

A Proposed C++ Toolkit for Kalman Filtering in Larsoft

H. Greenlee

March 25, 2012

Contents

1	Introduction	2
2	C++ Classes Overview	2
3	Linear Algebra	3
4	Surfaces	4
4.1	Surface Base Class	4
4.2	Planar Surfaces	5
4.3	Bounded Surfaces	6
5	Track State	6
5.1	Track State Without Errors	6
5.2	Track State With Errors	7
5.3	Track State with Fit	7
5.4	Track State With Measurements	7
6	Measurements	8
6.1	Measurement Base Class	8
6.2	General Measurement Class	9
6.3	Concrete Measurement Classes	10
6.3.1	Wire-Time Measurement	10
6.3.2	Space Point Measurement	11
7	Measurement Containers	11
8	Propagators	11
9	Interactors	12
10	Persistent Tracks	13

A Kalman Filter Algorithm	13
B Propagation	15
B.1 Coordinate Transformation	16
B.2 Chaining Propagation	16
C Global Fits and Kalman Smoothing	16

1 Introduction

The subject of this document is a proposed set of c++ classes for implementing the Kalman Filter track reconstruction algorithm [1, 2, 3, 4, 5] in larsoft.

The Kalman Filter algorithm has been widely used in high energy physics experiments since the 1980s. The Kalman Filter is useful both as an efficient method of estimating track parameters, and as a road-following type pattern recognition tool. The proposals contained in this document are informed by the Kalman Filter implementation used by the D0 experiment, called TRF [6], as well as by the GENFIT framework [7], of which a copy exists in larsoft, without being a copy of either one. A brief description of the Kalman Filter algorithm is given in Appendices A–C.

2 C++ Classes Overview

There are various types of detector-specific classes that are needed or useful for implementing a Kalman Filter algorithm, including the following.

- Linear algebra
- Surfaces.
- Track state (i.e. information about a track on a single surface).
- Measurements.
- Measurement containers.
- Propagators.
- Interactors.
- Persistent tracks.

In the following sections, we will consider various aspects of the above types of classes, including use cases, interfaces, attributes, methods, and persistence.

3 Linear Algebra

The Kalman filter algorithm makes heavy use of linear algebra objects (vectors, matrices, symmetric matrices). The choice of linear algebra package will affect the performance of the Kalman filter, in terms of both time and memory use. Since it will be hard to change linear algebra package after the Kalman filter is implemented, it is important to make a good choice at the outset.

We will want a linear algebra package that is optimized for handling large numbers of small dimensional matrices (up to dimension five, probably). One area where we may expect differences among linear algebra packages is memory management. It is well known that in these kinds of cases (large numbers of small objects), that performance can be very adversely affected by excessive numbers of memory allocations and deallocations. Ideally, we would like a linear algebra package that avoids memory allocations entirely.

Three linear algebra packages that are readily available, and which are already included among standard larsoft dependencies (included in art externals) are ROOT, CLHEP, and UBLAS (UBLAS is part of the BOOST software library). Each of these packages is aware of the memory allocation problem, and has solved it in a different way.

ROOT would be a natural choice for linear algebra package, since ROOT is already familiar and heavily used throughout larsoft and high energy physics generally. Furthermore ROOT objects have the advantage being easily persistable. ROOT solves the memory allocation problem by preallocating memory on the stack for objects up to dimension five (that is, five doubles for vectors, and 25 doubles for matrices, including symmetric matrices). This approach saves memory allocations at the expense of wasting memory. I think larsoft would unwise to emulate this unconcern about memory usage, since larsoft is already grappling with memory bloat.

CLHEP has the same philosophy as ROOT. CLHEP stores data in STL vectors, but it provides a custom allocator that preallocates memory on the stack (25 doubles for matrices, the same as ROOT). Thus, CLHEP and ROOT have the same non-ideal time vs. memory trade off.

I had never heard of UBLAS before, but it seems to be the best choice among the three linear algebra packages listed above. UBLAS gives you the choice of different storage models, one of which is to specify the exact sizes of linear algebra objects at compile time via template parameters, which memory is preallocated on the stack. I believe that this is the most ideal solution. Furthermore, UBLAS avoids virtual methods entirely (unlike either ROOT or CLHEP), and thus avoids the time and memory overheads associated with virtual methods. My overall impression about UBLAS and BOOST is that they are much more sophisticated, in a computer science sense, than either ROOT or CLHEP.

