SLS-TME-TA-1999-0014
September 1999

Practical Guidelines for Lattice Design

Andreas Streun

Paul Scherrer Institut
CH-5232 Villigen PSI
Switzerland

Foreword

This paper is based on the “Lattices” course held by Michel Martini (CERN) and the author at the Cern Accelerator School at Bénodet (France) from Aug.30–Sep.9, 1999. The course consisted of seven units of each 90 minutes. Units 2,3,4 covering transverse dynamics, the *BeamOptics* code, analytic and computer aided design of lattice sections as FODO cell, triplet, dispersion suppressor etc. were held by M.Martini, and the contents are well documented in refs.[32] and [2]. Units 1,5,6,7 covering introduction, magnets, global quantities, acceptance and lattice errors were held by the author and thus are subject of this paper.

Chapter 1

Introduction

1.1 Intention

For a beginner the problem of lattice design appears as a jungle of many parameters nested in complicated ways with all kinds of constraints superimposed. It is difficult to obtain an overview, to sort the priorities and to develop a methodical design process. There exist many good lecture notes, articles and books on beam dynamics, but how to build a bridge across from theory to practical lattice design is not obvious for the beginner. It is the intention of this note to close that gap by taking a very pragmatic approach. We will largely do without derivations but rather quote the equations we need and trust those who derived them. However we will always try to be aware of the underlying approximations and limiations of validity. We also include technological constraints like poletip fields or coil sizes of magnets in order to avoid unrealistic designs right from the beginning. Thus we hope to offer a useful complement to theoretical beam dynamics.

1.2 Lattice design

1.2.1 Task definition

The arrangement of magnets and some other devices to guide and focus the beam is called the lattice. Given the purpose of the machine (particle factory, light source...) and its desired performance (energy, luminosity, brightness, lifetime...) lattice design has to find a solution within limitations of available area, budget and technology.

Beyond performance and feasibility the basic requirements to lattice design are, that the solutions should be

simple: few different component types, stadardization, high lattice symmetry

robust: tolerance to errors in alignment and component manufacturing

fail-safe: generally conservative, taking technological risks only when it is really necessary to achieve unprecedented performance

cheap: concerning both installation and operation costs

reliable: thoroughly testing in order to enable significant performance predictions

Of course all these requirements contradict performance and a careful weighting has to be done.

1.2.2 Methodology

We may divide the lattice design process into four phases:

1. **Preparation:** Definition of performance issues (→sec.4.1) and boundary conditions. Acquisition of information on available building blocks (magnets) for composing the lattice concerning their properties and technical limitations (→ch.2).

2. **Linear lattice:** Arrangement of linear building blocks (quadrupoles and bending magnets) to obtain the desired global (i.e. concerning the lattice as a whole) quantities like circumference, emittance, tune, etc. This phase is most creative and deals with inventing periodic cells, matching sections, insertions etc. in linear approximation (→ch.3 and 4).

3. **Nonlinear lattice:** Introduction of sextupoles and RF cavities for stabilization of particels with momentum deviations. Due to the nonlinearity of these elements dynamic acceptance, i.e. stability limits for transverse and longitudinal deviations from the ideal reference particle becomes the main design issue (→ch.5).

4. **Real lattice:** Investigation of lattice performance in presence of magnet misalignments, multipolar errors, vibrations etc. and development of correction schemes (→ch.6). This final design phase ends with a significant prediction on the performance of the lattice including the estimated imperfections of reality.

Little of lattice design can be done analytically, most tasks require the aid from a computer code. In phase 2 the designer needs a visual code to “play” with lattices and optimize them interactively. In phase 3 special tools for acceptance optimization and for tracking in order to check the results will be needed. Finally in phase 4 an extended tracking code for simulation of all kinds of errors is required.

At the end of lattice design construction of the machine may start. Of course lattice design continues for existing machines too in order to improve them further or to understand problems.

1.2.3 Interfacing

Lattice design provides the “skeleton” of the machine, telling how single particles move around the ring. However the “flesh” to be added is the beam current. The maximum beam current is a fuzzy subject to be shared between beam dynamics, vacuum and RF

departments and difficult to predict precisely. In the framework of this paper we assume the beam current as given and do not consider it further. Instead we recommend that the lattice designer should be in contact with his/her colleagues from other departments for different reasons:

Vacuum: Impedance of vacuum chamber affects maximum beam current, pressure affects lifetime, pumps absorbers and flanges require space.

Radiofrequency: RF parameters determine momentum acceptance and other parameters affecting directly the lattice design.

Diagnostics: Beam position monitors have to be inserted at the right locations and also require space.

Magnet Design: Technological limits and geometrical properties of magnets determine maximum magnet strength to be used in lattice design, magnet coils need space, multipolar errors affect the acceptance.

Alignment: Misalignments cause orbit distortions affecting the performance and requiring correction schemes.

Mechanical engineering: Grouping of magnets on stiff girders improves the robustness of the lattice to misalignments and vibrations.

Construction: Devices to be installed by the different departments meet on the design engineers blueprint and conflicts reveal there.

Most important for the lattice designer is to include the space requirements for all devices right from the beginning.

Fig.1.1 gives a comparison of the lattice designers and the design engineers view of a lattice section.

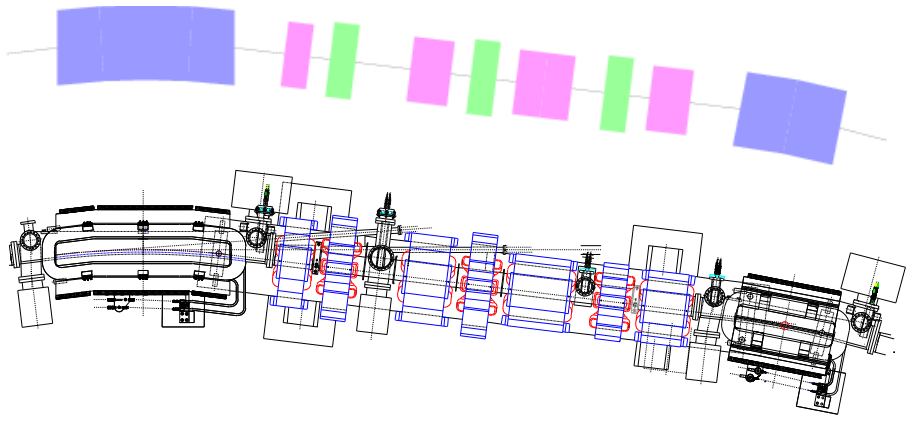

Figure 1.1: A lattice sections as seen by the lattice designer (top) and the design engineer (bottom). Note how the space between ideal magnets is consumed by coils, BPMs, absorbers, pumps, etc.

Chapter 2

Lattice building blocks

2.1 Lattice composition

Lattice design is the art of choosing the suitable building blocks and concatenate them in the appropriate way in order to obtain the desired machine performance within given limitations.

It is important to make a clear distinction between local properties of the building blocks and global properties of the lattice as pronounced by Forest and Hirata [17]:

> "A quantity is called *local* if it is derivable from the individual magnet irrespective of the magnet position in the ring and even irrespective of the ring itself. For example a trajectory of a particle through the magnet is local. [...] *Global* information, on the contrary, is derivable only after the full ring is produced. For example the dynamic aperture has no meaning whatsoever if we cannot iterate the one-turn map (i.e. circulate particles in the machine)."

Concatenation of building blocks is done by coordinate transformations, i.e. translations and rotations. For example a vertical bending magnet is inserted by a horizontal bending magnet preceded by a 90°rotation around the s-axis.

A building block may have any coordinate system, however in practise it is either a cartesian geometry with parallel entrance and exit planes ((x,y) planes, perpendicular to s) and length L or a cylindric geometry with an angle ϕ between the entrance and axis planes and arc length L. Obviously cylindric geometry is more convenient for the description of bending magnets and cartesian geometry for other magnet types and drift spaces.

After assembling all building blocks and closing the ring the one turn map can be calculated, it is just the mapping of the particle vector $\vec{X} = (x, p_x, y, p_y, \delta, \Delta s)$ from one turn to the next $\vec{X}_n \xrightarrow{\text{map}} \vec{X}_{n+1}$ with the order of the map corresponding to the highest power in the coordinates. Thus the closed orbit is a fixpoint of the one-turn map and the transfermatrix is a linearization of the one-turn map around the closed orbit.

For the practical lattice design as subject of this note we will need only a small, idealized subset of the formalism with following approximations:

- Linear beam dynamics with convenient quantities: emittance (the Courant-Snyder invariant), betafunctions, betatron phases, etc.
- Decoupling of subspaces: Synchrotron motion, i.e. dynamics in (δ,Δs) subspace is slow and thus treated as constant parameters over the timescales of betatron motion. Coupling between horizontal (x,p_x) and vertical (y,p_y) subspaces is considered to be small ($\rightarrow$sec.6.4).
- Nonlinearties are treated as perturbations ($\rightarrow$ch.5).
- The so called "design orbit" is just a coincidence of the closed orbit with (most) magnets' symmetry axis for the ideal lattice but is not defined a priori.

It is important to be aware of these simplifications.

Eventually any lattice design has to be tested by tracking particles through the lattice, since tracking concatenates all local transformation of the particle vector from block entrance to block exit including the coordinate transformations between blocks and thus implicitely applies the full one-turn map.

The properties of the building blocks and their limitations will be investigated in this chapter, the properties of the lattice are subject of the following chapters.

2.2 Magnetic fields

2.2.1 Multipole definition

In the local coordinate system of a magnet the field is given as multipole expansion around the local reference axis (x=y=0) by

$$B_y(x,y) + iB_x(x,y) = (B\rho)\sum_n (ia_n + b_n)(x+iy)^{n-1} \tag{2.1}$$

with n the multipole order and $2n$ the number of poles in the magnet, i.e. $n = 1, 2, 3, \ldots$ indicating dipole, quadrupole, sextupole, etc. The b_n are the regular multipoles ($B_x = 0$ for x=0) and a_n the skew multipoles, obtained through a rotation around the s-axis by $90°/n$.

The quantity $(B\rho)$ is called the magnetic rigidity. From the Lorentz force equation it is directly derived as ratio of momentum over charge:

$$B\rho = \frac{p}{q} = \frac{\beta E/e}{n_e c} = \frac{\beta}{n_e} 3.3356 E[\text{GeV}] \tag{2.2}$$

with n_e the number of elementary charges per particle and $\beta = v/c$. For electrons and positrons is $\beta \approx 1$ usually.

By differentiation of eq.2.1 we obtain a useful expression for a pure, regular multipole:

$$b_n = \frac{1}{B\rho}\frac{1}{(n-1)!}\left.\frac{\partial^{(n-1)} B_y(x,y)}{\partial x^{n-1}}\right|_{y=0} \tag{2.3}$$

2.2.2 Pole tip field

The radius of a cylinder around the symmetry axis touching the magnet poles is called the pole inscribed radius or magnet aperture radius R. In case of dipoles instead the full gap $g = 2R$ is used for characterization. The poletip field of a regular magnet is then given from eqs.2.1 and 2.3 by

$$B_{\rm pt} = (B\rho)b_n R^{n-1} = \frac{R^{n-1}}{(n-1)!}\frac{\partial^{(n-1)}B_y(x,y)}{\partial x^{n-1}}\bigg|_{y=0} \tag{2.4}$$

2.2.3 Conventions

Unfortunately there are different definitions of multipole strength. The quadrupole strength, also often called k is defined with different polarities: $k = b_2$ or $k = -b_2$. For the sextupole (and higher multipole) strength, called m or k_s or other, definitons with and without the factorial are used: $m = \pm b_3$ or $m = \pm 2b_3$. In order to avoid misunderstandings it is safer to characterize sextupole strength in terms of $B_{\rm pt}/R^2$. In our definition a quadrupole of positive strength provides horizontal focussing to a beam of positive particles.

2.2.4 Elementary particle tracking

A particle traversing the magnetic field is deflected. If we assume that the motion is paraxial (small angle of inclination x', y' with the axis), that the magnet is short and not too strong (i.e. x, $y \approx$ constant during passage) we obtain for the kick applied to the particle from integration of Lorentz force

$$\Delta x' = -\frac{1}{B\rho}\int B_y dl \qquad \Delta y' = \frac{1}{B\rho}\int B_x dl \tag{2.5}$$

Examples of kicks from pure regular multipoles with effective length L:

$$\begin{array}{lllllll} \text{Quadrupole:} & \Delta x' & = & -b_2 L x & \Delta y' & = & b_2 L y \\ \text{Sextupole:} & \Delta x' & = & -b_3 L(x^2 - y^2) & \Delta y' & = & 2b_3 L x y \end{array} \tag{2.6}$$

Dividing a magnet in many thin slices and applying a series of kicks and drifts would be a first order symplectic integrator of the particle's motion through the magnets. Higher order symplectic integrators apply kicks and drifts in a deliberate way in order to reduce the integration error.

2.3 Building blocks

The elementary building blocks for linear lattice design are bending magnets and quadrupoles for guiding and focusing the beam. The nonlinear lattice design also includes sextupoles used for correction of the quadrupoles' chromatic aberrations. The real

lattice with alignment and other errors also includes small corrector dipoles and skew quadrupoles. In addition every ring needs injection devices (kickers and septa) and one or more RF cavities for acceleration and longitudinal focusing.

Special devices not used in every machines are undulators in light sources, electrostatic beam separators in some colliders, and seldom solenoids for focussing and beam rotation.

2.3.1 Bending magnet

The bending magnet is a block of cylindrical symmetry with a curvature, resp. radius $h_{\rm ref} = 1/\rho_{\rm ref}$, arc length L and bend angle $\Phi = Lh_{\rm ref}$.

The dipole moment $b_1 = B_y/(B\rho)$ provides a curvature $h = b_1$ of a particle's trajectory. The magnetic field is adjusted that $h = h_{\rm ref}$ for the particular energy of the reference particle. For particles at other energies the trajectories' curvatures do not match the coordinate system's curvature, they thus leave the bend off-axis even if they entered on-axis, an effect called dispersion. It is important not to mix $h_{\rm ref}$ which is given by geometry and h which is a function of magnet current.

In a pure sector bend the entrance and exit edges of the magnet are orthogonal to the arc, in the general case the edges may be rotated by angles ζ_1, ζ_2 to be included in the coordinate transformation. Rectangular bends have parallel entrance and exit edges, $\zeta_1 = \zeta_2 = \Phi/2$. Laminated magnets like used in synchrotrons for reasons of eddy current suppression are always rectangular since manufacturing is done by stacking the laminates.

If the bending magnet also has a quadrupole moment b_2 it is called a combined function magnet. Weak focusing synchrotrons used combined function magnets, the definition of the "field index"

$$n = -\frac{\rho}{B_y}\frac{\partial B_y}{\partial x} \tag{2.7}$$

goes back to these times. Recently combined function magnets became again attractive for low emittance booster synchrotrons [34].

The magnet gap g has most impact on magnet design (see below) and also affects slightly the fringe field optics [8].

Summarizing the parameters for description of a bending magnet are
L, $h_{\rm ref} \stackrel{!}{=} b_1$, ζ_1, ζ_2, g, [$b_2 \ldots$].

2.3.2 Quadrupole

Quadrupoles are used for focusing, they are treated in the next chapters. The field gradient is given by $G = \partial B_y/\partial x|_{y=0} = (B\rho)b_2$.

Parameters: L, b_2, R.

2.3.3 Sextupole

The sextupole is used for correction of quadrupole chromatic aberrations, and will be treated in chapter 5.

Parameters: L, b_3, R.

2.3.4 RF cavity

For the purpose of lattice design the cavities are treated as thin elements where particles receive a change in energy depending on the time of arrival expressed as the lag Δs compared to the reference particle: $\Delta E = qV_{\rm rf}\sin(\varphi_s - 2\pi\Delta s/\lambda_{\rm rf})$

Parameters: Wavelength $\lambda_{\rm rf}$ and peak voltage $V_{\rm rf}$.

2.3.5 Correctors

Closed orbit correction (→sec.6.3) requires small horizontal (regular) and vertical (skew) dipole magnets of cartesian geometry. Sometimes they are embedded as additional coils in quadrupoles or sextupoles, however it is important to subtract the correctors maximum field from the maximum pole tip field allowed for the quadrupoles, resp. sextupoles.

Parameters: $\Delta x'(=\int b_1 dl)$, $\Delta y'(=\int a_1 dl)$, g.

2.3.6 Monitors

Beam positions monitors are passive elements for observation of beam position. They are only mentioned because they require space, have to be inserted into the lattice at special locations and there are many of them.

2.3.7 Skew Quadrupoles

Correction of horizontal-vertical coupling due to roll errors in the real lattice requires small skew quadrupole correctors (→sec.6.4). Often they are included into the sextupoles as additional coils.

Parameters: $\int a_2 dl$, R

2.3.8 Injection kickers

Injection requires either a fast kicker to bend the injected beam to the axis (“on-axis injection”) or a closed bump of slower kickers for multiturn injection. Important is time structure and maximum kick to the beam.

Parameters: $\int b_1(t) dl$

2.3.9 Septum

For injection or for beam separation in a collider septum magnets are required. A septum is a copper sheet of high current attached to a dipole magnet in order to provide a sharp cut-off of the field. Distance to the axis and thickness of septum are important for acceptance estimates.

Parameters: L, b_1, distance to axis, thickness

2.3.10 Solenoid

Solenoids appear mostly as parasitic elements in lattice design since the detectors at the interaction regions of colliders often use large solenoidal fields affecting the beam. Focussing by a solenoid is a second order effect while beam rotation is the first order effect, thus the focusing strength goes down with E^2. Use of solenoids is thus restricted to low energies. From the longitudinal field B_s follow rotation angle and focusing strength approximately to $\theta_{\text{sol}} = B_s L/(2B\rho)$ and $k_{\text{sol}} = (B_s L/(2B\rho))^2$. The exact treatment however is rather cumbersome since the transverse vectorpotentials in the Hamiltonian cannot be neglected like for other elements [21].

