2 Transverse Optics Measurement and Correction

In order to preserve the beam quality, accurate knowledge of the transverse optics and its correction is most often mandatory. For example, if the distribution of a beam injected into a storage ring is not matched to the ring optics, the emittance will grow due to filamentation. Or, if there is a significant optics error, e.g., induced by a strength error in a quadrupole magnet, the beam envelope may vary strongly. The resulting reduction in dynamic aperture may then lead to enhanced beam loss.

There are several approaches for verification of the beam optics. Some of these are based on exciting the beam using a transverse kicker magnet, others vary the strength of individual quadrupoles. In either case, the optics can be evaluated using measurements of the beam response, which is measured on one or more transverse electrodes ("pick ups").

In this chapter, we discuss several techniques used to measure the betatron tune, betatron phase advances, and beta functions. We then describe how the source of an optics error may be localized, and, alternatively, how an imperfect optics can be matched to the design values by the use of orthogonal 'multiknobs'. We next present general and powerful data processing techniques. The beam response to (large) transverse deflections is discussed next which includes a combination of effects including nonlinear detuning, radiation damping, head-tail damping, and chromaticity. The last section is devoted to measurements and correction techniques for betatron coupling.

2.1 Betatron Tune

2.1.1 Introduction

In a storage ring the betatron tune, or Q value, is defined as the number of betatron oscillation periods per revolution[1]:

$$Q = \frac{\phi(C)}{2\pi} = \frac{1}{2\pi} \oint_C \frac{\mathrm{d}s}{\beta(s)}\,, \tag{2.1}$$

where the integral is taken around the ring of circumference C. A schematic of a betatron oscillation is shown in Fig. 2.1.

[1] often, especially in the American literature, it is also denoted by ν

M. G. Minty et al., *Measurement and Control of Charged Particle Beams,*
https://doi.org/10.1007/978-3-662-08581-3_2

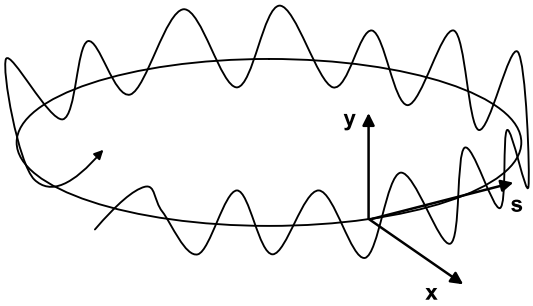

Fig. 2.1. Schematic of a betatron oscillation around a storage ring

The integer part of the tune is easily inferred from the orbit distortion induced by exciting a single steering corrector. The resulting distortion produces a coherent betatron oscillation and is easily measured by taking the difference in the orbits as measured using beam position monitors (BPMs) both with and without the kick; counting the number of oscillation periods around the ring circumference determines the integer value of the tune. A more intricate method, discussed in Sect. 2.2, involves performing a harmonic analysis of the betatron oscillations recorded by multi-turn BPMs. In this case the betatron phase advance between adjacent BPMs can be determined. The total phase advance around the ring gives the tune. If the integer part of the tune agrees with model predictions, large optics errors can be excluded. As important as the integer value of the tune is its fractional part, since the latter can have a strong effect on the beam lifetime or emittance.

Tune measurements are useful for quite a variety of applications: the tune shift with quadrupole strength gives the local beta function, the tune shift with rf frequency the chromaticity, the tune shift with current the effective transverse impedance, and the tune shift with betatron amplitude

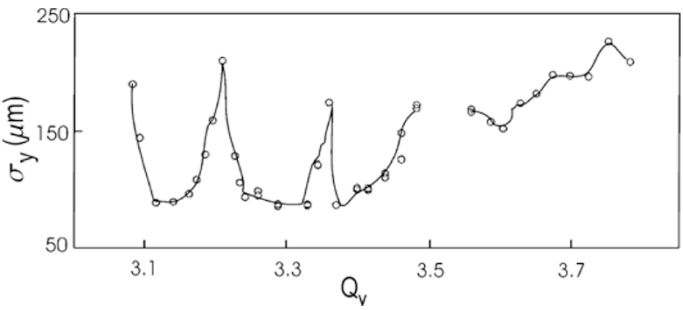

Fig. 2.2. Rms vertical size of the electron beam extracted from the SLC damping ring as a function of the vertical betatron tune. This measurement was performed under unusually poor vacuum conditions [1]

the strength of nonlinear fields. Furthermore, optimizing and controlling the tunes improves the beam lifetime and the dynamic aperture, and can reduce beam loss or emittance growth during acceleration. For example, Fig. 2.2 shows the variation of the extracted vertical beam size as a function of the vertical betatron tune which was measured at the SLC electron damping ring. Other effects including space charge, ionized gas molecules, the beam-beam interaction, and radiation damping can affect the tune signal, which may be seen for example in the shape of the beam response to a swept-frequency excitation. An example showing the dramatic effect of the nonlinear beam-beam force is shown in Fig. 2.3.

Fast decoherence and filamentation, head-tail damping or instabilities may make it difficult to extract a clean and reproducible tune signal. Conversely, the strong influence of various phenomena on the tune signal also implies that all these processes can be studied by means of tune measurements.

In the following we will describe three methods of measuring the fractional part of the tune. These approaches fall into two different categories: (1) precision tune measurements and (2) tune tracking (the latter aims at monitoring and controlling fast changes, e.g., during acceleration). For simplicity, the fractional part of the tune will also be denoted by Q.

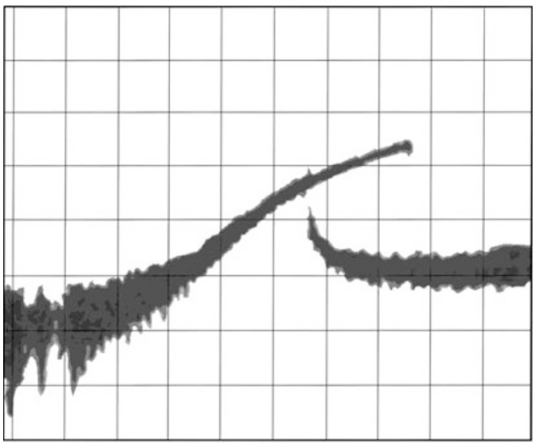

Fig. 2.3. Transverse tune measurement (swept-frequency excitation) with 2 colliding bunches at Tristan. Vertical axis: 10 dB/div., horizontal axis: 1 kHz/div. [2] (Courtesy K. Hirata and T. Ieiri, 1998)

2.1.2 Fast Fourier Transform (FFT) and Related Techniques

A common method to measure the fractional part of the betatron tune involves exciting transverse beam motion and detecting the transverse beam position over a number of successive turns N. The excitation may consist

either of white noise or of a single kick. Beam oscillations after injection are also often measured as these are often naturally present so the measurement does not interfere with accelerator operation. The power density of the detected signal is computed via a Fourier transformation and the betatron tunes are identified as the frequencies with the highest amplitude peak (this is not always the case as sometimes the beam is strongly excited at other frequencies). Figure 2.4 shows typical multi-turn BPM measurements. Alternatively, a spectrum analyzer can be used to frequency analyze the signal detected by a pick up.

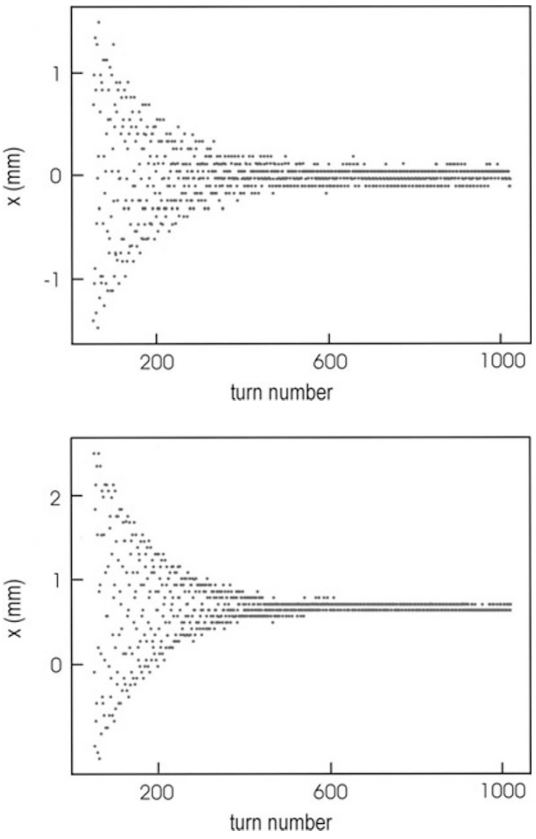

Fig. 2.4. Multi-turn orbit measurement for the motion of a single bunch in a train of 3 bunches at LEP-I. Shown are horizontal BPM orbit readings as a function of turn number: (*top*) BPM in a dispersive arc region; (*bottom*) BPM in a non-dispersive straight section. At the start of the measurement the bunch was deflected by a kicker. The corresponding FFT spectra are displayed in Fig. 2.5 (Courtesy A.-S. Müller, 2001)

A Fourier analysis uses a time series $x(1),\ x(2), ..., x(N)$ of N orbit measurements for consecutive turns as input. This time series is expanded as a linear combination of N orthonormal functions,

$$x(n) = \sum_{j=1}^{N} \psi(Q_j)\ \exp(2\pi \mathrm{i} n Q_j)\,. \tag{2.2}$$

where the coefficients $\psi(Q_j)$ are calculated from the inverse formula

$$\psi(Q_j) = \frac{1}{N}\ \sum_{n=1}^{N} x(n) \exp(-2\pi \mathrm{i} n Q_j)\,. \tag{2.3}$$

The expansion can be done efficiently with a Fast Fourier Transform algorithm. The frequency corresponding to the largest value of ψ is taken as the approximate tune (see Fig. 2.5). The frequency error is given by the finite sample size, N.

$$|\delta Q| \leq \frac{1}{2N}\,. \tag{2.4}$$

Thus, to obtain a tune value with a resolution of 0.001 or better requires orbit data for about $2^{10} \approx 1000$ turns. As an illustration, Fig. 2.5 displays FFT spectra of the orbit motion measured at the two BPMs of Fig. 2.4. The FFT confirms that a large part of the orbit motion in the dispersive region is due to synchrotron oscillations.

The resolution can be improved by an interpolated FFT. If we use a simple Fourier analysis based on the peak amplitude of ψ in (2.2), typically we need about 1000 turns of orbit data to obtain an adequate tune resolution. During this time the beam could filament or the oscillation amplitude could decrease significantly, giving rise to spurious results. Fortunately, interpolating the shape of the Fourier spectrum around the main peak improves the resolution quite dramatically [3]. Thereby the same resolution can be achieved by processing data over considerably fewer turns.

The underlying assumption is that the shape of the Fourier spectrum is known and corresponds to that of a pure sinusoidal oscillation with tune Q_{int},

$$|\psi(Q_j)| = \left| \frac{\sin N\pi(Q_{\mathrm{int}} - Q_j)}{N \sin \pi(Q_{\mathrm{int}} - Q_j)} \right|\,. \tag{2.5}$$

The formula for the interpolated tune Q_{int} reads [4]:

$$Q_{\mathrm{int}} = \frac{k}{N} + \frac{1}{\pi} \arctan \left(\frac{|\psi(Q_{k+1})| \sin\left(\frac{\pi}{N}\right)}{|\psi(Q_k)| + |\psi(Q_{k+1})| \cos\left(\frac{\pi}{N}\right)} \right)\,, \tag{2.6}$$

where $|\psi(Q_k)|$ is the peak of the Fourier spectrum in (2.2), and $|\psi(Q_{k+1})|$ its highest neighbor. In other words, instead of using only the peak value of the

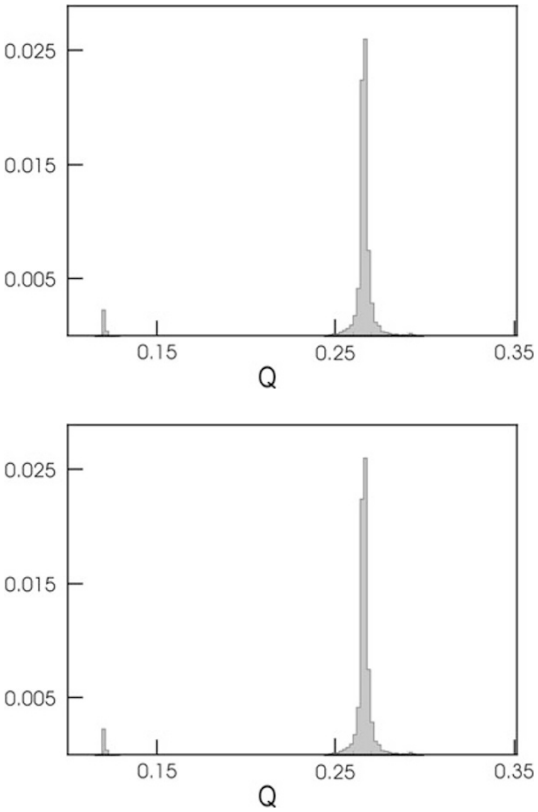

Fig. 2.5. FFT power spectra for the two BPM measurements of Fig. 2.4 in a dispersive (*top*) and in a nondispersive region (*bottom*). The horizontal scale is in tune units, the vertical scale in arbitrary units (Courtesy A.-S. Müller, 2001)

FFT, one interpolates between the two highest points. For large N the error decreases as

$$|\delta Q| \propto \frac{1}{N^2} \,. \tag{2.7}$$

So, the resolution improves quadratically with the number of turns and already from a beam signal recorded over 30–60 turns, fairly accurate tune values can be obtained. For $N \gg 1$, (2.6) may be approximated by the simpler form [3]

$$Q_{\text{int}} \approx \frac{k}{N} + \frac{1}{N} \arctan\left(\frac{|\psi(Q_{k+1})|}{|\psi(Q_k)| + |\psi(Q_{k+1})|}\right) . \tag{2.8}$$

A further refinement is an interpolated FFT with data windowing, which further increases the accuracy of the Fourier analysis [4, 5]. Here, the data

$x(n)$ are weighted with filter functions $\chi(n)$ before the interpolated FFT is applied. The Fourier coefficients of the filtered signal are

$$\psi(Q_j) = \frac{1}{N} \sum_{n=1}^{N} x(n)\chi(n) \exp(-2\pi i n Q_j) \,. \tag{2.9}$$

Note that the regular FFT, (2.3), corresponds to (2.9) without filter, or to $\chi(n) = 1$. Applying instead a Hanning-like filter of order l, $\chi_l(n) = A_l \sin^l(\pi n/N)$, where A_l is a normalization constant, in the limit $N \gg 1$ the interpolated tune reads

$$Q_{\mathrm{Han}} = \frac{k}{N} + \frac{1}{N} \left(\frac{(l+1)\psi(Q_{k+1})}{\psi(Q_k) + \psi(Q_{k+1})} - \frac{l}{2} \right) . \tag{2.10}$$

The resolution improves with the $(l+2)$th power of the number of samples N:

$$|\delta Q| \propto \frac{1}{N^{l+2}} \,. \tag{2.11}$$

An example comparing the precision of different FFT procedures is shown in Fig. 2.6 [4], which clearly demonstrates the superiority of the interpolated FFT with data windowing (Hanning filter). Unfortunately, the beneficial effect of the Hanning filter disappears when the signal contains a small noise component [4], in which case the resolution decreases as $\sim N^{-2}$ only, just as with the simple interpolated FFT. Another potential problem may arise if the signal measured in the control room is far from a purely sinusoidal shape, e.g., due to a nonzero chromaticity, filamentation, or collective effects.

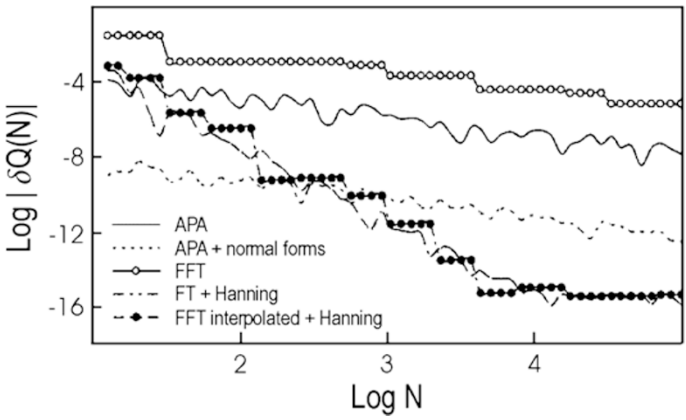

Fig. 2.6. Tune precision vs. number of turns, considering different FFT techniques applied to tracking data for the 4-D Hénon map [4]. The abbreviation 'APA' refers to a calculation of the average phase advance, which can be computed either in the original phase-space coordinates or in so-called normal-form coordinates. See [4] for more details on these alternative methods (Courtesy M. Giovannozzi, 1998)

Accurate computations of fundamental oscillation frequencies, using techniques similar to those described above, allow for a global view of the underlying beam dynamics, and in particular the identification of chaotic region of phase space. Such 'frequency map analysis' [6] has been applied to both simulation data [7, 8, 9] and measurements [10].

