Review of linear optics measurement and correction for charged particle accelerators

Rogelio Tomás
CERN, CH-1211 Geneva 23, Switzerland

Masamitsu Aiba
Paul Scherrer Institut, CH-5232 Villigen PSI, Switzerland

Andrea Franchi
ESRF, CS 40220, 38043 Grenoble Cedex 9, France

Ubaldo Iriso
ALBA-CELLS, C. de la Llum 2-26, 08290–Cerdanyola, Spain

(Received 24 September 2016; published 31 May 2017)

Measurement and correction of charged particle beam optics have been a major concern since the advent of strong focusing synchrotron accelerators. Traditionally, particle colliders have led the development of optics control based on turn-by-turn beam centroid measurements, while lepton storage rings have focused on closed-orbit-response matrix techniques. Recently, considerable efforts are being invested in comparing these techniques at different synchrotron radiation sources and colliders. An emerging class of less invasive techniques based on the optimization of performance-related observables is demonstrating a great potential. In this paper, a review of existing techniques is presented highlighting comparisons, relative merits and limitations.

DOI: 10.1103/PhysRevAccelBeams.20.054801

I. DAWN OF A NEW DISCIPLINE

Perturbations from field imperfections and misalignments of accelerator components became a concern for accelerator design and operation with the conception of the strong focusing theory in 1957 [1]. However, the assumed approach was to specify design tolerances that would not impact machine performance. For example, in [1] it was envisaged that with 1% rms gradient errors any particular machine would be unlikely to have more than 8% peak β-beating. In the following decade the AGS experienced tune drifts and aperture limitations due to quadrupolar errors at injection energy which were mitigated with dedicated correction circuits [2]. Actually, it seems that the modern nomenclature of tune to designate the betatron frequency [1] or betatron number [3] can be attributed to the intense efforts to tune this quantity. In 1972 the term tune was already widely used [4]. Tunes are of critical importance since resonant motion must be avoided [1,5–10]. An entertaining method to find resonances in the tune diagram can be found in [11].

In 1975 first documented beam-based measurements of the average beta function over independently powered quadrupoles took place in the ISR using the tune change due to a quadrupole gradient variation [12]. This technique is referred to as K modulation in the following.

In the same year first beam-based measurements and corrections of transverse coupling was performed at the AGS [13] while a more refined technique was being developed for the ISR [14]. These techniques are based upon turn-by-turn beam centroid measurements at a single location in the accelerator.

The plans to build larger colliders and the use of low-β⁰ insertions triggered the need to measure and correct chromaticity [4,15,16] in the mid 1970s.

In 1983 a major achievement took place in the ISR. Beam position monitors (BPMs) around the collider were used to measure betatron phase advance and beta functions from betatron oscillations [17]. This was the first realization of an optics measurement from turn-by-turn BPM data around the ring. This technique has been greatly developed since then.

In the late 1980s a new optics control technique emerged using closed orbits excited by horizontal and vertical orbit correctors [18–22]. The optics model of the machine was numerically adjusted to reproduce the measured closed orbit response matrix. Successful optics corrections based on these measurements were demonstrated for the first time.
at SPEAR in 1993 [21] and soon after at the NSLS and ALS [22]. This technique is referred to as the orbit response matrix (ORM) method in the following.

A third class of accelerator optics control may be introduced by citing the first sentence of [23] (1991): “For future linear colliders, […] with demanding tolerances on final focus system alignment and magnet errors, it becomes increasingly important to use the beam as a diagnostic tool.” Extrapolating to any accelerator, a beam-based optimization of machine-performance-related observables is a universal approach for the mitigation of lattice imperfections. This technique can be considered as noninvasive when the required size of the perturbations might be tolerated during machine operation. We refer to this set of techniques as optimization methods.

These first realizations of the techniques presented above (K modulation, turn-by-turn, ORM and optimization methods) occurred between 1975 and 1993, setting the ground for a new discipline: “Optics measurement and correction in charged particle accelerators.” The materialization of this discipline came a decade later with the publication of a book [24]. The development of the optics measurement and correction techniques is illustrated by the evolution of the β-beating over time for many circular accelerators, see Fig. 1. The measured, or inferred, β-beating values have been extracted from the bibliography in this paper. The accuracy demonstrated through comparisons of different measurement techniques is marked with a dashed line at the 1% level as it is discussed in Sec. VI. Similarly Fig. 2 shows how the coupling control has improved over time. Here, the closest tune approach in hadron machines and the emittance ratio $\epsilon_y/\epsilon_x$ in lepton rings are used as figures of merit for the coupling.

The following section gives an introduction to the theory of optics perturbations. K modulation and turn-by-turn optics measurement techniques are described in Secs. III and IV with a general approach to correction given in Sec. IV E. The ORM technique is described in Sec. V. Section VI reports on recent comparisons between turn-by-turn and ORM techniques. Optimization techniques are described in Sec. VII.

II. OPTICS PERTURBATIONS FROM LATTICE IMPERFECTIONS

Quadrupolar gradient errors, $\delta k(s)$, distort accelerator optics parameters. To first order in these errors the β-beating at location $s$ can be expressed as

$$\frac{\Delta \beta}{\beta}(s) = \pm \frac{1}{2} \sin(2\pi Q) \int ds' \beta_{(\text{theo})}(s') \delta k(s') \times \cos[2(\phi_{(\text{theo})}(s) - \phi_{(\text{theo})}(s')) - 2\pi Q],$$

where $Q$ is the betatron tune, $\beta_{(\text{theo})}$ and $\phi_{(\text{theo})}$ are the design beta function and phase advance, respectively, and the sign is positive for the horizontal plane and negative for the vertical plane. The phase beating, $\Delta \phi(s_0, s)$, between the locations $s_0$ and $s$ can be derived from the β-beating using the relation between $\phi$ and $\beta$,

$$\Delta \phi(s_0, s) = \int_{s_0}^{s} ds' \frac{1}{\beta_{(\text{theo})}} \left( 1 + \frac{\Delta \beta}{\beta} - 1 \right).$$

First order expansions on $\delta k$ can be found in [25–27].