Parameters: L, θ_{sol}, k_{sol}, R

2.3.11 Undulators

Light sources use undulators to force the electrons on a "slalom course" for emission of bright synchrotron light. In contrary to other elements most undulators use permanent magnets. They are characterized by peak field B_u, wave length of the slalom motion λ_u and the number of periods N_u. The gap g can be very small in modern light sources causing acceptance problems. The most primitive model for an undulator is a series of small dipole magnets of length $\lambda_u/4$ of alternating polarity with empty spaces of same length in between. Wigglers are similar to undulators, but have fewer poles and higher field.

Parameters: λ_u, N_u, B_u, g

2.3.12 Separators

Colliders with particle/antiparticle beams of same energy cannot seperate the beams by magnetic devices but need electrostatic plates. Since Lorentz force is given by $\vec{F} = q(\vec{E} + \vec{v} \times \vec{B})$ electric fields compared to magnetic fields act weakly on high energy beams. Thus the devices are unwieldy and need very high voltage. An alternative are RF separators.

Parameters: L, $\vec{E}$, gap

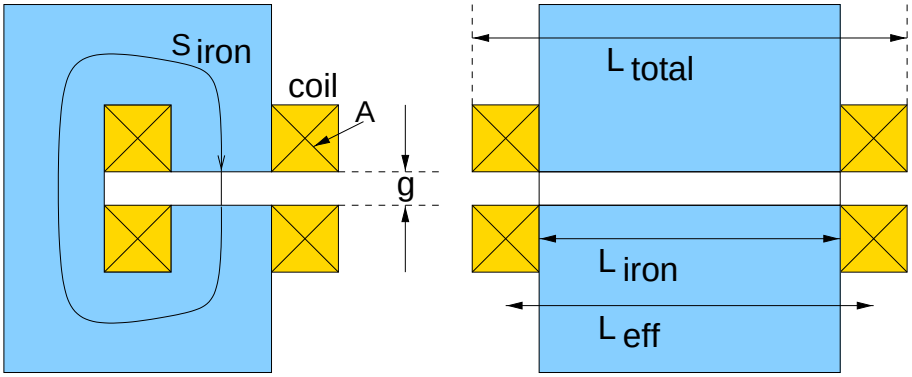

Figure 2.1: Iron dominated dipole magnet. Integrating Maxwell's equation $\oint Hds = \int\int jda$ along the path as shown at the left gives the cross section area A of the coils required to create the field B: $A = B/(2j_c)[S_{\text{iron}}/(\mu_o\mu_r) + g/\mu_o]$ with j_c the current density in the coil. Since the iron permeability $\mu_r \gg 1$ this simplifies to $A \approx Bg/(2j_c\mu_o)$. The figure at right shows the distinction between effective magnet length, iron length and total length due to addition of coil width to iron length.

2.3.13 Space

A very important element not to be forgotten during lattice design is free space. Space requirements for RF systems, diagnostic equipment, vacuum pumps, absorbers, flanges and magnet coils have to be taken into account right from the beginning. The lattice designer should consult the colleagues from other departments early in order to avoid revisions of his beautiful compact lattice...

2.4 Interferences from magnet design

A brief detour into magnet design is required in order to include limitations on magnet strengthes and requirements for distances between magnets into the lattice design process.

2.4.1 Coil size

For an iron dominated magnet like shown in fig. 2.1 the required coil cross section area is given by $A = Bg/(2j_c\mu_o)$. A conservative value for the gross average current density in a water cooled coil (including water pipes, insulations, etc.) is $j_c \approx 2\ldots3$ A/mm^2. Assuming a realistic bending magnet of 1.5 T field and 40 mm gap the cross section is $A \approx 100$ cm^2 and the coil might be quadratic with 10 cm width and height. Although the iron length is shorter than the effective length of the magnet by $L_{\text{iron}} \approx (L_{\text{eff}} - g)$ the coil in our example would add 16 cm to the effective magnet length as used in lattice design. Same considerations can be done for quadrupoles and sextupoles. Figure 1.1 makes it clear.

Some data for the effective coil width, i.e. $(L_{\rm total} - L_{\rm eff})/2$, to be added on both sides of a magnet as free space are listed in table 2.1.

Table 2.1: Some data for effective coil width maximum poletip fields and aperture inscribed radii obtained from a survey on light source magnets [27]

	eff. coil width	max. poletip field	aperture radius
Bending magnets:	6.5 . . . 15 cm	1.5 T	20. . . 35 mm (=g/2)
Quadrupoles:	4 . . . 7 cm	0.75 T	30. . . 43 mm
Sextupoles:	4 . . . 8 cm	0.6 T	30. . . 50 mm

2.4.2 Maximum poletip field

Some data for the maximum poletip field for normalconducting iron dominated magnets are also given in table 2.1. The poletip field is limited due to saturation effects somewhere in the iron. Although magnet iron saturates fully around 2.2 T it becomes nonlinear at lower fields already. In order to maintain the high field homogeneity required for modern machines and to ensure predictability and reproducability the magnet design should avoid saturation. The limits on poletip fields for quadrupoles and sextupoles are lower then for bending magnets since flux lines are compressed due to pole geometry and higher field values appear somewhere else in the yoke. Saturation of a $2n$-pole will create a parasitic $6n$-pole (→sec.6.5).

2.4.3 Magnet apertures

Calculations as done in figure 2.1 can be done for any iron dominated multipole and lead to the result that the required current per coil expressed as windings × current NI or current density × coil cross section $j_c A$ is proportional to the poletip field × aperture radius $B_{\rm pt} R$.

Keeping the multipole strength constant we thus obtain from eq. 2.4 a proportionality of $NI \propto R^n$, e.g. quadrupole currents increase with the square of aperture. The power consumption rises even steeper as with square of current since larger currents require larger coils with longer windings and thus higher resistance, etc. Thus from the magnet design point of view the apertures should be as small as possible.

On the other hand the apertures have to be as large as necessary for the desired lattice performance, in particular for the acceptance (→ch.5).

Since poletip field is the limited quantity, larger aperture decreases the maximum acceptable multipole strength. With larger aperture a quadrupole has to be built longer for maintaining the integrated focusing strength, when the maximum poletip field has been reached. In this way the whole machine size increases with magnet apertures.

Some data for pole inscribed radii are also listed in table 2.1.

2.4.4 Superconductivity

Superconducting magnets reach higher fields and thus can be built much more compact than normalconducting ones: Compact light sources routinely use fields of about 4...5 T (→text following eq.4.6) with conventional NbTi-wires at boiling helium temperature of 4.2 K. Even higher fields of 8...10 T require other, less common materials for the superconducting wires like Nb_3Sn or cooling with superfluid helium [18]. The wiggler for VEPP-2M reached 7.5 T [24], and the final focussing solenoids of the Novosibirsk Φ-factory are designed for 11 T [3].

However superconducting magnets have serious disadvantages compared to normalconducting, iron dominated magnets:

- Far beyond saturation iron behaves magnetically like air, i.e. $\mu_r = 1$, thus the magnet field geometry cannot be shaped with iron poles but only with the coils. This imposes complicated design constraints on the coil geometry.
- Nevertheless the same field homogeneity like with iron dominated magnets cannot be reached usually.
- The forces in the coils are enormous and require thorough mechanical engineering.
- The devices are generally much more expensive, in particular when the cryogenics is not yet in house but has to be built up for the new machine.
- The devices are not as reliable, they may quench due to inhomogeneities in the wires or even due to radiation.

Thus lattice design should prefer normalconducting magnets and request for superconductivity only when it is unavoidable due to space restrictions, for example in the interaction region of a collider.

Chapter 3

Transverse dynamics

For an introduction to linear transvese dynamics we refer to refs. [32, 50, 52, 48, 59] covering the linear Hamiltonian of betatron motion with appropriate approximations (small curvature, paraxial motion, piecewise constant fields), the equations of motion, Hill's equation and introduction of betafunctions, betatron phases and dispersion, and how to obtain these quantities from a transfer matrix. These references also give examples of transfermatrices for particular magnets and of basic lattice cells like a triplet, a FODO cell, etc., obtained by multiplication of the concatenated magnets' transfermatrices.

For our purposes we only quote the essential formulae to be frequently used:

The linear betatron motion of a particle as its coordinate at location s is described by

$$x(s) = \sqrt{2J_x \cdot \beta_x(s)}\cos(\phi(s)) + D(s) \cdot \delta \tag{3.1}$$

with β_x, ϕ_x and D betafunction, betatron phase and dispersion function, $\delta = \Delta p/p_o$ the relative momentum deviation and $2J_x$ the betatron amplitude. The same equation applies to the vertical y, with $D(s) \equiv 0$ in case of a flat lattice.

Eq. 3.1 is derived from the canonical transformation of the linear Hamiltonian to the form $H = J/\beta$ with invariant betatron amplitude J and phase ϕ as new variables (thats the reason why it is $\mathbf{2}J_x$), see refs. [48, 59] for details. δ is assumed as a constant since it is only slowly varying ("adiabatic approximation").

α, β, γ are called the twiss parameters (not to be confused with the relativistic parameters!) and related with each other and the betatron phase by

$$\phi = \int (1/\beta) ds \qquad \alpha = -\frac{\beta'}{2} \qquad \gamma = \frac{1+\alpha^2}{\beta} \tag{3.2}$$

The general transfermatrix $M_{a\to b}$ from some location a to another location b in the lattice is described by a transformation to action angle variables (circle transformation to Courant-Snyder coordinates) at location a, followed by a rotation in phase space by the betatron phase advance $\Delta\phi = \phi_b - \phi_a$ and by a backtransformation at location b:

$$M_{a\to b} = T_b^{-1} \begin{pmatrix} \cos\Delta\phi & \sin\Delta\phi \\ -\sin\Delta\phi & \cos\Delta\phi \end{pmatrix} T_a \quad \text{with} \quad T = \begin{pmatrix} \sqrt{\beta} & 0 \\ \frac{\alpha}{\sqrt{\beta}} & \sqrt{\beta} \end{pmatrix} \tag{3.3}$$

Multiplication gives

$$M_{a\to b} = \begin{pmatrix} \sqrt{\frac{\beta_b}{\beta_a}}(\cos\Delta\phi + \alpha_a \sin\Delta\phi) & \sqrt{\beta_a\beta_b}\sin\Delta\phi \\ \frac{(\alpha_a-\alpha_b)\cos\Delta\phi-(1+\alpha_a\alpha_b)\sin\Delta\phi}{\sqrt{\beta_a\beta_b}} & \sqrt{\frac{\beta_a}{\beta_b}}(\cos\Delta\phi - \alpha_b\sin\Delta\phi) \end{pmatrix} \tag{3.4}$$

For a periodic structure, i.e. $b = a$, this simplifies to the one-turn matrix

$$M_a = \begin{pmatrix} \cos 2\pi Q + \alpha_a \sin 2\pi Q) & \beta_a \sin 2\pi Q \\ -\gamma_a \sin 2\pi Q & \cos 2\pi Q - \alpha_a \sin 2\pi Q \end{pmatrix} \tag{3.5}$$

The transformation of twiss parameters along the lattice is given by a 3×3 matrix composed of elements of $M_{a\to b}$, here expressed in term of sine and cosine type solutions:

$$\begin{pmatrix} \beta \\ \alpha \\ \gamma \end{pmatrix}_b = \begin{pmatrix} C^2 & -2SC & S^2 \\ -CC' & S'C + SC' & -SS' \\ C'^2 & -2S'C' & S'^2 \end{pmatrix} \cdot \begin{pmatrix} \beta \\ \alpha \\ \gamma \end{pmatrix}_a \quad \text{with} \quad M_{a\to b} = \begin{pmatrix} C & S \\ S' & C' \end{pmatrix} \tag{3.6}$$

Chapter 4

Global quantities

Equipped with knowledge on and tools for lattice design we may face now the purpose of the machine and find out how the lattice has to be shaped in order to obtain best performance.

First we will see how the figures of merit depend on quantities we can manipulate by lattice design, then we will examine the limitations and interconnections of these quantities and how to optimize them.

4.1 Figures of merit

4.1.1 Light source: Brightness

Brightness is the performance issue for undulator based light sources used for research in biology, chemistry, surface physics, etc. Brightness is defined as photon flux density in phase space and within a certain bandwidth of wavelength. On one of the peaks of the undulator's line spectrum and on axis the peak brightness $\mathcal{B}$ is given by [35]

$$\mathcal{B} = \frac{\text{Flux}}{4\pi\Sigma_x\Sigma_{x'}\Sigma_y\Sigma_{y'}} \qquad \left[\frac{\text{photons}}{\text{s}\cdot\text{mm}^2\cdot\text{mrad}^2\cdot 0.1\%\,\text{W}}\right] \tag{4.1}$$

with the Σ indicating the photon beam's rms radius and divergence obtained from a convolution of the electron beam and the photon diffraction parameters by

$$\Sigma_z = \sqrt{\sigma_z^2 + \sigma_r^2} \;\text{ and }\; \Sigma_{z'} = \sqrt{\sigma_{z'}^2 + \sigma_{r'}^2} \qquad z = x \text{ or } y \tag{4.2}$$

with the diffraction source size and divergence given by

$$\sigma_r = \frac{\sqrt{\lambda L_u}}{4\pi} \;\text{ and }\; \sigma_{r'} = \sqrt{\frac{\lambda}{L_u}} \tag{4.3}$$

with L_u the undulator length and λ the photon wavelength under consideration given by

$$\lambda = \frac{\lambda_u}{2n\gamma^2}(1 + K_u^2/2) \;\text{ with }\; K_u = \frac{eB_u\lambda_u}{2\pi m_o c} \tag{4.4}$$

Table 4.1: High brightness light sources in operation or under construction [35] (Status 1996). (The lattice factor F corresponds to eq.4.18, see below.)

Name	Country	E[GeV]	N_{mag}	ϵ_{xo} [nm rad]	F
ALS	USA	1.5	36	3.4	8.9
MAX-2	Sweden	1.5	20	8.7 (+D)	3.9
BESSY-2	Germany	1.7	32	5.2	7.5
ELETTRA	Italy	2.0	24	7.0	3.1
PLS	S.Korea	2.0	36	12	18
SLS	Switzerl.	2.4	36	4.8	4.9
ESRF	Europe	6.0	64	4.0 (+D)	3.7
APS	USA	7.0	80	8.2	11
Spring-8	Japan	8.0	96	5.6	9.6

(+D) indicates dispersive beams in undulators, SLS data 1998

and finally λ_u, B_u the undulator's period and peak field and n the selected harmonic of the spectrum.

The photon flux again is a function of undulator properties, of accepted bandwidth and of course of the beam current. For more detailed and improved brightness formulae see ref. [28].

Since we are dealing with the lattice design we are not interested in absolute brightness numbers but only want to know what we can do in order to increase brightness; increasing the photon flux which depends on undulator design and the maximum beam current is not our task.

Anticipating the results on minimum emittance we consider a modern, few GeV light source with a flat lattice designed for low horizontal emittance ϵ_x of some nm·rad and with the vertical emittance $\epsilon_y = g\epsilon_x$ in the range of some pm·rad due to low coupling related parameter $g \approx 10^{-3}$ ($\rightarrow$sec.6.4). From eqs.4.3 we may see that horizontally the convoluted beam parameters from eq.4.2 are dominated by the electron beam's emittance but vertically it depends on the wavelength: Radiation in the VUV range (photon energy ≈ 100 eV) will be diffraction dominated whereas in the X-ray region (> 10 keV photons) the electron beam is dominating. For the brightness we thus obtain an approximate scaling as

$$\text{VUV:} \quad \mathcal{B} \propto \frac{1}{\epsilon_x} \qquad \text{X-ray:} \quad \mathcal{B} \propto \frac{1}{\epsilon_x^2 g} \tag{4.5}$$

We assumed here that the electron beam has a focus at source point, i.e. $\alpha_z = 0$ and thus $\sigma_z \cdot \sigma_{z'} = \epsilon_z$. The relations 4.5 give a clear directive to light source lattice design: Make the emittance as low as possible. The means to achieve that we will investigate below. The choice of beam energy is guided by experimental demands and undulator technology in order to provide high brightness at photon energies of interest. Table 4.1 gives an overview of high brightness light sources.

Compact light sources for industrial applications as X-ray lithography of wafers used for microchip production have completely different requirements: Brightness is no concern but flux at a photon energy of one or a few keV. The critical photon energy is given by [35]

$$\varepsilon_c[\text{keV}] = 0.665\, B[\text{T}]\, (E[\text{GeV}])^2 \tag{4.6}$$

Since this type of light source has to be as compact as possible and maybe even light weighted in order to make it as a whole an interchangable component in a wafer factory, the choice is made in favour of high magnetic field and low beam energy. Thus compact light sources usually operate below 1 GeV and have a small number of superconduting bending magnets with many radiation ports.

A typical example would be the HELIOS machine, operating at relatively low energy of 700 MeV and using two 180° bending magnets of 4.5T field in a simple racetrack lattice for producing radiation at a suitable energy of 1.47 keV [13].

4.1.2 Collider: Luminosity

Luminosity is the key issue for colliders since the production rate for a particle of interest is given as product of luminosity and cross section for the desired reaction.

Luminosity is generally defined as 4-dimensional overlap integral of two colliding bunches in space and time and in the ultrarelativistic limit given by

$$\mathcal{L} = f_c \int\int\int\int_{-\infty}^{+\infty} \varrho^+(x,y,s+ct)\,\varrho^-(x,y,s-ct)\; 2c\,dt\,ds\,dx\,dy \tag{4.7}$$

with ϱ^- and ϱ^+ the particle distributions within the bunches. Note that the relative velocity of the colliding bunches in lab space is $2c$. $f_c = c/b$ is the collision frequency with b the distance between successive bunches in the beam.

We assume Gaussian distributions for ϱ^- and ϱ^+, as they develop rapidly in electron-positron colliders due to the synchrotron radiation equilibrium (see below). For simplification we also assume identical rms beam parameters for both beams. If we also allow the beams to collide not head-on but at a small crossing angle $2\theta \ll 1$ we may write

$$\varrho^{\pm}(x,y,s\pm ct) = \frac{N^{\pm}}{(2\pi)^{3/2}\sigma_x(s)\sigma_y(s)\sigma_s} e^{-\frac{(x\pm s\theta)^2}{2\sigma_x(s)^2} - \frac{y^2}{2\sigma_y(s)^2} - \frac{(s\pm ct)^2}{2\sigma_s^2}}. \tag{4.8}$$

We assumed the crossing to be horizontal, in case of a vertical crossing x and y have to be exchanged of course.

