Paul Avery
CBX 92–39
August 20, 1992

Applied Fitting Theory V
Track Fitting Using the Kalman Filter

1. Introduction

I present in this note a relatively new approach to track fitting that uses an iterative algorithm to correctly account for the effects of multiple scattering and energy loss along the track trajectory. This technique, known generally as a Kalman filter, was first applied to track fitting by P. Billoir[1] and is now being used by several CERN experiments. Some useful discussions have recently started appearing in the literature[2,3]. I implemented it in the FTMONTE package some time ago but it is not used in our standard data reduction yet. My intention is that this note will serve as a mathematical reference for an installation into DUET sometime later this year.

My discussion here focuses exclusively on track fitting rather than track finding, although it is not hard to figure out how the filter can be used to select hits as well as fit them. I have also left out a nice discussion of how the method can be applied to vertex finding and fitting. I will discuss that topic in a future note, but I will say in passing that I have implemented a very powerful set of vertex fitting routines in KNLIB based on the filter.

An extensive discussion of the motion of charged particles in magnetic fields can be found in CBX 92–45[4]. Appendices 1 and 2 provide a useful set of formulas for dealing with multiple scattering and energy loss when fitting tracks with the Kalman filter.

2. **How Tracks are Measured**

Charged particle tracks in high energy physics experiments are measured through a two step process. First, detector measurements (e.g., drift chambers, straws, MWPCs, silicon strips, etc.) are put through a pattern finding algorithm to select a subset that seems consistent with belonging to a single track. This set of measurements is then fit statistically through a maximum likelihood method to determine the most probable set of track parameters consistent with the measurements. Errors in these parameters are estimated from the measurement uncertainties and other factors, as discussed below. Measurements may be further eliminated during the track fit if they are found to be of sufficiently poor quality or not consistent with belonging to the track. Because of its well understood properties, the least squares algorithm is most commonly used to fit the track parameters and estimate the parameter errors.

In a constant magnetic field oriented along the z axis (such as found in CLEO) a charged particle moves along a helical trajectory, a path consisting of circular motion in the $x-y$ plane combined with a constant velocity along the z axis. Unfortunately, several detector related effects complicate the trajectory and therefore the fitting process: (1) inhomogeneous magnetic fields, (2) energy loss and (3) multiple scattering. I consider only dE/dx and multiple scattering corrections in this note.

Energy loss from particle interactions with detector walls and gases causes the particle to gradually decrease its radius of curvature, resulting in a spiralling-like effect for very soft tracks. The energy loss has a deterministic (average) component that can be calculated via the Bethe–Bloch dE/dx formula and a stochastic (fluctuating) piece caused by the large Landau tail of the dE/dx distribution. Correcting the mean energy loss is straightforward and is implemented at the DST level in CLEO, not in the DUET fit (see CBX 92-40 for a discussion[5]). This "after the fact" correction is approximate but works pretty well since most of the energy loss occurs before the outer drift chamber where the momentum is mainly determined. It works less well for soft tracks where the momentum determination depends more equally on measurements in the inner tracking chambers and the original track fit attempted to fit the whole track to a single helix.

As a charged particle traverses a medium it is deflected by many small angle scatters, mostly due to coulomb interactions with atomic electrons. The scattering occurs in the two perpendicular planes to the direction of motion and for small angles has a gaussian distribution in each plane (uncorrelated), hence it is known as multiple coulomb scattering or just multiple scattering. For larger angles, one should also consider the hard scattering component which adds a non-gaussian tail to the multiple scattering distribution, but we won't worry about this effect here.

3. The Standard Track Fit

To appreciate how the Kalman fitting algorithm works, let's compare it to the method used in DUET which uses a general least squares fit. The helix is specified by the 5 parameters $\alpha = (c, \phi_0, D, \lambda, z_0)$ where c is 1/2 the inverse curvature of the track, ϕ_0 is the ϕ coordinate of the momentum at the $r-\phi$ point of closest approach to the origin, D is the signed distance of closest approach to the origin in $r-\phi$, $\lambda = \cot\theta$ and z_0 is the z position at the distance of closest approach to the origin. The equations of motion of tracks using these parameters are described in CBX 92–45. The general least squares algorithm was discussed in detail in CBX 91–72[6] and is summarized as follows using matrix notation. Suppose we want to fit a set of n measurements $\mathbf{y} = (y_1,\, y_2,\, \ldots\, y_n)$ to a set of m parameters α ($m = 5$ here) through the relation $y_l = f_l(\alpha)$ for $1 \le l \le n$. In track fitting, the meaning of y_l depends on the type of detector, i.e., it can be the distance of closest approach to a wire (drift chambers and straws), the z coordinate (cathode strip) or a distance from a metal strip in a plane (silicon strip detector).

