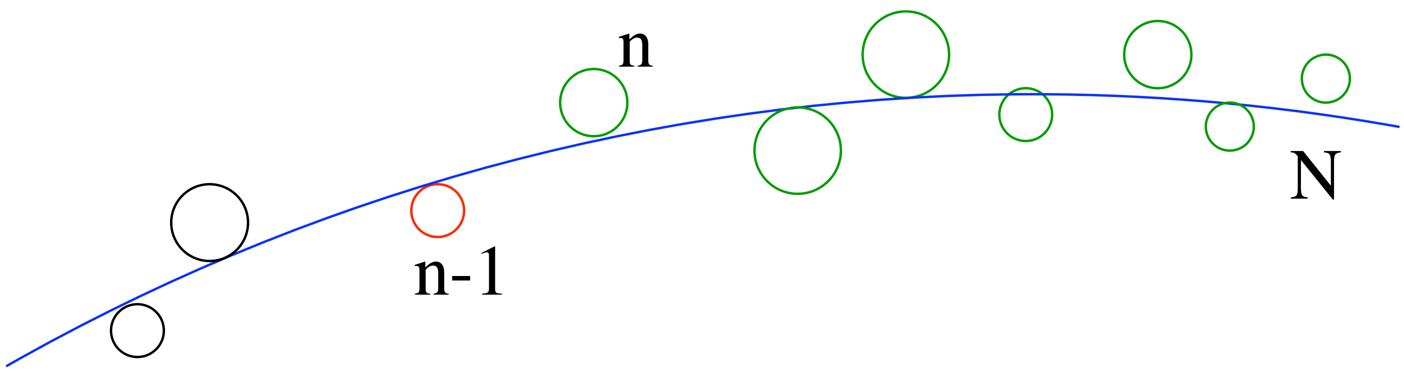
Mu2e-doc-557-v2

Kalman Filters:
As Told by a Physicist to Physicists



Rob Kutschke, Fermilab

Food For Thought Seminar

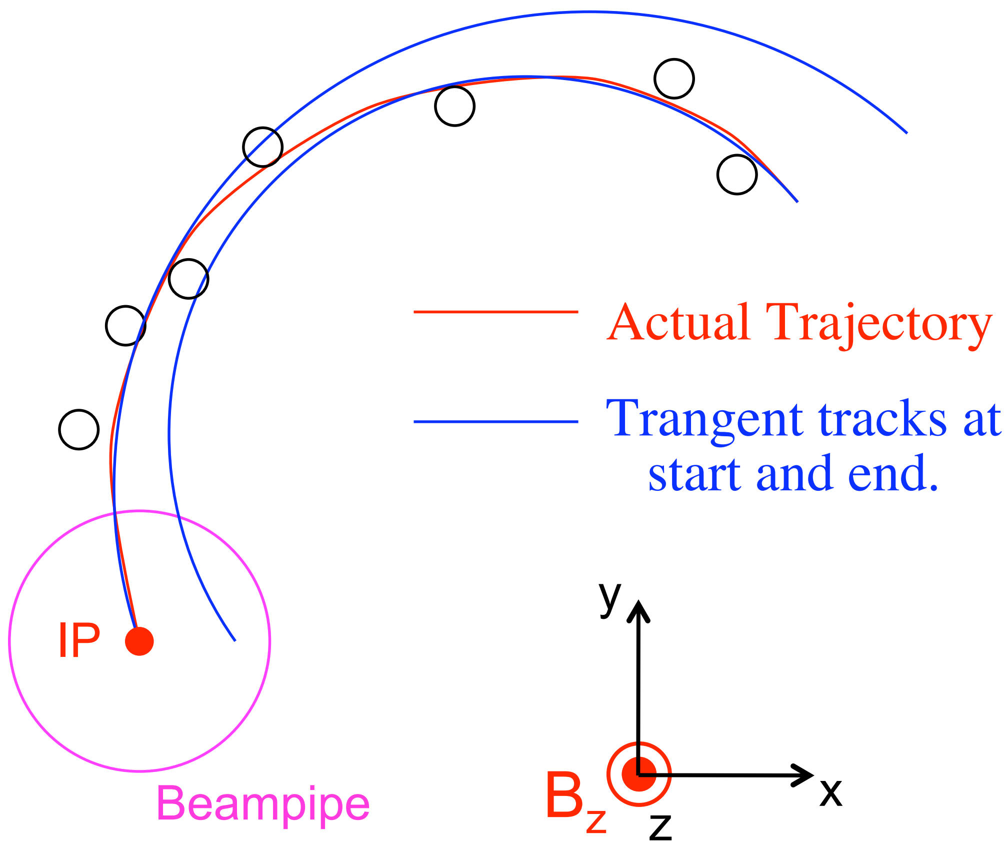
June 9, 2009

Outline

- Introduction
- History
- Helical parameters.
- A series of illustrative exercises.
- A few pages of math.
- Gory details: a straight line in 2D.

Forget "The" Track Parameters

- Why?
  - Scattering and Eloss.
  - Non-uniform fields.
  - depends on mass hypothesis: (e,μ,K,π,p)
- Helix: a locally valid approximation.
- Need a description:
  - Valid everywhere.
  - Smooth.



Some Caveats

- I will discuss use as a final fitter.
  - Someone else did the pattern recognition.
- First few concrete examples will use helical tracks in a drift chamber in a pure solenoidal field.
  - All these can easily be generalized.

Treatment of MS

- Invert $N_{hit} \times N_{hit}$ weight matrix.
  - Slow $O(N^3)$.
  - Only gives parameters valid at one point on the track!
  - Hard to follow non-uniform magnetic fields.
- Helix fit without MS + ad hoc MS correction.
  - CLEO – 1980’s
  - Need to “detune” the measurement errors!
  - Still get biased results.
  - Curlers give huge biases.
- Insert explicit constrained kinks (+energy loss).
  - ARGUS, 1980’s.
  - Works but not optimal.
- Kalman filter.

