PARTICLE DEFLECTIONS BY MAGNETS

Richard Talman

Laboratory of Nuclear Studies
Cornell University
Ithaca, NY 14853

ABSTRACT

This is one packet of notes accompanying a course *Mechanics and Electromagnetism in Accelerators,* offered as part of the U.S. Particle Accelerator School, Yale University, summer, 2002. This packet discusses the deflections caused by magnets and the use of Fourier analysis to describe the resultant motion.

Chapter 1.
Deflections Caused By Magnetic Fields

Magnetic Multipole Expansions

Particle kicks are administered by magnets. Here we introduce some typical ideal magnet types, describe their field dependencies and introduce Taylor series expansions, to be called multipole exansions, that describe both their ideal fields and nonideal deviations from them.

Though this section is not long, it purports to contain all that is required to represent all practical accelerator magnets except solenoids. *Ideal* accelerator magnetic fields are produced by currents parallel to the design orbit. These currents cause no longitudinal fields and, in complete generality, their magnetic fields can be expanded in a multipole series, to be exhibited shortly. It describes the two transverse magnetic field components in a plane perpendicular to the current elements, (and hence also perpendicular to the design orbit.) For definiteness, we will take this plane as passing through the longitudinal center of the magnet. This multipole plane will be so prominent in the analysis of accelerator orbits that we will use the terminology "multipole planes bend particles" rather than "magnets bend particles." Though this grates on the ear, it will do so less than the endless repetition of an expression such as "the bend at the magnet is calculated from its multipole series".

Later the use of particular pure magnet types such as quadrupoles or sextupoles will be described, but formally, at least for now, there is only one generic magnet, a superposition of all elemental types. Practical details about the construction and special use of pure magnet types are available from any number of sources.

Of course, some transverse currents at the magnet ends are unavoidable, and they will cause small longitudinal (also known as solenoidal) fields. We will usually ignore these small solenoidal fields, or compensate for them only crudely, with the disclaimer that this is almost always satisfactory. Special (almost certainly numerical) methods are called for when it is not.

This section, and the two that follow it, concentrate on a magnet present at a particular lattice location, say the one labeled (i). In order to make the equations appear

less formidable, we suppress the index (i), keeping only the $\pm$ label to distinguish output and input variables. The effect of a thin element is to make any typical output quantity $f^{(i+)}$ differ from the corresponding input quantity $f^{(i-)}$ by a calculated quantity $\Delta f = f^{(i+)} - f^{(i-)}$.

A prototypical magnet is the quadrupole.† With the quadrupole centered on the design orbit, the "true" (indicated by superscript T) magnetic field $\mathbf{B}^T(x,y,z)$, has no z component (at least on axis), no transverse field on axis, and purely linear dependence on x and y; it is given by

$$\begin{aligned} B_x^T(y,z) =& \frac{\partial B_x^T}{\partial y}(z)\ y = \frac{\partial B_y^T}{\partial x}(z)\ y, \\ B_y^T(x,z) =& \frac{\partial B_y^T}{\partial x}(z)\ x, \\ B_z^T =& 0, \end{aligned} \tag{1.1}$$

where the equality of partial derivatives follows from Ampere's law. The field components are more or less constant over the length L of the magnet, and presumeably fall to zero quickly, but smoothly, outside the magnet ends. We now idealize this field, replacing it by a field $\mathbf{B}$ for which the gradient $\partial B_y/\partial x$ is independent of z within the length L, but vanishes outside that. Of course, the field discontinuities at the ends violate Maxwell's equations.‡ Nevertheless this is often an excellent approximation. We will return later to a method for improving it. The field $\mathbf{B}$ is selected to give the same "field integral" as does $\mathbf{B}^T(z)$;

$$\frac{\partial B_y}{\partial x} = \frac{1}{L}\int_{-\infty}^{\infty} \frac{\partial B_y^T}{\partial x}(z)\, dz. \tag{1.2}$$

The product $L\,\partial B_y/\partial x$ is called the "quadrupole field integral". Finally we obtain the thin element idealization by reducing L to zero while holding the field integral constant. Rather than introducing a new symbol for the field integral we will leave the factor L explicit. This will leave open for the time being whether the magnet is "thick", with exactly constant derivatives over its finite length L, or "thin" in which case the $L = 0$ limit has been taken. Historically in accelerator theory the thick approach has usually been taken, but we will

† The dipole is even more essential, but we pass over that temporarily, since we prefer to defer facing the complication caused by curvature of the design orbit.

‡ Actually, such a magnetic field could be produced by placing current sheets at the ends, but that is never done in practice.

almost always take the thin approach. (As mentioned before, by subdividing elements where that is required for improved precision, we will be able to obtain results as accurate as can be obtained by the thick approach, without compromising symplecticity.)

The analysis just completed for a quadrupole can be generalized to any magnetic field caused by currents parallel to the z-axis. The steps in a general argument will be reduced to the following series of problems.

Problem 1..1. Steady current I flows along the infinite line $x = y = 0$ from $z = -\infty$ to $z = \infty$. Show (using Ampère's law) that the magnetic field is given by

$$B_x = -\frac{ky}{r^2}, \quad B_y = \frac{kx}{r^2}, \tag{1.3}$$

where $r^2 = x^2 + y^2$ and $k = \mu_0 I/(2\pi)$ (or $1/c$ in Gaussian units.)

Problem 1..2. Excluding the origin, show that the magnetic field of the previous problem can be derived from a magnetic scalar potential $\Phi_m = k \tan^{-1}(x/y)$ according to

$$\mathbf{B} = -\nabla \Phi_m. \tag{1.4}$$

Problem 1..3. Again excluding the origin, show that Φ_m satisfies the 2D Laplace equation,

$$\frac{\partial^2 \Phi_m}{\partial x^2} + \frac{\partial^2 \Phi_m}{\partial y^2} = 0. \tag{1.5}$$

Problem 1..4. Show that either the real or complex part of any analytic function $u(z) + iv(z)$ of the complex variable $z = x + iy$ satisfies the 2D Laplace equation. Hint: Use the Cauchy-Riemann conditions; $\partial u/\partial x = \partial v/\partial y$, $\partial u/\partial y = -\partial v/\partial x$.

Problem 1..5. Instead of treating a single line current at the origin, we now permit any number of line currents everywhere *except* in a tube centered on the z-axis (say in the beam tube of an accelerator.) There is still only current parallel to the z-axis however. The magnetic field will necessarily still be derivable from a magnetic scalar potential Φ_m satisfying Eq. (1.5). Why? Show that the real or imaginary parts of any non-negative power $(x+iy)^0, (x+iy)^1, (x+iy)^2, \ldots$ satisfies Eq. (1.5) in the current free region near the origin, and that the same can be said for any Taylor series in powers of x+iy. Such a Taylor expansion is called a "multipole series" because pure fields corresponding to different terms can be generated with magnets having multiple poles.

Problem 1..6. Now remove the restriction that the line currents be infinitely long (but they are still parallel to the z-axis); without loss of generality it is sufficient to consider current I flowing only in line element $\hat{z}dz$ at the origin. Along a line parallel to the z-axis, passing through the point $(x, y, 0)$, evaluate (using the Biot-Savart law) the magnetic field components $B_x(x, y, z)$ and $B_y(x, y, z)$. Integrating over z, show that "field integrals" satisfy

$$\int_{-\infty}^{\infty} B_x\,(x, y, z)\,dz = -\frac{k'y}{r^2}, \quad \int_{-\infty}^{\infty} B_y\,(x, y, z)\,dz = \frac{k'x}{r^2}. \tag{1.6}$$

where k' is a constant. This "closes the circle" of this series of problems in that this dependence on x and y is the same as was derived in Problem 1..1 and the following problems relied only on that dependence.

The result of Problem 1..6 is that the integrals over field components due to finite-length current elements have the same transverse spatial dependence as do the corresponding field components for infinite-line currents, calculated in Problem 1..1 It follows the field integrals can be expressed as Taylor series, as in Problem 1..5 These results will be used immediately, and then again later in Section 1.4 when analysing the focusing action caused by slanted pole faces.

The currents in all practical accelerator magnets (except solenoids) flow predominantly parallel to the design orbit. (The same statement even applies to the "bound currents" in ferromagnets.) Transverse currents necessary to convey the current to return legs are usually unimportant—the beam can be shielded from their fields which, as a result, tend to cause negligible deflection. Because the magnet lengths are finite, sometimes not much longer than the transverse magnet dimensions, it is not necessarily valid to treat the fields as two-dimensional. But the above series of problems showed that field integrals are subject to the same multipole analysis as applies to a purely two dimensional field, even though the fields themselves may have appreciable longitudinal components and complicated dependence on z. As we proceed, all magnetic deflections will be expressed in term of these magnetic field integrals. To the extent particle dynamics is expressible this way, detailed longitudinal field variation is unimportant. In particular the motion is insensitive to the details of the field shapes in the end regions where the fields fall to zero.

On the basis of this discussion one proceeds with confidence using a "standard" magnetic field expansion, valid over the multipole plane;

$$(LB_y) + i\,(LB_x) = (LB_0) \sum_{n=0}^{M} (b_n + ia_n) \left((x - \underline{\Delta x}) + i\left(y - \underline{\Delta y}\right) \right)^n . \tag{1.7}$$

The upper limit M of the multipole summation is chosen large enough to assure adequate accuracy over the "good field region" of the magnet. In our treatment, we will restrict M so that $M \leq N_T$ (the truncation order of the calculation).

Transverse element positioning misalignments ($\underline{\Delta x}$, $\underline{\Delta y}$) have been allowed for explicitly in Eq. (1.7). In this formalism "roll" misalignments, *i.e.* rotations around the design orbit can be represented cleanly by appropriate modification of the a_n and b_n. (See exercises. 1. Dipole roll in Eq. (1.7). 2. Effect of roll on all multipoles in Eq. (1.7).) Also it is easy to represent a misalignment that is a pure translation along the local longitudinal z-axis. Being a solid object, a magnet is also subject to two other orientation errors, which could be described as small rotations around the local x or y axes. These errors cannot be represented in the form of Eq. (1.7). Fortunately, with B_z vanishing except as an end effect, and with the direction of particle motion predominantly in the z-direction, the effects of such errors are extremely small. Should a situation be encountered with fields so strong, angles so great, *etc.*, that such forces were important, the present formalism would have to be refined. In that case, numerical solution of the differential equation of motion through that region would probably be the best way to proceed. As has been implied repeatedly, we seek an excellent idealization which contains the essential physics, and is subject to rigorous mathematical analysis, but we do not insist on perfection.

Dipole fields are described by the $n = 0$ terms. In fact, the notation was designed with dipoles in mind, which is why the factor B_0, the nominal dipole field, has been factored out. The field nonuniformity is quantified by the multipole coefficients. Ordinary steering dipoles have only a B_y component which means that

$$a_0 = 0, \quad b_0 = 1, \quad \text{for an erect dipole.} \tag{1.8}$$

With these definitions, the coefficients a_n and b_n for $n > 0$ give the multipole field errors as fractions of the nominal dipole field.

Since magnet types other than dipoles have vanishing dipole fields, it is necessary to replace B_0 if Eq. (1.7) is still to be used. Though the expansion (1.7) is reasonably standard for dipoles,† there appears to be no established standard for other elements. A natural choice for the erect quadrupole is to write

$$\left(LB_y^Q\right) + i\left(LB_x^Q\right) = \left(L\frac{\partial B_y^Q}{\partial x}\right)\left[x + iy + 10^{-4}\sum_{n=2}^{M}\left(b_n^Q + ia_n^Q\right)\frac{(x+iy)^n}{R_r^{n-1}}\right], \qquad (1.9)$$

where R_r is a reference radius, say 1 cm. The value of R_r and the factor 10^{-4} are normally chosen such that the numerical values of a_n^Q and b_n^Q are of order 1 for "bad", low order, multipoles and much less than 1 for high order multipoles.

Equations (1.7) and (1.9) can be reconciled with each other and with Eqs. (1.1) by making the identifications $B_0 = \partial B_y^Q/\partial x$ and $a_n(b_n) = 10^{-4}a_n^Q(b_n^Q)/R_r^{n-1}$. Often, even for dipoles, a factor R_r like this is introduced, with a value comparable to the bore radius of the magnet. For the Fermilab Tevatron a value $R_r = 2.54$ cm, better known as an inch, was selected. For the SSC, $R_r = 1$ cm. Leaving the factor R_r out completely, as in Eq. (1.7), is equivalent to setting it to 1 meter, since we use M.K.S. units. This being an unrealistically large amplitude, it also causes the coefficients a_n and b_n to become numerically very large for large values of n, even for magnets having excellent field uniformity. This is connected with the slightly inelegant fact that the coefficients a_n and b_n have dimensions depending on n. Nevertheless, for brevity, all our subsequent analysis will be based on Eq. (1.7). Expansions for all basic magnet types can be expressed in that form.

In order to coalesce common factors, scaled multipole coefficients are defined by

$$\tilde{a}_n = \frac{LB_0}{p_0/e}a_n, \qquad \tilde{b}_n = \frac{LB_0}{p_0/e}b_n. \qquad (1.10)$$

(A common mnemonic $p_0/e \rightarrow$ "$B\rho$" will be justified below.) In terms of the multipole coefficients, scaled magnetic field components are

$$\begin{aligned}\tilde{B}_y = \frac{B_y L}{p_0/e} &= \Re\sum_{n=0}^{M}\left(\tilde{b}_n + i\tilde{a}_n\right)\left((x-\underline{\Delta x}) + i\left(y-\underline{\Delta y}\right)\right)^n,\\ \tilde{B}_x = \frac{B_x L}{p_0/e} &= \Im\sum_{n=0}^{M}\left(\tilde{b}_n + i\tilde{a}_n\right)\left((x-\underline{\Delta x}) + i\left(y-\underline{\Delta y}\right)\right)^n.\end{aligned} \qquad (1.11)$$

† Actually expansion (1.7) is standard only in America. In Europe, coefficients are defined initially in an expansion of a vector potential which is then differentiated to obtain a series similar to (1.7). As a result the coefficients can be placed in one-to-one correspondance, but the indices are shifted by one, and the coefficients acquire extra factors of $n!$. Otherwise the expansions are equivalent.

In particular the factor $\frac{LB_0}{p_0/e}$ is the bend angle in radians of a particle of momentum p_0 passing through a dipole with field B_0 and arc length L. The factor p_0/e is often called "B-rho" and measured in Tesla-meters, because bend angle $= \frac{LB_0}{p_0/e} = \frac{LB_0}{B_0\rho} = \frac{L}{\rho}$, which is correct for motion in a circle of radius ρ. See Eq. (1.35) below.

To recapitulate, both for field components, B_x, B_y, and multipole coefficients b_n, a_n, the overhead tilde signifies conversion from the "thick element" continuum, absolute magnetic field representation of the element to its impulsive, "thin element", energy independent, representation. The factors have been chosen in a way that will simplify the formulas for "kinks", *i.e.* impulsive deflections, due to magnetic fields. (See Eqs. (1.32) below.)

The coefficients in multipole series Eq. (1.7) can be related to other conventional magnet strength parameters as shown in Table 1.1. Real and imaginary coefficients R_n and I_n are defined by

$$(x+iy)^n = R_n + iI_n, \tag{1.12}$$

where the misalignment errors ($\underline{\Delta x}$, $\underline{\Delta y}$) can be restored later. The factors 1!, 2!, 3! entering the definitions of quad strength q, sextupole strength S, octupole strength O, *etc.* are conventional. Formulas relating transverse momentum discontinuites to the magnetic field components ($\Delta x' \equiv \Delta f = -\tilde{B}_y$ and $\Delta y' \equiv \Delta g = \tilde{B}_x$) will be derived in Eqs. (1.32) of the next section. (Here the usual slope symbols x' and y' have been abbreviated by f and g in order to suppress the primes.) Anticipating those formulas we have for pure multipoles

$$\begin{aligned} \Delta f_n &= -\tilde{B}_y|_n = -\tilde{b}_n R_n + \tilde{a}_n I_n, \\ \Delta g_n &= \tilde{B}_x|_n = \tilde{b}_n I_n + \tilde{a}_n R_n. \end{aligned} \tag{1.13}$$

A typical horizontal steering magnet is illustrated in Fig. 1.1 which also shows the same magnet rotated to steer vertically. The measured, midplane, vertical magnetic field B_y is plotted in Fig. 1.2. It has the form

$$B_y(x, y=0) = B_{y0}\left(1 + b_2 x^2 + b_4 x^4\right), \tag{1.14}$$

where B_{y0} is the nominal value of B_y. For this magnet the dominant multipole imperfection coefficient happens to be decapole b_4. Referring to Table 1.1 it can be seen that this form can be continued off the median plane by

$$B_y(x, y) = B_{y0}\left(1 + b_2\left(x^2 - y^2\right) + b_4\left(x^4 - 6x^2y^2 + y^4\right)\right), \tag{1.15}$$

	n	R_n	I_n	$\tilde{b}_n$	$\tilde{a}_n$	$\Delta f = -\tilde{B}_y$	$\Delta g = \tilde{B}_x$
Horizontal bend	0	1	0	$\Delta\theta_x$	0	$-\Delta\theta_x$	0
Vertical bend				0	$\Delta\theta_y$	0	$\Delta\theta_y$
Erect quadrupole	1	x	y	$q = 1/f$	0	$-qx$	qy
Skew quadrupole				0	$q_s = 1/f_s$	$q_s y$	$q_s x$
Erect sextupole	2	$x^2 - y^2$	$2xy$	$S/2$	0	$-\frac{S}{2}(x^2 - y^2)$	$\frac{S}{2}2xy$
Skew sextupole				0	$S_s/2$	$\frac{S_s}{2}2xy$	$\frac{S_s}{2}(x^2 - y^2)$
Erect octupole	3	$x^3 - 3xy^2$	$3x^2y - y^3$	$O/6$	0	$-\frac{O}{6}(x^3 - 3xy^2)$	$\frac{O}{6}(3x^2y - y^3)$
Skew octupole				0	$O_s/6$	$\frac{O_s}{6}(3x^2y - y^3)$	$\frac{O_s}{6}(x^3 - 3xy^2)$
Erect decapole	4	$x^4 - 6x^2y^2$	$4xy(x^2 - y^2)$	$D/24$	0	$-\frac{D}{24}(x^4 - 6x^2y^2 + y^4)$	$\frac{D}{24}4xy(x^2 - y^2)$
Skew decapole		$+y^4$		0	$D_s/24$	$\frac{D_s}{24}4xy(x^2 - y^2)$	$\frac{D_s}{24}(x^4 - 6x^2y^2 + y^4)$

Table 1.1: Deflections caused by standard magnets and notation for their strengths

as well as yielding the other field component,

$$B_x(x,y) = B_{y0}\left(b_2 2xy + b_4 4xy\left(x^2 - y^2\right)\right). \tag{1.16}$$

Note that the field uniformity of the actual magnet is somewhat better at small amplitudes than either multipole would give by itself. This illustrates the possibility that truncation of the multipole series can, by defeating desirable cancellation, yield overly pessimistic field values.

