The Physics and Math of Bumps in Courant/Snyder

M. Fischler

September 22, 2003

Abstract

A collection of formulae and concepts relevant in doing Tevatron orbit analysis and beam smoothing is presented. These are from the viewpoint of a relative newcomer to Tevatron beams physics; there is no claim that useful results are derived here for the first time.

A rather long section on the results of trying to use the DFG's to shift the entire orbit (or large fractions it) by a fixed small amount.

This document is likely to evolve a bit further as additional useful information is derived in the course of developing cusp-finding software.

Contents

1 Overview and Problems

This technical memo is intended for an audience of readers who have a Ph. D. in physics and working experience in particle physics, but who are not by profession accelerator physicists, immersed in beam orbit theory. We also assume that the reader has a casual knowledge of Courant-Snyder theory of beam orbits[1], as could be obtained from a semi-serious perusal of the first few chapters of Edwards and Syphers[2]. Thus the informal title of this memo: *"Bumps For Dummies."*.

For a highly localized corrector or field distortion, the relevant quantity describing how much the corrector will affect the orbit is, of course, the integral of the dipole field strength over the extent of the correction field. If that integral were 1, we would say this is a corrector of unit strength.

The following fundamental equation, applying to perturbations of a closed Courant-Snyder orbit, describes the effect of a unit strength dipole field corrector (or a unwanted dipole field distortion of unit strength) at position j, on the beam location at position i:

$$a_{ij} = \frac{\sqrt{\beta_i \beta_j}}{2\sin(\pi Q)} \cdot \cos(|\mu_i - \mu_j| - \pi Q) \tag{1}$$

where β_i (β_j) is the β-function and μ_i (μ_j) is the betatron phase at position i (j). For this document, the coefficient is moot, and we will re-write this as

$$a_{ij} = A\cos(|\mu_i - \mu_j| - \pi Q) \tag{2}$$

where Q is the beam tune.

The first difficulties come because of the circular nature of the ring: Given two points i and j, it is ambiguous to say whether i is "before" or "after" j. Given the presence of an absolute value in expression (2), this becomes troublesome. We shall resolve the conceptual difficulties raised by this in the next section.

Another challenges is to determine the mathematics of collections of such distortions. We shall discuss the nature of single distortions and certain special cases of 2, 3, and 4 distortions (2-bumps, 3-bumps and 4-bumps). Then we will present the effect of a continuous (flat) distribution of distortion in some range.

Armed with this mathematics, we can then discuss the math involved in correcting distortions. We include a section on the special case of attempting to move the entire beam by some uniform small distance.

2 The Absolute Value and Tune Multiples

Consider a point i which is a small distance to the right of j, such that $\mu_i - \mu_j = \epsilon > 0$. Then (2) becomes

$$a_{ij} = A\cos(\epsilon - \pi Q) \tag{3}$$

Now consider moving point i along the ring, until ultimately it is at i', a small distance to the left of i, such that $\mu_j - \mu_i = \kappa > 0$. Then (2) becomes

$$a_{i'j} = A\cos(|(\mu_j - \kappa) - \mu_j| - \pi Q) = A\cos(\kappa - \pi Q) \tag{4}$$

But what if we had moved the point only about halfway around the ring, such that it is unclear whether to say that i' is to the right or to the left of j? Clearly, unless the two expressions (3) and (4) match, there is something fundamentally wrong with equation (2).

In fact, this worry can be seen even when i' is a small distance to the left of j, and at postion i which is a small distance to the right of j. Consider that i' has betatron phase of $\mu_j - \kappa$, but one could have taken i' to be a large distance to the right of j–in particular, to have a betatron phase of $\mu_j + (2\pi Q - \kappa)$. Equation (2) had better not lead to a different displacement when this equivalent expression for the phase was used!

In fact equation (2) leads to no inconsistency here because

$$\begin{aligned} &\cos\left(|\mu_j - (\mu_j + (2\pi Q - \kappa))| - \pi Q\right) \\ &= \cos(|\kappa - 2\pi Q| - \pi Q) \\ &= \cos(2\pi Q - \kappa - \pi Q) \\ &= \cos(\kappa - \pi Q) \end{aligned} \tag{5}$$

And now we can see that this consistency comes about precisely because the correct amount is subtracted in equation (2). If the subtracted term were anything other than πQ, this equality of results would not hold, and the value for change in beam location would indeed be inconsistent. But since the correct value is subtracted, for any position around the ring, the change in beam location will be the same whether you consider the betatron phase offset to be some negative number in $(-2\pi, 0)$ or its positive equivalent obtained by adding 2π.

Finally, what if one were to subtract $4\pi Q$ from a positive offset, or add some multiple of $2\pi Q$? Then the answer gotten from equation (2) would indeed change. Why is this not a problem? Because we know that the magnets in the ring will correct the (small) distortion, so as to end up with a new (slightly displaced) orbit–that is, we know that the design orbit is stable. These magnets, among other effects, have the effect of advancing the phase by πQ as a particle goes around the ring *once.* Equation (2) represents the change in the orbit due to a distortion; if you tried to take $\delta\mu$ to be multiple times around the ring, then you also would have to include multiple contributions of πQ arising from traversing the ring multiple times. So the fact that (2) does not hold in the most naive manner if you consider the betatron phases $\mu, \mu + \pi Q, \mu + 2\pi Q, ...$ does not constitute any physical inconsistency.

3 Discrete Distortions

Obviously, a discrete distortion is only an approximation, saying that the distorting field is concentrated at one or several discrete points. As we will see, the mathematics of the cosine form of equation (2) make this approximation very solid.

3.1 Single Distortions

The relevant physics for how the orbit perturbations work, given a particular (highly localized) magnetic field, is dictated by how much the particle path will be bent by that field. The discussion in the 1996 Particles and Fields book[3] in section 24.8:

> The trajectory of a particle with momentum p (in GeV/c) and charge ze in a constant magnetic field $\vec{B}$ is a helix, with radius of curvature R and pitch angle λ. The radius of curvature and momentum component perpendicular to $\vec{B}$ are related by
>
> $$p\cos\lambda = 0.3zBR \tag{6}$$
>
> where B is in tesla and R is in meters.

From this, one can compute the bend in a field, based on its integrated strength. And the bend translates into a cusp in the orbit.

A field at x_j with integrated field strength H will (absent any other fields) perturb the orbit positions $y(x)$ into

$$y(x) = H\frac{\sqrt{\beta(x)\beta(x_j)}}{2\sin(\pi Q)}\cos(|\mu_i - \mu_j| - \pi Q) \tag{7}$$

This distortion has (at x_j) a change in slope $\Delta(y')$ given by

$$\begin{aligned}
\Delta(y') &= \lim_{x\to x_j} H\frac{\sqrt{\beta(x_j)}}{2\sin(\pi Q)}\frac{d}{dx}\left[\sqrt{\beta(x)}\cos(x - x_j - \pi Q) - \sqrt{\beta(x)}\cos(x_j - x - \pi Q)\right] \\
&= H\frac{\sqrt{\beta(x_j)}}{2\sin(\pi Q)}\lim_{x\to x_j}\sqrt{\beta(x)}\left[-\sin(x - x_j - \pi Q) - \sin(x_j - x - \pi Q)\right] \\
&= H\frac{\sqrt{\beta(x_j)}}{2\sin(\pi Q)}\sqrt{\beta(x_j)}\,(2A\sin(\pi Q)) \\
&= H\beta(x_j)
\end{aligned}$$

In the derivative step above, we have not disregarded the fact that $\beta(x)$ depends (in fact, depends rather strongly at the DFG points) on x; the two terms involving the derivative of $\beta(x)$ cancel each other.

So we see that a distorting or correcting highly localized field produces a kick in the derivative which is equal to the (integrated) field strength times the β-function at the point of the distortion.

Turning this around, if we are given that the distortion causes a change in slope of some specified value $\Delta(y')$, then $H = \Delta(y')/\beta(x_j)$ and the orbit equation, caused by that perturbation alone will be

$$\begin{aligned}
y(x) &= \frac{\Delta(y')}{\beta(x_j)}\frac{\sqrt{\beta(x)\beta(x_j)}}{2\sin(\pi Q)}\cos(|x - x_j| - \pi Q) \qquad (8)\\
&= \frac{\Delta(y')}{2\sin(\pi Q)}\sqrt{\frac{\beta(x)}{\beta(x_j)}}\cos(|x - x_j| - \pi Q)
\end{aligned}$$

In rescaled units, this of course has the form

$$\tilde{y}(x) = \frac{y(x)}{\sqrt{\beta(x)}} = A\cos(|\mu_i - \mu_j| - \pi Q) \tag{9}$$

where

$$A = \frac{\Delta(y')}{2\sin(\pi Q)\sqrt{\beta(x_j)}} \tag{10}$$

An important feature is that this "kick" depends only on the field strength and the β-function at j, not on the betatron phase.

Because the β-function varies, the effect of a distortion on the actual orbit is not a pure "cusped-cosine" curve. However, if we work with the "rescaled distortion" defined above, then the form of $x(\mu)$ for a single distortion is, in fact, a pure "cusped-cosine" curve. Henceforth we will work in terms of these more convenient rescaled distortions.

Because any distorting field at a single point leads to this form of beam location distortion, our fit to the beam location curves will always be of the form

$$A\cos(x + \delta) + k \tag{11}$$

That constant k represents an overall position drift (either radially in the orbit plane or vertically), but also has another justification when we consider distorting fields which are constant over a region in betatron phase space – see section 4.

3.1.1 Numbers For a Single Cusp in the Tevatron

This sub-section is present in order to get a feel for numbers involved in finding cusps in Tevatron orbit differences, and in using a "snapshot generator" to produce controlled semi-realistic snapshots to test the analysis software.