Another approach to obtain higher accuracy than an FFT is Lomb's method [11], which was applied at LEP with great success [12]. This method can be applied to an arbitrary data sample without any constraints on the number of data points or on the time interval between two successive points. It weights the data on a 'per point basis' and not 'per time interval' as an FFT [12].

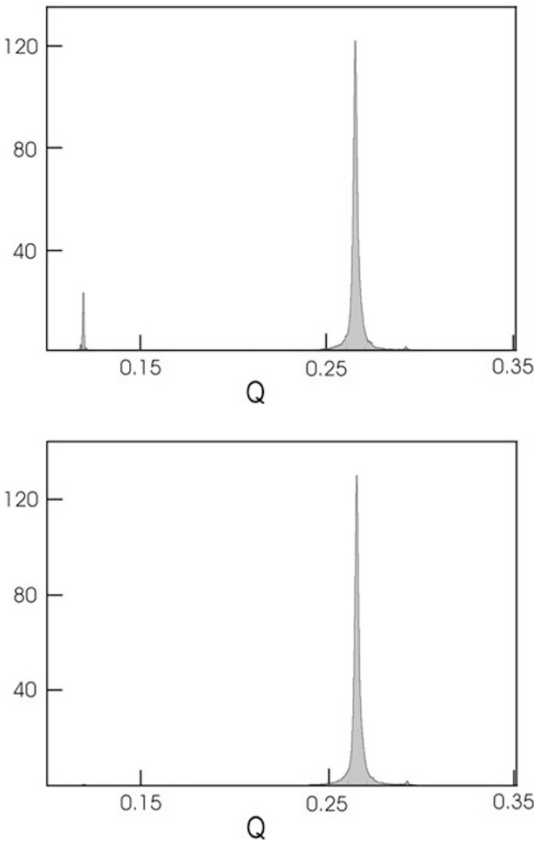

Fig. 2.7. Lomb normalized periodogram [12] for the two BPM measurements of Fig. 2.4 in a dispersive (*top*) and in a nondispersive region (*bottom*). The horizontal scale is in tune units, the vertical scale in arbitrary units (Courtesy A.-S. Müller, 2001)

For N data points h_i measured at times t_i, the so-called 'Lomb normalized periodogram' $P_N(Q)$ is defined by

$$P_N(Q) = \frac{1}{2\sigma^2}\left\{\frac{\left[\sum_n \cos(2\pi Q(n-n_0))\right]^2}{\sum_j \cos^2(2\pi Q(n-n_0))} + \frac{\left[\sum_n \sin(2\pi Q(n-n_0))\right]^2}{\sum_n \sin^2(2\pi Q(n-n_0))}\right\} \tag{2.12}$$

with

$$d_n = h_n - \frac{1}{N}\sum_{n=1}^{N} h_n \ , \tag{2.13}$$

$$\sigma^2 = \frac{1}{N-1}\sum_{n=1}^{N} d_n^2 \ , \tag{2.14}$$

$$\tan(4\pi Q n_0) = \frac{\sum_n \sin(4\pi Q n)}{\sum_n \cos(4\pi Q n)} \ . \tag{2.15}$$

For any test frequency Q the Lomb method determines the contents of that frequency in the data set. The constant n_0, which in general is not an integer, is computed so as to eliminate the phase of the original harmonic in the expression for the power spectrum $P_N(Q)$. Since the phase dependence is removed, the Lomb periodogram can be more accurate than the FFT.

As an example, Fig. 2.7 displays the Lomb normalized periodograms for the two LEP data sets of Fig. 2.4. Comparison with the corresponding FFT power spectra in Fig. 2.5 reveals the higher quality of the Lomb transformation.

2.1.3 Swept-Frequency Excitation

A different method to measure the tune involves exciting the beam with a steady sinusoidal wave and detecting the amplitude and phase of the beam response. The excitation frequency is increased in steps, or at a constant rate. The strength of the harmonic excitation is adjusted so as to produce beam oscillations of adequate amplitude at the resonant frequency.

The result of this measurement is a 'transverse beam-transfer function', which is the (complex) response of the beam to a harmonic excitation as a function of frequency. The beam-transfer function contains important information, for example, about the transverse impedance or about radiation damping [13]. It is easy to see from (2.2) that, in frequency domain, the tune signal from a single bunch repeats itself in frequency intervals corresponding to multiples of the revolution frequency f_{rev} (i.e., a spectrum analysis of the signal from one pick up contains no information about the integer part of the tune). If n_b equidistant bunches are stored in a ring and the combined signal

of all bunches is detected, the periodicity of the FFT signal is $n_b f_{\mathrm{rev}}$. In addition, the tune spectrum from 0 to $n_b f_{\mathrm{rev}}/2$ and that from $n_b f_{\mathrm{rev}}/2$ to $n_b f_{\mathrm{rev}}$ are mirror images of each other. Therefore, for the study of multibunch instabilities, it is sufficient to measure the beam transfer function around each revolution harmonic between zero and $n_b f_{\mathrm{rev}}/2$.

The concept of the beam-transfer function can be extended to higher-order beam excitations. At the CERN Antiproton Accumulator a quadrupole pick-up was used to measure the quadrupole mode beam-transfer function of an antiproton beam [14].

The frequency-sweep method as discussed so far requires a relatively long time in order to measure the response at each frequency with sufficient accuracy. However, there exists a fast version of this method, called a chirp excitation. Here the frequency of the excitation is ramped rapidly across the tune resonance, while the beam response is observed [15]. This is useful for monitoring fast tune changes as is done for example during acceleration in the SPS [16, 17].

2.1.4 Phase Locked Loop

Exactly at the betatron tune the amplitude of the beam-transfer function has zero slope as a function of excitation frequency, whereas the phase of the beam-response exhibits maximum slope. The phase difference between excitation and beam motion changes from 0 degree to 180 degree when the excitation frequency is ramped through the resonance. Directly at the betatron tune, the phase difference is 90 degrees. The phase difference can be monitored continually by a phase locked loop (PLL) circuit (see, for example, [15, 18]) and may be used in feedbacks which regulate the betatron tune [19].

The signal flow diagram of a phase locked loop is sketched in Fig. 2.8. The phase detector compares the frequency of a beam position signal, e.g., from a BPM, with the frequency of a local oscillator. The phase-detector output voltage is a measure of the frequency difference of its two input signals. After low-pass filtering and amplification, this signal is used to adjust the frequency

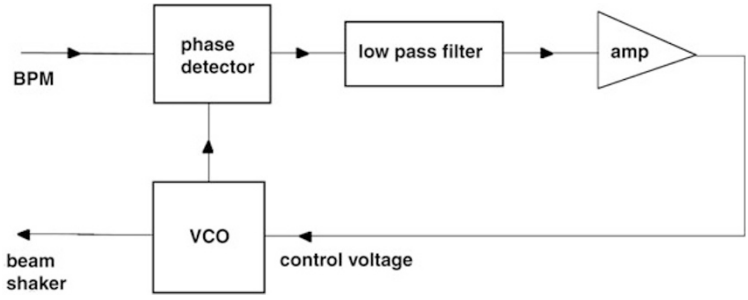

Fig. 2.8. Schematic of phase locked loop for continuous tune control

of the local oscillator (VCO), such that the oscillator 'locks' to the frequency of the input beam signal. The oscillator frequency serves as the betatron tune signal which is displayed or processed by the accelerator control system. Sometimes the oscillator signal is also used to excite the beam, in which case the phase locked loop becomes part of a 'lock-in amplifier'. PLL circuits allow a continuous tracking of the time evolution of the betatron tune.

2.1.5 Schottky Monitor

All the techniques reported so far measure the coherent betatron tune, i.e., the oscillation frequency of the beam centroid. In the case of unbunched proton, antiproton or ion beams it is also possible to measure the incoherent betatron tune, i.e., the oscillation frequency of individual particles in the beam (in the absence of centroid motion). The incoherent signal is proportional to $\sqrt{\epsilon N \Delta f}$, where ϵ is the beam emittance, N the number of particles in the beam, and Δf the frequency bandwidth. Though this signal is small, it can be detected with sensitive 'Schottky monitors' [20].

2.1.6 Multi-Bunch Spectrum

Spectrum analysis of the beam position signal reveals differences between single bunch and multi-bunch operation. For a single bunch, the spectrum analyzer displays two betatron side bands (upper and lower) at frequencies $nf_{\mathrm{rev}} \pm f_\beta$ around each revolution harmonic. There are two sidebands because negative frequencies are not displayed, and instead appear at positive value (as mirrored about the zero-frequency axis). All sidebands have equal amplitudes. Only at very high frequencies, comparable to the bunch frequency $f \sim c/(2\pi\sigma_z)$ does the amplitude decrease.

For n_b bunches uniformly distributed around the ring, the different betatron sidebands may have different amplitudes, which are determined by the strength of different possible multibunch modes. There are a total of n_b independent multibunch modes, which can all be found in the frequency range extending from 0 to $n_b f_{\mathrm{rev}}/2$. The amplitudes of the sidebands are in general unequal.

In the simplest multi-bunch case, we have two bunches diametrically opposed to each other. The position as a function of time seen by a beam position monitor is

$$x(t) \propto \sum_{n=-\infty}^{\infty} \delta(t - nT_{\mathrm{rev}})\, \mathrm{e}^{\mathrm{i}\omega_x t} + r \sum_{n=-\infty}^{\infty} \delta(t - nT_{\mathrm{rev}} - T_{\mathrm{rev}}/2)\, \mathrm{e}^{\mathrm{i}(\omega_x t + \phi)} \ , \quad (2.16)$$

where r is a coefficient which relates the oscillation amplitudes of the two bunches. We assume that the beam position $x(t)$ is obtained from two pick ups, which are located on the horizontally inward and outward sides of the

beam pipe, respectively, by taking the difference signal $(A - B)$ and normalizing it by the sum signal $(A + B)$, so that the measurement $x(t)$ does not vary with the beam current. Without such normalization, the coefficient r will also depend on the intensity ratio of the bunches.

The Fourier transform of (2.16) is

$$\tilde{x}(\omega) \propto \sum_{k=-\infty}^{\infty} \delta(\omega - \omega_x - k\omega_{\mathrm{rev}}) \left(1 + re^{\mathrm{i}(k\pi+\phi)}\right). \tag{2.17}$$

If $r = 1$ and $\phi = 0$, the two bunches oscillate in phase. This is the so-called 0 mode or σ mode. If this mode is excited, only betatron sidebands around the even revolution harmonics are visible. On the other hand if $r = 1$ and $\phi = \pi/2$, the bunches oscillate out of phase. This is the so-called π mode, related to the odd-harmonic betatron sidebands.

A multi-bunch oscillation mode can be excited and its amplitude grow, due to an interaction with the accelerator impedance at the mode frequency. The accelerator impedance entails, e.g., the resistive chamber wall, or cavity-like objects in the beam pipe. Usually, the dominant impedance has a narrow frequency band width, and, therefore, in multi-bunch operation with n_b bunches only a few of the n_b multi-bunch modes can be driven by the impedance.

2.2 Betatron Phase

2.2.1 Harmonic Analysis of Orbit Oscillations

By exciting transverse oscillations, sampling the beam position over N turns, and performing a harmonic analysis, one can determine the betatron phase at the location of the pick up as follows [21].

The oscillation detected by the BPM is a harmonic function

$$x_{km} = A_k \cos(2\pi Q_x m + \phi_k), \tag{2.18}$$

where the index k specifies the BPM, m is the turn number, and A_k the measured amplitude, which depends on the local beta function, on the magnitude of the oscillation, and on the BPM calibration. Here ϕ_k is the measured phase of the oscillation, which may be aliased (using a single BPM the integer part of the tune is undetermined).

In the limit of large N, the two Fourier sums

$$C_k = \sum_{m=1}^{N} x_{km} \cos(2\pi m Q_x), \qquad S_k = \sum_{m=1}^{N} x_{km} \sin(2\pi m Q_x), \tag{2.19}$$

approach the asymptotic values

$$C_k \approx \frac{A_k N}{2} \cos\phi_k \,, \qquad S_k \approx \frac{A_k N}{2} \sin\phi_k \,. \tag{2.20}$$

The betatron phase at the kth monitor can therefore be expressed as

$$\phi_k \approx \tan^{-1}\left(\frac{S_k}{C_k}\right) . \tag{2.21}$$

The amplitude is given by $A_k \approx 2\sqrt{C_k^2 + S_k^2}/N$. Figure 2.9 shows 5 consecutive measurements of the betatron phase advance around the PEP-II HER using (2.21). The phase advance predicted by the model was subtracted from the measured phase. The figure demonstrates that the measurement is highly reproducible, and that, for this example, it is in good agreement with the model. The offset of about 40° is due to different reference points in model and measurement.

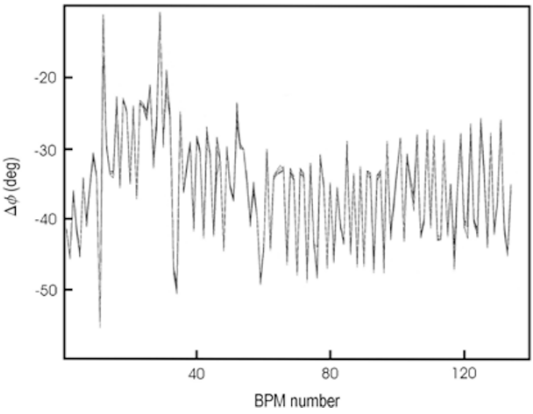

Fig. 2.9. Difference between measured and predicted betatron phase advance (in degrees) as a function of position around the PEP-II HER (BPM number) for 5 consecutive measurements; the 5 curves are superimposed. The total phase advance around the ring is about 9000° (Courtesy M. Donald, 1998)

Application: Transverse Impedance Measurement. Measuring the betatron phase advance for different bunch currents provides information about the effective transverse impedance, a quantity which describes the electromagnetic coupling of the beam to its environment [22]. A measurement of the current-dependent phase advance around the LEP ring is shown in Fig. 2.10. Clearly visible as step changes are the locations of the rf cavities in the straight sections.

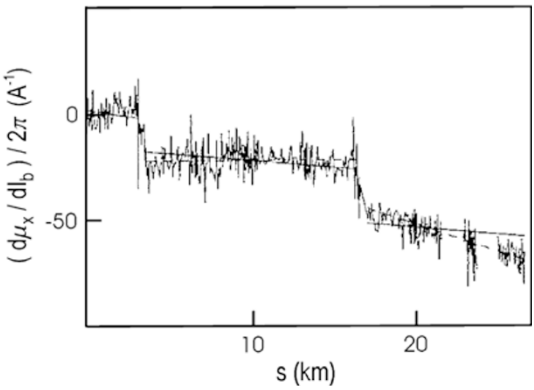

Fig. 2.10. Dependence of the horizontal betatron phase advance on the bunch current, $d\phi/dI_b/(2\pi)$ in units of A^{-1}, measured at LEP [23] (Courtesy A. Hofmann, 1998)

2.3 Beta Function

2.3.1 Tune Shift Induced by Quadrupole Excitation

Presumably the simplest beta-function measurement is to detect the shift in the betatron tune as the strength of an individual quadrupole magnet is varied. This shift can be computed using the 'thin-lens' approximation of (1.10). The tune shift induced by a gradient change for a long quadrupole can then be obtained by linear superposition.

A focusing quadrupole exerts a restoring force on the particle trajectory. For example, as discussed in Chap. 1, the linear equation for the horizontal particle motion inside a quadrupole of strength k is

$$\frac{\mathrm{d}^2x}{\mathrm{d}s^2} = -kx \;, \tag{2.22}$$

where

$$k = \frac{B_T}{(B\rho)a} \tag{2.23}$$

is measured in units of m^{-2}, and B_T denotes the pole-tip field, a the pole-tip radius, and $(B\rho)$ the magnetic rigidity of the beam: $B\rho[\mathrm{Tm}] \approx 3.356p[\mathrm{GeV/c}]$. We already mentioned in Chap. 1 that the effect of the quadrupole can often be integrated over its length, and represented as a deflection or 'kick':

$$\Delta x' = -Kx \;. \tag{2.24}$$

The integrated quadrupole strength K, in units of m^{-1}, is obtained by multiplying k with the length l_Q of the quadrupole.