Higher order expressions of the β-beating can be found in [28] using an infinite determinant method and in [27] using the resonance driving terms (RDTs) $f_{2000}$ and $f_{0020}$ as

$$\frac{\Delta \beta_s}{\beta_s}(s) = 2 \sinh(4|f_{2000}|) \left[ \sinh(4|f_{2000}|) \right]$$

$$+ \cosh(4|f_{2000}|) \sin q_{2000},$$

FIG. 1. Measured or inferred β-beating versus time for many circular accelerators as found in the bibliography of this paper. Three stages are differentiated: (i) commissioning, when magnet powering mistakes are expected, (ii) after fixing these mistakes but before careful optics corrections and (iii) after final optics corrections.

FIG. 2. Coupling evolution versus time for many circular accelerators as found in the bibliography of this paper. The closest tune approach in hadron machines and the emittance ratio $\epsilon_y/\epsilon_x$ in lepton rings are used as figures of merit for the coupling.
where \( q_{2000} \) represents the phase of the RDT at location \( s \). An equivalent relation applies to the vertical plane, after replacing \( f_{2000} \) with \( f_{0020} \). Reference [27] also shows how Eq. (1) descends from Eq. (3) after a series of approximations.

Coupling between the horizontal and vertical motion can be described by the off-diagonal blocks of one-turn or N-turn maps reconstructed from BPM data [29–31] or by two RDTs, \( f_{1001} \) and \( f_{1010} \), corresponding to the difference and sum resonances, respectively. The equivalence between these two descriptions was established in [32].

The main coupling figure of merit in lepton machines is represented by the vertical emittance divided by the horizontal emittance, \( \epsilon_y/\epsilon_x \), which grows with vertical dispersion and coupling RDTs [33]. Vertical emittance can be either measured by beam profile monitors or inferred from the radiation emitted from vertical undulators [34].

Betatron coupling determines a stop band around the difference resonance \( Q_x - Q_y = M \), where \( M \) is any integer. This implies that the fractional tunes cannot get closer than \( \Delta Q_{\text{min}} \), the closest tune approach, given by [24]

\[
\Delta Q_{\text{min}} = \left| \frac{1}{2\pi} \int ds j(s) \sqrt{\beta_x\beta_y e^{-i(\phi_x - \phi_y) + i(\dot{\phi}_x - \dot{\phi}_y)s/R}} \right| .
\]

where \( j(s) \) is the skew quadrupolar gradient around the ring, \( R \) is the machine radius and \( \dot{Q}_{x,y} \) are the fractional tunes. \( \Delta Q_{\text{min}} \) can also be computed from the difference RDT \( f_{1001} \) around the ring by [35,36]

\[
\Delta Q_{\text{min}} = \left| \frac{4(\dot{Q}_x - \dot{Q}_y)}{2\pi R} \int ds f_{1001} e^{-i(\phi_x - \phi_y) + i(\dot{Q}_x - \dot{Q}_y)s/R} \right| .
\]

### III. K MODULATION

In [12] an elegant approximation for the average \( \beta \) in a quadrupole is derived taking into account only the beta functions at the magnet edges (\( \beta_1, \beta_2 \)) and the length of the quadrupole (\( L \)),

\[ \bar{\beta} \approx \frac{1}{3} \left( \beta_1 + \beta_2 + \sqrt{\beta_1 \beta_2 - L^2} \right). \]

The exact calculation of \( \bar{\beta} \) requires the quadrupole strength leading to a considerably longer equation [37]. Changing the normalized integrated gradient of a quadrupole by \( \delta K \) results in a tune change of \( \delta Q_{x,y} \) that can be used to measure \( \bar{\beta} \) as given by

\[
\bar{\beta}_{x,y} = \pm \frac{2}{\delta K} \left\{ \cot(2\pi Q_{x,y}) \left[ 1 - \cos(2\pi \delta Q_{x,y}) \right] + \sin(2\pi \delta Q_{x,y}) \right\} \\
\approx \pm 4\pi \frac{\delta Q_{x,y}}{\delta K}.
\]

\( K \) modulation has been successfully used to measure average beta functions in almost every accelerator, for example, ISR [12,17], LEP [38,39], HERA [40], RHIC [41,42], SLS [43], Tevatron [44], ALBA [45] and LHC [46,47].

This technique is limited by tune measurement resolution, residual betatron coupling, accuracy of the quadrupole integrated gradient versus current, quadrupole fringe fields, and unwanted tune shifts due to possible orbit changes at nonlinear magnets during the quadrupole modulation. Magnetic hysteresis effects may also require attention to recover the same tune value before and after the quadrupole modulation. Since the \( K \) modulation requires varying individual quadrupoles, the procedure can be very time consuming.

Furthermore, the tune measurement is limited by the fluctuation of the magnet power supplies. An example is shown in Fig. 3 from [48], where a tune spread of \( 2 \times 10^{-6} \) among the 120 BPMs in one single acquisition is observed at the ALBA storage ring, whereas the spread over 100 acquisitions increases to about \( 10^{-4} \). This spread is compatible with the noise in the power supplies, which is specified to \( 0.01\% \). This tune jitter affects the precision of the \( K \) modulation technique.