Let's get a handle on the practical numbers involved in a cusp in the actual machine, and how that translates to "beta rescaled" numbers. We will let $y(s)$ stand for the beam displacement (in mm) at position s meters around the ring, $\psi(s)$ for the betatron phase at s, and $\tilde{y} = y/\beta(s)$ for the rescaled displacement. Then a glance at the lattice description file tells us that

$$\frac{d\psi}{ds} \approx \frac{1\text{ radian}}{50\text{ meters}} \tag{12}$$

We also see that near the BPM's and DFG's $\beta(\psi)$ is generally around 100meters, although we know that between those areas, β drops to less than half that.

When the beams people talk about a kick given by a DFG, they talk in terms of some number of mrads, say a "1mrads kick." This means that the change of slope, that is, the change in dy/ds, due to the effect of this DFG, is a milliradian. Notice two things:

1. The factor of $2\sin(\pi Q)$ in the denominator of equation (7) does not appear in this statement; that is, whatever that factor was, the net change in slope is 1 mrad.

2. Similarly, the factor of $\sqrt{\beta}$ is already taken into account.

How big a fluctuation would such a kick cause? Well, the fluctuation is $\sqrt{\beta(\psi)}$ times a cosine curve with magnitude

$$\frac{\Delta(y')}{2\sin(\pi Q)}\sqrt{\frac{1}{\beta(\psi_{\text{kick}})}} \tag{13}$$

Now we will make the patently incorrect assumption that $\sqrt{\beta(\psi)}$ is constant at $\sqrt{\beta(\psi_{\text{kick}})}$; this assumption will introduce a error by a factor of about 1.5 in many places, but here we are only trying to get a gross estimate of the range of the fluctuation, so that is acceptable. All that is left to get the desired answer is to translate the change in slope, which is given with respect to s as the independant variable, to a change in $y' = dy/d\psi$. And this is accomplished by

$$y' = \frac{ds}{d\psi}\frac{dy}{ds} \approx \frac{50 \text{ meters}}{\text{radian}} \cdot \frac{dy}{ds} \tag{14}$$

Or, using our number of 1 mrad for the kick, we get a fluctuation of magnitude

$$\frac{1}{2\sin(\pi Q)}\frac{50 \text{ meters}}{\text{radian}} \cdot 1 \text{ mrad} \tag{15}$$

and if we plug in the tune value of $Q \approx 20.58$ so that $\sin(\pi Q) \approx 1$ we find a fluctuation of magnitude 25 mm.

In the cusper, or to generate data to test the cusper, we of course deal with the rescaled value $\tilde{y}$. In the case of a 1mrad, what would our magnitude be in rescaled $\tilde{y}$? We merely divide by $\sqrt{\beta}$, giving $0.5 \cdot 10^{-3}\sqrt{\text{meter}}$. And the kick, in rescaled units, is about

$$\frac{d\tilde{y}}{d\psi} = \frac{dy}{ds}\frac{d\tilde{y}}{dy}\frac{ds}{d\psi} = 10^{-3} \cdot \frac{1}{10\sqrt{\text{meter}}} \cdot 50 \text{ meter} = 5 \cdot 10^{-3}\sqrt{\text{meter}} \tag{16}$$

3.2 Special Cases of Multiple Distortions

3.2.1 Restricted-range 2-bumps

If two distortions at μ_1 and $\mu_2 > \mu_1$ are separated by precisely $k\pi$, and their field strengths (times their β-functions) are related by $f_2 + (-1)^k f_1 = 0$ then for all phases outside of $\mu_1 < \mu < \mu_2$, the distortions caused exactly cancel. This is a "two-bump," and the concept is only of limited use in beam smoothing since in general, one has control of DFG's at fixed phases which generally don't differ by precise multiples of π.

At any rate, the shape of such a "two-bump" in rescaled postion measurement is that of k lobes of a sine curve, starting at μ_1 and ending at μ_2, and flat zero outside that range. And the peak magnitude is twice the peak magnitude of either single bump.

3.2.2 3-bumps

For three discrete distortions at three fixed positions, no matter what their relative betatron phases $\mu_1 < \mu_2 < \mu_3$ you can always adjust the magnitudes of the three kicks such that the sum of the distortions is zero outside the range $\mu_1 < \mu < \mu_3$. If we let the (fixed) phase differences be δ_{12}, δ_{13}, and δ_{23}, we find that for this to happen,

$$\frac{A_1}{A_2} = -\frac{\sin\delta_{23}}{\sin\delta_{13}}$$

$$\begin{aligned}\frac{A_1}{A_3} &= -\frac{\sin\delta_{23}}{\sin\delta_{12}}\\ \frac{A_2}{A_3} &= -\frac{\sin\delta_{13}}{\sin\delta_{12}}\end{aligned} \tag{17}$$

These three equations are not linearly independent; they represent only two constraints. Still, for a given set of three correctors at fixed positions, one has left only one free parameter. The shape of the bump is fixed (but depends of course on δ_{12}, δ_{13}, and δ_{23}); the magnitude (at some given point) can be varied.

The shape of a 3-bump looks like a pair of sections of two different sine curves, joined at the middle position μ_2. Every such shape (with the proviso that the frequency of each sine curve is the universal frequency, 1 in our definition of μ) can be achieved, but this represents little freedom since two curves of arbitrary magnitude won't necessarily meet at μ_2. In real life, one cannot adjust the positions of the DFG correctors, so only a very limited selection of 3-bump shapes can be attained.

When introducing intentional displacements, it will be useful to know how far a beam is displaced at μ_0 due to a 3-bump at positions (μ_-, μ_0, μ_+), given a distortion of unit magnitude at μ_0. This is calculated by setting A_2 in equations (17) to 1, using (17) to find the $(A_- = A_1, A_+ = A_3)$ that produce a 3-bump with $A_2 = 1$, and adding three contributions of the form of equation (2) with $\mu_j = \mu_0$ and $\mu_i = \{\mu_-, \mu_0, \mu_+\}$.

$$\Delta(\mu_0) = \cos Q\pi - \frac{\sin(\mu_+ - \mu_0)}{\sin(\mu_+ - \mu_-)}\cos((\mu_0 - \mu_-) - Q\pi) - \frac{\sin(\mu_0 - \mu_-)}{\sin(\mu_+ - \mu_-)}\cos((\mu_+ - \mu_0) - Q\pi) \tag{18}$$

Using a trigonometric identities for the cosine of sums of angles, this can be simplified:

$$\Delta(\mu_0) = -\frac{2\sin Q\pi \sin(\mu_0 - \mu_-)\sin(\mu_+ - \mu_0)}{\sin(\mu_+ - \mu_-)} \tag{19}$$

Important consequences of equation (19) are seen when taking the two outside positions to be out of phase by nearly π. When $\mu_+ - \mu_- \approx n\pi$ the denominator in (19) blows up.

- There is no possible 3-bump with non-zero kick at μ_0 if $\mu_+ - \mu_- = n\pi$. Instead, one can form a 2-bump at $(\mu_,\mu+)$.
- If $\mu_+ - \mu_- \approx n\pi$ but not exactly out of phase by $n\pi$, then a three-bump can be formed, but the kick at μ_0 will be tiny relative to the overall displacement of the bump.

A related situation with even more important consequences arises when the center μ_0 is very close to (or 180 degrees out of phase with) either μ_- or μ_+. Then one of the sine functions in the numerator becomes very small.

- If $\mu_0 \approx \mu_-$ then although a 3-bump with unit magnitude at μ_0 can be formed, it requires very large distortions at μ_- and μ_0.
- If $\mu_0 - \mu_- \approx n\pi$ for non-zero n, then the 3-bump which gives a unit magnitude at μ_0 not only requires very large distortions at μ_- and μ_0, it also produces very large displacements between μ_- and μ_0.
- Similar consequences arise if $\mu_+ - \mu_0 \approx n\pi$.

3.2.3 4-bumps

For four discrete distortions, there is one additional degree of freedom, which can be used to adjust the magnitude and slope at a given point. This vastly increases the number of bump shapes which can be achieved with DFG's at fixed positions.

The question arises: Can every 4-bump be expressed as the superposition of two 3-bumps? The degrees-of-freedom counting suggests this, and a naive argument would imply this is so: When condition on the amplitudes of corrections for a 3-bump at (μ_1, μ_2, μ_3) , that ensures that the distortion is zero to the left of μ_1, is met, that same condition ensures that the distortion is zero to the right of μ_3. Similarly for (μ_2, μ_3, μ_4); so one might think that in order to be zero outside of (μ_1, μ_4) the 4-bump must be a sum of those two 3-bumps.

The hole in that argument is that one can imagine a "bad" 3-bump at (μ_1, μ_2, μ_3) which gives sine-curve distortions outside of (μ_1, μ_3), and a similarly flawed 3-bump at (μ_2, μ_3, μ_4) such that the outside distortion on the left and right sides precisely cancel with those of the first 3-bump. Can this possibility be used to grant additional freedom, to create a valid 4-bump which is not the sum of two valid 3-bumps?

In general the answer is no. For any given three correctors, there is a relationship between the positions of the correctors, and the phase shifts of the leftover sine curves of the left and right sides. In general, this phase shift will be different for (μ_1, μ_2, μ_3) than for (μ_2, μ_3, μ_4), so that if the left side flaws cancel, the right side flaws cannot.

There are, however, possible placements of $(\mu_1, \mu_2, \mu_3, \mu_4)$ such that the phase shifts would be equal; in particular, uniform spacing has this property. In such a case, can one form a 4-bump which is not the sum of two valid 3-bumps? In that case, what becomes of the degrees-of-freedom counting?

