DAΦNE TECHNICAL NOTE

INFN - LNF, Accelerator Division

Frascati, April 18, 1997

Note: **G-46**

HARMONIC CLOSED ORBIT CORRECTION WITH THE LAGRANGE MULTIPLIERS METHOD

M. Bassetti, S. Guiducci

1. Introduction

The closed orbit correction scheme presented in this paper is based on the calculation of the Fourier spectrum of the measured closed orbit and of the orbit created by the correctors with a limited number M of harmonics. The purpose of the truncation is to correct only the harmonics of the real orbit and not those caused mainly by the unavoidable monitor errors. The correction of the filtered orbit is obtained by minimizing the strengths of the correctors with the Lagrange multipliers method.

The correction scheme is described in the first sections of this note and an example of its application is given in Section 6.

2. Correction procedure

The closed orbit correction scheme presented in this paper is an iterative procedure made up of the following steps:

1) Computation of the Fourier spectrum of the measured closed orbit.

2) Elimination of the harmonics which are mainly generated by the unavoidable monitor errors rather than by the misalignment errors and choice of the M harmonics to correct.

3) Reconstruction of a filtered orbit with M harmonics.

4) Correction of the filtered orbit by minimizing the corrector strengths with the Lagrange multipliers method.

Steps 1) and 3) are performed with the same algorithm, which is described in Section 3.

This method, due to the truncation and the minimization procedure, has two main advantages with respect to other methods:

- first, it allows to reduce the contribution of the monitor errors to the correction, producing a good correction of the effective closed orbit due to misalignment errors with low corrector strengths.
- second, all the calculations are performed by solving linear systems of order 2M, whatever the number of correctors and monitors.

The second point is particularly important for large machines. In fact, as we show in Section 6, a small number of harmonics M, centered around the value of the betatron tune, is needed to obtain a good orbit correction.

3. Determination of the reconstructed orbit with M harmonics

We expand the closed orbit in Fourier series up to order M and impose that it is equal to the measured values at the monitors. The coefficients of the M harmonics are calculated with the least square method.

We use Courant-Snyder [1] variables to simplify the calculations. The Fourier expansion, up to order M, of the measured closed orbit at the monitors, in Courant-Snyder variables η and ϕ, is:

$$\eta_i = \frac{x_i}{\sqrt{\beta_i}} = \frac{a_0}{2} + \sum_{m=1}^{M}\left[a_m \cos(m\phi_i) + b_m \sin(m\phi_i)\right]; \qquad \text{i=1,...H} \qquad (1)$$

where x_i are the measured values of the orbit in the horizontal or vertical plane and H is the number of monitors.

The constant term can be neglected because it depends on the length of the orbit which is determined only by the RF frequency.

By applying the least square method to the above equations with:

$$M \leq H/2$$

the 2M coefficients a_m and b_m can be calculated:

$$\frac{\partial}{\partial a_j}\left\{\sum_{i=1}^{H}\left[\eta_i - \sum_{k=1}^{M}\left(a_k \cos(k\phi_i) + b_k \sin(k\phi_i)\right)\right]^2\right\} = 0 \qquad \text{j=1,..M} \qquad (2)$$

$$\frac{\partial}{\partial b_j}\left\{\sum_{i=1}^{H}\left[\eta_i - \sum_{k=1}^{M}\left(a_k \cos(k\phi_i) + b_k \sin(k\phi_i)\right)\right]^2\right\} = 0 \qquad \text{j=1,..M} \qquad (3)$$