Recent measurements at SLS and ALBA obtained an rms statistical error in the \( 1\%–2\% \) level [43,45] with a similar systematic error. In the Duke storage ring a \( 1\% \) statistical resolution is demonstrated [49]. In hadron colliders, the \( K \) modulation technique is mostly used to infer interaction point (IP) \( \beta^* \) functions from nearby quadrupoles. Nevertheless, simulations for the HL-LHC [37] show that the pushed interaction region optics challenges the ability of this technique to infer \( \beta^* \) values accurately.

In general, average beta functions from \( K \) modulation can be included in the optics correction scheme presented in Sec. IV E.

![FIG. 3. Tune jitter measurement at ALBA. In a single acquisition, the rms tune spread over the 120 BPMs (in red) is only \( 2 \times 10^{-6} \), but this increases to \( 10^{-4} \) when the measurement is extended over 100 acquisitions due to the power supplies noise [48].](image-url)
IV. TURN-BY-TURN TECHNIQUES

Free horizontal betatron oscillations sampled turn by turn at a given location along the accelerator, $s$, are described by

$$x(N, s) = \sqrt{2J_2}\beta_x(s) \cos[2\pi Q_x N + \phi_x(s) + \phi_y]$$

(8)

where $N$ is the turn number and $J_2$ is the horizontal action. In the presence of both normal and skew quadrupole error field the RDTs $f_{2000}$, $f_{1010}$ and $f_{1001}$ can be used to describe the motion up to first order by [50,51]

$$x(N, s) = \sqrt{\beta_x(\text{theo})(s)} \Re\{\sqrt{2J_2} e^{i(2\pi Q_x N + \phi_x(s) + \phi_y)}$$

$$- 4if_{2000} \sqrt{2J_2} e^{-i(2\pi Q_x N + \phi_x(s) + \phi_y)}$$

$$- 2if_{1010} \sqrt{2J_2} e^{-i(2\pi Q_x N + \phi_x(s) + \phi_y)}$$

$$- 2if_{1001} \sqrt{2J_2} e^{i(2\pi Q_x N + \phi_x(s) + \phi_y)}\}.$$  

(9)

Equations to higher order in the RDTs can be found in [27,33]. Turn-by-turn beam position data contain all the information concerning the linear optics (as long as the oscillation amplitude is kept sufficiently low, so to avoid mixing with nonlinear terms [27]). In the following sections typical techniques used to excite betatron oscillations and thereby extract and correct optics parameters from the data are presented.

A. Excitation of betatron oscillations

The quality of an optics measurement depends, among other aspects, on the type of excitation used to induce betatron oscillations around the closed orbit. In lepton rings this is traditionally done with a fast kicker (whose pulse length is usually less than one revolution period) and measurements are performed while the particle beam distribution decoheres in phase space. The beam emittance is eventually restored by radiation damping. Analysis based on the digital Fourier transform, singular value decomposition, or independent component analysis is then performed on the recorded data to extract the tune spectral lines and the other harmonics of Eq. (9). Both decoherence and radiation damping modify the characteristics of the spectral modes of the turn-by-turn data and set limits on the achievable spectral resolution. Studies at SPS showed how analytic decoherence factors are sufficient to remove the effects of decoherence on the amplitude of spectral lines [52,53]. Analytical expressions derived in [54] estimate the impact of decoherence on the phase of the fast Fourier transform (FFT) main spectral line.

To reduce the decoherence from chromaticity and amplitude detuning special optics may be used. At ALBA, for example, the vertical normalized chromaticity during normal operation is large ($\sim$4) to fight against beam instabilities. Therefore, linear optics using turn-by-turn methods are performed by changing the machine lattice from the operational one to a lattice with small ($\sim$0) normalized chromaticity [48,55].

In hadron machines single-kick excitations lead to irreversible transverse emittance blowup. Important progress occurred in 1998 when ac dipoles were proposed to excite (forced) nondestructive coherent betatron oscillations [56]. The adiabatic excitation of the ac dipole minimizes the emittance growth after the measurement [57]. A first direct application to optics measurements was demonstrated in RHIC [58]. In [59] it is shown analytically how both linear and nonlinear RDTs are modified by an ac dipole excitation with experimental results obtained in RHIC [51]. In the Tevatron [60] it was demonstrated how the excitation generated by an ac dipole introduces a perturbation of the linear optics equivalent to a quadrupole-like error at the same location with a gradient depending only on the betatron tune and the ac dipole frequency.

A comparison between forced ac-dipole oscillations and free motion after a single kick in the LHC is shown in Fig. 4. Here, the ac dipole was kept at a constant value for 6600 turns, during which the BPM data were taken.

The ac dipole technique has been essential to commission the lattice optics in the LHC [62,63], where 20 measurements had to be analyzed and corrected within tolerance in the shortest possible time during the energy ramp and the $\beta^*$ squeeze. The use of destructive single kicks would have required an unaffordable number of cycles. The ac dipole technique also has applications beyond linear optics measurements, extending into the evaluation of collective effects and nonlinear dynamics [26,64–66].

For coupling measurements the interpretation of ac dipole measurements is less intuitive. The corresponding analytical equations are derived in [67]. Coupling measurements can be repeated off momentum to measure chromatic coupling with or without the ac dipole [68,69].

![FIG. 4. Comparison between the vertical beam response (top) to a single kick in the LHC at injection energy, and to an ac dipole excitation (bottom) [61].](image-url)
B. Analysis techniques

Two fundamental processing techniques are commonly applied to the analysis of turn-by-turn BPM data: the Fourier transform (FT) and matrix decomposition methods, such as singular value decomposition (SVD) or independent component analysis (ICA).

