fields, and we assume $E_x = -Px$. Since $\nabla \cdot \vec{E} = 0$ (there are no charges inside the field generating element), we can infer $E_y = Py$. The potential to this field is given by $V = \frac{1}{2}P(x^2 - y^2)$.

To produce this field, we can carve metal electrode along one set of equipotential lines, producing a set of boundary conditions in which to solve the Laplace equation. Since to a given boundary condition, there is a unique solution, we automatically obtain the right one. Because of the four fold symmetry, the resulting device is called the "Electrostatic Quadrupole".

Plugging in the fields into the equations of motion, we obtain linear equations which have the following solutions:

$$
\begin{aligned}
x &= x_0 \cos KL + \frac{x_0'}{K} \sin KL && (33)\\
x' &= -K x_0 \sin KL + x_0' \cos KL && (34)\\
y &= y_0 \cosh KL + \frac{y_0'}{K} \sinh KL && (35)\\
y' &= -K y_0 \sinh KL + y_0' \cosh KL && (36)\\
& && (37)
\end{aligned}
$$

where we have used the quantities

$$
\begin{aligned}
K^2 &= (e/mv^2)P\\
\chi_e &= mv^2/e
\end{aligned}
$$

where χ_e is called the "electric rigidity".

$$
\begin{pmatrix} x_2 \\ x_2' \\ y_2 \\ y_2' \end{pmatrix} = \begin{pmatrix} \cos KL & \frac{1}{K}\sin KL & 0 & 0 \\ -K\sin KL & \cos KL & 0 & 0 \\ 0 & 0 & \cosh KL & \frac{1}{K}\sinh KL \\ 0 & 0 & -K\sinh KL & \cosh KL \end{pmatrix} \begin{pmatrix} x_1 \\ x_1' \\ y_1 \\ y_1' \end{pmatrix} \quad (38)
$$

So differing from the simple glass lens, here the x and y directions are treated quite differently; Focusing in one direction comes at the expense of defocussing in the other. Note however that by choosing P of the other sign, K becomes imaginary and the trig functions and hyperbolic functions exchange their roles. This is important because it allows to combine quadrupoles with different P to make a simultaneous image in x and y directions.

3.3 Electric Case, $E_z \neq 0$

It turns out that by giving up the condition $E_z = 0$, we can obtain a field that treats the x and y directions equal. Let us assume that we want $E_x = -Px$,

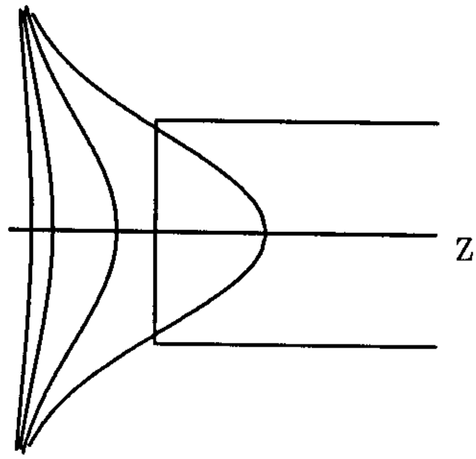


Figure 11: Electric round lens

$E_y = -Py$, then from $\nabla\vec{E} = \vec{0}$ we infer $\frac{d}{dz}E_z = 2P$. So there must be a field in z direction, which is indeed larger as the fields in x and y directions. Since only the latter ones are suitable for focusing, the bulk of the field is wasted. Furthermore, the element cannot be made very long because of the need for an ever increasing z field, and so their use is limited to low energy beams. Since rotational symmetry can be maintained, these elements are called round lenses.

3.4 Magnetic Case , $B_z = 0$

To obtain linear motion in the magnetic cases, from the equations of motion one can infer $B_y = Px$. Since $\nabla \times \vec{B} = 0$, we get $B_x = Py$, and there is a scalar potential which has the form $V = -Pxy$.

The procedure to produce this field is similar as in the last section: carve iron following equipotential lines. An important difference is that while in the electric case, the electrode surface was always a perfect equipotential surface, in the magnetic case this is no longer the case because of the not infinite permeability of the iron.

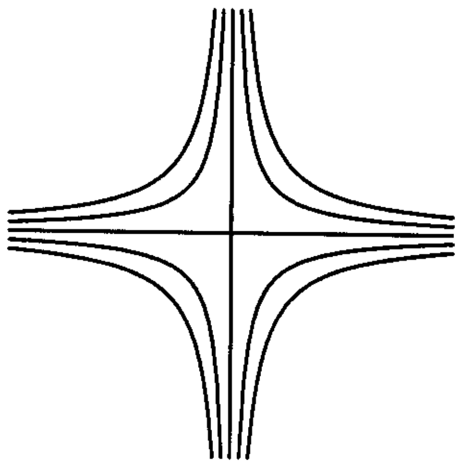


Figure 12: Magnetic quadrupole

Altogether, the solutions of the resulting linear equations of motion can be written in the following way:

$$\begin{pmatrix} x_2 \\ x_2' \\ Y_2 \\ y_2' \end{pmatrix} = \begin{pmatrix} \cos KL & \frac{1}{K}\sin KL & 0 & 0 \\ -K\sin KL & \cos KL & 0 & 0 \\ 0 & 0 & \cosh KL & \frac{1}{K}\sinh KL \\ 0 & 0 & -K\sinh KL & \cosh KL \end{pmatrix} \begin{pmatrix} x_1 \\ x_1' \\ y_1 \\ y_1' \end{pmatrix} \tag{39}$$

where $K^2 = Pe/mv$, $mv/e = \chi_m$, and χ_m is called magnetic rigidity.