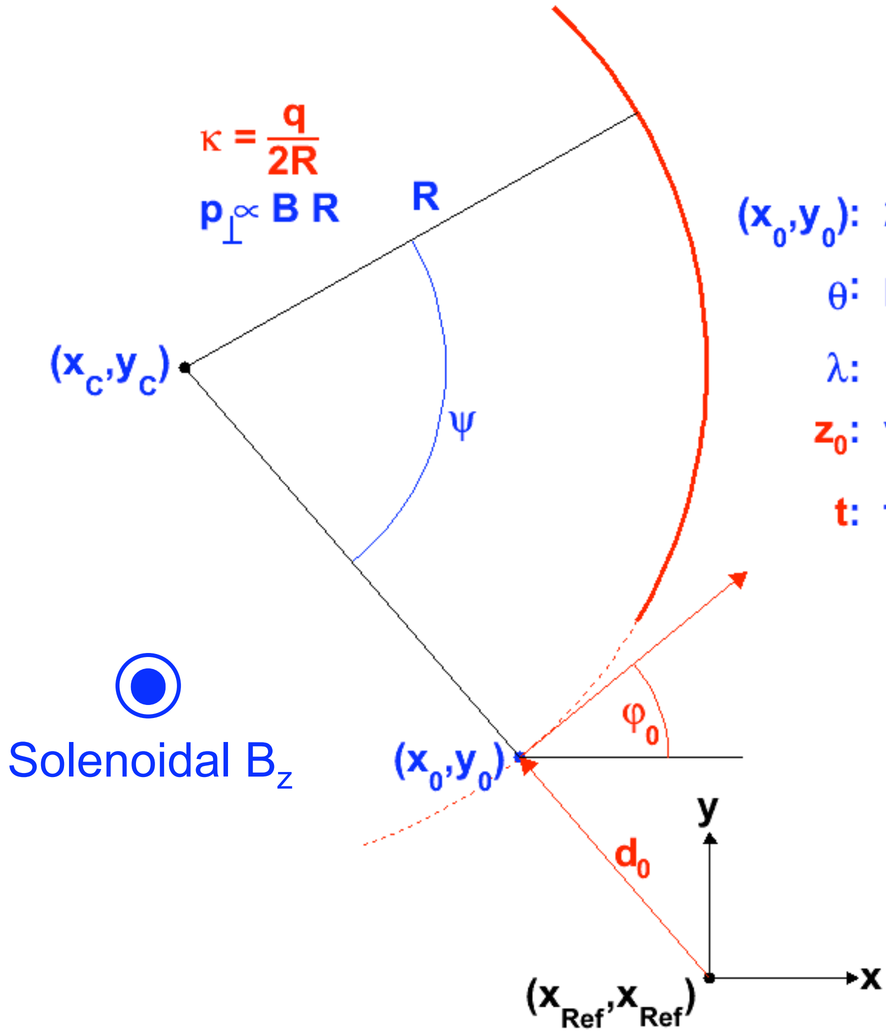
Kalman Filter

- "Filters" the MS "noise" from a measured trajectory.
  - MS = multiple scattering
- If measurement errors and scattering processes are gaussian, KF is an optimal estimator of trajectory:
  - Unbiased central value.
  - Errors correct. Confidence level correct.
  - Minimum error of all unbiased estimators.
- Computes in order (N_{hit}) .
  - No matrices of order $N_{hit} \times N_{hit}$.
- Can follow an arbitrary magnetic field.
- For uniform field: equivalent to inverting an NxN weight matrix that includes all correlations due to scattering.

A Crude History

- ≈1960 Kalman spacecraft tracking.
- ≈1970 Noticed by a Finn: “impractical” for HEP
- Billoir: NIM 225 (1984) 352. “Billoir Fitter”
  - Rediscovered independently applied to HEP tracking.
  - Very intuitive presentation.
  - Not all problems solved.
- Late 1980’s
  - Fruhwirth recognized Billoir fitter as a Kalman filter.
  - Applied the extensive literature
- Used in digital signal processing all along.

Helix Parameters: $\eta = \{\kappa, \varphi_0, d_0, t, z_0\}$



(x_0,y_0): 2D PCA (x_{Ref},y_{Ref})

θ: polar angle at (x_0,y_0)

λ: lattitude 90-θ

z_0: value of z at (x_0,y_0)

t: $\tan\lambda = \cot\theta$

5 parameters + turning angle define a point.

Conventions:
- sign of d_0
- factor of 2 in κ

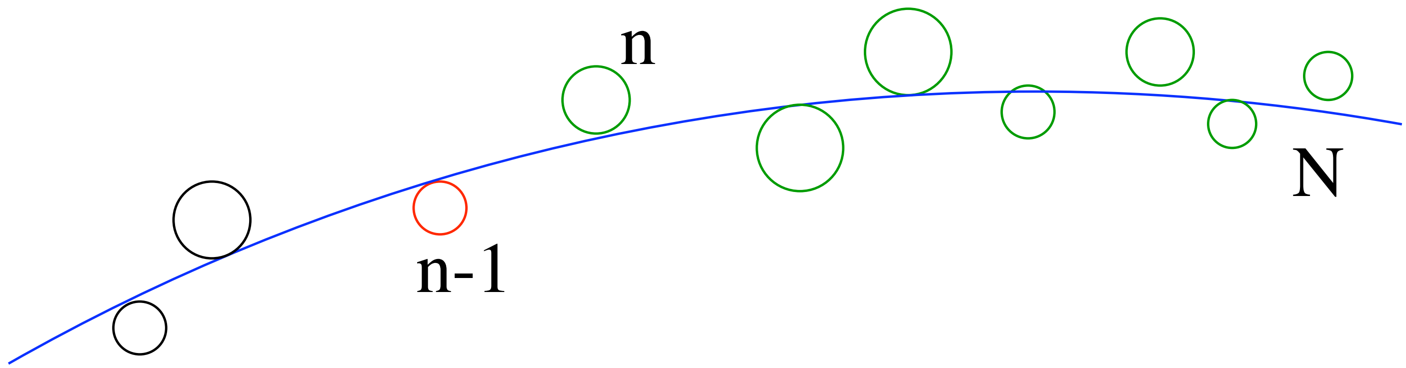
Why These Parameters?