Evaluating the derivatives and defining a 2Mx2M matrix $\hat{M}$:

$$\hat{M} = \begin{vmatrix} \hat{M}_{11} & \hat{M}_{12} \\ \hat{M}_{21} & \hat{M}_{22} \end{vmatrix}$$

$$M_{11jk} = \sum_{i=1}^{H}\cos(j\phi_i)\cos(k\phi_i) \quad ; \quad M_{12jk} = \sum_{i=1}^{H}\cos(j\phi_i)\sin(k\phi_i) \qquad (4)$$

$$M_{21jk} = \sum_{i=1}^{H}\sin(j\phi_i)\cos(k\phi_i) \quad ; \quad M_{22jk} = \sum_{i=1}^{H}\sin(j\phi_i)\sin(k\phi_i)$$

and two column vectors $\overline{A}$ and $\overline{V}$ with 2M components:

$$A_j = a_j \qquad A_{M+j} = b_j \qquad \text{j=1,...M}$$

$$V_j = \sum_{i=1}^{H}\eta_i \cos(j\phi_i) \qquad V_{M+j} = \sum_{i=1}^{H}\eta_i \sin(j\phi_i) \qquad \text{j=1,...M} \qquad (5)$$

equations (2) and (3) can be written:

$$\overline{V} = \hat{M}\overline{A} \,. \tag{6}$$

The Fourier coefficients of the expansion of the measured closed orbit are obtained by inverting the matrix $\hat{M}$:

$$\overline{A} = \hat{M}^{-1}\overline{V} \,. \tag{7}$$

It has to be noted that the dimension of the $\hat{M}$ matrix is twice the number of harmonics M independently on the number of monitors.

4. Orbit created by a single corrector and constraints to correct the orbit

Let us consider the betatron equation in Courant-Snyder variables with an unitary corrector kick placed at the azimuth ϕ_k:

$$\frac{d^2\eta_c}{d\phi^2} + Q^2\eta_c = \delta(\phi - \phi_k) \tag{8}$$

The delta function can be written:

$$\delta(\phi-\phi_k) = \frac{1}{2\pi} + \frac{1}{\pi}\sum_{m=1}^{\infty} \frac{\cos(m(\phi-\phi_k))}{2} \,. \tag{9}$$

Solving term by term we obtain:

$$\eta_c = \frac{1}{2\pi Q^2} + \frac{1}{\pi(Q^2-m^2)} \cos(m(\phi-\phi_k)) \tag{10}$$

Defining:

$$C_{mk} = \frac{1}{\pi\left(Q^2 - m^2\right)}\cos\left(m\phi_k\right)$$
$$S_{mk} = \frac{1}{\pi\left(Q^2 - m^2\right)}\sin\left(m\phi_k\right) \,. \tag{11}$$

and keeping only the oscillating terms the orbit of a corrector can be written as:

$$\eta_c = \sum_{m=1}^{\infty}\left(C_{mk}\Delta_k \cos(m\phi) + S_{mk}\Delta_k \sin(m\phi)\right) \tag{12}$$

where Δ_k is the kick of the corrector in Courant-Snyder variables:

$$\Delta_k = Q\sqrt{\beta_k}\,\frac{(Bl)_k}{B\rho} \tag{13}$$

In order to cancel with the correctors M harmonics of the measured orbit we must satisfy the following 2M equations:

$$a_m - \sum_{k=1}^{N} C_{mk}\Delta_k = 0 \qquad \text{m=1,...M} \tag{14}$$

$$b_m - \sum_{k=1}^{N} S_{mk}\Delta_k = 0 \qquad \text{m=1,...M} \tag{15}$$

where N is the number of unknown kicks.

5. Lagrange multipliers method to minimize the average quadratic kick

The closed orbit due to misalignment errors is produced by many small kicks distributed along the ring. Therefore the correction is more efficient if it is obtained with many low intensity correctors instead of few high strength ones. This is obtained by minimizing the correctors strengths with the Lagrange multipliers method. The minimum number of correctors is twice the number of harmonics but in practice the number of correctors is always larger than 2M. We use all the available correctors minimizing the sum of the kicks squared, i.e. the function:

$$F_0 = \sum_{k=1}^{N} \Delta_k^2 \tag{16}$$

A well known way to satisfy at the same time (14), (15) and (16) is the Lagrange multipliers method. We add to F_0 the 2M constraints (14) and (15) multiplied respectively by λ_{am} and λ_{bm}. The total function to minimize by using the 2M multipliers λ_{am} , λ_{bm} and the N kicks Δ_k is:

$$F = \sum_{k=1}^{N} \Delta_k^2 + \sum_{m=1}^{M}\left[\lambda_{am}\left(a_m - \sum_{k=1}^{N} C_{mk}\Delta_k\right) + \lambda_{bm}\left(b_m - \sum_{k=1}^{N} S_{mk}\Delta_k\right)\right] \tag{17}$$

where:

$$2M \le N$$

Imposing that the derivatives of F with respect to Δ_k must vanish we obtain:

$$\frac{\partial F}{\partial \Delta_k} = 2\Delta_k - \sum_{m=1}^{M}\left(\lambda_{am}C_{mk} + \lambda_{bm}S_{mk}\right) = 0 \qquad \text{k=1,...N} \tag{18}$$

namely:

$$\Delta_k = \frac{1}{2}\sum_{m=1}^{M}\left(\lambda_{am}C_{mk} + \lambda_{bm}S_{mk}\right) \tag{19}$$

From the 2M derivatives of F with respect to λ_{am} and λ_{bm} we reproduce equations (14) and (15). Substituting eq. (19) into (14) and (15) gives:

$$2a_m = \sum_{k=1}^{N}\sum_{j=1}^{M}\left[C_{mk}C_{jk}\lambda_{aj} + C_{mk}S_{jk}\lambda_{bj}\right] \tag{20}$$

$$2b_m = \sum_{k=1}^{N}\sum_{j=1}^{M}\left[S_{mk}C_{jk}\lambda_{aj} + S_{mk}S_{jk}\lambda_{bj}\right] \tag{21}$$

It is useful to define a 2Mx2M matrix $\hat{G}$ and a vector $\overline{\lambda}$ with 2M components; the vector $\overline{A}$ is defined in (5):

$$G_{11mj} = \sum_{k=1}^{N} C_{mk}C_{jk} \qquad G_{12mj} = \sum_{k=1}^{N} C_{mk}S_{jk}$$
$$G_{21mj} = \sum_{k=1}^{N} S_{mk}C_{jk} \qquad G_{22mj} = \sum_{k=1}^{N} S_{mk}S_{jk} \tag{22}$$

$$\hat{G} = \begin{vmatrix} \hat{G}_{11} & \hat{G}_{12} \\ \hat{G}_{21} & \hat{G}_{22} \end{vmatrix} \qquad \begin{array}{l} \lambda_j = \lambda_{aj} \\ \lambda_{M+j} = \lambda_{bj} \end{array} \qquad \text{j=1,...M} \tag{23}$$

Equations (20) and (21) can be written:

$$2\overline{A} = \hat{G}\overline{\lambda} \tag{24}$$

and the values of $\overline{\lambda}$ can be found by inverting the matrix $\hat{G}$:

$$\overline{\lambda} = 2\hat{G}^{-1}\overline{A}\,. \tag{25}$$

Finally the angles of the correctors are obtained by inserting the values of $\overline{\lambda}$ in equation (19).

The correction is calculated by inverting a square matrix 2Mx2M, where M is the number of harmonics chosen by truncating the Fourier spectrum of the orbit. As the closed orbit is resonant with the betatron tune, a good correction is achieved with a relatively small number of harmonics centered around the tune value.

This is an advantage for large machines with respect to other methods which find the solution by inverting large matrices with the dimension of the number of monitors or correctors.

6. Application to a storage ring

The closed orbit measured in a storage ring is always affected by unknown monitor errors; the main advantage of this method is the possibility of filtering out the errors of the monitors and of reducing the strengths of the correctors. To show this we simulate a closed orbit due to quadrupoles misalignment, then we add to the value of the displacement at each monitor a monitor error extracted at random with a gaussian distribution. In this way we simulate a measured closed orbit for which we know exactly the contribution of the monitor errors.

The lattice used is that of the DAΦNE accumulator ring with Q_x =3.1, Q_y = 1.15 and the error orbit is simulated by MAD [2]. We have applied the harmonic correction scheme varying the number of harmonics used to reconstruct the orbit and to calculate the correctors. The rms and peak values of the orbit at the monitors, before correction, are given in Table I for an horizontal orbit with and without monitor errors (σ_{err} is the standard deviation of the monitor errors).