At collision point the beams have a sharp focus with rms beam radii of σ_x^*, σ_y^*. The beam radii before and after this point are given by

$$\sigma_z(s) = \sigma_z^* \sqrt{1 + \left(\frac{s}{\beta_z^*}\right)^2}, \qquad z = x; y \tag{4.9}$$

Integration of eq.4.7 using eqs.4.8,4.9 gives for the luminosity

$$\mathcal{L} = \frac{f_c N^+ N^-}{4\pi\ \sigma_x^* \sigma_y^*} \cdot S \tag{4.10}$$

We may interpret this equation as a a particle stream of $f_c N^+$ (particles/second) hitting a target of density $N^-/4\pi\ \sigma_x^* \sigma_z^*$ (particles/m^2), thus the 2σ beam cross section is the interaction area.

The "luminosity suppression factor" S describes the increase of effective interaction area due to the divergence from eq.4.9 [33] and to the crossing angle between the beams.

$$S = \frac{2}{\sqrt{\pi}\,\sigma_s} \int_0^\infty \frac{e^{-\left(\frac{s}{\sigma_s}\right)^2} e^{-\left(\frac{\theta s}{\sigma_x(s)}\right)^2}}{\sqrt{1+\left(\frac{s}{\beta_x^*}\right)^2}\sqrt{1+\left(\frac{s}{\beta_y^*}\right)^2}}\, ds \tag{4.11}$$

Usually the suppression factor is about $S = 0.8 \ldots 0.95$.

Eq.4.10 gives us still no guideline for lattice design except that we should have a sharp focus at interaction point. In order to obtain a more useful formula we introduce the space charge parameters [19]

$$\xi_z^\pm = \frac{r_e}{2\pi\gamma^\pm} N_b^\mp \frac{\beta_z^*}{\sigma_z^*(\sigma_x^* + \sigma_y^*)} \qquad z = x; y \tag{4.12}$$

r_e is the classical electron radius. The space charge parameters are related to the maximum tune shift a particle may experience during collision due to the focusing from the oncoming bunch. Experience at many machines tells that these parameters are generally limited to values < 0.05 otherwise the beams become unstable. For our purpose of lattice design they thus nicely hide the beam current, we accept that they are limited and do not consider them anymore.

Actually eq.4.12 is two equations for the two transverse planes, however it is of course the same particle number in both equations. Eliminating it we can relate the ratios of emittances and betafunctions to the aspect ratio of the beam at collision $V = \sigma_y^*/\sigma_x^*$:

$$g = \frac{\epsilon_y}{\epsilon_x} = \frac{\xi_x}{\xi_y} V \qquad\qquad \frac{\beta_y^*}{\beta_x^*} = \frac{\xi_y}{\xi_x} V \tag{4.13}$$

Introducing eqs.4.12,4.13 into eq.4.10, also using eq.6.7 and assuming same energy and current for both beams gives

$$\mathcal{L} = \frac{\pi c \gamma^2}{r_e^2} \frac{(1+V)^2}{(1+g)} \frac{\epsilon_{xo}}{\beta_y^* b} \xi_x \xi_y S \tag{4.14}$$

We see that the use of round beams ($V = g = 1$) would provide double luminosity, however we don't consider that option here, leading to rather exotic lattices [3][9]. Instead we assume flat lattices with $V \ll 1$ and $g \ll 1$.

It is at first puzzling that luminosity increases with emittance, since large emittance widenes the focus, however remember that the current limitations depend on the beam sizes too and have been eliminated by eq.4.12. Actually one can easily show for ideal flat or round beams that $\xi \propto N/\epsilon_{xo}$, i.e. the space charge parameter expresses a limitation of phase space particle density. Thus emittance enters in square via N but only linear via $\sigma_x^* \sigma_y^* = \epsilon_{xo} \beta_y^* \xi_x / \xi_y$.

For the lattice design we thus obtain three guidelines:

- Emittance should be as large as possible, i.e. as large as the beam pipe permits.

- β_y^* should be very small but larger than the bunchlength σ_s, see eq.4.11. β_x^* follows from eq.4.13. With β_y^* as small as 1 cm a special lattice section, the so called "mini-beta-insertion", equipped with extremely strong focusing elements, has to be designed.

- The distance between successive bunches b should be minimized. With crossing beams every RF bucket could be used in principle thus making $b = \lambda_{rf}$. In head-on collision the beam have to be united and sepearated as close to the interaction point as possible in order to avoid parasitic crossings. In case of particle/antiparticle beams of same energy the separation has to be done with electrostatic plates or RF devices, else magnets too can be used. An empirical law for the minimum distance in parasitic crossings, i.e. close encounters of vertically separated flat beams is given by $\Delta y > 2.5\sigma_x$ [43].

There are many possible layouts of colliders depending on the purpose:

The electron-positron particle factories go for highest luminosity at a fixed energy, either with same beam energies like the Φ-factories [5][3] or with different energies like the B-factories preferring a boost of the B-mesons. The center of mass energy to be supplied in beam collision is given by the particle to be produced of course, ranging from 1.02 GeV for the Φ-meson to 161 GeV for W-pairs.

High energy machines are in search of new physics in new energy scales and thus are made for beam energies as high as possible like the electron-positron collider LEP operating at 100 GeV or the new hadron collider LHC operating at 7 TeV [18] There are also hybrid colliders like HERA for collisions of electrons and protons and heavy ion colliders like RHIC [39].

The ring layout is a double ring in case of different energies or particles, however with particle/antiparticles at same energy both beams could share the same vacuum chamber, avoiding parasitic crossings by electrostatic excitation of betatron oscillations.

This is the so called "pretzel" scheme as applied at CESR [44].

Table 4.2 gives an overview over some colliders.

Table 4.2: Circular colliders existing or under construction [22]

Name	country	particles	E [GeV]	$\mathcal{L}$ [10^{32}cm^{-2}s^{-1}]
DAΦNE	Italy	e^+e^-	2×0.7	5
BEPC	China	e^+e^-	2.8	0.2
VEPP-4M	Russia	e^+e^-	6.0	1
CESR	USA	e^+e^-	6.0	6
PEP-2B	USA	e^+e^-	3.1 + 9.0	
KEK-B	Japan	e^+e^-	3.5 + 8.0	100
PETRA	Germany	e^+e^-	2×40	
LEP	Europe	e^+e^-	100	0.6
HERA	Germany	e^-p	12 + 40	0.1
RHIC	USA	pp, Ions	2×250(p)/$\frac{100}{u}$(Au$^+$)	0.1 (p)
Tevatron	USA	$p\bar{p}$	2×1800	1.8
UNK	Russia	pp	2×3000	
LHC	Europe	pp	2×7000	100

4.1.3 Lifetime

Performance in terms of brightness and luminosity is only useful if the beam lifetime is long enough to do the experiments. Of course both high brightness and high luminosity lead directly to particle losses. The most important processes are:

- Touschek effect is scattering of particles within the bunch, leading to a transfer of transverse to longitudinal momentum exceeding the momentum acceptance. The loss rate is inversely proportional to the bunch volume, thus light sources with their low emittance beams suffer most from Touschek effect. Touschek lifetime is sensitive to momentum acceptance thus making momentum acceptance of the lattice an important design issue to be discussed in detail in the next chapter.

- Beam beam bremsstrahlung is scattering of particles between colliding bunches also leading to a momentum change exceeding the acceptance. However the loss cross section σ_b depends only weakly on the momentum acceptance. With the half lifetime simply given by $T = N_o/(\sigma_b \mathcal{L})$ [36] and since luminosity is what we want, the only way out is to start with a large inventory of particles in many bunches.

- Residual gas contributes mainly by bremsstrahlung and elastic scattering on nuclei to losses, the first effect leading to a transverse deflection and the second to a momentum change of the particles. Most sensitive to elastic scattered particle losses are light sources due to the narrow gap of the undulators. From the lattice design little can be done. It is the task of the vacuum department to reduce the gas pressure to a level of 1 nTorr.

- Quantum lifetime in electron storage rings is just the loss of particles at the tails

of the Gaussian to the beampipe walls or to the RF bucket's separatrix. As a rule of thumb the beampipe resp. the bucket should at least accomodate 6.5σ of the transverse distribution in the ideal lattice or better 10σ taking into account orbit distortions. This effect limits the maximum emittance of colliders but is of no concern for light sources.

For more information on lifetime see refs.[60, 61, 36, 45].

4.2 Emittance ϵ_{xo}

For light sources emittance is *the* design criterion.

The natural horizontal emittance of a flat lattice in practical units is given by

$$\epsilon_{xo}[\text{nm·rad}] = 1470\ (E[\text{GeV}])^2 \frac{\langle \mathcal{H}/\rho^3 \rangle}{J_x \langle 1/\rho^2 \rangle} \tag{4.15}$$

with $\mathcal{H}$ the so called lattice invariant or dispersion's emittance,

$$\mathcal{H}(s) = \gamma_x(s) D(s)^2 + 2\alpha_x(s) D(s) D'(s) + \beta_x(s) D'(s)^2 \tag{4.16}$$

and $\langle \ldots \rangle$ an average over the lattice.

The horizontal damping partition number J_x is given by eq.4.25 below, in most cases it is close to unity.

In case of an isomagnetic lattice, i.e. all magnets having same bending radius, eq.4.15 simplifies to

$$\epsilon_{xo}[\text{nm·rad}] = 1470\ (E[\text{GeV}])^2 \frac{\langle \mathcal{H} \rangle_{\text{mag}}}{\rho J_x} \tag{4.17}$$

with $\langle \ldots \rangle_{\text{mag}}$ an average taken over the magnets only.

4.2.1 Minimum emittance

Solving the integral over $\mathcal{H}$ and minimizing the result in respect to the values of α_x, β_x, D and D' at magnet centre or entrance gives the theoretical minimum emittance [56, 16, 55].

Assuming that there are identical cells with only one type of bending magnet with deflection angle Φ, and further assuming that $\Phi/2 \ll 1$, which is valid for most light sources ($\Phi < 20^\circ$ gives $< 1\%$ error), the emittance can be written as

$$\epsilon_{xo}[\text{nm·rad}] = 1470\ \frac{(E[\text{GeV}])^2}{J_x} \frac{\Phi^3 F}{12\sqrt{15}} \tag{4.18}$$

with Φ the deflection angle per bending magnet in radian and $F \geq 1$ a factor depending on the lattice type. The theoretical minimum emittance is achieved for $F = 1$. Table 4.1 gives the F values for some existing light sources: The region of operation generally is given by $F > 3$.

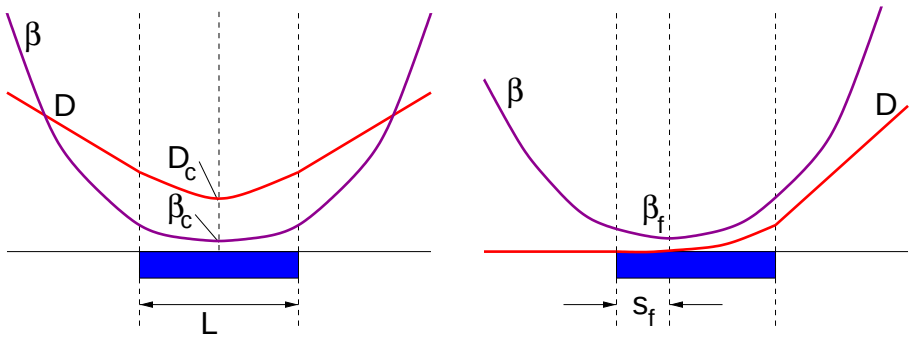

Figure 4.1: Requirements to obtain minimum emittance from a center bend (left) or from an end bend (right)

Note that emittance is independant from the bending radius, resp. the magnetic field, but increases cubically with the angle per bending magnet. That's why light sources have many cells with relatively short bending magnets.

For two basic situations as shown in fig. 4.1 the constraints on betafunction and dispersion for obtaining lowest emittance and the minimum factor F were calculated:

- Center bending magnet: The beam has a focus ($\alpha_{xc} = D'_c = 0$) at the magnet center, dispersion and betafunction are symmetric in respect to the bend center and the dispersion is nonzero everywhere. Then we get, with $L = \rho\Phi$ the magnet length:
$$\beta_{xc} = \frac{1}{2\sqrt{15}}L \quad D_c = \frac{1}{24\rho}L^2 \quad \Longrightarrow \quad F = 1$$

- End bending magnet: The beam enters the bending magnet with zero dispersion. Then we get a constraint for the distance s_f of the focus (where $\alpha_x = 0$) from the entrance edge and for the betafunction at that focus:
$$s_f = \frac{3}{8}L \quad \beta_{xf} = \sqrt{\frac{3}{320}}L \quad \Longrightarrow \quad F = 3$$

It is interesting to note the following:

- Zero dispersion in the center of a bending magnet does *not* provide the minimum emittance.

- A lattice consisting of only center bends with dispersion everywhere provides the lowest emittance achievable, it is 3 times lower than in a DBA (double bend achromat) lattice where each cell is made from two end bending magnets. As a consequence the DBA lattice of ESRF starting with dispersion free straights was later tuned into a dispersive mode in order to reduce the emittance further [35]. SOLEIL plans for a dispersive lattice right from the beginning [30]. However the local *effective* emittance relevant for the brightness has to include the projection of momentum

spread σ_δ to the horizontal dimension via dispersion and is given by

$$\epsilon_{x,\text{eff}}(s) = \sqrt{\epsilon_{xo}^2 + \epsilon_{xo}\mathcal{H}(\int)\sigma_\delta^2} \tag{4.19}$$

- In TBA and higher bend achromats the end magnets contribute most to emittance, consequently they should be made shorter by a factor $\sqrt[3]{3}$ in order to compensate for the factor 3 larger value of F [29].
- FODO lattices are not suitable for light sources since $F \approx 100$.
- As to be seen from table 4.1 ELETTRA almost operates at the minimum emittance achievable for a DBA lattice with dispersion free straight sections. MAX-2 and ESRF also achieving low F-values take into account slightly dispersive straights.

We will now investigate the deviations from ideal emittance conditions. We restrict this consideration to a lattice with center bends like shown in fig.4.1 (left). Defining dimensionless parameters

$$b = \frac{\beta_{xc}}{\beta_{xc,\min}} \qquad d = \frac{D_c}{D_{c,\min}}$$

with the index $_{\min}$ denoting the ideal values to obtain the minimum $F = 1$. Introducing b and d into $\mathcal{H}$ after some algebra leads to the equation of an ellipse

$$\frac{5}{4}(d-1)^2 + (b-F)^2 = F^2$$

shown in fig. 4.2. In order to learn more about the cell providing a factor F we impose constraints on periodicity, i.e. $\alpha_x = D' = 0$ at the entrance and exit of cell. Due to symmetry β_x and D have same values at entrance and exit anyway. In the approximation of small deflection angle $\Phi/2 \ll 1$ the transfer matrix B from center to exit of the bending magnet is given by

$$B = \begin{pmatrix} 1 & L/2 & \rho\phi^2/8 \\ 0 & 1 & \phi/2 \end{pmatrix}$$

with the third column describing the dispersion production. Now let the rest of the cell, from bend exit to cell exit, be described by a matrix M from which we only know that it contains no other bending magnet and that it is symplectic of course, i.e. $|M| = 1$.

Starting from the optical parameters described by b and d in the magnet center we get the constraint on M that the matrix $M \cdot B$ zeroes α_x and D'.

We are interested in the full cell betatron phase advance Ψ and calculate it from

$$\cos\Psi = \text{Trace}\left(M \cdot B \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} (M \cdot B)^{-1} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\right)$$

since the first half of the cell is the mirror image of the second half. Finally after some algebra we arrive at

$$\Psi = 2\arctan\left(\frac{6}{\sqrt{15}}\frac{b}{(d-3)}\right)$$

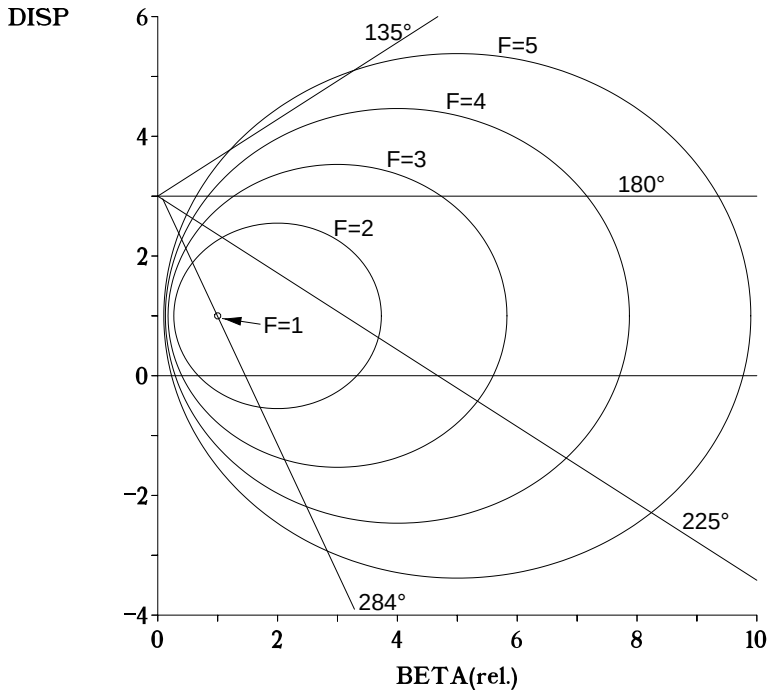

Figure 4.2: Ellipses of constant ratios F of emittance to minimum emittance as a function of the deviation from ideal dispersion and β_x values for obtaining minimum emittance. Also shown are lines indicating the phase advance per cell [26].

This equation describes lines of constant Ψ value in the (b, d) plane intersecting at d=3, b=0, as shown in fig.4.2[26]. Reaching the minimum emittance requires a phase advance per cell of 284.5°. Ideal lattices based on those cells have been studied [15]: They require "empty cells" alternating with the magnet cells in order to accomodate the additional focus for obtaining high phase advance. An example of a minimum emittance cell is shown in fig.4.3.