Since the $f_l(\alpha)$ are nonlinear we expand them about an approximate solution $\alpha = \alpha_A$ (corresponding to a first guess of the track parameters): $y_l = f_l(\alpha_A) + (\partial f_l/\partial\alpha_i)(\alpha_i - \alpha_{A\,i})$. This linearization permits us to define the χ^2 statistic

$$\begin{aligned}
\chi^2 &= \sum_l (y_l - (\alpha_A) - A_{li}(\alpha_l - \alpha_{A\,l}))^2/\sigma_l^2 \\
&= (\mathbf{y} - \mathbf{f}(\alpha_A) - \mathbf{A}(\alpha - \alpha_A))^T \mathbf{V}_y^{-1}\, (\mathbf{y} - \mathbf{f}(\alpha_A) - \mathbf{A}(\alpha_A - \alpha)) \\
&\equiv (\mathbf{\Delta y} - \mathbf{A}(\alpha - \alpha_A))^T \mathbf{V}_y^{-1}\, (\mathbf{\Delta y} - \mathbf{A}(\alpha - \alpha_A))
\end{aligned}$$

where $\mathbf{\Delta y} = \mathbf{y} - \mathbf{f}(\alpha_A)$, $A_{li} = \partial f_l(\alpha)/\partial \alpha_i|_{\alpha_A}$ is a constant matrix of derivatives, and

$$\mathbf{V}_y^{-1} = \begin{pmatrix} 1/\sigma_1^2 & 0 & \cdots & 0 \\ 0 & 1/\sigma_2^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1/\sigma_n^2 \end{pmatrix}$$

is the inverse of the covariance matrix of the measurements. It is diagonal for independent measurements. Since the χ^2 measures how much the measurements "miss" the function, the solution we want is that which minimizes χ^2, i.e., $\partial \chi^2/\partial \alpha_i = \mathbf{0}$. The solution was found to be $\alpha = \alpha_A + \mathbf{V}_A \mathbf{A}^T \mathbf{V}_y^{-1} \mathbf{\Delta y}$ with covariance matrix $\mathbf{V}_\alpha$, where $\mathbf{V}_\alpha = \mathbf{V}_A = (\mathbf{A}^T \mathbf{V}_y^{-1} \mathbf{A})^{-1}$. In theory one should iterate this procedure to guarantee that the final track is close to the χ^2 minimum.

4. Problems with the Standard Fit

The standard fit simultaneously fits *all* the measurements and so can be regarded as a global method. When the measurement errors σ_l are independent of each other it executes in a time proportional to n, the number of measurements. The situation, however, changes drastically when we introduce multiple scattering because a single scattering event affects the drift distance measurements at all subsequent points. The measurement covariance matrix is now $\mathbf{V}_y = \mathbf{V}_{\text{meas}} + \mathbf{V}_{\text{MS}}$, where $\mathbf{V}_{\text{meas}}$ is the original contribution from independent measurement errors and $\mathbf{V}_{\text{MS}}$ is the correlated component due to the presence of the multiple scattering event and contains off diagonal terms (see CBX 91–74[7] for the calculation of $\mathbf{V}_{\text{MS}}$). The problem now is obtaining $\mathbf{V}_y^{-1}$ when $\mathbf{V}_y$ is no longer diagonal: the inversion of the covariance matrix is prohibitively time consuming when the number of measurements n is large (execution time is $\mathcal{O}(n^3)$). In the CLEO central detector, for instance, a track can have as many as 71 measurements. Inverting a 71×71 matrix at least once per track to account for multiple scattering is enormously expensive in computer time and clearly unacceptable.

