

Goal programming algorithms for closed orbit correction

W. Herr and J. Miles
CERN SL
CH-1211 Geneva 23

Abstract

The goal of most popular closed orbit correction algorithms is to find a set of correctors which either minimise the orbit deviations in a mean-squared sense or suppress unwanted frequencies in the orbit spectrum. In both cases some difficulties can be expected if, in addition, the solution is required to satisfy certain constraints.

A new look at the problem is presented in which constraints appear quite naturally. The method is based on the principle that even when the desired goal is unreachable, it should be useful to find a solution which approaches the goal as close as possible. Two algorithms are presented which differ only in how they measure the distance from the goal. In one case the l_1 -norm is chosen, while in the other it is the l_∞ or Chebyshev norm. In both cases the computation is carried out using the Simplex method of linear programming.

1 STATEMENT OF THE PROBLEM

A large number of papers have already been written on the subject of closed orbit correction and the associated implementation details [1, 2, 3]. In all of these the commonly used correction strategies assume a single objective or goal which can be stated in simple terms as "... minimise the rms orbit deviation at the BPM's ...". In most cases this is certainly adequate from both an operations and a beam-dynamics point of view. However, there are occasions when more may be required. For example, it could be that the orbit must follow a specific path over a particular section of the machine, while at the same time the correction dipole currents do not exceed their hardware limits and the orbit deviation is minimised over the rest of the machine. These constraints are not necessarily commensurate, which means that they cannot be directly combined or compared. It is also clear that they could be conflicting. Trade-offs must be made in the sense that sacrificing the requirements on any one constraint will tend to produce greater returns on the others. Another scenario, likely to occur during machine startup or commissioning, is the need to reduce peak-to-peak orbit excursions in an essentially uncorrected machine. It will be shown that in a machine with suitably distributed BPM's and correction dipoles (a weak condition in practice) it should be possible to compute a *unique* solution at each iteration provided that the problem is formulated in a particular way.

Following the usual notations, the general problem of correcting the closed orbit subject to (optional) constraints

can be stated as :

$$\text{Minimise } \|r\| = \|y - Ak\| \quad (1.1)$$

$$\text{subject to } Ck = d \quad (1.2)$$

$$Ek \leq f \quad (1.3)$$

where the $m \times n$ transfer matrix A has elements :

$$a_{ij} = \frac{\sqrt{\beta_i \beta_j}}{2 \sin(\pi Q)} \cos(|\mu_i - \mu_j| + \pi Q)$$

The m -dimensional column vector y contains BPM readings and the n -dimensional column vector k contains the unknown correction dipole strengths. The norm $\|\cdot\|$ is quite arbitrary, though the discussion will be limited to l_p -norms, $p = 1, 2$ and ∞ . The constraints (1.2) can be used to express orbit path requirements while correction dipole limitations can be incorporated in (1.3). Note that in the case of the SPS (Sp̄S) and LEP, $m > n$, i.e. the system is overdetermined, and this will be assumed in the remainder of the paper. The case $p = 2$ i.e. least squares, is well documented [4, 5] and will not be considered further.

2 TWO ALGORITHMS

With an appropriate choice of norm, the constrained closed orbit correction problem can be solved using algorithms which are well known in the management science and numerical approximation literature [6, 7]. The problem is reformulated as extended linear programs (LP) which are solved either in primal or dual form using the Simplex algorithm [8]. This approach is not unknown in accelerator control. An early reference [9] deals with field trimming in a cyclotron and compares LP results with a least squares solution. A later paper [10] describes the computation of optimum septa and bumper currents and septa positions for the SPS extraction system.