The answer to the dilemma is that even for specialized spacing of the correctors, one cannot form a 4-bump which is not the sum of two valid 3-bumps. Yes, you can choose amplitudes (A_1, A_2, A_3) for the bump at (μ_1, μ_2, μ_3), and (B_2, B_3, B_4) for the bump at (μ_2, μ_3, μ_4), such that the left and right flaws cancel out. When you do this, however, you will always be able to transfer some amplitude from A_2 to B_2 and from A_3 to B_3 such that although you have the identical 4-bump, now each 3-bump is no longer flawed.

Thus, we come to the conclusion that *every 4-bump which is zero outside the range $\mu_1 < \mu < \mu_4$ can be expressed as the sum of a 3-bump which is zero outside the range $\mu_1 < \mu < \mu_3$ and another 3-bump which is zero outside the range $\mu_2 < \mu < \mu_4$.*

4 Distortion Due to a Distributed Dipole Field

Consider a dipole field which is not concentrated at a single point, but instead is distributed in betatron phase. Then many contributions from equation (1) must be integrated to get the net distortion.

In particular, say the kick from each point x is given by equation (1) times some $2\sin(\pi Q) \cdot \rho(x)/\sqrt{\beta(x)}$, where we divide out the square root of the β-function so that a given value of $\rho(x)$ represents the same strength of distortion at any point x. Then when you integrate over x, equation (2) takes the form

$$a_j = \int dx\, \rho(x) \cos(|\mu(x) - \mu_j| - \pi Q) \tag{20}$$

or, working with betatron phase μ instead of position x and assuming $\rho(\mu)$ is non-zero only in the range $a \le \mu \le b$,

$$a_j = \int_a^b \rho(\mu)\cos(|\mu - \mu_j| - \pi Q)\, d\mu \tag{21}$$

Outside the range $a \le \mu_j \le b$, the expression for which absolute value is taken in equation (21) is either strictly positive or strictly negative, and if $\rho(\mu)$ is some simple known function, the integral can be performed analytically. For example, if $\rho(\mu)$ is linear (or linear in each of N sub-regions) then the integral is easy to do, and works out to some factor times $\cos(\mu_j + \delta)$ where the factor and phase angle depend on the details of $\rho(\mu)$. An interesting case is to take a Gaussian shape: although this is strictly speaking not confined to a finite range, it dies off quickly and if we are far enough away from the "origin" the integral for $\rho(\mu) = e^{-c\mu^2}$ is $e^{-c}\cos(\mu + \pi Q)$. The general pattern persists, that outside the range, the effect of the distortion is always of the form $A\cos(\mu_j + \delta)$, which is equivalent to that of a single discrete distortion somewhere inside the range. This equivalence theorem can in fact be proven.

In fact, outside the range of an arbitrary distorting dipole field of finite extent, you can always mimic that field by a single discrete distortion, positioned somewhere within the range and of magnitude less than or equal to the integrated field. An easy way to visualize this is to pretend the distributed field were a large number of discrete distortions and express all distortions as a vector where the angle is the phase is the phase of the distortion. Then (outside the range) the net distortion is represented by the vector sum of the individual distortions. The properties asserted above immediately follow from that fact.

Inside the range $a \le \mu_j \le b$, the integral must be broken into two separate integrals, since the absolute value expression changes sign at $\mu = \mu_j$. This has two effects: The answer is no longer a simple cosine curve, and the integrals are a bit less likely to have simple answers. (For example, the Gaussian-shaped distortion gives bumps whose shape is a horrible combination of error functions and exponential integrals.) This motivates the consideration of only the simple linear case: $\rho(\mu) = m\mu + c$.

A priori, one would expect that within the range of a distributed distortion, the phase of the distortion would vary with x, and that for slowly varying fields, this could be expressed as a cosine curve with a slightly "incorrect" frequency which varies slowly. It would be tempting to fit BPM data to a function of this form, to take into account the possibility of a weak but widespread distortion. This would probably not bear fruit, for the following reason: The combination of an underlying cosine curve plus another cosine curve of different frequency gives a total of at least 8 parameters to be determined: The 3 parameters for the underlying curve, plus the magnitude, frequency and phase of the second curve, plus starting and ending points for the distortion. Thus one needs at least 8 points to simply determine the fit, and since the whole idea is rejection of outlier measurements based on statistics of nearby measurements, probably at least 12 points to do anything practical. This amounts to more than a tenth of the ring; the assumption that the distorting field is almost constant over that large a region is shaky.

But if you do the algebra for $\rho(\mu) = m\rho + c$, you find that the expression is:

$$\begin{aligned} m\cos(b - \pi Q - \mu) + (bM + c)\sin(b - \pi Q - \mu) \\ -m\cos(a + \pi Q - \mu) - (aM + c)\sin(a + \pi Q - \mu) \\ +2(m\mu + c)\sin(\pi Q) \end{aligned} \tag{22}$$

In this expression, all terms but the last can be absorbed into a cosine curve with frequency one and some phase and amplitude. There is no "incorrect frequency" term. The last term is merely a linear correction. In fact, if m were zero, the entire distortion would not change the form–equation (11)–to which we were already fitting. The last term contributes to k in that equation.

This is fortuitous, because although the equivalence theorem mentioned above always allows you to substitute discrete set of distortions, one between each BPM, for continuous distortion fields, such a course would have rendered all analysis of BPM measurements (to reject outliers) useless. On the other hand, fitting to the constrained form of equation (11) gives an excellent framework in which to search for outliers.

The bottom line for distributed distorting dipole fields is:

- They don't alter the form $A\cos(x+\delta)+k$ to which we should fit the BPM measurements.
- When you cross the end or beginning or such a region, A, δ and k of the fit will change–but the same is true when you cross the location of a discrete distortion.
- To some extent, one can identify the k found in such a fit with an indication of a distorting field distributed in betatron phase.

5 Correcting For Distortions

5.1 The Tevatron

The Tevatron ring has a tune $Q = 20.585$ (the horizontal and vertical tunes are very slightly different, but for purposes here we will just look at horizontal). In the work in this section, we can always ignore the even-integer part of the tune; thus we will work with $Q = .585$.

The roughly 118 DFG/BPM pairs are spaced at intervals of about 67 degrees. If one just looks at the β-function at the positions of the DFG's and BPM's, the appear to vary fairly smoothly; and if you examine α (representing $-1/2$ times the derivative of β) it naively seems to be much too large to be consistent with this smooth variation.

Looking at the lattice (MAD) file for TEV15 (for example), we see that aside from major increases in the β-function near CDF and D0, the nature of $\beta(\mu)$ is a sequence of dips that look like parabolas. Each parabola meets its neighbors in a cusp around $\beta = 100$ and dips down to about $\beta = 30$; there are typically about 10% variations in the heights of both the peaks and the troughs, so there is an appearance of slight asymmetry.

The placements of the BPM's and DFG's straddle the peak. Each BPM is slightly before the peak, with a value of α which is high and negative. Each DFG is slightly after the peak, where α is high and positive.

The nature of the lattice is that of a long bending dipole, then a BPM, then a thin focusing quadrupole, then a DFG to correct steering errors, and repeat. Thus β is highest (and its derivative is given a huge kick) right at each quadrupole.

Steering errors (what I call distortions) can originate in the quads or in the long dipoles. The former sort of error will be well approximated by a discrete distortion, which occurs very near the DFG. It may be better to consider the distortions in a dipole as being smoothly distributed.

5.2 Discrete Distortions

Always working with positions normalized by dividing out the β function, a discrete distortion can be recognized by a sinusoidal perturbation to the left of the distortion, and another to the right. The distortion (or 1-bump) at μ_0 is of the form

$$A\cos(\mu_0 - \mu - Q\pi) \equiv A\cos(\mu - (\mu_0 - Q\pi)) \tag{23}$$

to the left of μ_0 and

$$A\cos(\mu - \mu_0 - Q\pi) \equiv A\cos(\mu - (\mu_0 + Q\pi)) \tag{24}$$

to the right. From this we deduce that

- The magnitudes of the two cosine curves are equal.
- The phases are shifted by $2\pi Q$.

Actually, any distributed distortion which is confined to a region not including any of the BPM's themselves will exhibit the properties just described. If we assume the distortion is a 1-bump, however, we can also deduce the position of the cusp if we know the phase of the curve on the left (or right):

$$\mu_0 = \delta_{\text{left}} + Q\pi = \delta_{\text{right}} - Q\pi \tag{25}$$

Then we know how to use the nearest DFG on each side of μ_0 (which we say are at μ_- and μ_+) to "perfectly" correct for this distortion (in the sense that outside the range (μ_-, μ_+) the corrected perturbation is zero). This is done by choosing amplitudes A_- and A_+ such that the three distortions caused by the two DFG's and the distortion at μ_0 from a 3-bump:

$$\begin{aligned} A_- &= -A\frac{\sin(\mu_+ - \mu_0)}{\sin(\mu_+ - \mu_-)} \\ A_+ &= -A\frac{\sin(\mu_0 - \mu_-)}{\sin(\mu_+ - \mu_-)} \end{aligned} \tag{26}$$

When we apply this "perfect" correction to a distortion which happens to be at the midpoint between two DFG's separated by 67 degrees, we see (figure 1) that in the region (μ_-, μ_+) the correction is not very good.