Also plotted in Fig. 1.2 is the "wrong field component" $B_x(x, y = 10)$ plotted as a function of x, along a line displaced to positive $y = 10\,\mathrm{mm}$.

If the steering magnet is rotated by 90 degrees (anti-clockwise to an observer looking downstream so that positive horizontal deflection becomes positive vertical deflection) the new multipole expansion can be obtained from series (1.15) and (1.15) by transformations suggested by the labels on Fig. 1.1,

$$\begin{aligned} -B_x(x,y) &= (-B_{x0})\left(1 + b_2\left(y^2 - x^2\right) + b_4\left(y^4 - 6y^2x^2 + x^4\right)\right), \\ B_y(x,y) &= (-B_{x0})\left(b_2 2y(-x) + b_4 4y(-x)\left(y^2 - x^2\right)\right), \end{aligned} \tag{1.17}$$

where the nominal field is now B_{x0} with sign opposite to B_{y0}. Hence we have

$$\begin{aligned} B_y(x,y) &= B_{x0}\left(b_2 2xy - b_4 4xy\left(x^2 - y^2\right)\right), \\ B_x(x,y) &= B_{x0}\left(1 - b_2\left(x^2 - y^2\right) + b_4\left(x^4 - 6x^2y^2 + y^4\right)\right). \end{aligned} \tag{1.18}$$

To make this form match expansion Eq. (1.11) it is necessary to introduce skew coefficients a_2 and a_4 into Eq. (1.13)

$$\begin{aligned} B_y(x,y) &= B_{x0}\left(-a_2 2xy - a_4 4xy\left(x^2 - y^2\right)\right), \\ B_x(x,y) &= B_{x0}\left(1 + a_2\left(x^2 - y^2\right) + a_4\left(x^4 - 6x^2y^2 + y^4\right)\right). \end{aligned} \tag{1.19}$$

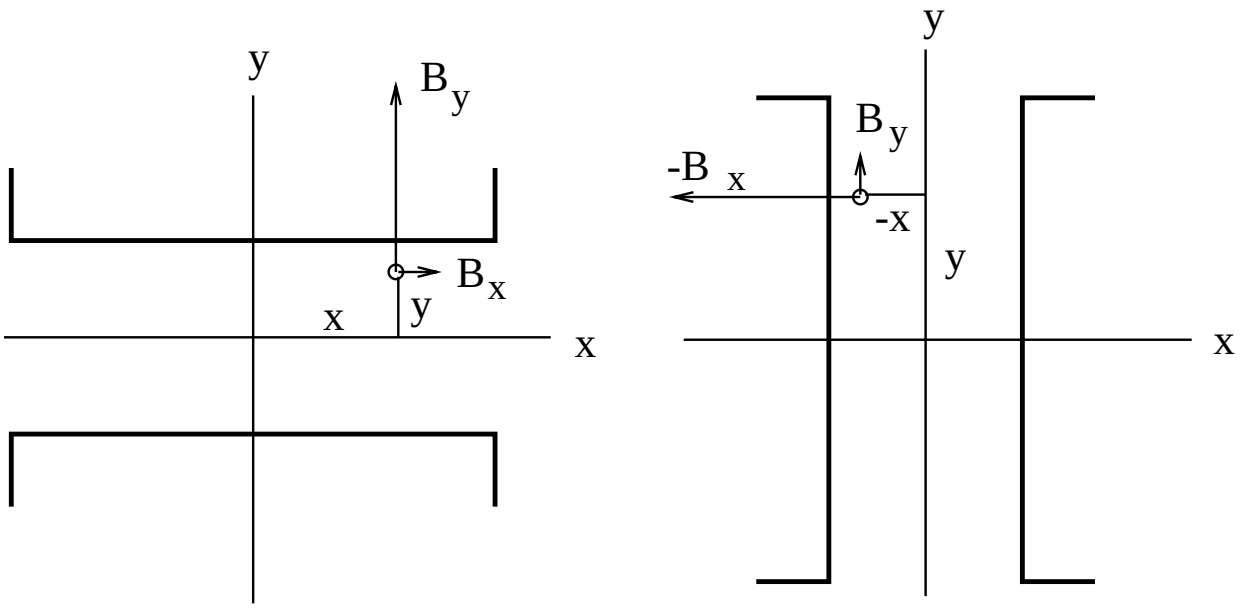

Figure 1.1: A horizontal steering dipole and the same magnet rotated so that it steers vertically. The median-plane field and its multipole approximation is shown in Fig. 1.2. The analytic description of the rotated fields is given in the text.

If the field of the erect magnet is described by the series Eq. (1.7) with the parameter set $(B_0 = B_{0y}, b_2, a_2 = 0, b_4, a_4 = 0)$ then field of the same magnet, anti-clockwise rotated by 90 degrees, will be described by the same series Eq. (1.7) with the parameter set $(B_0 = B_{0x}, b_2 = 0, a_2 = -b_2(\text{old}), b_4 = 0, a_4 = b_4(\text{old}))$

Problem 1..7. Check all entries in Table 1.1.

Problem 1..8. An erect quadrupole is misaligned from its design orientation by a small roll angle $\Delta\phi << 1$ around the longitudinal axis. In its natural (x', y') coordinates, its multipole expansion is

$$B_{y'} + iB_{x'} = b_1' \left(x' + iy'\right), \tag{1.20}$$

and in the design lattice coordinates it is

$$B_y + iB_x = (b_1 + ia_1) \left(x' + iy'\right). \tag{1.21}$$

Show, to lowest order in $\Delta\phi$, that

$$b_1 = b_1', \quad a_1 = -2b_1'\Delta\phi. \tag{1.22}$$

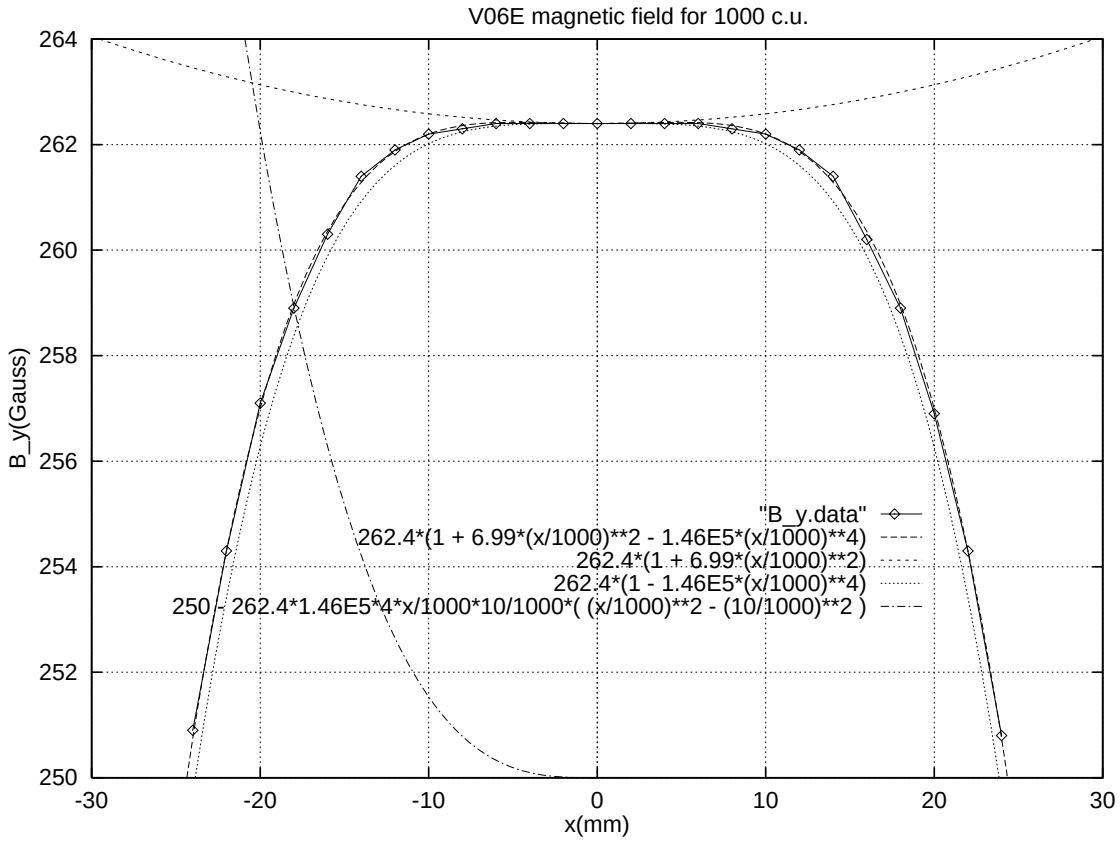

Figure 1.2: Midplane magnetic field B_y in a transversely-limited erect dipole magnet. Curves with sextupole or decapole "turned off" are also plotted, as well as a curve showing the decapole contribution to the off-median-plane, "wrong field component" $B_x(x, y = 10\,\mathrm{mm})$ (displaced upwards by 250 Gauss for plotting purposes.). The multipole coefficients are $b_2 = 6.99\,\mathrm{m}^{-2}$ and $b_4 = -1.46 \times 10^5\,\mathrm{m}^{-4}$. When the same magnet is used for vertical steering the non-vanishing coefficients are $a_2 = -6.99\,\mathrm{m}^{-2}$ and $a_4 = -1.46 \times 10^5\,\mathrm{m}^{-4}$. The field calculation is due to Sasha Temnyck.

In modeling the effect of a misaligned quad the factor of 2 in Eq. (1.22) must not be overlooked. A mnemonic for remembering it is that a quadrupole need only be rotated through angle $\pi/4$ (not $\pi/2$) for pure b_1 to become pure a_1.

Transverse Momentum Impulses Occurring at Multipole Planes

Deflection by a General Multipole

Following the orbit of a particle in an accelerator amounts to keeping track of the effects of transverse kicks or, if there is acceleration, of longitudinal kicks. Commonly the longitudinal coordinate s is adopted as independent variable and the abbreviations $x' = dx/ds$ and $y' = dy/ds$ are used. The pair x, x' are often called "1D phase space coordinates" even though, technically, this terminology should be reserved for x, p_x where p_x is the true Hamiltonian variable conjugate to x. In some formalisms the "momentum variable" is taken to be transverse velocity v_x or θ_x, the angle of the trajectory. To avoid the primes, and to provide for the the various possible definitions, the symbol f, approximately dx/ds, and g, approximately dy/ds, will be used as transverse "momenta", so the transverse "phase space" coordinates will be x, f, y, g. For linearized (otherwise known as paraxial or Gaussian) treatment, the propagation from beginning to end of a sector can be represented by a 4×4 "transfer matrix". If nonlinear effects are to be included the matrix has to be generalized to a "transfer map" which gives the output coordinates as functions of the input coordinates.

The evolution of the transfer map is more general than the orbit of a particular particle. In effect the map represents the orbits of all possible particles. Allowing for the possibility of the reference orbit being slightly displaced, by introducing phase space displacements $\Delta = (\Delta x, \Delta f, \Delta y, \Delta g, \Delta l, \Delta h)$ the trajectory can be expressed

$$\begin{aligned} x^{(i)} &= \overline{x}^{(i)} + \widehat{X}^{(i)}_{\beta}\Delta^{\beta}, \quad f^{(i_-)} = \overline{f}^{(i_-)} + \widehat{F}^{(i_-)}_{\beta}\Delta^{\beta}, \\ y^{(i)} &= \overline{y}^{(i)} + \widehat{Y}^{(i)}_{\beta}\Delta^{\beta}, \quad g^{(i_-)} = \overline{g}^{(i_-)} + \widehat{G}^{(i_-)}_{\beta}\Delta^{\beta}. \\ l^{(i)} &= \overline{l}^{(i)} + \widehat{L}^{(i)}_{\beta}\Delta^{\beta}, \quad h^{(i_-)} = \overline{h}^{(i_-)} + \widehat{H}^{(i_-)}_{\beta}\Delta^{\beta}. \end{aligned} \tag{1.23}$$

The overhead tildes, as in $\widehat{X}^{(i)}_{\beta}$, allows formally for the possibility that the "matrix elements" depend on the coordinates (i.e. the "mapping" is nonlinear.) Symbols such as $\overline{f}^{(i_-)}$ indicate that the the momentum components are expressed "just before" element (i). The "linear transfer map" part of Eqs. (1.23) describes "linearized" particle evolution

from some arbitrary starting location to lattice location (i) by

$$\begin{pmatrix} \Delta x^{(i)} \\ \Delta f^{(i)} \\ \Delta y^{(i)} \\ \Delta g^{(i)} \\ \Delta l^{(i)} \\ \Delta h^{(i)} \end{pmatrix} = \begin{pmatrix} X_1^{(i)} & X_2^{(i)} & X_3^{(i)} & X_4^{(i)} & X_5^{(i)} & X_6^{(i)} \\ F_1^{(i)} & F_2^{(i)} & F_3^{(i)} & F_4^{(i)} & F_5^{(i)} & F_6^{(i)} \\ Y_1^{(i)} & Y_2^{(i)} & Y_3^{(i)} & Y_4^{(i)} & Y_5^{(i)} & Y_6^{(i)} \\ G_1^{(i)} & G_2^{(i)} & G_3^{(i)} & G_4^{(i)} & G_5^{(i)} & G_6^{(i)} \\ L_1^{(i)} & L_2^{(i)} & L_3^{(i)} & L_4^{(i)} & L_5^{(i)} & L_6^{(i)} \\ H_1^{(i)} & H_2^{(i)} & H_3^{(i)} & H_4^{(i)} & H_5^{(i)} & H_6^{(i)} \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta f \\ \Delta y \\ \Delta g \\ \Delta l \\ \Delta h \end{pmatrix}. \tag{1.24}$$

Using index notation and summation convention, the evolution is given by

$$\Delta^{(i)\alpha} = X_\beta^{(i)\alpha} \Delta^\beta. \tag{1.25}$$

where X, as in $X_\beta^{(i)\alpha}$, stands for any one of the matrix elements.

With m, e, $\mathbf{p}$, $\mathbf{v}$, E, γ, and δ being the particle rest mass, charge, momentum, velocity, energy, relativistic factor, and fractional momentum offset, respectively, we have

$$\begin{aligned} v &= \sqrt{v_x^2 + v_y^2 + v_z^2} \\ p &= \sqrt{p_x^2 + p_y^2 + p_z^2} \\ E &= \sqrt{p^2c^2 + m^2c^4} = m\gamma c^2, \\ \mathbf{p} &= \frac{E\mathbf{v}}{c^2} = \mathbf{p}_0 \left(1 + \delta\right), \end{aligned} \tag{1.26}$$

where $\mathbf{p}_0$ is the momentum of a reference particle traveling along the ideal closed orbit, passing through the center of every element's design location; $\mathbf{v}_0 = \mathbf{p}_0 c^2/E_0$ is the corresponding velocity. Though the variable δ is normally small compared to 1, we will not use it as an expansion parameter, and Eqs. (1.26) are all exact.