With a quadrupole excitation of ΔK, the 2×2 transport matrix for the entire ring is the product of the original transport matrix, (1.12), with $f = i$,

$$\begin{pmatrix} \cos(2\pi Q_{x,y}) + \alpha_{x,y}\sin(2\pi Q_{x,y}) & \beta_{x,y}\sin(2\pi Q_{x,y}) \\ -\sin(2\pi Q_{x,y}) & \cos(2\pi Q_{x,y}) - \alpha_{x,y}\sin(2\pi Q_{x,y}) \end{pmatrix}, \quad (2.25)$$

and a perturbation matrix representing the effect of the change in gradient,

$$\begin{pmatrix} 1 & 0 \\ -(\pm\Delta K) & 1 \end{pmatrix}, \quad (2.26)$$

where Q_x is the original tune, $\beta_{x,y}$ and $\alpha_{x,y}$ the optical functions at the quadrupole, and the plus or minus sign refers to the horizontal and vertical plane, respectively. The function $\beta_{x,y}$ is to be determined.

The trace of the product matrix must be equal to $2\cos(2\pi\bar{Q}_{x,y})$, where $\bar{Q}_{x,y} = (Q_{x,y} + \Delta Q_{x,y})$ is the new tune, and $\Delta Q_{x,y}$ is the tune shift induced by a quadrupole excitation of ΔK. Explicit evaluation of the trace gives the equation

$$2\cos(2\pi(Q_{x,y} + \Delta Q_{x,y})) = 2\cos(2\pi Q_{x,y}) - \beta_{x,y}(\pm\Delta K)\ \sin(2\pi Q_{x,y})\,. \quad (2.27)$$

Solving for $\beta_{x,y}$ we find [24, 25]:

$$\beta_{x,y} = \pm\ \frac{2}{\Delta K}\ \left[\cot(2\pi Q_{x,y})\left\{1 - \cos(2\pi\Delta Q_{x,y})\right\} + \sin(2\pi\Delta Q_{x,y})\right], \quad (2.28)$$

where the $\pm$ sign refers to the horizontal and vertical planes, respectively. For a small tune change, (i.e., $2\pi\Delta Q_{x,y} \ll 1$), and far from the integer or half integer resonance (i.e., $\cot(2\pi Q_{x,y}) \leq 1$), we can further simplify and obtain

$$\beta_{x,y} \approx \pm 4\pi\ \frac{\Delta Q_{x,y}}{\Delta K}\,, \quad (2.29)$$

which is commonly applied.

Figure 2.11 illustrates the error involved in approximating (2.28) by (2.29). The difference between the two expressions becomes important if $Q_{x,y}$ is close to an integer or half integer resonance, and for large changes ΔK [24].

A recent beta function measurement at the Fermilab Recycler is depicted in Fig. 2.12. The nonlinear dependence is well described by the complete (2.28).

Care has to be taken that the applied change in quadrupole strength does not alter the beam orbit, which happens if the beam is off-center in the quadrupole whose strength is varied. If the orbit changes, part of the measured tune shift could then be caused by the closed-orbit variation at the sextupole magnets elsewhere in the accelerator. If a strong effect on the orbit is observed, the orbit should first be corrected with the help of steering correctors before the new (shifted) tune value is measured. Sometimes, several magnets are connected to the same power supply, and then the strengths K_i $(i = 1, ..., m)$ of m quadrupoles must be changed simultaneously, all by the same amount ΔK. The above result is easily generalized to this case: the induced tune change is related to the average beta function at the m quadrupoles via $\langle\beta_{x,y}\rangle_m \approx \pm 4\pi\ \Delta Q_{x,y}/(m\Delta K)$. However, the disadvantage of averaging over several quadrupoles is that beta beating may be less evident.

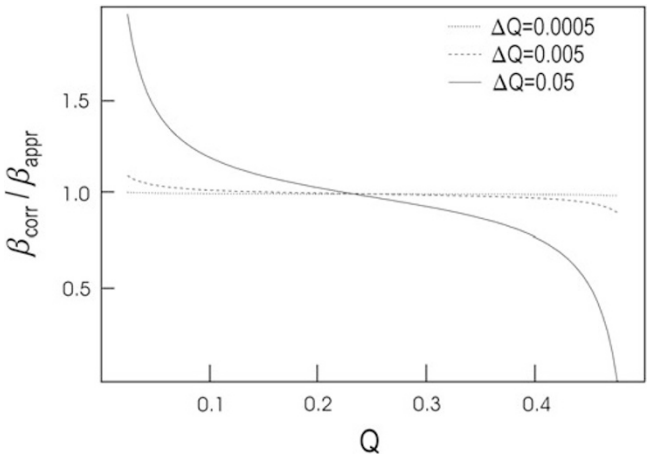

Fig. 2.11. The ratio $\beta_{\rm corr}/\beta_{\rm appr}$ of the correct beta function $\beta_{\rm corr}$, inferred from (2.28), to the approximation $\beta_{\rm appr}$ given in (2.29) as a function of the nominal tune Q. The three curves correspond to different magnitudes of ΔQ

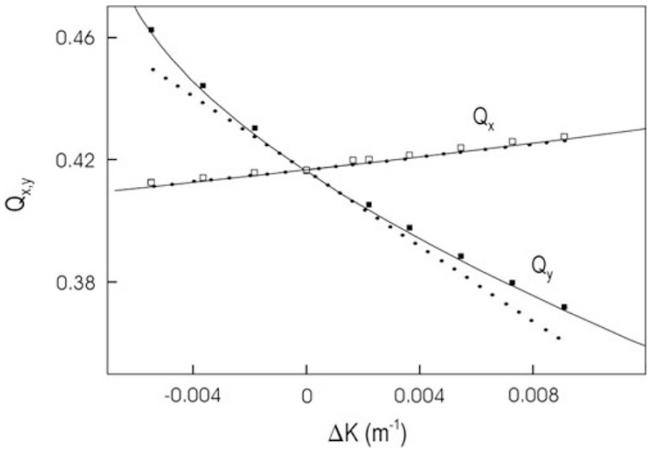

Fig. 2.12. Betatron tunes in the Fermilab Recycler Ring are plotted versus the strength of quadrupole QT601 [26]. The measurements are compared with theoretical predictions using the exact nonlinear dependence of (2.28) [*solid lines*], or its linear approximation (2.29) [*dotted lines*], and taking the beta functions at quadrupole QT601 to be equal to their design values

2.3.2 Betatron Phase Advance

A different method determines the beta function from betatron oscillations measured by multi-turn beam position monitors (BPMs). The underlying idea is to calculate the beta function from the betatron phase advance between three adjacent BPMs. The betatron phase at each BPM can be obtained with

a high precision using (2.19) and (2.21) [27]. Since the oscillation amplitude may be subject to calibration errors, it is not taken as an input to this calculation. Instead, the computed beta functions can be used to check and correct the BPM calibration.

The first row of the matrix $\mathbf{R}_{fi}$ in (1.25) can be rewritten as

$$\tan\phi_{fi} = \frac{R_{12}}{R_{11}\beta(s_i) - R_{12}\alpha(s_i)}\,, \tag{2.30}$$

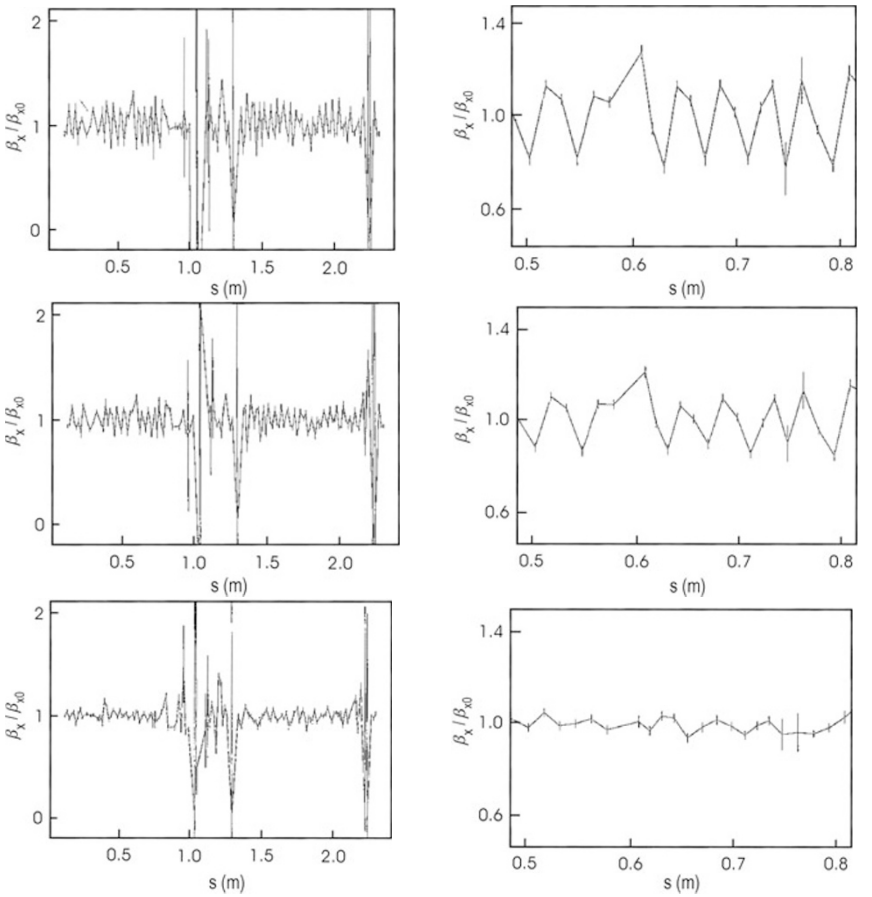

Fig. 2.13. Ratio of the horizontal beta function inferred from phase advance measurements to the model beta function: (*left column*) for the entire PEP-II High Energy Ring (HER); (*right column*) for a limited region only; (*top row*) with all magnets at nominal strength; (*middle row*) for an increased strength of a single quadrupole pair (called QF5) by 0.1%; (*bottom row*) for a strength increase of 0.15%. The "fliers" evidencing large error bars correspond to either malfunctioning BPMs or to a phase advance between successive BPMs equal to 0, $\pi/2$ or π (Courtesy M. Donald, 1998)

where ϕ_{fi} is the phase advance from monitor i to monitor f, and the R_{kl} are transport matrix elements between the same two locations. These matrix elements can be calculated from the geometry of the beam line assuming that the quadrupole magnets located between the BPMs are at their nominal strength. For a set of three BPMs, there are two independent equations of the form (2.30), which we can solve for the two unknowns α and β [27].

To be more explicit, let us denote the transport matrix from BPM 1 to 2 by $\mathbf{M}$ and the matrix from BPM 1 to 3 by $\mathbf{N}$:

$$\mathbf{M}(1 \to 2) = \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix} \qquad \mathbf{N}(1 \to 3) = \begin{pmatrix} n_{11} & n_{12} \\ n_{21} & n_{22} \end{pmatrix} \tag{2.31}$$

and denote the phase advance from BPM 1 to 2 and that from 1 to 3 by ϕ_{21} and ϕ_{31}, respectively. Applying (2.30) twice, we arrive at two expressions for the values of α and β at the first BPM [27]:

$$\beta(s_1) = \left(\frac{1}{\tan\phi_{21}} - \frac{1}{\tan\phi_{31}} \right) \Big/ \left(\frac{m_{11}}{m_{12}} - \frac{n_{11}}{n_{12}} \right) \tag{2.32}$$

$$\alpha(s_1) = \left(\frac{n_{11}}{n_{12}\tan\phi_{21}} - \frac{m_{11}}{m_{12}\tan\phi_{31}} \right) \Big/ \left(\frac{m_{11}}{m_{12}} - \frac{n_{11}}{n_{12}} \right) . \tag{2.33}$$

Figure 2.13 gives an example of beta functions obtained by this method, and it also shows the optics correction achieved by changing the strength of a quadrupole pair. The method encounters numerical degeneracies if the phase advance between adjacent BPMs is $\frac{\pi}{2}$, since then $\phi_{31} = \pi$ and $\tan\phi_{31} = 0$.

2.3.3 Orbit Change at a Steering Corrector

A simple method to measure the local beta function at a steering corrector magnet next to a BPM consists of exciting the corrector and detecting the orbit change at that BPM [28].

The formula for the closed-orbit distortion Δx_{co} of a relativistic beam induced by a single dipole kick (measured by measuring the difference in beam orbits obtained with and without the kick) is

$$\Delta x_{\mathrm{co}}(s) = \Delta\theta \, \frac{\sqrt{\beta(s)\beta(s_0)} \, \cos(|\phi(s) - \phi(s_0)| - \pi Q)}{2 \, \sin\pi Q} + \Delta\theta \, \frac{D_x(s) D_x(s_0)}{(\alpha_c - 1/\gamma^2) C} \tag{2.34}$$

where s is the location of the BPM, C the ring circumference, α_c the momentum compaction factor defined as the relative change in circumference per relative momentum deviation or $\alpha_c \equiv (\Delta C/C)/(\Delta p/p)$, and s_0 the location where the kick ($\Delta\theta$) is applied. The last term is a small correction reflecting the change in beam energy induced by a kick at a dispersive location, for a

bunched beam and constant rf frequency. If the locations s and s_0 are the same, and if we ignore the small correction due to the energy change, the formula simplifies, and the beta function at the BPM-corrector pair can be obtained from

$$\beta_{\mathrm{BPM/cor}} \approx 2 \ \tan \pi Q \ \frac{\Delta x_{\mathrm{co}}}{\Delta \theta} \ . \tag{2.35}$$

2.3.4 Global Orbit Distortions

One can also infer the beta functions at all BPMs simultaneously by measuring the average (static) orbit response to two, or more, steering correctors [29]. We call x_{ia} and x_{ib} the orbit change measured at the ith BPM when deflections θ_a or θ_b are applied at the correctors a or b. The beta function is computed from the relation [29]

$$\beta_i = \frac{4 \sin^2 \pi Q}{\sin^2 \Delta} \left(\frac{x_{ia}^2}{\beta_a \theta_a^2} + \frac{x_{ib}^2}{\beta_b \theta_b^2} - \frac{2 x_{ia} x_{ib} \cos \Delta}{\sqrt{\beta_a \beta_b} \theta_a \theta_b} \right) , \tag{2.36}$$

where $\Delta = |\phi_a - \phi_b|$ is the phase advance between the two correctors, which should not be a multiple of π. Prior to applying this equation, the three quantities Δ, θ_a and θ_b are determined by fitting a few BPM readings $x_{ja,b}$ in the vicinity of the correctors to the model optics. The computed beta functions can be verified by exciting other corrector pairs in different sections of the ring and comparing the results. Figure 2.14 shows an example from the KEK B factory, where this method was developed.

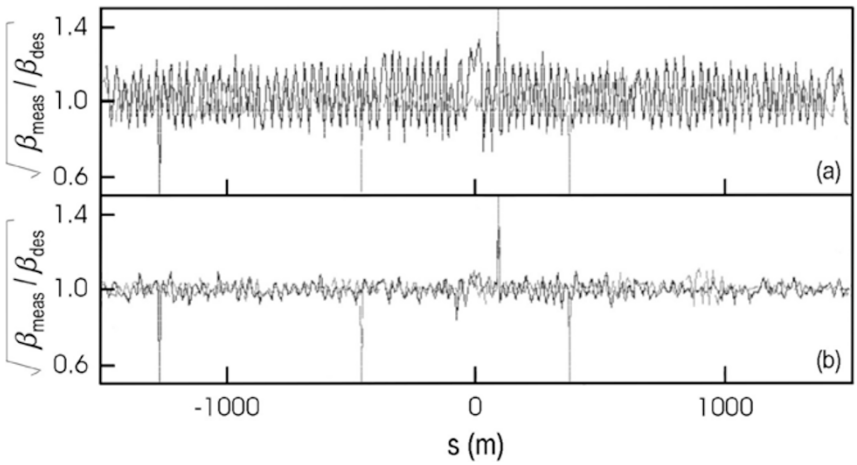

Fig. 2.14. Beta function measurement at KEKB, based on (2.36). Ratio of measured β function to the design value is shown before (*top*) and after optics correction (*bottom*) (Courtesy H. Koiso, 2000)

Fitting difference orbits to a betatron oscillation is a simpler variant, which has been used at many accelerators. An example from the heavy-ion collider RHIC is shown in Fig. 2.15.