The FT is applied independently to each BPM to extract frequency, amplitude and phase of the spectral modes of the recorded motion. Refined FTs as NAFF [70] and Sussix [71] are more accurate in the measurement of main frequencies than the regular fast Fourier transform (FFT). However, for the evaluation of the phase advance between 2 BPMs the FFT is preferred [26], due to the fact that the tunes measured at different BPMs fall into the same FFT bin. Alternatively, after accurately measuring the main tune with NAFF or Sussix at all BPMs, the phases can be computed via an FT by looking at the spectral line corresponding to the average tune.

As far as the matrix decomposition approach is concerned, a $P \times N$ matrix can be constructed from the data measured at $P$ BPMs over $N$ turns. SVD analysis of this matrix can be performed to detect the singular modes appearing in the motion. This analysis was successfully applied for the first time to the Stanford linear collider (SLC) using single-pass BPM data [72]. In order to overcome numerical problems of the SVD in the presence of coupled eigenmodes, a different decomposition approach based on ICA was proposed and applied for the first time to optics measurements in the Fermilab Booster [73]. Since then, both SVD and ICA were applied in many circular machines [74–81].

Large-scale BPM systems inevitably contain malfunctioning units, which can manifest in very subtle ways, such as, e.g., "ghost data" appearing intermittently on random turns [40,52] or BPMs featuring synchronization problems. It is important to identify and remove these faulty units to avoid nonphysical measurements and erroneous corrections. Efficient ways to detect defective BPMs were developed for both the SLC [72] and SPS [52], based on SVD and FT decompositions, respectively. A comparison of these techniques was performed later at RHIC [82]. Figure 5 illustrates typical U-vectors extracted from the SVD of the $P \times N$ matrix $R = USV^T$ which show both fundamental betatron motion and uncorrelated signals which originate from malfunctioning BPMs.

In order to extract the most meaningful information from the two techniques, it is beneficial to first condition, or clean, the BPM data matrix $R$ with SVD by filtering the singular value spectrum and then to apply FT analysis to the result.

C. Measurement of beta functions

Early measurements of beta functions from turn-by-turn BPM signals were performed in the ISR [17] in 1983. The beta functions were inferred from the amplitude of induced betatron oscillations of the stored beam. This technique is referred to as “β from amplitude” and its weakest point is the unknown scale factor in the BPM gains. For instance, if a BPM has a linear scale error of, e.g., 5% rms, the resulting uncertainty in the calculated $\beta$ is of 10%. In 1988 turn-by-turn optics measurements in LEAR [83] focused on the betatron phase advance since the BPM gain calibration was unknown. Another source of error in the $\beta$-from-amplitude approach comes from the need to normalize the measured $\beta$ to the average $\beta$ from the model. Indeed, the as-built, perturbed lattice features an average beta function which tends to increase with the rms $\beta$-beating [27,84]. An optics-based measurement of the BPM calibration factors has been recently demonstrated in the LHC [85,86] by switching off the...
quadrupoles in the interaction region and utilizing the parabolic behavior of the beta function in a drift space. BPM gain calibration uncertainties of about 0.5% were obtained. Recently, similar $\beta$-from-amplitude measurements have demonstrated satisfactory results both in ALBA [55,87,88] and RHIC [81].

To avoid the aforementioned limitations of the $\beta$-from-amplitude method, beta functions were measured at LEP in 1993 by using the ideal and measured phase advance between three neighboring BPMs via the following equation [89,90]:

$$
\beta_{11}(\text{exp}) = \beta_{11}(\text{theo}) \frac{\cot \phi_{12}(\text{exp}) - \cot \phi_{13}(\text{exp})}{\cot \phi_{12}(\text{theo}) - \cot \phi_{13}(\text{theo})},
$$

(10)

where the integer numbers refer to the BPM index, while the subscripts “theo” and “exp” denote the values from the model and the experimental measurement, respectively. Equation (10) has the great advantage of being independent of any BPM calibration and tilt errors and of being insensitive to betatron coupling to first order. Nevertheless it assumes perfect timing synchronization and no optics errors between the 3 BPMs.

This method, named “$\beta$ from phase,” was also applied to CESR [30] in 2000, revealing $\Delta \beta / \beta$ errors at the 100% level before correction. The CESR optics correction was based on fitting a model to the measured phases, ultimately reaching a $\Delta \beta / \beta$ of 2% rms, thanks to the availability of independent quadrupole power supplies.

The 3-BPM method developed at LEP has been recently extended to consider any number $N$ of BPMs [91], hence becoming the N-BPM method: This approach considerably increases the measurement resolution, as the experimental results in the LHC have shown (see Fig. 6). In order for the N-BPM method to work correctly, knowledge of the uncertainties in the optics model is fundamental. This method has been also applied in ALBA [55,88] and ESRF [92].

A recent study of the systematic errors involved in the beta function measurements from turn-by-turn data is presented in [27]. An extension of Eq. (10) is derived to first order in quadrupolar gradient errors, $\delta k$, between the BPMs:

$$
\beta_{11}(\text{exp}) = \beta_{11}(\text{theo}) \frac{\cot \Delta \phi_{12}(\text{exp}) - \cot \Delta \phi_{13}(\text{exp})}{\cot \Delta \phi_{12}(\text{theo}) - \cot \Delta \phi_{13}(\text{theo})} + (h_{12} - h_{13}),
$$

(11)

$\tilde{h}_{ij} = \mp \frac{1}{\sin^2 \Delta \phi_{j}(\text{theo})} \int_{s_i}^{s_j} ds \beta_{1}(\text{theo})(s) \delta k(s)

\times \sin^2 [\phi_{i}(\text{theo})(s) - \phi_{i}(\text{theo})(s)],

(12)

where $s_i$ is the longitudinal location of the $i$th BPM. The sign on the right-hand side of Eq. (12) is negative in the horizontal plane and positive for the analysis in the vertical plane. Known quadrupole gradient uncertainties from magnetic measurements can be incorporated in Eq. (11) to estimate systematic errors in the beta function measurements. This approach should speed up the implementation of the N-BPM method, which currently includes systematic errors derived from Monte Carlo simulations. Analytical formulas are also given in [27] to account for the perturbations to phase advance and coupling measurements from nonlinear magnetic elements. These effects are known to limit the accuracy of turn-by-turn optics measurements at the ESRF storage ring operating in ultralow betatron coupling mode.