Note again that by choosing P of a different sign, the trig and hyperbolic solutions switch.

3.5 Magnetic Case , $B_z \neq 0$

If we demand a rotationally symmetrical arrangement of the fields, we have to satisfy $B_x = Px$, $B_x = Py$, which can be achieved in a similar way as in the electric case by setting $\frac{d}{dz}B_z = -2P$. Note that per se, the rotationally

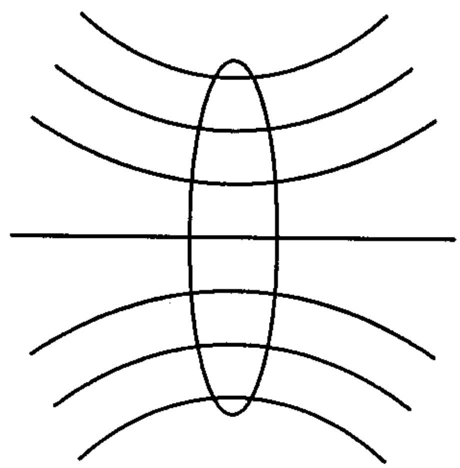


Figure 13: Magnetic round lens

symmetric field does not provide any focusing, but rather a coupling of the x and y motions of the equations of motion. The bulk result is a rotation in the x-y plane; as a side effect, the azimuthal velocities, combined with the radial field, generate radial focusing.

So similar to the electric case, the bulk of the field is wasted, and the resulting focusing is rather weak. Since again the elements can only be rather short, they are only suitable for low energy particles. They are used in electron microscopes and TVs. Similar to the case of the electrostatic round lenses, in reality the fields are quite nonlinear, leading to substantial nonlinearities in the true transfer map.

4 Bending Elements

4.1 The Magnetic Dipole

The most important bending element is the dipole which in the inside has just a constant magnetic field. So the particles move in circles:

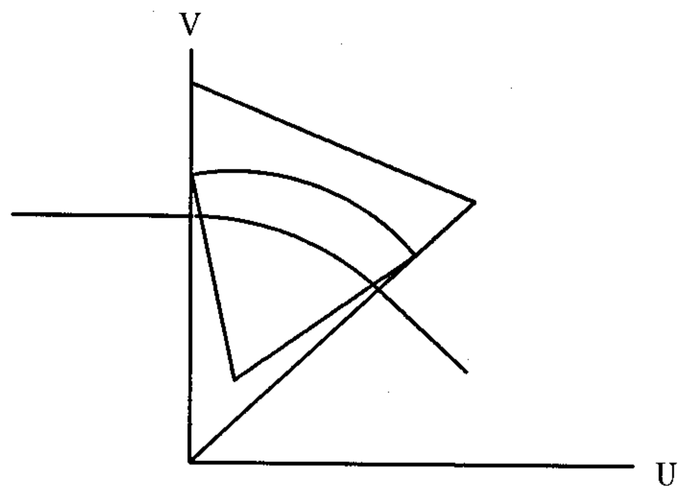


Figure 14: Sketch of a magnetic dipole

$$\frac{mv^2}{R} = evB \tag{40}$$

From which we have

$$RB = \frac{mv}{e} = \chi_m \tag{41}$$

where $[\chi_m] = 1Tm$.

Because there are just circles, the transfer map of the dipole can be computed analytically with just a little geometry. Consider particle with x_i,x'_i,y_i,y'_i, $\delta_p = (\chi_m - \chi_{m0})/\chi_{m0}$. Compared with the reference particle, $r = R(1+\delta_p)$, and the center of the circle lies at $\vec{P} = (r \sin\alpha_i, x_i + R - r\cos\alpha_i)$. Where $\vec{P}$. On the exit line, we have

$$u_e = (R + x_f)\sin\phi \tag{42}$$
$$v_e = (R + x_f)\cos\phi \tag{43}$$

So we have to solve

$$[(R+x_f)\sin\phi - r\sin\alpha_i]^2 + [(R+x_f)\cos\phi - (x_i + R - r\cos\alpha))]^2 = r^2 \quad (44)$$

and keep only terms linear in x_i, $\alpha_i = x'_i$, δ_p. After some arithmetic, this leads to

$$x_f = x_i \cos\phi + x'_i \cdot R\sin\phi + \delta_0 R(1-\cos\phi) \quad (45)$$

$$x'_f = \frac{1}{R}\frac{d}{d\phi}x_f(\phi) = -\frac{1}{R}x_i \sin\phi + x'_i \cos\phi + \delta_0 \sin\phi \quad (46)$$

Again, the simple linear relation represents quite an approximation. Firstly, there are obviously higher order terms which appear with an exact solution of the geometry. Furthermore, the entrance and exit regions of the dipole do have nonuniform fields, which lead to non-circular motion.