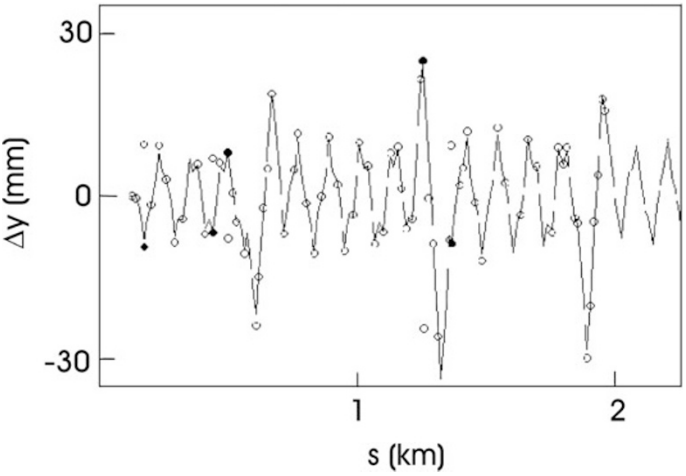

Fig. 2.15. Measured difference orbit in RHIC (*circles*) together with a fit using the online model (*dashed line*) [30] (Courtesy S. Peggs, 2000)

2.3.5 β^* at Interaction or Symmetry Point

To determine the beta function at the interaction point of a collider ring, or at any other symmetry point, e.g., in a light source), one can excite a pair of symmetrically placed quadrupoles, by an amount $\pm\Delta K$ (asymmetric excitation), where K is the integrated quadrupole gradient in units of m^{-1}. From (2.29), the total tune shift is given by

$$\Delta Q_{\mathrm{tot}} = \Delta Q_+ - \Delta Q_- \approx \frac{\Delta K}{4\pi} \left[\langle\beta_+\rangle - \langle\beta_-\rangle\right] , \tag{2.37}$$

where $\langle ... \rangle$ indicates the average over the effective length of the quadrupole and the $\pm$ sign refers to the left or right quadrupole. The advantage of the asymmetric excitation of two quadrupoles is that if the phase advance between the two quadrupoles is about 180 degrees, which is usually the case, almost no beta-beating is induced around the accelerator. In addition, if the optics is perfect and the beam waist is centered at the collision (or symmetry) point, the beta functions at the two quadrupoles are the same and, to first order, there is no net tune change: $\Delta Q_+ - \Delta Q_- = 0$.

The beta function at the collision (symmetry) point β^* is a quadratic function of the ratio

$$\eta = \langle\beta_+\rangle - \langle\beta_-\rangle = 4\pi \, \frac{\Delta Q_+ - \Delta Q_-}{\Delta K} , \tag{2.38}$$

which takes the form [31]

$$\beta^* = \beta^*_{\text{design}} \left(1 + a_{\text{optics}} \eta^2\right) , \tag{2.39}$$

where β^*_{design} is the nominal interaction-point beta function. The coefficient a_{optics} depends on the optics between the quadrupoles which are being varied and the interaction point and can be calculated with using an optics program, e.g., MAD [32]. For the LEP low-β insertions, $a_{\text{optics}} \approx 1/15$ [31]. The optics is optimally adjusted when $\Delta Q_{\text{tot}} = 0$.

2.3.6 R Matrix from Trajectory Fit

Consider a set of three BPMs, which are not a multiple of π apart in betatron phase and with nonzero dispersion for at least one. The horizontal orbit readings at these three BPMs, which we denote by $x^{(1)}$, $x^{(2)}$, and $x^{(3)}$, then contain complete information about the betatron motion (x and x') and the energy offset (δ) of each trajectory. This means we can express the orbit at every other BPM as a linear combination of the orbit reading for these three BPMs:

$$x(s) = B(s)x^{(1)} + C(s)x^{(2)} + D(s)x^{(3)} \tag{2.40}$$

If the three BPMs are adjacent and the optics between them is known, (2.40) is equivalent to the more familiar form given in terms of the transport matrix elements as

$$x(s) = R_{11}^{s_0 \to s} x(s_0) + R_{12}^{s_0 \to s} x'(s_0) + R_{16}^{s_0 \to s} \delta, \tag{2.41}$$

because then the three variables $x(s_0)$, $x'(s_0)$ and δ are known linear combinations of $x^{(1)}$, $x^{(2)}$ and $x^{(3)}$.

If multiple data sets for many different turns (in a ring) or for many bunch passages (in a transport line) are acquired, the coefficients $B(s)$, $C(s)$, and $D(s)$, or equivalently the R matrix elements R_{11}, R_{12} and R_{16} may be obtained by a fitting procedure. However, care has to be taken: a simple least squares fit may produce spurious results if the BPM readings on the left and right side of (2.41) have noise components.

The effect of the noise in the horizontal coordinates can be illustrated by a simple example, taken from [33]. Consider a linear fit of the form $y = px+q$, where p and q are to be determined and both x and y are smeared stochastically. Figure 2.16 shows the reconstructed slope normalized to the true slope as a function of the signal-to-noise ratio in the horizontal coordinate, R_x. Even for a signal-to-noise ratio of 3 the fitted slope still has a 10% error. This result is independent of the noise in the y coordinate.

A better approach, which takes into account the noise in the horizontal coordinates, is to 'find the principal axes of the set of data points and then turn the parameter vector parallel to the principal axis along which the data points fluctuate the least' [33]. The general problem and its solution are as

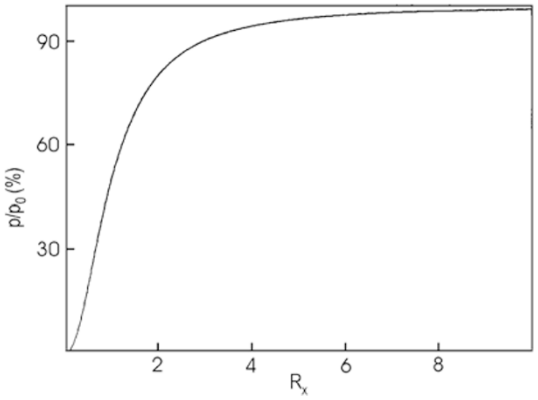

Fig. 2.16. Systematic slope error introduced in a linear χ^2 fit of the form $y = px+q$, neglecting the noise smearing in the x measurement [33]. Shown is the reconstructed slope normalized to the true slope as function of signal-to-noise ratio R_x in the horizontal coordinate (Courtesy P. Emma, 1998)

follows [33]. Let x_n be a measured variable which is linearly correlated with $(n-1)$ other measurements $x_1, ..., x_{n-1}$, and suppose there are a total of N data sets. Introducing normalized coordinates,

$$z_i = \frac{x_i - \langle x_i \rangle}{\sigma_i}, \tag{2.42}$$

the fit equation (2.40) or (2.41) may be reexpressed as

$$\mathbf{u}^T \cdot \mathbf{z} = 0\,. \tag{2.43}$$

Introducing the symmetric covariance matrix

$$C_{ij} = \sum_{l=1}^{N} z_{li} z_{lj}, \tag{2.44}$$

Equation (2.43) may then be solved in a least squares sense by

$$\mathbf{C} \cdot \mathbf{u} = \lambda u\,, \tag{2.45}$$

$$|\mathbf{u}|^2 = 1\,, \tag{2.46}$$

$$\chi^2 = \lambda\,, \tag{2.47}$$

where the solution $\mathbf{u}$ is simply the normalized eigenvector corresponding to the smallest eigenvalue λ of $\mathbf{C}$, where λ gives also the χ^2 of the fit.

If we assume that $\lambda = \lambda_1$ (the smallest eigenvalue) is not degenerate, and consider a scalar function $f(\mathbf{u})$ of the fit parameters $\mathbf{u}$, the rms fit error in f is given by

$$\sigma(f)^2 = (\nabla_{\mathbf{u}} f)^T \cdot \mathbf{T} \cdot (\nabla_{\mathbf{u}} f)\,, \tag{2.48}$$

where $\mathbf{T}$ is a symmetric $n \times n$ matrix defined by

$$T_{ij} = \sum_{r=2}^{n} \frac{\lambda_r + \lambda}{(\lambda_r - \lambda)^2} \, (u_r)_i (u_r)_j \, . \tag{2.49}$$

In particular, the rms error of the coefficient u_i in the normalized equation (2.43) is $\sigma(u_i) = \sqrt{T_{ii}}$.

The reconstruction of lattice parameters from orbit and energy fluctuations can be studied by computer simulations. Figure 2.17 presents simulation results for the SLC final focus [33] with an assumed BPM resolution of 20 μm, employing both a standard χ^2 fit and a principal axes transformation. The results of the former differ strongly from the underlying model parameters despite the good fits and small error bars. The improvement using the principal axes method shows that the optics are reconstructed almost perfectly.

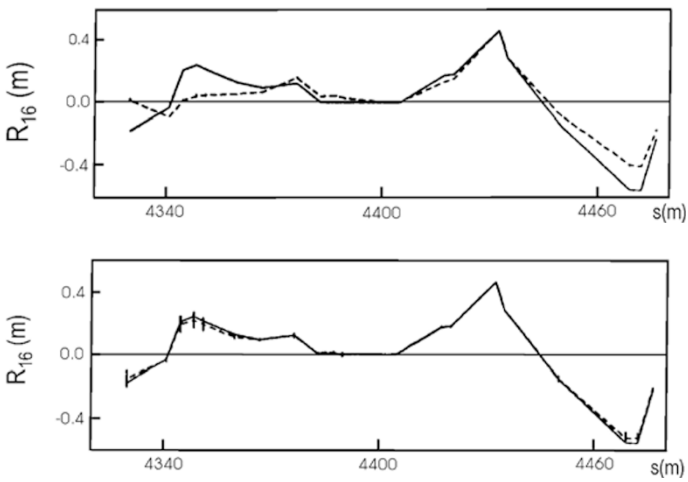

Fig. 2.17. Reconstructed R matrix element R_{16} in the SLC final focus from a sample of 100 simulated trajectories with fluctuations in betatron orbit and energy assuming a 20 μm BPM resolution [33]: (*top*) standard χ^2 fit; (*bottom*) principal axes transformation. The fit results (*dashed*) are compared to the model used for the trajectory generation (*solid*) (Courtesy P. Emma, 1998)

A Fitting Example. As a much simpler illustration, let us consider the single linear equation

$$y = px \, , \tag{2.50}$$

and assume that N pairs of data (x_i, y_i) are taken. The slope p is to be determined from a fit. A standard fit minimizes the χ^2 defined as

$$\chi^2 = \frac{1}{N} \sum_{i=1}^{N} (y_i - px_i)^2 \, . \tag{2.51}$$

Setting the derivative $\mathrm{d}\chi^2/\mathrm{d}p$ equal to zero, one obtains

$$\sum_i x_i y_i = \sum_i x_i^2 p = 0 \tag{2.52}$$

or

$$p = \frac{\sum_i x_i y_i}{\sum_i x_i^2} \,. \tag{2.53}$$

In the alternative 'eigenfit' method [33] we instead minimize

$$\chi^2 = \frac{1}{N} \sum_{i=1}^{N} (u_1 y_i - u_2 x_i)^2 + \lambda(1 - u_1^2 - u_2^2) \,, \tag{2.54}$$

where λ is a Lagrange multiplier. Now, three derivatives are set to zero: $\mathrm{d}\chi^2/\mathrm{d}u_1 = \mathrm{d}\chi^2/\mathrm{d}u_2 = \mathrm{d}\chi^2/\mathrm{d}\lambda = 0$. The last equation constrains the magnitude of the vector $\boldsymbol{u} = (u_1, u_2)$: $u_1^2 + u_2^2 = 1$. The others can be written as a system of two linear equations:

$$\begin{pmatrix} \sum_i y_i^2 & \sum x_i y_i \\ \sum_i x_i y_i & \sum x_i^2 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \lambda \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \,. \tag{2.55}$$

Suppose now that the real slope is $p = 1$, that the real x assumes the values $+1$ and -1 with equal probability, and that the data points for x and y have a random measurement error of ± 1. As an example, this is fulfilled for an 8-component real vector $\boldsymbol{x}$ equal to

$$\boldsymbol{x} = (-1, -1, -1, -1, +1, +1, +1, +1) \,, \tag{2.56}$$

a measured vector $\boldsymbol{x}_{\mathrm{meas}}$

$$\boldsymbol{x}_{\mathrm{meas}} = (-2, -2, 0, 0, 0, 0, +2, +2) \,, \tag{2.57}$$

and a measured vector $\boldsymbol{y}_{\mathrm{meas}}$

$$\boldsymbol{y}_{\mathrm{meas}} = (-2, 0, -2, 0, 0, +2, 0, +2) \,. \tag{2.58}$$

The relevant sums over the measured values are $\sum_i x_i^2 = 16$, $\sum_i y_i^2 = 16$, and $\sum_i x_i y_i = 8$. Hence, the standard fit (2.53) yields a slope $p = 1/2$, which is off by a factor of two. The eigenfit yields the characteristic equation $(16 - \lambda)^2 - 64 = 0$. Choosing the smallest solution $\lambda = 8$, one finds $u_2 = -u_1$, from which we can deduce that the slope is 1. This is the correct solution.

2.4 Detection of Quadrupole Gradient Errors

Once the beta functions have been measured and a significant difference from the model has been found, the source of the discrepancy must be determined. In most cases, the difference from the model beta function will be a beta beat

(an oscillation of the measured beta function around the design beta function at twice the betatron frequency) and the source will be a gradient error in one (or more) of the quadrupole magnets.

A gradient error ΔK (in units of m^{-1}) at location s_0 will result in a beta beat of the form

$$\Delta\beta(s) = \frac{\beta(s)\beta(s_0)}{2\sin(2\pi Q)}\ \Delta K(s_0)\ \cos(2|\phi(s) - \phi(s_0)| - 2\pi Q)\,. \tag{2.59}$$

2.4.1 First Turn Trajectories

In a storage ring, a first attempt to find the error may consist of exciting steering correctors (or changing the amplitude of the injection kicker) and fitting first-turn difference trajectories to an on-line or off-line optics model. The difference of two trajectories measured for different injection amplitudes should match a betatron oscillation as predicted by the model. Using corrector excitations, this method applies also to gradient error detection in linear accelerators and transport lines.

The initial conditions can be determined by fitting the difference orbit, using only a few BPMs, to the model. The oscillation so obtained is then propagated along the beam line. It will agree with the measured difference trajectory until it passes the location of the gradient error, at which point the propagated betatron oscillation and the measurement will start to disagree. The location of the gradient error thus identified can be confirmed by back-propagating the solution obtained with an independent set of BPMS further downstream. If the previous hypothesis is correct, the backward fit should begin to deviate from the model at the same point as the forward fit.

In principle, by analyzing first-turn orbits gross optics errors are easily identified. In practice, it is not always so simple, as beam loss, BPM spray (from lost particles), or kicker noise may corrupt the BPM measurements on the first few turns.

2.4.2 Closed-Orbit Distortion

A variant of this method is to make use of the fact that, except for the location of the corrector, a closed-orbit distortion for a stored beam has exactly the same pattern as a betatron oscillation. Thus, in much the same manner as for the first turn, the model can be used to fit the change in the closed orbit (with and without corrector excitation) to a betatron oscillation and then to propagate this oscillation around the ring. Again, the location where a noticable disagreement starts identifies the magnet with a gradient error. The excitation of this magnet can be changed, and the measurement repeated, until the agreement with the model is satisfactory. Figure 2.18 shows an example of this method from the PEP-II commissioning. A gradient error close to the interaction point was clearly identified.