- Very stiff tracks are well behaved:
  - Curvature goes to 0, not infinity.
  - Smooth behavior if field is not strong enough to determine the charge!
- Covariance Matrix is closer to diagonal than it is for other parameterizations.

More Notation

$$
\begin{array}{lcl}
\eta & = & \text{track parameters } (\kappa, \phi_0, d_0, \cot\theta, z_0). \\
V & = & 5 \times 5 \text{ covariance matrix of } \eta. \\
d_m & = & \text{the quantity that is measured by some} \\
 & & \text{tracking device.} \\
\sigma_m & = & \text{the error on } d. \\
\Delta\eta & = & \text{the change to } \eta \text{ made by adding a} \\
 & & \text{hit to the track.}
\end{array}
$$

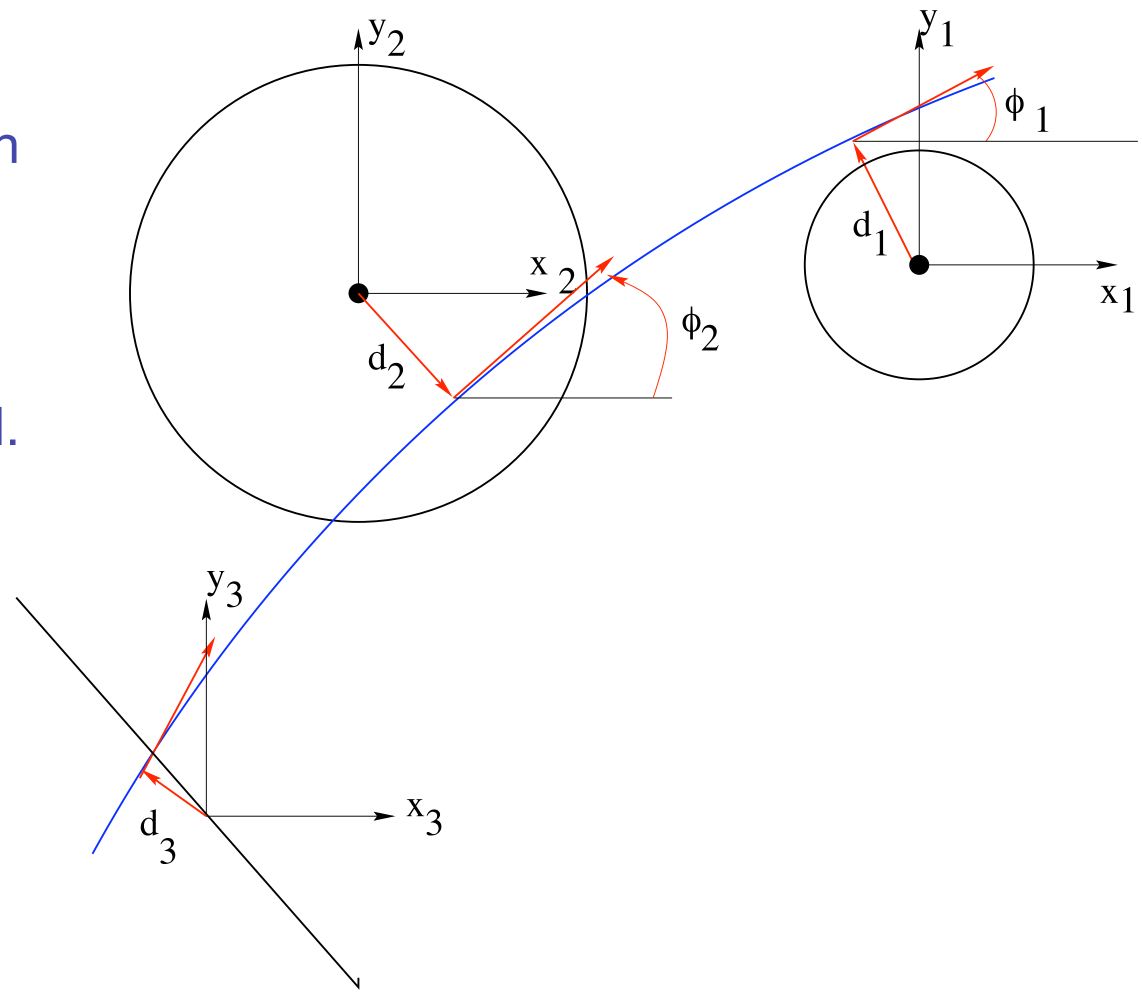
Exercise 1: No MS or Eloss



- Given:
  - η, V valid at n using (N….n); d_{n-1},σ_{n-1} (defined for point n-1).
- Question:
  - Compute best estimate of d, σ at (n-1) using N…(n-1).
- Answer
  - Compute $d_{n-1}(\eta)$, $\sigma_{n-1}(\eta,V)$.
  - d = weighted mean d_{n-1} and $d_{n-1}(\eta)$
  - σ = geometric mean of σ_{n-1} and $\sigma_{n-1}(\eta,V)$.

Transport as a Basis Transformation

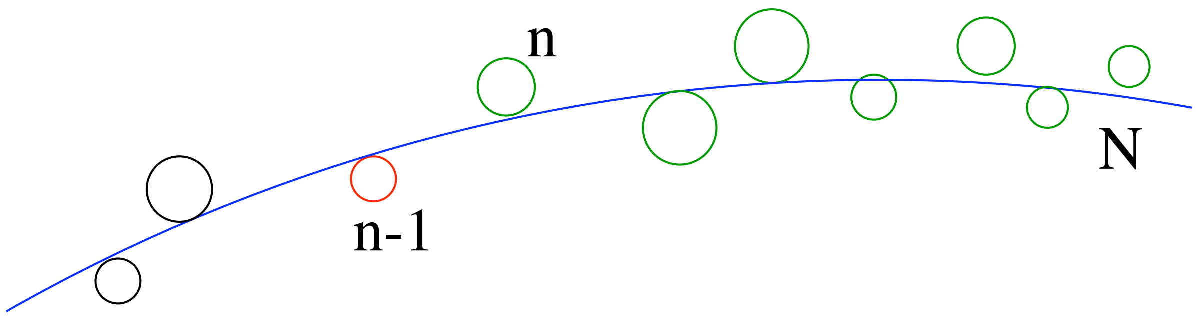
- The same trajectory (circle) can be defined in many bases.
- Angles and impact parameters change.
- Curvature is unchanged.