There is an extensive literature about the multiple scattering problem in high energy physics. DUET partially accounts for multiple scattering by including a kink in the $r - \phi$ plane at the entrance to the DR drift chamber. The rest of it is included by a method described in CBX 91–74[7] in which multiple scattering is ignored in the fit but its effects on the covariance matrix is accounted for by propagation of errors. The implementation does not work well for soft tracks which bend significantly and in any case is non-optimal because the multiple scattering information is not used in the determination of the track parameters.

The standard fit also does not account for energy loss as the track propagates. Implementing dE/dx corrections cannot be done by linear expansion because soft tracks lose large amounts of energy in detector walls. In any case, the equations become very difficult to write down.

5. Track Fitting with the Kalman Filter

The Kalman filter method, in contrast to the standard algorithm, is an iterative procedure that is, in some sense, intuitively obvious (in fact, parts of the idea have been used for years in track fitting). The procedure iteratively traces the track from its outermost point to the origin, picking up measurements and accounting for multiple scattering and dE/dx losses along the way. Unlike global methods which fit all the measurements to a single set of track parameters, the filter causes the track to "follow the measurements" through the detector.

Let the n hits on the track trajectory be divided among r regions, labelled 1, 2, $\ldots r$ starting from the outermost one, where the regions are separated from one another by material boundaries. A particularly simple case is the one where every hit lies in a separate region because, as we shall see, no matrix inversion need be done. Assume for the time being that the parameters α_1 and their covariance matrix $\mathbf{V}_{\alpha\,1}$ have somehow been determined in region 1 (this avoids the tricky subject of starting the fit, a topic that I defer till later). The track parameters in region 2 are obtained by minimizing a χ^2 which includes contributions from the measurements in region 2 and the projected track from region 1. The projection must include corrections to the track parameters from the expected energy loss in the material ($\alpha_1 \to \alpha_1'$) and to the track covariance matrix to account for multiple scattering and fluctuations in the energy loss ($\mathbf{V}_{\alpha\,1} \to \mathbf{V'}_{\alpha\,1}$). The formulas for these corrections can be found in Appendices 1 and 2.

The change in χ^2 for region 2 can be written as a sum of two terms, the first accounting for the measurements in region 2 and the second allowing for the pull of the track parameters from their previously fitted values (after dE/dx corrections), e.g.:

$$\Delta\chi^2 = (\mathbf{\Delta y}_2 - \mathbf{A}_2(\alpha_2 - \alpha_1'))^T \mathbf{V}^{-1}_{\mathrm{meas}\,2}(\mathbf{\Delta y}_2 - \mathbf{A}_2(\alpha_2 - \alpha_1')) + (\alpha_2 - \alpha_1')^T \mathbf{V}'^{-1}_{\alpha\,1}(\alpha_2 - \alpha_1')$$

where $A_{2\,li} = \partial f_{2\,l}(\alpha)/\partial\alpha_i$ are the derivatives of calculated drift distance in region 2 with respect to the five parameters, α_1' are the track parameters from region 1 corrected for energy loss and $\mathbf{V}'_{\alpha\,1} = \mathbf{V}_{\alpha\,1} + \mathbf{V}_{\alpha\,\mathrm{MS}} + \mathbf{V}_{\alpha\,\mathrm{energy}}$ is the covariance matrix of the track from region 1 corrected by the contributions from multiple scattering and energy fluctuations. As before, $\mathbf{\Delta y}_2 = \mathbf{y}_2 - \mathbf{f}(\alpha_1')$ represents the vector of differences between the measurements and the predictions based on the parameters before fitting. The solution, obtained by setting $\partial\chi^2/\partial\alpha_i$ to zero, is

$$\begin{aligned}
\alpha_2 &= \alpha_1' + \mathbf{V}_{A\,g}\mathbf{A}_2^T\mathbf{V}^{-1}_{\mathrm{meas}\,2}\mathbf{\Delta y}_2, \\
\mathbf{V}_{\alpha\,2} &= \mathbf{V}_{A\,g}, \\
\Delta\chi^2 &= (\mathbf{\Delta y}_2)^T\mathbf{V}^{-1}_{\mathrm{meas}\,2}\mathbf{\Delta y}_2 - (\mathbf{\Delta y}_2)^T\mathbf{V}^{-1}_{\mathrm{meas}\,2}\mathbf{A}_2\mathbf{V}_{A\,g}\mathbf{A}_2^T\mathbf{V}^{-1}_{\mathrm{meas}\,2}\mathbf{\Delta y}_2,
\end{aligned}$$

where

$$\mathbf{V}_{A\,g} = (\mathbf{A}_2^T\mathbf{V}^{-1}_{\mathrm{meas}\,2}\mathbf{A}_2 + \mathbf{V}'^{-1}_{\alpha\,1})^{-1} = \mathbf{V}'_{\alpha\,1}(1 + \mathbf{A}_2^T\mathbf{V}^{-1}_{\mathrm{meas}\,2}\mathbf{A}_2\mathbf{V}'_{\alpha\,1})^{-1}.$$

