

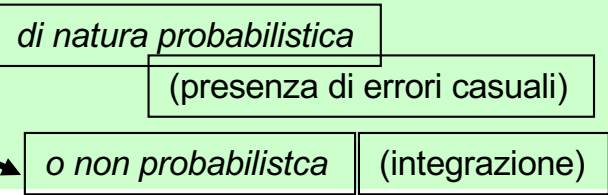
Monte Carlo

* N.Metropolis, S.M.Ulam, "The Monte Carlo Method", J.Am.Stat.Ass., 44 (1949) 335-341

{ T. Lund, "An Introduction to the Monte Carlo Method", CERN report, HS-RP/067, 1981. }

To solve problems in Physics, Mathematics, Finances, Social Science, Biology, etc.:

- in an approximate way
- through the simulation of the problem
- using random numbers



Computers ! → in USA from 1950.... Random numbers = sequences of numbers

• true random	→	non predictable, non reproducible	(radioactive source)
• pseudo-random	→	reproducible, with a period	(mathematical formula)
• quasi-random	→	reproducible, minimization	(ad hoc for the problem)

Von Neumann: x_n with 2t digits on base R,
 x_n^2 with 4t digits
 x_{n+1} given by the 2t "central" digits

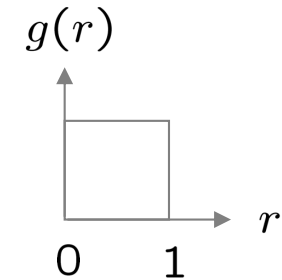
$$x_{n+1} = \left[\frac{x_n^2}{R^t} \right] - \left[\frac{x_n^2}{R^{3t}} \right] \cdot R^{2t}$$


The Monte Carlo method

What it is: a numerical technique for calculating probabilities and related quantities using sequences of random numbers.

The usual steps:

- (1) Generate sequence r_1, r_2, \dots, r_m uniform in $[0, 1]$.
- (2) Use this to produce another sequence x_1, x_2, \dots, x_n distributed according to some pdf* $f(x)$ in which we're interested (x can be a vector).
- (3) Use the x values to estimate some property of $f(x)$, e.g., fraction of x values with $a < x < b$ gives $\int_a^b f(x) dx$.
→ MC calculation = integration (at least formally)



MC generated values = 'simulated data'

→ use for testing statistical procedures

*one may also use the cumulative function known by steps...

Random number generators

Goal: generate uniformly distributed values in $[0, 1]$. *

Toss coin for e.g. 32 bit number... (too tiring).

→ 'random number generator'

= computer algorithm to generate r_1, r_2, \dots, r_n .

Example: multiplicative linear congruential generator (MLCG)

$$n_{i+1} = (a n_i) \bmod m, \quad \text{where}$$

n_i = integer

a = multiplier

m = modulus

n_0 = **seed** (initial value)

N.B. mod = modulus (remainder), e.g. $27 \bmod 5 = 2$.

This rule produces a sequence of numbers n_0, n_1, \dots

* In the fifties sequences of several millions were generated, truly random !

$r_i = n_i/m$ are in $[0, 1]$ but are they ‘random’?

Choose a, m so that the r_i pass various tests of randomness:

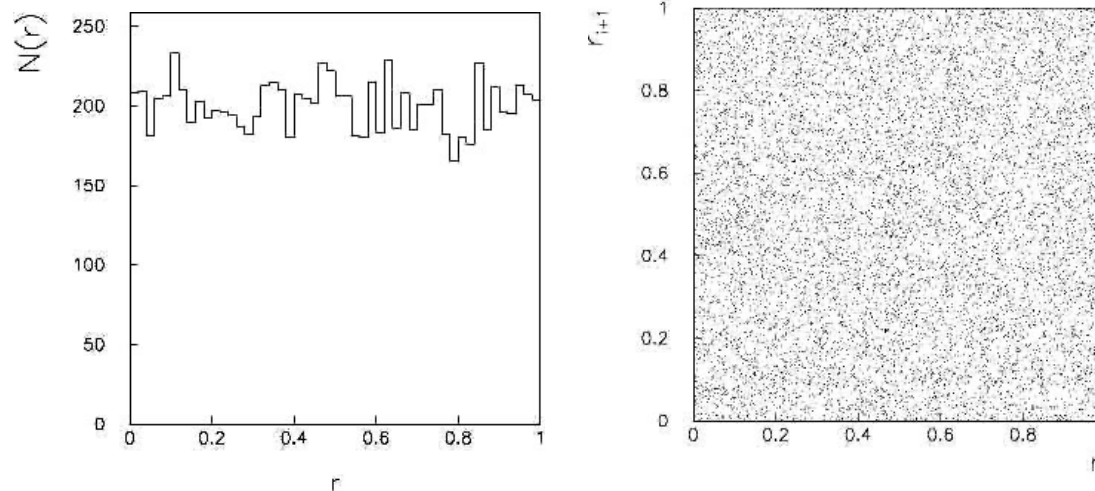
uniform distribution in $[0, 1]$,

all values independent (no correlations between pairs),

e.g. L’Ecuyer, Commun. ACM **31** (1988) 742 suggests

$a = 40692$

$m = 2147483399$



Far better generators available, e.g. **TRandom3**, based on Mersenne twister algorithm, period = $2^{19937} - 1$ (a “Mersenne prime”).