$$V' = A\ V\ A^T$$

$$A_{ij} = \frac{\partial \eta'_i}{\partial \eta_j}$$

Exercise 2: No MS or Eloss



- Given: η, V, d_{n-1},σ_{n-1} as before.
- Question: compute estimate of η, V at n-1 using N… (n-1).
- Answer:
  - New basis at n-1: η, V using info from N … (n).
  - d = weighted mean d_{n-1} and $d_{n-1}(\eta)$ (Trival!).
  - σ = geometric mean of σ_{n-1} and $\sigma_{n-1}(\eta,V)$.
  - Remaining η_i get pulled by V.
  - Remaining V_{ij} get pulled by V.

Minimize:

$$(\Delta\eta)^T V^{-1} (\Delta\eta) + \left(\frac{d(\eta) - d_{n-1}}{\sigma_{n-1}}\right)^2$$

The Real Math

Minimize:

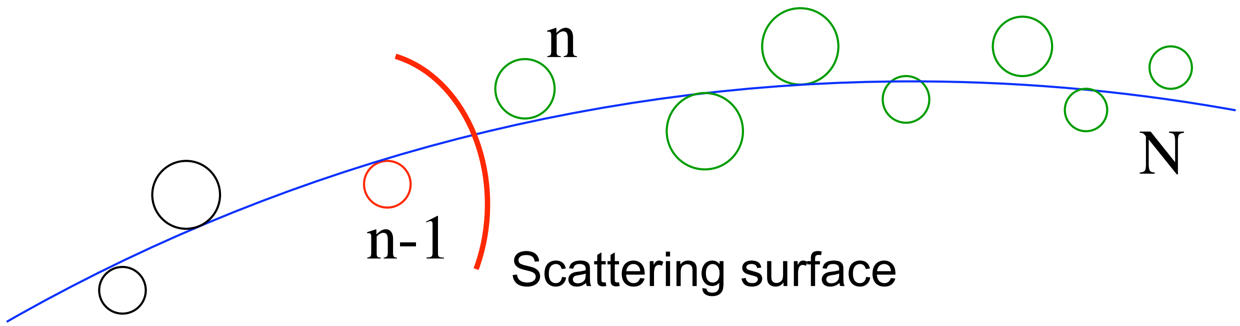
$$(\Delta\eta)^T V^{-1}(\Delta\eta) + \left(\frac{d(\eta) - d_{n-1}}{\sigma_{n-1}}\right)^2$$

Solution by linearizing about η:

$$\begin{aligned} V' &= V - \frac{VDD^TV}{\sigma^2 + D^TVD} \\ \eta' &= \eta + V'D\frac{d_m - d(\eta)}{\sigma^2} \\ D_i &= \frac{\partial d}{\partial \eta_i} \end{aligned}$$

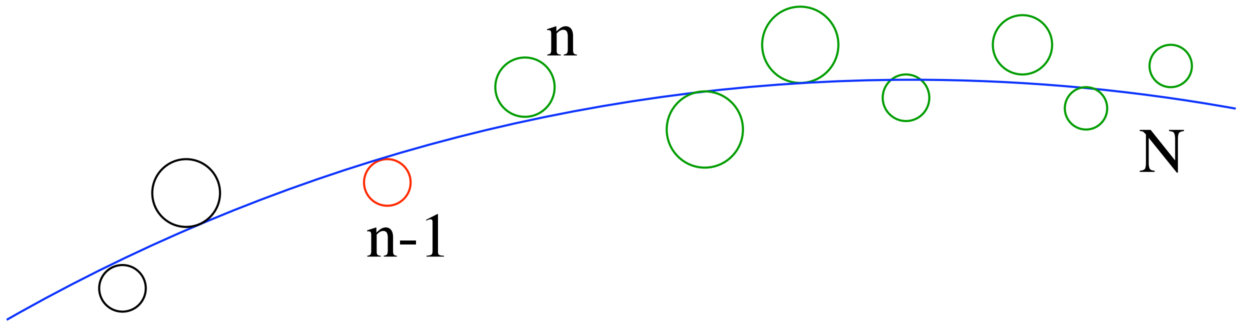
- $V' < V$
- If d is one of the track parameters, D is trivial.
- d can be any measurable thing, not necessarily a track parameter!

Exercise 3: Include MS and Eloss



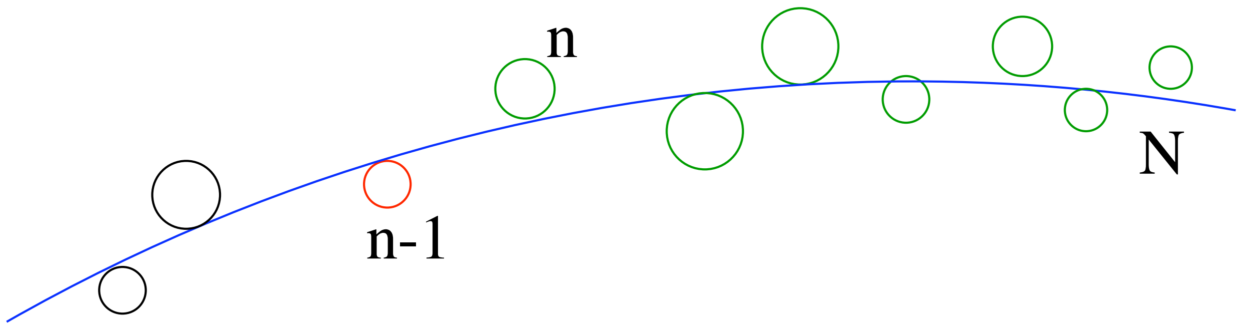
- Question:
  - Compute estimate of η, V at n-1 using N… (n-1).
- Answer:
  - New basis at scattering surface: η, V.
  - Gain energy. Increase error on slopes (thin scatter).
    - Also update correlation coefficients!
    - Do a thick scatter if needed.
  - New basis at measurement n-1: η, V.
  - Continue as in example 2.