Noting that the motion in y direction is just a drift, we obtain for the total transfer matrix

$$\begin{pmatrix} x_f \\ x'_f \\ y_f \\ y'_f \\ \delta_f \end{pmatrix} = \begin{pmatrix} \cos\phi & R\sin\phi & 0 & 0 & R(1-\cos\phi) \\ -\frac{1}{R}\sin\phi & \cos\phi & 0 & 0 & \sin\phi \\ 0 & 0 & 1 & R\phi & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_i \\ x'_i \\ y_i \\ y'_i \\ \delta_i \end{pmatrix} \quad (47)$$

4.2 Edge Focusing

Frequently the entrance into a dipole magnet is not perpendicular, but deliberately at a certain angle. For a positive angle, the net effect in x direction is focusing; particles above the reference particle enter the field earlier and are thus bent down, and the opposite happens to particles below the reference particle. This can be approximated by a thin lens action. Interestingly, there is also an effect in the y direction: Since the fringe field is now

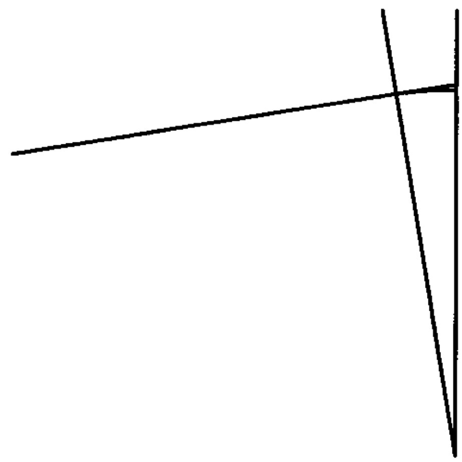


Figure 15: The Effect of Edge Focusing

entered at an angle, there is a net defocussing. Altogether, the effects are approximately:

$$x_2 = x_1 \tag{48}$$
$$x_2' = x_1' - x_1 \cdot \frac{\tan\varepsilon}{r} \tag{49}$$
$$y_2 = y_1 \tag{50}$$
$$y_2' = y_1' + y_1 \cdot \frac{\tan\varepsilon}{r} \tag{51}$$
$$\tag{52}$$

So the transfer matrix is

$$M = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ -\frac{\tan\varepsilon}{r} & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & \frac{\tan\varepsilon}{r} & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \tag{53}$$

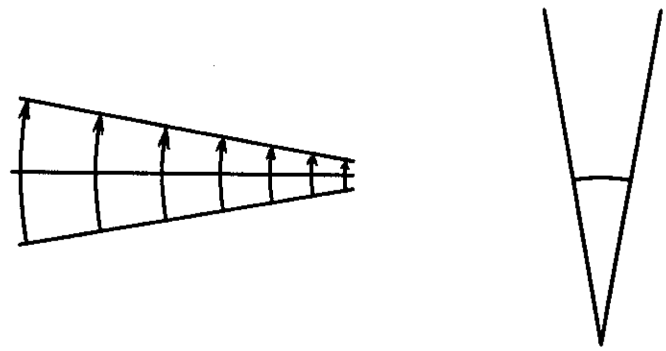


Figure 16: Magnetic dipole with inhomogeneous field

4.3 The Inhomogeneous Magnet

In many cases, it is advantageous to deviate from the concept of a linear homogeneous magnet. Often there is a deliberate nonlinearity of the form

$$B = B_0(1 - n_1 \frac{x}{R}) \tag{54}$$

We have to postpone a quantitative discussion of this effect; qualitatively, trajectories with positive x go "more straight" if $n_1 > 0$, and trajectories with negative x are bent more if $n_1 > 0$. Trajectories are focused in y-direction. The transfer matrix is

$$\left(\begin{array}{cc} \cos\sqrt{1-n_1}\phi & (R/\sqrt{1-n_1})\sin\sqrt{1-n_1}\phi \\ -(\sqrt{1-n_1}/R)\sin\sqrt{1-n_1}\phi & \cos\sqrt{1-n_1}\phi \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{array} \right.$$

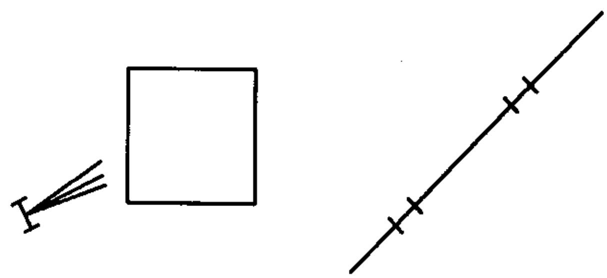


Figure 17: Sketch of a momentum spectrograph

$$
\left.\begin{array}{ccc}
0 & 0 & R(1-\cos\sqrt{1-n_1}\phi)/\sqrt{1-n_1} \\
0 & 0 & \sin\sqrt{1-n_1}\phi \\
\cos\sqrt{n_1}\phi & (R/\sqrt{n_1})\sin\sqrt{n_1}\phi & 0 \\
-(\sqrt{n_1}/R)\sin\sqrt{n_1}\phi & \cos\sqrt{n_1}\phi & 0 \\
0 & 0 & 1
\end{array}\right)
$$

An interesting consequence is that an inhomogeneous magnet treats x and y identically for $n_1 = \frac{1}{2}$.