D. Measurement of coupling

Coupling RDTs are measured from the spectral lines detected at dual-plane BPMs via the following equations [52,93]:

$$
|f_{1001}| = \frac{1}{2} \sqrt{\frac{H(0,1)V(1,0)}{H(1,0)V(0,1)}},
$$

$$
|f_{1010}| = \frac{1}{2} \sqrt{\frac{H(0,-1)V(-1,0)}{H(1,0)V(0,1)}},
$$

(13)

where $H(1,0)$ and $V(0,1)$ are the amplitudes of the betatron tune lines in the horizontal and vertical planes, respectively, whereas $H(0, \pm 1)$ and $V(\pm 1,0)$ indicate the coupling harmonics. The phase of the two RDTs is also measurable [93]. For hadron machines, and any circular accelerator with the transverse tunes having similar integer parts, $|f_{1010}| \ll |f_{1001}|$ and betatron coupling is well described by $f_{1001}$ only. Third generation light sources and damping rings with stronger horizontal focusing
require the inclusion of the sum RDT $f_{1010}$ to fully describe coupling [33,94].

For the “closest tune” approach, $\Delta Q_{\text{min}}$ is measured by monitoring the tunes (or the transverse emittances [95]) while varying quadrupole field strengths to cross the difference resonance $Q_v - Q_x = M$, where $M$ is the difference between the integer parts of the betatron tunes. This technique has been used since the first synchrotrons [13,14], with an illustrative example from the CERN PS shown in Fig. 7. If BPMs with turn-by-turn capabilities are available, $\Delta Q_{\text{min}}$ can be approximated from Eq. (5) by the following averages involving the RDT $f_{1001}$ measured at the BPMs:

$$\Delta Q_{\text{min}} \approx |4(\hat{Q}_x - \hat{Q}_y) \frac{1}{f_{1001}} e^{-i(\phi_x - \phi_y)}| \approx 4|\hat{Q}_x - \hat{Q}_y||f_{1001}|.$$  \hfill (14)

### E. Optics corrections

A powerful numerical approach to linear optics correction is to minimize the deviation between measured and model phase advance between adjacent BPMs. According to Eq. (11) this technique minimizes the $\beta$-beating when the $\delta h_{ij}$ terms are negligible or when the correction of the BPM phase advance error is accompanied by the minimization in the nonmeasurable $\delta h_{ij}$ terms (as in a local correction). LHC simulations, for instance, have shown both rms and peak $\beta$-beatings are linearly related to the rms and peak phase advance errors, respectively [96]. A response matrix describing the shift in the BPM phase advance induced by independent quadrupole perturbations can be computed directly from the computer model and then pseudoinverted to solve the following system:

$$\begin{pmatrix} \delta \Delta \tilde{\phi}_x \\ \delta \Delta \tilde{\phi}_y \\ \delta \tilde{D}_x \\ \delta \tilde{Q} \end{pmatrix} = \mathbf{P}_{(\text{theo})} \cdot \delta \vec{k},$$  \hfill (15)

where $\mathbf{P}_{(\text{theo})}$ denotes the response matrix, $\delta \vec{k}$ represents the vector containing the available quadrupoles, and $\delta \Delta \tilde{\phi}$ contains the phase advance error between BPM pairs, $\delta \tilde{D}_x$ is the vector containing the dispersion errors measured at the BPMs, or its equivalent calibration-independent form $D_x/\sqrt{\beta_x}$ [84]. The vector $\delta \tilde{Q}$ contains the tune shifts. The $\beta$-beating measured either from the tune line amplitude (Sec. IV C) or by $K$ modulation (Sec. III) can also be included in the above system for a more accurate correction as done in LHC [97]. The use of appropriate weights in Eq. (15) is fundamental to produce accurate results. These weighting factors have been omitted here for simplicity. Optimization of the optics correction, together with the actual variable set and corresponding weights can be normally investigated using model-based computer simulations, though it is eventually achieved heuristically with experimental measurements.

FIG. 7. Example of closest tune approach applied to the CERN PS after powering a strong skew quadrupole [95]. Top: The model betatron tunes programmed to cross the difference resonance. Bottom: Corresponding measured tune values. $\Delta Q_{\text{min}}$ is inferred from the closest approach between the tunes at the crossover point.
(\begin{pmatrix} \tilde{f}_{1001} \\ \tilde{f}_{1010} \\ \tilde{D}_\gamma \end{pmatrix} = T_{\text{theo}} \cdot \delta \tilde{f}, \tag{16}

where $T_{\text{theo}}$ denotes the response matrix, $\delta \tilde{f}$ represents the vector containing the skew quadrupoles available for correction, and $\tilde{D}_\gamma$ is the vertical dispersion measured at the BPMs. Similarly to the system of Eq. (15), weights are to be included to obtain efficient solutions. In hadron machines with equal integer tunes one family of skew quadrupoles is enough to control $\Delta Q_{\text{min}}$, whereas, in general, two orthogonal families are required. $\Delta Q_{\text{min}}$ values as low as $2 \times 10^{-4}$ have been reported in [101].