If (f, g, h) are true canonical variables (normalized by the nominal momentum) they are related to true velocities $(v_x = dx/dt, v_y = dy/dt, v_z = dz/dt)$ and true momenta (p_x, p_y, p_z) by the relations

$$\begin{aligned} f &\equiv \frac{p_x}{p_0} = \frac{c}{v_0}\frac{\gamma}{\gamma_0}\frac{v_x}{c} = \frac{dx}{dl}, \\ g &\equiv \frac{p_y}{p_0} = \frac{c}{v_0}\frac{\gamma}{\gamma_0}\frac{v_y}{c} = \frac{dy}{dl}, \\ h &\equiv \frac{p_z}{p_0} = \frac{c}{v_0}\frac{\gamma}{\gamma_0}\frac{v_z}{c} = \frac{dz}{dl}; \end{aligned} \tag{1.27}$$

$$\mathbf{p} = p_0 \left(f\widehat{\mathbf{x}} + g\widehat{\mathbf{y}} + h\widehat{\mathbf{s}}\right), \tag{1.28a}$$

$$p^2 = p_0^2 \left(f^2 + g^2 + h^2\right). \tag{1.28b}$$

where γ and γ_0 are the usual relativistic factors of the particle being tracked and of the ideal design particle respectively. To linear order l and s are equivalent. To that order then, Eq. (1.27) shows that f and g are identical to the "traditional" transverse slope variables $x' = dx/ds$ and $y' = dy/ds$.

As mentioned before, with alignment and field errors present, the perturbed closed orbit deviates from the ideal closed orbit. Systematically in these notes, overhead bars label parameters of the reference particle traveling on the perturbed closed orbit (which from now on will simply be called the "closed orbit".) Hence the closed orbit momentum is $\overline{\mathbf{p}}$, and the corresponding velocity is $\overline{\mathbf{v}}$.

At this point one is tempted to treat transverse and longitudinal variables asymetrically, by treating the variable p, or its equivalent δ, as a parameter rather than as a dependent variable. Its treatment then would be essentially different from that accorded the quartet of transverse dependent variables (x, f, y, g) being tracked. Such a separation is firmly established in the "lore" of accelerators. It is "justified" by physical considerations according to which longitudinal oscillations (normally called "synchrotron oscillations") occur on a much slower time scale than transverse ("betatron") oscillations. This makes it tempting to regard δ as varying so slowly that it can be temporarily treated as constant. The very term "chromaticity" defines a lattice property that is applicable to a particle for which δ is constant. Unfortunately, after this approximate picture has been too wholeheartedly adopted, it is hard to figure out which results are independent of this assumption or to repair those that are not. For example it is difficult to describe "synchrobetatron" effects or "transition crossing" rigorously. At least for now we will proceed more carefully.

Introducing l as the longitudinal coordinate of a particle, relative to a central, or reference, particle, and h as longitudinal momentum the phase space becomes 6 dimensional with a particle position specified by a 6 component vector (x, f, y, g, l, h). This complicates the use of the fractional momentum offset δ. When (f, g, h) are expressed as the output of a map we must, in principle, regard δ also as an output of the map. This means that the map itself has to be such as to preserve relations (1.26). When tracking individual particles it is easy to enforce this constraint since one can update f and g and then, using the constancy of p, update h by using Eqs. (1.26). In the map approach p must be calculated from updated values of (f, g, h); its constancy can used as a check of the calculation.

One must resist being erroneously reassured by the fact that δ, being conserved in purely magnetic elements, can be treated as a simple constant. To see this, contemplate the use of the map to generate the time evolution of a "beam" made up of a number of particles, each having its own individual value of δ. Since the same map applies to all particles, it must contain within itself dependencies on δ (such as the fact that quadrupoles are effectively weaker or stronger for off-momentum particles). It is clear then that δ, γ, p, E, *etc.*, cannot be regarded as constant, uniformly over all elements of the map. On the other hand, for the map to be correct, it must conserve, to its claimed accuracy, the values of δ, γ, p, E, *etc.*, for each individual particle passing through a pure magnetic element. Below we will use terminology "δ is not constant *over* the map" to recall this discussion. The same restriction applies to each of the variables v, p, E, (p/p_0), γ, and δ; they are all essentially equivalent in that any one can be caculated starting from any other. In contrast, quantities like mass and charge, quantities with subscript 0, like p_0, and all quantities with overlines, like $\overline{p}$, or underlines, like $\underline{\Delta x}$, *are* constant over the map.

The step about to be taken, though simple, is probably the most important ingredient of the thin element description. It amounts to justifying the limiting process in which the magnet length is allowed to approach zero: $L \to 0$. With L finite, the magnet field is large and roughly constant only in the range $-L/2 < z < L/2$. The particle trajectory through the magnet must be obtained by integrating the equation of motion over this range, or somewhat beyond, if the more accurate fields $\mathbf{B}^T$ are used. The equation of motion is obtained from the Lorentz force equation of electricity and magnetism;

$$\frac{d\mathbf{p}}{dt} = e\,\mathbf{v} \times \mathbf{B}^T = e \det \begin{vmatrix} \hat{\mathbf{x}} & \hat{\mathbf{y}} & \hat{\mathbf{z}} \\ v_x & v_y & v_z \\ B_x^T & B_y^T & 0 \end{vmatrix}, \tag{1.29}$$

which, using $v_z = dz/dt$, reduces to

$$dp_x = -eB_y^T\left(x, y, z\right) dz, \quad dp_y = eB_x^T\left(x, y, z\right) dz. \tag{1.30}$$

Integrating these equations along the particle trajectory and completely through the magnet length yields

$$\begin{aligned} \Delta p_x =& -e\int_{\infty}^{\infty} B_y^T\left(x\left(z\right), y\left(z\right), z\right) dz \overset{L\to 0}{\longrightarrow} -eLB_y, \\ \Delta p_y =& e\int_{\infty}^{\infty} B_x^T\left(x\left(z\right), y\left(z\right), z\right) dz \overset{L\to 0}{\longrightarrow} eLB_x. \end{aligned} \tag{1.31}$$

The unknown dependencies $x(z)$ and $y(z)$ would have prevented the direct evaluation of the integrals if L were finite. But in the $L = 0$ limit, though the integrands continue to depend on x and y, their dependence on z can be ignored—though the slope of the trajectory is altered, the transverse displacement does not change along the trajectory in this limit, because of the infinitely foreshortened longitudinal range. The integrals remaining are the same as those used in defining thin element multipole strengths, such as in Eqs. (1.2).

Substituting the multipole expansions (1.11) into these equations yields equations for the transverse deflections caused by a thin element;

$$\begin{aligned}
\Delta f = f^{(+)} - f^{(-)} = -\tilde{B}_y &= -\Re \sum_{n=0}^{M} \left(\tilde{b}_n + i\tilde{a}_n\right) \left((x - \underline{\Delta x}) + i \left(y - \underline{\Delta y}\right) \right)^n , \\
\Delta g = g^{(+)} - g^{(-)} = \tilde{B}_x &= \Im \sum_{n=0}^{M} \left(\tilde{b}_n + i\tilde{a}_n\right) \left((x - \underline{\Delta x}) + i \left(y - \underline{\Delta y}\right) \right)^n .
\end{aligned} \tag{1.32}$$

Having determined $f^{(-)}$ and $g^{(-)}$ previously, these "kink equations" enable us to obtain $f^{(+)}$ and $g^{(+)}$ immediately. Knowing these quantities amounts to knowing the particle direction as it exits the thin element. From that point the orbit will be the straight line having that direction until the next thin element is encountered.

Deflection by a Dipole

Because dipole elements determine the closed orbit, and the expansion Eq. (1.7) singles out b_0 and relates it to B_0 as in Eq. (1.8), our treatment of dipoles will not be quite consistent in form with that of all other multipoles. To represent imperfect dipoles a new coefficients Δb_0 will be introduced, as well as a new abbreviation, $\tilde{b}_{00} = LB_0/\text{"}B\rho\text{"}$. (The extra subscript on $\tilde{b}_{00}$ will be justified shortly.) First we consider the design orbit. The entrance momentum components $(f_0^{(-)}, g_0^{(-)}, h_0^{(-)})$, in the local coordinates (horizontal but skewed by the direction of $\hat{\mathbf{x}}$,) are assumed to be known. Applying Eq. (1.32) with yields

$$f_0^{(+)} = f_0^{(-)} - \tilde{b}_{00}, \quad g_0^{(+)} = g_0^{(-)}. \tag{1.33}$$

Referring to the geometry of Fig. 1.3 (in which angular deflection $\Phi^{(+)}$ occurs in the multipole plane, causing the trajectory laboratory angle to change from $\Theta^{(-)}$ to $\Theta^{(+)}$) we obtain

$$\frac{\Phi^{(+)}}{2} = \sin^{-1} \frac{p_0 f_0^{(-)}}{p_0} = \sin^{-1} \frac{\tilde{b}_{00}}{2}. \tag{1.34}$$

For small bend angles,

$$\Phi^{(+)} \simeq b_{00} = \frac{LB_0}{\text{“}B\rho\text{”}}. \tag{1.35}$$

(This would be an exact equality for arbitrary bend angles if L were the arc length of the circular arc through the magnet, but L is usually taken to be the straight line distance through the magnet. By breaking a magnet into short enough segments this distinction can be made insignificant.) Because the dipole bend is horizontal g_0 is unaffected and, by symmetry, neither is h_0. The outgoing momentum of the design orbit is

$$f_0^{(+)} = -\frac{1}{2}\tilde{b}_{00}, \quad g_0^{(+)} = 0, \quad h_0^{(+)} = h_0^{(-)}. \tag{1.36}$$

Repeated application of this formula, alternated with accounting for drift regions, completes the determination of the closed orbit.

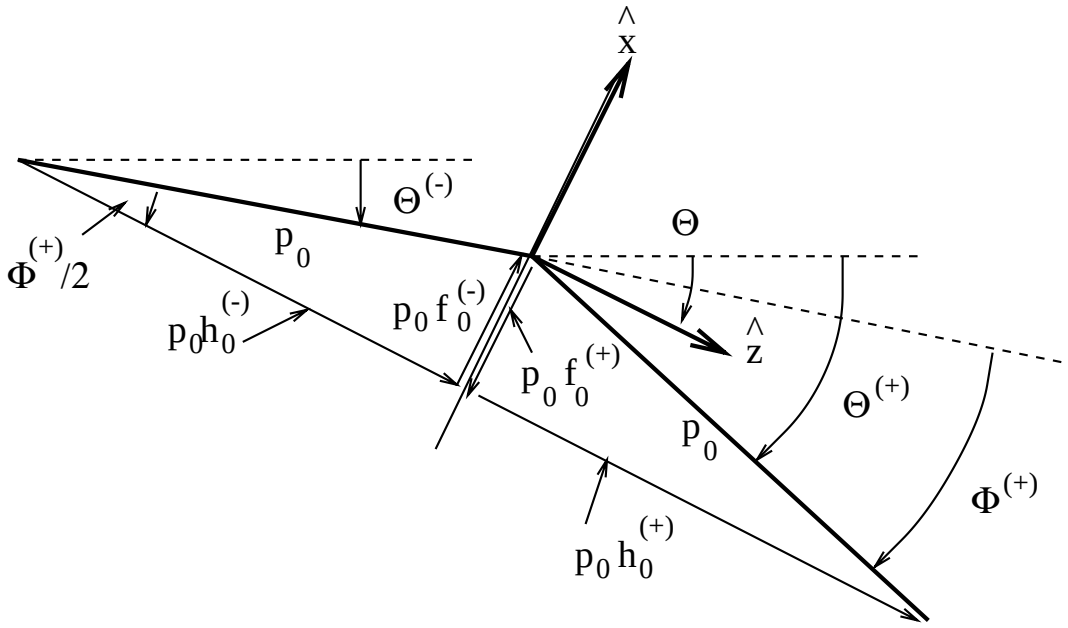

Figure 1.3: Geometry of the design orbit as it passes through a thin element bending through angle $\Phi^{(+)}$. Parameters of the design orbit are denoted by subscript 0.

Before discussing the effect of the dipole on a general trajectory we admit the possibility that the true bend deviates slightly from its ideal value (as is always the case at some level.) To handle this define

$$b_0 = 1 + \Delta b_0, \quad \tilde{b}_0 = \tilde{b}_{00} + \widetilde{\Delta b_0}. \tag{1.37}$$

The definitions of all other multipole coefficients are unaffected by Δb_0, but even if they were, the effect would typically be negligible as Δb_0 rarely exceeds 0.001. As stated previously, the design orbit is not permitted to have vertical deflections, but a small vertical error bend can be handled by permitting $\tilde{a}_0$ to be non-vanishing. The effects of $\widetilde{\Delta b_0}$ and $\tilde{a}_0$ can be treated together, starting with Eq. (1.11)

$$\tilde{B}_y = \tilde{b}_0, \quad \tilde{B}_x = \tilde{a}_0. \tag{1.38}$$

Substituting these and Eqs. (1.23), (1.24) and (1.32) yields for the evolution of the phase space coordinates,

$$\begin{aligned} \Delta\overline{f} &= -\,\tilde{b}_0, \\ \Delta\overline{g} &= \tilde{a}_0, \\ \Delta\widehat{F} &= \text{to be determined}, \\ \Delta\widehat{G} &= 0. \end{aligned} \tag{1.39}$$

The final two entries here have been included as a reminder that it will be important also to determine the effect of the dipole on the transfer map. In particular the coefficient $\widehat{F}$ is changed but, by symmetry $\Delta\widehat{G} = 0$, at least for the leading power in a Taylor expansion. Other than a small focusing effect to be determined in Eq. (1.45), the dipole element has no direct effect on the transfer matrix; but a dipole does alter the true closed orbit, which indirectly alters the transfer matrix through feed-down from nonlinear elements.

Problem 1..9. A "roll" error is an azimuthal rotation around the the design longitudinal axis. Calculate values of $\widetilde{\Delta b_0}$ and $\tilde{a}_0$ that model a roll error of magnitude $\Delta\phi$ for a dipole of strength $\tilde{b}_0$.

Deflection by a Quadrupole

Next consider the map discontinuity at an ideal, erect quadrupole. In this case the factors $\tilde{B}_y$ and $\tilde{B}_x$ are given by

$$\tilde{B}_y = \tilde{b}_1 \left(x - \underline{\Delta x}\right), \qquad \tilde{B}_x = \tilde{b}_1 \left(y - \underline{\Delta y}\right). \tag{1.40}$$

Substituting into Eq. (1.32) and matching coefficients order by order yields

$$\begin{aligned} \Delta\overline{f} &= -\tilde{b}_1 \left(\overline{x} - \underline{\Delta x}\right), \\ \Delta\overline{g} &= \tilde{b}_1 \left(\overline{y} - \underline{\Delta y}\right), \\ \Delta\widehat{F}_\beta &= -\tilde{b}_1 \widehat{X}_\beta, \\ \Delta\widehat{G}_\beta &= \tilde{b}_1 \widehat{Y}_\beta. \end{aligned} \tag{1.41}$$

Recognizing the first two equations as "thin lens" equations one sees that $\tilde{b}_1$ is just $1/f$, the "inverse focal length" of the quadrupole, positive in one plane, negative in the other. From now on, for brevity, we will use the term "focusing strength" to describe $\tilde{b}_1$, though the longer term, inverse focal length, is more descriptive. It has been defined in such a way as to "scale away" the momentum dependence. Notice that a positive value of $\tilde{b}_1$ corresponds to horizontal focusing, vertical defocusing. Like the bending, the focusing effect is being treated as purely geometric, independent of momentum. We know, of course, that the focal length varies proportional to momentum. Since this effect has been accounted for by factoring out central momentum p_0 in definition (1.10) of $\tilde{b}_1$ we might be tempted to make a "chromatic" correction proportional to $\delta = (p - p_0)/p_0$ for off-momentum values—that would be a traditional procedure. But in our formalism it would be wrong since Eqs. (1.32) yield the correct momentum deviation, independent of momentum. In fact Eqs. (1.41) include all chromatic effects as they stand—finally, we get a modest return in simplicity for the large investment that has been made in formalism!

This separation has been especially simple because the quadrupole field has only linear terms; as a result, order by order, $\widehat{X}_\beta$ influences only $\widehat{F}_\beta$ and $\widehat{Y}_\beta$ influences only $\widehat{G}_\beta$. These formulas will be re-expressed in more familiar language in the next chapter.

Already the phenomenon called "feed-down" has made an appearance. For example, the term $\tilde{b}_1(\overline{x} - \underline{\Delta x})$ in the expression for $\overline{f}^{(+)}$ represents a "steering effect", due to the positioning error of the quadrupole. This is non-ideal behavior since quadrupoles are

supposed to focus, not steer. This steering affect would already have been noted and accomodated in the preliminary closed orbit determination.

The complication of linear coupling intrudes when skew quadrupoles (or nonlinear elements, in most cases) are present. Such elements will mix (*i.e.* bring in linear combinations of) X_β and Y_β in the expressions for F_β or G_β. These terms alter the linearized description, but without bringing in nonlinearity. As a result they are rather easily manageable.

Higher orders terms of Eq. (1.7), due to sextupole, octupole, and so on, being quadratic or higher in (x, y), introduce nonlinearity. For sufficiently small amplitudes they are negligeable, but for large amplitudes they become dominant, and always destructive to particle stability at some point. They also cause feed-down, in particular feed-down into the linear regime. For example, a displaced sextupole causes a steering and in addition, a focusing effect, which is reflected in the linear transfer matrix.