Table I - rms and peak values of the simulated closed orbit at the monitors

monitor errors	monitor readings	
σ_{err} (mm)	x_{rms} (mm)	x_{max} (mm)
0.	5.05	8.65
1.	5.01	8.17

The Fourier spectrum of the orbit is calculated, as described in Section 3, with five harmonics centered around m=3, which is the harmonic nearest to the betatron tune. The amplitude of the different harmonics :

$$\rho = \sqrt{\langle\beta_x\rangle}\sqrt{a_m^2 + b_m^2} \ ,$$

where $<\beta_x>$ is the mean value of the beta-function over the ring, is shown in Fig. 1 for the simulated closed orbit with and without monitor errors and for the monitor errors only.

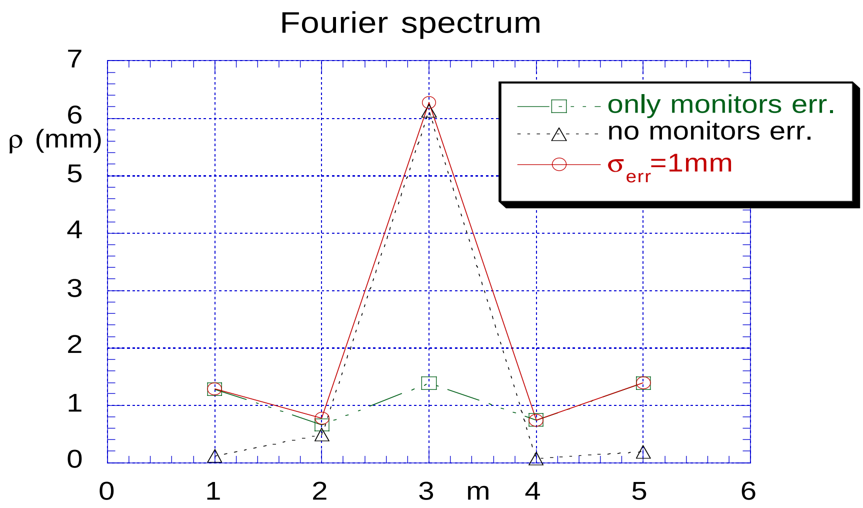

Fig. 1 - Fourier spectrum of the reconstructed closed orbit
with and without monitor errors and of only the monitor errors.

The closed orbit is resonant with the betatron tune of the ring. Therefore its Fourier spectrum has a sharp peak on the third harmonic (the nearest to Q_x = 3.1). The ratio between the third harmonic and the two nearby ones is of the order of $1/\Delta Q$, where ΔQ is the fractional part of the machine tune, a factor 10 in this case.

The spectrum of the monitor errors only is nearly flat.

The spectrum of a real orbit is always affected by the monitor errors and is similar to the curve for σ_{err} = 1 mm shown in Fig. 1: the harmonic nearest to the betatron tune is dominated by the misalignment closed orbit while the other harmonics are dominated by the monitor errors.

The correction of the third harmonic reduces the closed orbit by a factor nearly ten, which is already satisfactory. The correction of further harmonics does not reduce the closed orbit, on the contrary makes it larger because those harmonics are mainly produced by the monitor errors.

The correction of a few harmonics beside the main one is convenient only if the contribution of the monitor errors for those harmonics is smaller than that of the closed orbit.

We apply the formulas given in Section 5 to the simulated orbit to obtain the strength of the correctors and then calculate the orbit at the monitors after correction for different values of the number of harmonics N_{harm}.

This is done for two different values of the standard deviation of the monitor errors:

a) σ_{err} = 1 mm the harmonics, except the third, are dominated by the monitor errors

b) σ_{err} = .2 mm for the third and the second harmonic the monitor errors are smaller than the closed orbit.

For each case the correction is calculated for 100 orbits corresponding to the same alignment errors and different random extraction of the monitor errors.

The average, over 100 extraction of the monitor errors, of the rms value of the closed orbit after correction, as a function of $N_{harm,}$, is shown in Fig. 2 for the two cases. Correspondingly the average of the rms values of the corrector strengths is given in Fig. 3. The index m of the N_{harm} harmonics is listed in Table 2.