V. CLOSED ORBIT RESPONSE MATRICES TECHNIQUES

For circular accelerators the closed orbit distortion (COD) $\delta r(s)$ due to a dipolar kick of angle $\theta$ at location $s_c$ is given by

$$\delta r(s) = \frac{\sqrt{\beta(s)} \theta \sqrt{\beta(s_c)} \cos |\phi(s) - \phi(s_c)| - \pi Q|}{2 \sin (\pi Q)},$$

when dispersion effects are neglected. The distortion can be purposefully introduced by a dipole corrector. The orbit response matrix (ORM) is measured by changing the corrector excitation one by one and recording the corresponding orbit shifts from the BPM readings. Of significance, the ORM contains detailed information concerning the spatial field structure of magnets in the accelerator lattice. Since the information is recorded under static orbit conditions, the BPM data has a high degree of resolution.

Starting in the 1980s purposely introduced COD was used to measure beta functions assuming that the BPM was close enough to the corrector so their $\beta$ and $\phi$ values were identical, leaving only one unknown, $\beta(s_c)$, on the right-hand side of Eq. (17). This was referred to as the “cusp” technique [19].

These early approaches were later extended to include the orbit response at all available BPMs using two or more orbit correctors [19,20]. $\beta$ and $\phi$ were found at every BPM by fitting Eq. (17) to the measured ORM. This technique was successfully used at KEKB and SuperKEKB with a reduced ORM (six orbit correctors) [102,103]. It is noted that the beta functions obtained this way are directly influenced by BPM gain errors.

This limitation was mitigated by using a full ORM and fitting parameters in the optics model directly rather than using the analytical COD of Eq. (17). A comprehensive parameter set including quadrupole gradients, BPM gain errors and orbit corrector strengths is used as proposed in [21]. The fitting parameters can be extended to include a range of magnet errors (for example, longitudinal positions of correctors taking into account magnet cross-talk effects [22]). The design optics model is often taken as the starting point for the ORM fitting procedure. Conversely, a modified optics model can be used to establish a new machine optics other than the design optics. A realistic model of the accelerator can be found by numerically fitting component parameters in the accelerator model such that the model ORM reproduces the measured one [104]. The accuracy of the resulting accelerator model parameters depends on residual differences between the model and the accelerator, which are not included as fitting parameters.

The beta functions of the machine are inferred from the accelerator model. Optics corrections can be performed by applying the differences in quadrupole gradients found through the fitting $\delta k$ with opposite sign to the machine. The removal of all inferred quadrupole errors would then correct the machine $\beta$-beating. Several iterations are typically required for convergence.

The ORM approach was implemented into the computer program LOCO (in FORTRAN), and it was successful to correct the optics parameters of the NSLS vacuum ultra-violet (VUV) ring and also ALS [22]. Since then LOCO has been applied to many accelerators worldwide during the initial commissioning phase, to control accelerator optics during machine development and to maintain optimal conditions for standard machine operations.

Variations on the ORM method include cases where not all the quadrupoles are individually powered, i.e. quadrupoles are grouped in families, and/or a smaller set of quadrupole correctors is available (e.g. only dedicated quadrupole correctors are used). In these cases, the system of linear optics equations is reduced in size to include only available correctors and again is inverted to minimize the difference between model and measurement. For instance, correctors can be powered to reduce the difference between the measured and model diagonal blocks of the ORM (together with horizontal dispersion). Alternatively, the $\beta$-beating computed from the accelerator model fitted with errors at all individual quadrupoles can be put into the left-hand side of Eq. (1), whose linear system can be pseudoinverted by using only the available quadrupole correctors. At the ESRF another approach is routinely followed. Focusing errors over all 256 quadrupoles (grouped in six families) are first inferred by fitting the measured diagonal ORM blocks, and then used to compute the focusing RDTs $f_{2000}$ and $f_{0020}$ at the BPMs. The 32 quadrupole correctors are then set to minimize these RDTs along with the dispersion beating and the tune shifts. It is worthwhile noticing that the minimization of the above RDTs ensures the correction of the beta functions, according to Eq. (3), though it does not guarantee the minimization of the phase advance errors.
Correction of betatron coupling (along with vertical dispersion) has been carried out in three different ways. In the first technique, skew quadrupoles are used to minimize the off-diagonal blocks of the measured ORM. Alternatively, the fitted model can be used to compute a skew quadrupole correction by using the model value for vertical emittance as a figure of merit. Finally, the fitted lattice model can be used to evaluate the coupling RDTs \(f_{1001} \) and \(f_{1010} \) around the ring and the system described by Eq. (16) can be pseudoinverted to generate a solution. The minimization of RDTs and \(D_y \) ensures a simultaneous abatement of the vertical emittance [33].

In 2002, the Fortran version of LOCO was rewritten in MATLAB [107,108] using the MATLAB ACCELERATOR Toolbox [109] for lattice calculations and incorporating a graphical user interface to ease the data analysis. In recent years LOCO (in MATLAB) has become a workhorse for optics corrections of storage ring optics in modern light sources such as Diamond [110], SOLEIL [111] and ALBA [112] as well as the CLS, SSRF, Australian Synchrotron, TLS-II, NSLS-II, SPEAR3 and the PLS.

Several fitting algorithms are now available including Gauss-Newton, (Scaled) Levenberg-Marquardt, and constrained fitting. The last one is used to alleviate degeneracy problems, e.g., when two or more quadrupoles are situated between adjacent BPMs. Third generation light sources often face this problem, and the quadrupole gradient errors found from LOCO fitting tend to be too vigorous [108] when a simple unconstrained fitting algorithm is employed. The use of SVD with a proper cut of the eigenvalue spectrum can reduce degeneracy problems provided the initial rms \(\beta \)-beating is not larger than \(\sim 10\% \) [43]. It is noted that the dispersion function is separately measured and included in the fitting procedure in these applications. In this way, the dispersion errors are also minimized. The ORM method usually produces an accelerator model with an rms \(\beta \)-beating at the 1% level or below with accurate betatron tune correction.