Problem 1..10. Show that the steering errors suffered by the design orbit as it passes through a quadrupole of strength $q = \tilde{b}_1$ with displacement $(\underline{\Delta x}, \underline{\Delta y})$ from its design location are given by

$$\Delta\theta_x = q\underline{\Delta x}, \text{ and } \Delta\theta_y = -q\underline{\Delta y}.$$

Problem 1..11. Show that the linearized relations giving $(\Delta x^{(+)}, \Delta x'^{(+)})$ output particle coordinates just after a quadrupole of strength $q = \tilde{b}_1$ in terms of inputs $(\Delta x^{(-)}, \Delta x'^{(-)})$ are given by

$$\begin{pmatrix} \Delta x^{(+)} \\ \Delta x'^{(+)} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -q & 1 \end{pmatrix} \begin{pmatrix} \Delta x^{(-)} \\ \Delta x'^{(-)} \end{pmatrix},$$

and the corresponding vertical relation is

$$\begin{pmatrix} \Delta y^{(+)} \\ \Delta y'^{(+)} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ q & 1 \end{pmatrix} \begin{pmatrix} \Delta y^{(-)} \\ \Delta y'^{(-)} \end{pmatrix}.$$

Dipole Focusing and Dipole-Edge Focusing

There are two focusing effects of dipoles that our formalism does not yet encompass. Some people might consider these as two manifestations of the same thing, but we treat them quite differently. One of these effects, edge-focusing, will be easily represented by the artificial introduction of an effective quadrupole, and analysed like any true quadrupole. The best example for which this focusing is important is the ZGS (zero gradient synchrotron.) Before analysing that case we study a horizontal focusing effect that is significant in small accelerators for which the radius of curvature is comparable to or less than the focal length of the main focusing quadrupoles. We will call it "curvature focusing". In the extreme case of the weak-focusing synchrotron, there are no quadrupoles and curvature focusing is dominant.

We take the opportunity in this section to illustrate the philosophy that permeates our entire treatment of accelerator modeling. Our motto is "use fairly good physics and perfect mathematics". "Fairly good physics" certainly includes the requirement that element idealizations not violate the laws of physics. But consistent with that requirement we will use casual, almost slipshod, arguments in approximating the two new dipole effects, and will make no attempt to go beyond lowest order. This may seem inconsistent with our stated purpose of writing "exact" formulas. But it is not. We analyse only idealizations of the actual lattice, but we restrict the idealization to the "physics" and do not let it intrude into the "mathematics". If there is an element present in the ring that is approximately an ideal thin quadrupole, we treat it as if it were exactly an ideal thin quadrupole. The resulting simplification is sufficiently great that subsequent mathematical analysis can be "exact". It should be obvious to everyone that lattice descriptions are necessarily only approximate; we consider it important to segregate physics assumptions (which might in some cases be called conjectures) from mathematics approximations (which should be admitted only when the errors they are capable of making are under control.) If our physical idealizations turn out to be too crude we will have to come back and improve them later.†

† This discussion of the interplay of mathematics and science should be entirely superfluous. But in accelerator physics there is perhaps too great a tendency to confuse mathematical necessity or convenience with physical assumption.

Fig. 1.4 illustrates particle trajectories in two standard bending elements, a sector bend and a rectangular bend. The design trajectory enters and leaves a "sector bend" normal to the pole faces. This means the pole faces are necessarily not parallel. If the faces *are* parallel the magnet is called a "rectangular bend". One reason rectangular magnets are in common use is that they (especially short, laminated ones) are easier to fabricate than sector magnets. It is traditional in both cases to let L stand for distance along the design orbit. Our procedure is to let L approach zero, holding $B_0 L$ constant. It should be clear that this limiting process is hazardous since the two pictures look different for large L but the same for small L.

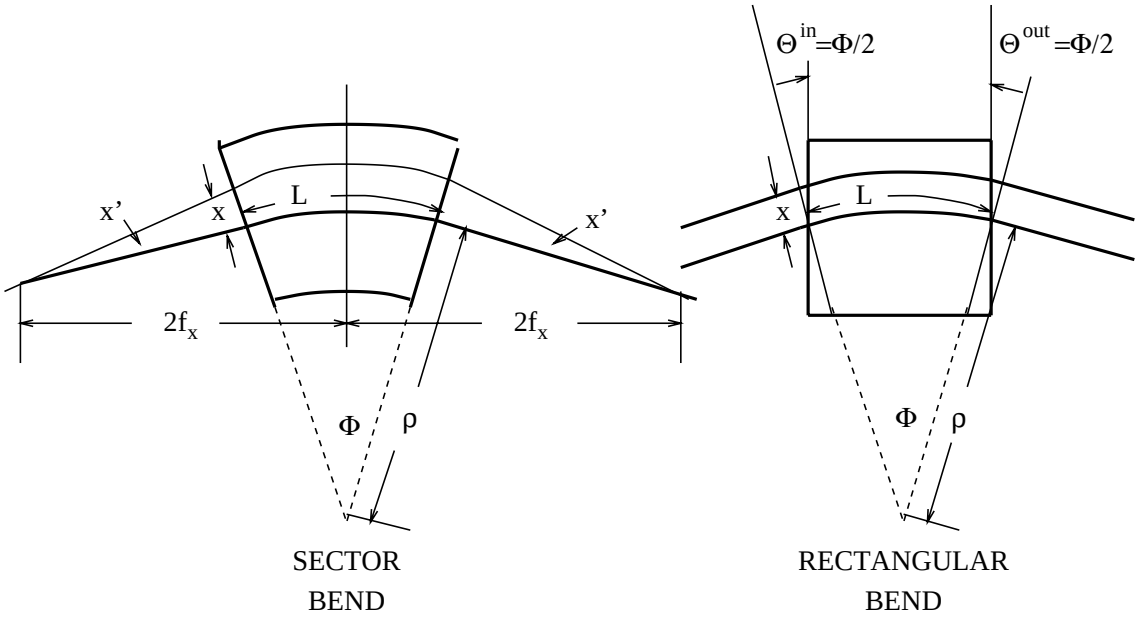

Figure 1.4: Comparison between a sector bend and a rectangular bend magnet. The central orbit is identical, with arc length L and radius of curvature ρ in both cases. The off-axis orbit is focused in one case but not in the other.

Because edge focusing can later be handled separately it is sufficient to restrict the formalism to sector bends. For modeling rectangular magnets (or magnets with other pole face angles) we will use sector bends with extra elements at both ends to model the pole-face rotations. Just by glancing at the trajectories drawn in Fig. 1.4 it should be obvious that there is a horizontal focusing effect for the sector bend that is not present for the rectangular bend. Since our discussion of dipole deflections in Section 1.2 included

no curvature focusing effect, it is clear that discussion has to be modified—provision for a term to be added later was included in Eq. (1.39).

To understand the following discussion, Fig. 1.4 should be contemplated under the assumption that the magnet length is much less than both the radius of curvature ρ and the focal length f, *i.e.* $\Phi = L/\rho << 1$. This requires a bit of mental contortion since, as drawn, the figure does not really meet this requirement. Another manifestation of the distorted geometry is that it violates a remarkable theorem according to which object, image, and center of curvature lie on the same straight line for a sector bend.[1] Also it is possible, though rare, for Φ to be negative, which would correspond to an outwards bend. In that case there would be possible sign changes in some of the following equations. The focusing effect of the sector bend is due to the extra distance traveled in the magnetic field by the trajectory with displacement x. (In our approximation, this displacement can be treated as constant through the magnet.) This causes bend $\Phi + 2x'$ which is greater than the nominal bend Φ in the ratio $(\rho + x)/\rho$; or (to lowest order)

$$x' = \frac{\Phi}{2\rho} x. \tag{1.42}$$

The extra deflection $\Delta x' = -2x'$ bends the trajectory back toward the focus; this focusing action of the sector bend can be ascribed to an effective focusing strength $\tilde{b}_1^\rho = -\Delta x'/x$, which, with Eq. (1.42), yields

$$\tilde{b}_1^\rho = \frac{\Phi}{\rho}. \tag{1.43}$$

Here we use Eq. (1.41) which identified $\tilde{b}_1$ as the inverse focal length, or focusing strength, of a quadrupole. From these two equations, referring to Eqs. (1.10) and (1.35), we can introduce an effective multipole coefficient b_1^ρ by

$$b_1^\rho = \frac{\text{“}B\rho\text{”}}{LB_0} \tilde{b}_1^\rho \simeq \frac{1}{\Phi} \frac{\Phi}{\rho} = \frac{1}{\rho}. \tag{1.44}$$

The positive numerical value of b_1^ρ (except in the rare case of an outward bend, which would be represented by negative ρ) is consistent with the effect being *horizontally focusing*, which was obvious from Fig. 1.4.

An extreme case, with $\Phi = \pi/2$, is shown in Fig. 1.5. For the image and object distances to work out as shown, it is necessary to replace Eq. (1.43) by $\tilde{b}_1^\rho = \sin\Phi/\rho$. Even if that is done, the thin and thick element representations differ in arc length as the can

be seen from the figure. For these reasons, it would be foolhardy to represent a $\pi/2$ bend by one thin element. But by breaking the ninety degree bend into, say, ten equal sectors, the distinction between Φ and $\sin\Phi$ becomes unimportant, and the thin and thick lens arc lengths become equal to high accuracy.†

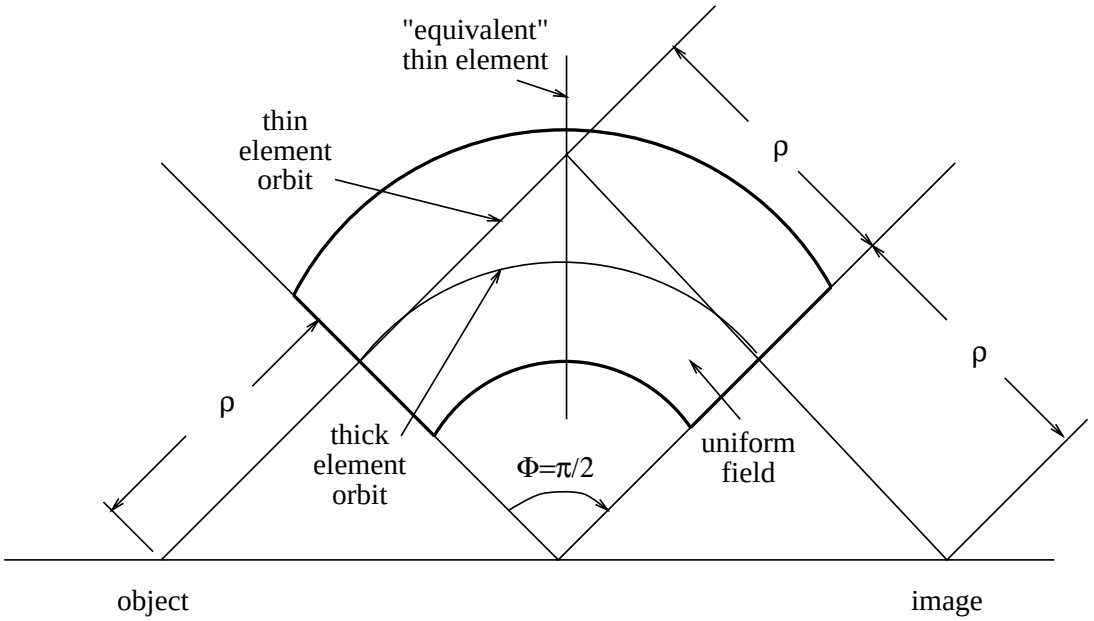

Figure 1.5: Trajectories through a 90 degree sector bend. Such an angle is far too great to be sensibly replaced by a thin element. Though replacing Φ by $\sin\Phi$ makes Eq. (1.43) give the correct object/image geometry, the length of "thin element orbit" is appreciably longer than the correct " thick element orbit length".

We are now in a position to complete Eq. (1.39), which updates the map at a sector bend, and which we repeat here

$$\begin{aligned}\Delta\overline{f} &= -\tilde{b}_0,\\ \Delta\overline{g} &= \tilde{a}_0,\\ \Delta\widehat{F} &= -\frac{\tilde{b}_0}{\rho}\widehat{X}_\beta,\\ \Delta\widehat{G} &= 0.\end{aligned} \tag{1.45}$$

† In *TEAPOT*, curvature focusing is modeled by $\tilde{b}_{10} \equiv \tilde{b}_1^\rho = \sin\Phi/\rho$. This was obtained by matching to the thick lens transfer matrix of a dipole. This mild disagreement with Eq. (1.43) should not cause concern since the disagreement is only in cubic terms. This is all the more true since even the curvature effect is usually almost negligible, and the cubic terms can be made fractionally less important by subdividing the sector bend.

The sector bend can be subdivided into shorter sector bends in order to reduce errors due to the use of thin elements, but that does not reduce the fractional importance of the curvature focusing. That is because $\tilde{b}_0$, which reduces proportional to L as the magnet is subdivided, appears linearly in both the bend term (first equation) and the focusing term (third equation.)

The focusing properties of a sector bend are traditionally represented by transfer matrices (remember that the upper case coefficients relate deviations from the central orbit);

$$\begin{aligned}
\begin{pmatrix} \Delta x^{(+)} \\ \Delta x'^{(+)} \end{pmatrix} &= \begin{pmatrix} 1 & \rho\Phi/2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -\Phi/\rho & 1 \end{pmatrix} \begin{pmatrix} 1 & \rho\Phi/2 \\ 0 & 1 \end{pmatrix} \\
&\simeq \begin{pmatrix} 1-\Phi^2/2 & \rho\Phi\left(1-\Phi^2/4\right) \\ -\Phi/\rho & 1-\Phi^2/2 \end{pmatrix} \begin{pmatrix} \Delta x^{(-)} \\ \Delta x'^{(-)} \end{pmatrix}, \qquad (1.46) \\
\begin{pmatrix} \Delta y^{(+)} \\ \Delta y'^{(+)} \end{pmatrix} &= \begin{pmatrix} 1 & \rho\Phi \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \Delta y^{(-)} \\ \Delta y'^{(-)} \end{pmatrix}.
\end{aligned}$$

A small bend angle, $\Phi << 1$, has been assumed, but one must still be careful to approximate elements consistently to assure that each determinant value is exactly 1.

There is a certain lack of elegance in our treatment, which replaces a curvature effect by a focusing effect. It would avoid this, and be conceptually somewhat more satisfactory, to model the trajectory in the sector bend as the arc of a circle. We have three reasons for not doing this: it would bring in transcendental functions of the dynamical variables; it would make it difficult to model nonlinear dipole-field-imperfections; and the straight line model is simpler and has proved to be adequate.

The curvature focusing effect is restricted to the horizontal plane, which differentiates it from the focusing of an ordinary quadrupole which defocuses in one plane if it focuses in the other. In other words, the effect cannot be represented using the multipole expansion (1.7). One might worry that introducing focusing in one plane with none in the other would be "non-symplectic", but the worry would be unjustified. Reassurance can be obtained, for example, by realizing that ordinary optical lenses focus (or defocus) simultaneously in both planes.

We can now expand upon the statement made above, that curvature focusing is important only at low energies, E. For typical accelerators (with the same maximum magnetic

field), $\rho \sim E$, focusing strength $\tilde{b}_1 \sim E^{-1/2}$, and bend per half cell $\Phi \sim E^{-1/2}$. Combining these dependencies with Eq. (1.43), one finds that $\tilde{b}_1^{\rho}/\tilde{b}_1 \sim E^{-1}$, which supports the contention that curvature focusing becomes less important in higher energy accelerators.

The preceding discussion leads naturally into a discussion of dipole edge focusing. Looking at the right hand figure in Fig. 1.4, it is obvious that there is no net horizontal focusing in passing through a rectangular bending magnet. It is inescapable that the ends are causing defocusing just strong enough to cancel the horizontal focusing we have just become persuaded is caused by the uniform field. Furthermore this is made plausible by the little entrance and exit wedges, where the upward displaced trajectory in the rectangular bend misses bending field that it would see in the sector bend. Knowing the pole rotation angles to be $\Theta^{in} = \Theta^{out} = \Phi/2$, and assuming the focusing strength is proportional to the pole rotation angle, we obtain directly formulas for the rotated pole focusing effect. To cancel the curvature focusing, we model the edges with focusing strengths $\tilde{b}_1^{\rm in} = -\tilde{b}_1^{\rho}/2$ and $\tilde{b}_1^{\rm out} = -\tilde{b}_1^{\rho}/2$. We obtain then $\tilde{b}_1^{in} = -\frac{\Theta^{in}}{\rho} = -\Phi/(2\rho)$ and $\tilde{b}_1^{out} = -\frac{\Theta^{out}}{\rho} = -\Phi/(2\rho)$. The pole face angles are measured away from the sector bend face, with polarity (agreeing with *MAD*) such that positive angles lead to horizontal defocusing. We can corroborate these formulas by calculating the deflection discrepancies caused by the little missing field regions mentioned above. They are both $x\tan(\Phi/2)/\rho)$ which agrees provided $\tan(\Phi/2) \simeq (\Phi/2)$, which is satisfied in all practical cases. Though it might be considered "apple-polishing" to make this replacement, it costs nothing to use $\tan\Phi$, and that is what is usually done;

$$\tilde{b}_1^{in} = -\frac{\tan\Theta^{in}}{\rho}, \quad \tilde{b}_1^{out} = -\frac{\tan\Theta^{out}}{\rho}. \tag{1.47}$$

Naturally these formulas are not restricted to rectangular bend magnets, they apply to any dipole magnets with non-normal entry of the design orbit. Like the curvature focusing, the pole-edge focusing is mainly important at low energy. But unlike that case, we cannot improve the approximation by using thinner elements. For too-low energy, too-oblique entry, this approximation simply breaks down.