2.1 The l_1 solution

The idea of setting approximation problems in this norm appears to originate with Laplace and is considered to be important in the analysis of data which includes gross inaccuracies or outliers, for the effective weight given to such values is smaller with any other l_p -norm. The LP formulation is more recent [11, 12, 13]. Recall that the l_1 -norm is defined as :

$$\|r\|_1 = \sum_{i=1}^n |r_i|$$

Then the closed orbit correction problem (1.1) can be restated as :

$$\text{Minimise } e(u + v) \quad (2.1)$$

$$\text{subject to } A(k' - k'') + u - v = y \quad (2.2)$$

$$C(k' - k'') = d \quad (2.3)$$

$$E(k' - k'') + u' = f \quad (2.4)$$

$$\text{and } \begin{array}{l} k', k'' \\ u, v \\ u' \end{array} \begin{array}{l} \geq 0 \\ \geq 0 \\ \geq 0 \end{array}$$

The formulation of the problem is such that

$$u_i + v_i = |u_i - v_i| = |r_i|$$

hence it is entirely equivalent to minimising $\|r\|_1$. In general this problem will not have a unique solution [7]. Note that two kinds of constraints occur. Equation (2.2) defines goal constraints, which are also known as "soft constraints" and may be violated if necessary. Here u and v are m -dimensional column vectors which are called deviation variables. At optimality at least one of the pair u_i, v_i will always be zero [6]. The variable u represents underachievement while v represents overachievement.

Equations (2.3) and (2.4) define system constraints, which are also known as "hard constraints" and cannot be violated. The column vector u' is a slack variable which is used to transform the inequality constraints into equations. The variable $k = k' - k''$ is defined in terms of two nonnegative n -dimensional column vectors k', k'' . The m -dimensional row vector $e = [1, 1, \dots, 1]$ is used to construct the objective function (2.1) which must be made as small as possible to satisfy as closely as possible the goal constraints. The remaining constraints ensure nonnegativity (which is not strictly necessary).

2.2 The l_∞ solution

The analysis of data with errors drawn from distributions with sharply defined extremes (as typified by the uniform distribution) is best carried out in this norm, for the large peak-to-peak excursions are not considered to be outliers and they must be dealt with correctly. The Chebyshev norm is defined as :

$$\|r\|_\infty = \max_i |r_i|$$

Using this definition the closed orbit correction problem (1.1) can be restated as the following *ansatz* :

$$\text{Minimise } \mathcal{P}_1 z + \mathcal{P}_2 u_1 + \mathcal{P}_3 v_1 \quad (2.5)$$

$$\text{subject to } A(k' - k'') + u - v = y \quad (2.6)$$

$$C(k' - k'') + u_1 - v_1 = d \quad (2.7)$$

$$E(k' - k'') + u' - v_1 = f \quad (2.8)$$

$$Iz - u - v \geq 0 \quad (2.9)$$

$$\text{and } \begin{array}{l} k', k'' \\ u_1, v_1 \\ u, v \\ u', v' \end{array} \begin{array}{l} \geq 0 \\ \geq 0 \\ \geq 0 \\ \geq 0 \end{array}$$

where I is an $m \times m$ identity matrix. In the objective function (2.5), the terms \mathcal{P}_k serve merely to indicate priorities,

with \mathcal{P}_1 denoting highest priority, and so on [6, 14]. It follows that the scalar quantity z must be minimised with first priority. The formulation of the problem is such that

$$z \geq u_i + v_i$$

so that minimising z is completely equivalent to minimising $\|r\|_\infty$. Note that the computational difficulty of an LP is approximately proportional to $m^2 n$ so increased efficiency will result from solving the problem in its dual form.

The column vector u_1 denotes underachievement in trying to meet orbit path requirements and v_1 denotes overachievement while exceeding correction dipole limitations. Both of these quantities are minimised as lower priority goals for obvious practical reasons. The column vectors u' and v' are merely slack and surplus variables, respectively.

Normally each $n \times n$ submatrix of the transfer matrix A will have full rank and the Haar condition will be satisfied. This means that the \mathcal{P}_1 iteration of the LP will have a unique solution [7] and it is unclear whether the lower priority iterations will yield anything useful. However, if the Haar condition is not satisfied then these additional iterations could be worthwhile. The question of l_∞ approximation subject to either general or restricted constraints has been examined [17, 18] and has applications in the design of digital filters.