See F. James, Rep.Prog.Phys. 43 (1980), F. James, Comp. Phys. Comm. 60 (1990) 111;
Good references in Max Sioli: http://www.bo.infn.it/~sioli/didattica/1415/asd/asd_sioli.htm

Period 4194304 - 10¹⁸

(2²² per RAND.f - 2⁶⁰ for RANECU in GEANT)



Initial generation of random di numbers ξ in the interval [0,1]

... and for any function $f(x)$?

Example for one Gaussian

Impossible for a single Gaussian, possible for two independent Gaussians:

$$f(x_1, x_2) = \frac{1}{2\pi} \cdot \exp\left(-\frac{(x_1^2 + x_2^2)}{2}\right)$$

One can define the transformation:

$$x_1 = \sqrt{-2 \ln \xi_1} \cos(2\pi \xi_2)$$

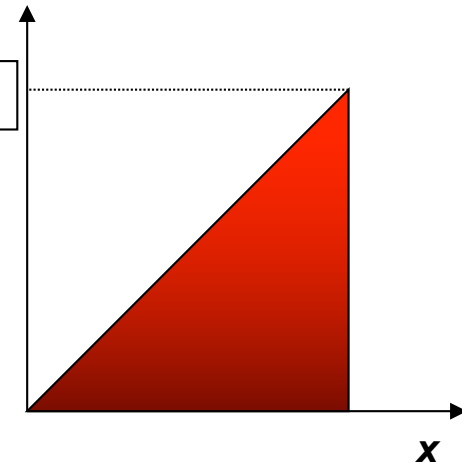
$$x_2 = \sqrt{-2 \ln \xi_1} \sin(2\pi \xi_2)$$

being

$$d\xi_1 d\xi_2 = |J| \cdot dx_1 dx_2 = f(x_1, x_2) \cdot dx_1 dx_2$$

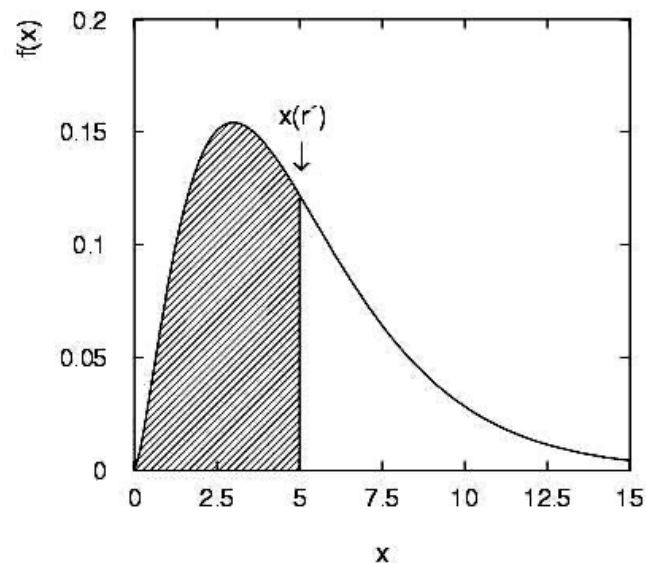
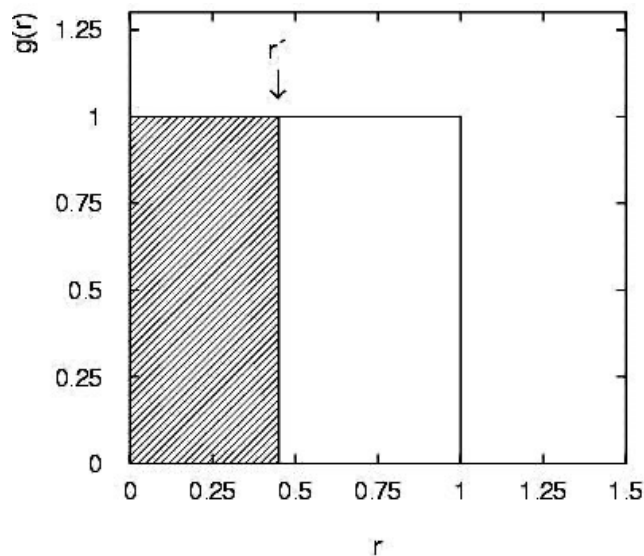
1. Direct method
2. Rejection method
3. Composition method
4. Von Neumann-Forsythe

Another example



The transformation method

Given r_1, r_2, \dots, r_n uniform in $[0, 1]$, find x_1, x_2, \dots, x_n that follow $f(x)$ by finding a suitable transformation $x(r)$.