It is easily seen that pole edge focusing is like the focusing of an ordinary quadrupole—focusing in one plane is necessarily accompanied by equal and opposite focusing in the other plane. For an ordinary quadrupole this equal but opposite effect followed from Ampère's law, as in Eq. (1.1). It also followed from the multipole series (1.7).

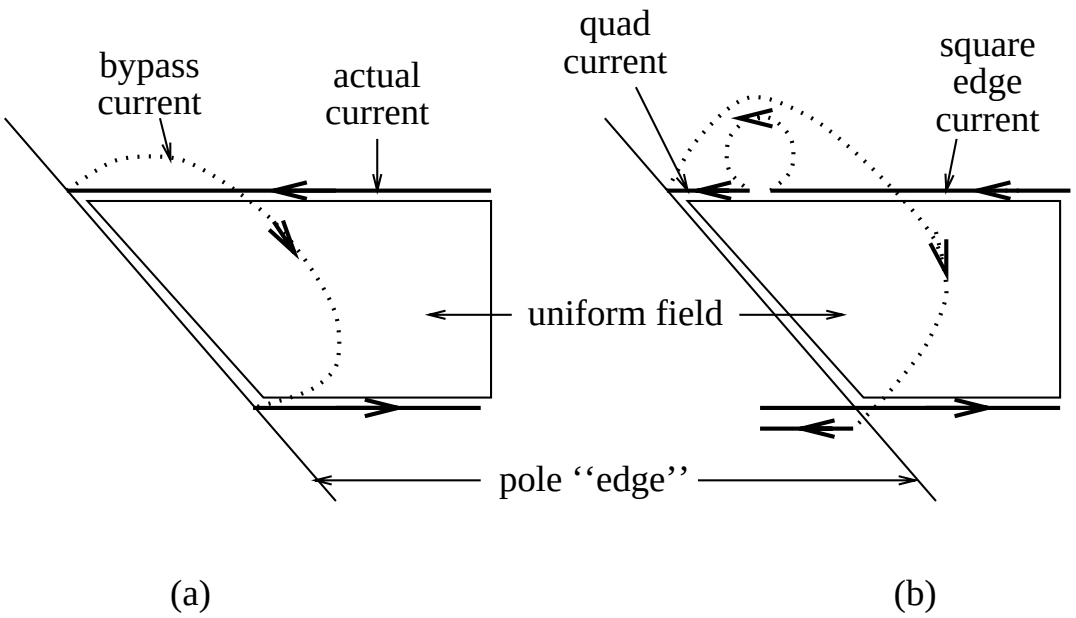

Figure 1.6: The magnetic field of an iron pole, uniform field magnet can be regarded as being primarily due to surface current sheets, shown in (a) that define the sides of the uniform field regions. With the entrance face slanted, these currents can be regarded as a superposition of “quadrupole” current sheets and square-pole current sheets, as shown in (b). Transverse bypass currents, shown dashed, are being neglected.

Referring to Fig. 1.6, it can be seen that the magnetic field near the end of a uniform-field magnet with non-normal entry can be regarded as due to short elements of longitudinal current, flowing (parallel) on the two sides of the beam. The remaining (anti-parallel) currents generate a sector bend, normal entry, face. The short parallel-directed end currents may or may not look to you like quadrupole windings, but we know from the above that they give horizontal focusing like a quadrupole. Also these currents are parallel to the axis. From the series of problems, in Section 1.1, dealing with the fields due to short current elements, we know that the field integrals of B_x and B_y are subject to the same relations as any multipole. In particular, there must be vertical defocusing with strength equal to the horizontal focusing.

By these arguments we conclude that the effect of non-normal entry into, and exit from, a dipole magnet can (to lowest order) be treated exactly as if there were actual quadrupoles at the ends, with the uniform field magnet itself being a pure sector bend. The focusing strengths of these quadrupoles are given by Eqs. (1.47). The map evolution

caused by the slanted poles is described by Eqs. (1.41), with $\tilde{b}_1$ obtained from Eqs. (1.47). Using this procedure we can, without loss of generality, treat any uniform field magnet, whatever its pole-face angles, as a sector bend, by placing artificial quadrupoles of correct strengths at entry and exit.

Another way of analysing the focusing due to rotated poles is to analyse Maxwell's equations near the pole. The bulging out of the fringe field at the magnet end, causes a field component B_x proportional to y, which is what causes the vertical focusing action. Integrating from well outside the magnet to well inside the the equality of the vertical defocusing strength to the horizontal focusing strength follows directly. See Steffen Ref. 1 for details.

There is an all-too-likely source of error due to curvature focusing or rotated pole-face focusing that can accompany the treatment of dipole magnets. If one halves the length of a quadrupole and doubles its strength, the effect on the lattice is usually small. But doing the same thing with dipole magnets can lead to curious focusing effects. The effect of $L \to L/2$ and $B_0 \to 2B_0$ is to leave the bend angle Φ unchanged. But it also has the effect of halving the radius of curvature $\rho \to \rho/2$ which, by Eq. (1.43), causes $\tilde{b}_1^\rho \to 2\tilde{b}_1^\rho$, doubling the curvature focusing. For rectangular bend magnets, Eqs. (1.47) show that cutting the length in half and doubling the field also doubles the edge focusing.

Just as the effect of a slanted dipole face is equivalent to the effect of a quadrupole, the effect of a slanted quadrupole face is equivalent to a sextupole, and similarly for higher multipoles.

Problem 1..12. Justify the previous sentence using the reasoning of the series of problems in Section 1.1. Also calculate the sextupole strength needed to model (horizontal) non-normal entry of the design orbit into a quadrupole of focusing strength $\tilde{b}_1$ and length l_Q. Hint: first calculate the excess deflection blameable on the pole slant angle θ_H for a particle with displacement x and then, using Table 1.1, determine the equivalent sextupole strength.

Problem 1..13. Determine the effective sextupole for modeling entrance at vertical angle θ_V into the quadrupole of the previous problem.

Problem 1..14. For a rectanglar bend dipole with bend radius of curvature ρ, assuming small deflection angle $\Phi << 1$, show that the linearized transfer matrix relations are

$$\begin{pmatrix} \Delta x^{(+)} \\ \Delta x'^{(+)} \end{pmatrix} = \begin{pmatrix} 1 & \rho\Phi \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \Delta x^{(-)} \\ \Delta x'^{(-)} \end{pmatrix}, \quad \begin{pmatrix} \Delta y^{(+)} \\ \Delta y'^{(+)} \end{pmatrix} = \begin{pmatrix} 1 & \rho\Phi \\ -\Phi/\rho & 1 \end{pmatrix} \begin{pmatrix} \Delta y^{(-)} \\ \Delta y'^{(-)} \end{pmatrix}.$$

Longitudinal Fields at Quadrupole Ends

This section follows Steffen.[2] Everywhere except in this section and in a later section discussing solenoids, longitudinal magnetic fields have been or will be neglected. One configuration for which this neglect is hard to justify is at the ends of "IR quads" which are quadrupoles adjacent to low-β intersection regions. Such quads are rarely thin and the central trajectory is often off-axis and slanted at that point. Since the β-functions are large, and very different at the two ends, the end effects are not at all correctly included in the length-strength product that normally gives an adequate representation of the magnetic element. Also it only makes sense to treat each end individually. In problem b.3 the effect of a slanted pole is estimated. In this section, we give formulas for estimating the effect of the longitudinal field on entrance to or exit from a quadrupole. The word "estimate" is to be taken seriously here; if the estimate is that the end effect is important then some more accurate method such as numerical tracking through a grid of measured field values is called for. Unlike all other calculations, the symplecticity of this calculation is not assured.

The ideal central field in a quadrupole, according to Eq. (1.1) is

$$B_x = Gy, \quad B_y = Gx, \quad \text{where} \quad G = \left.\frac{\partial B_y}{\partial x}\right|_0. \tag{1.48}$$

An end field that increases from zero, say at the origin $z = 0$, to match this field at $z = b$ is

$$B_x = \frac{G}{b}yz, \quad B_y = \frac{G}{b}xz, \quad B_z = \frac{G}{b}xy. \tag{1.49}$$

It is easy to check that this field satisfies both $\nabla \cdot \vec{B} = 0$ and $\nabla \times \vec{B} = 0$, which makes it a valid free space magnetostatic field. Quadrupoles carefully designed for spectrometers have iron "mirror plates" that match to this field at $z = 0$ and the quadrupole ends are contoured as space hyperboloids to match the fields in the range from $z = 0$ to $z = b$. (A true mirror plate would consist of a ferromagnetic slab perpendicular to the beam

centerline. Such a slab would be unacceptable in practice as the beam would have to pass through the slab. As a result there is necessarily a hole bored through the slab. This invalidates Eq. (1.49) over a longitudinal range comparable with the bore radius.) Steffen gives curves illustrating the accuracy of this approximation. With mirror plate the agreement is quite good. Without mirror plate the disagreement is appreciable, but we will use Eq. (1.49) even knowing that IR quads in high energy accelerators rarely have mirror plates or correctly contoured ends. If the distance b, undetermined in the absence of mirror plate, is taken equal to the bore diameter of the quadrupole the end fields are roughly correct.

Let us assume that b is sufficiently short that the end can be represented as an impulse. Then, integrating from $z = 0$ to $z = b$ the field integrals corresponding to Eqs. (1.49) are

$$\begin{aligned}\tilde{B}_x =& \frac{B_x L_{\text{eff}}}{p_0/e} = \frac{Gb}{2p_0/e} y,\\ \tilde{B}_y =& \frac{B_y L_{\text{eff}}}{p_0/e} = \frac{Gb}{2p_0/e} x,\\ \tilde{B}_z =& \frac{B_z L_{\text{eff}}}{p_0/e} = \frac{G}{p_0/e} xy.\end{aligned} \tag{1.50}$$

Comparing with Eq. (1.40), we see that the x and y components match the ideal quadrupole components. We drop them on the assumption that they have already been accounted for by appropriate choice of quadrupole length. By Eq. (1.29) the remaining deflection is

$$\begin{aligned}\Delta f =& \frac{1}{L}\frac{GL}{p_0/e}\frac{v_y}{c} xy = \frac{1}{L}\tilde{b}_1 \frac{v_y}{c} xy,\\ \Delta g =& -\frac{1}{L}\frac{GL}{p_0/e}\frac{v_x}{c} xy = -\frac{1}{L}\tilde{b}_1 \frac{v_x}{c} xy.\end{aligned} \tag{1.51}$$

Here the factor L, the quadrupole length, has been introduced artificially so the strengths are expressed in terms of the fundamental quadrupole parameters, focal length $\tilde{b}_1 = 1/f$ and length L. Notice though that the only circumstance in which the value L can affect this calculation is if the quadrupole length is so short as to invalidate the approximations on which this formula is based.

The deflection caused by the longitudinal end field can be seen to be cubic in the small quantities v_x/c, x, and y; one might say "octupole order". As stated before the effects are likely to be negligible except at sensitive close-to-IR locations, where these factors as well as the β-functions are unusually large. Note though that the deflection depends on the

slope, which is possibly non-symplectic (or rather hard to approximate without destroying symplecticity) and would be likely to invalidate calculation of long term stability.

Why Momenta are Preferable to Velocities

This section can be skipped without loss of continuity.

While discussing quadrupole focusing we have seen how chromatic effects are included automatically. Now we digress to show that another choice of dependent variables, using (v_x, v_y) instead of (f, g), would complicate this as well as causing "mixing of orders", both feed-down and feed-up, even for dipoles and quadrupoles. Though having manageable effect on the linearized description this would seriously complicate the formulation, and impair the convergence in higher orders.

Using Eqs. (1.27) and (1.11), the momentum discontinuities due to a multipole plane, given by Eqs. (1.32), can also expressed as discontinuities of the velocity components;

$$\begin{aligned} \frac{v_x^{(+)}}{c} - \frac{v_x^{(-)}}{c} &= -\frac{v_0}{c}\frac{\gamma_0}{\gamma}\tilde{B}_y \quad \text{also} \quad = \quad -\frac{v/c}{1+\delta}\tilde{B}_y, \\ \frac{v_y^{(+)}}{c} - \frac{v_y^{(-)}}{c} &= \frac{v_0}{c}\frac{\gamma_0}{\gamma}\tilde{B}_x \quad \text{also} \quad = \quad \frac{v/c}{1+\delta}\tilde{B}_x. \end{aligned} \tag{1.52}$$

With v being constant in magnetic elements, the last expression is simple, and as we have seen, exact. Note that $\Delta v_x/v$ varies as $(1+\delta)^{-1}$. It is this form that justifies the omnipresence of the variable δ. For dipoles, this dependence of deflection on δ causes dispersion. For quadrupoles it causes chromaticity. This is the form used for particle tracking in *TEAPOT*.

When these equations are used to advance the map through a multipole plane, the extra factor $(v/c)/(1+\delta)$ has to be included, already making the analysis somewhat more complicated than was true for momenta. As it happens, at least in the relativistic regime, the factor v/c is quite accurately constant and the denominator factor is easily Taylor expanded into a polynomial in δ in the numerator. Still, neither of these approximations is particulary attractive, and the former is simply unacceptable for less than fully relativistic machines, such as typical proton "booster" accelerators in the few GeV range.

We proceed to try to improve the non-relativistic situation. For this it is preferable to use the middle expressions of Eqs. (1.52), since the only factor varying over the map is

$1/\gamma$. In preparation for replacing power series division by power series multiplication we define a function $\gamma_I(v_x, v_y, v_z) = 1/\gamma$, where the "I" is mnemonic for inverse. In order to proceed beyond linear order it is necessary to express γ_I as a Taylor series in the "small variables" (v_x, v_y) as well as a longitudinal deviation away from design. For transverse transverse velocity $v_T = \sqrt{v_x^2 + v_y^2}$, and setting $c = 1$,

$$\begin{aligned}\frac{\gamma_0}{\gamma} = \gamma_0\gamma_I &= \sqrt{\frac{1-\left(v_T^2+v_z^2\right)}{1-v_{z0}^2}} \\ &\stackrel{e.g.}{=} \left(1-\frac{1}{2}\left(v_T^2+v_z^2\right)-\frac{1}{8}\left(v_T^2+v_z^2\right)^2-\cdots\right)\left(1+\frac{1}{2}v_{z0}^2+\frac{3}{8}v_{z0}^4-\cdots\right) \\ &= 1-\frac{1}{2}v_T^2-\frac{1}{2}\left(v_z+v_{z0}\right)\left(v_z-v_{z0}\right)+\cdots.\end{aligned} \tag{1.53}$$

This manipulation has been performed in hopes of obtaining a rapidly convergent power series for substitution in Eqs. (1.52), where it could be multiplied by the (also rapidly convergent) series expansions for $\tilde{B}_y$ and $\tilde{B}_x$, yielding, after truncation, a useful, improved approximation. The manipulation has been only partly successful. The second term is quadratically small in the small quantity v_T, which is good. It is only linearly small in the small quantity $v_z - v_{z0}$, but that factor itself will turn out to be quadratically small. But the cancelation between a numerator term and a denominator factor does not occur in higher orders. Unfortunately this degrades the convergence of subsequent terms of Eq. (1.53). This behavior is characteristic of the balky nature of the relativistic factor γ.

Problem 1..15. Work out the next term in the expansion (1.53) and confirm the poor convergence of this series.

It can now be understood that the choice of momenta (f, g, h) was made precisely to avoid power series expansion of the relativistic factor γ. We have arrived at the conclusion (which may not surprise those knowledgable in relativistic mechanics) that the dynamical equations are simpler when expressed in terms of momentum rather than velocity. We can worry, having suppressed this factor in the dynamics, that it will pop back up in the kinematics. This fear will turn out to be unjustified—though velocity components determine the intersection of a trajectory with the next multipole plane, it is only their ratios that matter.

Chapter 2.
Fourier Analysis of 2D Nonlinear Motion

Analysis of a 4×4 Symplectic Matrix

In Section 2.5 explicit evaluation of the eigenvalues of a fully general 6×6 symplectic matrix has been exhibited. For this section, in the interest of reduced complexity, we drop back to 2D motion, either pure transverse (x, y) motion, or, replacing y by z, horizontal-longitudinal motion. A difference equation, satisfied by the linearized transfer matrix, and needed for Fourier analysis of the motion, is the most important result to be obtained.