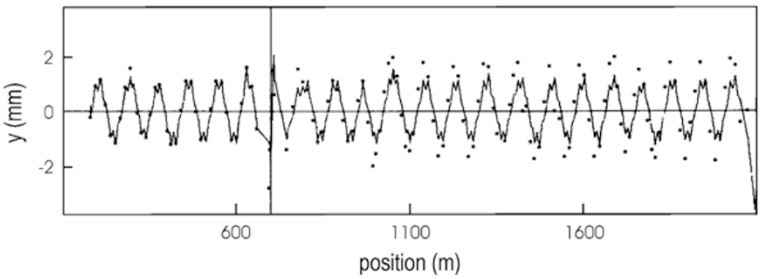

Fig. 2.18. Finding quadrupole gradient errors by fitting betatron oscillations to closed-orbit distortions: an example from the PEP-II HER commissioning using the codes LEGO [34] and RESOLVE [35]. The induced orbit change is fitted to a betatron oscillation over a small number of BPMs (*towards the left*). The betatron oscillation so obtained was propagated along the beam line using the model optics (*solid line*) and compared with the actual orbit variation (*symbols*). In this example, the measurements and fit agree well up to a region close to the interaction point, near $s = 700$ m on the horizontal axis. It was later verified that two quadrupole pairs in this region had gradient errors on the order of 0.1% (Courtesy Y. Cai, 1998)

It is possible to considerably extend this simple closed-orbit distortion scheme. For example, the response of all BPMs to every single steering corrector may be combined into a big matrix, which can be used as an input to a sophisticated statistical fitting program, such as LOCO [36, 37]. LOCO then varies the individual gradients of the quadrupoles in a computer model, e.g., using MAD [32], to find the modified quadrupole gradients that best reproduce the measured orbit response data. The advantage of using multiple data sets and multiple fits lies therein that the numerical solutions are overconstrained (which reduces the influence of systematic errors) and that multiple error sources, if present, can be more accurately ascertained.

2.4.3 Phase Advance

Instead of fitting trajectories, one can also use (2.32) to compute the beta functions from the measured phase advance around the ring. Then either the quadrupoles fields may be adjusted in the model or the actual magnet settings of the accelerator may be changed to improve the agreement between the measured and predicted phase advance and so identify the source of the discrepancy.

An example from PEP-II has been presented in Fig. 2.13 [38]. From top to bottom the improved agreement of model and measurement is clear as the strength of a quadrupole pair in the interaction region was changed by a total of 0.15%. For each quadrupole value, the left column shows the entire ring while the right column shown an expanded view of a particular section. As can be seen, the final quadrupole strength (bottom) yields a satisfactory agreement with the model.

2.4.4 π Bump Method

Another method which can be used to identify local gradient errors is the π-bump technique applied at Tristan and at the ATF [39, 40, 41]. Here, local orbit bumps are induced, one by one, across each quadrupole magnet (or across small groups of quadrupole magnets), and the orbit is measured. A non-closure of a theoretically closed bump is indicative of an optics error within the region of the bump. Of course, an error in the calibration of the bump dipole magnets may also result in non-closure of the bump, but the effects of dipole errors and optics errors can be separated by their betatron phase. In particular, for an ideal π-bump, the bump "leakage", or non-closure, due to a gradient error and that due to a dipole error are perpendicular to each other. In more complex situations, computer programs can be used to process multiple measurements involving overlapping bumps to determine both the dipole and the focusing errors.

2.5 Multiknobs, Optics Tuning, and Monitoring

A problem frequently encountered in practice is to correct an aberration or to match one optical function without degrading others. For example, we may want to cancel a residual dispersion without affecting the beta functions. This can be done very efficiently by using a "multiknob", which is a combination of quadrupoles and skew quadrupoles (or possibly dipole correctors, sextupoles, even octupoles,...) that are simultaneously changed with the proper ratio and relative sign in such a way that only the aberration of interest is generated, i.e., corrected.

Such knobs are very powerful. For example, by scanning a dispersion knob, one can minimize the beam size on a downstream profile monitor or wire scanner, thus eliminating any residual dispersion. In a ring, similar knobs may be used to correct skew coupling and dispersion: by minimizing the beam size on a synchrotron light monitor with these knobs, the vertical equilibrium emittance can be reduced.

Multiknobs were used to minimize the spot sizes at the SLC interaction point (IP) and so produce the maximum luminosity. The spot sizes at the IP, which could be estimated from scans of the beam-beam deflection angle versus the offset between the two beams [42, 43, 44, 45], were routinely optimized by correcting the most important low-order aberrations including the waist shift, dispersion and skew coupling of either beam using the multiknobs which consisted of orthogonal linear combinations of quadrupoles and/or skew quadrupoles. For many years, the fine-tuning with colliding beams was performed on an hourly basis by measuring the spot size for different (typically 5–7 knob values) and adjusting each knob to the best value, as determined by fitting the square of the measured spot size as a function of the knob setting to a parabola.

A typical set of optimization scans for both beams is illustrated in Fig. 2.19. Small waist shifts were obtained by changing the strength of the last three quadrupoles (the 'final triplet') and of one quadrupole further upstream. Part of the linear coupling was corrected by means of a skew quadrupole at the same betatron phase as the final triplet. The residual dispersion at the IP was corrected using two normal and two skew $-I$ quadrupole pairs in the chromatic correction system (CCS), excited equally with opposite sign. Chromaticity was compensated by the two pairs of CCS-sextupoles. The strength of these sextupoles could also be varied to minimize the vertical spot size and, thus, to minimize the change in focusing with particle energy. As it turned out, the IP spot sizes proved to be relatively insensitive to the exact

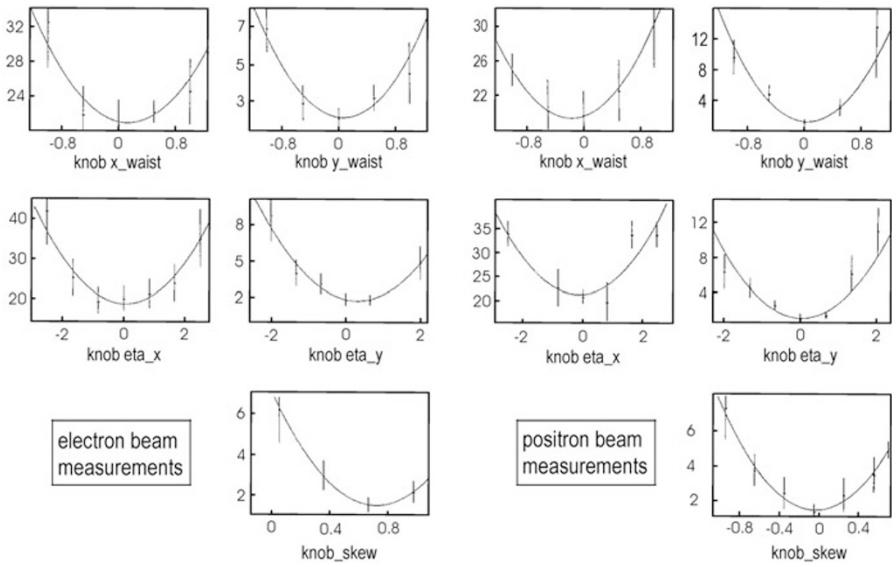

Fig. 2.19. Aberration scans at the SLC interaction point. Shown is the square of the measured horizontal or vertical convoluted spot size of the colliding electron and positron beams in units of μm^2 as a function of the multiknob setting in arbitrary units. In total 10 aberrations were controlled: the waist position, the horizontal and vertical dispersion, and the coupling from the horizontal into the vertical plane (for both beams). All these aberrations were corrected using different combinations of quadrupole and skew quadrupole magnets, which formed orthogonal multiknobs. Each multiknob was adjusted so as to optimize the IP spot size, which was determined using a parabolic fit to the beam sizes (squared). Until 1997, the beam size was inferred from the beam-beam deflection scans, with results as shown in the figure. The errors were significant [46]. In the last years of SLC operation, the aberrations were controlled and the luminosity was optimized using a new automatic 'dither' feedback, which made up/down changes to the multiknobs, and determined their effect on the luminosity from the correlated variation of the beamstrahlung-induced energy loss or other luminosity-related signals, such as the number of radiative Bhabha events [47]

settings of these chromatic sextupoles. In addition, several normal and skew sextupoles at non-dispersive locations in the final transformer were used for the correction of second-order geometric aberrations.

In general, the multiknob coefficients may be calculated in a variety of ways. A straightforward approach is to use the matching functions of the MAD program [32] to determine the relative changes in quadrupole strengths, ΔK_i, required to vary the parameter of interest (such as the beta function or the dispersion at some position).

A second approach to generate multiknobs involves using a singular value decomposition singular valued decomposition (SVD), which is a procedure commonly used for solving systems of linear equations with either too many or too few variables. The problem can be cast into a matrix equation of the form

$$\begin{pmatrix} \Delta\beta_x \\ \Delta\alpha_x \\ \Delta\mu_x \\ \Delta\eta_x \\ \Delta\eta'_x \\ \Delta\beta_y \\ \Delta\alpha_y \\ \Delta\mu_y \\ \Delta\eta_y \\ \Delta\eta'_y \\ \cdots \end{pmatrix}_{s=s_0} = \begin{pmatrix} B_{11} & B_{12} & \ldots & B_{1N} \\ B_{21} & B_{22} & \ldots & B_{2N} \\ & \cdots & \cdots & \\ B_{M1} & B_{M2} & \ldots & B_{MN} \end{pmatrix} \begin{pmatrix} \Delta K_1 \\ \Delta K_2 \\ \cdots \\ \Delta K_N \end{pmatrix} = BK\,. \tag{2.60}$$

The sensitivity matrix **B** may be obtained using an optics model or it may be determined empirically by measurement. To optimally constrain the solution, the number of adjustable parameters N should be larger or equal to the number of constraints M. So unless there is an optical symmetry that can be exploited, one should require $N \geq M$.

We postpone the detailed discussion of the SVD algorithm to Chap. 3. Here, we only note that an explicit solution using SVD will be presented for (3.28). Since (2.60) is of exactly the same form, it can be solved accordingly. If, for example, we want to create a dispersion knob which changes ΔD_x by 1 mm at location $s = s_0$ while keeping all other parameters constant, we can solve this problem by SVD. The latter will determine a set of changes in the quadrupole strengths $(\Delta K_1, \ldots, \Delta K_N)$, which fulfills the objective and which simultaneously minimizes the overall magnitude of the changes, i.e., the sum $\sum_i (\Delta K_i)^2$. This scheme can be generalized to include higher-order optics in an obvious way.

The knob coefficients are calculated only for small changes to the intermediate optics. If many quadrupoles are part of the knob, the R matrices between them will change as the knob is being varied. This means the coefficients determining the knob are not constant, but change depending on

the knob setting. If needed, this problem can be overcome by using nonlinear knobs, where the differential change in quadrupole strengths is recalculated in many small steps, effectively performing an integration of (2.60), in order to determine the final quadrupole values at the desired knob setting. Such nonlinear knobs were developed for the SLC final focus [48].

In many cases, the beam optics varies diurnally, or is sensitive to outside temperature, air pressure, etc., with associated variations in the accelerator performance. A continuous optics monitoring on relevant time intervals can yield a better understanding of the source of the perturbation and may also serve as a first step towards a stabilizing feedback. As an example, in the SLC linac a diagnostic pulse was used to monitor, quasi-continuously, the stability of the linac optics [49]. Every few minutes fast kickers induced betatron oscillations for individual selected bunches, prior to their injection into the linac. The propagation of the measured betatron oscillations through the linac was recorded, and then decomposed into an amplitude and a phase component. The betatron phase advance along the linac, thus inferred, proved a sensitive indicator of optics changes, which were caused, e.g., by variations in the linac energy profile due to drifts in the rf phase reference system, or by variations in beam current, bunch length, and bunch distribution.

2.6 Model-Independent Diagnostics

Some accelerators may be so complex and their optics so variable that it is not possible to establish a model which reproduces reality sufficiently well over longer periods of time as to be useful for diagnostics and monitoring purposes. A good example of such a situation is the SLAC linac. There, for example, the beam energy profile along the linac, which changed due to time-dependent drifts in the klystron (power source) phases, was impossible to determine precisely as the beam energy was measurable only in a few select locations. Therefore, the required (energy-matched) quadrupole gradients were not fully determined. Also transverse and longitudinal wakefields contributed to changes in the beam optics. These depended on the bunch charge, the bunch distribution, and the beam orbit.

A method employed to evaluate relative changes in the optics (determined in this case not only by lattice parameters but also by the beam parameters which were influenced strongly by hard to quantify beam-environment effects) entailed model-independent analysis of the orbit data. In such applications, no attempt is made to accurately determine the parameter set for an optics model, but the beam information and the beam response to certain perturbations are used directly to monitor the accelerator stability, to determine misalignments of accelerator structures and so on.

An interesting approach has been developed at SLAC by J. Irwin, C.-X. Wang, Y. Yan *et al.* [50, 51]. The primary quantity on which the analysis is based is a matrix of BPM readings $\mathbf{B}$, where

$$\mathbf{B} = \begin{pmatrix} b_{11}\ b_{12}\ \ldots\ b_{1m} \\ \ldots\ \ldots \\ b_{p1}\ b_{p2}\ \ldots\ b_{pm} \end{pmatrix} \tag{2.61}$$

has m columns, representing m different BPMs, and p rows, for p different beam pulses. In an actual application m may be of the order of 100 and p can be several 1000. There are many contributions to $\mathbf{B}$, for example changes in the initial conditions, changes in the beamline components, ground motion, BPM noise etc.

One can assume a linear (or quadratic) expansion of the form

$$\mathbf{b_i} = \mathbf{b_0} + \sum_{s=1}^{S} \Delta\hat{q}_i^s \left[\sigma_s \frac{\partial \mathbf{b}}{\partial q^s}\right] + \text{noise}\,, \tag{2.62}$$

where $\mathbf{b}$ denotes a row vector of the matrix $\mathbf{B}$, $\hat{q}^s$ represents the sth variable affecting the BPM readings (such as an incoming betatron oscillation, or klystron amplitude), and σ_s the rms variation of this perturbation. The variable $\hat{q}_i^s$ is normalized so that the rms value over time is one, or $\langle \hat{q}^s \rangle_{\text{rms}} = 1$, and $\langle \Delta q_s \rangle_{\text{rms}} = \sigma_s \langle \hat{q}^s \rangle_{\text{rms}} = \sigma_s$.

The above equation can be rewritten in matrix form as

$$\mathbf{B} = \mathbf{Q} \cdot \mathbf{F}^t + \eta\,, \tag{2.63}$$

where now η represents the statistical variations in the BPM readings, and the rows of the matrix $\mathbf{F}^t$ represent the sensitivity vectors $\mathbf{f} = [\sigma_s \partial \mathbf{b}/\partial q^s]$, weighted by the amount of variation (σ_s) detected in each variable.

The relative contribution of different variables q^s to the observed orbit variation can be identified fairly easily, provided certain tagging signals $\mathbf{t_i}$ (e.g., the bunch length, energy, or the incoming betatron oscillation inferred from the first few BPMs) are available and correlated to one (or more) of the source perturbations $\mathbf{f}_j$. For example, suppose the beam phase (arrival time) ψ and the current I are the tagging signals, Further suppose that their mutual correlation is not zero ($\langle \psi I \rangle \neq 0$) and that they are uncorrelated with all other perturbations. In this case one can multiply the matrix $\mathbf{B}$ by the two normalized tagging vectors $(\hat{\psi}, \hat{\mathbf{I}})$, obtaining

$$\begin{pmatrix} \hat{\mathbf{I}} \\ \hat{\psi} \end{pmatrix} \mathbf{B} = \begin{pmatrix} 1 & \langle \hat{\mathbf{I}} \hat{\psi} \rangle\ \ldots \\ \langle \hat{\mathbf{I}} \hat{\psi} \rangle & 1\ \ \ldots \end{pmatrix} \begin{pmatrix} \mathbf{f}_I \\ \mathbf{f}_\psi \end{pmatrix} + \mathcal{O}\left(\frac{\sigma}{\sqrt{p}}\right). \tag{2.64}$$

Equation (2.64) can be inverted, up to terms of order $\sigma/\sqrt{p}$. This inversion is of great interest because it provides the explicit 'space pattern' (shape) and the magnitude of the two orbit perturbations corresponding to changes in current or phase: $\mathbf{f_I}$, and $\mathbf{f}_\psi$.