For the above reasons, I propose to use UBLAS linear algebra for internal use within the Kalman filter, and to use ROOT for persistent linear algebra objects.

4 Surfaces

Surface classes must be able to be used in the following ways.

- As a destination for track propagation.
- Each kind of surface must have a defined set of track parameters.
- Any kind of surface that can hold a measurement must have a defined set of measurement coordinates.

Regarding the last two bullets, it will be convenient to have a local three-dimensional coordinate system defined with respect to the surface, such that track parameters and measurement coordinates are meaningful in the local coordinate system.

Despite the suggestive name “surface,” a surface in the sense of any of the above bullets need not be a two-dimensional geometric surface, although that is the most common use case. Equally valid are lines and curves (propagate to line) and points (propagate to vertex, propagate to space point). However, regardless of the dimensionality of the surface, we will need to have a three-dimensional local coordinate system associated with the surface.

Classes described in this section are summarized in Fig. 1.

4.1 Surface Base Class

The surface base class should not have any data, and should have the following methods.

- Pure virtual methods for converting 3-vectors between local and global coordinate systems.
- Pure virtual method to convert track parameters (passed as linear algebra vector) to space point and/or momentum vector (without error).
- Pure virtual methods for testing equality and near equality within some tolerance. The near equality method will be used to determine whether a propagation is necessary to reach a specified destination surface.

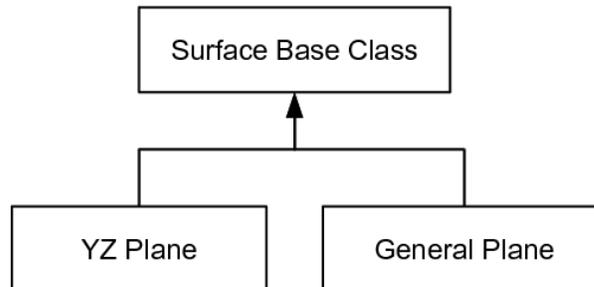


Figure 1: Inheritance diagram for surface classes.

- Pure virtual method for testing whether surfaces are parallel, within some tolerance (assuming that makes sense for a particular kind of surface).
- Pure virtual method for returning perpendicular separation between two parallel surfaces.
- Virtual method(s) to test whether track is in-bounds (with or without error).

It might seem odd that we don't want errors for the space point and momentum vector methods. The reason for this is that we will stipulate that any time we need to do calculations involving errors, we will use track objects. These methods are rather intended for things like event displays where precise errors are not needed, and adding errors will complicate things too much.

4.2 Planar Surfaces

In larsoft we will initially (and maybe only ever) be concerned with planar surfaces. Note that plane surfaces are good enough to implement the propagation-to-line and propagation-to-point use cases, simply by making a destination surface that passes through the destination point (or line) with some convenient orientation. Therefore, we will assume that there is no need to specify separate surface classes for these use cases.

All of the use cases described at the start of this section can be achieved by defining a general Cartesian coordinate transformation from the global (x, y, z) coordinate system to a local (u, v, w) coordinate system (not to be confused with the three wire readout views). Such a general transformation will include both a translation and a rotation. The plane itself is defined as $w = 0$ in the local coordinate system. The u and v coordinates are used for measurement coordinates and track parameters.

The attributes of the plane surface class are:

- Displacement vector specified in global coordinated (x, y, z) .
- Rotation matrix. This could be specified using Euler angles, or redundantly by storing the entire rotation matrix.

The track parameters on the plane surface are as follows.

1. u .
2. v .
3. du/dw .
4. dv/dw .
5. $1/p$ (q/p for magnetized detector).
6. Boolean direction parameter specifying whether dw/ds is positive or negative.

In the case of a measurement surface defined by a readout wire and the drift velocity, we take the u -axis parallel to the drift velocity, and the v -axis parallel to the wire.

It may be convenient to define a separate class for the special case of a plane surface parallel to the x -axis (this is the class “YZ Plane” in Fig. 1), in addition to a general plane surface. Such a special plane surface will require only a single rotation angle, thus saving memory and execution time compared to the general plane, in addition to having a simpler implementation. This type of plane surface is appropriate in the case of a nonmagnetic liquid argon TPC where the drift velocity is parallel to the x -axis.