Table II - Index m of the N_{harm} harmonics used for the correction

N_{harm}	1	2	3	5
m	3	2,3	2,3,4	1,2,3,4,5

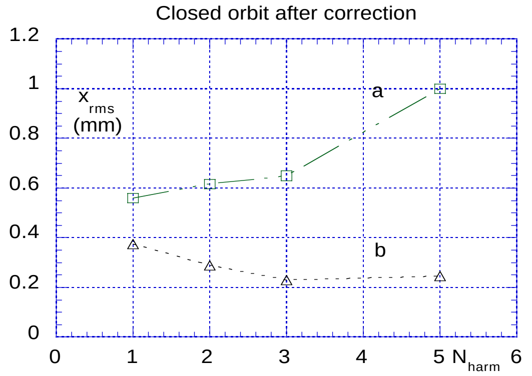

Fig. 2 - Average, over 100 extraction of the monitor errors, of the rms value of the orbit, after correction with N_{harm} harmonics, versus N_{harm}.
a) σ_{err} = 1 mm; b) σ_{err} = .2 mm

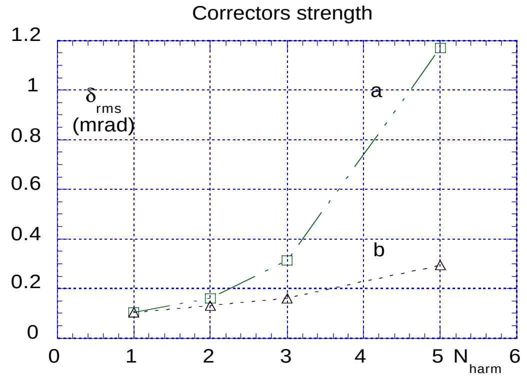

Fig. 3 - Average, over 100 extraction of the monitor errors, of the rms value of the deflection angles of the correctors versus N_{harm}.
a) σ_{err} = 1 mm; b) σ_{err} = .2 mm

For the orbit with larger monitor errors (a) it is clearly convenient to correct only the third harmonic because correcting more harmonics gives a larger closed orbit and, moreover, increases very rapidly the strength of the correctors.

For the orbit with smaller monitor errors (b) the correction with only the third harmonic reduces the closed orbit by a large factor (x_{rms} before corr./after corr. = 13). Correcting also the second and fourth harmonic gives only an orbit reduction of a factor 1.3 at the expense of an increase of the correctors angles by a factor 1.6.

It has to be noted that the rms deflection angles required to correct the third harmonic do not depend on the monitor errors because the closed orbit is resonant with the betatron tune and the influence of the monitor errors on the peak harmonic is negligible.

7. Conclusions

With this method a large reduction of the closed orbit can be obtained, by using only the harmonic nearest to the Q value, with the minimum strength of the correctors independently from the magnitude of the monitor errors. To further reduce the closed orbit the number of harmonics used in correction can be increased; the optimal correction is obtained by truncating the Fourier spectrum when the monitor errors became dominant with respect to the alignment closed orbit.

In a real machine the monitor errors are not known; therefore an iterative procedure is needed for the choice of the best set of harmonics.

A possible procedure is to calculate the Fourier spectrum with a given number of harmonics (always less than half the number of monitors) and consider the ratio between the peak and the neighbors harmonics: if it is small with respect to the theoretical value ($1/(Q^2-m^2)$) it means that the harmonic is dominated by the monitor errors and is not useful for correction.

As it is a good rule to try to keep low the correctors strengths a good procedure is to try to correct the orbit with a small number of harmonics, then check the goodness of the correction and in case increase the number of harmonics. A possible method to check the goodness of the correction is to measure two orbits at different tunes and subtract one from the other: when the difference vanishes the orbit is corrected.

References

[1] E. D. Courant, H. S. Snyder, Annals of Physics: 3,1 48 (1958).

[2] H. Grote, F.C. Iselin, "The MAD Program. User's Reference Manual", CERN/SL/90-13(AP).