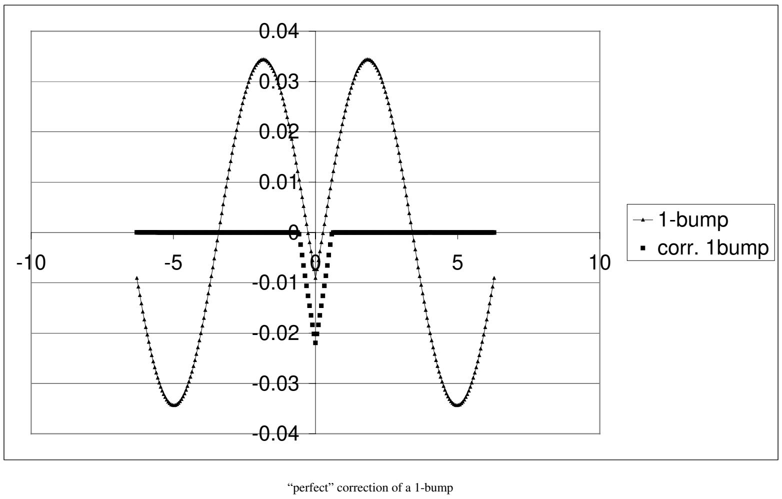

Figure 1: 1-bump in center of 67-degree range

(The straight-line appearance of the two sides of the "V" in figure 1 comes from a pair of sine curves–below 30 degrees a sine curve does look very straight.)

If the figure of merit is how much the mean squared perturbation has been reduced, this correction has done quite well, since in the 20 cycles around the ring it has quashed it completely. But if the figure of merit is the most extreme perturbation, then this correction has only reduced that by a factor of .64. Reality is somewhere between these goals, but in the rest of this discussion I will consider the figure of merit to be reduction of the most extreme perturbation–this is the harder goal to achieve.

What has happened is that in correcting the exterior shape, we have pushed too hard on the interior shape. Instead, we can under-correct, balancing a remaining exterior sinusoid against the interior V-like shape to optimize the worst-case deviation. In figure 2 we have

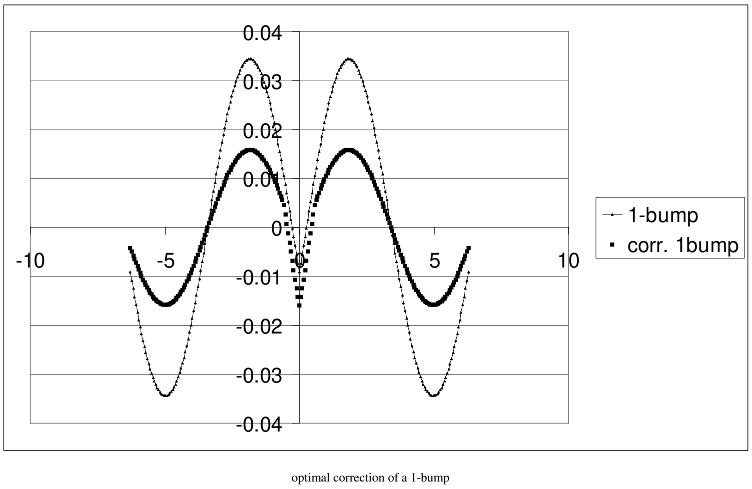

Figure 2: 1-bump, under-corrected by a factor of .54

applied only about half the “perfect” correction, and thereby reduced the most extreme perturbation by a factor of .47 instead of .64.

If the distortion is near to one of the correcting DFG's, the situation is much better. For example, if we place the distortion at 5 degrees from the left-side corrector, then the naive "perfect" correction (shown in figure 3) achieves a reduction of the worst deviation by a factor of .157.

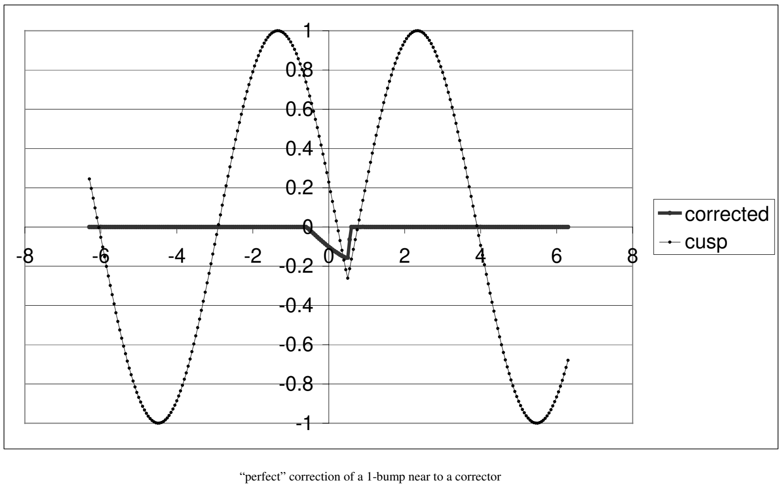

Figure 3: 1-bump at left side minus 5 degrees, "perfectly" corrected

In that situation, we can improve the factor to .150 by over-correcting by 12% (see figure 4), but this improvement is much smaller.

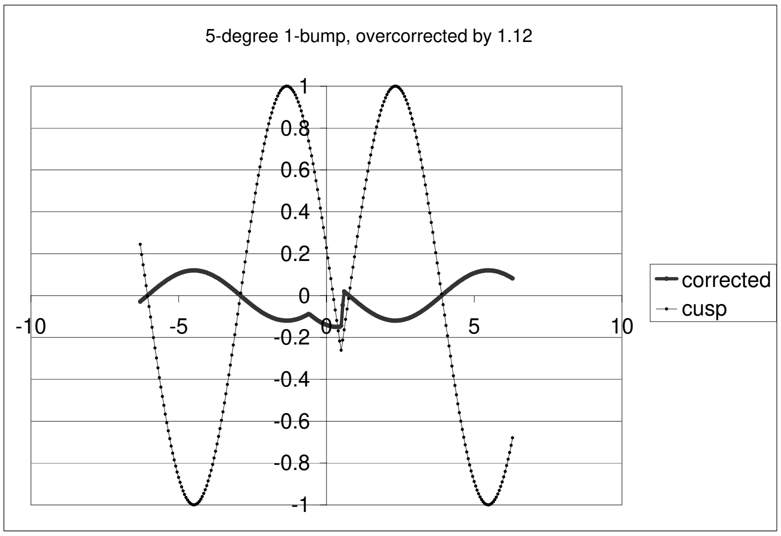

Figure 4: 1-bump at 5 degrees, over-corrected by a factor of 1.12

5.3 Distributed Distortions

It is plausible that most distortions occurring some distance from the DFG/BPM pairs would be due to some sort of mis-alignment or mis-setting or flaw in a dipole. Since the dipole magnet is a long object, these distortions ought to be considered as being distributed in space, rather than as delta functions. However, since a single dipole magnet is completely between two consecutive measurement points, as mentioned earlier, the effect of the distortion on the orbit outside that range can be *precisely* mimicked by a single delta function distortion somewhere between those measurement points.

On the other hand, it is conceivable that due to some systematic effect there could be a distorting field which extends across multiple measurement points. For example, if the earth below some fraction of the ring has shifted by a few millimeters, such that the magnets tilt or roll by a tiny amount, that could manifest as a small distortion in the vertical or horizontal field which extends over a large spacial region.

The mathematics of a distributed distortion is not intractable if we can consider the shape of the distorting field as roughly piecewise linear. And this assumption is not that unreasonable; even if we had to consider in detail the effective distortions between measurement points, if we assume some sort of constant deviation in the dipole field, then after correcting for the β-function, the approximation of a pair of linear distortions is excellent.

Figure 5 shows $\sqrt{\beta(\mu)}$ in typical areas of the ring; the pair of lines can approximate this to 10% accuracy or better. And since to first order the distortion contributions within each line are nearly in phase, this means that if the mean value in the line matches that of $\sqrt{\beta(\mu)}$, the perturbation will be approximated with accuracy on the order of 1%, which is more that adequate for our purposes.

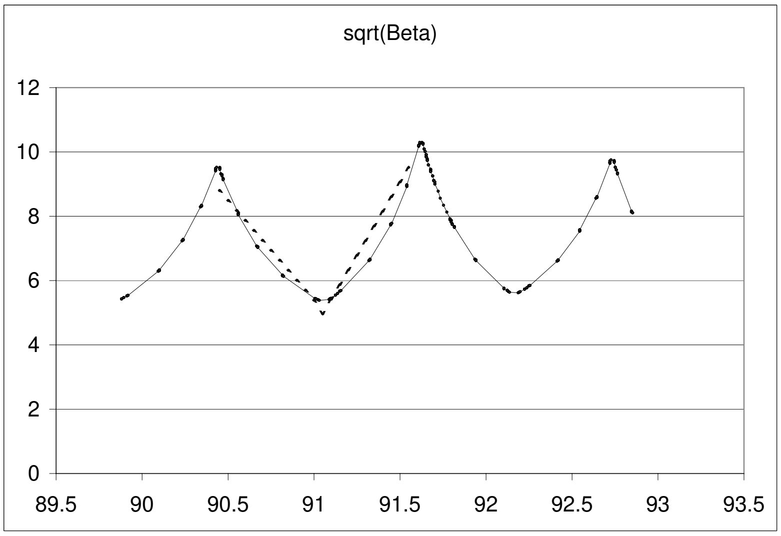

Figure 5: Square root of beta, fit to a pair of lines

In the end, however, we don't even need to go to this level of fidelity, since (as before) any configuration of field distortion confined to the region between two consecutive measurement points can, for purposes of computing the orbit outside those points, be faithfully replaced by a single delta function distortion. The only thing we need to consider at all is the behavior of the β-function-correction-averaged distortion at a scale larger than the inter-BPM distance.