The second form of $\mathbf{V}_{A\,g}$ is preferred because it avoids the calculation of $\mathbf{V}'^{-1}_{\alpha\,1}$.

This procedure is repeated for the subsequent regions, each iteration improving the track parameters and their covariance matrix, and terminates when the beginning of the track is reached (usually when it breaks through the beam pipe). The net effect of the Kalman filter is to continuously update the track parameters and their covariance matrix.

If we include the measurements a few at a time the fitting process can be speeded up by avoiding the 5×5 matrix inversion in each iteration of the above formulas. This follows from applying the Woodbury matrix inversion theorem[6] to the calculation of $\mathbf{V}_{A\,g}$ in the above equations. The theorem states that if A is an invertible $n \times n$ matrix and U and V are $n \times p$ matrices (with $p < n$), then

$$(A + UV^T)^{-1} = A^{-1} - A^{-1}U(1 + V^TA^{-1}U)^{-1}V^TA^{-1}$$

where $(1 + V^TA^{-1}U)$ is a $p \times p$ matrix. Here $n = 5$ and p is the number of hits being added from a region. The formula for $\mathbf{V}_{A\,g}$ now becomes

$$\mathbf{V}_{A\,g} = \mathbf{V}_{\alpha\,1} - \mathbf{V}_{\alpha\,1}\mathbf{A}_2^T \left(\mathbf{V}_{\mathrm{meas}\,2} + \mathbf{A}_2\mathbf{V}_{\alpha\,1}\mathbf{A}_2^T\right)^{-1} \mathbf{A}_2\mathbf{V}_{\alpha\,1}$$

We must be careful here to include the multiple scattering terms in the equation for $\mathbf{V}_{\mathrm{meas}\,2}$ as worked out in reference 7. Of course, if a large number of hits are included at one time the inversion of the measurement covariance matrix becomes prohibitively expensive in computer time. Appendix 3 shows a faster way of adding the hits that relies on a careful choice of step size in the gas.

For the special case $p = 1$, i.e., hits added one at a time, the entire process becomes very fast since the matrix inversion reduces to a simple arithmetic inverse. The expressions for α_2 and its covariance matrix $\mathbf{V}_{\alpha\,2}$ now become

$$\alpha_2 = \alpha_1' + \mathbf{V}_{A\,g}\mathbf{A}_2^T \frac{\delta y_i}{\sigma_i^2},$$

$$\mathbf{V}_{\alpha\,2} = \mathbf{V}_{A\,g} = \mathbf{V}_{\alpha\,1} - \frac{\mathbf{V}_{\alpha\,1}\mathbf{A}_2^T\mathbf{A}_2\mathbf{V}_{\alpha\,1}}{\sigma_i^2 + \mathbf{A}_2\mathbf{V}_{\alpha\,1}\mathbf{A}_2^T}.$$

These equations assume that the track has been traced as closely to the point of closest approach as possible so that further multiple scattering and energy corrections need not be done.

There are several interesting points about the Kalman filter that should be noted:

1. The Kalman filter method uses *all* the information and cannot, if used correctly, give poorer track parameters by adding more measurements. For example, the contribution from the back half of a curler would be small because multiple scattering errors would reduce the significance of the parameters determined in this region. Including these hits would *not* cause the track parameters to worsen, unlike many other track fitting methods.

2. Because the algorithm traces the track backwards, the parameters on the outer part of the track are much more poorly determined than the ones on the inner part. Although this is good for physics reasons (we want the parameters at the production point), it means that the projection of the track to the outer region is unreliable. To do the

projection properly, the track must be refit by tracing it outward from the production point to the outermost part of the drift chamber. It might be necessary to store these parameters in ROAR, but probably not the covariance matrix. A study should be performed to determine how accurately we need to carry out the projection.