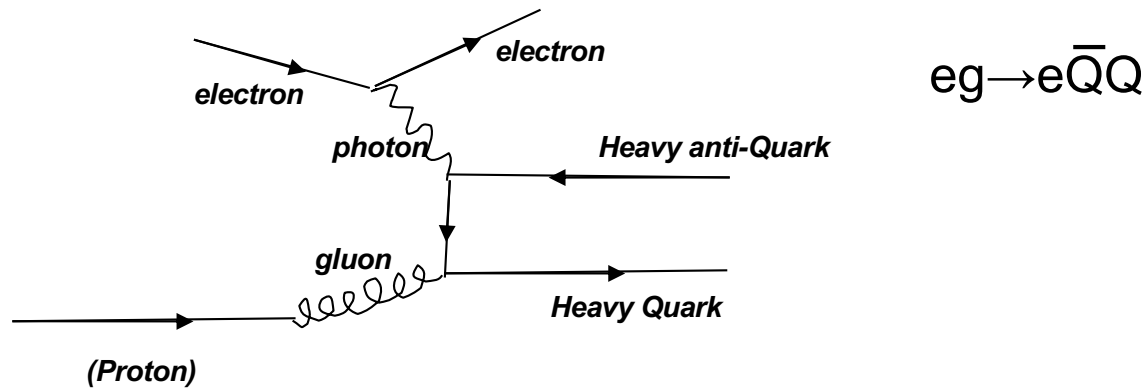


Require: $P(r \leq r') = P(x \leq x(r'))$

$$\text{i.e. } \int_{-\infty}^{r'} g(r) dr = r' = \int_{-\infty}^{x(r')} f(x') dx' = F(x(r'))$$

That is, set $F(x) = r$ and solve for $x(r)$.

Using the Monte Carlo method to simulate events produced events by a physical process:



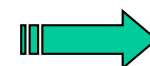
Modelled by the differential cross section $\frac{d\sigma}{dydQ^2d\eta dzd\phi}$
 (d.o.f. = 3 trimomenta – 4 conservation laws)

$P_{\#}$: point in the phase space
 $W(P_{\#})$: weight (cross section)
 W_{\max} : maximum weight
 $R_{\#} \in [0,1]$: pseudo-random number

Events NOT weighted if it holds:



$$W(P_{\#}) \geq W_{\max} \cdot R_{\#}$$



generic configuration Q
 with probability $W(Q)$

Phase 1. Initialization and computing of W_{\max}
Phase 2. Generation of the single event
efficiency $\varepsilon = \frac{N_{ev.gen.}}{N_w}$

One possibility is to subdivide the phase space in some parts.
 Another one stays in the “direct” method.

To generate x in $[x_0, x_1]$, having the total cross section: $\sigma = \int_{x_0}^{x_1} dx \frac{d\sigma}{dx} \rightarrow \sigma = \int_0^1 d\xi W(\xi)$

$$x = x_0 + (x_1 - x_0) \cdot \xi$$

$$\xi \in [0,1]$$

$$W(\xi) = \frac{dx}{d\xi} \frac{d\sigma}{dx}$$

with $\frac{dx}{d\xi} = x_1 - x_0$

complication: $\frac{d\sigma}{dx} \propto \frac{1}{x} \rightarrow$ We generate : $\log x = \log x_0 + (\log x_1 - \log x_0) \cdot \xi$

Id est: $x = x_0 \left(\frac{x_1}{x_0} \right)^\xi$

$$\frac{dx}{d\xi} = \log \left(\frac{x_1}{x_0} \right) \cdot x$$

In general the Jacobian is computed: $J = \left| \frac{\partial(y, Q^2, \eta, z, \phi)}{\partial(\xi_1, \xi_2, \xi_3, \xi_4, \xi_5)} \right|$

$$\sigma = \int_{\Omega} dy dQ^2 d\eta dz d\phi \frac{d^5\sigma}{dy dQ^2 d\eta dz d\phi} = \int_0^1 d\xi_1 d\xi_2 d\xi_3 d\xi_4 d\xi_5 \cdot W(\xi_1, \xi_2, \xi_3, \xi_4, \xi_5)$$

$$W(\xi_1, \xi_2, \xi_3, \xi_4, \xi_5) = J \cdot \frac{d\sigma}{dy dQ^2 d\eta dz d\phi}$$

and the extraction is done following the integral of the Jacobian.

Thus, it results:

$$W(\xi) = \frac{d\sigma}{d\xi} \propto \log \left(\frac{x_1}{x_0} \right)$$

Kalman filter

Applications in the HEP field

- Filter
- Kalman
- Markov (*process*)
- Non-recursive/recursive (*filter*)
- High Energies (*track detectors*)
- Fit (*tracks/vertices*)
- Multiple Scattering
- Filter di Kalman (*tecnica*)
- Applications
- Comparisons with the Filter of Wiener in presence of dE/dx (*non-linear systems*)

} some hints

Given a Physics System

(submarine, rocket, satellite, electricity grid, ... particle in a non destructive detector)

subjected to **random** noises,

the issue to determine its **STATE** “pulling out” from a set of measurements

(*observations*) of the system itself,

is defined as the ESTIMATION problem, or **FILTERING**.

Physics System → **(model)** Process of Markov

(i.e.) output of a stochastic differential equation
(or finite differences)

Probabilistic point-of-view

opposed to

Statistical point-of-view ⇒ STATE → mean, variance
(deterministic)

Kalman, R.E. *del Research Institute for Advanced Study, Baltimore,
Maryland, USA.*

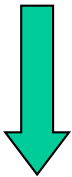


(e Bucy, R.C.)