3 IMPLEMENTATION AND STATUS

Early closed orbit correction experiments in the l_1 and l_∞ -norms (unconstrained) were performed on the Sp̄pS collider using codes which were NODAL [19] adaptations of Algol procedures [12]. These were later replaced by FORTRAN versions [15, 16] running in the NODAL environment and included the possibility of handling constraints in the l_1 case. This was used successfully to correct the orbit in the Sp̄pS low- β insertions which contained correctors of very limited strength at high energy. The FORTRAN routines are now incorporated into the standard closed orbit correction package COCU [2] of the SPS and LEP but they are not fully tested and remain only partially operational. Small scale offline experiments using prioritised optimisation in the l_∞ -norm are being conducted using a software package for PC's [14].

4 REFERENCES

- [1] W. Herr *et al.*; Proc. IEEE Particle Accelerator Conference, Chicago, (1989), IEEE Conf. Proc. 3, (1989) 1681.
- [2] D. Brandt *et al.*; *A New Closed-Orbit Correction Procedure for the CERN SPS and LEP*, Nuclear Instruments and Methods, A293 (1990) 305-307.
- [3] A. Chao *et al.*; *On-Line Orbit and Dispersion Correction Schemes for PEP*, PEP-Note 318, PEP, (1979).
- [4] B. Autin and Y. Marti; *Closed Orbit Correction of A.G. Machines Using a Limited Number of Magnets*, CERN ISR-MA/73-17, CERN, Geneva, (1973).

- [5] V. Ziemann; *Corrector Ironing in the SLC Final Focus*, Proc. Particle Accelerator Conference, Washington, (1993).
- [6] G. D. Eppen *et al.*; *"Quantitative Concepts for Management"*, 3rd edition, Prentice-Hall, New Jersey, (1989).
- [7] G. A. Watson; *"Approximation Theory and Numerical Methods"*, Wiley, New York, (1980).
- [8] L. Schrage; *"LINDO, An Optimisation Modeling System"*, 4th edition, Scientific Press, San Francisco, (1991).
- [9] C. M. Davisson and R. B. Theus; *Linear Programming vs Least Squares Method in Determining the Best Trim Coil Currents for the NRL Cyclotron*, Particle Accelerators, 7, (1976) 103-110.
- [10] M. Gyr; *Computation of Optimum Septa and Bumper Currents and Septa Positions for the SPS Extraction System*, CERN SPS/AOP/78-20, CERN, Geneva, (1978).
- [11] H. M. Wagner; *Linear Programming Techniques for Regression Analysis*, J. Amer. Stat. Assoc., 54, (1959) 206-212.
- [12] I. Barrodale and A. Young; *Algorithms for Best L_1 and L_∞ Linear Approximation on a Discrete Set*, Num. Math., 8, (1966) 295-306.
- [13] I. Barrodale and F. D. K. Roberts; *An Improved Algorithm for Discrete l_1 Linear Approximation*, SIAM J. Num. Anal., 10, (1973) 839-848.
- [14] Y-L Chang and R. S. Sullivan; *"QSOB+, Quantitative Systems for Business Plus"*, version 2, Prentice-Hall, New Jersey, (1991).
- [15] I. Barrodale and F. D. K. Roberts; *Algorithm 478: Solution of an Overdetermined System of Equations in the l_1 -norm*, Comm. ACM, 17, (1974) 319-320.
- [16] I. Barrodale and C. Phillips; *Algorithm 495: Solution of an Overdetermined System of Linear Equations in the Chebyshev Norm*, A.C.M. Trans. Math. Software, 1, (1975) 264-270.
- [17] M. J. D. Powell; *The Minimax Solution to Linear Equations Subject to Bounds on the Variables*, Proc. of the Fourth Manitoba Conference on Numerical Mathematics, eds. B. L. Hartnell and H. C. Williams, Utilas Mathematica, Winnipeg, (1975) 53-107.
- [18] N. N. Abdelmalek; *The Discrete Linear Restricted Chebyshev Approximation*, BIT, 17, (1977) 249-261.
- [19] M. C. Crowley-Milling and G. C. Shering; *The NODAL System for the SPS*, CERN 78-07, CERN, Geneva, (1978).