For more refined studies, the SVD technique is again of great use. For example, via SVD we can decompose the matrix $\mathbf{B}$ as

$$\mathbf{B} = \mathbf{U}(\mathbf{W}\mathbf{V}^t)\,, \tag{2.65}$$

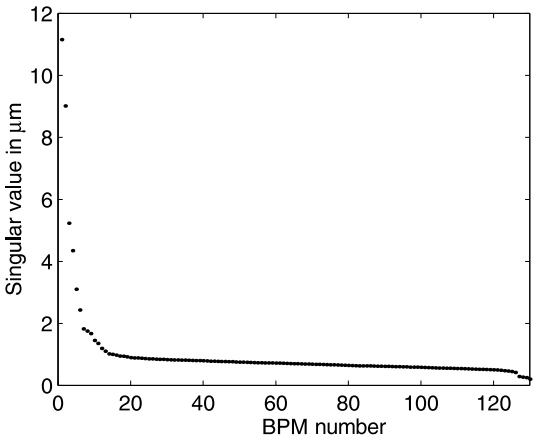

Fig. 2.20. Horizontal eigenvalues versus BPM number along the SLAC linac. In this plot 5000 pulses and 130 BPMs were analyzed

where $\mathbf{U}$ now represents the 'time patterns' (equivalent to $\mathbf{Q}$) and $(\mathbf{WV}^t)$ the 'space patterns' (equivalent to $\mathbf{F}^t$) [50]. If we calculate the significant SVD eigenvalues w_j of the matrix $\mathbf{W}$ for an increasing number of BPMs (in the order of their position along the beamline), we can find locations where additional eigenvalues become large. An example is shown [52] in Fig. 2.20. The large eigenvalues indicate locations where either the trajectory measurement has become faulty or an additional orbit jitter has been introduced. The latter may be caused for by improper regulation of corrector dipole or quadrupole fields or by a structure misalignment, which translates an incoming bunch length variation into an orbit variation. Online monitoring of the number of nonzero eigenvalues may prove useful in future accelerators to, among other things, ensure functionality of the BPMs particularly those used in orbit feedback.

This method was been applied to determine the alignment of all structures along the SLAC linac [53]. Similar methods can of course be used at storage rings.

2.7 Coherent Oscillations and Nonlinear Optics

2.7.1 Beam Response to a Kick Excitation

If a beam is deflected, e.g., by a kicker, a coherent oscillation is observed on the beam-position monitors. This oscillation will ultimately decrease to zero. There are at least three effects contributing to this decrease:

- radiation damping in an electron storage ring;
- head-tail damping, if the chromaticity is positive (above transition)

- filamentation due to a spread in betatron frequencies within the bunch. If the frequency spread results from nonzero chromaticity or energy spread, the signal will recohere after one synchrotron period. Such a frequency spread may also arise from amplitude-dependent detuning, in which case there is no recoherence.

2.7.2 Coherent Damping

Beautiful data from LEP are shown in Figs. 2.21–2.24, where the beam was kicked and the resulting oscillation was fitted to a harmonic function with exponentially decreasing amplitude, $A = A_0 e^{-t/\tau}$ and amplitude-dependent frequency [54]. In LEP the decrease of centroid oscillations was dominated by the head-tail damping and by radiation damping. The damping rate could thus be written as

$$\frac{1}{\tau} = \frac{1}{\tau_{SR}} + \frac{1}{\tau_{HT}} + \dots , \tag{2.66}$$

where the head-tail damping rate $1/\tau_{HT}$ is proportional to the bunch current, the chromaticity and the real part of the impedance. The contribution from head-tail damping was examined by varying the chromaticity [54]; see Fig. 2.21. Taking data for various beam currents, and extrapolating to zero determined the radiation damping time, as illustrated in Figs. 2.21 and 2.22.

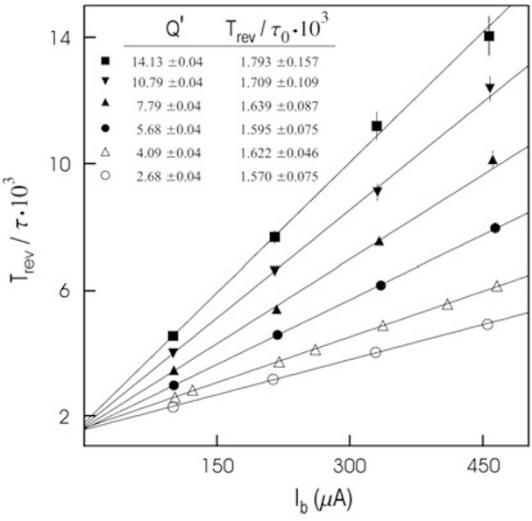

Fig. 2.21. Coherent damping rate in LEP as function of bunch current measured at a beam energy of 45.625 GeV for several chromaticities Q' [54]. The *straight lines* are fits to the individual samples. The *table* gives measured chromaticities and 'zero current' damping rates (Courtesy A.-S. Müller, 2001)

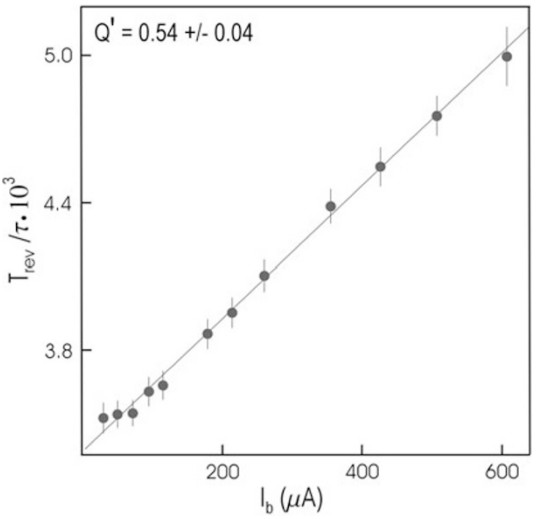

Fig. 2.22. Coherent damping rate in LEP as function of bunch current at one BPM with 60 GeV beam energy [54]. The *straight line* is a fit to the data (Courtesy A.-S. Müller, 2001)

2.7.3 Detuning with Amplitude

The tune shift with amplitude conveys information about the nonlinear fields experienced by the beam. Nonlinear fields have strong influence on beam stability and on the beam lifetime. Nonlinear fields result in an amplitude-dependent tune shift, since the average focusing experienced by a particle executing large betatron oscillations is changed.

Since, unlike filamentation, the radiation and head-tail damping change the real oscillation amplitude of individual particles, at LEP a single measurement could give the tune at various amplitudes (Fig. 2.23) and, from this, the detuning (see Fig. 2.24).

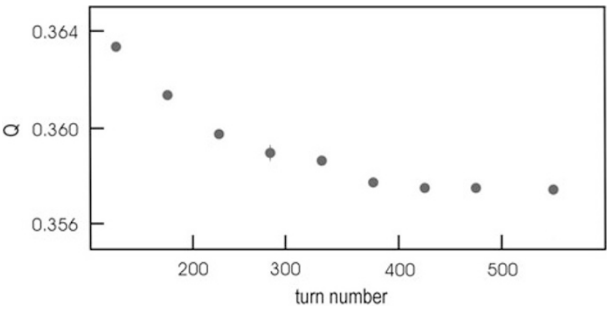

Fig. 2.23. Tune evolution inferred from a damped oscillation in LEP [54] (Courtesy A.-S. Müller, 2001)

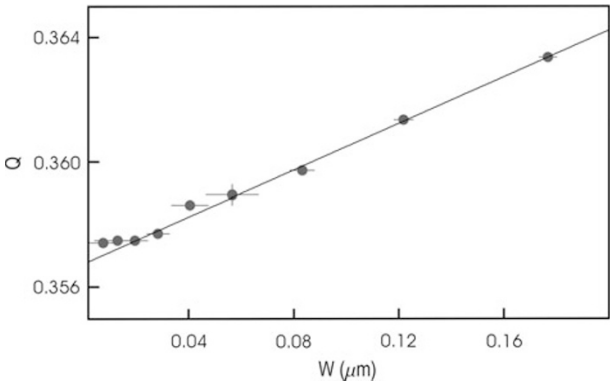

Fig. 2.24. Tune as a function of Courant-Snyder invariant with the 108/90 degree phase advance optics at an arc monitor in LEP [54]. The *straight line* is a fit to the data (Courtesy A.-S. Müller, 2001)

Earlier measurements at LEP and SPEAR also benefitted from the coherent damping, and determined the tune shift with betatron amplitude by applying an interpolated FFT with data windowing [55], as discussed in Sect. 2.1.2. In these LEP experiments, the beam was kicked in the vertical plane, and the tune was calculated over successive short time windows, of 32 turns each, as the beam oscillation damped rapidly. Figure 2.25 shows a result for LEP at 20 GeV. In Fig. 2.25, the vertical beam position is displayed as a function of time demonstrating the fast damping, over 200 turns. The shift of the horizontal and vertical betatron tune with the vertical action variable, as computed over time windows of 32 turns in length, is also shown. The vertical action was inferred from the os-

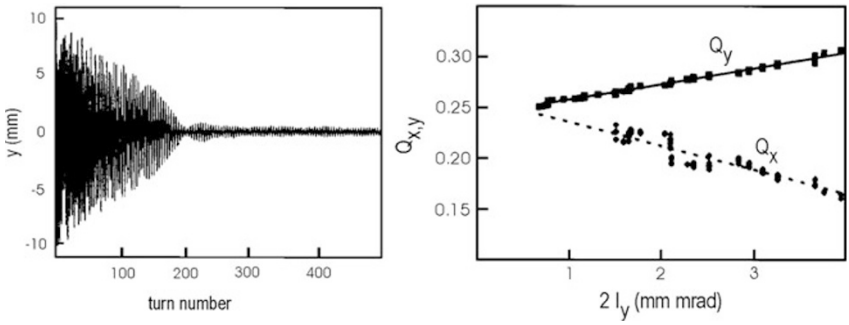

Fig. 2.25. Measurement of tune shift with amplitude in LEP at 20 GeV, using a high-precision FFT tune analysis [55]: (*left*) vertical oscillation amplitude after a kick; (*right*) horizontal and vertical betatron tunes vs. twice the vertical action variable I_y of the beam centroid motion. The observed tune shift with amplitude was consistent with the expected effect of the sextupole and octupole field components in the dipole magnets (Courtesy R. Bartolini, 1998)

cillation amplitudes in each time window. The amplitude-dependent tune shifts calculated by an off-line model and the measurement agreed to within 5%.

2.7.4 Filamentation due to Nonlinear Detuning

In other cases, notably at proton accelerators, the amplitude-dependent detuning will lead to a rapid reduction in the measured oscillation amplitude due to filamentation. In this case, to extract the amplitude-dependent detuning several measurements with different kick amplitudes may be required, and care is needed in the analysis of the signal.

An elegant mathematical formalism was developed by R.E. Meller and coworkers [56]. Considering a beam with a Gaussian transverse distribution, and further assuming that the decoherence of the centroid oscillation is due to a quadratic tune shift with amplitude a,

$$Q = Q_0 - \mu a^2 \,, \tag{2.67}$$

where a denotes the oscillation amplitude in units of σ and μ characterizes the strength of the nonlinear detuning, the centroid oscillation after a kick $\Delta x'$ is

$$\bar{x}(N) = \sigma_x Z A(N) \cos[2\pi Q_0 N + \Delta\bar{\phi}(N)] \,, \tag{2.68}$$

where N is the turn number (the kick is applied at $N = 0$), $Z = \beta\Delta x'/\sigma_x$ is the magnitude of the kick in units of σ_x, A the decoherence factor [56]

$$A(N) = \frac{1}{1+\theta^2} \exp\left[-\frac{Z^2}{2}\frac{\theta^2}{1+\theta^2}\right] \tag{2.69}$$

and $\bar{\phi}$ the centroid phase shift [56]

$$\Delta\bar{\phi}(N) = -\frac{Z^2}{2}\frac{\theta}{1+\theta^2} - 2\arctan\theta \,. \tag{2.70}$$

The parameter θ contains the time dependence of the decoherence:

$$\theta = 4\pi\mu N \,. \tag{2.71}$$

Figure 2.26 shows the decoherence factor $A(N)$, (2.69), as a function of θ for four different normalized kick strengths. For moderate kicks (up to 1σ), we can estimate the detuning parameter μ from the number of turns, $N_{1/2}$, after which the initial signal amplitude has decreased by a factor of two, as

$$\mu \approx \frac{1}{4\pi N_{1/2}} \,. \tag{2.72}$$

For larger kicks, this approximation overestimates the value of μ.

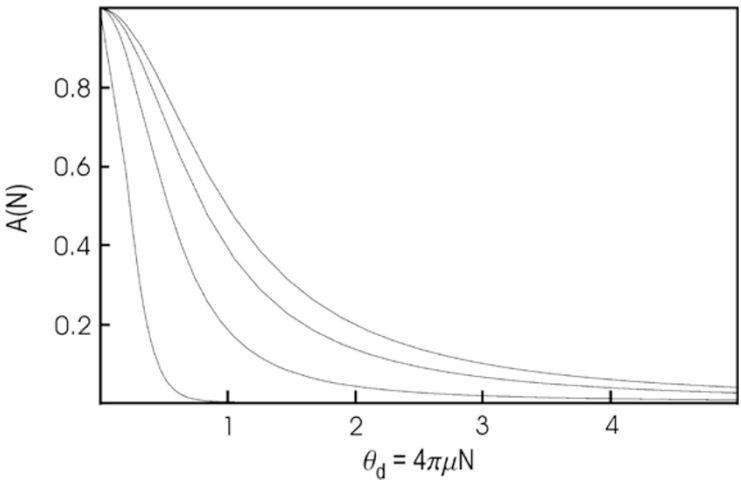

Fig. 2.26. Decoherence factor $A(N)$ as a function of $\theta = 4\pi\mu N$ for normalized kick strengths of $Z = 0.2$, 1, 2, and 5 (ordered from *top* to *bottom*)

2.7.5 Decoherence due to Chromaticity and Momentum Spread

Even without nonlinear detuning the beam signal can decohere due to the finite energy spread σ_δ and a non-vanishing chromaticity $Q' \neq 0$. In this case, the amplitude of the centroid motion after a kick, defined in (2.68), evolves as [56]

$$A(N) = \exp\left[-\frac{2\sigma_\delta^2 Q'^2}{Q_s^2}\sin^2(\pi N Q_s)\right] . \tag{2.73}$$

Note that (2.73) significantly differs from (2.69), so that the two effects can be experimentally separated. The signal described by (2.73) recoheres every full synchrotron period. From the signal modulation after a kick, one can infer the momentum spread if the chromaticity is known, and vice versa. In the case of chromaticity, the tune does not change: $\Delta\bar{\phi}(N) = 0$.

2.7.6 Resonance Driving Terms

Nonlinear magnetic fields not only induce a tune shift with amplitude, but they also excite higher-order resonances, which are visible as additional lines in the tune Fourier spectrum. A general betatron resonance is defined by the condition

$$kQ_x + lQ_y = p\,, \tag{2.74}$$

where k, l, and p are integers. Spectral analysis in the presence of nonlinear resonances [57] shows that in the Fourier spectrum of the horizontal coordinate $x(n)$, the above resonance gives rise to lines at the two frequencies

$(k \pm 1)Q_x + lQ_y$, and in the Fourier spectrum of the vertical signal $y(n)$ it generates lines at $kQ_x + (l \pm 1)Q_y$. Note that there is no line at $kQ_x + lQ_y$ [57], as one might have naively expected.

From amplitude, phase and frequency of the various spectral lines the dominant nonlinearities affecting the beam motion can be reconstructed [58, 59]. An example is shown in Fig. 2.27.

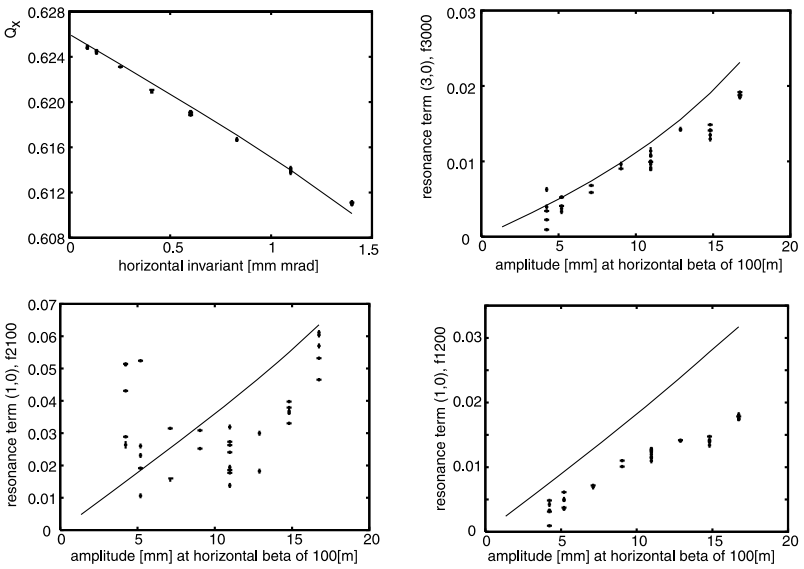

Fig. 2.27. Detuning and first-order resonance driving terms, measured at the SPS [60]. Shown are the betatron tune [*top left*] and the relative amplitude of several resonance lines with respect to the main tune line [*other*] as a function of the oscillation amplitude. Plotting symbols are experimental data; lines are from a tracking simulation. 170 turns were sampled per point (Courtesy F. Schmidt, 2000)

The nonlinearities may also be probed by adiabatically exciting large coherent betatron oscillations using an ac dipole [61]. This second method will be tested at the Relativistic Heavy Ion Collider (RHIC).