Papers: Journal of Basic Engineering, March 1960 and March 1961

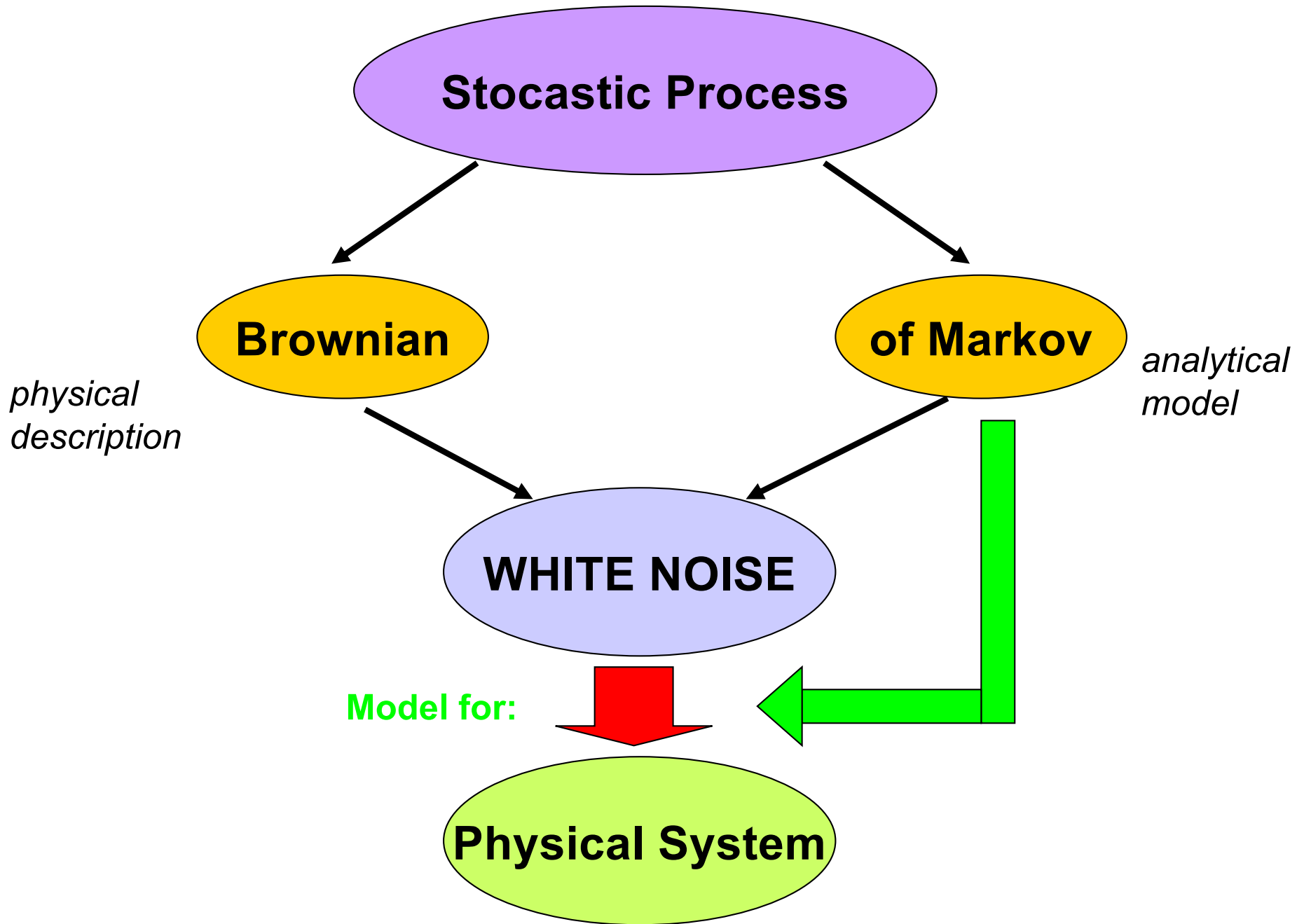
Under contracts: *United States Air Force AF 49(638)-382, AF33(616)-6952
Bureau of Naval Weapons NORD-73861*

Furthemore: re-analysis of the Kalman filter with ***statistical methods***
(through Least Squares and Maximum Likelihood)



Extensive use for the determination:

- satellites' orbit
- navigation of submarines and rockets
- space flights



STOCHASTIC Process (**random**) : $\{ \mathbf{x}_t, t \in T \}$, with \mathbf{x}_t scalar or vector

→ Family of *random* variables parametrized in t

→ \mathbf{x}_t : *state vector*

STOCHASTIC Process at INDEPENDENT INCREMENTS **if:**

$(\mathbf{x}_{t2} - \mathbf{x}_{t1}); (\mathbf{x}_{t3} - \mathbf{x}_{t2}); \dots (\mathbf{x}_{tn} - \mathbf{x}_{t(n-1)})$: independent *random* variables

and even STAZIONARY **if** : $(\mathbf{x}_{t+h} - \mathbf{x}_{t+h})$ e $(\mathbf{x}_t - \mathbf{x}_t)$ own the same distribution $\forall h$

STOCHASTIC BROWNIAN Process **if:**

- $\{ \mathbf{x}_t, t \geq 0 \}$ owns *increments independent and stationary*,
- $\forall t \geq 0$, \mathbf{x}_t owns a distribution *NORMAL*, centered at ZERO.