5.3.1 Mathematics of Distributed Distortions

Assume, now, that the center of the range of a distortion is at $\mu = 0$, and that the distortion at any given betatron phase x is given by:

$$\begin{cases} -m_L x + c & a < x < 0 \\ m_R x + c & 0 < x < b \\ 0 & \text{elsewhere} \end{cases} \tag{27}$$

Then the perturbation at any point μ is given by the sum of two or three integrals over the point of the distorting field x:

$$\begin{cases} \int_a^0 (-m_L x + c)\cos(x-\mu-Q\pi) & \text{if } x<a \\ \int_a^\mu (-m_L x + c)\cos(x-\mu-Q\pi) + \int_\mu^0 (-m_L x + c)\cos(\mu-x-Q\pi) & \text{if } a\le x<0 \\ \int_0^\mu (-m_R x + c)\cos(x-\mu-Q\pi) + \int_\mu^b (-m_R x + c)\cos(\mu-x-Q\pi) & \text{if } 0\le x<b \\ \int_0^b (-m_R x + c)\cos(x-\mu-Q\pi) & \text{if } b\le x \end{cases} \tag{28}$$

Outside the "active" distortion range (a, b), each of the two integrals contributes a sum of sines and cosines, each argument of which is the betatron phase plus some phase offset. We can use the trigonometric identities for functions of sums of two angles to decompose into some μ-independent factor times $\cos\mu$ plus another μ-independent factor times $\sin\mu$. For $\mu < a$, the perturbation p is given by

$$p = \begin{array}{l} \left\{ \begin{array}{l} m_R\cos(Q\pi-b) - m_L\cos(Q\pi-a) \\ +(am_L+c)\sin(Q\pi-a) - (bm_R+c)\sin(Q\pi-b) \\ +(m_L-m_R)\cos(Q\pi) \end{array} \right\} \cos\mu + \\ \left\{ \begin{array}{l} m_L\sin(Q\pi-a) - m_R\sin(Q\pi-b) \\ +(am_L+c)\cos(Q\pi-a) - (bm_R+c)\cos(Q\pi-b) \\ +(m_L-m_R)\sin(Q\pi) \end{array} \right\} \sin\mu \end{array} \equiv C_L\cos\mu + S_L\sin\mu \tag{29}$$

For $\mu > b$,

$$p = \begin{array}{l} \left\{ \begin{array}{l} m_R\cos(Q\pi+b) - m_L\cos(Q\pi+a) \\ -(am_L+c)\sin(Q\pi+a) + (bm_R+c)\sin(Q\pi+b) \\ +(m_L-m_R)\cos(Q\pi) \end{array} \right\} \cos\mu + \\ \left\{ \begin{array}{l} m_R\sin(Q\pi+b) - m_L\sin(Q\pi+a) \\ +(am_L+c)\cos(Q\pi+a) - (bm_R+c)\cos(Q\pi+b) \\ +(m_L-m_R)\sin(Q\pi) \end{array} \right\} \sin\mu \end{array} \equiv C_R\cos\mu + S_R\sin\mu \tag{30}$$

In some sense, formulas (29) and (30) are not very useful in real life, since nobody is going to hand you the parameters describing the distortion (m_L, n_R, c). But they are useful in that they allow you to study what happens when a distorted bump is corrected, because knowledge of C_L and S_L (or of C_R and S_R) is sufficient to determine the amplitude $(\sqrt{C_L^2 + S_L^2}$ and position in betatron phase $(Q\pi + \arctan(C_L/S_L))$ of the "equivalent" one-bump. Knowing the position and amplitude of the one-bump that matches the exterior behavior of this distortion, we can apply equations (26) to determine the necessary correction strengths at μ_- and μ_+. We can then add that to the distortion in the interior region, to see the shape of the resulting corrected perturbation.

In the interior region, for $x < 0$ the perturbation is:

$$p = \begin{array}{l} \{m_R\cos(Q\pi-b) - (bm_R+c)\sin(Q\pi-b) - m_R\cos(Q\pi)\}\cos\mu \\ +\{-m_R\sin(Q\pi-b) - (bm_R+c)\cos(Q\pi-b) - m_R\sin(Q\pi)\}\sin\mu \\ -m_L\cos(\mu-Q\pi-a) + m_L\cos(\mu+Q\pi) \\ +(am_L+c)\sin(\mu-Q\pi-a) - c\sin(\mu+Q\pi) + 2(m_L\mu+c)\sin Q\pi \end{array} \tag{31}$$

And in the interior region, for $x > 0$ the perturbation is:

$$p = \begin{array}{l} \{-m_L\cos(Q\pi-a) + (am_L+c)\sin(Q\pi-a) + m_L\cos(Q\pi)\}\cos\mu \\ +\{m_L\sin(Q\pi-a) + (am_L+c)\cos(Q\pi-a) + m_L\sin(Q\pi)\}\sin\mu \\ +m_R\cos(\mu+Q\pi-b) - m_R\cos(\mu-Q\pi) \\ -(bm_R+c)\sin(\mu+Q\pi-b) + c\sin(\mu+Q\pi) + 2(m_R\mu+c)sinQ\pi \end{array} \tag{32}$$

As noted in the discussion of equation (22), all terms except the $\sin Q\pi$ terms can be absorbed into a cosine curve with frequency one, but unlike the case for outside (a, b), we have found no advantage in grouping the various frequency one terms into the form $C \cos \mu + S \sin \mu$.

5.3.2 Correcting Distributed Distortions

Take as a first example a perturbation caused by a distorting field extending all along the dipole magnet–in our example, 91% of the way from μ_- to μ_+. We will take the bottom of the β function to be a third of the peak value, and also assume that $m_L = m_R$.

As discussed above, it is easy to determine the "perfect" correction for this bump, and figure 6 shows the perturbation and the corrected trajectory.

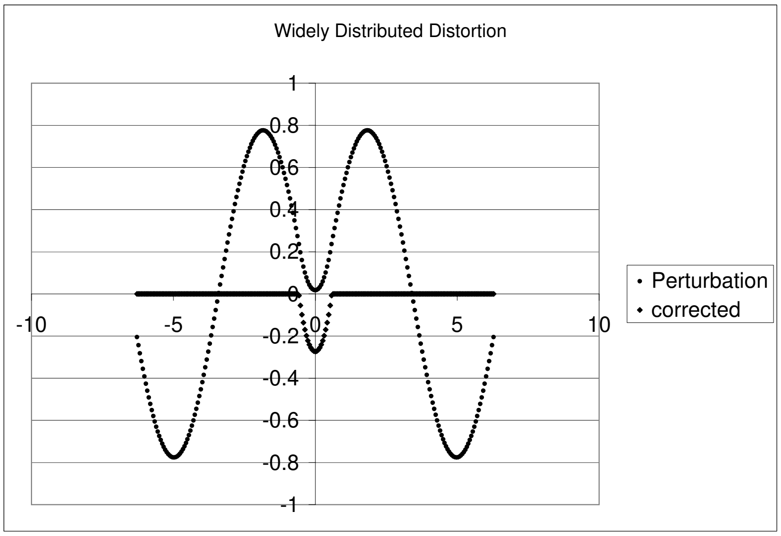

Figure 6: Perfect correction of a distortion spread over 91% of range

The "perfect" correction achieves a reduction in extreme perturbation by a factor of .35. This is 1.8 times better than in the case of the delta-function distortion (true 1-bump).

One item which might be of practical interest in the Tevatron: For some stretch, it is desirable to move the orbit position by some distance in one direction, due to obstructing equipment in one small part of the ring. Observe the shape of the corrected perturbation in figure 6. It is just about the right shape to give the desired orbit change, assuming the region of desired change doesn't span a DFG. What this says is that one could achieve this sort of orbit modification much more accurately if one were free to intentionally introduce a small distributed dipole kick (that is, by an intentional very slight mis-alignment of one dipole). We can say this in a more negative way: Although you can fool yourself into thinking you have introduced a desired local shift of the center orbit by adjusting DFG's (creating a 3-bump), in fact the correction is not good at all. With the actual 67 degree spacing between DFG's, the 3-bump looks like a sharp cone, and not at all like a flat shift in position.

As in the case of the true 1-bump, we can improve matters by under-correcting. Figure 7 shows the effect of a 75% correction, which achieves a reduction factor of .26.

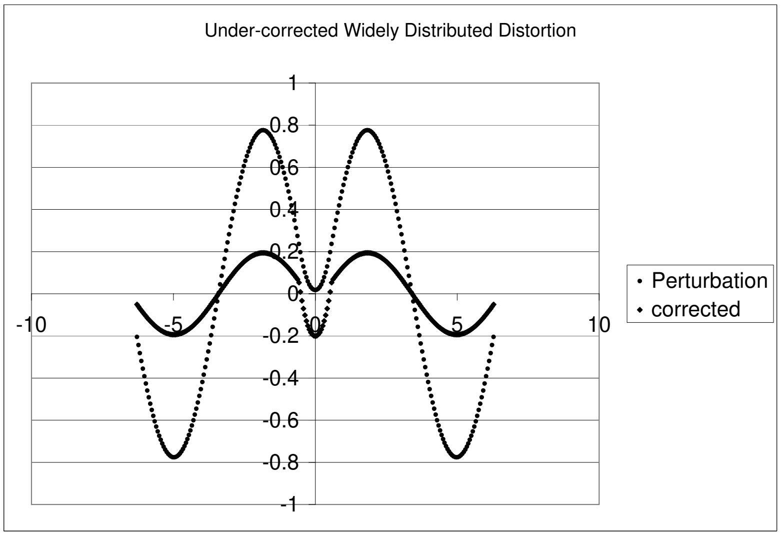

Figure 7: Optimal correction of a distortion spread over 91% of range

5.4 Conclusion–How Well Can Distortions Be Corrected?

We will assume that perturbations have causes other than mysterious monkeys inserting a small but powerful dipole magnet at an arbitrary position; that is, we will consider distortions which have something to do with the dipole or quadrupole ring magnets. We are interested in correcting perturbations specifically

5.4.1 Extreme-value Norm

Consider first the toughest goal, that of minimizing the worst-case perturbation.