4.3 Bounded Surfaces

Kalman Filter packages generally include classes for bounded surfaces. Although liquid argon TPCs lack internal structure, it could be useful to have bounded plane surfaces to implement the concept of TPC or fiducial volume boundaries in the local coordinate system of the plane. Alternatively, boundary calculations could be done exclusively in the global coordinate system.

Bounded plane surface classes would inherit from the unbounded plane class, add their own attributes for specifying the boundary, and override the bounds checking methods.

5 Track State

Track state classes encapsulate information about a track on a single surface. The track state class does not need to be polymorphic in the same sense that one would expect to have different implementations of the same interface. Rather, I envision a cascading set of derived classes where each derived class adds attributes and methods. The class hierarchy described in this section is summarized in Fig. 2.

5.1 Track State Without Errors

This is the most basic track state classes. It has two attributes.

- Surface (pointer to Surface base class).
- Track state vector.
- Track direction (boolean). Specifies whether momentum vector is pointing in forward or reverse direction compared to surface.

This class includes methods for calling the underlying virtual methods of its surface.

- Return position vector (without errors).
- Return momentum vector (without errors).
- Return in-bounds status (without errors).

5.2 Track State With Errors

This class inherits from the Track State Without Errors class and adds the following attribute.

- Track parameter covariance matrix.

And it adds the following method.

- Return in-bounds status (with errors).

5.3 Track State with Fit

This class inherits from Track State With Errors class and adds additional information obtained from the Kalman fit.

- Cumulative fit chisquare.
- Fit status (i.e. whether fit is optimal or not).
- Propagation distance.

5.4 Track State With Measurements

This class inherits from the Track State With Fit class and adds information about which measurements contributed to this track.

- A collection of measurements (pointers to measurement base class).

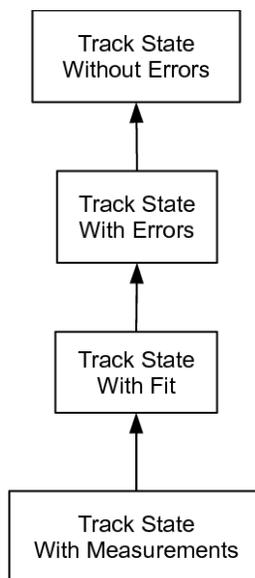


Figure 2: Inheritance diagram for track state classes.

One of the measurements is designated as the latest or current measurement (i.e. the last measurement added to the collection).

The basic Kalman Filter algorithm is carried out with this kind of track using the following steps.

1. Start with a track containing a set of successfully fit measurements.
2. Identify a candidate measurement for addition to the track.
3. Make a trial track by copying Track State With Errors part of the track.
4. Propagate the trial track to the candidate measurement surface.
5. Call the prediction method of the candidate measurement.
6. Calculate the incremental chisquare.
7. If the incremental chisquare fails the incremental chisquare cut, discard this measurement and look for a new candidate measurement (go to step 2).
8. If the predicted incremental chisquare passes the incremental chisquare cut, then do the following:
 - (a) Add the candidate measurement to the collection.
 - (b) Add the incremental chisquare to the cumulative chisquare.
 - (c) Update the track state vector and covariance matrix using the Kalman filtering formula (Eq. 7 and one of Eqs. 9–11).

To carry out these kinds of steps, there will need to be the following kinds of methods.

- Add a measurement to the collection.
- Update the track parameter state vector and covariance matrix using the Kalman updating formula.

6 Measurements

Measurement classes encapsulate data and methods related to measurements, predictions, and residuals (all specified in measurement coordinates). Classes described in this section are summarized in Fig. 3.

6.1 Measurement Base Class

The measurement base class should be totally independent of any details of specific kinds of measurements, including the dimensionality of the measurement.

The measurement base class should have the following attributes:

- Measurement surface.

- Art pointer to associated `RecoBase` object (art associations could possibly be used instead).

The measurement surface attribute should be held via pointer to the surface base class.

The measurement base class should have the pure virtual methods.

- A pure virtual method for calculating a prediction. This method should accept a track state as argument. The purpose of this method is to calculate the measurement coordinates corresponding to the track state. These predicted measurement coordinates should be cached in the class that implements this method. After calling this method, the incremental chisquare should be available also.
- A pure virtual method to access the cached incremental chisquare calculated by the prediction method.
- A pure virtual method for updating a track state using the Kalman updating formula. This method can assume that the prediction method has been called previously, and can access attributes cached by the prediction method.