4.4 The Momentum Spectrograph

Using a bending magnets, it is possible to build a devices that sorts momenta by combining fields suitably to let the final position only depend on the momentum, and not the initial velocity. These are called momentum spectrographs. We thus require

$$(x|x') = 0, \quad (x|\delta_p) \text{ large.} \tag{55}$$

Determinant unity implies that $(x|x) \neq 0$,but it should not be too large. The initial width D_i has to be made as small as possible. The final width is

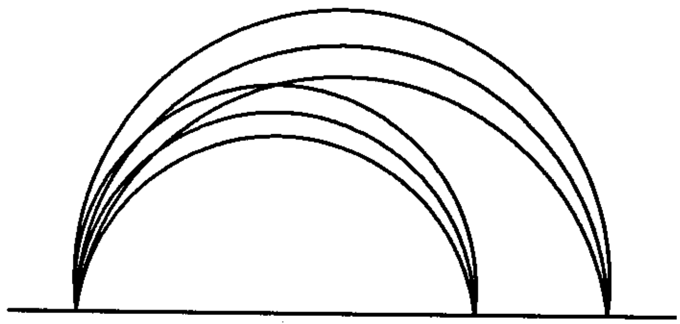


Figure 18: Isotope separator

$|(x|x)D_i|$, and thus to separate δ_{p1} and δ_{p2}, we must have $|(x|\delta)(\delta_{p1} - \delta_{p2})| > |(x|x)D_i|$.

We define the resolution of the device as $r = \chi_m/\Delta\chi_m = 1/\Delta\delta = |(x|\delta)/(x|x)D_i|$

A momentum spectrograph can also be used as a mass spectrograph if the incoming energy is constant. An example for a simple isotope separator is a 180 degree bending magnet (see picture).

4.5 Electrostatic Deflector

Besides magnetic deflectors, sometimes also some electrostatic equivalents are used.

To force a particle on a circle, we have $mv^2/R = eE \Rightarrow E = \chi_e/R$, $\chi_e = mv^2/e$ "electric rigidity". The unit is $[\chi_e] = 1V$. The quantitative derivation of the transfer matrix has to be postponed until later. The result is

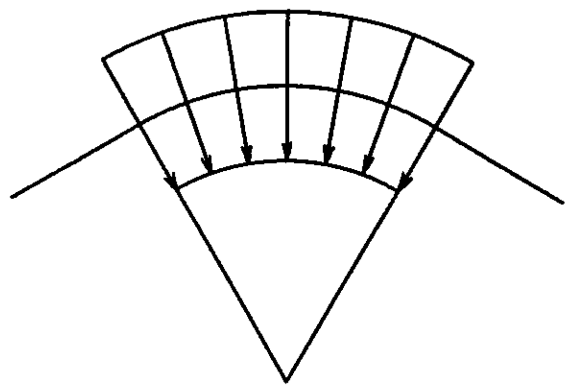


Figure 19: Sketch of an Electrostatic Deflector

$$\begin{pmatrix} \cos\sqrt{1-n_1}\phi & (R/\sqrt{1-n_1})\sin\sqrt{1-n_1}\phi & 0 & 0 & R(1-\cos\sqrt{1-n_1}\phi)/\sqrt{1-n_1} \\ -(\sqrt{1-n_1}/R)\sin\sqrt{1-n_1}\phi & \cos\sqrt{1-n_1}\phi & 0 & 0 & \sin\sqrt{1-n_1}\phi \\ 0 & 0 & \cos\sqrt{n_1}\phi & (R/\sqrt{n_1})\sin\sqrt{n_1}\phi & 0 \\ 0 & 0 & -(\sqrt{n_1}/R)\sin\sqrt{n_1}\phi & \cos\sqrt{n_1}\phi & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

where again n_1 describes the inhomogeneity of the electric field. We note that a Condenser with spherical sheets has potential $1/r$ and can be done analytically similar to the homogeneous bending magnet. (Kepler problem).