The column vector of coordinates $\mathbf{x} = (x, f, y, g)^T$ represents small transverse deviations from the reference orbit. Linearized evolution of $\mathbf{x}$ from longitudinal coordinate s_0 to s is described by a transfer matrix $\mathbf{M}(s, s_0)$,

$$\mathbf{x}\,(s) = \mathbf{M}\,(s, s_0)\,\mathbf{x}\,(s_0)\,. \tag{2.1}$$

The fact that $\mathbf{M}$ is symplectic, critical to the derivation of the difference equation being sought, can be expressed using the matrix

$$\mathbf{S} = \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix}; \tag{2.2}$$

For $\mathbf{M}$ to be symplectic, its inverse must be equal to its "symplectic conjugate" $\overline{\mathbf{M}}$,

$$\mathbf{M}^{-1} = \overline{\mathbf{M}} = -\mathbf{S}\mathbf{M}^T\mathbf{S}. \tag{2.3}$$

Partitioning the 4×4 matrix $\mathbf{M}$ into 2×2 elements, it and its symplectic conjugate are

$$\mathbf{M} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix}, \quad \overline{\mathbf{M}} = \begin{pmatrix} \overline{\mathbf{A}} & \overline{\mathbf{C}} \\ \overline{\mathbf{B}} & \overline{\mathbf{D}} \end{pmatrix}. \tag{2.4}$$

A 2×2 matrix $\mathbf{A}$ and its symplectic conjugate are related by

$$\overline{\mathbf{A}} = \overline{\begin{pmatrix} a & b \\ c & d \end{pmatrix}} = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \mathbf{A}^{-1} \det|\mathbf{A}|, \tag{2.5}$$

provided the determinant $\det|\mathbf{A}|$ is non-vanishing.

Especially important for analysing the state of (x, y) coupling is a particular off-diagonal combination from Eq. (2.4), $\mathbf{E} = \mathbf{C} + \overline{\mathbf{B}}$ and its determinant $\mathcal{E} = \det|\mathbf{E}|$.

$$\mathbf{E} \equiv \mathbf{C} + \overline{\mathbf{B}} \equiv \begin{pmatrix} e & f \\ g & h \end{pmatrix} = \begin{pmatrix} c_{11} + b_{22} & c_{12} - b_{12} \\ c_{21} - b_{21} & c_{22} + b_{11} \end{pmatrix}, \quad \det|\mathbf{E}| = eh - fg \equiv \mathcal{E}. \tag{2.6}$$

For a stable lattice, eigenvalues λ_A and λ_D, of $\mathbf{M}$ (with their complex conjugate twins) satisfy the relations

$$\begin{aligned} \Lambda_A \equiv \lambda_A + 1/\lambda_A &= \exp(i\mu_A) + \exp(-i\mu_A) = 2\cos\mu_A \\ \Lambda_D \equiv \lambda_D + 1/\lambda_D &= \exp(i\mu_D) + \exp(-i\mu_D) = 2\cos\mu_D, \end{aligned} \tag{2.7}$$

where $\mu_A \equiv 2\pi\nu_A$ and $\mu_D \equiv 2\pi\nu_D$ are real angles. The quantities Λ_A and Λ_D, eigenvalues of $\mathbf{M} + \overline{\mathbf{M}}$, satisfy

$$\begin{aligned} \Lambda_A + \Lambda_D &= \operatorname{tr}\mathbf{A} + \operatorname{tr}\mathbf{D} \\ \Lambda_A\Lambda_D &= \operatorname{tr}\mathbf{A}\ \operatorname{tr}\mathbf{D} - \mathcal{E}. \end{aligned} \tag{2.8}$$

For motion at small amplitude the linearized transfer matrix description gives a thoroughly satisfactory description of the motion. In the presence of coupling the tunes ν_A and ν_D are only approximately equal to the ideal, (or nominal, or design) tunes ν_x and ν_y. But ν_A and ν_D are readily measurable, no matter how badly coupled the lattice is. For that reason, they can be regarded as known, or at least operationally measurable, quantities. In fact the greatest practical application of Fourier analysis of particle motion (as measured with beam position monitors), is for the operational measurement of these tunes. The formulation of this section can be used to support that procedure. But that is peripheral to our main purpose which is to prepare for subsequent harmonic analysis of nonlinear motion.

As partially seen already, the combination

$$\mathbf{M} + \overline{\mathbf{M}} = \mathbf{M} + \mathbf{M}^{-1} = \begin{pmatrix} \operatorname{tr}\mathbf{A} & 0 \\ 0 & \operatorname{tr}\mathbf{D} \end{pmatrix} + \begin{pmatrix} 0 & \overline{\mathbf{E}} \\ \mathbf{E} & 0 \end{pmatrix} \tag{2.9}$$

has simpler properties than $\mathbf{M}$. Using the fact that $\mathbf{M}^{-1}$ can be used to propagate backwards in time, this relation can be used to obtain four third-order, coupled difference equations that relate the coordinates on three successive turns (labeled $-$, 0, $+$):

$$\begin{aligned} x_+ - \operatorname{tr}\mathbf{A}\, x_0 + x_- &= hy_0 - fg_0 \\ f_+ - \operatorname{tr}\mathbf{A}\, f_0 + f_- &= -gy_0 + eg_0 \\ y_+ - \operatorname{tr}\mathbf{D}\, y_0 + y_- &= ex_0 + ff_0 \\ g_+ - \operatorname{tr}\mathbf{D}\, g_0 + g_- &= gx_0 + hf_0. \end{aligned} \tag{2.10}$$

(Please forgive the clash of symbols; the coefficients e, f, g, h, introduced for brevity in Eq. (2.6), and not to be confused with momentum coordinates, are retained only to maintain contact with decoupling algorithms of *TEAPOT*, and will be eliminated straight away.)

It is possible to uncouple these equations. Start by squaring Eq. (2.9), subtracting 2I, and using Eqs. (2.8);

$$\mathbf{M}^2 + \mathbf{M}^{-2} = \begin{pmatrix} \mathrm{tr}^2\,\mathbf{A} + (\mathcal{E}-2)\,\mathbf{I} & 0 \\ 0 & \mathrm{tr}^2\,\mathbf{D} + (\mathcal{E}-2)\,\mathbf{I} \end{pmatrix} + (\mathrm{tr}\,\mathbf{A} + \mathrm{tr}\,\mathbf{D}) \begin{pmatrix} 0 & \overline{\mathbf{E}} \\ \mathbf{E} & 0 \end{pmatrix}. \quad (2.11)$$

From Eqs. (2.9) and (2.11), form the combination that eliminates the off-diagonal blocks,

$$\mathbf{M}^2 + \mathbf{M}^{-2} - (\Lambda_A + \Lambda_D)\left(\mathbf{M} + \mathbf{M}^{-1}\right) + (2 + \Lambda_A\Lambda_D)\,\mathbf{I} = 0. \quad (2.12)$$

Using this equation to obtain a difference equation for the phase space coordinates on successive turns yields

$$\mathbf{x}_{++} + \mathbf{x}_{--} - (\Lambda_A + \Lambda_D)\,(\mathbf{x}_+ + \mathbf{x}_-) + (2 + \Lambda_A\Lambda_D)\,\mathbf{x}_0 = 0. \quad (2.13)$$

This is the equation we have been seeking. Before applying it to practical problems such as closed-orbit finding and feedback control, we note the simpler equations that hold in case there is no cross-plane coupling. In that case, Eqs. (2.7) and (2.8) reduce to

$$\Lambda_A = \mathrm{tr}\,\mathbf{A} = 2\cos\mu_x, \qquad \Lambda_D = \mathrm{tr}\,\mathbf{D} = 2\cos\mu_y; \quad (2.14)$$

the right hand sides of Eqs. (2.10) vanish; and the first equation, for example, becomes

$$x_+ - 2\cos\mu_x\, x_0 + x_- = 0. \quad (2.15)$$

It is left as an exercise to show that this equation and the corresponding y-equation are consistent with Eq. (2.13) when there is no coupling.

Finding the Closed Orbit

Koutchouk[3] has described a closed-orbit finding procedure, based on Eq. (2.15), which he ascribes to Verdier and Risselada.[4] That method, which assumes purely linear motion, will now be described and then generalized. Much the same description applies whether one is discussing operational procedures applied in the control room of an actual accelerator or simulation in a computer. In either case, finding the closed orbit is usually performed by starting with a guess and iteratively improving it.

In the derivations of the preceeding section it was implicitly assumed that transverse coordinates were measured relative to an unambiguous, previously-known, closed orbit. In order to treat the closed orbit as unknown we can make the replacement $\mathbf{x} \to \mathbf{x} - \mathbf{x}_{\rm CO}$ in Eq. (2.13) to give, after simplification,

$$\mathbf{x}_{++} + \mathbf{x}_{--} - (\Lambda_A + \Lambda_D)(\mathbf{x}_+ + \mathbf{x}_-) + (2 + \Lambda_A\Lambda_D)\,\mathbf{x}_0 = (2 - \Lambda_A)(2 - \Lambda_D)\,\mathbf{x}_{\rm CO}. \quad (2.16)$$

and correspondingly, make the replacement $x \to x - x_{\rm CO}$ in Eq. (2.15)

$$x_+ - 2\cos\mu_x\, x_0 + x_- = 2(1 - \cos\mu_x)\, x_{\rm CO}. \quad (2.17)$$

The parameters Λ_A and Λ_D in Eq. (2.16) are simple functions of operationally measurable tunes, as is $\cos\mu_x$ in Eq. (2.17). In the control room of an actual accelerator, if circulating beam can be obtained, they can be measured by spectrum analysis of beam position monitor data. Similarly, in a computer simulation, if multiple turns survive, the tunes can be obtained by FFT analysis. Unfortunately the "if's" in the two previous sentences are often not satisfied. For that reason we need a robust procedure for extracting tunes that makes minimal operational demands. Following Verdier and Risselada we obtain another equation like Eq. (2.17) by incrementing the indices by one, (which leaves the right hand side unchanged), and then eliminate $x_{\rm CO}$ from the two equations,

$$x_3 - 2\cos\mu_x\, x_2 + x_1 = x_2 - 2\cos\mu_x\, x_1 + x_0. \quad (2.18)$$

Solving for μ_x yields

$$\cos\mu_x = \frac{x_3 - x_2 + x_1 - x_0}{2(x_2 - x_1)}. \quad (2.19)$$

From this equation, starting with x_0, if the particle (or beam) can survive three full turns, and the displacement measured on each passage through the origin, the tune can

be obtained. Once the tune is known, the closed orbit is obtained from Eq. (2.17),

$$x_{\rm co} = \frac{x_2 - 2\cos\mu_x\, x_1 + x_0}{2\,(1-\cos\mu_x)}. \tag{2.20}$$

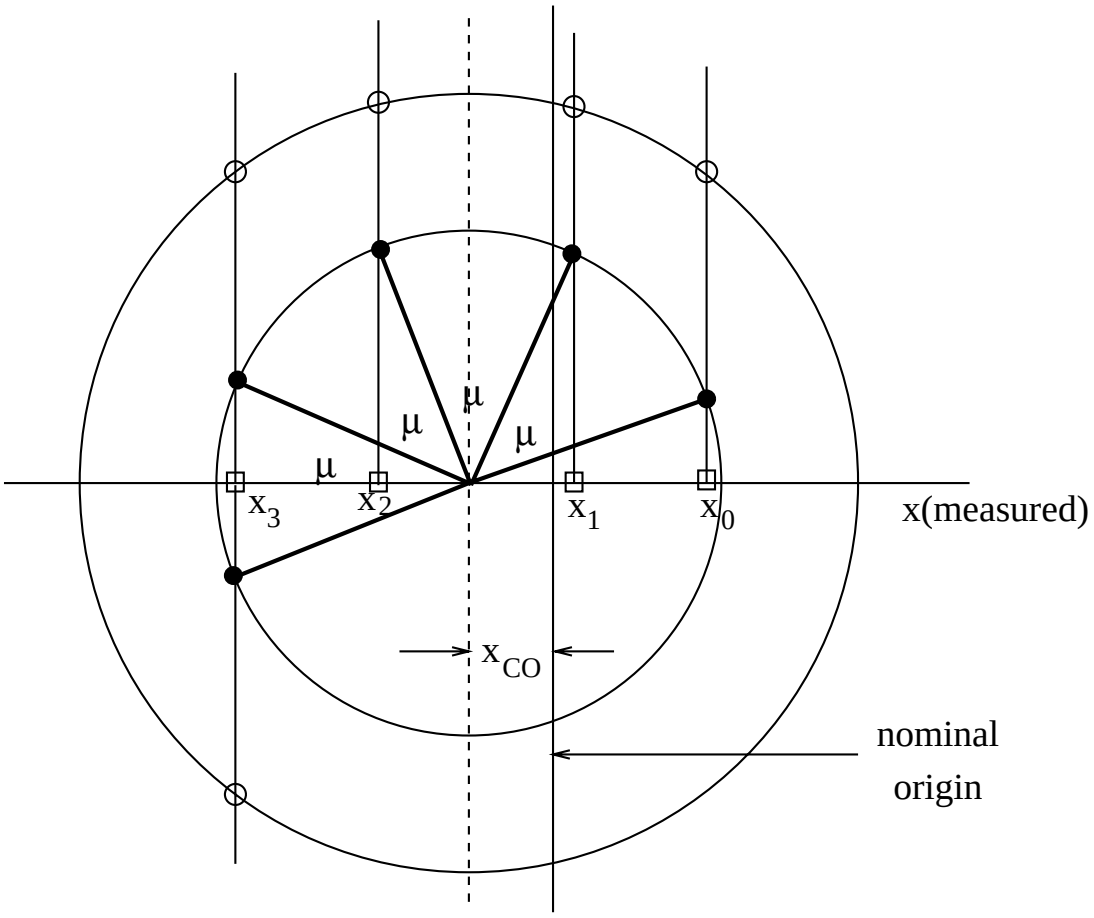

Figure 2.1: Geometric construction indicating how tune and closed orbit can be found from measuring the transverse displacement for several successive turns. Open squares are measured. Closed circles lie on and define the correct phase space circle.

This prescription can be foiled by measurement errors, by the presence of coupling, or by the presence in the lattice of nonlinear elements that violate the conditions used in deriving the difference equation. There is nothing we can do about measurement errors except complain about the instrumentation. Before proceeding to discuss what can be done about coupling we consider nonlinearity.

Because of nonlinearity, Eqs. (2.19) and (2.20) will be not quite satisfied and the closed orbit not quite found. This performance is typical of almost all operational accelerator

procedures. The universal attempted fix is to proceed by iteration. In this case, having found a tentative value for $x_{\rm co}$ we launch another particle from that point. Assuming a sensible prescription for picking the initial slope is available, the ability of nonlinearity to foil this approach will rapidly decrease with each succeeding iteration, as the orbit will stay in progressively reduced, and hence more linear, regions. It can certainly happen however, that the first iteration fails due to nonlinearity. Either the particle is lost completely, (a possibility the derivation excluded,) or the errors make the "improved" closed orbit worse than the tentative starting value. In either case an alternate approach must be found. In practice the alternate approach is usually trial-and-error or "knob-twiddling", which normally succeed eventually. From that point rapid convergence employing Eqs. (2.19) and (2.20) is typical.

We now wish to generalize this prescription in two ways in order to make its convergence more robust. First, as mentioned in the previous paragraph, a procedure for improving the starting slope is required. The derivation of the previous section shows that the slope variables satisfy the same difference equations as the displacements. As a result we obtain

$$f_{\rm co} = \frac{f_2 - 2\cos\mu_x\, f_1 + f_0}{2\,(1-\cos\mu_x)}. \tag{2.21}$$

Assuming slope values are available (which is certainly true in a computer simulation, but would only be true by bringing in another beam position monitor in the laboratory,) this equation can be used to improve the tentative closed orbit initial conditions.

In practice, the presence of coupling seriously compromises the effectiveness of this closed orbit finding procedure. Because coupling is a "linear effect", its fractional importance does not reduce with succeeding stages of iteration. For too great coupling the above approach simply does not converge. For that reason we contemplate using the more general Eq. (2.16) to obtain simultaneous convergence in both planes. As in the uncoupled case, there are two stages, the first to find the tune(s), the second to improve the initial conditions, (the tentative closed orbit).

Four alternative approaches to finding the tunes suggest themselves. The first two are applicable only if the coupling is weak, which is almost always the case since its presence is unintentional, and hence its effect on the tunes is likely to be negligible. (Because tune shifts depend quadratically on skew quadrupole strengths.) In that case the "design" tunes

can be used, or the determination of μ_x using Eq. (2.19), and a corresponding determination of μ_y may be adequate. A third, more robust approach is to obtain equations for Λ_A and Λ_D in a manner analogous to the derivation of Eq. (2.19). We write the vector equation

$$\begin{aligned}\mathbf{x}_4+\mathbf{x}_0 - (\Lambda_A+\Lambda_D)(\mathbf{x}_3+\mathbf{x}_1) + (2+\Lambda_A\Lambda_D)\,\mathbf{x}_2 = \\ \mathbf{x}_5+\mathbf{x}_1 - (\Lambda_A+\Lambda_D)(\mathbf{x}_4+\mathbf{x}_2) + (2+\Lambda_A\Lambda_D)\,\mathbf{x}_3.\end{aligned} \tag{2.22}$$

Collecting terms yields

$$\begin{pmatrix} x_4-x_3+x_2-x_1 & -x_3+x_2 \\ y_4-y_3+y_2-y_1 & -y_3+y_2 \end{pmatrix} \begin{pmatrix} \Lambda_A+\Lambda_D \\ \Lambda_A\Lambda_D \end{pmatrix} = \begin{pmatrix} x_5-x_4+2x_3-2x_2+x_1-x_0 \\ y_5-y_4+2y_3-2y_2+y_1-y_0 \end{pmatrix}. \tag{2.23}$$

These equations can be solved for Λ_A and Λ_D if data from five consecutive full turns is available.