The prediction and update methods should ensure that the track being predicted or updated has been propagated to the measurement surface.

6.2 General Measurement Class

The general measurement class derives from the measurement base class. It can own and use measurements and predictions in measurement coordinates as opaque linear algebra objects. That is, the general measurement knows the dimensionality of the measurement, but nothing else about the measurement. The general measurement base class should have the following attributes:

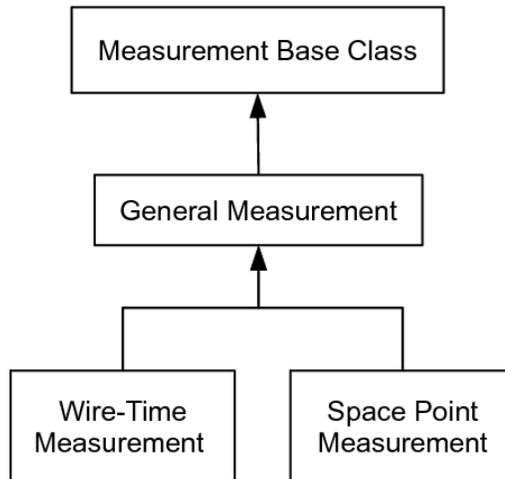


Figure 3: Inheritance diagram for measurement classes.

- Measurement vector and covariance matrix.
- Prediction vector and covariance matrix.
- Residual vector and covariance matrix.
- Inverses of residual and measurement covariance matrix.
- H-matrix.
- Predicted incremental chisquare (using Eq. 12).

The general measurement class should have the following methods.

- A pure virtual method for calculating the prediction vector, prediction error matrix, and H-matrix.
- A partial implementation of the prediction method inherited from the measurement base class. Since the general measurement class does not know what the track parameters or measurement coordinates mean, it can not fully implement the prediction method. The calculation of the prediction vector, prediction error matrix, and H-matrix must be left to a further derived class via the previous method. The general measurement class can implement some parts of the prediction method, including the calculation of the residual, the residual error matrix, and the incremental chisquare.
- An accessor for the incremental chisquare (virtual method inherited from base class).
- The general measurement class should fully implement the Kalman updating formula, as this does not require knowledge of the meaning of measurement or track parameters.

6.3 Concrete Measurement Classes

Concrete measurement classes derive from the general measurement class. Concrete measurements do not have any attributes of their own.

Concrete measurement classes should override the prediction virtual method that calculates the prediction vector, prediction error matrix, and H-matrix.

6.3.1 Wire-Time Measurement

The wire-time measurement is a one-dimensional measurement based on `RecoBase/Hit`. Namely, it consists of a drift time (the measurement coordinate) measured on a particular wire. The measurement plane is defined by the wire and the drift velocity.

6.3.2 Space Point Measurement

The space point measurement is a two-dimensional measurement consisting of the coordinates of a point on the measurement surface. The corresponding `RecoBase` object is `RecoBase/SpacePoint`. One feature of this type of measurement is that there is no single uniquely defined measurement plane. In principle, any plane passing through the space point will do. There are a couple of ways this class could be implemented.

- The measurement surface could be chosen once and for all at the beginning of the event and never changed (e.g. a z -plane).
- The prediction method could update the measurement surface in some way, for example, by setting it to the plane perpendicular to the track. In this case, the prediction method would have to update the measurement vector and covariance matrix also.
- A third possibility is to define pointlike surface (destination of propagate-to-point) and use that as the measurement surface. We said in Sec. 4 that this wouldn't be necessary, but we mention it as a possibility.

7 Measurement Containers

A measurement container has ownership of all of the measurements of a particular kind in the event and makes them available to other objects in the form of abstract pointers to the measurement base class (measurements are never copied once the measurement container is filled). Conceptually, a measurement container is nothing more than a collection of pointers. The reason that I mention measurement containers as a separate category is because there is plenty of room for inventing clever containers with the goal of efficiently finding candidate measurements that are “near” the endpoint of the currently growing track (to help with step two of the track finding algorithm described in Sec. 5.4). I do not have any particular clever ideas to put forward in this proposal, though.

8 Propagators

Propagators accept a track state (tracks parameters and error matrix on specified surface), and calculate a new track state on a destination surface. The propagation distance to reach the destination will also be calculated. Refer to Appendix B for a general discussion of propagation.