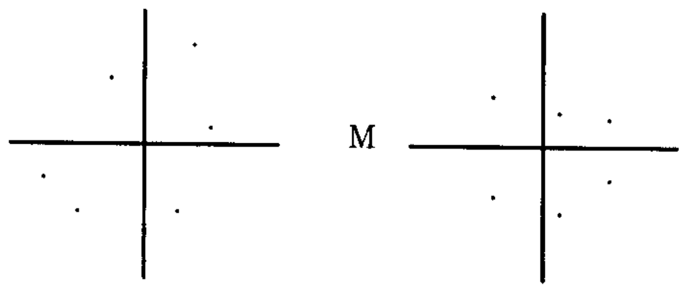


Figure 20: Mapping of Phase Space Points

5 Motion in Phase Space

After having discussed important particle optical elements and some devices, we consider the effect of these elements on x, x' space, the so-called phase space. We point out several important aspects of the motion.

- In our approximation, the action of elements is a linear matrix with unity determinant.
- Since the matrix is nonsingular, different initial points have different final points.
- Because of linearity, straight lines stay straight lines.
- Because of linearity, continuous curves stay continuous curves.
- Because of unity determinant, areas are preserved.

5.1 Emittance

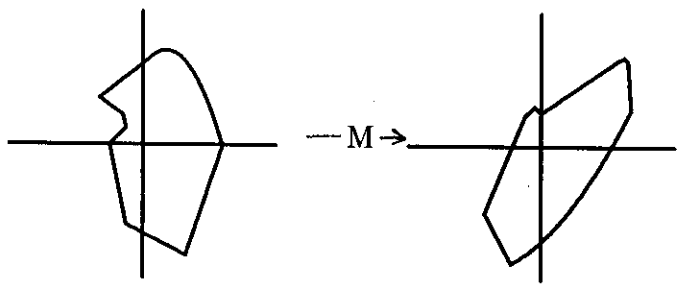


Figure 21: Mapping of the emittance

Since the areas of phase space are preserved under the action of an element, it makes sense to introduce a name for this area; we call it emittance. We note that the emittance is a property of the incoming beam, and is invariant as it goes through the system. Usually a system can only accept a certain maximum emittance without having particles hit certain boundaries, and this maximum emittance is called the acceptance.

We note the concept of the preservation of emittance does also hold in the nonlinear case; there it is connected to Liouville's theorem and the Hamiltonian type of the motion.

A nice consequence of the above said is that in order to study the motion, it suffices to study a closed curve containing all particles, because particles cannot penetrate this surface as they go through the system.

5.2 The Phase Space Action of Drifts and Lenses

From the transfer matrices, it becomes apparent that the effect of a drift is a shearing parallel to the x-axis, and the action of a lens is a shearing parallel to the x'-axis.

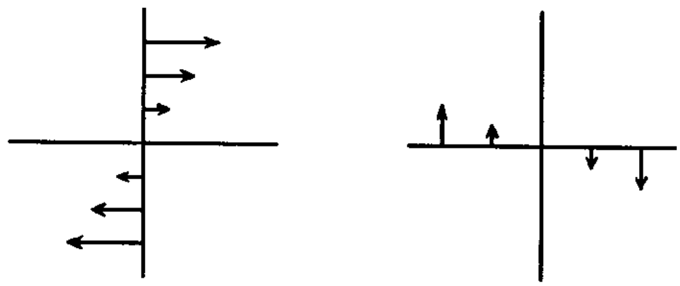


Figure 22: Motion of phase space points through drift(left) or lens(right)

5.3 Parallelogram-Like Phase Space

An important example for phase spaces is the parallelogram-shaped phase space. In this case, the occupied space is included in a rectangle. Since straight lines stay straight lines and particles do not penetrate, after the action of a transfer matrix is still contained in a quadrangle, and for it suffices to study what happens to the corners alone. In particular, the width of the total beam can be computed by just looking at the maxima of the positions of the four corner points.

5.4 Elliptic Phase Space

Another important case is the elliptic phase space. Let us assume that our phase space is circular at the beginning:

$$(x_1, x_1') \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_1' \end{pmatrix} = \varepsilon \tag{56}$$

After a transformation $(x_2, x_2') = M(x_1, x_1') \Rightarrow (x_1, x_1') = M^{-1}(x_2, x_2')$