A fourth, somewhat noise-immune approach, applicable when multiple turns can be obtained, has been applied to the LEP lattice.[5] Define the expectation value $< f >$ of N samples f_i by $\sum_i^N f_i/N$. Multiplying the x and y components of (2.13) by x_0 and y_0 respectively, taking expectation values, and rearranging to express as equations for Λ_A and Λ_D yields

$$\begin{aligned}\begin{pmatrix} <(x_+ + x_-)\,x_0> & -<x_0^2> \\ <(y_+ + y_-)\,y_0> & -<y_0^2> \end{pmatrix} \begin{pmatrix} \Lambda_A+\Lambda_D \\ \Lambda_A\,\Lambda_D \end{pmatrix} \\ = \begin{pmatrix} <(x_{++}+x_{--})\,x_0> +2<x_0^2> \\ <(y_{++}+y_{--})\,y_0> +2<y_0^2> \end{pmatrix}.\end{aligned} \tag{2.24}$$

When this equation was applied for 512 turn data, and solved for Λ_A and Λ_D, accuracies of approximately ± 0.003 were obtained for the tunes ν_x and ν_y.

Once the tunes are known, the coefficients in Eq. (2.16) can be evaluated, and improved values for all four closed orbit coordinates can be obtained;

$$\mathbf{x}_{\rm CO} = \frac{\mathbf{x}_{++}+\mathbf{x}_{--} - (\Lambda_A+\Lambda_D)(\mathbf{x}_+ + \mathbf{x}_-) + (2+\Lambda_A\Lambda_D)\,\mathbf{x}_0}{(2-\Lambda_A)(2-\Lambda_D)}. \tag{2.25}$$

This formula has not yet been incorporated into an operational program, but its effectiveness is theoretically assured.

Use of Spectral Analysis to Identify Sources of Nonlinearity

Correlating Spectral Lines With Multipole Terms.

In this section we develop a semi-quantitative method for diagnosing likely sources of nonlinearity by looking for peaks in the spectra derived from multiturn beam position monitor output. These spectra can be obtained by spectrum analyser hardware or by digital Fourier transformation.

For sufficiently small amplitudes all nonlinear terms become negligible and the motion is described approximately by the pure (uncoupled by preference) betatron motions $x_t = a_x \cos(2\pi\nu_x t)$ and $y_t = a_y \cos(2\pi\nu_y t)$ for $t = 0, 1, \ldots$. These "fundamental oscillations" can be regarded as the "zero'th" approximation to the motion. To obtain an improved approximation the multipole deflection terms can be treated, perturbatively and artificially, as being due to "external drive" deflections. These deflections are nonlinear functions of the fundamental oscillations amplitudes, but since they are periodic with the same period, they can be expessed as "nonlinear harmonics" of the fundamental oscillations. The purpose of this Fourier expansion is to again "linearize" the problem so the perturbing higher harmonics can be obtained as the well known driven response of a linear system. This procedure can be iterated, but the possible harmonics proliferate badly. Higher iteration is however justified in the (quite common) case of harmonics absent only in lowest order.

Observing which peaks are present, and with what strength, and then correlating with Fourier expansions of particular multipole fields can give clues as to which fields are causing the motion to be nonlinear. Conversely the importance of nonlinearities known (from magnetic measurements) can be assessed. For interpreting spectra in this way it is necessary to write Fourier expansions of the multipole expressions for deflections $\Delta x' \simeq \Delta f$ and $\Delta y' \simeq \Delta g$ appearing in Table 1.1. The following formulas are needed for those

expansions:

$$
\begin{aligned}
\left|{}^{m_x}_{m_y}\right. &\equiv \cos\left(\left(m_x\nu_x + m_y\nu_y\right)2\pi t\right), \qquad \left|{}^{m_x}_{\pm m_y}\right. \equiv \left|{}^{m_x}_{m_y}\right. + \left|{}^{m_x}_{-m_y}\right., \qquad 1 = \left|{}^{0}_{0}\right. \\
x &= a_x \cos\left(2\pi\nu_x t\right) = a_x \left|{}^{1}_{0}\right. \\
y &= a_y \cos\left(2\pi\nu_y t\right) = a_y \left|{}^{0}_{1}\right. \\
x^2 - y^2 &= \frac{a_x^2}{2}\left|{}^{2}_{0}\right. - \frac{a_y^2}{2}\left|{}^{0}_{2}\right. + \frac{a_x^2 - a_y^2}{2}\left|{}^{0}_{0}\right. \\
2xy &= a_x a_y \left|{}^{1}_{\pm 1}\right. \\
x^3 - 3xy^2 &= \frac{a_x^3}{4}\left|{}^{3}_{0}\right. - \frac{3a_x a_y^2}{4}\left|{}^{1}_{\pm 2}\right. + \frac{3a_x^3 - 6a_x a_y^2}{4}\left|{}^{1}_{0}\right. \\
3x^2y - y^3 &= -\frac{a_y^3}{4}\left|{}^{0}_{3}\right. + \frac{3a_x^2 a_y}{4}\left|{}^{2}_{\pm 1}\right. - \frac{3a_y^3 - 6a_x^2 a_y}{4}\left|{}^{0}_{1}\right. \\
x^4 - 6x^2y^2 + y^4 &= a_x^4\left|{}^{4}_{0}\right. - \frac{6a_x^2 a_y^2}{8}\left|{}^{2}_{\pm 2}\right. + \frac{a_y^4}{4}\left|{}^{0}_{4}\right. \\
&\quad + \frac{4a_x^4 - 12a_x^2 a_y^2}{8}\left|{}^{2}_{0}\right. + \frac{4a_y^4 - 12a_x^2 a_y^2}{8}\left|{}^{0}_{2}\right. + \frac{3a_x^4 - 12a_x^2 a_y^2 + 3a_y^4}{8}\left|{}^{0}_{0}\right. \\
4x^3y - 4xy^3 &= \frac{4a_x^3 a_y}{8}\left|{}^{3}_{\pm 1}\right. - \frac{4a_x a_y^3}{8}\left|{}^{1}_{\pm 3}\right. + \frac{12a_x^3 a_y - 12a_x a_y^3}{8}\left|{}^{1}_{\pm 1}\right.
\end{aligned}
\tag{2.26}
$$

The notation $\left|{}^{m_x}_{m_y}\right.$ with x-harmonic number m_x on top and y-harmonic number m_y on the bottom is intended to help in correlating with spectra having spectral lines labeled the same way.

Since the factors a_x and a_y are presumeably, in some sense, "small", the dominant lines tend to be those having minimal powers of these factors.

m_x	0	1	0	2	1	0	3	2	1	0	4	3	2	1	0
m_y	0	0	1	0	±1	2	0	±1	±2	3	0	±1	±2	±3	4
b_0	1														
a_0															
b_1		a_x													
a_1			a_y												
b_2	a_x^2, a_y^2			a_x^2		a_y^2									
a_2					$a_x a_y$										
b_3		$a_x^3, a_x a_y^2$					a_x^3		$a_x a_y^2$						
a_3			$a_x^2 a_y, a_y^3$					$a_x^2 a_y$		a_y^3					
b_4	$a_x^4, a_x^2 a_y^2, a_y^4$			$a_x^4, a_x^2 a_y^2$		$a_x^2 a_y^2, a_y^4$					a_x^4		a_x^2, a_y^2		a_y^4
a_4					$a_x^3 a_y, a_x a_y^3$							$a_x^3 a_y$		$a_x a_y^3$	

Table 2.1: Spectral lines in X-spectrum (horizontal) caused by particular multipoles. a_x and a_y are "fundamental" amplitudes. There are also numerical factors, of order one, not shown.

m_x	0	1	0	2	1	0	3	2	1	0	4	3	2	1	0
m_y	0	0	1	0	±1	2	0	±1	±2	3	0	±1	±2	±3	4
b_0															
a_0	1														
b_1			a_y												
a_1		a_x													
b_2					$a_x a_y$										
a_2	a_x^2, a_y^2			a_x^2		a_y^2									
b_3			$a_x^2 a_y, a_y^3$					$a_x^2 a_y$		a_y^3					
a_3		$a_x^3, a_x a_y^2$					a_x^3		$a_x a_y^2$						
b_4					$a_x^3 a_y, a_x a_y^3$							$a_x^3 a_y$		$a_x a_y^3$	
a_4	$a_x^4, a_x^2 a_y^2, a_y^4$			$a_x^4, a_x^2 a_y^2$		$a_x^2 a_y^2, a_y^4$					a_x^4		a_x^2, a_y^2		a_y^4

Table 2.2: Spectral lines in Y-spectrum (vertical) caused by particular multipoles. a_x and a_y are "fundamental" amplitudes. There are also numerical factors, of order one, not shown.

Problem 2..16. Show that the effect of a horizontal closed orbit displacement $\underline{\Delta x}$ is to produce "feed-down" such that the presence of multipole coefficients b_n and a_n leads to multipoles

$$b_{n-1}^{(\underline{\Delta x})} = -n b_n \underline{\Delta x}, \quad a_{n-1}^{(\underline{\Delta x})} = -n a_n \underline{\Delta x}. \tag{2.27}$$

Problem 2..17. Show that the effect of a vertical closed orbit displacement $\underline{\Delta y}$ is to produce feed-down such that the presence of multipole coefficients b_n and a_n leads to

multipoles

$$b_{n-1}^{(\underline{\Delta y})} = n a_n \underline{\Delta y}, \quad a_{n-1}^{(\underline{\Delta x})} = -n b_n \underline{\Delta y}. \tag{2.28}$$

Problem 2..18. The Courant-Snyder invariant ϵ of a particle executing one dimensional betatron oscillations is given by $\gamma_x x^2 + 2\alpha_x x x' + \beta_x x'^2$. With proper axis-scaling, if the motion is linear, the point in phase space with coordinates (x, x') lies on a circle and rotates at uniform rate. The effect of a deflection $x' \to x' + \Delta x'$ will sometimes be to increase ϵ and sometimes to decrease it. Show however that on the average there is a net increase given by

$$< \Delta\epsilon >= \beta_x \left(\Delta x'\right)^2. \tag{2.29}$$

Problem 2..19. Consider the one dimensional motion of a particle with amplitude a_x through a bending element that bends the central trajectory through angle $\Delta\theta$. The field nonuniformity of the magnet is described by a multipole coefficient b_n. Show that he deflection suffered is $b_n a_x^n \Delta\theta$ times an oscillatory factor in the range from -1 and 1. Continuing to drop a numerical factor of order 1, show that the average fractional increase in the Courant-Snyder invariant in passing through the magnet is

$$\frac{< \Delta\epsilon >}{\epsilon} \sim \left(b_n a_x^{n-1} \beta_x \Delta\theta\right)^2 \sim \left(b_n \Delta\theta\right)^2 \epsilon_x^{\frac{n-1}{2}} \beta_x^{\frac{n+1}{2}} N_x^{n-1}. \tag{2.30}$$

where a_x is quoted as $N\sigma_x$ where ϵ_x is the horizontal beam emittance and $\sigma_x = \sqrt{\epsilon_x \beta_x}$ is the r.m.s. horizontal beam size. This formula is not valid for $n = 1$. Why not?

The convergence (*i.e.* the extent to which succeeding terms become less important) of the multipole series as a formula for magnetic field at displacement a_x can be assessed by the numerical value of ratios $\frac{b^{n+1} a_x^{n+1}}{b^n a_x^n} = \frac{b^{n+1}}{b^n} a_x$. But to estimate the absolute influence on accelerator performance of a particular multipole an estimator like that calculated in Problem 2..19 is needed. In practice, as mentioned before, there is a strong tendency for lower powers of a_n and b_n to dominate.

Examples and Problems

Exercises Concerning Symplectic Requirements and FFT Analysis

The following "problems" were developed using an HP-48 hand calculator; this makes an HP-48 calculator especially appropriate for performing them. But no use is made of the graphics capabilities, and it is likely that any programmable calculator can be used to automate many of the operations. Some of the calculations assume that your calculator has matrix handling capabilities, such as multiplying, transposing, and inverting. If your calculator lacks these capabilities (the normal situation) it will be necessary to skip some parts. To allow for that, typical output data sets are given for possible use in subsequent parts. The purpose of this series of problems/exercises is to make concrete the symplectic turn-by-turn mapping of particle coordinates, the difference equation the coordinates satisfy, and their tune/frequency analysis using Discrete Fourier Transform's.

Depending on the calculator that is available, this "making concrete" will take different forms. With the HP-48, the codes given can simply be keyed in and various cases tried. Variable names begin with upper or lower case alphabetic characters and can contain decimal integers. By (my) convention program names begin with a $ symbol. Programs themselves are contained within << and >>. Hence

```
$PLUS: << A 3 + 'A' STO >>
```

is a program that loads `A` then `3`, adds them, and stores the result in `A`'s memory location `'A'`.

With a programmable computer lacking matrix manipulations, only some of the codes can be generated. In any case, it is not intended that all problems be completed. Rather they represent a somewhat coherent train of thought and the purpose is to understand the sequence, repeat some calculations, and spot check others. This is intended to enhance familiarity with, and confidence in, the methods. Some of the examples include analytic work and/or reasonably short calculations; they are identified as "problems" rather than "examples".

FFT.1 Generate a random 4×4 transfer matrix.

```
$RAN: << 1 16 FOR I RAND 2 * 1 - NEXT { 4 4 } ->ARRY 'R44'
STO >>
```

$$R44 = \begin{pmatrix} -0.9215 & -0.8827 & 0.5889 & 0.8173 \\ 0.0614 & 0.5857 & -0.3813 & 0.7203 \\ 0.1137 & 0.1917 & -0.6056 & 0.9430 \\ 0.4199 & -0.3086 & 0.7853 & -0.0874 \end{pmatrix}$$

Problem 2..20. Generate a simple, obviously symplectic, unperturbed 4×4 transfer matrix using, for example, tunes `QX=0.1` and `QY=0.2`.

```
$M0: <<
     QX PI * 2 * DUP COS     SWAP SIN    0 0
     QX PI * 2 * DUP SIN NEG SWAP COS    0 0
     0 0                                 QY PI * 2 * DUP COS     SWAP
SIN
     0 0                                 QY PI * 2 * DUP SIN NEG SWAP
COS
     { 4 4 }
     ->ARRY 'M0' STO
     >>
```

$$M0 = \begin{pmatrix} 0.8090 & 0.5878 & 0 & 0 \\ -0.5878 & 0.8090 & 0 & 0 \\ 0 & 0 & 0.3090 & 0.9511 \\ 0 & 0 & -0.9511 & 0.3090 \end{pmatrix}$$

Problem 2..21. Algorithm for converting an almost symplectic matrix into a symplectic matrix. Define

$$S = \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

The "symplectic conjugate of matrix M is defined by

$$\overline{M} = -SM^T S,$$

where M^T is the transpose of M. One can write a calculator routine to perform this operation

```
$BAR: << TRN S * S SWAP * NEG >>
```

A matrix M_s is symplectic if and only if

$$\overline{M}_s = M_s^{-1}.$$

Suppose that M is "almost" symplectic. Define a new matrix, close to M by

$$M_s = M + \frac{M - M\overline{M}M}{2}.$$

Neglecting terms quadratic in $M - M_s$, show that M_s is approximately symplectic.