The DFG's and BPM's come in pairs, only about 5.6 degrees apart in betatron phase. The quadrupole focusing magnets are in the center of that gap, and thus are about 2.8 degrees from a correcting DFG. If a distortion is cause by some mis-alignment or mis-strength or other defect in a quad, then we can't consider it to be the (more favorable case of) a broadly distributed distortion. Still, even considering a delta-function distortion (a true 1-bump), we can achieve a correction by a factor of 0.09 with the naive "perfect" correction, and by a slight over-correction we can improve that to a factor of 0.082.

A distortion due to a dipole magnet is tougher to cure, since it can have an effective position which is not close to either of the straddling DFG's. If the problem were a delta-function kick right in the middle of the dipole magnet, then one could only correct matters by a factor of 0.64. Things are much better than that, for two reasons:

- The distributed bump intrinsically can be corrected 1.8 times better than the delta function.
- The freedom to under-correct helps (in the worst case of a centered distortion) by a factor of about 1.3, whether the distortion is spread or a delta function.

Overall, one can correct a realistic distortion caused by a problem in a dipole magnet, by a factor of about 0.26. It is sobering to look at this in a more negative light: *Using the extreme-value criterion, you can never beat down the steering errors (of this sort) by more than a factor of four.*

5.4.2 RMS Norm

Matters are much more favorable if the objective is to reduce root mean squared deviation from the normative orbit. Since the region outside the distortion/correctors range amounts to about 120 times the size of the range, trading improvements inside the range for degradation outside the range is not worthwhile. The "perfect" correction which zero's the value of the function outside the active range is the best correction to use. Assuming that a bump lying far from either DFG will be caused by a distributed distortion, this achieves an improvement by a factor of .35 for the size of the bump in the active range.

The lever arm of two orders of magnitude buys excellent results. The perturbation arising from any single discrete or distributed distortion can be reduced (in RMS measure) by a factor of roughly .35/120, that is, corrections can improve matters by a factor of 300 or more.

5.4.3 Multiple Distortions

Say, first, that we are interested in the RMS norm. Assuming that the perturbations in the ring can accurately be modeled by a series of discrete or distributed distortions, then we can achieve that factor of 300 for each distortion. This is obvious in the case where the distortions are relatively sparse (such that we can ignore interference between corrections for distortions in two adjacent ranges), but also holds (due to superposition) for any set of distortions.

However, the net improvement factor in the case of multiple distortions is not as good as in the case of a single distortion, for the following reason: The RMS perturbation due to the sum of N randomly chosen distortions grows as $\sqrt{N}$, not as N. Thus if we observe some pattern of deviation from the normative orbit, we can expect proper corrections to improve matters by a factor of $.35\sqrt{N}/120$, which ranges from a factor of thirty if there are lots of distortions, to a much greater factor if the deviation was caused by just a few identifiable distortions.

Now consider the extreme-value norm. Computing the distribution of the worst-case perturbation due to the sum of N randomly chosen distortions is not an easy mathematical problem, but we can get a decent estimate of the potential correction improvement factor as follows:

- For more than one or two distortions, we immediately lose the gains associated with under-correcting, because the variations outside the active range potentially can add to the variations associated with other corrected distortions, leading to worse deviations at at least one point.
- Thus we can expect corrections again by a factor of .35 in each active range.
- If each distortion were approximately of equal size, then the typical perturbation in any range due to the sum of N distortions would be $\sqrt{N}$ times the typical contribution from any single distortion.
- This estimate is corrected in two opposing ways: The perturbations are not of equal size, so the $\sqrt{N}$ factor is overly optimistic; but the sum needs to be done for N different cases (since as you cross each cusp you have a different summation) and this increases the estimate of the original worst deviation. Thus we can in fact assume that the worst-case perturbation of the overall sum is about $\sqrt{N}$ times the worst-case perturbation due to a single distortion.
- So the correction factor is approximately $.35/\sqrt{N}$.

We see that using the extreme-value norm, the potential correction ranges from a factor of about 4 for a single distortion (where under-correction is useful), up to a factor of about 30 in the case of a maximal number of distortions.

Note that in the case of lots of distortions, both the correction improvement and the specific corrections to use match for the RMS and extreme-value norms. It is in the case of very few distortions that the choice of norm profoundly affects matters. If there are very few distortions, one can correct perturbations quite well if judging by the RMS norm, but not very well if judging by the extreme-value norm; and the optimal corrections will differ.

5.4.4 Intermediate Norms

It is probably important to determine what the cost (in reduced luminosity) is for a given perturbation of the normative orbit at a given location in the ring. For example, assuming a Gaussian profile of particle density, the proper figure of merit might best be approximated by integrating some power series in the perturbation, plus some very steeply rising function which comes into play when the beam is unacceptably close to hitting something. And it is conceivable that some points in the ring will have tighter apertures than others (so that the penalty function depends on position).

However, accurately determining this information would require a considerable program of intentionally introducing small local deviations and measuring the decrease of luminosity per unit time in response to each fluctuation. If and when such program were done, it would still be difficult to say anything analytical about how the corrections should be applied.

So for the purposes of this memo, we content ourselves with studying the two norms at either end of the spectrum: The "worst deviation" norm and the root mean square norm.

For practical purposes, one or the other of these cases must give at least the qualitative and likely also the quantitative information we need.

5.4.5 Non-modelable Perturbations

This entire section about correcting for distortions has been predicated on the assumption that one can accurately model the deviation from the normative orbit, as perturbations due to a sequence of discrete and/or distributed Courant-Snyder distortions.

This is not necessarily the case. Since we are forced (by the fact that we have only a finite number of BPM measurements) to consider only distortions which have equivalent expressions as a finite set of distortions, one per inter-BPM region, there are perturbations which cannot possibly be expressed as such a sum of 1-bumps. To the extent that this is the case, any correction will leave an additional error equal to this modeling error.

However, many shapes which naively might be considered tough to model actually can be well approximated, if one is willing to overlook deviations occurring on the scale of the inter-DFG betatron phase advance.

Consider, for example, a case where the desired or golden "normative" orbit contains a long-ish region in which one wants to move the beam by a few millimeters off the perfect central orbit. So relative to the desired orbit, when the beam is on the central orbit, it looks like we have a rectangular perturbation. Such a perturbation is *not* expressible as the sum of 1-bumps, and is not well-correctable by any settings of the finite number of DFG's. But you can (with DFG's spaced at 67 degrees apart) correct for a (very nearly) trapezoidal-shaped distortion, where leading and trailing edges go up and down, respectively, in a 67 degree interval. This is done by introducing symmetric 3-bumps, each of which look like triangular shapes centered on one DFG, and superposing them such that the odd and even originated bumps add up to $\sin x + \sin(67 - x)$, which is only about 10% away from being perfectly flat, being about .92 at each DFG and 1.10 at each midpoint. So you can, in fact, use the DFG's to dodge an obstruction along one side of the beam pipe. (However, when the actual Tevatron β-function is factored in, we will see that the fluctuations from perfect flatness become 20-30%.)

6 Shoving Protons: Nearly Constant Displacement of the Entire Beam

Consider the following task: To use the horizontal DFG's in the ring to displace the entire beam by some small target distance T in the radial (horizontal) direction. This section is not really of practical importance, since one would not try to move the whole beam using the DFG's; this is somewhat of a toy problem. But it does impart a good feel for the math involved in orbit changing, and it shows to what extent you can achieve the sort of orbit correction that you *think* you have achieved based on the BPM measurement points.

The strategy will be to create, centered on a given DFG, a 3-bump of the desired central amplitude T, then to create another 3-bump of the same central magnitude centered on the next DFG. Assuming the distances between DFG's are less than 90 degrees, between those two selected DFG's, the shape of the pair of 3-bumps will be a curve descending from T to 0, and a curve ascending from 0 to T, respectively. If the inter-DFG gaps are not too close to 90 degrees, the curves can be approximated by straight lines, in which case the sum is a flat line at displacement T in the entire interval between the two central DFG's. In the Tevatron, most DFG's are about 67 degrees apart, which leads to deviation from linearity of about 10%.

In the last part of this section will will lift our policy of working only with β-scaled displacements, because we will wish to see the effect of the highly non-uniform β-function on the purportedly uniform displacement we can achieve. And we will show figures resulting working with the DFG placements in the actual ring, which are highly irregular at a few key places.

6.1 Calculating the Distortions Needed at Each DFG

The sequence of calculations to form a "moved" beam is as follows:

1. Using equation (19), for each DFG at position μ_i, if a unit 3-bump with magnitude $A_i = 1$ at positions $(\mu_{i-1}, \mu_i, \mu_{i+1}$ will have displacement at μ_i of

$$\Delta(\mu_i) = -\frac{2\sin Q\pi \sin(\mu_i - \mu_{i-1})\sin(\mu_{i+1} - \mu_i)}{\sin(\mu_{i+1} - \mu_{i-1})} \tag{33}$$

2. The strength of the correction at μ_i needed to hit a target displacement at μ_i of T_i will be

$$C(\mu_i) = \frac{T_i}{\beta(\mu_i)\Delta(\mu_i)} \tag{34}$$

   The β-function appears in the denominator because this DFG will have an effective A of its correction strength $C(\mu_i)$ times $\sqrt{\beta(\mu_i)}$, and then each point μ will get an effect from μ_i which is the effective A times $\sqrt{\beta(\mu)}$.

3. The DFG at μ_i also participates as the right-hand member of a 3-bump centered at μ_{i-1}. To participate properly, it must be assigned correction strength

$$C_R(\mu_i) = -C(\mu_{i-1})\frac{\sin(\mu_{i-1} - \mu_{i-2})}{\sin(\mu_i - \mu_{i-2})}\sqrt{\frac{\beta(\mu_{i-1})}{\beta(\mu_i)}} \tag{35}$$

   The ratio square roots of the β-functions reflects the fact that you need to adjust the corrections such that the *effective* strengths have the proper ration.