2.7.7 Tune Scans

The beam lifetime is often related to the dynamic aperture of the storage ring, where the term 'dynamic aperture' denotes the maximum stable betatron amplitude beyond which particles are lost after a certain finite number of turns. In the case of colliding beams, the lifetime is likely limited by the beam-beam interaction. Both dynamic aperture and the beam-beam interaction are sensitive to the value of the betatron tune. Measuring and plotting the beam lifetime as a function of the horizontal and vertical betatron tunes, Q_x and

Q_y, yields a tune diagram, in which higher-order resonances, given by (2.74), are evident as stripes with reduced lifetime. Figure 2.28 compares a typical beam-lifetime tune scan performed during the commissioning of the PEP-II High Energy Ring with the result of a dynamic-aperture simulation [62].

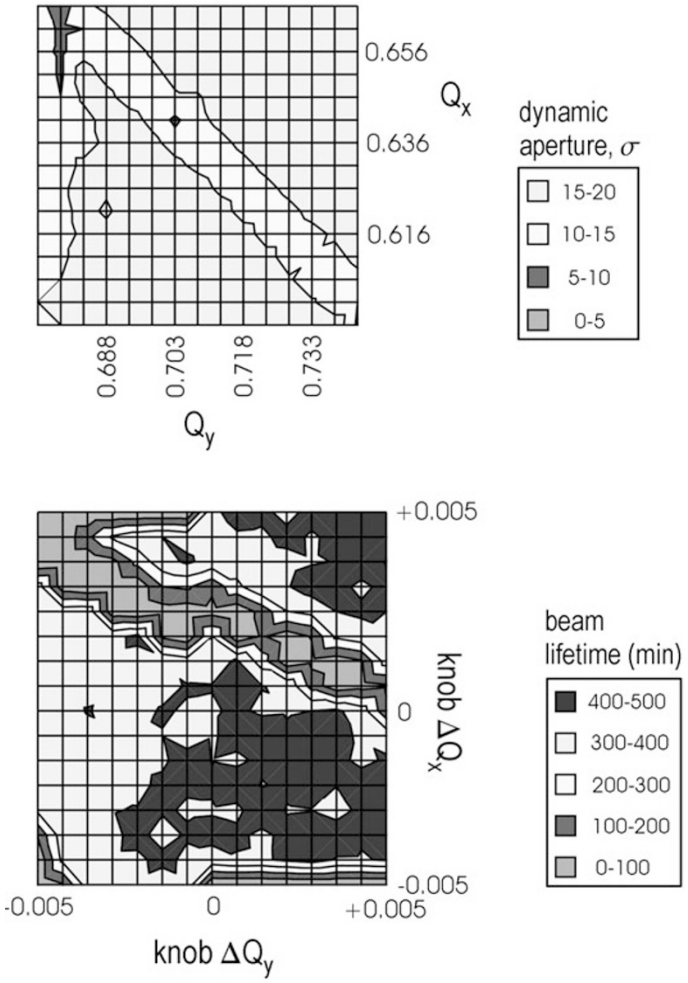

Fig. 2.28. Tune scan in PEP-II centered at $Q_x = 24.709$ and $Q_y = 23.634$ [62]: (*top*) simulated dynamic aperture (for a momentum offset $\Delta p/p = 10\ \sigma_\delta$, where δ is the rms momentum spread) as a function of the horizontal and vertical betatron tunes, Q_x and Q_y; (*bottom*) measured beam lifetime as a function of Q_x and Q_y. Total scan range is ± 0.005 on both axes. The different slope of the resonance line, as compared with the top figure, is attributed to a miscalibration of the tune knobs (Courtesy Y. Cai, 1998)

2.8 Betatron Coupling

Skew quadrupole field errors and detector solenoids generate betatron coupling between the horizontal and vertical planes of motion. Spurious betatron coupling is a concern, since it may reduce the dynamic aperture [63] and since, in electron accelerators, it contributes to the vertical equilibrium emittance. The coupling of horizontal and vertical oscillations generates two new eigenmodes of oscillation. These eigenmodes are no longer purely vertical or purely horizontal, but rather they correspond to oscillations whose reference planes are tilted and rotate with the azimuthal position s. In this case, new coupled beta functions can be defined [64, 65, 66, 67].

To illustrate the fundamental coupling phenomenon, we consider two coupled linear oscillators described by the equations

$$\begin{aligned} \frac{\mathrm{d}^2 x}{\mathrm{d}\theta^2} + Q_0^2 x &= -\kappa y\,, \\ \frac{\mathrm{d}^2 y}{\mathrm{d}\theta^2} + Q_0^2 y &= -\kappa x\,, \end{aligned} \tag{2.75}$$

where κ is a constant describing the strength of the coupling. We can decompose the oscillation into two new normal modes. The normal-mode coordinates are

$$u = \frac{x+y}{2} \tag{2.76}$$

$$v = \frac{x-y}{2}\,. \tag{2.77}$$

Using (2.75) it is easily verified that the normal coordinates fulfill the uncoupled equations of motion

$$\frac{\mathrm{d}^2 u}{\mathrm{d}\theta^2} + (Q_0^2 + \kappa)u = 0\,, \tag{2.78}$$

$$\tag{2.79}$$

$$\frac{\mathrm{d}^2 v}{\mathrm{d}\theta^2} + (Q_0^2 - \kappa)v = 0\,. \tag{2.80}$$

The squared eigenmode frequencies are $Q_u^2 = (Q_0^2 + \kappa)$ and $Q_v^2 = (Q_0^2 - \kappa)$. Although the original oscillation frequencies were equal to Q_0 in both planes, x and y, the coupling introduces a frequency split between the two eigenmode frequencies $(Q_u^2 - Qv^2) = 2\kappa$, which is proportional to the coupling parameter κ.

2.8.1 Driving Terms

In a storage ring, the coupling between the planes is not constant around the ring, and as a consequence there are two important coupling parameters. These are the so-called two driving terms for the sum and difference resonances, which are given by [63, 67, 68]:

$$|\kappa_{\pm}| = \left| \frac{1}{2\pi} \oint ds \; k_s(s) \sqrt{\beta_x(s)\beta_y(s)} \; e^{\mathrm{i}(\phi_x \pm \phi_y - (Q_x \pm Q_y - q_{\pm})2\pi s/C)} \right| , \quad (2.81)$$

where k_s is the normalized gradient of the skew quadrupole (in units of m^{-2}), $\phi_{x,y}$ the horizontal and vertical betatron phase, C is the circumference, $\beta_{x,y}$ are the uncoupled beta functions, and, in (2.81), we assume that the betatron tunes are near the sum or difference resonance

$$Q_x \pm Q_y + q_{\pm} = 0 \, , \quad (2.82)$$

where $q_{\pm}$ is an integer. The dynamic aperture or the beam lifetime of colliding beams can be increased by measuring and minimizing the two driving terms $|\kappa_{\pm}|$.

In an electron storage ring, the vertical emittance contribution due to weak betatron coupling is [69]

$$\gamma\epsilon_y = \frac{C_q \gamma^3}{16 \oint G^2 ds} \oint \mathcal{H}_x |G^3| \left[\sum_{\pm} \frac{|W_{\pm}(s)|^2}{\sin^2 \pi(\Delta Q_{\pm})} + \frac{2\mathrm{Re}\left\{ W_+^*(s) W_-(s) \right\}}{\sin \pi(\Delta Q_+) \; \sin \pi(\Delta Q_-)} \right] \mathrm{d}s \, , \quad (2.83)$$

where $C_q = 3.84 \times 10^{-13}$ m, $\mathcal{H}_x$ is the horizontal dispersion invariant ($\mathcal{H}_x = (D_x^2 + (\alpha_x D_x + \beta_x D_x')^2)/\beta_x$), $G = 1/\rho$ the inverse bending radius, the star $*$ denotes the complex conjugate, Re gives the real portion of its argument, $\Delta Q_{\pm} = Q_x \pm Q_y + q_{\pm}$ characterizes the distance to the resonance (which should not be too small for the perturbation theory to be valid), and

$$W_{\pm}(s) = \int_s^{s+C} \mathrm{d}z \; k_s(s') \sqrt{\beta_x(s')\beta_y(s')} e^{\mathrm{i}[(\phi_x(s) \pm \phi_y(s)) - (\phi_x(s') \pm \phi_y(s')) + \pi(Q_x \pm Q_y)]} \quad (2.84)$$

are the s-dependent driving terms, including all Fourier components. Note that $|W_{\pm}(0)| = |\kappa_{\pm}|2\pi$, if there is a single skew quadrupole in the ring.

Equation (2.83) shows that, in order to minimize the vertical emittance, the driving terms $W_{\pm}(s)$, for the two nearest sum and difference resonances, should be corrected.

2.8.2 First Turn Analysis

Large coupling sources can be identified as locations where a horizontal orbit change generates a vertical kick and vice versa. In order to find such locations, the orbit is changed in one plane, by exciting steering correctors or by changing injection conditions (for example, the kicker amplitude), and the effect on the orbit in the perpendicular plane is measured. The same type of analysis can be applied to a transport line.

Large numbers of orbits and BPM data for excitations of different correctors can be fitted to determine the skew quadrupole component of each magnet in the beam line.

2.8.3 Beam Response after Kick

The driving term $|\kappa_-|$ may be measured by first kicking the beam, and then observing its response in the plane of the kick over *many turns*.

In the vicinity of the difference resonance, the envelopes of the oscillations in the horizontal and vertical plane exhibit a beating (energy exchange between the two planes) with a characteristic total modulation amplitude of [67, 70]

$$S = \frac{\hat{x}^2_{\min}}{\hat{x}^2_{\max}} \,. \tag{2.85}$$

Here $\hat{x}$ denotes the envelope of the betatron oscillation in the plane in which the kick was applied; $\hat{x}_{\min}$ is its minimum value, and $\hat{x}_{\max}$ its maximum value; these two extreme values are assumed alternately, with a modulation (or beating) period T. The driving term for the difference resonance, $|\kappa_-|$ of (2.81), is given by [70]

$$|\kappa_-| = \frac{\sqrt{1-S}}{f_{\rm rev} T} \,. \tag{2.86}$$

Thus measuring the modulation period T and the squared envelope ratio S after a kick is sufficient to infer $|\kappa_-|$.

An example from the ATF Damping Ring is shown in Fig. 2.29. The frequency spectrum from a horizontal BPM signal is viewed over a wide frequency range on a spectrum analyzer (top figure), and the frequency of the betatron signal is identified as the peak of the spectrum. The span of the spectrum analyzer is then set to zero, and its center set to the betatron frequency. This produces a signal proportional to the square of the betatron-oscillation amplitude. The output signal of the spectrum analyzer can be viewed on an oscilloscope, with results as displayed in Fig. 2.29 (bottom). The slow oscillation in this picture corresponds to synchrotron motion (the BPM is at a location with nonzero mismatched dispersion), while the fast beating reflects the transverse coupling. The picture was taken for a tune separation of $|Q_x - Q_y + q_-| \approx 0.02$. If the two tunes are separated further, the modulation period increases and the modulation amplitude decreases. Using (2.86) with $T \approx 17.6$ μs and $S \approx 0.3$–0.7, we infer a coupling term of $|\kappa_-| \approx 0.02$, consistent with other measurements [71].

It is of course possible to perform a much more detailed analysis of multi-turn BPM data. For example, one can determine the evolution of the coupled optical functions (e.g., the tilt angle of the two transverse eigenplanes) around the ring. An example may be found in [72].

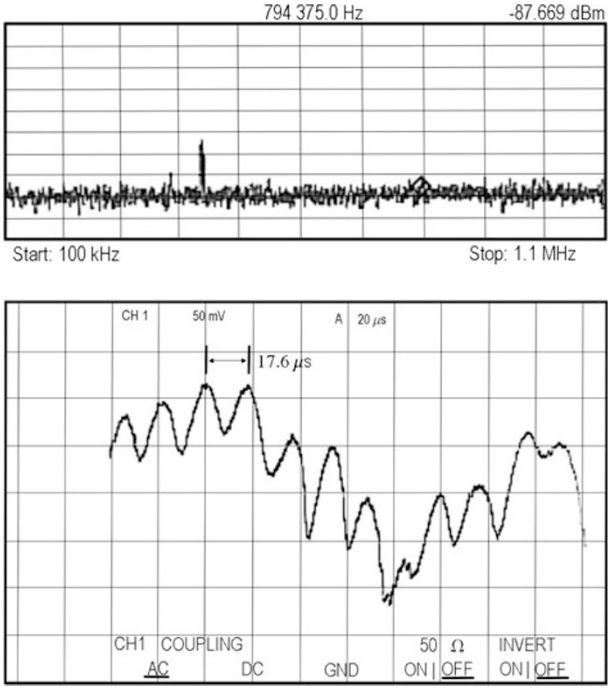

Fig. 2.29. Monitoring betatron coupling at the ATF Damping Ring [71]. (*top*) Frequency spectrum of a horizontal pick up on a spectrum analyzer; (*bottom*) evolution of the peak signal in the frequency spectrum as a function of time, as viewed on an oscilloscope; the slow variation reflects synchrotron motion, the fast beating with a period of about 17.6 μs is due to the transverse coupling; the amplitude and period of the modulation can be used to determine the driving term $|\kappa_-|$

2.8.4 Closest Tune Approach

Near the difference resonance, the tunes of the two coupled eigenmodes in the vertical plane are [67, 70]

$$Q_{I,II} = \frac{1}{2}\left(Q_x + Q_y + q \pm \sqrt{(Q_x - Q_y + q)^2 + |\kappa_-|^2}\right) , \qquad (2.87)$$

where Q_x and Q_y are the tunes which one would expect without coupling. A similar formula, with the same fractional values of $Q_{I,II}$, describes the coupled tunes in the horizontal plane. Equation (2.87) shows that the measured fractional parts of Q_I and Q_{II} are never exactly equal, but can only approach each other up to a distance $|\kappa_-|$. Figure 2.30 illustrates this with an example from the PEP-II High Energy Ring (HER), which is consistent with our simple example in (2.79) and (2.80). A common technique for correcting the betatron coupling in a storage ring is to minimize the distance of closest approach using at least two skew quadrupole magnets. It is often the

only correction necessary, especially if the tunes are close to the difference resonance $Q_x - Q_y + q = 0$.

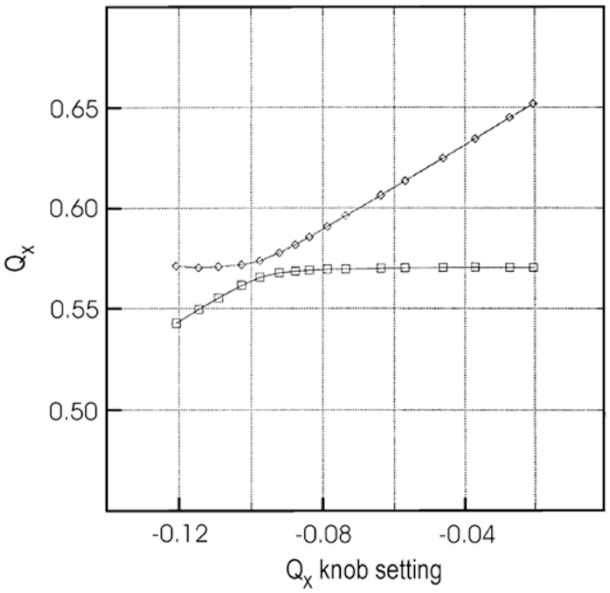

Fig. 2.30. Closest tune approach in the PEP-II HER before final correction [73]. Shown are the measured fractional tunes as a function of the horizontal tune 'knob' (which would only change Q_x if the accelerator were uncoupled), in dimensionless units. The minimum tune distance is equal to the driving term $|\kappa_-|$ of the difference resonance (Courtesy Y. Cai, 1998)

2.8.5 Compensating the Sum Resonance

In the vicinity of the difference resonance, while there is a continuous energy exchange between the two transverse planes, the beam or particle motion remains bounded. In contrast, close to the sum resonance for which $|Q_x + Q_y + q| < |\kappa_+|$, the motion is unstable. The total width of the stop band around the sum resonance is equal to twice the driving term $|\kappa_+|$ of (2.81) [67]. In practice, while perhaps more difficult, than determining the driving term $|\kappa_-|$ for the difference resonance using the closest-tune approach, the driving term for the sum resonance $|\kappa_+|$ can be compensated by adjusting two skew quadrupoles so as to minimize the stop band width.