Due to the large number of measured data points, the LOCO fit is quite robust against statistical measurement errors. For the NSLS VUV ring, for instance, ten repeated measured ORMs were independently analyzed, and the random uncertainty of the inferred beta functions was only 0.1%. Important computational details for the success of the ORM approach are found in [22,108]. These include the following.

(i) In electron storage rings, the beam energy shifts due to corrector excitations must be also included as fitting parameters. The orbit distortion at BPMs situated in dispersive section is due not only to the corrector kick but also to the energy shift.

(ii) Weighting factors taking into account the measurement noise at each BPM can be included in the evaluation of \(\chi^2 \) of the fit.

(iii) When the measured ORM is not reproduced within the measurement noise level, it is often an indication of systematic errors (e.g. magnet cross talk as previously discussed). Corresponding fitting parameters must be included to improve the fit if identified.

(iv) Corrector excitations used to measure ORM must be adjusted to reach a trade-off between the BPM signal-to-noise ratio and the lattice nonlinearities.

(v) A lower number of fit parameters tends to increase numerical accuracy of the fitted parameters. When only the beta functions are corrected, the coupling elements of ORM can be dropped and the related parameters are omitted in the fit.

Nevertheless, even when the LOCO fit is successful, the measured ORM after optics correction may not converge towards the model ORM of the design optics. This was observed at the SLS [43] and illustrates limitations of the parametrized lattice model to represent the real machine. Therefore, the uncertainty of the inferred \(\beta \)-beating should be comparable to the inferred \(\beta \)-beating when the convergence is not satisfactory. The quality of the fit and the convergence of the measured and model ORMs have to be carefully checked. Applying ORM to localized sections of the ring proved successful to detect possible locations of the differences between the model and the machine in SLS [113].

Depending on the size of the accelerator, ORM measurement can be a lengthy procedure as it involves varying the corrector excitation current one by one while recording the closed-orbit response. A faster measurement method that takes about one minute using the fast orbit feedback network is under development at the Diamond light source [114].

In small and medium-size hadron machines ORM techniques have proved successful only in estimating the \(\beta \)-beating [115–118] while in large colliders first attempts to use LOCO lead to impractical [119] or unrealistic corrections [120].

Techniques based on closed orbit bumps were successful in identifying gradient errors in Tristan and RHIC [121,122]. Another technique that has been applied to excite closed orbits is the already mentioned action and phase analysis [123]. In this case, a single quadrupole error could be identified with excellent accuracy [124].

VI. TURN-BY-TURN AND ORM COMPARISONS

Large efforts are under way to compare turn-by-turn and ORM methods for measurement and correction of accelerators optics. Turn-by-turn measurements are potentially faster and less invasive than ORM measurements. However, first attempts to perform turn-by-turn measurements in modern light sources like SLS or Soleil initially faced important limitations when the BPM electronics is switched to turn-by-turn acquisition mode [43,125]. Once
these limitations were overcome, the measured beta functions showed an agreement between turn-by-turn and ORM results at about the 1% rms level in four different synchrotrons [55,87,88,92,126,127]. This 1% can be regarded as the experimentally achieved accuracy. Nevertheless there are ongoing efforts in the ESRF [128] addressing the intrinsic accuracy inferred from simulations (systematic error) and the reproducibility over several acquisitions (statistical error). The turn-by-turn N-BPM method shows a precision in both planes of about 0.4%. The ORM technique features a precision of 0.6% in the horizontal plane and 0.4% in the vertical plane. Orbit and energy stability, along with the power supply noise, represent the main culprits limiting the measurement reproducibility, because of the quadrupolar feed-down generated by off-centered orbit across sextupoles. Simulations of the ESRF storage ring show that a rms closed-orbit distortion of 24 μm would induce a change in the β-beating of about 0.5% rms.

The performances of BPM electronics in light sources are constantly being improved [129]: Latest developments not only include turn-by-turn acquisition modes, but also long buffer memories with subturn resolutions [130] which pave the way for faster and more precise accelerator physics studies, such as tune shift with amplitude or localized impedance measurements [131].

Concerning coupling measurements, a direct comparison of the two techniques is presented in [126]. In this work the betatron coupling measured with both techniques yields similar average emittance ratios at the 1% level (before correction) and 0.3% (after correction). However significant discrepancies appear between the two techniques for the projected emittance ratios around the ring. In [27], it is estimated that turn-by-turn measurement techniques with the present generation of BPM electronics are challenged for emittance ratios at per-mill level and below.

VII. OPTIMIZATION METHODS

Since the commissioning of the first accelerators performance is optimized by scanning available parameters. An illustrating example of linear optics correction is found in the linear collider SLC [132], where orthogonal multiparameter knobs were developed to counteract phase space aberrations at the IP. The knobs were scanned until a minimum beam size was found with the help of a parabolic fit to mitigate measurement errors. Similar techniques are currently in use in final focus systems (FFS) such as ATF2 [133] and foreseen at future linear colliders. Other optimization algorithms have been also frequently used, such as the Simplex at the KEKB injector linac since 1998 [134]. Simulations show that the Simplex is also needed in the CLIC FFS to achieve acceptable performance [135]. In circular colliders the luminosity is typically maximized using multivariate optimization algorithms over long operational periods [136,137].