4. Similarly, the DFG at μ_i also participates as the left-hand member of a 3-bump centered at μ_{i+1}. To participate properly, it must be assigned correction strength

$$C_L(\mu_i) = -C(\mu_{i+1})\frac{\sin(\mu_{i+2} - \mu_{i+1S})}{\sin(\mu_{i+2} - \mu_i)}\sqrt{\frac{\beta(\mu_{i+1})}{\beta(\mu_i)}} \tag{36}$$

   Note that the target displacements at neighboring DFG's come in indirectly, since T_{i+1} affects $C(\mu_{i+1})$ which in turn affects $C_L(\mu_i)$.

5. The effective strength of kick (A in equation (2)) due to the DFG at μ_i is

$$A_i = [C_R(\mu_i) + C(\mu_i) + C_L(\mu_i)]\sqrt{\beta(\mu_i)} \tag{37}$$

6. The net displacement at every sample point μ in the ring is then the sum of the displacements at μ caused by each DFG:

$$d(\mu) = \sum_i A_i \cos(|\mu_i - \mu| - Q\pi) \tag{38}$$

6.2 The Ideal Case: Uniform Beta

Ignoring β-function variations, one can do a very good job of moving the beam. The equations above apply to this ideal case, by setting $\beta(\mu) = 1$ (or any constant).

Consider first two consecutive 3-bumps, with target strength of 1, centered on two consecutive DFG's at μ_0 and μ_+. Between μ_0 and μ_+ the sum of those two 3-bumps takes the form of the sum of two segments of sine curves, one descending from 1 to zero, the other ascending from 0 to 1. Since for angles less than 67 degrees this sum of sine curves looks very much like a flat line (it actually has a hump in the center which is roughly 20% higher than the two ends, overall we have a shape which is decently approximated by a trapezoid, rising at μ_-, flat between μ_0 and μ_+, the falling until μ_{++}. This is illustrated for an actual section of the Tevatron:

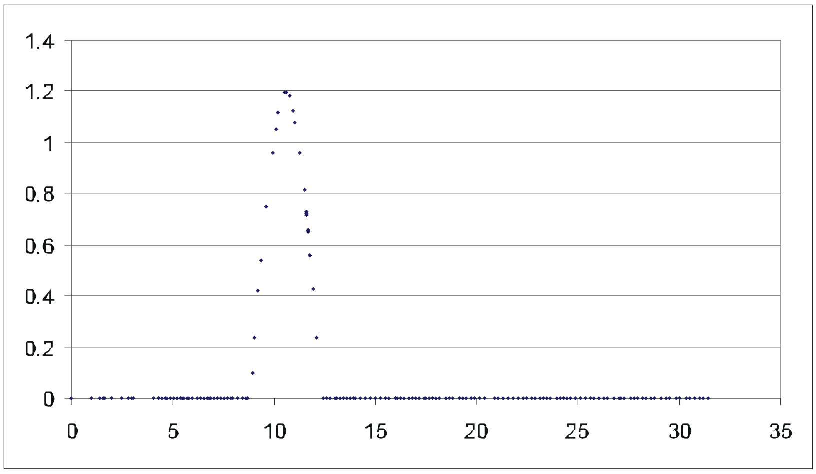

Figure 8: Four DFG's forming two 3-bumps

This does not look much like a flat line on the top, but note that between the center DFG's the variation is just 20%.

If we extend this to 8 DFG's instead of 4, the picture looks much more like we have shoved the beam:

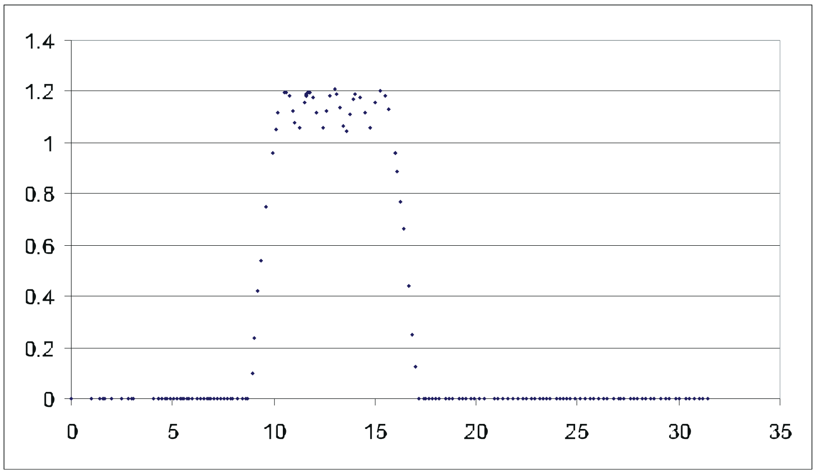

Figure 9: Eight DFG's forming a short plateau

Now let's move the target down by 10% to center on 1, and move the beam in a region spanning a quarter of the ring. Most of the deviations from flatness on top are confined to the 10% zone, but notice that at a betatron phase of about 21, the deviation is slightly greater. At this value, there is a gap where two DFG's are 77 degrees apart rather than 67; and this increase in deviation stems from that larger spacing.

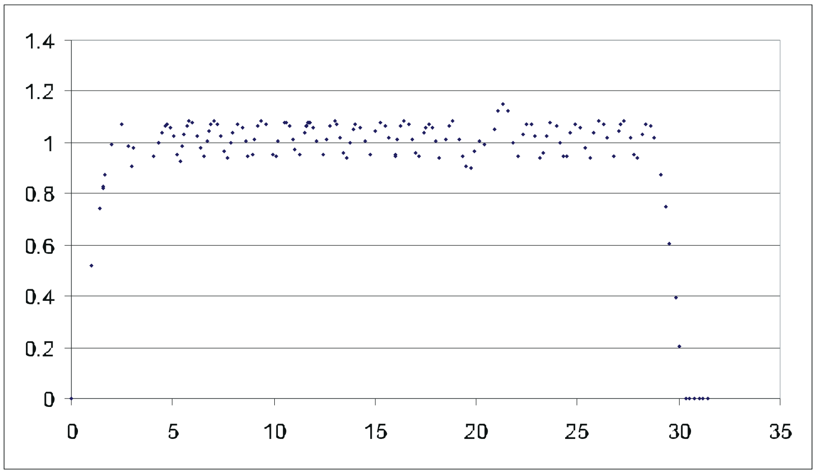

Figure 10: Moving the beam in a quarter of the ring

Next we include the second quarter of the ring, and notice utter insanity around $\mu = 60$. This is near HA48, where we have consecutive gaps between DFG's of 120, 0.65, 5 and 182 degrees. The huge swing is because of the 182 degree gap (consecutive DFG's nearly perfectly out of phase); as discussed earlier, when we try to force a unit-height 3-bump centered on one side of that gap, the deviations inside the gap are huge.

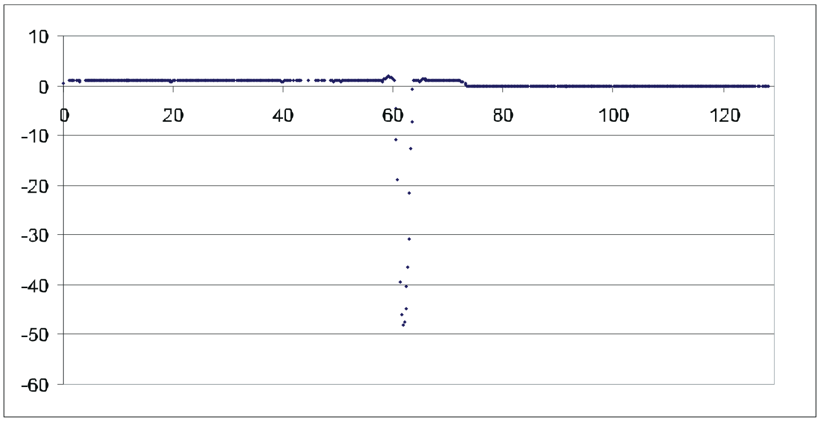

Figure 11: Naive attempt at beam moving across 182 degree gap

The best you can do in this situation, if we insist on following the mantra of letting each DFG act as a 3-bump center, involves lowering the targets assigned to C:HA48 and T:HA48 by a factor of 3, and changing the targets for the two DFG's straddling the gap (HA49 and HB11 to -1/40. This replaces the flat combination of rising and falling lines with a cosine curve which dips near zero, but does not go crazy anywhere. This gives:

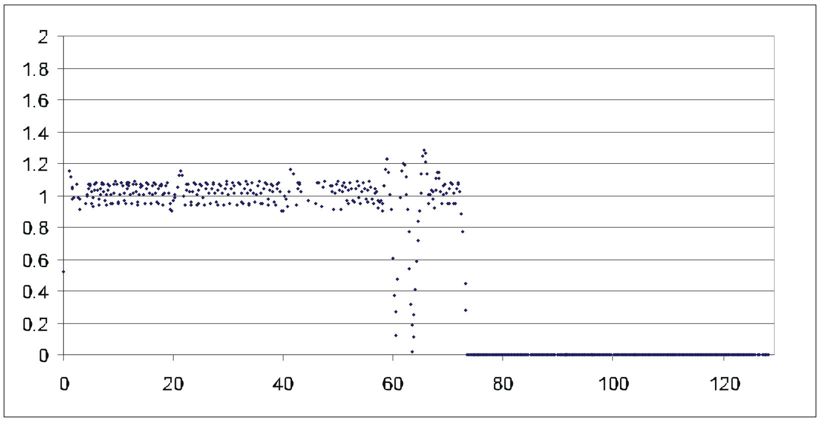

Figure 12: Improved beam moving across 182 degree gap

We can do a bit better by ignoring the presence of T:HA48, and setting targets of 0.2, 0, -0.05, and 0.8 for C:HA48, HA49, HB11 and HB12 respectively. Here we show that configuration, extending the "on" portion of the ring to about 100 to make it easier to contrast the well-behaved areas with this critical area. Note that though the top of the pattern looks better, the dips to nearly zero at either side of the 182 degree gap are still present.