We will definitely want to have an abstract propagator base class that has basically a single pure virtual propagation method. Then we will want one or more concrete propagator classes that derive from the propagator base class and implement the propagation method (Fig. 4).

There are many possible use cases for propagators. The various use cases may be implemented by different classes, or the same class with different configuration parameters. Some possible use cases are.

- Propagate without error.
- Propagate with error, but without noise
- Propagate with error and noise.
- Propagate with or without magnetic field (uniform field or map).
- Propagate with or without average energy loss.
- Propagate with or without energy loss fluctuations.
- Propagate with or without multiple scattering.
- Propagate with or without short distance approximations.

9 Interactors

Interactors accept a track state (track parameters on specified surface), and calculate the noise matrix (see Appendix B) resulting from propagation over a specified distance through a specified material. In the case of larsoft, the material will usually be liquid argon, but could potentially be other kinds of deal material like walls of containment vessels. Interactors would be called from a propagator when propagation with noise was needed.

We should have an abstract interactor base class with a pure virtual interaction method. Various concrete derived classes will implement the interaction method (see Fig. 5). Here are some use cases, which probably will want different derived classes.

- Gaussian thin or thick multiple scatterer.
- Gaussian energy loss fluctuator.

As previously noted, the interaction associated with average energy loss is not noise, and would have to be handled within the propagator proper.

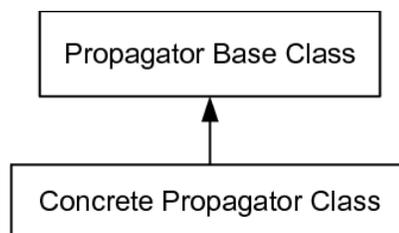


Figure 4: Inheritance diagram for propagator classes.

10 Persistent Tracks

I do not think we will want to save most of the classes described above as data products directly. In particular, I do not think we will ever save measurement objects. Rather, we will associate persistent track objects with the corresponding `RecoBase` objects (`Hit` or `SpacePoint`). Therefore, a persistent track object would have the following kinds of attributes.

- Track state information (Track State With Fit class) on one or more surfaces, including at least the endpoints.
- Art pointers or associations to `RecoBase` objects that are used as measurements.

A Kalman Filter Algorithm

Let \mathbf{x} stand for the five-dimensional state vector of a track on a given surface S . The state vector can be regarded as being a function of the surface,

$$\mathbf{x} = \mathbf{x}(S). \tag{1}$$

If we know the state vector on some surface S , then it is possible to calculate the state vector on any other surface S' via propagation. The basic Kalman Filter algorithm factorizes almost completely with propagation algorithms. The latter are detector-specific and are often quite complicated, whereas the Kalman Filter algorithm per se is simple and detector-independent. In this appendix, we will limit ourselves to describing the Kalman Filter algorithm on a single surface. In Appendix B we examine some general aspects of track propagation.

We denote by the symbol \mathbf{C} the covariance matrix of the estimated state vector \mathbf{x} with respect to the true state vector $\hat{\mathbf{x}}$.

$$\mathbf{C} = \langle (\mathbf{x} - \hat{\mathbf{x}})(\mathbf{x} - \hat{\mathbf{x}})^T \rangle. \tag{2}$$

Let \mathbf{m} be a measurement vector on a surface S . The dimensionality of the measurement vector can vary depending on the type of measurement. We assume that there is a function $\mathbf{h}(\mathbf{x})$ that maps track parameters to predicted measurement coordinates for any measurement surface.

$$\mathbf{m}_{\text{pred}} = \mathbf{h}(\mathbf{x}). \tag{3}$$

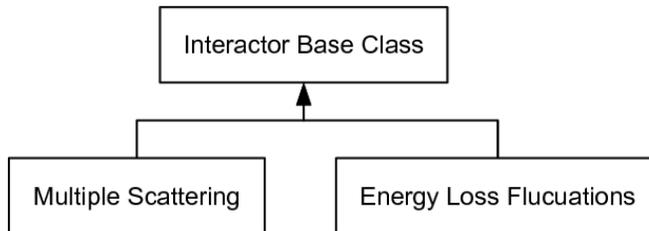


Figure 5: Inheritance diagram for interactor classes.