Processo STOCASTICO di MARKOV **se**:

$\{x_{t1}, x_{t2}, \dots, x_{t(n-1)}, x_{tn}\}$ è tale per cui: $Pr\{x_{tn} \leq \lambda \mid x_{t1}, x_{t2}, \dots, x_{t(n-1)}\} = Pr\{x_{tn} \leq \lambda \mid x_{t(n-1)}\}$
 $\forall \lambda$ reale e $\forall n$ intero.

Si legge: Probabilità che la variabile casuale x_{tn} possa essere più piccola di λ
(funzione di distribuzione) dati i precedenti valori $x_{t1}, \dots, x_{t(n-1)}$
(probabilità condizionata)

La proprietà di Markov dice che la legge di probabilità del processo stocastico $\{x_t, t\}$ nel futuro, una volta che si trova in un dato **STATO**, **NON DIPENDE** da come il processo è arrivato lì.

Principio generalizzato della causalità

Ne segue che LA LEGGE DI PROBABILITÀ DI UN PROCESSO DI MARKOV è specificata da

$p(x_t)$ e $p(x_t | x_{t-1})$, $\forall t > \tau$, con p = densità di probabilità di transizione

NOTA: PROCESSI DI MARKOV \equiv analogo STOCASTICO delle eq. differenziali ordinarie:

$$\frac{dx(t)}{dt} = f(x(t))$$

RUMORE BIANCO:

Un processo STOCASTICO **WHITE** è un processo di MARKOV per cui:

$$p(\mathbf{x}_k | \mathbf{x}_l) = p(\mathbf{x}_k) \quad (k > l) \text{ cioè: tutte le } \mathbf{x}_k \text{ sono mutuamente indipendenti.}$$

Si dimostra che: **un processo bianco gaussiano** (cioè con distribuzione normale)
è la derivata di un processo di Markov

In definitiva: *Molti processi fisici soggetti a disturbi casuali i cui STATI possono essere rappresentati da vettori a dimensioni finite, possono essere modellati via equazioni alle differenze finite*

$$\mathbf{x}_{k+1} = f(\mathbf{x}_k, \mathbf{w}_{k+1}, \mathbf{t}_k)$$

vettore di stato noise

equazione stocastica alle differenze, che rappresenta un sistema dinamico stocastico discreto

Se $\{\mathbf{w}_k\}$ è **white**, la soluzione dell'equazione è una sequenza di MARKOV !

MARKOV CHAIN

da A. Romano (Università di Napoli)

Il **metodo delle differenze finite** rappresenta il metodo più semplice per la ricerca di una **soluzione numerica** di un'equazione o di un sistema di equazioni differenziali, ossia **dei valori approssimati della soluzione in un numero discreto di valori della variabile indipendente** (v. Fig.3).

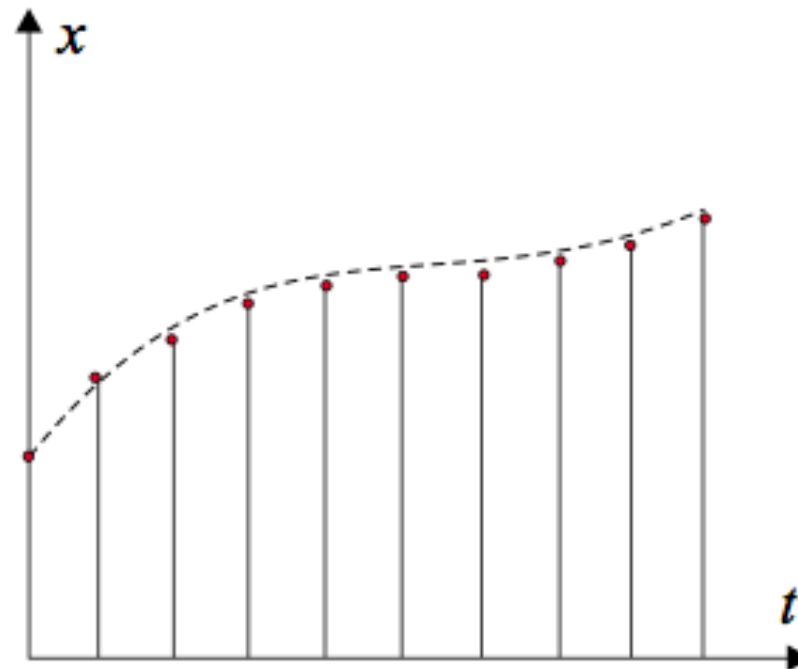


Fig. 3

Sistemi lineari: $x_{k+1} = \Phi_{k+1,k} \cdot x_k + \Gamma_{k+1,k} \cdot \omega_{k+1,k}$

matrici a valori costanti

con equazioni di evoluzione (*soluzione "statistica"*) per la media e la varianza:

$$\begin{cases} \hat{x}_{k+1} = \Phi_{k+1,k} \cdot \hat{x}_k \\ P_{k+1} = \Phi_{k+1,k} \cdot P_k \cdot \Phi_{k+1,k}^T + \Gamma_{k+1} \cdot Q_{k+1} \cdot \Gamma_{k+1}^T \end{cases}$$

matrice di covarianza

matrice di covarianza del rumore bianco

SOLUZIONE DEL PROBLEMA DI **PREDIZIONE**

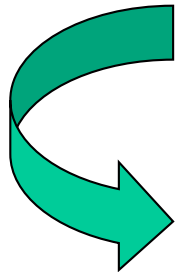
Problema della **stima** di x_k :

$k = l \Rightarrow$ **FILTERING**
 $k < l \Rightarrow$ **SMOOTHING**
 $k > l \Rightarrow$ **PREDICTION**

(interpolazione)

FILTRO: trovare le **EQUAZIONI** di evoluzione della **MEDIA**
e della **VARIANZA**, in presenza delle **MISURE**.