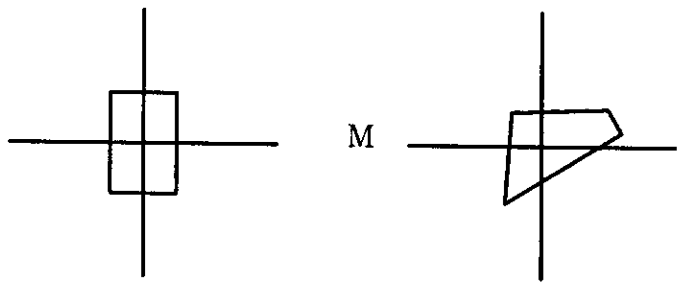


Figure 23: Change of a parallelogram-like phase area though mapping

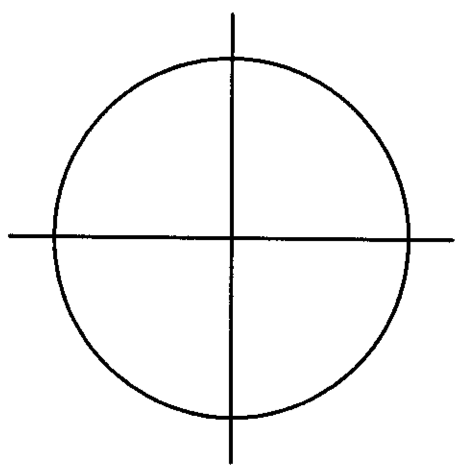


Figure 24: Circular phase area

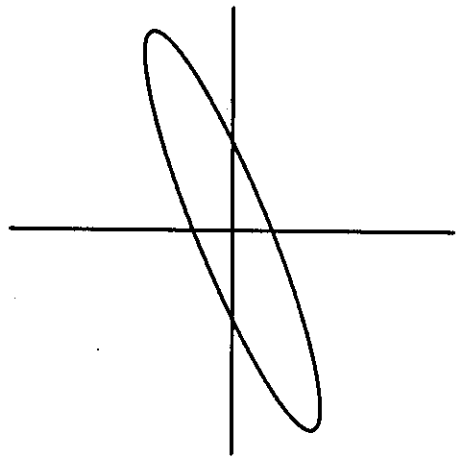


Figure 25: Elliptic phase area

Inserting this into the above equation, we obtain

$$\Rightarrow (x_2, x_2')(M^{-1t}M^{-1})(x_2, x_2') = \varepsilon \tag{57}$$

The matrix $(M^{-1t}M^{-1})$ is symmetric, so let $(M^{-1t}M^{-1}) = \begin{pmatrix} \gamma & \alpha \\ \alpha & \beta \end{pmatrix}$, and obtain

$$(x_2, x_2')\begin{pmatrix} \gamma & \alpha \\ \alpha & \beta \end{pmatrix}(x_2, x_2') = \varepsilon \tag{58}$$

This is an ellipse. Because $|M^{-1t}M^{-1}| = 1$, we infer $\beta\gamma - \alpha^2 = 1$, and we get

$$x^2\gamma + 2xx'\alpha + x'^2\beta = \varepsilon \tag{59}$$

from which we find out that the axis intersections are $x_0 = \sqrt{\frac{\varepsilon}{\gamma}}$, $a_0 = \sqrt{\frac{\varepsilon}{\beta}}$.

To determine the maxima in x and x' directions is a little more difficult. Let $f(x,x') = x^2\gamma + 2xx'\alpha + x'^2\beta$, then we have

$$\Rightarrow \vec{\nabla} f = (2x\gamma + 2x'\alpha, 2x\alpha + 2x'\beta) \tag{60}$$

Where we have maximum x, we have $(\vec{\nabla} f)_y = 0$. So $x\alpha = -x'\beta \Rightarrow x' = -\frac{\alpha}{\beta}x$ Plug x' back to the original equation

$$x^2\gamma + 2x\alpha(-\frac{x\alpha}{\beta}) + \beta(\frac{x\alpha}{\beta})^2 = \varepsilon \tag{61}$$

$$x^2\beta\gamma - x^2\alpha^2 = \varepsilon \tag{62}$$

$$x_m = \sqrt{\beta\varepsilon} \tag{63}$$

In a similar way, we obtain $a_m = \sqrt{\gamma\varepsilon}$

5.5 The Transformation of the Ellipse

If the initial emittance is an ellipse like below

$$(x_1, x_1') \begin{pmatrix} \gamma_1 & \alpha_1 \\ \alpha_1 & \beta_1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_1' \end{pmatrix} = \varepsilon \tag{64}$$

what happens after another transformation? We obtain

$$(x_2, x_2') \left[(M^{-1})^t \begin{pmatrix} \gamma_1 & \alpha_1 \\ \alpha_1 & \beta_1 \end{pmatrix} M^{-1} \right] \begin{pmatrix} x_2 \\ x_2' \end{pmatrix} = \varepsilon \tag{65}$$

And the relations between the initial and final parameters are

$$\begin{pmatrix} \beta_2 \\ \alpha_2 \\ \gamma_2 \end{pmatrix} = \begin{pmatrix} (x,x)^2 & -2(x,x)(x,x') & (x,x')^2 \\ -(x,x)(x',x) & (x,x)(x',x') + (x,x')(x',x) & -(x,x')(x',x') \\ (x',x) & -2(x',x)(x',x') & (x',x')^2 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \alpha_1 \\ \gamma_1 \end{pmatrix} \tag{66}$$

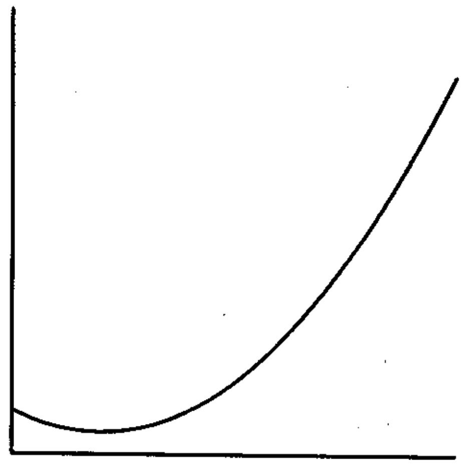


Figure 26: Behavior of β through drift L

As an example, let us consider the transform through a drift $\begin{pmatrix} 1 & L \\ 0 & 1 \end{pmatrix}$. We obtain

$$\beta = \beta_0 - 2L\alpha_0 + L^2\gamma_0 \tag{67}$$

The minimum of β occurs at $L = \frac{\alpha}{\gamma}$, and this point is called waist. It is worth pointing out that the location of the waist is does not have to coincide with the location of an image.