We furthermore assume that the prediction function can be linearized locally around some state vector \mathbf{x}_0 .

$$\mathbf{h}(\mathbf{x}) = \mathbf{h}(\mathbf{x}_0) + \mathbf{H}(\mathbf{x} - \mathbf{x}_0), \quad (4)$$

where \mathbf{H} is the matrix of partial derivatives of the components of the prediction with respect to the track parameters.

$$\mathbf{H}_{ij} = \frac{\partial h_i(\mathbf{x})}{\partial x_j}. \quad (5)$$

Define the residual vector \mathbf{r} as the difference between the measurement and the prediction on the same surface.

$$\mathbf{r} = \mathbf{m} - \mathbf{h}(\mathbf{x}). \quad (6)$$

Finally, define various covariance matrices in measurement coordinates: \mathbf{V} the intrinsic measurement error, \mathbf{T} the prediction error, and $\mathbf{R} = \mathbf{T} + \mathbf{V}$ the residual error.

The so-called filtered (i.e. updated) track parameters \mathbf{x}' are obtained by simultaneously minimizing the chisquare of the filtered track parameters with respect to the original estimated track parameters \mathbf{x} , and the chisquare of the filtered prediction vector $\mathbf{h}(\mathbf{x}')$ with respect to the measurement vector \mathbf{m} . The filtered track parameters are given by

$$\mathbf{x}' = \mathbf{x} + \mathbf{K}\mathbf{r}, \quad (7)$$

where

$$\mathbf{K} = \mathbf{C}\mathbf{H}^T\mathbf{R}^{-1} \quad (8)$$

is called the Kalman gain matrix.

The filtered track parameter covariance matrix is

$$\mathbf{C}' = (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{C}, \quad (9)$$

$$= (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{C}(\mathbf{I} - \mathbf{K}\mathbf{H})^T + \mathbf{K}\mathbf{V}\mathbf{K}^T. \quad (10)$$

Alternatively, the filtered covariance matrix can be expressed in terms of inverse covariance matrices as

$$\mathbf{C}'^{-1} = \mathbf{C}^{-1} + \mathbf{H}^T\mathbf{V}^{-1}\mathbf{H}. \quad (11)$$

In view of Eq. 11, it can be an interesting design choice to store the inverse of the track covariance matrix rather than the track covariance matrix as part of track state information. Doing this solves the tricky problem of setting the track errors to be initially “infinite.” On the other hand, it (slightly) complicates the calculation of propagation noise.

The incremental chisquare is

$$\chi^2 = \mathbf{r}^T\mathbf{R}^{-1}\mathbf{r}. \quad (12)$$

Equation 12 is invariant whether calculated in terms of the original or filtered residual. Therefore, it is easy to implement an incremental chisquare cut to select candidate measurements for inclusion in the track. Furthermore, hit selection can take place before updating the track parameters and error matrix.

B Propagation

The process of evolving the track parameter state vector \mathbf{x} and covariance matrix \mathbf{C} from one surface S to another surface S' is called propagation. In general, the propagation function is complicated and nonlinear. However, we assume that the propagation function can be linearized locally around some reference trajectory (\mathbf{x}_0, S) to (\mathbf{x}'_0, S') .

$$\mathbf{x}' = \mathbf{x}'_0 + \mathbf{F}(\mathbf{x} - \mathbf{x}_0). \quad (13)$$

\mathbf{F} is the forward propagation matrix, which is the matrix of partial derivatives of the propagated track parameters \mathbf{x}' with respect to the original track parameters \mathbf{x} .

$$F_{ij} = \frac{\partial x'_i}{\partial x_j}. \quad (14)$$

For the reverse propagation (from S' to S) we write

$$\mathbf{x} = \mathbf{x}_0 + \mathbf{B}(\mathbf{x}' - \mathbf{x}'_0), \quad (15)$$

where \mathbf{B} is the backward propagation matrix,

$$B_{ij} = \frac{\partial x_i}{\partial x'_j}, \quad (16)$$

and obviously $\mathbf{B} = \mathbf{F}^{-1}$.