Existing light sources operate at $\Psi < 180°$ and accept a larger emittance of $F \approx 3 \ldots 5$, see table 4.1.

The vertical emittance is produced by coupling and spurious vertical dispersion ($\rightarrow$sec.6.4).

4.2.2 Emittance in colliders

Colliders need large emittance, the upper limit primarily given by the aperture restriction from the beam pipe. In order to accomodate N_σ standard deviations in the beam pipe and to obtain a momentum acceptance at least as large as the RF momentum acceptance the vacuum chamber a_x at any location s of the lattice has to fulfill the requirement

$$a_x(s) \geq N_\sigma \sqrt{\epsilon_x \beta_x(s)} + D(s) \max(N_\sigma \sigma_e; 2\delta^{rf}_{\text{acc}}) \tag{4.20}$$

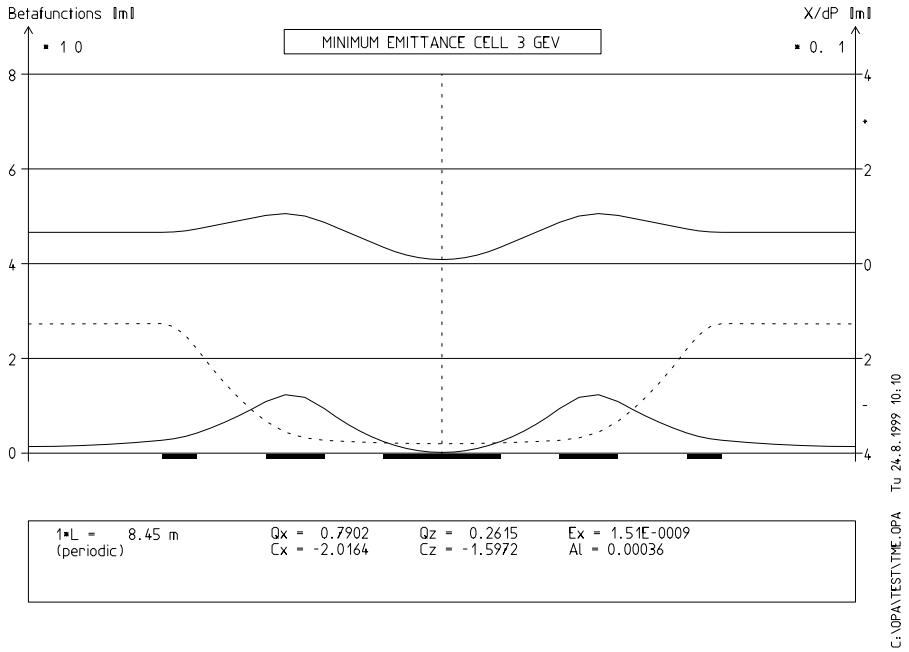

Figure 4.3: Minimum emittance cell: The 10°gradient free sector bending magnet with optimum betafunctions and dispersion in the center (b=d=1) creates the minimum emittance (Ex) of 1.5 nm ·rad at 3 GeV. The tune advance (Qx) of 0.7902 corresponds to the ideal phase advance of $\Psi = 360° \cdot \Delta Q_x = 284.5°$.

The factor 2 with $\delta_{\rm acc}^{rf}$ takes into account that the momentum acceptance has to keep particles from the beam core scattered in dispersive regions (→sec.5.3).

A minimization over all lattice locations gives the upper limit for the emittance.

4.3 Other lattice parameters

In the following we list constraints and dependancies for other lattice parameters which are usually not the primary target of lattice design. They nevertheless require thorough treatment because they need to be adjusted precisely like the tunes or they have to be kept within feasible ranges like the energy loss per turn. We will begin with lattice parameters common to all rings and continue with radiation related parameters which are only of concern for electron/positron rings.

4.3.1 Circumference C

For saving space and building costs one would like to make a compact machine. However small circumference is not always the best neither in lattice performance not in cost. Inserting some space may relax the optics, improve the acceptance, etc. It is also important to take into account from the very beginning of lattice design all kinds of spaces required for magnet coils, beam position monitors, corrector magnets, absorbers, pumps, flanges, etc.

The circumference also must be an integer multiple of the RF wavelength $C = h\lambda_{rf}$, with h called the harmonic number. h should have a nice prime factor decomposition allowing different filling patterns.

Circumferences of existing machines range from 98 cm in Ritsumeikan PSR[35] to 27 km in LEP.

4.3.2 Periodicity $N_{\rm per}$

Periodicity is the number of identical supercells making up the total lattice. High periodicity has fundamental advantages:

- simplicity: most optics calculations are based on the simple supercell
- stability: few systematic resonances (see below), better dynamic acceptance
- cost saving: larger series of a few different components

Periodicity of existing machines ranges from 1 in DORIS to 40 in APS[35].

4.3.3 Working point Q_x, Q_y: The tune diagram

The (Q_x,Q_y)-space, called the tune diagram is covered with resonances appearing as lines described by $aQ_x + bQ_y = p$. The resonance order is given by $a + |b|$. If p is an integer

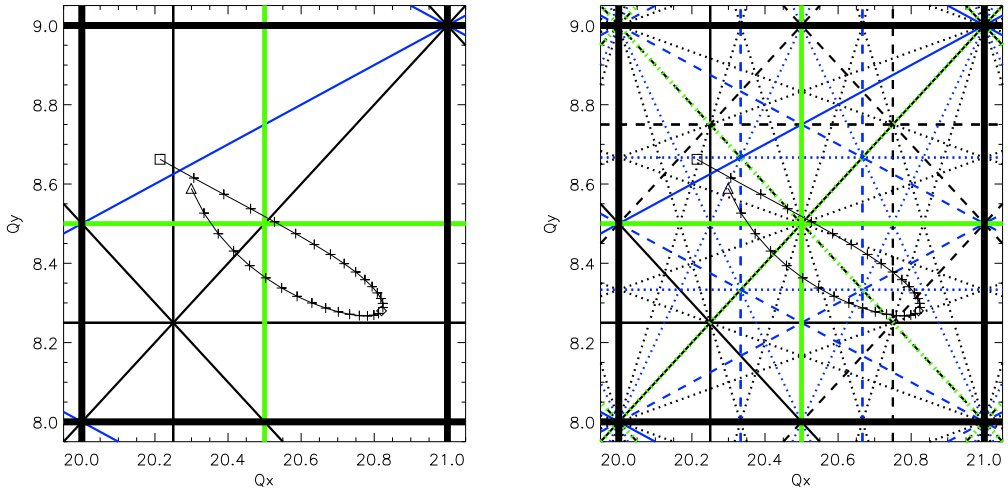

Figure 4.4: Tune diagram for ideal (left) and real (right) lattice of the Swiss Light Source. Solid lines are systematic, dashed non-systematic and dotted skew resonances. Larger thickness corresponds the lower order. Resonances up to 4^{th} order are shown. The curved line is the tune walk for relative momentum deviations of $\pm 8\%$, one segment corresponding to 0.5%.The beam stability limit is determined by reaching the half integer resonances $2Q_x = 41$, $2Q_y = 17$ and amounts to $\pm 5.5\%$, well beyond the RF-momentum acceptance of 4%. For the real lattice including multipolar and misalignment errors a breakdown in acceptance around -2% momentum deviation was observed in tracking and is probably due to crossing the $Q_x + Q_y = 29$ sum resonance.

multiple of $N_{\rm per}$ the resonance is systematic, i.e. amplified by the cell structure, else it is inhibited by periodicity. Even b identifies regular and and odd b skew resonances. The magnets in a flat lattice are all of regular type, see eq.2.1. Thus neither skew nor non-systematic resonances appear in the ideal latice, but show up in the real lattice including multipolar and misalignment errors ($\rightarrow$ch.6). Figure 4.4 gives an example.

There are many constraints for placing the working point:

- It must not be at integer to avoid closed orbit instability due to dipole errors.
- It must not be at half integer to avoid beam blow up due to gradient errors.
- It must not be at a sum resonance to avoid mutual amplification of horizontal and vertical betatron oscillations.
- It has to stay away from sextupole resonances (see next chapter on acceptances).
- Colliders should use a tune close above integer or half integer since then the tune shift for given space charge parameter is smallest and collective two-beam instabilities are not excited [14].

- Hadron colliders take must take into account resonances up to 12^th order while 5^th order is usually sufficient for electron machines due to radiation damping. Considerable "free spaces" in the tune diagram for hadron machines are only available close to the main difference resonances [20].

- Multiturn injection requires a fractional tune of > 0.2 in the plane of injection in order for the injected satellite to clear the septum in the turn following injection.

There are more constraints on the working point from collective instabilities. Since the beam also spreads out in the tune diagram due to amplitude and momentum (see figure 4.4) dependant tune shifts and due to beam beam collision, it is not easy to find a good location eventually.

Generally the fractional part of the tune is more important than the integer, but most important is the flexibility of the lattice to move the tunes independently in the tune diagram, since the best working point eventually is not found till operation. Thus lattice design has to ensure this flexibility.

4.3.4 Chromaticities ξ_x, ξ_y

Chromaticity is the variation of tune with momentum due to chromatic aberrations of the quadrupole, it is given by eq.5.7 and naturally negative. Strong quadrupoles at large betafunctions contribute most to chromaticity. Thus light sources and colliders both suffer from large chromaticity due to the strong focusing, which in case of light sources is distributed along the arcs and in case of colliders is localized at the mini-β-insertion. Chromaticity has to be corrected to zero or positive values by means of sextupoles. The nonlinear field of the sextupoles causes dynamic aperture problems to be discussed in the next chapter. Chromaticity cannot be avoided, however one may try during lattice design to relax the optics by carefully adjusting spaces to produce not more chromaticity than inevitable while keeping the lattice performance.

4.3.5 Momentum compaction factor α

The difference in pathlength travelled by a particle at given relative momentum deviation δ within one revolution of the reference particle is given by

$$\frac{1}{C}\frac{\Delta s}{\delta} = \alpha - \frac{1}{\gamma^2} \quad \text{with} \quad \alpha = \oint_C \frac{D(s)}{\rho} ds \tag{4.21}$$

with C the machine circumference and γ the normalized beam energy. Eq.4.21 shows two competing effects: For particles with higher momentum the pathlength in the magnet is longer (α) thus it will be delayed, but since it is faster it will catch up. The machine is isochronous, i.e. all particles need the same time for one turn, if $\gamma = 1/\sqrt{\alpha} =: \gamma_{tr}$, called the transition energy. With $\alpha \approx 10^{-1} \ldots 10^{-3}$ we get $\gamma_{tr} \approx 3 \ldots 30$, which is $1.5 \ldots 15$ MeV for electrons and $3 \ldots 30$ GeV for protons. Transition thus is only relevant for proton

synchrotrons, where special measures have to be taken when crossing transition energy during the ramp [25].

Concerning the lattice design one would like a large value of α in order to raise the synchrotron frequency and with it the threshold for turbulent bunch lengthening and other instabilities, and a low value for increasing the RF momentum acceptance or for obtaining short bunches. This has to be decided from case to case. However usually there are other dominant design criteria leaving for α not much to choose.

Lattices can be made isochronous by using negative dispersion in the bending magnets. In this case the next order $\alpha_1 = d^2s/d\delta^2$ has to be manipulated by means of sextupoles in order to obtain longitudinal stability again.

*

We now come to the radiation related parameters. This concerns only electron rings, where synchrotron radiation leads to an equilibrium distribution of the electrons in the bunch within damping times of a few milliseconds. The equilibrium values can be adjusted by the lattice design. Generally radiation damping is quite welcome since it stabilizes the beam and makes the electrons forget their history, however the power consumption due to radiation becomes prohibitive for electron energies beyond ≈ 100 GeV.

In hadron machines instead the particle motion may be damped by special installations like an electron cooler, transfering the entropy from the hadrons to a dense, low-energy electron beam.

In the following we display a collection of formulae, mostly taken from Sands [49], without derivations only in order to illuminate the mutual dependancies and implications on lattice design. Since we deal only with electron machines we set the velocity to the speed of light and make no distinction between momentum and energy.

4.3.6 Energy loss per turn U

The energy loss per turn is given by

$$U[\text{keV}] = 26.5(E[\text{GeV}])^3\, B[\text{T}] \tag{4.22}$$

Small machines either use high field bending magnets or install additional wigglers like in DAΦNE [5] or VEPP-2M [24] in order to increase the radiation and get shorter damping times from eq.4.25. In compact light sources high field magnets anyway meet the demand for higher critical energy of the radiation, see eq.4.6.

In contrary large machines operate with low field magnets in order to keep the power consumption within limits. Concerning beam energy LEP certainly represents the high end of electron rings: With a field of only 0.1 T making up the large circumference of 27 km the electrons and positrons loose 3 GeV per turn. With 2×2.5 mA beam currents this leads to a RF power consumption of 15 MW [23]. In normalconducting cavities most of the power is dissipated in the cavity walls and it thus increases with the square of the RF voltage V_{rf} (which has to be larger than U, see eq.4.23). With superconducting cavities the beam loading dominates and the power follows U linearly.

4.3.7 Synchronous phase φ_s

Like the closed orbit in the transverse the longitudinal fix point is given when the bunch is located at the so called synchronous phase, where the RF voltage exactly compensates the radiation loss:

$$\sin \varphi_s = U/V_{rf} \tag{4.23}$$

Every momentum or phase deviation results in synchrotron oscillation which return to the synchronous phase by radiation damping.

4.3.8 RF Momentum acceptance $\delta^{rf}_{\rm acc}$

The RF momentum acceptance is the height of the bucket at synchronous phase and given by

$$\delta^{rf}_{\rm acc} = \sqrt{\frac{2U\lambda_{\rm rf}}{\pi E \alpha C}(\cot \varphi_s + \varphi_s - \pi/2)} \tag{4.24}$$

A low value of the "longitudinal driftspace" αC increases the momentum acceptance from the lattice side. Everything else is subject to RF-design.

4.3.9 Damping times and partition numbers τ_i, J_i

The damping times depend on the energy loss per turn and are in flat lattices given by

$$\tau_i = \frac{2CE}{cUJ_i} \quad i = x; y; s \;\text{ with }\; Jx = 1 - \mathcal{D} \quad Jy = 1 \quad J_s = 2 + \mathcal{D} \tag{4.25}$$

where the quantity $\mathcal{D}$ is given by

$$\mathcal{D} = \frac{1}{2\pi}\int_{\rm mag} D(s)(b_1(s)^2 - 2b_2(s))ds \tag{4.26}$$

The integral is only to be taken over the bending magnets where $b_1 \neq 0$. If the bending magnets also use large gradients like in some low emittance lattice concepts [34], the gradients have to be adjusted carefully in order to ensure damping, in all dimensions: $-2 < \mathcal{D} < 1$. In a separate function lattice with no gradient in the bends, $\mathcal{D}$ is proportional to α and thus small in most lattices. Note that $\sum J_i = 4$, i.e. the damping may be shifted between the dimensions but the sum is limited.

4.3.10 Energy spread σ_e

The rms relative energy (or momentum) spread is given by

$$\sigma_e = 6.64 \cdot 10^{-4} \cdot \sqrt{B[T]\, E[GeV]/J_s} \tag{4.27}$$

Obviously lattice design can do little on manipulating the energy spread assuming that beam energy and bending field have been already chosen for other reason. This is a problem for Tau-charm factories because the J/ψ resonance at 3.1 GeV c.m. is only 87 keV wide [42], but the absolute energy spread from eq.4.27 would be $\sigma_e E \approx 1$ MeV.

4.3.11 Bunch length σ_s

Rms bunch length is given by

$$\sigma_s = \sigma_e \sqrt{\frac{\alpha C}{2\pi} \frac{E}{U} \lambda_{rf} \tan \varphi_s} \tag{4.28}$$

Again, like with the momentum acceptance, the RF-design has more influence than the lattice design since the lattice quantities α, E, U, C affecting σ_s are already assigned to other requirements.

Colliders prefer short bunches since the requirement $\beta_y > \sigma_s$ has to be fulfilled. Light sources prefer longer bunches for increasing the Touschek lifetime. Unfortunately both machine types get the opposite of what they want: The momentum compaction factor α from eq.4.21 depends on the average dispersion in the bending magnets, which again is related to the emittance (see below). Thus colliders with large emittance get long bunches and light sources with low emittance get short bunches. A way out for light sources is installation of a higher harmonic cavity for bunch lengthening [54][31].

Chapter 5

Acceptance

5.1 Acceptance definition

Closed orbit stability is not enough. The lattice has to accept particles with deviations from the ideal orbit in all six coordinates x, p_x, y, p_y, δ and Δs in order to provide sufficient beam lifetime and to allow multiturn injection.

We distinguish physical acceptances, determined by the beam pipe diameters, and dynamic acceptances, defined by the onset of chaotic or instable motion and particle loss beyond some critical amplitude due to nonlinear resonances. Generally acceptance is defined by the 6-dimensional phase space volume where particles are stable, i.e. where they perform bounded oscillations. In most machines the coupling between the subspaces is not too strong so we may separate horizontal, vertical and longitudinal acceptance as projections from 6D to 2D-spaces.

Then the acceptances are invariants of the lattice like the Courant-Snyder invariant in case of linear uncoupled betatron motion. Local projections of the transverse acceptance to real space $\{x, y\}$ give the dynamic apertures, which are not invariants. The transverse acceptance is usually measured in mm·mrad, the aperture in mm. Note that dynamic aperture plots like fig. 5.7 as presented frequently for characterization of lattice performance have little meaning if the local betafunction is not mentioned.

It is unusual to mention the 2D longitudinal acceptance of the RF bucket, instead its projection to the axis of relative momentum deviation δ, called RF momentum acceptance, is quoted.

Determination of dynamic acceptance is not trivial, since the equations of motion are nonlinear and thus not integrable usually. Stability of motion has to be proven by simulation, i.e. probing points in 6D-space on stability by tracking. A test particle is considered to be stable if it survives the tracking, i.e. its amplitudes stay within some limits. Of course this depends on the number of turns to be tracked. Electrons fortunately "forget" their history due to radiation damping, thus tracking one damping time usually is enough ($10^3 \ldots 10^4$ turns) . With protons many more turns need to be tracked ($> 10^6$). However during primal lattice design, even short term tracking ($\approx$100 turns) is sufficient to test the lattice acceptance, at least to reject bad designs, since of course the results

become only worse with more turns.