Recently, optimization techniques have been also applied in several storage ring light sources. Figure 8 shows how the vertical beam size at SLS was reduced by about 25% applying the random-walk optimization over a period of 1 h, demonstrating unprecedented levels of coupling minimization with a record low vertical emittance of 0.9 ± 0.4 pm [138]. Other applications of optimization techniques in light sources can be found in [139,140] setting a trend towards embedding optimization tools in accelerator operation.

VIII. SUMMARY AND OUTLOOK

Beam linear optics, understood as the arrangement of bending and focusing elements, is one of the fundamental pillars of modern accelerators. Machine performance and protection aspects rest upon linear optics parameters. Demands for even higher performances in modern accelerators have boosted the “optics measurement and correction” to grow into a discipline of its own. Table I summarizes the various techniques following the classification used above. The main challenge faced by all accelerators and measurement techniques is the required machine time. Conceptually optics correction could be as fast as orbit correction. First steps in this direction have been done in LHC [36,141] and NSLS-II [126] (turn by turn) and at the Diamond light source [114] (ORM). Turn-by-turn techniques are potentially the fastest but require a BPM system capable of quickly switching between slow and fast acquisition modes. Decoherence, radiation damping and nonlinearities introduce important limitations in the turn-by-turn optics measurements with single kick excitation. In principle, ac dipoles overcome all these limitations by exciting long-lasting and small betatron oscillations.

Large experimental programs have demonstrated a 1% accuracy in the beta function measurement from the various techniques. Equivalent comparative studies are still
required for coupling. Future projects, like HL-LHC, SuperKEKB, FCC, ESRF upgrade, MAX-IV, SLS-II, etc., will continue challenging optics control techniques in terms of accuracy, resolution, speed and instrumentation. In particular, diffraction-limited storage rings require design compromises between momentum aperture and tune footprint which could lead to half integer resonance crossing. For example, in HEPS the rms horizontal $\beta$-beating should be kept below 1.5% [142]. Developments in other disciplines like collective effects such as impedance, space charge and Touschek leading to particle loss also require improving the measurement and control of linear optics [48,143–147].

ACKNOWLEDGMENTS

We are extremely thankful to G. Benedetti, J. Cardona, M. Carla, F. Carlier, J. Coello, A. Garcia-Tabares, M. Giovannozzi, W. Guo, A. Langner, L. Malina, E.H. Maclean, Z. Marti, L. Nadolski, T.H.B. Persson, P. Skowronski and S. White for fruitful discussions and comments on the manuscript. This review paper has largely benefited from discussions in two workshops: Advanced Optics Control [148] and Beam Dynamics meets Diagnostics [149]. Special thanks for the organization of these workshops go to M. Bai, G. Franchetti, M. Giovannozzi, M. Lamont and F. Zimmermann.


---

<table>
<thead>
<tr>
<th>Excitation</th>
<th>Observable</th>
<th>Analysis</th>
<th>Parameter</th>
<th>Depends on</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Betatron oscillation free or forced</td>
<td>Centroid position turn-by-turn</td>
<td>FT, SVD, fit</td>
<td>$\phi$</td>
<td>$\cdots$</td>
<td>[17,30,43,63]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$\beta$ from $\phi$</td>
<td>M</td>
<td>[72,74,90,91]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$\beta$ from amplitude</td>
<td>C&amp;M</td>
<td>[27,52,81]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Action</td>
<td>C&amp;M</td>
<td>[59,60,67,98]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Coupling</td>
<td>C</td>
<td>[31,36,99,100]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>BPM calibration</td>
<td>C&amp;M</td>
<td>[85]</td>
</tr>
<tr>
<td>Betatron oscillation + RF freq</td>
<td>Centroid position turn-by-turn</td>
<td>FT, SVD, fit</td>
<td>$D_x/\sqrt{\beta}$</td>
<td>M</td>
<td>[84]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Chromatic coupling</td>
<td>C</td>
<td>[68,69]</td>
</tr>
<tr>
<td>Orbit distortion</td>
<td>Orbit</td>
<td>$\phi, \beta$ fit</td>
<td>$\phi, \beta$</td>
<td>C</td>
<td>[102]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Model fit</td>
<td>Any parameter</td>
<td>C&amp;M</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Fit</td>
<td>Arc Action</td>
<td>C&amp;M</td>
</tr>
<tr>
<td>Quadrupole gradient</td>
<td>Tune</td>
<td>Fit</td>
<td>$\Delta Q_{\text{min}}$</td>
<td>C</td>
<td>[12,17]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>On-line optimization of parameter</th>
<th>Depends on</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beam size</td>
<td>Coupling</td>
<td>$\cdots$</td>
</tr>
<tr>
<td>Loss rate</td>
<td>Dynamic aperture</td>
<td>$\cdots$</td>
</tr>
<tr>
<td>Luminosity</td>
<td>Integrated luminosity</td>
<td>$\cdots$</td>
</tr>
<tr>
<td>Lifetime</td>
<td>IP beam size</td>
<td>$\cdots$</td>
</tr>
</tbody>
</table>

---

TABLE I. Overview of measurement techniques. The meaning of acronyms and symbols follows. C: Calibration or tilt; FT: Fourier transform; M: Model; SVD: Singular value decomposition; $\phi, \beta$ and $D_x$: phase advance, beta function and dispersion; $\beta$: Average beta function over a quadrupole; $\Delta Q_{\text{min}}$: closest tune approach.


[61] Figure courtesy of Felix Carlier.


[104] The revised “model calibration” method for closed orbit analysis was a logical extension of the GOLD method originally developed for analysis of accelerator transport lines [105]. Recent development for the model calibration of transport lines is found in [106].


[148] Advanced Optics Control workshop, CERN, Switzerland, 2015, CERN.