

ORBIT CORRECTION USING VIRTUAL MONITORS AT JEFFERSON LAB*

Yu-Chiu Chao, Bruce Bowling, Johannes van Zeijts, William A. Watson III, Sue Witherspoon,
Thomas Jefferson National Accelerator Facility



Abstract

An orbit correction algorithm is developed to achieve the following goals for the CEBAF accelerator at Jefferson Lab: 1). Pre-processing of orbit input to account for estimated misalignment and monitor errors. 2). Automatic elimination of blind spots caused by response matrix degeneracy. 3). Transparency of exception handling to interchangeable generic steering engines. 4). CEBAF-specific demands on control of injection angle, path length, orbit effects on optics, simultaneous multiple pass steering, and orbit control at un-monitored locations. All of the above can be accomplished by the introduction of virtual monitors into the processed input orbit, whose theoretical basis is to be discussed in this report. Implementation of all or part of these features and operational experience during the CEBAF variable energy runs will also be discussed.

1 INTRODUCTION

At Jefferson Lab the particular optics configuration and operation constraints impose unique demands on the orbit correction program as elaborated in the abstract. These demands can be met with a universal orbit processing algorithm before specific steering engine is called and ideally the need for human intervention on exception handling is minimized. The algorithm interprets the underlying orbit from monitor and correction element data while performing estimates on misalignment and monitor errors by a well defined procedure. The resulting underlying orbit, including position, angle, path length and other orbit-dependent information, becomes source of steering constraints not only at locations of existing monitors, but at any location deemed necessary to have the orbit constrained or corrected, either due to need of blind spot elimination or due to explicit constraints on angle, path length etc. This enhanced set of orbit constraints, or *virtual monitors*, is fed into any generic steering engine with no a priori knowledge of either the specifics of the orbit control goals or the need of exception handling. In the following report this program is outlined and its theoretical basis discussed. Application at CEBAF is also described.

2 ORBIT CORRECTION ISSUES

2.1 The generic linear orbit correction problem

All linear orbit correction problems, regardless of the constraints, can be represented conceptually as

$$X_i = \sum_j M_{ij} \cdot C_j \quad (1)$$

where X_i is orbit error at the i -th BPM, C_j the strength of the j -th corrector, and M_{ij} the linear response linking X_i

and C_j . The core of any steering algorithm amounts to inverting or pseudo-inverting M such that the corrector changes ΔC_j can be found to cancel the orbit error X_i ,

$$\Delta C_j = \sum_i M_{ij}^+ \cdot X_i \quad (2)$$

with different algorithms having different emphases on the weighting scheme, response matrix near-singularity, corrector strength constraints, etc.

2.2 The more generalized orbit correction problem

For the most general orbit correction problem we generalize X_i to any of the coordinate errors including position, angle, path length, etc., as well as quantities derivable from orbit errors such as spurious dispersion caused by angle error at non-zero chromaticity. We can also generalize C_j to any orbit correction devices, as well as energy and devices having different effects on different beam passes. The response matrix elements M_{ij} therefore can be any of the linear transfer matrix elements or linearized second order transfer elements, possibly relating a single corrector to multiple pass beam coordinates. The generalized orbit errors need to be realized by quantities *not directly observable* by the BPM's. They serve as virtual BPM's when included in Eqn (1) with proper weighting.

2.3 Response matrix singularity

When the response matrix is near singular so that some linear combination of the correctors has almost undetectable response at all the monitors but large response elsewhere, naive correction schemes fail by producing excessive corrector strengths and undetected orbit bumps. Corrector reduction or strength constraint can solve this problem [1], as is done by many algorithms. It is not the optimal solution since usually a combination of correctors, not a single one, causes the problem. Indiscriminately limiting or eliminating correctors kills not only the offending combination, but also the useful ones, especially those needed for injection fixes. If however we impose a constraint on the orbit at a strategically chosen location which couples only to the offending combination, we would constrain the offending combination while allowing the useful ones to function. Such orbit constraints at arbitrary locations constitute another instance of the virtual monitors. A well defined algorithm to determine where these virtual monitors should be placed will be described later [1].

2.4 Transparency of singularity handling and other generalized constraints to generic steering algorithms

In light of the above discussion, the goal of the proposed algorithm is thus to construct the input generalized orbit error X_i , generalized corrector strength C_j , and generalized response matrix M_{ij} which, through the introduction of virtual monitors, take into account all the steering constraints described in section 2.2 and automatically remove singularities described in section 2.3. The input data for

* Work supported by the U.S. Department of Energy, contract DE-AC05-84ER40150.

subsequent steering algorithm contain both virtual and real orbit and response matrix, indistinguishable under the format of Eqn (1). They can be given to different steering algorithms with any level of singularity handling and a consistent singularity-free outcome conforming to all additional constraints should be expected.