It is a general criterion for lattice performance, that the pure dynamic acceptance, i.e. the phase space separatrix calculated excluding the cut off beyond beampipe, should be larger than the physical acceptance. To achieve this requires a lot of effort in modern lattices suffering from large nonlinear distortions of the dynamics due to strong sextupole magnets needed for correction of the large chromaticities. Pure trial and error will not succeed, a strategy for deliberate optimization of acceptance is required.

5.2 Physical acceptance

An initial lattice design composed from ideal quadrupoles and bending magnets has a purely linear dynamics and the dynamic acceptance is thus infinitely wide. (Actually this is not exactly true, since the equations of motions had been derived assuming paraxial motion, i.e. "small" deviations from the closed orbit, like $x \ll \rho$.)

A particle however will be lost if $|x(s)| \geq a_x(s)$ where $a_x(s)$ is the half width of the vacuum chamber. The linear betatron motion $x(s)$ of a particle is given by eq.3.1. Since we consider many turns we drop the betatron phase and find the physical acceptance as the maximum possible betatron amplitude, $A = 2J_{max}$, by identifying $x(s)$ with its limit $a_x(s)$ and taking the minimum from all locations in the lattice:

$$A_x = \min\left(\frac{(a_x(s) - |D(s)\cdot\delta|)^2}{\beta_x(s)}\right) \tag{5.1}$$

Since A_x is an invariant of the linear betatron motion we get the local projection of the acceptance, i.e. the minimum and maximum x-values that the particle is just not lost from eq. 3.1:

$$x(s) = \pm\sqrt{A_x \cdot \beta_x(s)} + D(s)\cdot\delta \tag{5.2}$$

5.3 Momentum acceptance

Momentum and phase acceptance as provided by the RF bucket height and length can be considered like longitudinal physical acceptances, although they are dynamic actually, because they are usually constant along the lattice and decoupled from the transverse dynamics.

But momentum acceptance is also restricted by the transverse acceptance of the lattice: From eq. 5.1 we also see that A_x disappears for momentum deviations $|\delta| > \min(a_x(s)/|D(s)|$, i.e. when the closed orbit hits the vacuum chamber.

Momentum acceptance of the lattice is relevant for the beam life time, usually determined by scattering events, either within the bunch (Touschek effect) or with a colliding bunch (beam beam bremsstrahlung) or with residual gas atoms or by emission of synchrotron radiation quanta. These events cause a sudden change in particle momentum while leaving the transverse coordinates almost constant. After the scattering event the

particle vector is given by $(\approx 0, \approx 0, \approx 0, \approx 0, \delta, 0)$ since theB scattered particles come from the beam core where the transverse coordinates are very small. Now the particle will start a betatron oscillation relative to the dispersive closed orbit. For a linear lattice the amplitude of this oscillation is given by

$$A_x = \gamma_{xo}(D_o\delta)^2 + 2\alpha_{xo}(D_o\delta)(D'_o\delta) + \beta_{xo}(D'_o\delta)^2 = \mathcal{H}_o\delta^2 \tag{5.3}$$

with $\alpha_{xo}, \beta_{xo}, \gamma_{xo}$ the twiss paramters at location "0" where the scattering event ocurred, D_o, D'_o the dispersion and its slope and $\mathcal{H}_o$ the lattice invariant from eq.4.16.

The particles will perform oscillations according to eq. 3.1 resulting at another location s in the maximum excursion

$$x(s) = \sqrt{A_x\beta_x(s)} + D(s)\delta = \left(\sqrt{\mathcal{H}_o\beta_x(s)} + D(s)\right)\cdot\delta \tag{5.4}$$

Like in derivation of eq. 5.1 we identify x with a_x as loss criterion and get the *local* momentum acceptance for location "0" as

$$\delta_{\mathrm{acc}}(s_o) = \pm\min\left(\frac{a_x(s)}{\sqrt{\mathcal{H}_o\beta_x(s)} + |D(s)|}\right) \tag{5.5}$$

Thus the momentum acceptance provided by the lattice varies for different locations and is the minor of the values from eq.5.5 and the RF momentum acceptance which does not vary along the lattice.

From eq.5.5 may be easily shown that the lattice momentum acceptance at location of maximum dispersion is approximately half of it in a dispersion free region.

Beam lifetime calculations finally have to integrate scattering rates and momentum acceptances over the lattice [53].

5.4 Chromaticity correction

Dynamic acceptance of a linear lattice is identical to the physical one and the momentum acceptance is determined by linear dispersion or by the RF bucket size. But this picture is spoiled completely by the chromaticity: Due to chromatic aberrations of the focusing elements the machine tune becomes a linear, falling function of momentum δ. The chromaticity $\xi = dQ/d\delta$ is a negative number which may be rather large if the lattice has strong focusing elements like it is the case in high brightness light sources as well as in high luminosity particle factories.

Large negative chromaticity must not be tolerated for two reasons:

- The machine tune has to stay away from integer or half integer numbers otherwise field or gradient errors will amplify coherently and destroy the beam (→sec.6.2). Thus even in the optimum case of frac(Q)=0.25 the momentum acceptance is restricted to $|\delta| < 1/(4\xi)$ which gives an unacceptable small number for most machines.

- Negative chromaticity excites the fundamental mode of the head tail instability, a collective oscillation of electrons in head and tail of the bunch leading to very fast beam loss. Suppression of the fundamental mode requires non negative chromaticity [11].

Sextupoles in dispersive regions of the lattice are used to compensate the chromaticity: From eq.2.6 we get the chromatic kick error for a particle with momentum deviation δ traversing a quadrupole of integrated strength $b_2 L$

$$\Delta x' = b_2 L \delta x \qquad \Delta y' = -b_2 L \delta y$$

In a dispersive region, where $x_\delta = D\delta + x$, $y_\delta = y$ the kick provided by a sextupole is given by (keeping up to second orders in products of δ, x and y)

$$\begin{aligned} \Delta x' &= -[2b_3 L D\delta]x - b_3 L(D\delta)^2 - b_3 L(x^2 - y^2) \\ \Delta y' &= [2b_3 L D\delta]y + 2b_3 L D x y \end{aligned} \tag{5.6}$$

Obviously the terms in square brackets act like a quadrupole and thus may compensate the chromatic errors from the quadrupoles if sextupole strength and dispersion are properly chosen. Neglecting contributions from bending magnets the chromaticity as is given by [48, 12]

$$\xi_x = \frac{1}{4\pi}\oint_C (2b_3(s)D(s) - b_2(s))\beta_x(s)\,ds \qquad \xi_y = \frac{1}{4\pi}\oint_C (-2b_3(s)D(s) + b_2(s))\beta_y(s)\,ds \tag{5.7}$$

Looking at eq.5.6 we notice that there are other terms than the desired ones, and in fact these terms cause serious problems and lead to the large subject of *dynamic acceptance*: With the introduction of sextupoles the equations of motion become nonlinear and usually can not be integrated any longer. In [4] this effect is aptly referred to as “opening of Pandora’s box”: From here we can only proceed by perturbation theory, derive new quantities that we optimize by introducing the sextupoles in a clever way, but nevertheless, beyond some transverse amplitude the motion will become unstable thus defining finite dynamic acceptances. It will be the goal of the lattice design to create a dynamic acceptance larger than the physical acceptance.

5.5 The sextupole Hamiltonian

Dynamic acceptance optimization will attack the sextupole Hamiltonian by suppressing some components and adjusting others for compensation of chromatic quadrupole errors. In the scope of practical lattice design we will not present the derivation of the required quantities but just quote the results from ref. [4]: The sextupoles are treated as perturbation of the linear Hamiltonian. Linear independant terms of different frequencies, i.e. different multiples of betatron phases, are collected and treated separately. Integrations

over the fields are replaced by sums over integrated strengthes under the assumptions that the magnets are short and the variation of beam parameters inside the magnets is negligible. In contrary to many authors no Fourier transformation for obtaining particular resonance drive terms is done, but the whole Hamiltonian as source of all resonance drive terms is suppressed right from the beginning. The single resonance approach is suitable for lattices where one or a few resonances determine the dynamics and has been widely and successfully applied [59, 1, 40, 46], however in constrained modern lattices with low periodicity and strong sextupoles many resonances contribute to dynamic aperture degradation.

The components of the first order (in sextupole strength!) Hamiltonian are given by [4]

$$\begin{aligned} h_{jklmp} \quad \propto \quad & \sum_n^{N_{sext}} (b_3 l)_n \beta_{xn}^{\frac{j+k}{2}} \beta_{yn}^{\frac{l+m}{2}} \eta_n^p e^{i\{(j-k)\phi_{xn}+(l-m)\phi_{yn}\}} \\ & - \left[\sum_n^{N_{quad}} (b_2 l)_n \beta_{xn}^{\frac{j+k}{2}} \beta_{yn}^{\frac{l+m}{2}} e^{i\{(j-k)\phi_{xn}+(l-m)\phi_{yn}\}} \right]_{p\neq 0} \end{aligned} \tag{5.8}$$

There are 4 chromatic terms, with also the quadrupoles contributing:

- h_{11001} and h_{00111} are just the chromaticities, compare to eq. 5.7: Quadrupoles and sextupoles contribute additive, independant of the betatron phase.

- h_{20001} and h_{00201} drive off-momentum $2Q_x$, $2Q_y$ resonances and cause betafunction beats $d\beta/d\delta$ and 2^{nd} order chromaticity as shown in [4]

There are 5 purely geometric terms:

- h_{21000} and h_{10110} drive integer resonances of type Q_x.

- h_{30000} drives 3^{rd} integer resonances $3Q_x$.

- h_{10200} and h_{10020} drive coupling resonances $Q_x \pm 2Q_y$.

It is instructive to draw vector diagrams as shown in figure 5.5 (upper right) with the contribution from every sextupole (and quadrupole) as a vector, the length given by local betafunctions and dispersion and the magnet integrated strength, and the angle given by the combination of betatron phases according to the resonance family considered.

Looking at these vector diagrams we may guess that appropriate distributions of sextupoles around the lattice could be found, where all first order terms of the sextupole Hamiltonian disappear. In order to find these patters we may write eq. 5.8 as linear system: With the N_{sext} sextupoles grouped into S families $N_1 \ldots N_S$ (a magnet family is a common circuit, i.e. all members have the same strength) we may write the 9 equations as a product of a $9 \times S$ matrix of beam parameters with the S-vector of sextupole strengthes, because the first order Hamiltonian is linear in sextupole strength. The 9-dimensional vector on the right side is either zero for the geometric terms of contains the sums over the

quadrupole contributions. The matrix elements in column s are sums over all sextupoles belonging to the family s ($s = 1 \ldots S$):

$$\left(\sum_{n \in N_s} (b_3 l)_n \beta_{xn}^{\frac{j+k}{2}} \beta_{yn}^{\frac{l+m}{2}} \eta_n^p e^{i\{(j-k)\phi_{xn} + (l-m)\phi_{yn}\}} \right)_{9 \times S} \times \left((b_3 l)_s \right)_{S \times 1} = \left(\sum_n^{N_{quad}} (b_2 l)_n \ldots \right)_{9 \times 1} \tag{5.9}$$

Looking at a symmetry point of the lattice where all imaginary terms cancel, we see that 9 independant sextupole allow us to solve the system and suppress the sextupole Hamiltonian in first order completely. In reality however we have to face several problems:

1. The linear system has the tendency to degenerate, the rank drops to 8 or 7 and no solution can be found, or the solution returns unacceptable values of sextupole strength. In particular this is the case for high brightness light sources, where the phase advance per cell is close to 180° for reasons of low emittance ($\rightarrow$ figure 4.2). In this case h_{20001} becomes proportional to h_{11001}, the chromaticity itself, the lattice will suffer from very low momentum acceptance due to 2^{nd} order chromaticity and no sextupole pattern exists to cure this.

2. The cancellation of sextupoles works on summation over several interleaved families of sextupoles. However locally in the lattice, between sextupoles, nonlinear distortions of betatron motion may appear and cause higher order sextupolar effects, corresponding to a crosstalk between the sextupoles. The Hamiltonian contains another 13 terms of second order in sextupole strength, as derived in [4]:

   - 3 terms are the linear tune shifts with amplitude:
     $\partial Q_x/\partial J_x$, $\partial Q_x/\partial J_y = \partial Q_y/\partial J_x$ and $\partial Q_y/\partial J_y$
   - 8 terms are octupole like and thus drive different modes of octupolar resonances:
     $4Q_x$, $2Q_x \pm 2Q_y$, $4Q_y$, $2Q_x$, $2Q_y$
   - 2 terms generate the 2^{nd} order chromaticities: $\partial^2 Q_x/\partial\delta^2$ and $\partial^2 Q_y/\partial\delta^2$

3. All terms are obtained from perturbation theory, however depending on the strength of the nonlinearities in the lattice the desired dynamic aperture limitation may be at amplitudes beyond the validity of perturbation theory. Thus it is never possible to guarentee that a good sextupole pattern in fact will provide large dynamic aperture. A bit of luck is always required of course, however in the next section we will plan our strategy.

5.6 Optimization of sextupole patterns

We will establish some guidelines on how to place the sextupoles:

5.6.1 Chromaticity correction

For chromaticity correction alone we need 2 families of sextupoles in regions of large dispersion. As obvious from eq. 5.7 the sextupoles for horizontal correction should be at locations were the horizontal beta is large and the vertical beta is small, and vice versa for the vertical correction, otherwise the 2 families would act against each other and push up the sextupole strength. It is essential to place the chromaticity sextupoles deliberately in order to keep their strength low and with it all terms of the Hamiltonian. In a relaxed lattice with low chromaticity like a booster synchrotron probably could have, the dynamic aperture might already be good enough with 2 families and we are done.

5.6.2 Sextupoles in quadrupoles

Looking at eq. 5.8 we may guess that the best places for sextupoles are at exactly the same location like the quadrupoles, either by building hybrid magnets or by mounting them very close to each other. Then all chromatic modes are corrected exactly where they are created. Additional small sextupoles could help to cancel the geometric modes too. However the strengthes of the multipoles would be correlated by $b_3 = b_2/D$ where the dispersion D of course is a global function of the lattice and affected by all magnets. Thus a lattice of this type would have only one mode of operation and no flexibility. Furthermore the purposes of the machines usually require dispersionfree sections where nevertheless chromaticity is created too. For example the interaction region of a collider is equipped with strong quadrupoles operating at large betafunctions (mini-β-insertion) thus creating the major part of the machines chromaticity which has to be transported into the arcs and corrected there.

5.6.3 Non-interleaved sextupoles

Another elegant way to suppress the sextupole Hamiltonian even in higher order is the use of non-interleaved sextupoles [37] like successfully applied to the KEK B-factory: Chromaticity correction is done with pairs of sextupoles apart 0.5 in tune (or 180° in betatron phase), the so called "$-I$-transformer". Then the kicks from the sextupoles on on-momentum particles cancel exactly. Since there are no sextupoles from other families between the two sextupole pairs all nonlinear distortion is hidden and no crosstalk and higher order effects can be generated. The drawback of this solution is, that it needs much space, i.e. increases the ring circumference and that the chromatic $2Q_x$ mode is amplified coherently leading to reduction of momentum acceptance. Thus it is not applicable to all machines.

5.6.4 Multicell cancellation

If there are large arcs in the lattice consisting of some number N of identical cells with tune advances ΔQ_x, ΔQ_y per cell, the sextupole vectors as shown in figures 5.5 will

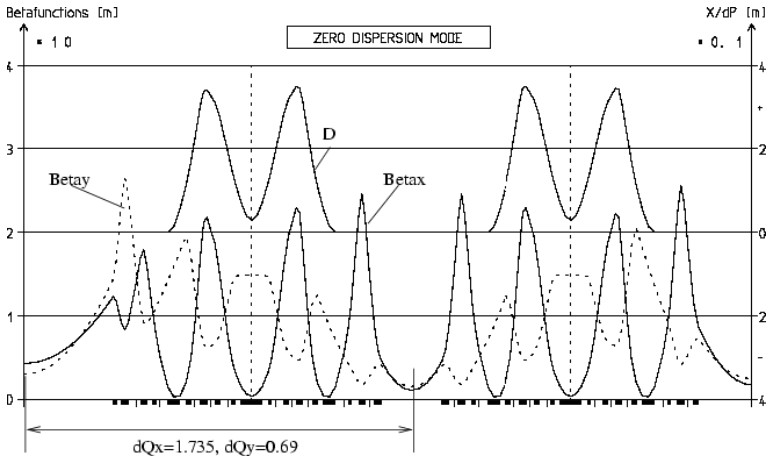

Figure 5.1: Triple Bend Achromat of the Swiss Light Source (two of 12 TBAs shown)

cancel to zero for all modes of the Hamiltonian if all products $N\Delta Q_x$, $3N\Delta Q_x$, $2N\Delta Q_x$, $2N\Delta Q_y$ are integers.

Example: The cell of a light source may have tune advances of $\Delta Q_x = 0.4$ in order to get low emittance and $\Delta Q_y = 0.1$. If there are 5 cells per arc all products become integers and we may hope for good dynamic acceptances. However if we push the tune advance per cell further to $\Delta Q_x = 0.45$ to get even lower emittance, the cancellation breaks down, in particular for the $2N\Delta Q_x$ amounting to a half integer now and we have to expect problems with transverse and momentum acceptance.

5.6.5 Cancellation between sections

If there are no multicell arcs we still may play sections of the lattice against each other. Instead of using pairs of sextupoles half integer apart we may use a whole lattice section and its mirror image.

Example: The Swiss Light Source (SLS) lattice consists of 12 TBAs alternating with dispersion free straight sections for the experiments [51], see figure 5.1.

An horizontal tune advance of about $\Delta Q_x = 1.75 = 7/4$ per TBA was found to provide low emittance and convenient beam parameters in the straights. The vertical tune was set to about $\Delta Q_y = 0.75 = 3/4$. Thus the $2Q_x$ and $2Q_y$ modes cancel between two TBAs and the Q_x and $3Q_x$ modes cancel between two TBA *pairs*. Of course the lattice must not have integer total tunes, furthermore, the TBAs are not identical, since the straights are of different length, so some trial and error was required to find a solution near the ideal cancellation conditions at $\Delta Q_x = 1.735$ and $\Delta Q_y = 0.69$ eventually.