How to Fit a Track



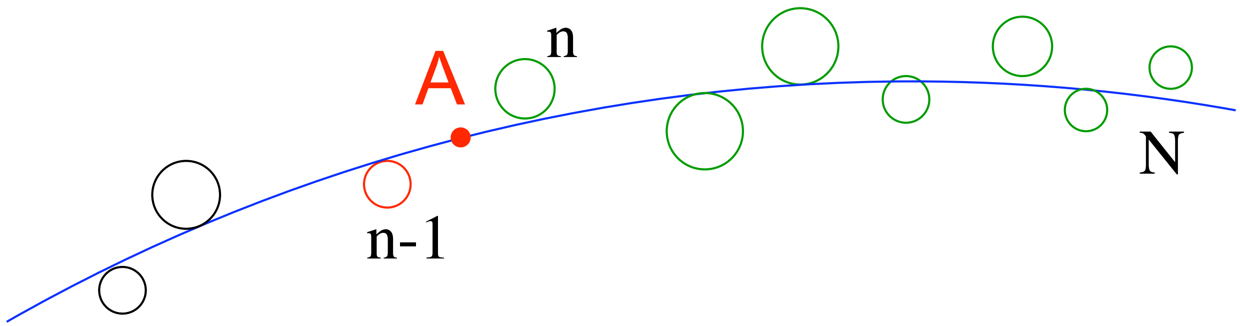
- Do pattern recognition:
  - List of hits, includes solution to sign ambiguities.
  - Initial value for η.
    - Neglibigle bias towards this starting value.
- V= diag(∞).
- Iterate through the hits and scattering surfaces.
- Change reference point to origin or average IP.
- Estimator only valid near IP!

Parameters Valid Near N



- Do everything as before but start the fit at the innermost measurement and run it outward towards hit N.
- Remember to lose, not gain energy, at MS surfaces.
  - MS does not change sign. It still increases V.

Parameters Valid Anywhere



- Question:
  - Find the best estimate of the trajectory that is valid in the neighbourhood of A.
- Answer:
  - Fit inwards (N … n): extrapolate to A.
  - Fit outwards (0 … n-1): extrapolate to A.
  - Compute the weighted mean!

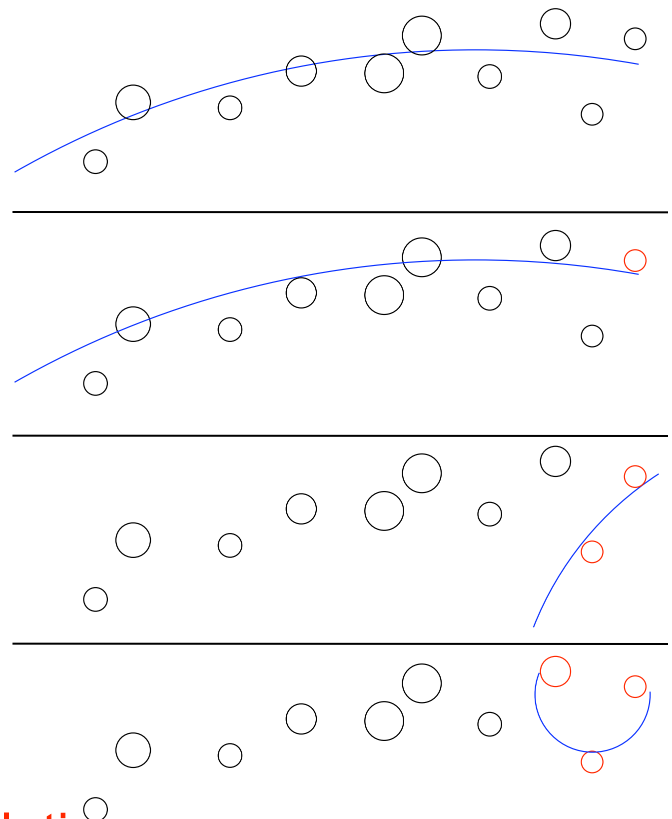
Computing the Weighted Mean

- Weighted mean of two tracks:
- Reminiscent of the weighted mean of two numbers.
- Works even if one or both tracks is under-determined!

$$
\begin{aligned}
V &= (V_1^{-1} + V_2^{-1})^{-1} \\
\eta &= V(V_1^{-1}\eta_1 + V_2^{-1}\eta_2)
\end{aligned}
$$

$$
\bar{x} = \frac{1}{\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}} \left(\frac{x_1}{\sigma_1^2} + \frac{x_2}{\sigma_2^2} \right)
$$

Startup Instability: Rare but Important



Parameters from pattern recognition.

Add first hit.

Add second hit.

Add third hit.
Transport to 4th hit is FUBAR.

Solution:

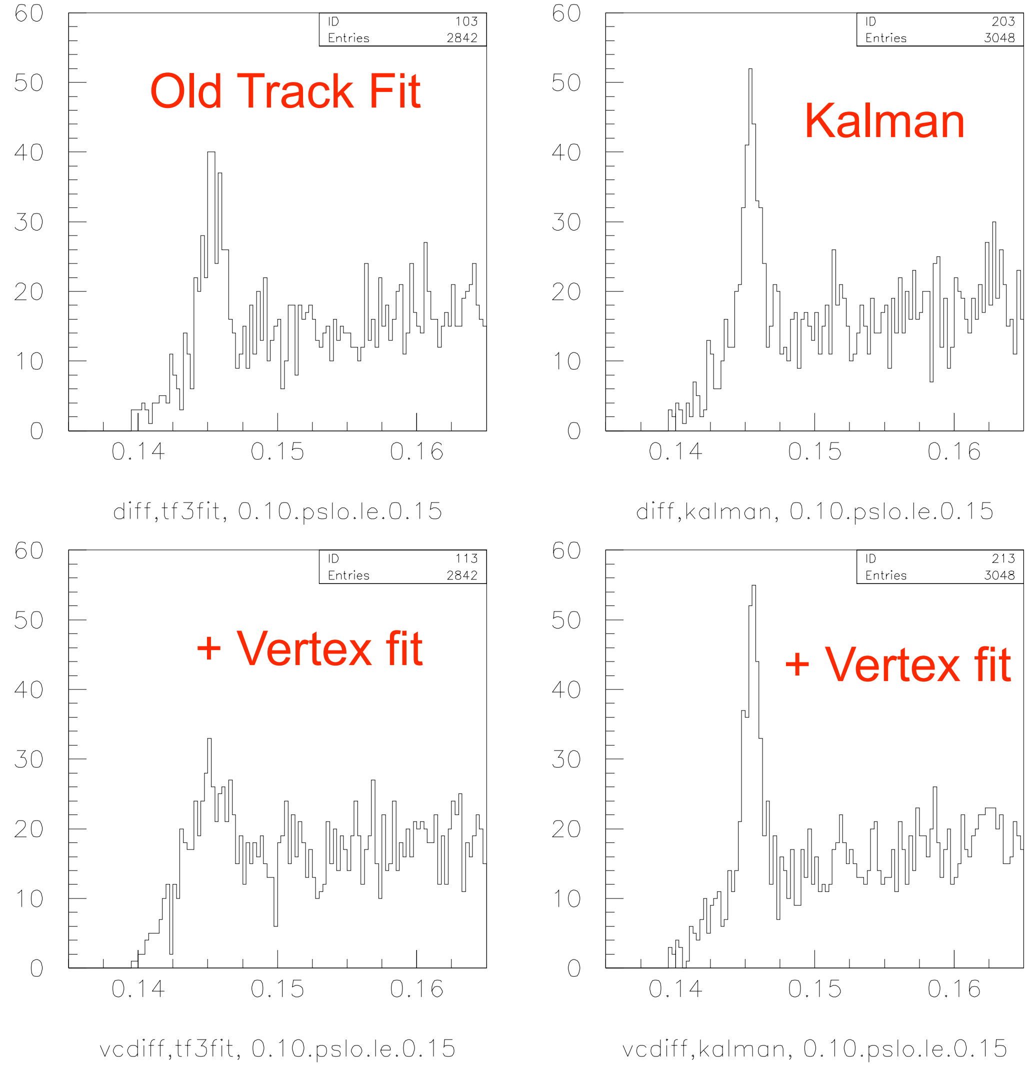
- Compute derivatives of first few points using the seed parameters but compute predicted measurement using full track info.
- Equivalent statement: Linearize fit about the seed parameters.

Comments

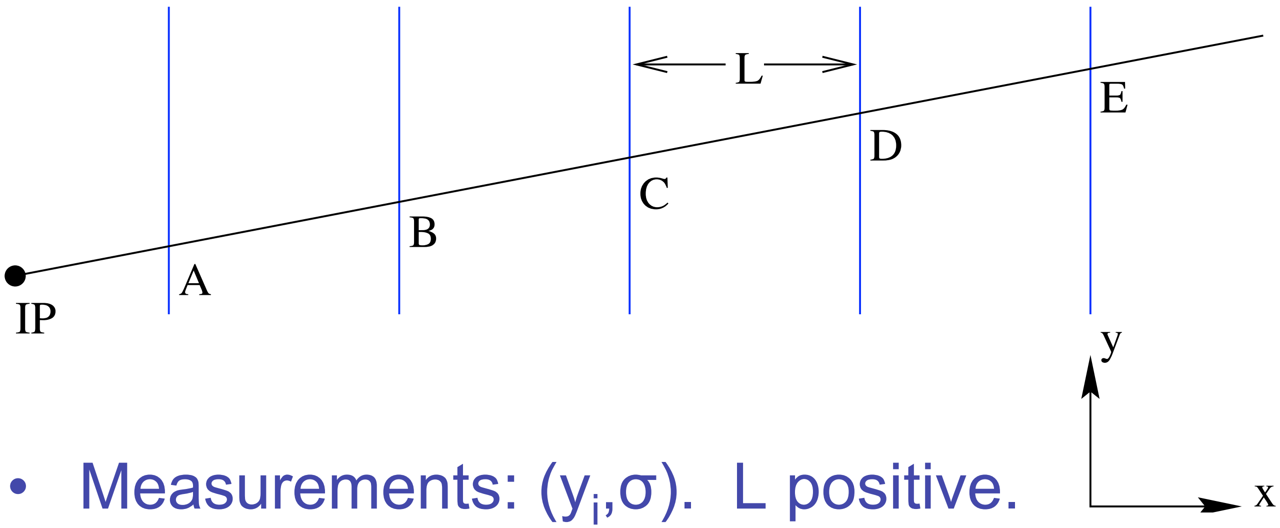
- If you use the formalism in which you make basis transformations to each hit or surface. Then there are two “hard” parts.
  - Writing the transport code.
    - But it’s just vacuum transport and its been done already.
  - Doing the derivatives of the measurements wrt to the local track parameters.
    - Its just a partial derivative; just do it.
- There are no hard parts.
- The rest is just a tedious bookkeeping exercise.

Game Changing Software

- $\Delta M(D^*-D)$
- CLEO II
  - No silicon.
- Same events.
- Key for $|V_{cb}|$.



Example 4: A Straight Line in 2D



$$y = mx + b$$

- Measurements: (y_i,σ). L positive.
- Want parameters at IP
- No multiple scattering or energy loss

$$\eta = \begin{pmatrix} m \\ b \end{pmatrix} \qquad V = \begin{pmatrix} V_{mm} & V_{mb} \\ V_{mb} & V_{bb} \end{pmatrix}$$

Starting Values

$$\eta = \begin{pmatrix} m_0 \\ b_0 \end{pmatrix} \qquad V = \begin{pmatrix} V_{0_{mm}} & 0 \\ 0 & V_{0_{bb}} \end{pmatrix} \qquad V_{0mm} >> \sigma^2$$

$$\boxed{V' = V - \frac{VDD^TV}{\sigma^2 + D^TVD}}$$

$$\begin{aligned} D &= \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ D^TVD &= V_{0_{bb}} \\ VDD^TV &= \begin{pmatrix} 0 & 0 \\ 0 & V^2_{0_{bb}} \end{pmatrix} \\ V' &= \begin{pmatrix} V_{0_{mm}} & 0 \\ 0 & \boxed{V_{0_{bb}} - \frac{V^2_{0_{bb}}}{\sigma^2 + V_{0_{bb}}}} \end{pmatrix} \end{aligned}$$

$$\begin{aligned} V' &\approx \begin{pmatrix} V_{0_{mm}} & 0 \\ 0 & \sigma^2 - \frac{\sigma^4}{V_{0_{bb}}} + \ldots \end{pmatrix} \\ &\approx \boxed{\begin{pmatrix} V_{0_{mm}} & 0 \\ 0 & \sigma^2 \end{pmatrix}} \end{aligned}$$

Numerical Precision!