One can write a calculator routine, using 4×4 unit matrix `I44`

```
$SYM: << DUP DUP $BAR * 2 / NEG I44 1.5 * + SWAP * >>
```

Problem 2..22. Generate a randomly perturbed, coupled 4×4 transfer matrix. Normalize it to have unit determinant and then make it symplectic. Find its characteristic polynomial, its eigenvalues λ and $\Lambda_A = \lambda_A + \lambda_A^{-1} = 2\cos(\mu_A)$ and $\Lambda_D = \lambda_A + \lambda_D^{-1} = 2\cos(\mu_D)$.

```
$M: <<
    M0 R44 FAC * + DUP 'M' STO
    DET DUP 'D' STO SQ 0.125 ^ M SWAP /
    $SYM $SYM $SYM $SYM $SYM 'M' STO
    >>
```

$$M = \begin{pmatrix} 0.7499 & 0.5207 & 0.0321 & 0.0675 \\ -0.6098 & 0.9076 & -0.0071 & 0.0439 \\ -0.0035 & -0.0250 & 0.2493 & 1.0499 \\ 0.0696 & -0.0438 & -0.8788 & 0.3028 \end{pmatrix}$$

$$\lambda^4 - 2.2096\lambda^3 + 2.9085\lambda^2 - 2.2096\lambda + 1 = 0.$$

$$\lambda_A = 0.8317 \pm 0.5552, \quad \lambda_D = 0.2731 \pm 0.9620.$$

$$\Lambda_A = 1.6635, \quad Q_A = 0.0937, \quad \Lambda_D = 0.5461, \quad Q_D = 0.2060.$$

FFT.5 For a starting displacement X_0 such as $x_0 = 1, f_0 = 0, y_0 = 1, g_0 = 0$, *i.e.* in the calculator `[1 0 1 0]`, iterate the matrix multiplication $X_{i+1} = MX_i$ `NTR=16` times. This generates simulated data at a single "BPM" that will subsequently by used to test the difference equation and for FFT analysis.

```
$x4: <<
     { } SWAP
     1 NTR START
     1 DUPN OBJ-> 1 GET ->LIST ROT SWAP + SWAP
(    1 DUPN EL GET ROT SWAP + SWAP                    )
(    1 DUPN 1 DUPN 1 GET SWAP 3 GET R->C ROT SWAP +
SWAP  )
     M SWAP *
     NEXT
     SWAP 'TRK' STO
     >>
```

Substituting the first line in parenthesis for the preceding line generates output for `EL = 1,2,3, or 4`. Using the following line generates output x_i+iy_i, useful for FFT below.

tracking output =

n	x_n	f_n	y_n	g_n
0	1.0000	0.0000	1.0000	0.0000
1	0.7821	−0.6169	0.2458	−0.8092
2	0.2186	−1.0741	−0.7756	−0.3796
3	−0.4459	−1.1192	−0.5659	0.6289
4	−0.8929	−0.7122	0.5487	0.7057
5	−0.9751	−0.0748	0.8986	−0.2995
6	−0.7616	0.5071	−0.0851	−0.9449
7	−0.3736	0.8838	−1.0233	−0.2865
8	0.1277	1.0246	−0.5767	0.7478
9	0.6612	0.8890	0.6152	0.6973
10	1.0256	0.4298	0.8609	−0.3225
11	0.9988	−0.2556	−0.1383	−0.8017
12	0.5574	−0.8752	−0.8732	−0.0406
13	−0.0685	−1.1298	−0.2404	0.8322
14	−0.5912	−0.9453	0.8423	0.5079
15	−0.8742	−0.4812	0.7690	−0.5861

Problem 2..23. Each of the components, and in particular x and y, satisfy fifth order difference equations

$$x_{i+4} + x_i - (\Lambda_A + \Lambda_D)(x_{i+3} + x_{i+1}) + (2 + \Lambda_A\Lambda_D)\, x_2 = 0$$

$$y_{i+4} + y_i - (\Lambda_A + \Lambda_D)(y_{i+3} + y_{i+1}) + (2 + \Lambda_A\Lambda_D)\, y_2 = 0$$

In theory the turn-by-turn data of the previous problem satisfy these equations, which are normally considered to be difference equations for unknowns x_i, y_i with known parameters, Λ_A and Λ_D but can alternatively be regarded as two nonlinear algebraic equations for unknowns Λ_A and Λ_D. Taking x_i, y_i pairs for any five successive turns as known, solve these two equations for Λ_A and Λ_D, and from them calculate tunes Q_A and Q_D. They should agree with values given above.

Problem 2..24. Using the fifth order difference equation applicable to symplectic propagation, write an expression to predict x_{i+4} from x_i, x_{i+1}, x_{i+2}, and x_{i+3}. Check this equation numerically using the tracking data of Problem 2..22. Check that f_i, y_i, and g_i satisfy the same equation.

```
x4 = '(LAMA+LAMD)*x3 - (2 + LAMA*LAMD)*x2 + (LAMA+LAMD)*x1 - x0'
        $x4: <<
             4 DUPN
             LAMA LAMD + * SWAP
             LAMA LAMD * 2 + * - SWAP
             LAMA LAMD + * + SWAP
             -
             >>
```

This difference equation code propagates any particular element just as well as the matrix code propagates the vector of 4 elements.

FFT.8 The following equations implement the $N = 8$ DFT (discrete Fourier transform) for sequences of 8 complex numbers $z_i \to Z_j, i, j = 0, 1, 2, \ldots, 7$. They are FFT's in the sense that they have been manipulated to require only 8 multiplications and 26 divisions.

$$
\begin{aligned}
&u = 2\pi/8,\\
&t_1 = z_0 + z_4, \quad t_2 = z_2 + z_6, \quad t_3 = z_1 + z_5,\\
&t_4 = z_1 - z_5, \quad t_5 = z_3 + z_7, \quad t_6 = z_3 - z_7,\\
&t_7 = t_1 + t_2, \quad t_8 = t_3 + t_5\\
&m_0 = t_7 + t_8, \quad m_1 = t_7 - t_8\\
&m_2 = t_1 - t_2, \quad m_3 = z_0 - z_4\\
&m_4 = \cos u\,(t_4 - t_6), \quad m_5 = j\,(z_0 - z_4)\\
&m_6 = j\,(z_6 - z_2), \quad m_7 = -j \sin u\,(t_4 + t_6)\\
&s_1 = m_3 + m_4, \quad s_2 = m_3 - m_4, \quad s_3 = m_6 + m_7, \quad s_4 = m_6 - m_7\\
&Z_0 = m_0, \quad Z_1 = s_1 + s_3, \quad Z_2 = m_2 + m_5, \quad Z_3 = s_2 - s_4\\
&Z_4 = m_1, \quad Z_5 = s_2 + s_4, \quad Z_6 = m_2 + m_5, \quad Z_7 = s_1 - s_3.
\end{aligned}
$$

But it is quite a bit easier to code the fundamental DFT formulas:

$$Z(k) = \frac{1}{N} \sum_{n=0}^{N-1} z(n) \exp\left(-\frac{j2\pi kn}{N}\right),$$

which takes 64 multiplications and additions. Starting parameters are calculated by

```
            $U: << 2 pi * N / DUP 'U' STO i * NEG EXP 'WM1' STO
    >>
```

A column of coefficients is calculated by

```
            $WC: << -> a << 1 0 R->C 1 N 1 - START DUP a * NEXT
    >> >>
```

The $N \times N$ array of coefficients is calculated (once only) by

```
$WE: <<
     $U 1 0 R->C 'W1I' STO
     1 N START
     W1I $WC WM1 * 'W1I' STORE* NEXT
     { N N } ->ARRAY 'WE' STO
     >>
```

The program $DFT accepts the array to be transformed from the stack lowest level and replaces it by the transformed array.

```
$DFT: << WE SWAP * N / >>
```

The inverse transform, the IDFT, is given by

$$z(n) = \sum_{k=0}^{N-1} Z(k) \exp\left(\frac{j2\pi kn}{N}\right).$$

Essentially the same program can be used for either DFT or IDFT because, with $*$ indicating complex conjugation,

$$x = N\left(DFT\{X^*\}\right)^*.$$

```
$IDFT: << CONJ WE SWAP * CONJ >>
```

Problem 2..25. The DFT of the 16 turns of data, $z_n = x_n + jy_n$, given in FFT.5 is

$$Z(k) = DFT\{z_n\} = \begin{array}{ccc} k & Re & Im \\ 0 & 0.0243 & 0.0939 \\ 1 & -0.0435 & 0.3666 \\ 2 & 0.0689 & -0.2762 \\ 3 & -0.2394 & 0.1148 \\ 4 & 0.2112 & -0.1149 \\ 5 & 0.1210 & -0.0393 \\ 6 & 0.0922 & -0.0093 \\ 7 & 0.0750 & 0.0096 \\ 8 & 0.0612 & 0.0238 \\ 9 & 0.0474 & 0.0354 \\ 10 & 0.0300 & 0.0451 \\ 11 & 0.0003 & 0.0501 \\ 12 & -0.0986 & 0.0219 \\ 13 & 0.4487 & 0.3939 \\ 14 & 0.1747 & 0.4274 \\ 15 & 0.0266 & -0.1429 \end{array}$$

In this form the Fourier transforms of x and y are "mixed together". Separate them as follows: suppose that $X(k)$ and $Y(k)$ (both complex) are the transforms of x and y, (both real). That is

$$\begin{aligned} z(n) &\leftrightarrow Z(k), \\ x(n) &\leftrightarrow X(k), \\ y(n) &\leftrightarrow Y(k). \end{aligned}$$

Show that

$$X(k) = \frac{1}{2}\left[Z^*(N-k) + Z(n)\right],$$
$$Y(k) = \frac{i}{2}\left[Z^*(N-k) - Z(n)\right].$$

Reversing the element order is the main complication in implementing this separation:

```
$SEP: <<
      DUP 'ZZ' STO OBJ-> DROP
      1 N FOR I I I + 1 -
      PICK CONJ NEXT
      { N 1 } ->ARRY 'ZZT' STO CLEAR
      ZZ ZZT + 2 / 'XX' STO
      ZZ ZZT - i * 2 / 'YY' STO
      >>
```

Commonly only absolute values are required in the tune domain.

```
$ABS: << 1 N FOR I I GETI SWAP DROP ABS I SWAP PUT NEXT
>>
```

$(|X(k)|, |Y(k)|) =$

k	$\lvert X(k)\rvert$	$\lvert Y(k)\rvert$
0	0.1211	0.0245
1	0.0723	0.4117
2	0.4234	0.1988
3	0.1752	0.0981
4	0.1341	0.1103
5	0.0865	0.0456
6	0.0733	0.0260
7	0.0685	0.0181
8	0.0685	0.0181
9	0.0733	0.0260
10	0.0865	0.0456
11	0.1341	0.1103
12	0.1752	0.0981
13	0.4234	0.1988
14	0.0723	0.4117
15	0.1211	0.0245

Here are some comments about these spectra. They are symmetric about the center; that is about tune=0.5. (Tune/bin = 1/N.) With only 16 turns the tune lines are not very sharp, and the tunes are not very well determined—not nearly as well as in Problem 2..23. (Also x and y have somehow been reversed.)

Eigenanalysis of the linear map

Before continuing with the analysis of $\mathbf{X} + \widehat{\mathbf{\Delta X}}$ we perform eigenanalysis of just the linear part. The linear transfer matrix $\mathbf{X}$, and its symplectic conjugate $\overline{\mathbf{X}}$ can be written in partitioned form as

$$\mathbf{X} = \begin{pmatrix} A & B & E \\ C & D & F \\ G & H & J \end{pmatrix}, \quad \overline{\mathbf{X}} = \begin{pmatrix} \overline{A} & \overline{C} & \overline{G} \\ \overline{B} & \overline{D} & \overline{H} \\ \overline{E} & \overline{F} & \overline{J} \end{pmatrix}. \tag{2.31}$$

We will assume that interplane coupling is sufficiently weak that the matrices A, D, and J, are "not too far from" the uncoupled 2×2 "design" transfer matrices corresponding to pure x, y, and z motion respectively. However, the purpose of this assumption is not to justify a perturbative expansion, since the formulas will be exact. Rather it is to resolve ambiguities in identifying the roots of the equations by considerations of continuity. Because $\mathbf{X}$ is necessarily symplectic, its symplectic conjugate, defined using block-diagonal matrix S, each of whose diagonal blocks is $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$,

$$\overline{\mathbf{X}} = -S\mathbf{X}^T S, \tag{2.32}$$

is also its inverse

$$\overline{\mathbf{X}} = \mathbf{X}^{-1}. \tag{2.33}$$

We define an auxiliary matrix,

$$\mathbf{\Xi} = \mathbf{X} + \overline{\mathbf{X}} = \mathbf{X} + \mathbf{X}^{-1}, \tag{2.34}$$

having much simpler properties than X. In particular, if (as it does) $\mathbf{X}$ has eigenvalue $\lambda = e^{i\mu}$ with μ real, then $\mathbf{\Xi}$ has real eigenvalue $\Lambda = \lambda + \lambda^{-1} = 2\cos\mu$. This implies that $\mathbf{\Xi}$ has three, real, double eigenvalues, Λ_x, Λ_y, and Λ_z, for a stable lattice.

Explicitly $\mathbf{\Xi}$ is given by

$$\mathbf{\Xi} = \begin{pmatrix} \operatorname{tr} A\ I & T & U \\ \overline{T} & \operatorname{tr} D\ I & V \\ \overline{U} & \overline{V} & \operatorname{tr} J\ I \end{pmatrix} \tag{2.35}$$

where

$$B + \overline{C} = T = \begin{pmatrix} h & -f \\ -g & e \end{pmatrix}; \qquad \overline{T} = \begin{pmatrix} e & f \\ g & h \end{pmatrix}, \tag{2.36}$$

$$E + \overline{G} = U = \begin{pmatrix} n & -l \\ -m & k \end{pmatrix}; \qquad \overline{U} = \begin{pmatrix} k & l \\ m & n \end{pmatrix}, \tag{2.37}$$

$$F + \overline{H} = V = \begin{pmatrix} s & -q \\ -r & p \end{pmatrix}; \qquad \overline{V} = \begin{pmatrix} p & q \\ r & s \end{pmatrix}. \tag{2.38}$$

Note especially that the 2×2 diagonal blocks of $\boldsymbol{\Xi}$ are proportional to identity matrix I. For simplifying formulas which follow, two relations, valid for 2×2 matrices, are useful:

$$\begin{aligned} A\bar{A} &= \det A \equiv |A| \\ A + \bar{A} &= \operatorname{tr} A \; I. \end{aligned} \tag{2.39}$$

The characteristic equation is

$$\Delta(\Lambda) = \det \begin{vmatrix} (\operatorname{tr} A - \Lambda) I & T & U \\ \bar{T} & (\operatorname{tr} D - \Lambda) I & V \\ \bar{U} & \bar{V} & (\operatorname{tr} J - \Lambda) I \end{vmatrix} = 0. \tag{2.40}$$

This determinant can be worked out by following Gantmacher.[6] To simplify the algebra it is useful to introduce symbols like

$$\alpha = (\operatorname{tr} A - \Lambda) I. \tag{2.41}$$

Though this is a 2×2 matrix it commutes with everything and can be treated just like a scalar factor. We obtain

$$\Delta(\Lambda) = \Lambda^3 - p_1 \Lambda^2 - p_2 \Lambda - p_3 \tag{2.42}$$

where

$$\begin{aligned} p_1 &= \operatorname{tr} A + \operatorname{tr} D + \operatorname{tr} J = \Lambda_A + \Lambda_D + \Lambda_J \\ p_2 &= -\operatorname{tr} A \;\operatorname{tr} D - \operatorname{tr} A \;\operatorname{tr} J - \operatorname{tr} D \;\operatorname{tr} J - |U| - |T| - |V| \\ &= -(\Lambda_A \Lambda_D + \Lambda_A \Lambda_J + \Lambda_D \Lambda_J) \\ p_3 &= -\operatorname{tr} D \;|U| - \operatorname{tr} J \;|T| - \operatorname{tr} A \;|V| + \operatorname{tr}\left(\bar{V}\bar{T}U\right) = \Lambda_A \Lambda_D \Lambda_J. \end{aligned} \tag{2.43}$$

The expression for p_3 has a suspicious-looking lack of symmetry, but it is invariant to reordering of the (x, y, z) coordinates; so also is its last term. For a stable lattice the three roots of Eq. (2.42) are all real, and an explicit formula can be written for them. Following Press et al.[7], and defining

$$\begin{aligned} Q &= \frac{p_1^2 + 3p_2}{9} \\ R &= \frac{-2p_1^3 - 9p_1 p_2 - 27p_3}{54} \\ \theta &= \arccos\left(R/\sqrt{Q^3}\right), \end{aligned} \tag{2.44}$$

the roots are given by

$$\begin{aligned}\Lambda_1 &= -2\sqrt{Q}\cos\left(\frac{\theta}{3}\right) + p_1/3 \\ \Lambda_2 &= -2\sqrt{Q}\cos\left(\frac{\theta+2\pi}{3}\right) + p_1/3 \\ \Lambda_3 &= -2\sqrt{Q}\cos\left(\frac{\theta+4\pi}{3}\right) + p_1/3.\end{aligned} \tag{2.45}$$

The eigenvalue triplet $(\Lambda_1, \Lambda_2, \Lambda_3)$ is some permutation of the triplets $(\Lambda_x, \Lambda_y, \Lambda_z)$. These can also be labeled $(\Lambda_A, \Lambda_D, \Lambda_J)$ assuming the perturbations away from design, uncoupled optics, leaves the tunes close to their design values. We will assume relabeling has been performed so that $(1, x, A)$ go together, as do $(2, y, D)$ and $(3, z, J)$.

References

1. Klaus G. Steffen, High Energy Beam Optics..
2. Klaus G. Steffen, High Energy Beam Optics, p. 56-62..
3. Koutchouk, J-P, *Trajectory and Closed Orbit Correction*, in Frontiers of Particle Beams; Observation, Diagnosis and Correction, Joint US-CERN School on Particle Accelerators, M. Month and S. Turner, (Eds.), Anacapri, Italy, 1988.
4. A. Verdier and T. Risselada, CERN/ISR-BOM-OP/82-19(1982).
5. G. Bourianoff, S. Hunt, D. Mathieson, F. Pilat, R. Talman, G. Morpurgo, *Determination of Coupled-Lattice Properties Using Turn-By-Turn Data*, SSCL-Preprint-181, 1992.
6. F.R. Gantmacher, *The Theory of Matrices*, Chelsea, New York, 1977, Section II§5..
7. W. Press, B. Flannery, S. Teukolsky, and W. Vetterling, *Numerical Recipes in C,* Cambridge University Press, Cambridge, 1988, p156..