FILTRO **OTTIMALE**: quello che rende minima la varianza (*errori quadratici*).
cioè con massima “efficienza”



FILTRO di KALMAN

è in alternativa all'approccio (*statistico*) globale (*per esempio*) dei minimi quadrati:

⇒ Minimizzazione del χ^2 sulle l osservazioni
(*estrazione/stima di $\hat{x}_1, \hat{x}_2, \dots, \hat{x}_l$ contemporaneamente*)

chiamato: **FILTRO DI WIENER**

Breve nota sul FILTRO di **WIENER**. *Stimatore ottimale NON-recursive*

Date le ℓ misure: $\mathbf{y}_k = \mathbf{H}\mathbf{x}_k + \varepsilon_k$, con $k=1,2,\dots,\ell$ si cerca $\hat{x} = \sum_{k=1}^{\ell} h_k y_k$
↑
coefficienti da trovare in modo
che $E(x - \hat{x})^2$ sia minimo

Soluzione (equazione di Wiener-Hopf): $\sum_{k=1}^{\ell} h_k p_{kj}^y = P_j^{xy}$; $j = 1,2,\dots,\ell$

Sistema lineare di ℓ equazioni in ℓ incognite.

Problemi:

- si richiede la conoscenza della matrice di covarianza P_{kj}^y iterazione
- il numero ℓ di misure deve essere fissato a priori
- se il numero di misure cambia il calcolo si deve ripetere
(importante in relazione agli “outliers”: misure fuori di 3σ)
- si richiede l’inversione della matrice (P_{kj}^y) di dimensione $\ell * \ell$

FILTRO di **KALMAN**: estimatore ottimale ricorsivo $\hat{\mathbf{x}}_k = \mathbf{a}_k \cdot \hat{\mathbf{x}}_{k-1} + \mathbf{b}_k \cdot \mathbf{y}_k$

Lo stimatore al punto \mathbf{k} , ($\hat{\mathbf{x}}_k$) è dato dalla media “opportunamente pesata” della stima ottimale ($\hat{\mathbf{x}}_{k-1}$) sulle ($\mathbf{k}-1$) misure precedenti e la \mathbf{k} -esima misura (\mathbf{y}_k).

Goal (*realizzato da Kalman*): trovare \mathbf{a}_k e \mathbf{b}_k in modo ottimale (minimo errore quadratico medio).

L'informazione viene aggiunta un punto alla volta (*real time processor*) in alternativa al filtro di WIENER (*batch processor*).

Solo alla fine, quando anche l'ultima misura è stata considerata ($\mathbf{k}=\mathcal{I}$), tutta l'informazione disponibile è stata utilizzata.

Se è richiesta una stima ottimale in $\mathbf{k} < \mathcal{I}$, si “interpola” (**smoothing**).

Rivelatore ZEUS

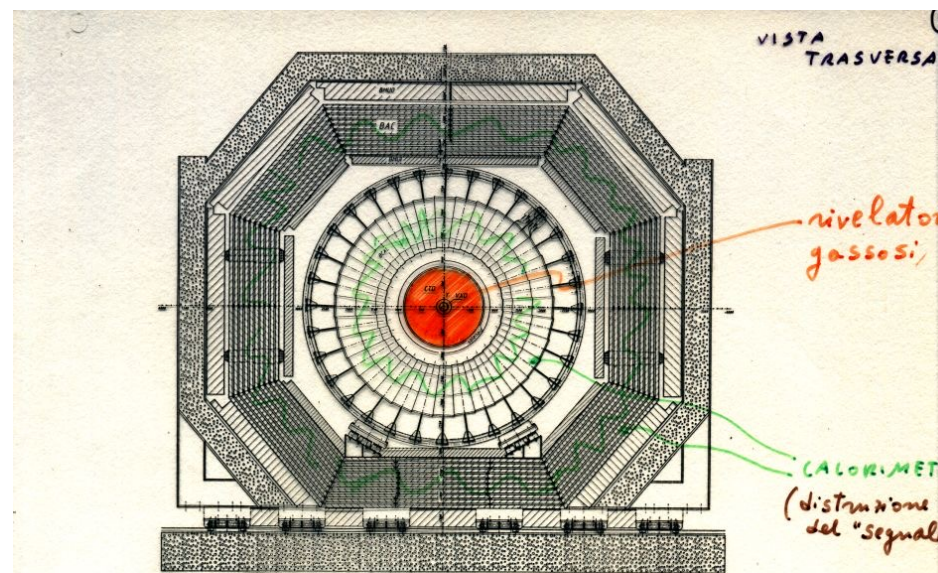


Fig. 2.2 Section of the ZEUS detector perpendicular to the beam.