In order to optimize the vertical emittance in an electron damping ring, in general it is necessary to correct both driving terms, $|\kappa_-|$ and $|\kappa_+|$. A minimum of six skew quadrupole correctors are needed to correct both coupling driving terms and the the vertical dispersion function since one sine-like and one cosine-like corrector are required for correction type of each aberration.

Even after perfect global coupling compensation, there can be a residual contribution of local coupling to the vertical emittance. Equation (2.83) suggests that local correction might be necessary. It is indeed employed at several light sources, the KEK ATF, and the B-factories.

2.8.6 Emittance near Difference Resonance

A third approach to determine the coupling parameter $|\kappa_-|$ is to measure the horizontal (or vertical) emittance as a function of the distance from the difference resonance [71]. Equation (2.83) does not apply close to the resonance. Instead, near the difference resonance, the horizontal emittance is described by [74]

$$\epsilon_x = \epsilon_{x0}\,\frac{(\Delta Q_-)^2 + |\kappa_-|^2/2}{(\Delta Q_-)^2 + |\kappa_-|^2}\,, \tag{2.88}$$

where ϵ_{x0} is the unperturbed horizontal emittance without coupling or far from the resonance. The tune difference $\Delta Q_- = (Q_x - Q_y - n)$ function is related to the measured tune difference $\Delta Q_{I,II} = |Q_{II} - Q_I|$ via

$$\Delta Q_- = \sqrt{\Delta Q_{I,II}^2 - |\kappa_-|^2}\,. \tag{2.89}$$

After inserting this into (2.88), the latter becomes

$$\epsilon_x = \epsilon_{x0}\,\left(1 - \frac{|\kappa_-|^2}{2\Delta Q_{I,II}^2}\right)\,. \tag{2.90}$$

A more recent and different derivation [75] describes the horizontal and vertical emittances near the difference resonance in the following modified form:

$$\epsilon_x = \epsilon_{x0} - (\epsilon_{x0} - \epsilon_{y0})\frac{|\kappa_-|^2/2}{(\Delta Q_-)^2 + |\kappa_-|^2 + \Delta Q_-\sqrt{(\Delta Q_-)^2 + |\kappa_-|^2}}\,, \tag{2.91}$$

$$\epsilon_y = \epsilon_{y0} + (\epsilon_{x0} - \epsilon_{y0})\frac{|\kappa_-|^2/2}{(\Delta Q_-)^2 + |\kappa_-|^2 + \Delta Q_-\sqrt{(\Delta Q_-)^2 + |\kappa_-|^2}}\,. \tag{2.92}$$

In an experiment at the KEK ATF, the horizontal emittance ϵ_x was inferred from the spot size σ_x measured with the ATF interferometric monitor [76], using the formula $\epsilon_x = (\sigma_x^2 - (D_x\delta)^2)/\beta_x$, where the design beta function ($\beta_x = 0.33$ m) and dispersion ($D_x = 4$ cm) at the point of light emission, and the calculated energy spread of $\delta = 7.14 \times 10^{-4}$ were assumed.

Performing a nonlinear fit to (2.90) of the emittance ϵ_x obtained for various tune differences $\Delta Q_{I,II}$, the driving term $|\kappa_-|$ can be extracted [71]. Figure 2.31 shows the measured horizontal emittance as a function of tune separation $\Delta Q_{I,II}$. The result of the nonlinear fit is also depicted. The coupling strength of $|\kappa_-| \approx 0.037$ inferred from the fit [71] agrees well with the coupling strength $|\kappa_-| \approx 0.042$ obtained from a measurement of the closest tune approach under the same conditions [77].

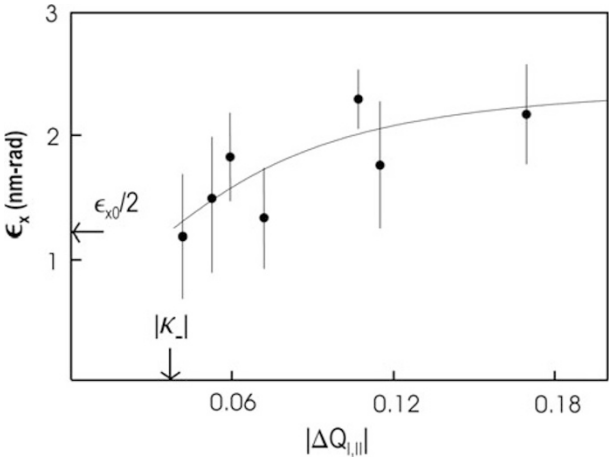

Fig. 2.31. Horizontal emittance as a function of the tune separation $\Delta Q_{I,II}$ at the ATF Damping Ring; the measured data [77] as well as the result of a nonlinear fit to (2.90) are shown; the fitted coupling strength of $|\kappa_-| \approx 0.037$ is consistent with the value $|\kappa_-| \approx 0.042$ inferred from a simultaneous measurement of the closest tune approach [77, 71]

2.8.7 Emittance near Sum Resonance

The driving term of the sum resonance κ_+ may also be inferred by measuring the emittance in the vicinity of this resonance [78]. However, we caution the reader that experimental experience with this scheme is scarce, and that there appear to be some uncertainties in the predictions by different theories and computer simulations.

Using the formulae in [74] and [79], the vertical emittance should depend on the measured distance to the sum resonance, $\Delta Q_{I,II,+} \equiv |Q_I + Q_{II} - p|$ (where p is the integer which minimizes the expression) as [78]

$$\epsilon_y \approx \epsilon_{x0} \frac{2|\kappa_+|^2}{5|\kappa_+|^2 + \Delta Q_{I,II,+}^2}\,, \tag{2.93}$$

while the horizontal emittance should vary as

$$\epsilon_x \approx \epsilon_{x0} \frac{3|\kappa_+|^2 + \Delta Q_{I,II,+}^2}{5|\kappa_+|^2 + \Delta Q_{I,II,+}^2}\,. \tag{2.94}$$

In deriving (2.93) and (2.94), we have used that the actual tune distance to the sum resonance, $\Delta Q_+ = (Q_x + Q_y - n)$, is related to the measured distance $\Delta Q_{I,II,+} = |Q_{II} + Q_I - p|$ via

$$\Delta Q_+ = \sqrt{\Delta Q_{I,II,+}^2 + |\kappa_+|^2}\,. \tag{2.95}$$

2.8.8 Coupling Transfer Function

A different method of measuring the coupling is through the 'coupling transfer function' [80]. Here, the beam is excited horizontally, while detecting the resulting vertical coherent motion. Such a technique was used to continually monitor and correct the coupling strength during collisions in the CERN ISR [80].

2.8.9 Excursion: Flat Versus Round Beams

Frequently the question is asked why flat beams are used in electron-positron colliders. One answer is that the beams naturally become flat due to synchrotron radiation, as will be discussed in a later chapter. However, as we have seen, one could make these beams round, by operating close to the linear difference resonance $Q_x = (Q_y + k)$, where k is an integer.

In general, whether round or flat beam collisions would give higher luminosity is not easy to answer. The luminosity is defined as the reaction rate divided by the cross section. Aside from the beam energy, it is the most important performance parameter of a high-energy collider. For a ring collider, assuming Gaussian beam distributions, the luminosity, typically measured in units of $\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$, can be expressed as (compare Chap. 1)

$$L = \frac{N_b^2 n_b f_{\mathrm{rev}}}{4\pi\sigma_x\sigma_y}\,, \tag{2.96}$$

where N_b is the number of particles per bunch, n_b the number of bunches per beam, and f_{rev} the revolution frequency. The beam sizes at the interaction point (IP), $\sigma_{x,y}$, can be expressed in terms of the IP beta function and the transverse emittances as

$$\sigma_x = \sqrt{\beta_x \epsilon_x}\,, \tag{2.97}$$

$$\sigma_y = \sqrt{\beta_y \epsilon_y}\,. \tag{2.98}$$

The sum of the emittances may be held constant, equal to ϵ_{x0}, while the ratio of the horizontal and vertical emittance can be varied via the betatron coupling according to

$$\epsilon_x = \frac{\epsilon_{x0}}{1+\kappa}\,, \tag{2.99}$$

$$\epsilon_y = \kappa\frac{\epsilon_{x0}}{1+\kappa}\,. \tag{2.100}$$

An important constraint of a collider is the beam-beam tune shift, i.e., the tune shift and tune spread due to the additional focusing at the collision point

by the field of the opposing beam. The beam-beam tune shift is characterized by the parameters $\xi_{x,y}$,

$$\xi_{x,y} = \frac{N_b r_e}{2\pi\gamma} \frac{\beta_x}{\sigma_{x,y}(\sigma_x + \sigma_y)} \,, \tag{2.101}$$

where r_e is the classical electron radius, and γ the relativistic factor. Best performance is achieved if the vertical and horizontal are both at the maximum possible value. If we assume that the maximum tune shift parameter is approximately equal for the two planes, $\xi_x = \xi_y$, it follows immediately that

$$\frac{\epsilon_x}{\epsilon_y} = \frac{\beta_x}{\beta_y} = \frac{1}{\kappa} \,. \tag{2.102}$$

Inserting this into (2.96), one can rewrite the luminosity as

$$L = \frac{1+\kappa}{2} \, \gamma\xi_{x,y} \, \frac{1}{r_e} \frac{N_b f_{\rm rev} n_b}{\beta_y} \,. \tag{2.103}$$

This expression shows that round beams may give two times higher luminosity, thanks to the factor $(1+\kappa)$, but only if the horizontal beta function can be made as small as the vertical, $\beta_x = \beta_y/\kappa = \beta_y$.

Recently a transformation was proposed [81] which can convert a round beam created in a solenoid field into a flat beam outside of the solenoid or vice versa. As discussed in Chap. 4, the transformation which achieves the conversion can be realized by a combination of skew and regular quadrupoles. As a practical example, using this scheme the round beam from an rf photoinjector can be transformed into a beam with a flat emittance ratio $\epsilon_y/\epsilon_x \ll 1$, which is more suitable for a linear collider [82]. A first experiment at the Fermilab A0 line has demonstrated the viability of this scheme [83] (see Chap. 5). Another application is the matching of a flat electron beam to a round proton beam, e.g., in an electron cooling device.

Exercises

2.1 Schottky Signals

The longitudinal spectrum of a single particle, denoted by k, of an unbunched beam circulating in a storage ring, considering positive frequencies only, is given [84] by

$$i_k(t) = e f_{\rm rev} + 2e f_{\rm rev} \sum_{n=1}^{\infty} \cos(n\omega_{\rm rev} t + \phi) \,, \tag{2.104}$$

where n represents the turn number and $\omega_{\rm rev} = 2\pi f_{\rm rev}$ is the angular revolution frequency of this particle.

a) Sketch the current spectrum for this particle in both the time and frequency domains.

b) Show that summing (2.104) over N particles and averaging in time gives the dc beam current $I_{\rm dc} = \left\langle \sum_{k=1}^{N} i_k \right\rangle_t = N e f_{\rm rev}$, assuming that different particles (index k) have slightly different angular revolution frequencies $\omega_{{\rm rev},k}$ and different initial phases ϕ_k, and that $f_{\rm rev}$ denotes the average revolution frequency.

c) The power spectrum one would measure with a spectrum analyzer is proportional to the rms current $I_{\rm rms} = \langle (\sum_k i_k)^2 \rangle_t^{\frac{1}{2}}$. Show that

$$I_{\rm rms} = 2 e f_{\rm rev} \sqrt{\frac{N}{2}} \,. \tag{2.105}$$

2.2 Betatron Tunes

a) From Fig. 2.5, what was the synchrotron tune in LEP?

b) Suppose that when the quadrupole family for the horizontal tune is increased in strength, the horizontal tune line is observed to move to the right in similar measurements. What is the fractional horizontal betatron tune?

c) After how many turns would a particle with the tunes found in a) and b) return to exactly the same place in the longitudinal or horizontal phase space?

2.3 Application of Multipole Field Expansion

A useful formula for fields produced by a magnet of order n ($n = 0$ represents a dipole) is[2]

$$B_y + {\rm i} B_x = B_0 \sum_{n=0} (b_n + {\rm i} a_n)(x + {\rm i} y)^n \,, \tag{2.106}$$

where for upright, or 'normal' magnets $b_n \neq 0$ and $a_n = 0$ while for magnets rotated by 45°, so-called 'skew' magnets, $b_n = 0$ and $a_n = 0$.

a) For the case of a upright, or 'normal', sextupole ($b_2 \neq 0$, $a_2 = 0$), show that the Lorentz force seen by the particle is

$$F_x \propto (x^2 - y^2) \quad \text{and} \quad F_y \propto xy \,. \tag{2.107}$$

b) For a nondispersive orbit ($x = x_{\rm co} + x_\beta$ and $y = y_{\rm co} + y_\beta$, where the subscripts co and β refer to the closed orbit and to the betatron orbit), show that a beam off-axis in x experiences a normal quadrupole field while a beam off-axis in y experiences a skew quadrupole field.

c) What transverse-to-longitudinal coupling terms arise if the dispersive contribution to the orbit ($x_\delta = D_x \delta$ and $y_\delta = D_y \delta$) are also included?

[2] We note that (2.106) refers to the American convention, and that in Europe the exponent of $(x + {\rm i} y)$ is chosen as $(n-1)$, so that a sextupole field would be given by b_3 and not by b_2.

2.4 Beta-Beat

Consider a quadrupole error. We split the quadrupole into two halves each of which is parametrized by the usual matrix for a quadrupole

$$R_q = \begin{pmatrix} 1 & 0 \\ -1/(2f) & 1 \end{pmatrix} . \tag{2.108}$$

where $1/f \equiv \Delta K$ in a thin-lens approximation, ΔK being the (integrated) quadrupole strength error. Let the one turn map (OTM) with the field error be given by $R = R_q R_0 R_q$ (where for algebraic convenience we have taken as a reference point the center of the quad), where R_0 is the OTM for the ideal ring:

$$M = \begin{pmatrix} \cos\phi_0 & \beta_0 \sin\phi_0 \\ -(1/\beta_0)\sin\phi_0 & \cos\phi_0 \end{pmatrix} . \tag{2.109}$$

a) Show that the perturbed phase advance ϕ is given in terms of the unperturbed phase advance ϕ_0 by

$$\cos\phi = \cos\phi_0 - \frac{\beta_0}{2f}\sin\phi_0 . \tag{2.110}$$

b) From the measurement of phase advance errors in Fig. 2.9, estimate the amplitude of the modulation on the beam size.

2.5 Quadrupole with a Shorted Coil

So-called difference orbit measurements are taken by saving a BPM reference orbit under nominal conditions, perturbing the beam (for example, by powering a single corrector) and forming the difference orbit consisting of measurements of the perturbed orbit minus that of the nominal orbit. This difference orbit may be easily compared with model expectation. Suppose such lattice diagnostic measurements are made using a negatively charged beam and reveal that the beam is errantly deflected up and outboard (away from ring center) at a focussing quadrupole.

a) Based on these observations, which of the quadrupoles' coils is suspect of a turn-to-turn short?

b) With a position resolution of 10 μm at a BPM located $l = 1$ m downstream of the quadrupole, how well can the relative change in the current through the offending coil be determined assuming $K = Bl_q/((B\rho)a) = 1.0\,\mathrm{m}^{-1}$ with a pole tip radius of $a = 1$ cm?

2.6 Quadrupole Gradient Errors

a) Design a closed π-bump spanning a focussing quadrupole of strength K assuming perfect calibration of the corrector dipoles.

b) Using this bump, suppose that the orbit is observed not to close with a residual amplitude A measured using a BPM. Find an expression relating the measured leakage A to the gradient error ΔK.

2.7 Multiknobs

A storage ring employs two families of quadrupoles for varying the tranverse focussing. As is common, each quadrupole family has influence on both the horizontal and vertical betatron tunes, Q_x and Q_y respectively. Design a multiknob that allows independent control of each of the betatron tunes; that is, find strength coefficients for each of the quadrupole families for the cases (i) $\Delta Q_x \neq 0$, $\Delta Q_y = 0$ and (ii) $\Delta Q_x = 0$, $\Delta Q_y \neq 0$.