The track parameter covariance matrix is modified by propagation as follows, for forward and backward propagation.

$$\mathbf{C}' = \mathbf{F}\mathbf{C}\mathbf{F}^T + \mathbf{Q}_F. \quad (17)$$

$$\mathbf{C} = \mathbf{B}\mathbf{C}'\mathbf{B}^T + \mathbf{Q}_B. \quad (18)$$

The \mathbf{Q} terms represent the irreversible noise contribution due to stochastic processes, such as multiple scattering and energy loss fluctuations (the effect of average energy loss is not irreversible, and is included in the propagation matrix and reference trajectory). In the absence of propagation noise, the propagation is fully reversible, with respect to both track parameters and error matrix. If there is propagation noise, the propagation is not reversible. However, the forward and backward noise are related as

$$\mathbf{Q}_B = \mathbf{B}\mathbf{Q}_F\mathbf{B}^T, \quad (19)$$

$$\mathbf{Q}_F = \mathbf{F}\mathbf{Q}_B\mathbf{F}^T. \quad (20)$$

Note that the noise contribution to the error matrix is positive regardless of the direction of propagation.

B.1 Coordinate Transformation

A special kind of propagation occurs when the track is already located on the destination surface, even though the destination surface may not be the same as the starting surface. In other words, one is simply transforming the track state from one surface to a different surface. A coordinate transformation of this type is perfectly well described using the formalism described in this appendix. In particular, the calculation of the propagation matrix and the transformation of the error matrix are according to Eqs 14 and 17, with zero noise.

B.2 Chaining Propagation

One frequently wants to divide a complex propagation into several shorter or simpler steps. Assume that we have n forward propagations, each with its own propagation matrix F_k ($k = 1, \dots, n$), and propagation noise matrix Q_{Fk} . The total propagation matrix F is

$$F = \prod_{k=1}^n F_k, \quad (21)$$

To find the total noise, calculate the cumulative noise matrix Q_{Fk} after k steps using

$$Q_{Fk} = F_k Q_{Fk-1} F_k^T + Q_{Fk}, \quad (22)$$

and the total noise matrix is $Q_F = Q_{Fn}$.

Note that neither the total propagation matrix F nor the total noise matrix Q_F depend on the initial error matrix C . Therefore, splitting a propagation into steps does not muddy the distinction between initial error and noise components in the final track parameter error matrix C' .

C Global Fits and Kalman Smoothing

In a typical track-fitting scenario, one has n measurements \mathbf{m}_k ($k = 1, \dots, n$), on surfaces S_k . Starting from an initial estimated track with “infinite errors,” one propagates to each measurement surface, starting with S_1 and ending with S_n . The fitted track parameters are optimal only at the final surface S_n , since that is the only surface that makes use of all available information. One can also do the Kalman fit in the reverse direction to obtain optimal track parameters at surface S_1 .

The accuracy of the fit can sometimes be improved by iterating the Kalman fit, in alternating directions. Iterating will improve the accuracy of the linearized prediction function (Eq. 4), by providing better estimates of the initial track parameters.

It may or may not be good enough to have optimal estimates of the track parameters at either or both endpoints. Sometimes it is desirable to have optimal estimates of track parameters at interior surfaces. In such cases, optimal estimates can be obtained by combining the predicted track parameters at surface S_k from the forward fit (\mathbf{x}_{Fk}) with the filtered track parameters from the backward fit (\mathbf{x}'_{Bk}), or alternatively, forward-filtered and backward-predicted track parameters can be combined. In

this way, one can obtain estimates of track parameters at interior surfaces that make use of all available information, and are therefore optimal. This type of algorithm is called Kalman smoothing.

References

- [1] R.E. Kalman, J. Bas. Eng. **82D**, 35 (1960); R.E. Kalman and R.S. Bucy, J. Bas. Eng. **83D**, 95 (1961);
- [2] P. Billoir, Nucl. Instr. and Meth. **A225**, 352 (1984).
- [3] R. Frühwirth (DELPHI exp.), Nucl. Instr. and Meth. **A262**, 444 (1987).
- [4] P. Billoir and S. Qian (ZEUS exp.), Nucl. Instr. and Meth. **A294**, 219 (1990).
- [5] E.J. Wolin and L.L. Ho, Nucl. Instr. and Meth. **A329**, 493 (1993).
- [6] From my personal experience. There is no public document, however, the D0 code repository is browsable via url <http://cdcvs.fnal.gov/cgi-bin/public-cvs/cvsweb-public.cgi/?cvsroot=d0cvs>, specifically, packages beginning with “trf.”
- [7] C. Höppner, S. Neubert, B. Ketzer, S. Paul, Nucl.Instrum.Meth. A620 (2010) 518-525 (arXiv:0911.1008 [hep-ex]).