6 Repetitive Systems

In many cases it is very important to see what happens when systems are iterated. This case happens in storage rings and circular accelerators. Suppose the transfer matrix for one turn is

$$M = \begin{pmatrix} (x,x) & (x,x') \\ (x',x) & (x',x') \end{pmatrix}$$

It turns out that the repetitive behaviour of the system is easiest studied by switching to new variables, the so-called normal form variables. These variables are connected to the eigenvectors of the matrix. We begin by finding the eigenvalues of M through the equation: $|M-\lambda E| = 0$, i.e. $(x,x)(x',x') - \lambda[(x,x)+(x',x')] + \lambda^2 - (x,x')(x',x) = 0$. So the eigenvalues are

$$\lambda_{1,2} = \frac{-[(x,x)+(x',x')] \pm \sqrt{[(x,x)+(x',x')]^2 - 4}}{2} \tag{68}$$

$$= -\frac{trM}{2} \pm \sqrt{\frac{trM}{2}^2 - 1} \tag{69}$$

It is important that the eigenvalues are either real or form a conjugate pair. Furthermore, if $trM \neq 2$, the two eigenvalues form a reciprocal pair, i.e. $\lambda_1 \cdot \lambda_2 = 1$.

6.1 The Unstable Case $|trM| > 2$

Here λ_1,λ_2 are real and distinct. Since they form a reciprocal pair, one of the λ is greater than unity. Considering motion in eigenvector coordinates $\vec{V_+}$, $\vec{V_-}$ (suppose $|\lambda_+| > 1$,$|\lambda_-| < 1$),we have

$$\vec{x} = a\vec{V_+} + b\vec{V_-} \tag{70}$$

$$M\vec{x} = \lambda_1 a\vec{V_+} + \lambda_2 b\vec{V_-} \tag{71}$$

Define $\bar{a} = \lambda_1 a$, $\bar{b} = \lambda_2 b$.

$$\bar{a}\bar{b} = ab \tag{72}$$

This mean that the phase space point moves along hyperbola. Thus the particles moves further and further out, and the same happens in the original coordinates. So the motion unstable and hence of little interest to us.

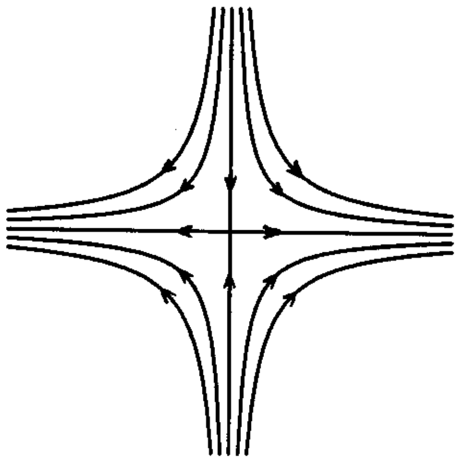


Figure 27: Trajectories of a phase space point in the eigenvector coordinates in case of $|trM| > 2$

6.2 The Stable Case $|trM| < 2$

In this case λ_1 and λ_2 are complex conjugates, $\lambda_2 = \bar{\lambda_1}$. And because $\lambda_2 = 1/\lambda_1$,we get $\lambda_{1,2} = e^{\pm i\mu}$.

The quantity μ is a very important characteristic of the accelerator, it is called the tune. The eigenvectors satisfy the same relation as the eigenvalues, $\vec{V_1} = \bar{\vec{V_2}}$. In the coordinates with unit vectors $\vec{V_+} = Re(\vec{V_1})$, $\vec{V_-} = Im(\vec{V_1})$, we have

$$\vec{x} = a\vec{V_+} + b\vec{V_-} \tag{73}$$

$$M\vec{x} = \begin{pmatrix} \cos\mu & \sin\mu \\ -\sin\mu & \cos\mu \end{pmatrix} \vec{x} \tag{74}$$

which means $\vec{x}$ moves in circles, which means the motion is stable under iteration. The angular advance in the normal coordinates is constant for each iterations and equals the angle μ, i.e. the tune.