5.6.6 Implications on linear lattice design

From the considerations on minimizing the sextupole Hamiltonian we obtain an essential guideline for lattice design:

For lattices with large chromaticites (high brightness light sources, high luminosity particle factories) linear and nonlinear lattice design are *not decoupled*. One must not proceed by designing the bending magnet and quadrupole arrangement first and add the sextupoles later. Instead from the early planning of the lattice the sextupole pattern has to be taken into account.

5.6.7 Design tool

An automated design tool is required to minimize the sextupole Hamiltonian for a given lattice. The tool should include 1st order terms and from the 2nd order terms at least the amplitude dependant tune shifts. The square sum of all terms with weight factors as chosen by trial and error forms a penalty function for the minimizer. The "knobs" to be adjusted by the minimizer are the sextupole strengthes of course. [38]

Generally one should not try to correct second order terms on expense of first order terms since the second orders are generated by crosstalk between first order terms. Only the amplitude dependant tune shifts are an exception to some amount since they may react very sensitive. In contrary the second order chromaticities are mostly rather "stiff".

Minimization can be rather time consuming in particular when second order terms are involved. Since they are double sums the number of summations goes with the square of the number of sextupoles in the lattice whereas the first orders go only linear. For calculation of second order chromaticites numerical differentiation of the off-momentum closed orbit performs superior to the evaluation of the analytic expressions [4].

Eventually tracking determines the dynamic acceptance and proves whether the optimization was successful.

5.7 Other sources of dynamic acceptance degradation

Sextupoles are usually the dominant source of nonlinearity. Their patterns are optimized based on the ideal lattice. In reality alignment and magnet multipolar errors also introduce nonlinearities and thus attack the dynamic aperture. This will be discussed in the next chapter.

In colliders of course the beam beam effect acts as an extremely strong nonlinear element, but this is a large subject on its own.

5.8 Example

5.8.1 ESRF DBA lattice

As an example we consider an early optics mode of the ESRF lattice as shown in figure 5.2: It consists of 32 DBAs tuned for low emittance and alternating straight sections with high and low beta optics, the periodicity thus is 16. One sextupole named SF for correction of horizontal chromaticity is placed at optimum location in the DBA centre where

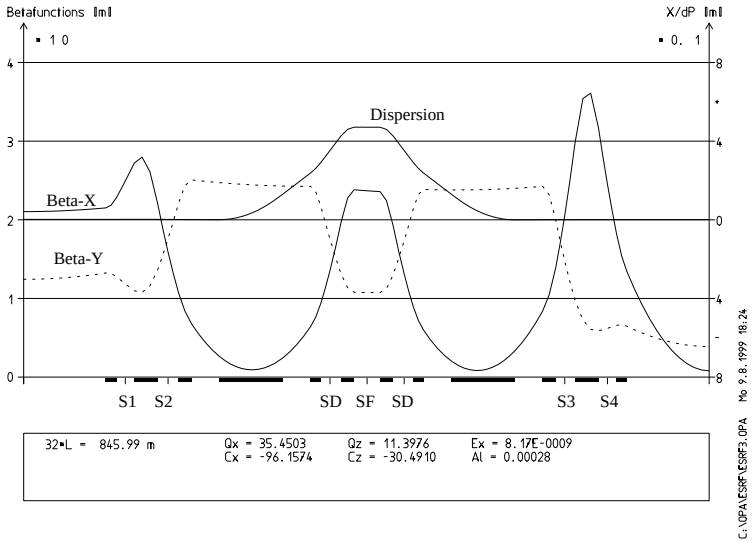

Figure 5.2: Double Bend Achromat of the European Synchrotron Radiation Facility in an early optics mode

dispersion and β_x are large and β_y is small. Two sextupoles SD for vertical correction had to be placed at less convenient locations with lower dispersion because the dispersion always follows β_x. In the straight sections are four auxiliary sextupole families S1-4 for cancellation of geometric terms.

5.8.2 Linear lattice

If all sextupoles are switched off, there are only chromatic terms generated by the quadrupoles in the Hamiltonian as shown in figure 5.3. The geometric terms are zero and the transverse dynamic acceptances are unlimited. However the momentum acceptance is less than $\pm 0.5\%$ in δ since the Q_x moves towards integer due to the large chromaticity as shown in fig. 5.4. Futhermore the head-tail instability would not allow to store a beam.

5.8.3 Chromaticity correction

If the two sextupole families SF and SD are set for zero chromaticity we only cancel the h_{11001} and h_{00111} terms and blow up all the others, see figure 5.5. As we see from figure 5.6 closed orbits should exist within $\pm 2.5\%$ of δ, the interval being limited by 2^{nd} order chromaticity. But the horizontal dynamic aperture as shown in fig 5.7 became very small and would not allow injection.

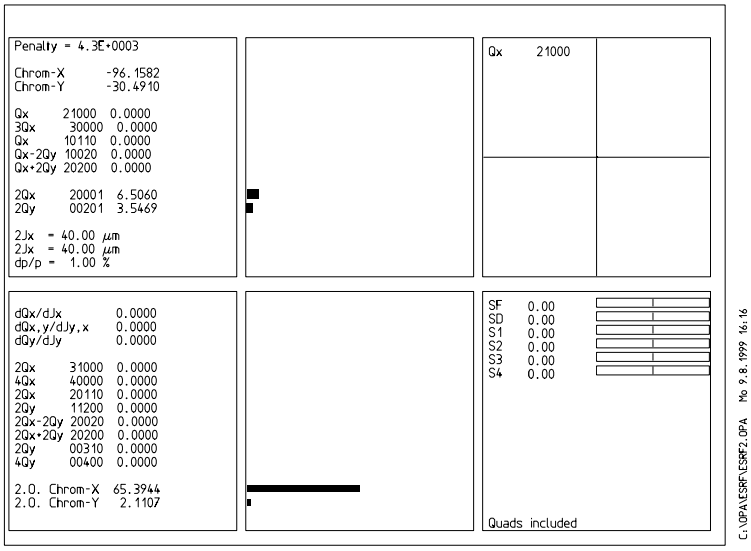

Figure 5.3: Sextupole Hamiltonian: All sextupoles switched off. First (top) and second (bottom) order terms of the sextupole Hamiltonian. Numbers refer to the values of $2J_x$, $2J_y$ and δ denoted in the upper left and multiplied by a factor 10^6 to get handy numbers. Lower right shows a control panel for the sextupole strengthes. Only the quadrupoles contribute to the chromatic terms.

5.8.4 First order optimization

Figure 5.8 shows the sextupole Hamiltonian after minimizing the 9 first order terms from eq. 5.8 by adjusting the four auxiliary families. The improvement in dynamic aperture is spectacular as seen from figure 5.10. Also the interval of closed orbit existence improved to almost $\pm 5\%$ as seen from figure 5.9 because the second order terms to some extent follow the first order terms since they are produced by crosstalks between them.

5.8.5 Second order optimization

Minimizing some 2^{nd} order terms, in particular the amplitude dependant tune shifts, on expense of the first orders as shown in figure 5.11 gains some more dynamic aperture as shown in figure 5.13 but not much in momentum acceptance, see figure 5.12. It can be seen that even higher orders of chromaticity become dominant now. Since ESRF has a large periodicity of 16 it was possible to obtain comparable results based on single resonance decomposition [46].

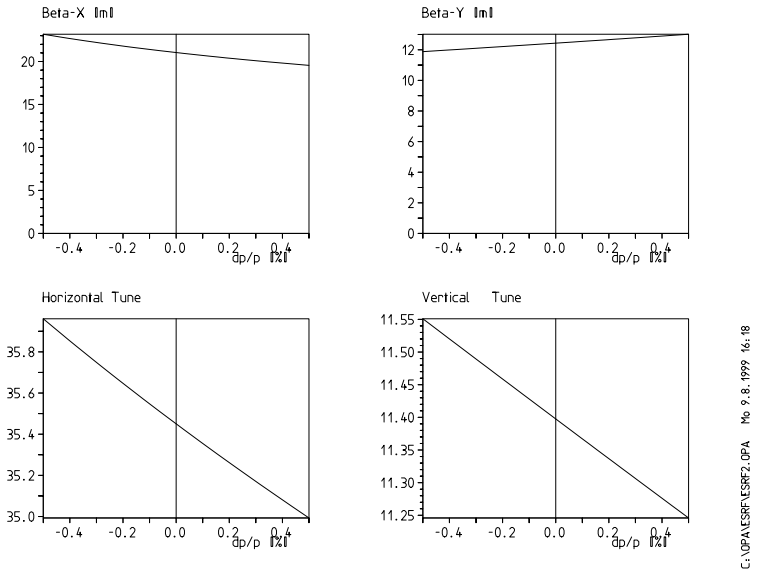

Figure 5.4: Tune vs. momentum: All sextupoles switched off. Without chromaticity correction, the horizontal tune shifts towards an integer for small momentum deviation already.

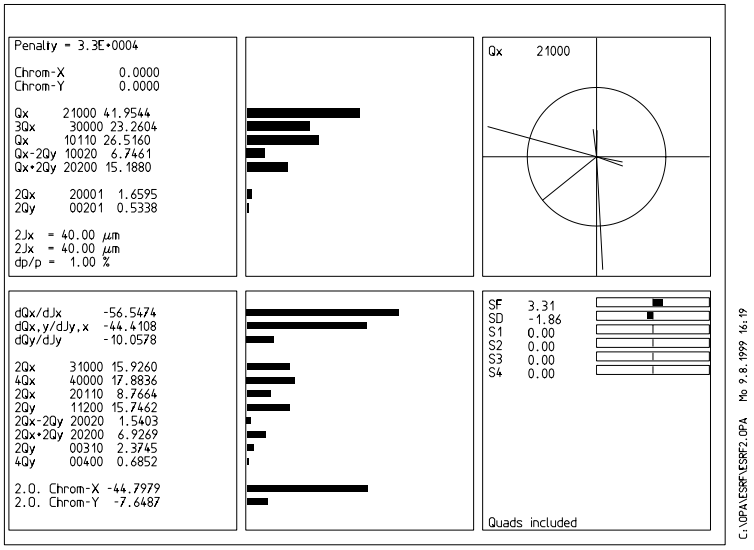

Figure 5.5: Sextupole Hamiltonian: Only chromatic sextupoles SF, SD switched on. The upper right presents the vector diagram for the integer driving term h_{21000}. Powering only the chromatic families corrects the chromaticity but excites all the geometric modes.

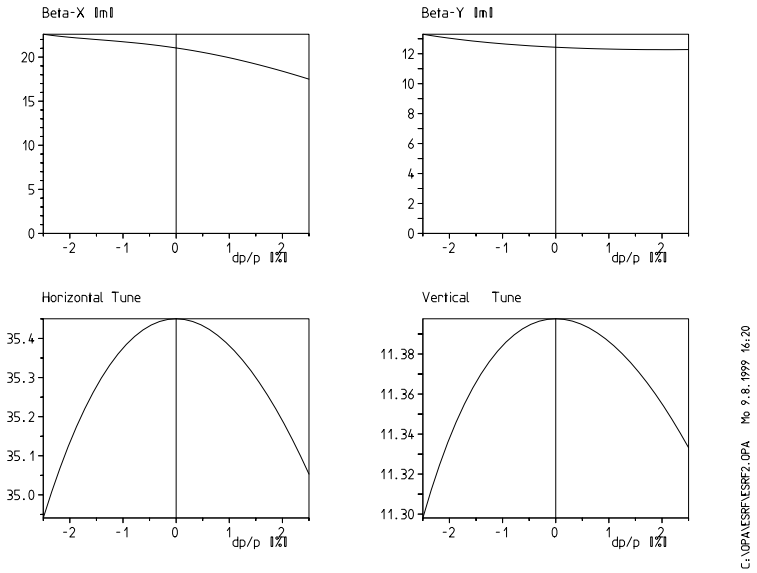

Figure 5.6: Tune vs. momentum: Only chromatic sextupoles SF, SD switched on. The linear chromaticity has been corrected to zero, i.e the curve of tune vs. momentum has no slope at $\delta = 0$. The closed orbit reaches integer at much larger momentum deviations now due to second order chromaticity.

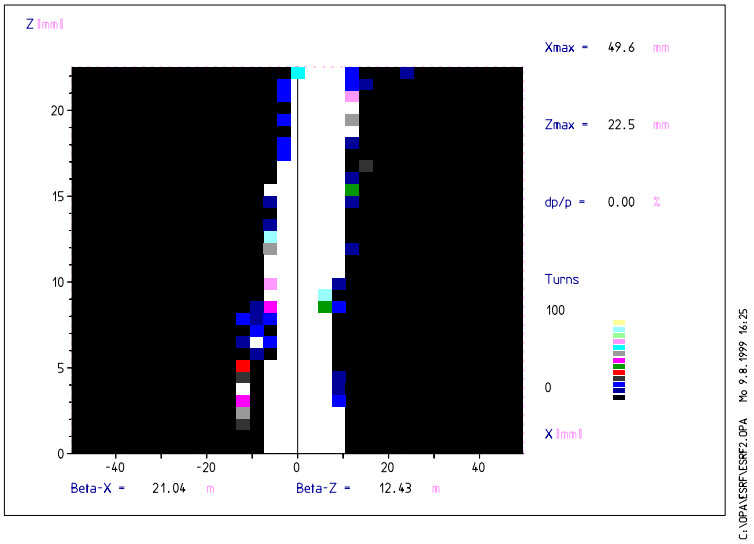

Figure 5.7: Dynamic aperture: Only chromatic sextupoles. The horizontal dynamic aperture does not reach at all the physical aperture. The framing rectangle corresponds to the *double* sized beam pipe transformed to the trackpoint according to eq. 5.1.

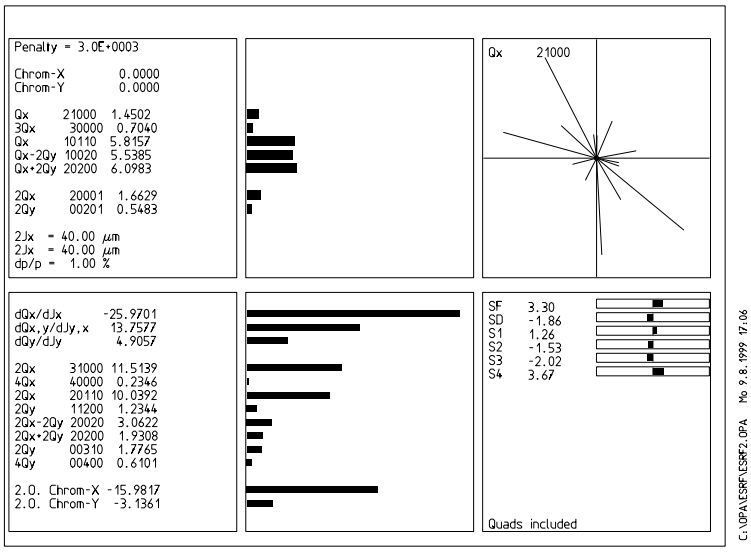

Figure 5.8: Sextupole Hamiltonian: First order optimization. A minimization of the 5 first order geometric terms was done, keeping the chromaticities at zero and neglecting the $2Q_x$, $2Q_y$ modes since they are not accessible by the families S1–S4 anyway (because they sit at dispersionfree locations), which were the "knobs" for the minimizer. Now the vector diagram vector diagram for the integer driving term h_{21000} has collapsed. Also note how the second order terms decreased although they were not controlled by the minimizer.

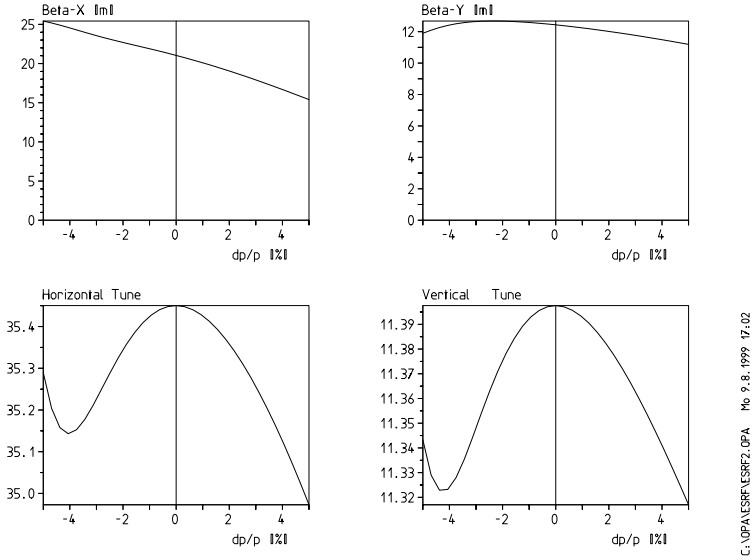

Figure 5.9: Tune vs. momentum: First order optimization. Due to decrease of the second order chromaticity during the minimization process the margin for the closed orbit to stay away from an integer has increased substantially.

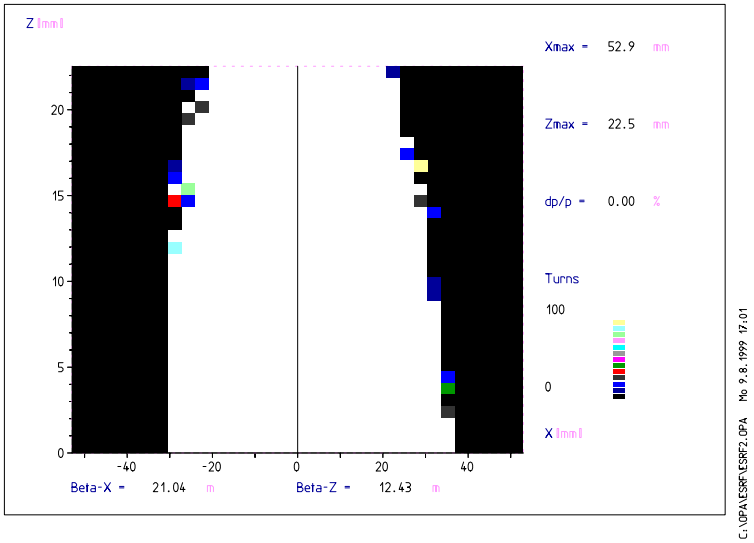

Figure 5.10: Dynamic aperture: First order optimization. Compared to figure 5.7 the improvement is significant: Now the horizontal dynamic aperture exceeds the physical aperture. The framing rectangle corresponds to transformation of the *double* beam pipe size according to eq. 5.1.