2.5 Errors in input data

A major problem affecting orbit correction is the inherent errors in the input data. Without an effective filtering mechanism, such errors can distort the underlying physical picture and compromise the correction. In the current context such filtering mechanism is also critical for constraints on the virtual monitors since the latter can inherit errors from the physical data. The success of using virtual monitors thus depends on the ability to screen out errors in interpreted orbit data, to be discussed in the next section.

3 INTERPRETING THE UNDERLYING ORBIT AND ERRORS

We have to answer the following question: Given a set of *physically* monitored orbit errors, how does one derive the generalized orbit errors for the *virtual* monitors? A complete set of orbit information of all beam coordinates at *any* location can be known given the following: corrector strengths, monitor readings, monitor errors, field and alignment errors, and injection errors. The last 3 errors, being unobservable, have to be estimated from the first 2 and the model. The distinction between the injection and all the other misalignment type errors is artificial, depending on choice of starting point. We therefore focus on two types of errors: monitor error and alignment type error. The effect of all alignment type error on the orbit X_1 at the p -th monitors is

$$K^P = \delta X_1^P = \sum_a \sum_i M_{1i}^{ap} \cdot \delta X_i^a \quad (3)$$

where δX_i^a is the misalignment induced error in the i -th coordinate at location indexed a . M_{1i}^{ap} is the transfer matrix element linking such error to the p -th monitor. The effect a monitor error has on the *apparent* orbit at the p -th monitor is simply its offset at that monitor

$$\Delta^P = \delta X_1^P \quad (4)$$

Although not part of the real orbit, this error will be included as input to any orbit correction algorithm.

It can be shown that with all the above errors included, the best orbit correction that can be achieved by any algorithm represented by Eqn (1) is limited by the residual error E :

$$\begin{aligned} E &= \Pi_{M_{CM}}^\perp \cdot K - \Pi_{M_{CM}}^\parallel \cdot \Delta \\ \Pi_{M_{CM}}^\perp &= M_{CM} \cdot (M_{CM}^T \cdot M_{CM})^{-1} \cdot M_{CM}^T \\ \Pi_{M_{CM}}^\parallel &= I - \Pi_{M_{CM}}^\perp \end{aligned} \quad (5)$$

where M_{CM} is the response matrix linking all correctors to all monitors, Π^\perp and Π^\parallel are projection operators mapping any vector into components outside and inside the subspace spanned by the column vectors of M_{CM} , and K and Δ are the error vectors of Eqns (3) and (4). Eqn (5) is intuitively obvious since no corrector system can eliminate an orbit error pattern which is outside the corrector's reach,

while only the part of a "fake" orbit error pattern due to monitor offsets that is *correctable* will lead to real but unobserved residual orbit errors after correction. An important task for any intelligent orbit correction algorithm is to distinguish the signature of these two very different sources of errors before applying correction. This can not be achieved exactly since the problem is usually underconstrained. An algorithm developed to obtain consistent estimates however proved successful in a wide range of cases. This is described in the following sections.

3.1 Alignment biased solution

We start with the extreme assumption that any discrepancy between observed orbit, corrector strength, and model

$$Y_i = X_i - \sum_j M_{ij} \cdot C_j \quad (6)$$

is predominantly attributable to monitor errors. We then perform singular value decomposition (SVD) on the response matrix M_{AM} from a wide range of alignment error locations and coordinates to all monitors

$$M_{AM} = U^T \cdot W \cdot V$$

The matrix V consists of orthonormal combinations of alignment errors which have decoupled effects in the monitor space represented by U . These effects U_i are then compared to Y_i of Eqn (6) in turn to identify the error combination responsible for Y_i . This is iterated to build a matrix N of combined error effects to account for Y_i

$$\begin{aligned} N_{ij} &= U_i^j & N^\dagger &= N \cdot (N^T \cdot N)^{-1} \\ S_j &= \sum_i N_{ji}^\dagger \cdot Y_i \\ Y_i &\rightarrow Y_i - \sum_j N_{ij} \cdot S_j \end{aligned} \quad (7)$$

It was indicated that the orbit discrepancy Y_i is reduced at each iteration as more alignment errors are included, and thus will represent progressively more contribution due to monitor errors. Y_i is also used to update the weighting of monitors at every iteration to speed up the distinction of major monitor errors above noise level. The iteration is terminated when the algorithm detects an un-natural jump in the alignment error magnitude S_j in Eqn (7), signaling a monitor error being mis-interpreted as alignment-induced.

3.2 BPM biased solution

We start with the extreme assumption that the orbit discrepancy of Eqn (6) is predominantly attributable to alignment type errors. We perform the same iteration process as described in the last section. The only difference is that the monitor weightings are not updated, and since the monitor errors are assumed to be more reliable, the iteration terminates when the orbit discrepancy is below a specified limit.

3.3 Interpolation between the biases

An entire spectrum of orbit interpretations can be interpolated between these two extremes. Such interpolation also represents a compromise between globally and locally-oriented interpretations of the data, since the more alignment error one introduces to account for the discrepancy, the more emphasis is given to satisfying the short range observation at the expense of global consistency. This algorithm

allows user defined criteria for choosing a desired interpolation, for example, the one with minimal combined RMS of alignment and monitor errors.