3. The Kalman filter is mathematically equivalent to the global fit when there is no multiple scattering or energy loss. That this is true can be seen from the first form of the solution for α_2 above.

4. Every region does not have to contain measurements. If a particular region lacks measurements, the track is simply projected through the region to the next one.

5. The natural iterativeness of this procedure allows tracks to be fit in pieces. For example, tracks can be found and fit in the drift chambers and then projected to the silicon with the full covariance matrix. The silicon hits can be added to the fit trivially without having to do the entire fit again. Many experiements already exploit this technique.

6. The Kalman filter technique is sensitive, especially at low momentum, to the particle hypothesis used in the fit (e, μ, π, K or p). For best results, we will probably want to fit each fit 5 times and store all the results, although we will need to devise a clever coding mechanism at the ROAR level to avoid storing excessive numbers of bytes (especially for the covariance matrices). An added benefit of separate hypothesis fitting is that we can adjust the cell drift times using the correct time of flight, rather than always using the π hypothesis. It should also be possible to perform some crude particle ID by comparing the pattern of measurement residuals.

7. The Kalman filter is well suited to object oriented programming (OOP) techniques. In OOP language, a region object takes an incoming track object, extrapolates the track across its boundaries, modifies the track parameters by adding its own measurements (if any), and extrapolates the track to the next boundary, where the next region takes over.

6. Starting the Kalman Fit

I have avoided until now the issue of how to start the Kalman filter fit. The problem is that in the outer region there are only a few measurements and the track parameters determined there are very poorly determined, causing the extrapolation into the next region to be unreliable. Moreover, the track parameters in this region become extremely sensitive to the vagaries of these few hits and, even worse, there may not be enough measurements to start the fit, especially if the hits lack z information.

There seems to be little published information on how to solve this problem. However, I have come up with a technique that might be robust enough work in data. The stratagem, which is used in FTMONTE, goes as follows. The track is first fit using all the hits in the outer drift chamber. I then multiply the 5×5 track covariance by a scale factor in the range of 200 to 1000. The track parameters with the new covariance matrix becomes the initial track that I use to start the fit in region 1 (I pretend that the track enters region 1 with these parameters). The scale factor must be small enough so that the track has some initial "stiffness" so that it cannot easily be pulled away from its starting value, but it must be large so that the contribution of the drift chamber hits is not double counted in the final fit.

I have not carried out any systematic studies on how to start the fit in an optimal way. Currently only the diagonal part of the covariance matrix are used and I scale the 5 terms by a different amount in FTMONTE (I use 200, 1000, 1000, 500, 500 for the 5 parameters). It might also be useful to choose the outermost region to include more hits so that the track is reasonably stabilized after the first region. These issues need to be explored more fully.

Appendix 1
Multiple scattering formulas

Multiple scattering causes the particle to scatter in the two planes perpendicular to its path without losing energy. The distribution of each angle is approximately gaussian with a standard deviation given by

$$\sigma_\theta = \frac{0.0141}{p\beta}\sqrt{\frac{L}{X_R}} \equiv \sqrt{HL}$$

where p is the particle momentum in GeV/c, β is its velocity, L is the length of the path and X_R is the radiation length of the material. The track also changes position in the two planes with a standard deviation in each plane given by

$$\sigma_x = \frac{0.0141}{p\beta}L\sqrt{\frac{L}{3X_R}} \equiv L\sqrt{\frac{HL}{3}}$$

The position and angle changes in each plane are correlated with a correlation coefficient $\rho = \sqrt{3}/2 \simeq 0.866$. The scatterings in each plane are independent of one another.

Suppose we want to trace the track to a radius r. In the absence of multiple scattering the track would end up at (x, y, z) (or (r, ϕ, θ) in spherical coordinates) and its direction would be described by the spherical angles $(\phi_t,\ \theta_t)$. Multiple scattering causes the track to scatter in the plane oriented perpendicular to its path. The unit vectors defining the plane are $\hat{\phi}_t = (-\sin\phi_t, \cos\phi_t, 0)$ and $\hat{\theta}_t = (\cos\theta_t\cos\phi_t, \cos\theta\sin\phi_t, -\sin\theta_t)$.