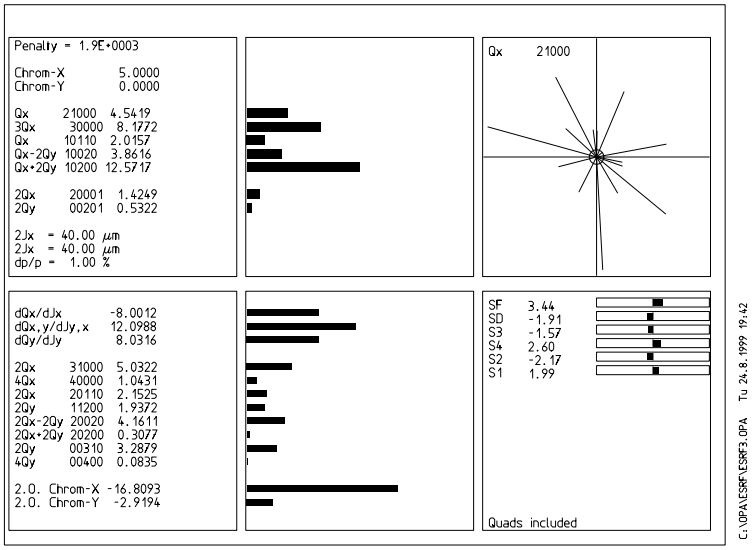

Figure 5.11: Sextupole Hamiltonian: Second order optimization. Also second order had been included in the minimization. Large weights had been set to the amplitude dependant tune terms and to the second order chromaticity. All second orders are reduced, however partially on the expense of the first orders. Since there are only 4 degrees of freedom (6 sextupole famlies, but 2 fixed by the demand for exact chromaticity values) trying different weight factors selects some terms of the Hamiltonian to be suppressed on the expense of others. Distributed weights lead to a compromise solution with the strengthes of the additional sextupoles relaxed.

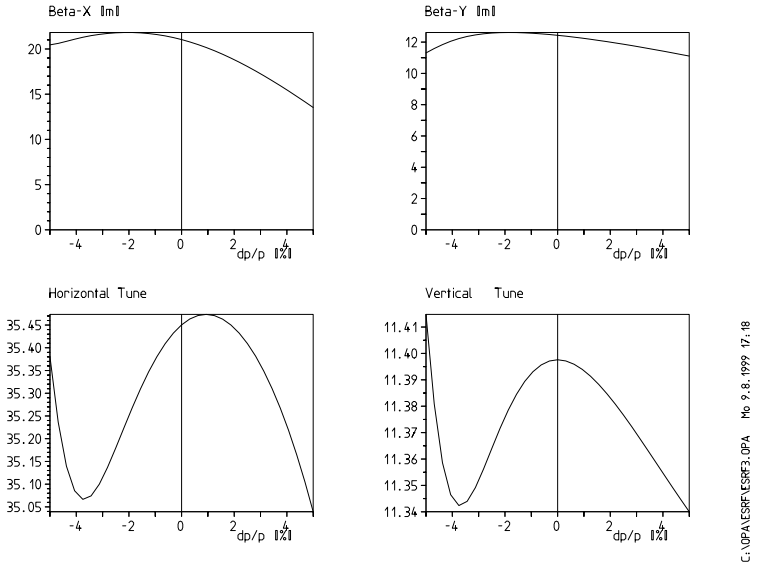

Figure 5.12: Tune vs. momentum: Second order optimization. Little has changed compared to figure 5.9. There was slight increase of the margin for the closed orbit to stay away from an integer by setting the horizontal chromaticity to $\xi_x = +5$ in order to "fold" the curve $Q_x(\delta)$ into the interval.

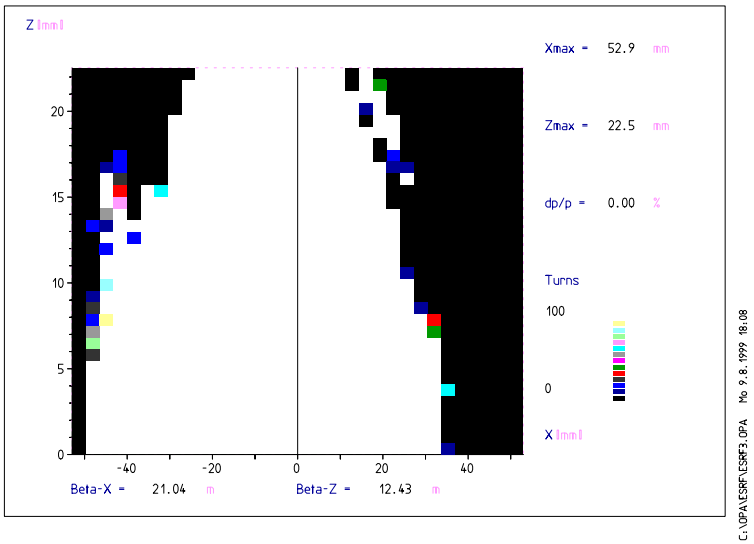

Figure 5.13: Dynamic aperture: Second order optimization. The dynamic aperture has improved again compared to figure 5.10, however in an asymmetric way since the integer term came up again, see figure 5.11. The framing rectangle corresponds to transformation of the *double* beam pipe size according to eq. 5.1.

Chapter 6

Lattice errors

6.1 Error sources

All lattice design work up to now was based on the ideal lattice. Eventually the final design phase tests the robustness of lattice performance and developes cures against various error sources related to the imperfection of reality:

Alignment errors cause translations and rotations of magnets. The naming derived from nautical terms is sway, heave and surge for horizontal, vertical and longitudinal displacements and pitch, yaw and roll for corresponding rotations. Sway (Δx) and heave (Δy) apply kicks to the beam because transverse displacements of all magnets provide a dipole downfeed and thus generate the close orbit distortion. Quadrupole rolls and sextupole heaves present skew quadrupole fields and thus lead to linear coupling. Dipole rolls also excite vertical orbit distortions. Surge, pitch and yaw errors are usually negligible. Misalignments are randomly distributed and static, but they may be partially correlated if magnet groups are mounted on girders, or they can be fully correlated and dynamic if caused by ground vibrations.

Multipolar errors in magnets excite higher order resonances and thus attack the dynamic aperture but do not distort the orbit. Some types of multipolar errors are systematic deviations from the ideal fields, others are random deviations different for each individual magnet.

Random errors of course do not follow the lattice periodicity and thus excite non systematic resonances that had been excluded for the ideal lattice by symmetry. Coupling and higher order skew multipole errors also excite skew resonances. The tune diagram becomes a dense web of resonance lines now. If the ideal lattice's working point had been set near a skew or non-systematic resonance unfortunately, a bad surprise may come up.

Finally dynamic acceptance tracking studies on the real lattice lead to tolerance limits for the mechanical engineering and for the magnet multipole contents. In modern machines the order of magnitudes are rather small, < 0.1 mm for the alignments and $< 10^{-3}$ for the multipole components in the magnets (expressed as relative field component at poletip). Nevertheless careful closed orbit correction or even correction of the coupling by means of skew quadrupoles is still mandatory.

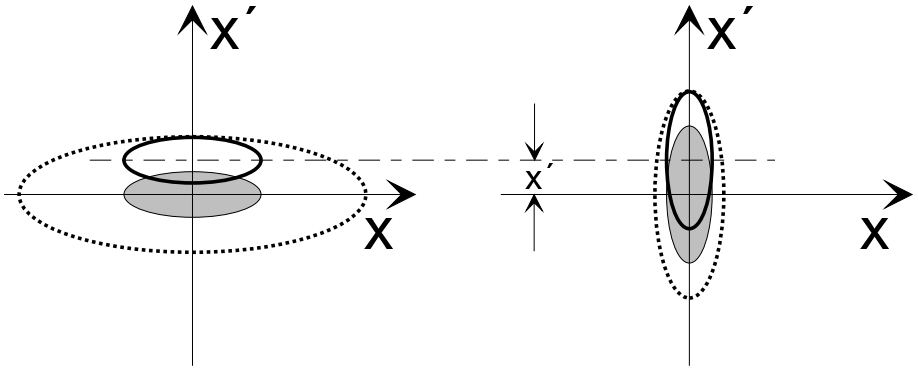

Figure 6.1: At large betafunction, i.e. flat phase space ellipse, (left) the beam is more sensitive to a transverse kick than at low betafunction (right). The dotted elllipse is required to accept the kicked beam.

6.2 Closed orbit distortions

For the ideal lattice the closed orbit is everywhere zero for $\delta = 0$ and given by $D(s)\delta$ for $\delta \neq 0$. In the real lattice transverse kicks caused by dipole errors, $\Delta x' = \int \Delta B_y ds/(B\rho)$ excite a nonzero closed orbit. If the kicks are too strong it may happen that no closed orbit exists.

6.2.1 Dipolar errors

The integrated kicks from a quadrupole with transverse displacements Δx and Δy are given from eq.2.6:

$$\Delta x' = -b_2 L \cdot \Delta x \qquad \Delta y' = b_2 L \cdot \Delta y \tag{6.1}$$

The kicks from a sextupole are given by

$$\Delta x' = -b_3 L \cdot (\Delta x^2 - \Delta y^2) \qquad \Delta y' = 2 b_3 L \cdot \Delta x \Delta y \tag{6.2}$$

A sextupole thus generates coupling between the transverse planes whereas a quadrupole does not.

A single kick $\Delta x'_i$ at location s_i will create a closed orbit oscillation given by[1]

$$x_{\mathrm{co}}(s) = \frac{\sqrt{\beta_x(s)\beta_{xi}}}{2\sin(\pi Q_x)} \Delta x'_i \cos(\phi_x(s) - \phi_{xi} - \pi Q_x) \tag{6.3}$$

with β,ϕ beta function and betatron phase and Q the machine tune. As we see the tune has to stay away from integer otherwise the closed orbit oscillation will become very large. We further see that a kick at location of large betafunction is most harmful. This is obvious from viewing figure 6.1 and has implications on girder design (see below).

[1] Eq. 6.3 is conveniently derived by using the circle transformation eq. 3.3. A nice geometric explanation is given in ref. [58].

6.2.2 Amplification factors

Since there are many kicks from all magnet misalignments, eq. 6.3 has to be replaced by a sum over all N lattice elements (assumed to be "thin"). The orbit distortion at lattice element k is thus given by

$$x_k = \frac{\sqrt{\beta_{xk}}}{2\sin(\pi Q)} \sum_{i=1}^{N} \sqrt{\beta_{xi}} \Delta x_i' \cos(\phi_{xk} - \phi_{xi} - \pi Q) \tag{6.4}$$

The misalignments are not known initially thus in modelling they will be set to random numbers, perhaps with some systematic offset if there are mechanical correlations between magnets. The error distribution is Gaussian with the standard deviation corresponding to the estimated precision of alignment but cut off beyond a threshold that can certainly be measured. The orbit is not observed everywhere but only on the beam position monitors (BPM, see below). The ratio of rms closed orbit deviation at BPMs to rms misalignments is called the closed orbit amplification factor. The amplification factors are highly sensitive to the tune, they range between values of 10 and 100 usually, i.e. 100μm displacements cause 1 to 10 mm closed orbit distortions.

Actually the amplification factor is not completely invariant against scaling since the closed orbit distortion is not linear dependant on the misalignment errors due to the presence of sextupoles, see eq.6.2. However the contribution from the sextupoles is small due to the square dependancy on Δx, Δy. Furthermore the relation between eqs.6.3 and 5.8 tells that at least the sextupole contribution to horizontal closed orbit distortion will be suppressed largely by optimizing the sextupoles for extermination of the integer resonance driving sextupole Hamiltonian h_{21000} for purposes of dynamic acceptance gain.

The tool for determination of amplification factors is a tracking code allowing magnet misalignments to be set and containing a closed orbit finder. Errors are set by a random generator. Many different "seeds" of random numbers have to be tried and averaged in order to obtain a statistically significant statement. One should also pay attention that there are only few seeds where no closed orbit can be found, because this gives the probability for serious commissioning problems. To decrease this set of seeds a threader module may precede the closed orbit finder.

6.2.3 Girders

One would like to move inevitable kicks to locations of low betafunctions since there the closed orbit is less sensitive. This can be done by girders: A group of magnets is aligned to very high precision ($< 50\mu$m) to the girder, which is a rather stiff piece of steel. The transitions between girders are at locations of low beta, the girder are connected by joints, which might be virtually established by optical systems, forming a so called "train link". The "kick" at girder transistion actually is a kink in the optical axis. For the precision of the girder alignment larger tolerances (few 100μm) are allowed than for individual magnets. We have to distinguish then between different types of amplification factors, depending if we consider misalignments of the magnets relative to the girder,

absolute misalignments of the girder joints or the play within the joint, i.e. violation of the train link. Apart from mechanical advantages a girder scheme reduces the closed orbit distortion substantially.

6.3 Closed orbit correction

6.3.1 Correctors and BPMs

In order to correct the orbit we look again at eq.6.4 and turn the tables: By application of appropriate dedicated kicks we can compensate the parasitic kicks due to misalignments. For these additional kicks we need small dipole corrector magnets of variable field since we have to adjust them to a yet unknown pattern of misalignments. We also need to observe the beam, i.e. install beam position monitors (BPMs) at many locations in order to control the success of the corrections.

With one corrector and one BPM at every magnet we certainly could compensate every misalingment locally, however that would be overdone. Generally it is sufficient to install BPMs and corrector at quarter betatron wavelength distance: With a phase advance of 90° a kick applied at a corrector results in a displacement in the subsequent BPM. There another corrector will adjust the beam to the next BPM, etc. In reality it is neither possible nor necessary to set BPM/corrector stations at exactly quarter betatron wavelengthes, but the phase advance has to substantially smaller than 180°, otherwise a corrector would not be able to create any change of displacement at the next BPM. Practically one or two BPM/corrector stations are required in every cell of the lattice.

This gives an important implication even for early lattice design: From the beginning the appropriate locations for BPMs and correctors and the required space for installing the devices has to be taken into account! The corrector really is a small dipole magnet, it has to provide deflections comparable to the kicks from eq.6.1. In order to save space at some machines the correctors are integrated as additional coils in the quadrupoles or sextupoles.

The BPM is made from four small insulated areas in the vacuum chamber which “see” a field proportional to the distance of the beam’s center of mass.

By means of the BPM/corrector system the orbit correction can be done. Naively steering from one corrector to the next BPM usually does not converge. We will briefly discuss the two methods of choice.

6.3.2 Sliding bump method

As long as the betatron phase advances between correctors are reasonably chosen, i.e $0^\circ < \Delta\phi < 180^\circ$, three correctors, let’s call them 1,2,3, always allow to zero the orbit in BPM 2 near the center corrector 2 without affecting the rest of the ring: Corrector 1 opens the orbit bump and 2 provides the necessary kick for 3 to close it again. The required corrector strengthes are obtained straightforward from matrix formalism as shown in [58].

In the next step correctors 2,3,4 are used to zero the beam in BPM 3. Since corrector 2 is close to BPM 2 changing its strength now will only little affect the previously optimised position in BPM 2. In this manner the closed orbit bump "slides" around the ring. Due to nonlinearities in betatron motion and small but nonzero phase advances between a corrector and the corresponding BPM the orbit will not be perfectly corrected after one turn and is iterated a few times or until all BPM readings are below some margin.

6.3.3 Singular value decomposition

With M BPMs in a lattice and N correctors, N and M not necessary equal, eq.6.4 can be written as multiplication of the N-dimensional vector of kicks to be applied by the correctors with a $M \times N$ matrix giving the M-dimensional vector of BPM readings:

$$(x_{\mathrm{bpm}})_M = A_{M\times N} \cdot (\Delta x'_{\mathrm{corr}})_N \quad \text{with} \quad a_{ki} = \frac{\sqrt{\beta_{xk}\beta_{xi}}}{2\sin(\pi Q)} \cos(\phi_{xk} - \phi_{xi} - \pi Q) \qquad (6.5)$$

A is called the response matrix because it provides the information how the beam responds to excitation of the correctors. Even exciting only one corrector will create nonzero readings in all BPMs as clear from eq.6.3. An example of a response matrix is shown in figure 6.3 (top).

Solving the linear system, i.e. inversion of the response matrix provides a vector of corrector strengthes and thus performs the closed orbit correction in one go. However the response matrix is not square and has the tendency to be singular or close to singular, if some BPMs or correctors are linear combinations of others. The method of choice to handle that is SVD (singular value decomposition) [41]: It is a robust technique to solve ill-conditioned, over– or underdetermined linear systems. For $N > M$ it will minimize the corrector strength. For $N < M$ it will find the solution with the smallest residual orbit in a least square sense. For $N = M$ it will simply invert the matrix. Furthermore it provides insight which correctors are most or least efficient by extracting an $N \times N$ diagonal matrix of weight factors from the response matrix.

Figure 6.3 (bottom) shows the inverted response matrix for a case with $N = M$ and correctors close to BPMs: It is a tridiagonal matrix, i.e. a nonzero reading at BPM k is produced by nonzero strengthes in the 3 adjacent correctors $k-1$, k and $k+1$. – Actually this is just the closed orbit bump used in the sliding bump method with the required coefficients for contructing the bump stored in the inverted response matrix elements. Thus for a square system sliding bump and SVD methods provide the same, unique solution for the orbit correction.

6.4 Coupling

Linear coupling between horizontal and vertical motion is introduced by skew quadrupole fields from vertically misaligned sextupoles or from quadrupoles with roll errors.

Solenoid fields as used in detectors cause a rotation of the beam and thus also introduce linear coupling.