Starting Values

$$\eta = \begin{pmatrix} m_0 \\ b_0 \end{pmatrix} \qquad V = \begin{pmatrix} V_{0_{mm}} & 0 \\ 0 & V_{0_{bb}} \end{pmatrix}$$

$$\boxed{\eta' = \eta + V' D \frac{d_m - d(\eta)}{\sigma^2}}$$

$$\begin{aligned} \eta' &= \begin{pmatrix} m_0 \\ b_0 \end{pmatrix} + \begin{pmatrix} V_{0_{mm}} & 0 \\ 0 & \sigma^2 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \frac{(y_E - b_0)}{\sigma^2} \\ &= \boxed{\begin{pmatrix} m_0 \\ y_E \end{pmatrix}} \end{aligned}$$

Numerical Precision

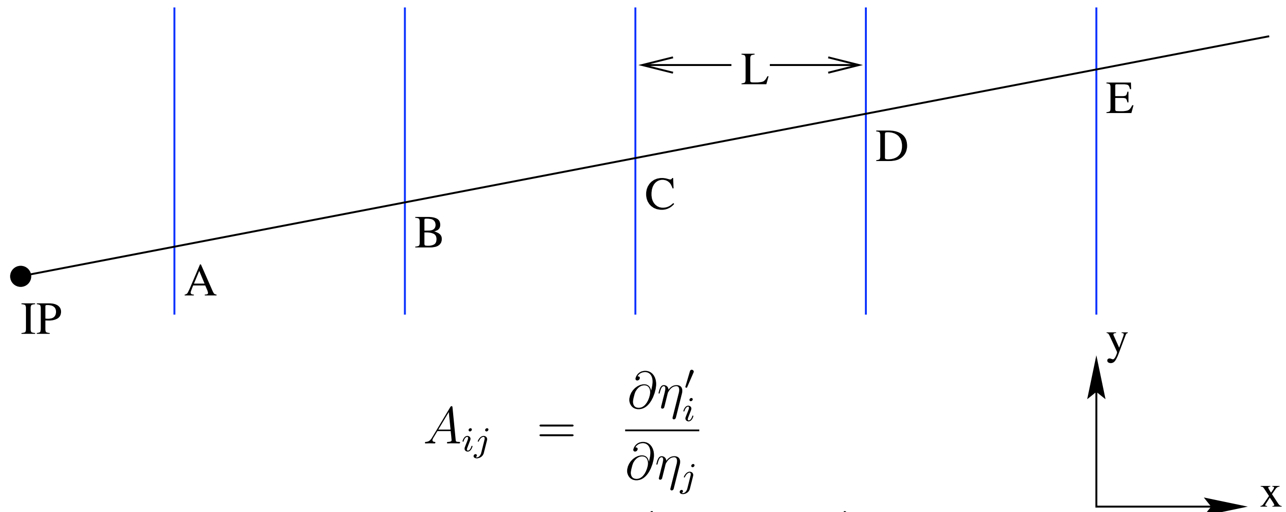
$$V'_{bb} = V_{0_{bb}} - \frac{V_{0_{bb}}^2}{\sigma^2 + V_{0_{bb}}}$$

$$\eta' = \begin{pmatrix} m_0 \\ b_0 \end{pmatrix} + \begin{pmatrix} V_{0_{mm}} & 0 \\ 0 & \sigma^2 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \frac{(y_E - b_0)}{\sigma^2}$$

- If V_{0bb} is too small, result is biased towards starting values.
- If V_{0bb} is too large the first equation has numerical precision problems.
  - Mostly a problem for 2D measurements in the middle of a track.
- Scale of bias is (not proven here):

$$O\left((y_E - b_0) \frac{\sigma^2}{V_{0bb}} \right)$$

Transport E to D



$$A_{ij} = \frac{\partial \eta'_i}{\partial \eta_j} = \begin{pmatrix} 1 & 0 \\ -L & 1 \end{pmatrix}$$

$$V'' = AV'A^T = \begin{pmatrix} V_{0_{mm}} & -LV_{0_{mm}} \\ -LV_{0_{mm}} & \sigma^2 + L^2 V_{0_{mm}} \end{pmatrix}$$