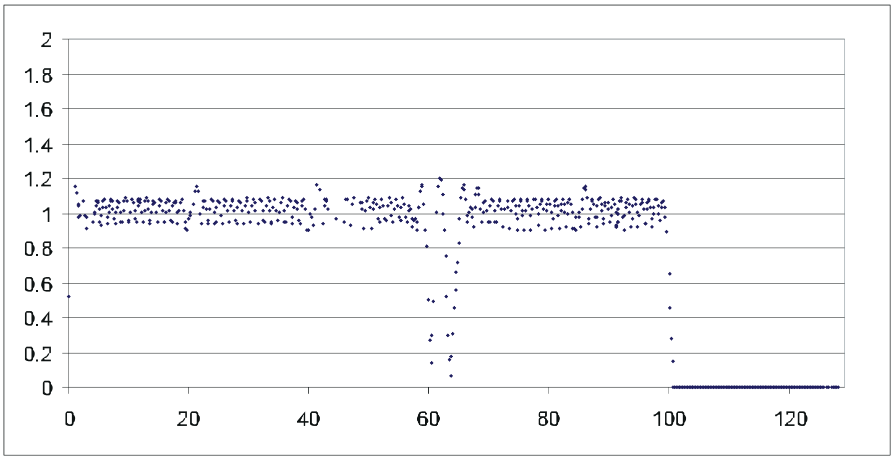

Figure 13: Best attempt near the 182 degree gap

Adding the final quarter of the ring, we again see a large deviation, this time near HC48. The DFG's there have the same highly irregular spacing as in the HA region; the same situation applies. Here we won't bother ignoring T:HC48 but will just use the 1/3 and -1/40 target settings, so we can see both sorts of corrections in these nasty areas.

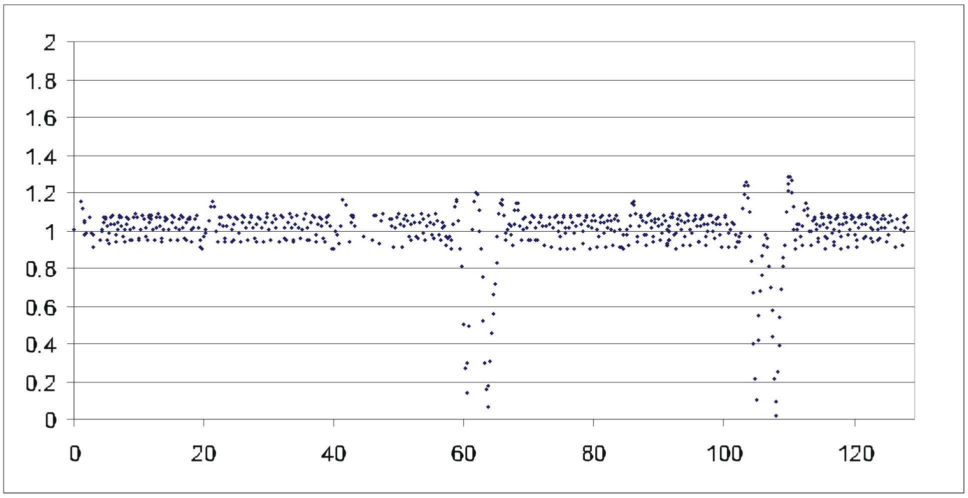

Figure 14: Beam moved along entire orbit (ignoring beta fluctuations)

Notice what this last figure represents: Other than the regions where there are large gaps between DFG's, we have, by adjusting all the DFG's (most of them to about the same strength of 0.62), moved the entire orbit by 1 unit with almost all deviations less than 10%. We can perhaps understand this physically: Had somebody altered the bending dipole magnet strengths, without changing the RF frequencies or the beam energy, then it is very plausible that the orbit radius would not remain the same. By introducing a set of discrete

DFG dipole fields that fairly well approximates a distributed dipole field around the whole ring, may may have mimicked just such a change.

6.3 Shoving Protons in the Actual Tevatron

Now we show the results when the actual β-function is included. (We used MAD data for lattice TEV15).

There is some reason to hope that we could achieve the same sort of flatness. Qualitatively, we note that $\sqrt{\beta}$ dips to a minimum midway between each pair of DFG's, and that is precisely where the naive correction has a maximum. One could hope that these bulges would largely cancel each other to give excellent flatness. But because $\sqrt{\beta}$ typically varies by a factor of about $\sqrt{3}$, while the bulge in the naive correction is only by a factor of 1.2, we are left with a 30-40% dip in the displacement between each pair of DFG's.

The first thing we can do is to apply the same set of targets that worked well in the constant-β computation above. Making no further corrections to this, we find we have not done a very good job of moving the beam by a constant amount.

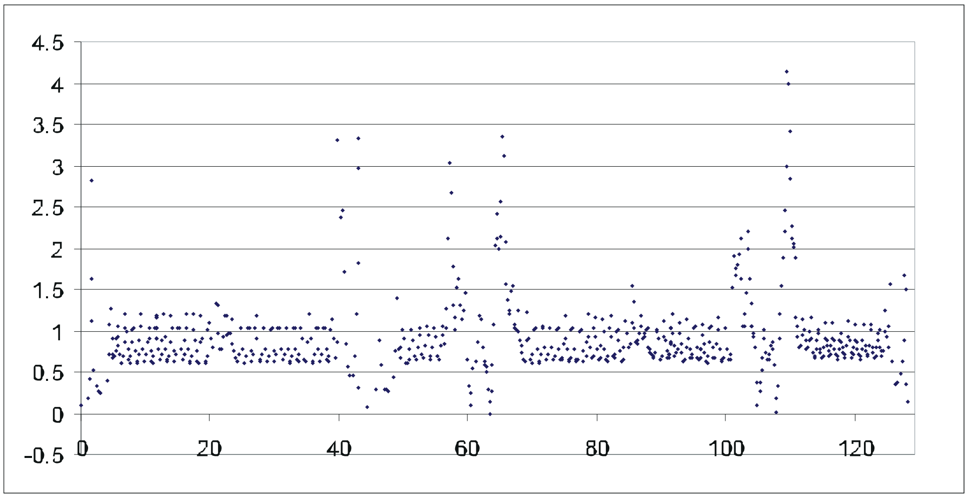

Figure 15: Moving the beta-affected beam, using naive movement targets

A fluctuation at the 30% level is unavoidable given the nature of the β-function, but we can do better on those points which skyrocket due to high β values, by carefully tuning the targets to minimize deviations. A decent set of targets has most targets set to 1.1 (to reflect the fact that due to β being smaller in most places than it is near the DFG's) the average displacement is somewhat less than 1 for unit targets. One then corrects specific DFG's to mitigate the skyrocketing displacement. When these adjustments are applied, a better pattern is achieved.

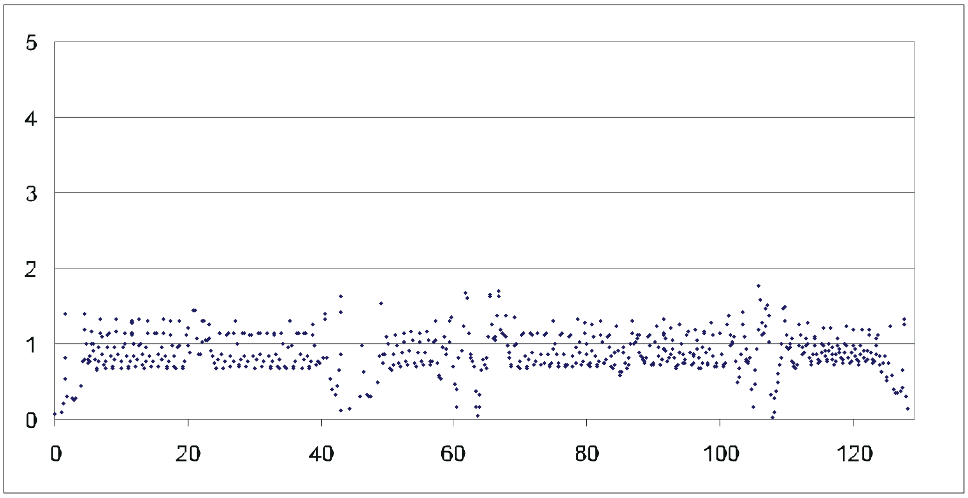

Figure 16: Moving the beta-affected beam, using adjusted targets

Here we have managed to move more than 90% of the points into an envelope of about 40% to either side of a unit displacement. But some points near irregular DFG regions remain undisplaced, while others at high β values get displaced by almost 2 units.

And this is the best we can do.

The moral is: The DFG's can be used to shift the beam, but one must remember that the variation in β prevents that movement from being very flat.

This teaches us a lesson which applies when the beam is intentionally corrected to avoid known protrusions–something which is actually done in the "golden orbits:" It may be possible to adjust the DFG settings such that if you just look at the beam position at the BPM measurement points, it appears you have nicely shifted the beam (over a short distance) by some desired distance. But because of the variations in β-function, and the nature of the corrections, the actual orbit shape may be much less flat that would appear from these measurements. This may be a danger worth worrying about.

References

[1] E.D. Courant and H.S. Snyder, "Theory of the Alternating Gradient Synchrotron," Ann. Phys. **3**, 1 (1958).

[2] D.A. Edwards and M.J. Syphers, *An Introduction to the Physics of High Energy Accelerators*, Wiley and Sons, 1993.

[3] Phys. Rev. D54, *Particles and Fields*, Amer. Isnt. of Physics, 1996.