In a flat lattice coupling is the source of a finite vertical emittance which ideally would be zero. It is important not to mix the coupling parameter κ and the emittance ratio $g = \epsilon_y/\epsilon_x$.

Coupling drives quadrupolar skew betatron resonances of type $Q_x \pm Q_y = p$. In sum resonances ("+" sign) the difference between the betatron amplitudes is constant and thus they both may grow infinitely leading to beam loss. In difference resonances ("$-$" sign) the sum of betatron amplitudes is constant and the motion remains confined, drawing Lissajous-figures in phase space. Thus the lattice's working point never should be close to a sum resonance.

Under the assumptions that the coupling is small, that only skew quadrupoles contribute, and that one single difference resonance is dominant the coupling constant is given by [21]

$$\kappa = \frac{1}{4\pi} \oint_C \sqrt{\beta_x \beta_y} a_2(s) e^{i(\phi_x(s) - \phi_y(s) - \frac{2\pi s}{C}\Delta)} ds \tag{6.6}$$

with $\Delta = (Q_x - Q_y - p)$ the distance to the difference resonance and a_2 the skew quadrupole strength. When moving the working point into the resonance horizontal and vertical tunes will not cross over but the betatron oscillations will merge into a coupled oscillation with two different tunes that cannot be assigned to one of the transverse planes. The closest tune approach corresponds to the width of the resonance and is given by $|\kappa|$.

If coupling is the only source of vertical emittance the emittance ratio g is given by [21]

$$g = \frac{|\kappa|/\Delta}{(|\kappa|/\Delta)^2 + 1/2} \quad \Longrightarrow \quad \epsilon_x = \frac{1}{1+g}\epsilon_{xo} \quad \epsilon_y = \frac{g}{1+g}\epsilon_{xo} \tag{6.7}$$

with ϵ_{xo} the natural horizontal emittance of the ideal flat lattice.

If a low vertical emittance is required, for example for reasons of brightness in a light source, dedicated skew quadrupoles may be added to the lattice in order to minimize the integral of eq.6.6. Another source of vertical emittance is spurious vertical dispersion caused by vertical dipole downfeed from quadrupole heaves and thus independant from coupling.

If the lattice is not flat it will of course provide a natural vertical emittance from synchrotron radiation in vertical bending magnets which has nothing to do with coupling either.

Coupling is usually considered as a parasitic effect and suppressed. However in operation some coupling resulting in $g \approx 10^{-3} \ldots 10^{-2}$ is often appreciated in order to prolong the Touschek lifetime which is inversely proportional to the bunch volume.

Some collider concepts go even further and use dedicated so called round beam, i.e. beams with equal horizontal and vertical emittance and equal betafunctions at the interaction point, in order to obtain higher luminosity: The Novosibirsk Φ-factory was designed to operate exactly on the main difference resonance but to produce horizontal and vertical

emittances independently by rotating the beam 90° per turn by means of superconducting solenoids also used for focussing at the interaction point [3]. Another concept was to do all focusing by skew quadrupoles and obtain uncoupled motions in the 45° rotated eigenplanes [9].

For further reading on coupling see refs. [21, 57, 10].

6.5 Multipolar errors

Finite approximation of ideal magnet pole contours, saturation effects and mounting errors of a magnet's components introduce parasitic multipoles.

Parasitic multipoles are obtained from magnet simulations or measurements, they are usually expressed as relative field component at some reference radius $B_m(r)/B_n(r)$ with the magnetic fields given by eq.2.1.

Due to finite approximation of ideal pole contours odd multiples of the fundamental pole appear systematically, e.g. a quadrupole ($n = 2$) has systematic dodekapoles ($n = 6$), ikosapoles ($n = 10$) and so on. These parasitic multipoles may also change with saturation. Quadrupoles are usually manufactured in two halfs in order to mount or unmount the vacuum chamber, thus alignment errors between the two halfs cause random octupoles, varying for individual magnets.

Bending magnets have systematic sextupoles caused by finite pole width and edge sextupoles caused by a curvature of the fringe field.

Dipolar, quadrupolar and sextupolar errors are not very harmful since they may be compensated by the corrector dipoles, by retuning the quadrupoles or reoptimizing the sextupoles. Higher multipoles would require introduction of a correction system of same multipole order if there is no other way to maintain sufficient dynamic acceptance. Tracking simulations are required before ordering magnets for a machine in order to define tolerances for the parasitic multipole content. For modern machines these tolerances expressed as relative field components at aperture inscribed radius are in the 10^{-4} range and thus require great skills from the magnet manufacturers.

6.6 Ground vibrations

Nature (tides, rivers, earthquakes) and civilization (machines, traffic, quarries) produce ground waves leading to dynamic, correlated misalignments of magnets. The relevant frequency range is from 0 to a few 100 Hz. Modelling should include the ground spectrum, damping from the concrete floor of the accelerator hall, mechanical resonances from the girders and finally the closed orbit response. Ground waves are too fast for the normal, DC based closed orbit correction to follow. As a a general rule of thumb the beam jitter due to vibrations should be less than 10% of the beam radius. However at the interaction region of a collider or at the undulator or a high brightness light source this tolerable jitter is in the micron range. If it is predicted to be larger the installation of a dedicated fast orbit feedback system has to considered. Figure 6.2 shows the closed orbit response to

ground waves for the Swiss Light Source indicating the benefit of girders to suppress the beam jitter. For further reading see ref. [47].

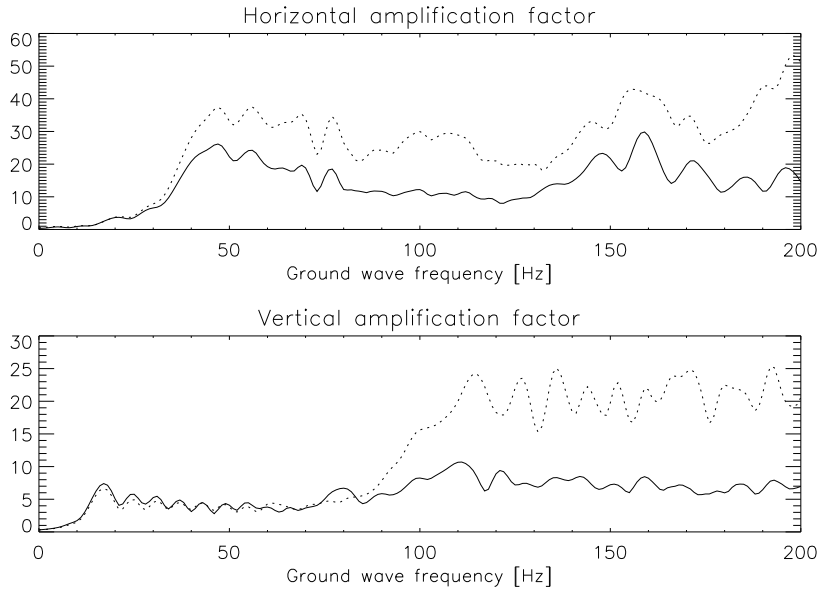

Figure 6.2: Amplification factors defined as ratio of closed orbit (averaged over ground wave incident angle and phase) to ground wave amplitude for single elements (dotted) and elements mounted on girders (solid). For high frequency the factors approach the values from random magnet misalignments. An increase of the factors is observed where λ equals the betatron wavelength, which occurs at 36 Hz for the horizontal and at 14 Hz for the vertical. Data are for the Swiss Light Source and were taken from ref.[7].

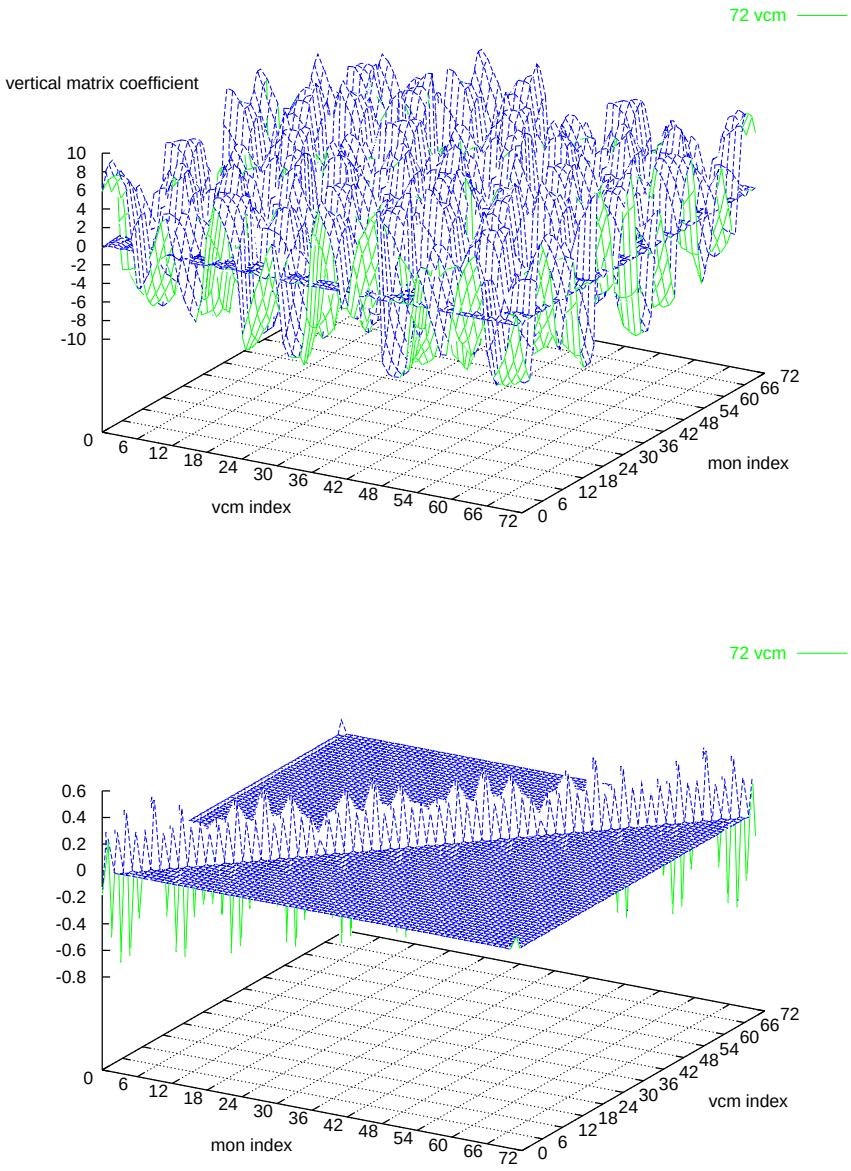

Figure 6.3: Singular value decomposition: Response matrix of BPM readings from corrector excitations (top) and its inverse calculated by singular value decomposition (bottom). Figures are taken from ref. [6].

Bibliography

[1] P. Audy, C. Leleux, A. Tkatchenko, First and second order effects of sextupolar perturbation, Internal report LNS 89-14, Laboratoire National Saturne, Saclay, 1989

[2] B.Autin (Ed.), *Beamoptics* A program for analytical beam optics, Internal report, CERN 98-06

[3] L.M.Barkov et al., Phi-factory project in Novosibirsk, Novosibirsk 1989

[4] J.Bengtsson, The Sextupole for the Swiss Light Source: An Analytic Approach, Internal Report SLS-Note 9/97, PSI, 1997

[5] M.E.Biagnini et al., DaΦNE main ring optics, in: Proc.EPAC'98, p.879

[6] M.Böge et al., Fast closed orbit control in the SLS storage ring, Internal report SLS-TME-TA-1999-0005, PSI 1999

[7] M.Böge et al., Studies on Imperfections in the SLS storage ring, Internal report SLS-TME-TA-1999-0004, PSI 1999

[8] K.L.Brown et al, Transport, Internal report SLAC-91, Rev.2, Stanford 1977

[9] J.Byrd et al.,An uncoupled round-beam electron accelerator lattice, in: Proc. of third advanced ICFA beam dynamics workshop on beam-beam effects, Novosibirsk 1989, p.87

[10] A.W.Chao, Evaluation of beam distribution parameters in an electron storage ring, J.Appl.Phys. 50.2 (1979) 595

[11] A. W. Chao, Coehrent instabilities of a relativistic bunched beam, AIP Conf. Proc. 105 (1982) 353

[12] E. D. Courant amd H. S. Snyder, The alternate gradient synchrotron, Ann. Phys. 3 (1958) 1

[13] N.C.E.Crosland et al., Recent Developments in HELIOS Compact Synchrotrons, in: Proc.EPAC'98, p.259

[14] N.S.Dikansky und D.V.Pestrikov, Colliding beams coherent instabilities, Part.Acc.12 (1982) 27

[15] D.Einfeld et al., A lattice design to reach the theoretical minimum emittance in a storage ring, in: Proc.EPAC'96, p.

[16] D.Einfeld and M.Plesko, A modified QBA optics for low emittance storage rings, NIM A 335 (1993) 402

[17] E.Forest and K.Hirata, A contemporary guid to beam dynamics, Internal report KEK 92-12, KEK August 1992

[18] L.R. Evans, LHC Accelerator Physics and Technology Challenges, in: Proc.EPAC'98, p.3

[19] C.R.Evans und J.Gareyte, Beam-beam-effects, Internal report CERN 87-03, p.159, CERN 1987

[20] J.Gareyte, priv.comm.

[21] G.Guignard, Betatron coupling with radiation, Internal report CERN-87-03, p.203

[22] HEACC'98, XVII International Conference on high energy accelerators, Dubna 1998

[23] K.Hübner, LEP-2 present and future performance and limitations, in: Proc.EPAC'98, p.38

[24] P.Ivanov et al., Luminosity and beam-beam effects on the storage ring VEPP-2M with superconducting wiggler magnet, in: Proc. of Third Advanced ICFA Beam Dyn. Workshop on Beam-Beam Effects, Novosibirsk 1989, p.26

[25] K.Johnson, Transition, Internal report CERN 85-19, p.178, CERN 1985

[26] W.Joho and L.Rivkin (PSI), priv.comm.

[27] W.Joho and A.Streun, unpublished, PSI 1993

[28] W.Joho, Radiation properties of an undulator, Internal report SLS-TME-TA-1995-0004, PSI 1995

[29] S.Y.Lee, Emittance optimization in three and multiple bend achromats, Phys.Rev.E 54.2 (1996) 1940

[30] M.P.Level et al., Status of the SOLEIL project, in: Proc. EPAC'96

[31] P.Marchand, Possible Upgrading of the SLS RF System for improving the Beam Lifetime, in: PAC-99, also Internal Report SLS-TME-TA-1999-0007, PSI 1999

[32] M.Martini, Transverse beam dynamics, Joint Universities Accelerator School 1996, Vol.3, Archamps 1996

[33] S.Milton, Calculation of how the ratio $\beta^*/\sigma_{bunchlength}$ affects the maximum luminosity obtainable, Internal report CBN 89-1, Cornell 1989

[34] G.Mülhaupt, A few design considerations for injector synchrotrons for synchrotron light sources, Internal Report ESRF/MACH-INJ/94-13, ESRF 1994

[35] J.Murphy, Synchrotron Light Source Data Booklet Version 4, Internal report BNL-42333, BNL 1996, download from
`http://www.nsls.bnl.gov/AccPhys/hlights/dbook/Dbook.Menu.html`

[36] H.Nesemann, Teilchenverlust und Lebensdauer, DESY 1970

[37] K. Oide and H. Koiso, Dynamic aperture of electron storage rings with noninterleaved sextupoles, Phys. Rev. E, 47.3 (1993) 2010

[38] OPA is the author's lattice design code developed and used for the design of the Swiss Light Source. As a DOS-program it is rather outdated but still may be useful for lattice design or for code developmet by incorporation of the algorithms into modern programs. For more information and free download go to
`http://slsbd.psi.ch/~ streun/opa/opa.html`

[39] S.Peggs, RHIC Progress Report, in: Proc.EPAC'98, p.13

[40] D. Poirier, Etude des défauts de champ magnetique de Mimas et correction de la chromaticité de Super-Aco, Thesis, Orsay 1984

[41] W.H.Press et al., Numerical recipes, Cambridge 1989

[42] Review of particle physics, Eur.Phys.J.C3 (1998) 580

[43] D.Rice (CESR), A.A.Zholents (VEPP-4), priv.comm.

[44] D.Rice, High luminosity operation at the Cornell electron storage ring, Proc. 1989 IEEE Part.Acc.Conf., p.444

[45] A.Ropert, Lifetime issues for third generation light sources, in:Proc. EPAC'98, p.62

[46] A. Ropert, Sextupole correction scheme for ESRF, Internal report CERN 88-04 (1988) 223

[47] J.Rossbach, Closed orbit distortions of periodic FODO lattices due to plane ground waves, Part.Acc.23 (1988) 121

[48] R.D.Ruth, Single particle dynamics in circularar yccelerators, AIP Conf. Proc. 153 (1987) 150

[49] M.Sands, The physics of electron-positron storage rings, Internal report SLAC-121, Stanford 1970

[50] J.Rossbach and H.Schmüser, Basic course on accelerator optics, Internal report CERN-94-01, p.17

[51] SLS design handbook, PSI

[52] K.Steffen, Basic course on accelerator optics, Internal report CERN 85-19, p.25, CERN 1985

[53] A.Streun, Momentum acceptance and Touschek lifetime, Internal report SLS-Note 17/87, PSI, 1987

[54] M.Svandrlik et al., Design of a 3rd harmonic superconducting cavity for bunch lengthening in ELETTRA, in: Proc.EPAC'98, p.1879

[55] L.C.Teng, Minimum emittance lattice for synchrotron radiation storage rings, Internal report LS-17, ANL 1985

[56] D.Trbojevic and E.Courant, Low emittance lattices for electron storage rings revisited, in: Proc. EPAC'94, p.1000

[57] F.Willeke and G.Ripken, Methods of beam optics, Internal report DESY 88-114, DESY 1988

[58] E.J.N.Wilson, Transverse beam dynamics, Internal report CERN 85-19, p.64

[59] E.J.N.Wilson, Nonlinearities and resonances, CERN 85-19 (1985) 96

[60] A.Wrulich, Single beam lifetime, CERN-94-01, p.409, CERN 1994

[61] M.S.Zisman et al., ZAP user's manual, Internal report LBL-21270, LBL 1986

Contents