$$\eta'' = \begin{pmatrix} m_0 \\ y_E - m_0 L \end{pmatrix}$$

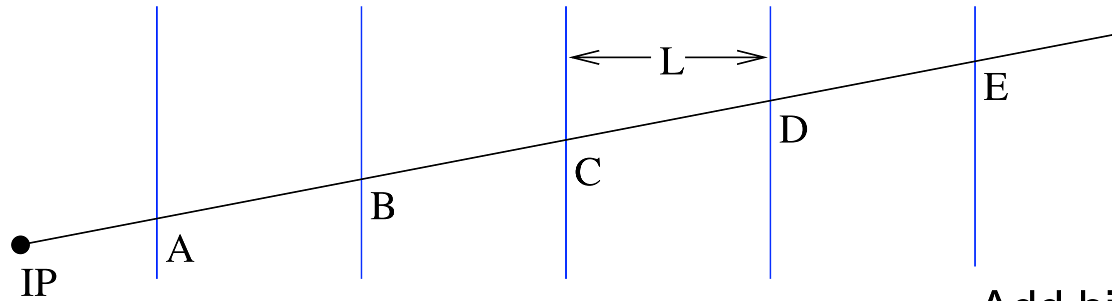
Add Hit at E

$$
\begin{aligned}
V''' &= \begin{pmatrix} V_{0_{mm}} & -LV_{0_{mm}} \\ -LV_{0_{mm}} & \sigma^2 + L^2 V_{0_{mm}} \end{pmatrix} \\
&\quad - \frac{1}{2\sigma^2 + L^2 V_{0_{mm}}} \begin{pmatrix} L^2 V_{0_{mm}}^2 & -LV_{0_{mm}}(\sigma^2 + L^2 V_{0_{mm}}) \\ -LV_{0_{mm}}(\sigma^2 + L^2 V_{0_{mm}}) & (\sigma^2 + L^2 V_{0_{mm}})^2 \end{pmatrix} \\
&\approx \begin{pmatrix} \frac{2\sigma^2}{L^2} & \frac{-\sigma^2}{L} \\ \frac{-\sigma^2}{L} & \sigma^2 \end{pmatrix}
\end{aligned}
$$

$$
\begin{aligned}
\eta''' &= \begin{pmatrix} m_0 \\ y_E - m_0 L \end{pmatrix} + V''' \begin{pmatrix} 0 \\ 1 \end{pmatrix} \frac{(y_D - (y_E - m_0 L))}{\sigma^2} \\
&= \begin{pmatrix} \frac{(y_E - y_D)}{L} \\ y_D \end{pmatrix}
\end{aligned}
$$

- Starting values have disappeared to leading order.
- Simplified equations are the intuitive ones you would expect.

Transport D to C and Add Hit



Transport:

$$V^{(iv)} = AV'''A^T = \begin{pmatrix} \frac{2\sigma^2}{L^2} & \frac{-3\sigma^2}{L} \\ \frac{-3\sigma^2}{L} & 5\sigma^2 \end{pmatrix}$$

$$\eta^{(iv)} = \begin{pmatrix} m''' \\ b''' - m'''L \end{pmatrix} = \begin{pmatrix} \frac{(y_E - y_D)}{L} \\ y_D - (y_E - y_D) \end{pmatrix}$$

Add hit:

$$V^{(v)} = \begin{pmatrix} \frac{\sigma^2}{2L^2} & \frac{-\sigma^2}{2L} \\ \frac{-\sigma^2}{2L} & \frac{5\sigma^2}{6} \end{pmatrix}$$

$$\eta^{(v)} = \begin{pmatrix} \frac{y_E - y_C}{2L} \\ \frac{2y_D - y_E + 5y_C}{6} \end{pmatrix}$$

Again the standad results are obtained.

Suggested Exercises

- Replace distance from D to C by L_2
  - Redo do the last page.
  - Verify that the correlation coefficient in V goes to -1 in the limit $L_2 \gg L$.
  - This circumstance can lead to precision problems.
- Finish the exercise all the way to the IP.

Add Multiple Scattering

- Add hit at E.
- Transport to D.
- Add hit at D. No change yet.
- Thin scatter at D by an angle δ:

$$V^{(vi)} = \begin{pmatrix} \frac{2\sigma^2}{L^2} \boxed{+\delta^2} & \frac{-\sigma^2}{L} \\ \frac{-\sigma^2}{L} & \sigma^2 \end{pmatrix}$$

- Transport to C:

$$V^{(vii)} = \begin{pmatrix} \frac{2\sigma^2}{L^2} \boxed{+\delta^2} & \frac{-3\sigma^2}{L} - \boxed{\delta^2 L} \\ \frac{-3\sigma^2}{L} - \boxed{\delta^2 L} & 5\sigma^2 + \boxed{\delta^2 L^2} \end{pmatrix}$$

Add Multiple Scattering - 2

- Add hit at C:

$$V^{(viii)} = \begin{pmatrix} \frac{2\sigma^2}{L^2} + \delta^2 & \frac{-3\sigma^2}{L} - \delta^2 L \\ \frac{-3\sigma^2}{L} - \delta^2 L & 5\sigma^2 + \delta^2 L^2 \end{pmatrix} - \frac{1}{6\sigma^2 + \delta^2 L^2} \begin{pmatrix} (3\sigma^2/L + \delta^2 L)^2 & -(3\sigma^2/L + \delta^2 L)(5\sigma^2 + \delta^2 L^2) \\ -(3\sigma^2/L + \delta^2 L)(5\sigma^2 + \delta^2 L^2) & (5\sigma^2 + \delta^2 L^2)^2 \end{pmatrix}$$

- Life is too complicated.
  - Give up trying to do write things explicitly.
  - Just run the program.

The Killer Precision Bug

- SiD: Coffee can all Silicon.
  - Inner 5 layers: pixels.
  - Outer 5 layers: strips.
- Inwards fit:
  - 1 or 2 endcap strips at zmax. For magic angles, no more z measurements for O(1.5m) of trajectory.
  - Correlations coefficients in V after transport are:
    - 0.99999999xxxx (8 9's).
  - Pixels are a 2D measurement.
  - Need $1 \neq 1-\varepsilon^2$ precision; dead on IEEE
  - Exacerbated by choice of bases

Summary and Conclusions

- I hope you have the basic idea.
- There are no hard parts.
  - Equations look intimidating but if you work in a convenient representation there are lots of 1's and 0's in the matrices.
  - It is possible to simplify the equations and see what is happening.
  - Pay attention to when errors increase or decrease.
  - Pay attention to correlation coefficients.
- Can get the parameters any where on the track by fitting inwards, fitting outwards and averaging.
- Bias towards the starting parameter values is negligible.
- Computes in O(N).

For Further Information

- Billoir: NIM 225 (1984) 352
  - Some trivial math errors but really good conceptually.
  - He's a physicist writing for physicists.
- Papers by Rudolph Fruhwirth (Vienna)
  - He's a math guy. A very different perspective.
- My home page: http://home.fnal.gov/~kutschke
  - Look at bottom section "Tracking Notes"
  - The "Simple Explanation …" note is the 2D linear example shown here.
  - The "Residuals with and without … " was not talked about here but it discusses an important idea.
- Some of the notes have a bit of a bibliography.