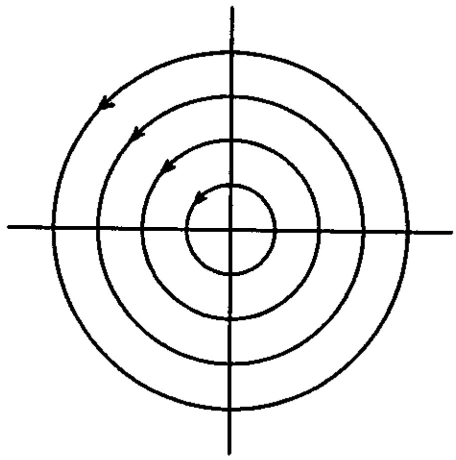


Figure 28: Trajectories in case of $|trM| < 2$

Note that in the original coordinates, the angular advance does not have to be constant from turn to turn; but one quickly sees that the average angular advance over many turns equals μ. For this purpose we observe that the angle advance in the original coordinates can be described by the angle advance due to the transformation to normal form coordinates plus the angle advance in normal form coordinates plus the angle advance of the reverse transformation. Since for many iterations, the transformation advances contribute only an insignificant amount, the average approaches the angle in the normal form coordinates.

It is very important in practice that even if the accelerator is perturbed a bit, i.e. the matrix elements change a little, the trace will still stay below two and so the qualitative behaviour does not change. In original coordinates the motion lies on an ellipse.

6.3 The Degenerate Case $|trM| = 2$

This case is degenerate and should be avoided. Since a small perturbation can change the trace to larger than 2, this case is potentially unstable under perturbation.

6.4 The Invariant Ellipse in the Stable Case

From above, we know $\lambda_{1,2} = e^{\pm i\mu}$; Here we choose sign($\mu$) =sign($(x,x')$).

Define

$$\alpha_i = \frac{(x,x)-(x',x')}{2\sin\mu} \tag{75}$$

$$\beta_i = \frac{(x,x')}{\sin\mu} \tag{76}$$

$$\gamma_i = -\frac{(x',x)}{\sin\mu}, \tag{77}$$

we note $\beta_i, \gamma_i > 0$; furthermore From the definitions, we rewrite the transfer matrix in terms of α_i,β_i,γ_i and μ:

$$M = \begin{pmatrix} \cos\mu + \alpha_i \sin\mu & \beta_i \sin\mu \\ -\gamma_i \sin\mu & \cos\mu - \alpha_i \sin\mu \end{pmatrix} \tag{78}$$

$$= I \cdot \cos\mu + K \sin\mu \tag{79}$$

$$\Rightarrow M^{-1} = I \cdot \cos\mu - K \sin\mu \tag{80}$$

Where I is the unit matrix, $K = \begin{pmatrix} \alpha_i & \beta_i \\ -\gamma_i & -\alpha_i \end{pmatrix}$.

It now turns out that the coefficients α_i, β_i and γ_i introduced above and determined by the map are the coefficients of an ellipse that stays invariant under the motion, which according to above means

$$M^{-1t} \begin{pmatrix} \gamma_i & \alpha_i \\ \alpha_i & \beta_i \end{pmatrix} M^{-1} = \begin{pmatrix} \gamma_i & \alpha_i \\ \alpha_i & \beta_i \end{pmatrix} \tag{81}$$

To prove this important relationship, we set $J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, $T = \begin{pmatrix} \gamma_i & \alpha_i \\ \alpha_i & \beta_i \end{pmatrix}$. Then

$$
\begin{aligned}
TK &= \begin{pmatrix} \gamma_i & \alpha_i \\ \alpha_i & \beta_i \end{pmatrix} \begin{pmatrix} \alpha_i & \beta_i \\ -\gamma_i & -\alpha_i \end{pmatrix} = J && (82) \\
K^t T &= K^t T^t = J^t = -J && (83) \\
\Rightarrow M^{-1t} T M^{-1} &= (J \cos\mu - K^t \sin\mu) T (J \cos\mu - K \sin\mu) && (84) \\
&= T && (85)
\end{aligned}
$$

This relation means that if the emittance at the beginning is

$$\vec{x} T \vec{x} = \varepsilon \tag{86}$$

The invariant ellipse of a stable repetitive system is an important characteristic of the system with an importance similar to the tune.

6.5 Matching and Beating

In order to utilize the acceptance of an accelerator most efficiently, it is important to match the ellipse delivered to the accelerator with the invariant ellipse of the accelerator. In this case, under each iteration, the occupied area in phase space falls together. If this is not the case, each time around another area will be occupied, and a much a larger acceptance is required. This effect is called beating.

6.6 A Glimpse at Nonlinear Effects

The previous sections fully solved the question of stability in an accelerator that is described by a linear transfer map. Unfortunately, all transfer maps are nonlinear, and nonlinear effects can make the situation much more complicated and often even impossible to fully solve. We have to refer to a future course to a more thorough discussion of these effects; here we want to limit ourselves to one observation.

If we go around a ring repeatedly, it is not desirable to have the particles come to the same locations in phase space after a small number of turns, because then they are particularly sensitive to the field nonlinearities at these particular points. So this requires to satisfy

$$n \cdot \mu_x + m\mu_y = 2\pi l \tag{87}$$

which is the famous tune resonance condition. In practice, the higher the n and m are allowed to become, the harder it will be to stay away from the corresponding resonances since the resonance lines lie dense in tune space.