RIVELATORE ZEUS ≈ 1000 metri cubi

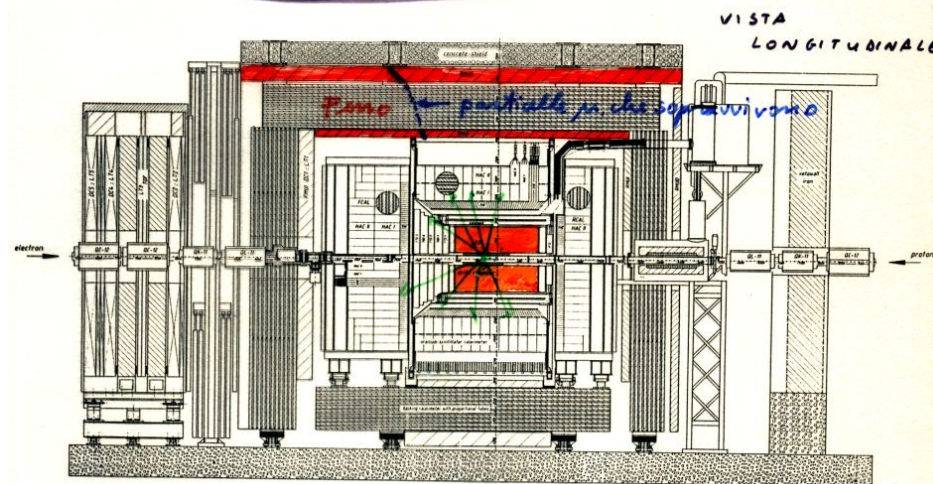
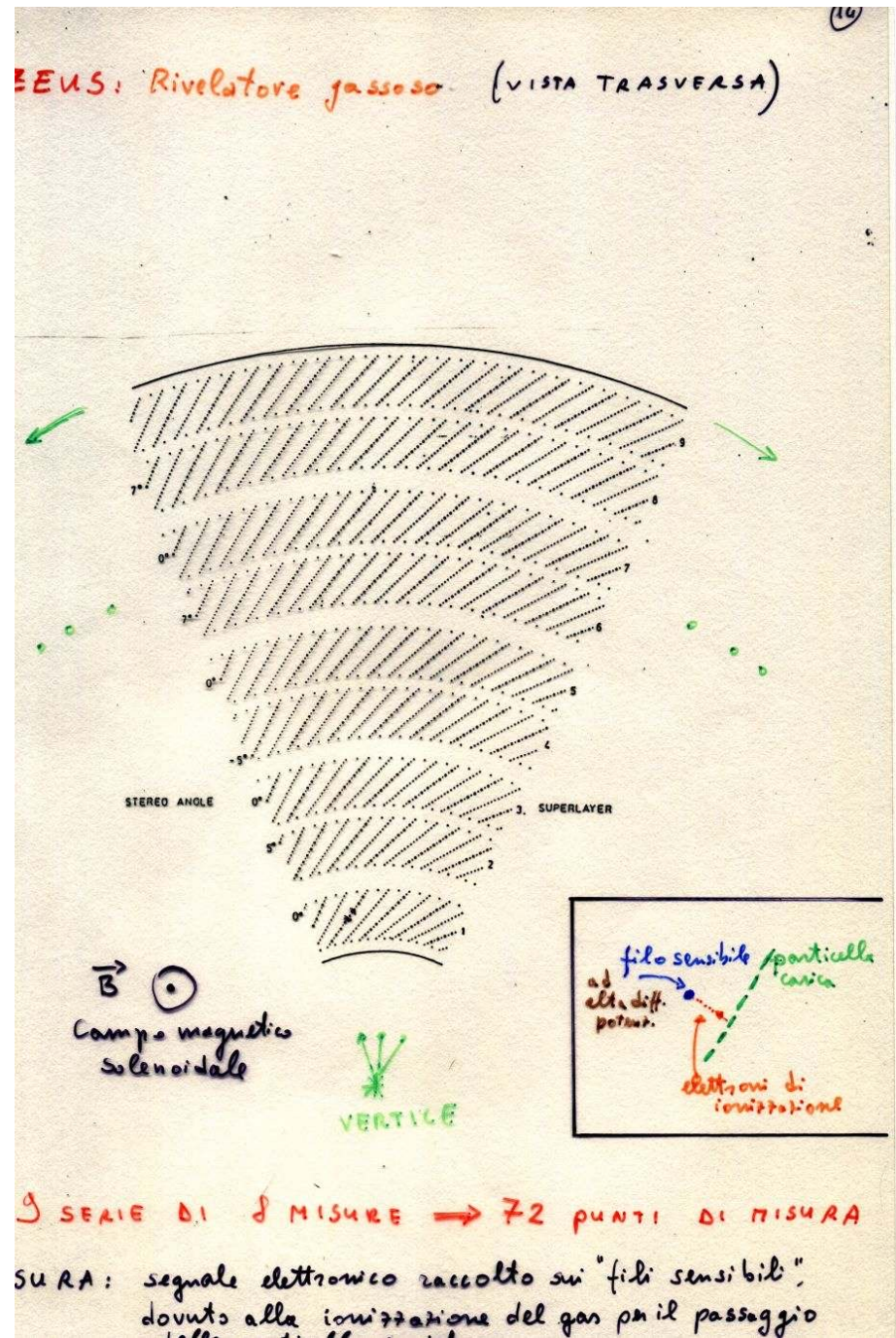