The introduction of alignment errors amounts to enhancing the fitting parameters in a rigid injection fit, with only responses from injection to monitors. We can represent this enhanced response with a matrix U_0 which contains at least the injection responses, and maybe any number of responses to alignment errors. The final residual correction error becomes, with input of the interpreted orbit:

$$E = \Pi_{M_{CM}}^\perp \cdot K + \Pi_{M_{CM}}^\parallel \cdot \Pi_{U_0}^\perp \cdot K - \Pi_{M_{CM}}^\parallel \cdot \Pi_{U_0}^\parallel \cdot \Delta \quad (8)$$

where the projection operators Π^\perp and Π^\parallel are as defined in Eqn (5). Depending on the predominance of alignment or monitor error, changing the content of U_0 , amounting to interpolation between the two biases, can help offset the effects of K or Δ and reduce the overall residual error.

4 IMPLEMENTING VIRTUAL MONITORS

Once the interpretation of the underlying orbit is decided, we have the complete information of beam coordinate at any location. The virtual monitors described below use this information as input to the steering algorithms.

4.1 Generalized steering constraints realized as virtual monitors

Very often not only the orbit at the monitors are targeted for correction, but also injection angle at the end of the line, orbit at un-monitored locations, or even overall path-length need to be constrained simultaneously. Knowing the interpreted beam coordinate everywhere, one can easily include these as orbit errors at virtual monitors, and feed this together with the relevant response matrix elements and weighting to a generic steering algorithm. The latter has no need to know a priori that the input contains more than pure position data at the real monitors.

Another type of steering constraint happens with orbit dependent optics. When the orbit is changed, the higher order optical elements pick up extra contributions. We are especially interested in the effects of chromaticity-like elements T_{ij6} , the orbit effect of which on the linearized elements M_{i6} between points a and b can be expressed as

$$\delta M_{i6}^{ab} = \sum_{j \neq 6} \sum_k \sum_c 2 T_{ijk}^{cb} M_{j6}^{ac} \delta X_k^c + \sum_j \sum_c 2 T_{ij6}^{cb} \delta X_j^c \quad (9)$$

where we sum over all the locations c with appreciable coordinate change and/or T_{ij6} . The second term in the sum of Eqn (9) becomes dominant with large chromaticity-like optical elements and a steering constraint on the overall M_{i6} can be realized again through virtual monitors which simply have as response matrix elements the product of these T_{ij6} 's and the ordinary M_{ij} 's.

4.2 Virtual monitors created for singularity control

To combat the problem of excessive correction due to response matrix singularities, an algorithm is developed which automatically places extra constraints in the form of

virtual monitors coupling strongly to the singular corrector combinations. It is outlined in the following.

1. Determine cutoff numbers R and S , R measures the evenness in corrector effect distribution among monitors reflected in SVD condition number. S with $0 < S < 1$ measures orthogonality of the corrector effects.
2. Form a set C_M of locations densely covering the beam line of interest, not necessarily tied to any physical elements. Form the response matrix M_{CA} from all correctors to these locations.
3. Perform SVD on the corrector-to-monitor response matrix M_{CM} , if the condition number $N_{M_{CM}}^{svd}$, i. e., the ratio between the largest and the smallest singular values, is greater than R , or the normalized Gram determinant $\bar{G}_{M_{CM}}$ [1] is less than S to the power of the number of correctors, continue.
4. Identify the row v of V with the smallest singular value.
5. Apply the matrix M_{CA} to the vector v . Identify its largest component with index j .
6. Place a virtual monitor at the j -th location in the set C_A .
7. Iterate steps 3-6 until $N_{M_{CM}}^{svd}$ and $\bar{G}_{M_{CM}}$ satisfy the conditions specified in step 3.

Once all the virtual monitors are identified, the interpreted orbit at these locations are added to the input orbit for the steering algorithm. The steering result should be automatically free of singularity.

5 STATUS OF APPLICATION

The algorithm described in this report has been prototyped in Mathematica and tested on a wide range of CEBAF machine data. From these tests the error analysis algorithm of Section 3 performed satisfactorily. It was used to identify BPM offsets and misalignments in various areas. Singularity-induced excessive correction has been observed in the lower arcs at CEBAF and verified by simulation. The algorithm of Section 4 for eliminating such problems using virtual monitors has been tested against simulation and found to perform extremely well. Inclusion of generalized steering constraints such as angle and path length has been tested in simulation. Inclusion of higher order optical effects has not been tested. Application to simultaneous multiple pass steering has been tested separately.

Implementation under the on-line control system [2] is being carried out. With the degree of versatility and robustness to be built into the production version, it is expected to meet the unique demands for global orbit control at CEBAF.

6 REFERENCES

- [1] 'Methods of Orbit Correction System Optimization', Y. Chao, these proceedings.
- [2] 'A Portable Accelerator Control Toolkit', William A. Watson III, these proceedings.