The parameters will be changed by an amount

$$\delta\alpha = \mathbf{M}\begin{pmatrix} \sin\theta_t\delta\phi_t \\ \delta\theta_t \\ \delta x_{\phi_t}/L \\ \delta x_{\theta_t}/L \end{pmatrix}$$

where $M_{i1} = \partial\alpha_i/\sin\theta_t\partial\phi_t$, $M_{i2} = \partial\alpha_i/\partial\theta_t$, $M_{i3} = L\partial\alpha_i/\partial x_{\phi_t}$ and $M_{i4} = L\partial\alpha_i/\partial x_{\theta_t}$, are the derivatives of the track parameters with respect to the scattering quantities and x_{ϕ_t} and x_{θ_t} are distances along the unit vectors $\hat{\phi}_t$ and $\hat{\theta}_t$, respectively. Note that the perpendicular plane is in general not tangent at radius r to the sphere centered at 0 because of the bending of the track.

Using the correlations, the contribution to the track covariance matrix due to multiple scattering is

$$\mathbf{V}_{\alpha\,\mathrm{MS}} \equiv \langle \delta\alpha\delta\alpha^T \rangle = (HL)\mathbf{M} \begin{pmatrix} 1 & 0 & 1/2 & 0 \\ 0 & 1 & 0 & 1/2 \\ 1/2 & 0 & 1/3 & 0 \\ 0 & 1/2 & 0 & 1/3 \end{pmatrix} \mathbf{M}^T.$$

The derivative matrix can be computed using the notation of CBX 92–45 with $r^2 = x^2 + y^2$ and $T = \sqrt{p_\perp^2 - 2a(xp_y - yp_x) + a^2r^2} = p_\perp + aD = p_\perp(1 + 2cD)$. In the interests of fitting the results on a single page we write $\mathbf{M}$ as the direct sum of two 5×2 matrices, e.g., $\mathbf{M} = M_{\mathrm{angle}} \oplus M_{\mathrm{position}}$, where M_{angle} and M_{position} account for the angular and positional multiple scattering effects, respectively. A straightforward but tedious calculation yields

$$M_{\mathrm{angle}} = \begin{pmatrix} 0 & -\frac{1}{2}a\lambda/p_\perp \\ (p/p_\perp)(p_\perp^2 - axp_y + ayp_x)/T^2 & a\lambda(xp_x + yp_y)/T^2 \\ -(p/p_\perp)(xp_x + yp_y)/T & -(a\lambda/2)(r^2 - D^2)/T \\ 0 & -(p/p_\perp)^2 \\ \lambda p(xp_y - yp_x - ar^2)/T^2 & s_\perp + \lambda^2 p_\perp(xp_x + yp_y)/T^2 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & -\frac{1}{2}a\lambda/p_\perp \\ pp_\perp(1 - \rho r\sin\alpha)/T^2 & a\lambda p_\perp r\cos\alpha/T^2 \\ -pr\cos\alpha/T & -(a\lambda/2)(r^2 - D^2)/T \\ 0 & -(p/p_\perp)^2 \\ -\lambda pp_\perp(D + c(D^2 + r^2)/T^2 & s_\perp + \lambda^2 p_\perp^2 r\cos\alpha/T^2 \end{pmatrix}$$

$$M_{\text{position}} = \begin{pmatrix} 0 & 0 \\ \rho p(xp_x + yp_y)/T^2 & -\rho\cos\theta(p_\perp^2 - axp_y + ayp_x)/T^2 \\ (p/p_\perp^2)(p_\perp^2 - axp_y + ayp_x)/T & \rho\cos\theta(xp_x + yp_y)/T \\ 0 & 0 \\ \rho\lambda p(xp_x + yp_y)/T^2 & [-T^2\sin\theta + \lambda\cos\theta(p_\perp^2 - axp_y + ayp_x)]/T^2 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & 0 \\ apr\cos\alpha/T^2 & -ap_z\sin\theta(1-\rho r\sin\alpha)/T^2 \\ p(1-\rho r\sin\alpha)/T & ar\cos\theta\cos\alpha/T \\ 0 & 0 \\ a\lambda pr\cos\alpha/T^2 & \sin\theta[-T^2 + p_z^2(1-\rho r\sin\alpha)]/T^2 \end{pmatrix}$$

where $\sin\theta = 1/\sqrt{1+\lambda^2}$ and $s_\perp$ is calculated from

$$\begin{aligned} T\sin\rho s_\perp &= \rho(xp_x + yp_y), \\ T\cos\rho s_\perp &= p_\perp - \rho(xp_y - yp_x). \end{aligned}$$

α is the angle between the radius vector and the track direction. Recall from CBX 92–45 that

$$\begin{aligned} \sin\alpha &= \frac{xp_y - yp_x}{rp_\perp} = rc - \frac{D}{r}(1+cD), \\ \cos\alpha &= \frac{xp_x + yp_y}{rp_\perp} = \epsilon\sqrt{(1 - D^2/r^2)[(1+cD)^2 - r^2c^2]}, \end{aligned}$$

where $\epsilon = +1(-1)$ refers to the outgoing (incoming track). It is equal to the sign of $xp_x + yp_y$.

Appendix 2

Energy loss formulas

There is a longer discussion of dE/dx corrections in CBX 92–40.[5] The corrected momentum p' is calculated from the initial momentum p by $p' = p + \Delta p$, where Δp is the energy loss correction calculated by integrating the Bethe-Bloch formula

$$\Delta p = \int_0^t \frac{dp}{dx} dx = \int_0^t \frac{1}{\beta}\frac{dE}{dx} dx = \int_0^t \frac{A}{\beta^3}\left[\ln\left(\frac{2m_e c^2 \beta^2 \gamma^2}{I_0}\right) - \beta^2\right] dx$$

where t is the thickness of the material to be traversed, A is a constant that depends on the material, Z is the atomic number, m_e is the electron mass, I_0 is the ionization potential and β and γ are the usual relativistic parameters. We use the approximation $I_0 = Z^{0.9}$eV, and take $Z = 9$ for representative media. A is determined from our knowledge of $(dE/dx)_{\text{min}}$ at minimum ionization, the quantity listed in the Particle Data Book. The minimum occurs at $\beta\gamma = 3.40$ and gives $A = (dE/dx)_{\text{min}}/11.528$.

It is easier to write the differential equation in terms of $\tau = p/m$ and the normalized thickness $x' = x/t$:

$$\frac{d\tau}{dx'} = \frac{\Delta E_{\text{min}}}{m} F(\tau)$$
$$F(\tau) = \frac{(1+\tau^2)^{3/2}}{11.528\tau^3}\left[9.0872 + 2\ln(\tau) - \frac{\tau^2}{1+\tau^2}\right]$$

where $\Delta E_{\text{min}} = (dE/dx)_{\text{min}} t$ is the average energy loss suffered by a minimum ionizing particle passing through a distance t of material (including path length corrections from non-normal tracks). The change in momentum can then be calculated from the integral

$$\Delta p = m\Delta\tau = \Delta E_{\text{min}} \int_0^1 F(\tau) dx'$$

For moderate velocities ($\beta > 0.7$), $F(\tau)$ is approximately constant giving $\Delta p \simeq \Delta E_{\text{min}} F(\tau)$. For low velocities, where $F(\tau)$ is changing rapidly, it is probably necessary to integrate the $d\tau/dx'$ differential equation numerically using a method such as the Runge–Kutte technique, as discussed in CBX 92–40.[5] The integration is aided by transforming to a new variable $\xi = \tau^4$ which removes the $1/\tau^3$ term from the differential equation and makes the numerical solution far more reliable. The new equation using this variable is

$$\begin{aligned}
\frac{d\xi}{dx'} &= \frac{\Delta E_{\text{min}}}{m} F'(\xi) \\
F'(\xi) &= \frac{(1+\sqrt{\xi})^{3/2}}{4 \cdot 11.528} \left[9.0872 + \tfrac{1}{2}\ln(\xi) - \frac{\sqrt{\xi}}{1+\sqrt{\xi}} \right]
\end{aligned}$$

Once the new momentum has been determined, the formulas in Section 1.4 in CBX 92–45 can be used to calculate the new helix parameters.

Appendix 3

Stepping Through Gases

If several hits are to be fit simultaneously in a region, it is important to add the multiple scattering corrections to the measurement covariance matrix $\mathbf{V}_{\rm meas}$ using formulas found in thelast section of CBX 91–74.[7] Since these corrections add off diagonal terms to $\mathbf{V}_{\rm meas}$, computing its inverse is time consuming if a moderately large number of measurements are added at one time. In this section I show that it is possible to ignore these corrections during the fit providing that we choose a step size small enough. After the fit is complete, one should add the multiple scattering contribution due to the material in the region to the new track covariance matrix using the formulas in Appendix 1.

The criterion I use is that the expected deviation of the track over the whole region from multiple scattering must be less than a certain amount, both in the $r-\phi$ and $r-z$ planes. Let Δ_ϕ and Δ_z be the maximum average deviation that will be tolerated in $r-\phi$ and along z, respectively. The RMS deviation in each plane from multiple scattering over a length L in material of radiation length X_R is $(0.0141L/p\beta)\sqrt{L/3X_R}$, where the step ΔL must satisfy

$$\frac{0.0141\Delta L}{p\beta}\sqrt{\frac{\Delta L}{3X_R}} < \Delta_\phi,$$

$$\frac{0.0141\Delta L}{p\beta\sin\theta}\sqrt{\frac{\Delta L}{3X_R}} < \Delta_z,$$

where $1/\sin\theta$ is a geometrical effect accounting for the error in z at fixed radius for steeply falling tracks. Rearranging, we find that ΔL must satisfy both

$$\Delta L < \Delta_\phi^{\rm max} = \left(\frac{3\Delta_\phi^2 p^2\beta^2 X_R}{0.0141^2}\right)^{1/3}$$

$$\Delta L < \Delta_z^{\rm max} = \left(\frac{3\Delta_z^2 p^2\beta^2 X_R\sin^2\theta}{0.0141^2}\right)^{1/3}$$

Note that the step size is much more sensitive to the momentum ($\sim (p\beta)^{2/3}$) than to the radiation length of the gas ($\sim X_R^{1/3}$). The following table shows maximum step lengths for 50–50 Argon-Ethane ($X_R = 166$ m) as a function of momentum (assuming $\sin\theta = 0.8$) with maximum allowed scattering $\Delta_\phi = 50\mu$m and $\Delta_z = 100\mu$m.

Maximum size step size (in cm) in 50–50 Argon–Ethane for π, K, p based on $\Delta_\phi = 50\mu$m ($\Delta_z = 100\mu$m)

p	π $\Delta_\phi^{\max}$ ($\Delta_z^{\max}$)	K $\Delta_\phi^{\max}$ ($\Delta_z^{\max}$)	p $\Delta_\phi^{\max}$ ($\Delta_z^{\max}$)
0.05	1.21 (1.66)	0.54 (0.74)	0.35 (0.48)
0.10	2.77 (3.79)	1.35 (1.85)	0.89 (1.22)
0.25	6.68 (9.14)	4.31 (5.89)	2.96 (4.05)
0.50	11.3 (15.5)	9.25 (12.7)	7.02 (9.61)
1.00	18.3 (25.1)	17.1 (23.5)	14.9 (20.4)
1.50	24.1 (33.0)	23.3 (31.9)	21.6 (29.6)
2.00	29.2 (40.0)	28.7 (39.2)	27.4 (37.5)
3.00	38.3 (52.4)	38.0 (52.0)	37.2 (50.8)
4.00	46.4 (63.5)	46.2 (63.2)	45.6 (62.4)
5.00	53.9 (73.5)	53.7 (73.5)	53.3 (72.9)

References

1. P. Billor, Track Fitting with Multiple Scattering: A New Method, NIM **225**, 352-366 (1984).

2. M. Regler and R. Fruwirth, *Reconstruction of Charged Particle Trajectories*, Plenum Publishing Company (1989).

3. R. Bock, H. Grote, D. Notz and M. Regler, *Data Analysis Techniques for High-Energy Physics Experiments*, Cambridge University Press (1990).

4. P. Avery, Applied Fitting Theory IV: Formulas for Track Fitting, CBX 92–45.

5. P. Avery, The Charged Particle Energy Correction in CLEO, CBX 92–40.

6. P. Avery, Applied Fitting Theory I: General Least Squares Theory, CBX 91–72.

7. P. Avery, Applied Fitting Theory III: Non-Optimal Least Squares Fitting and Multiple Scattering, CBX 91–74.