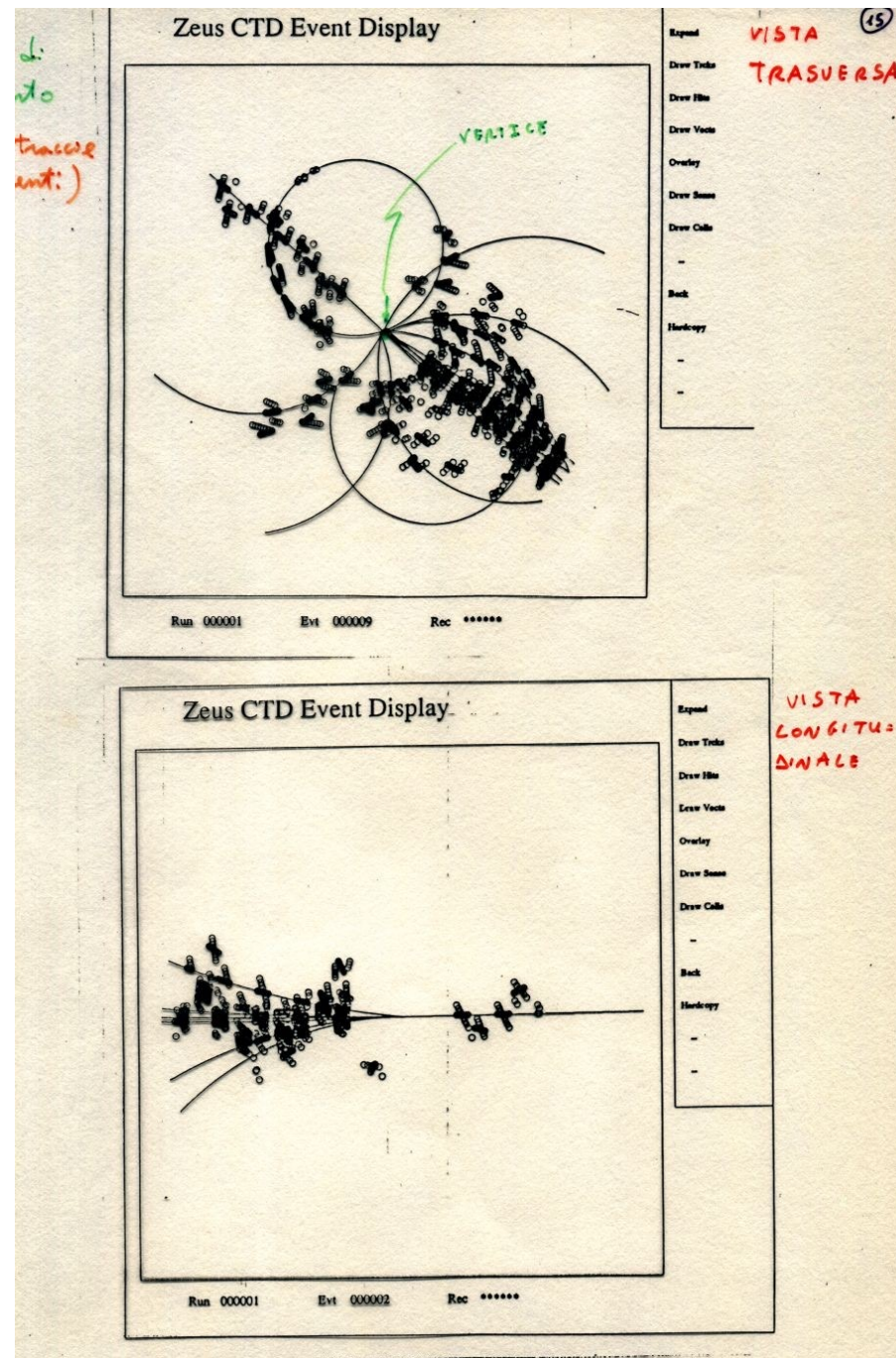
Fig. 2.1 Section of the ZEUS detector along the beam.

elettroni (30 GeV/c) protoni (120 GeV/c)

Tracciatore centrale di ZEUS



Esempio di tracce



Problematica in HEP (*relativamente alle tracce singole*).

- A. Riconoscimento delle tracce (definizione del set di misure relative alla singola traccia)
- B. Ricostruzione delle tracce (parametri che le definiscono, es. x, y, z, p_x, p_y, p_z). La conoscenza del “vettore di stato” è richiesta al vertice e all'estremo del rivelatore (*compatibilità con i rivelatori esterni*)
- C. Riconoscimento del vertice (*primario*) e di eventuali vertici secondari (*da decadimento delle particelle primarie*)

A \Rightarrow Pattern Recognition

B \Rightarrow Fit delle tracce

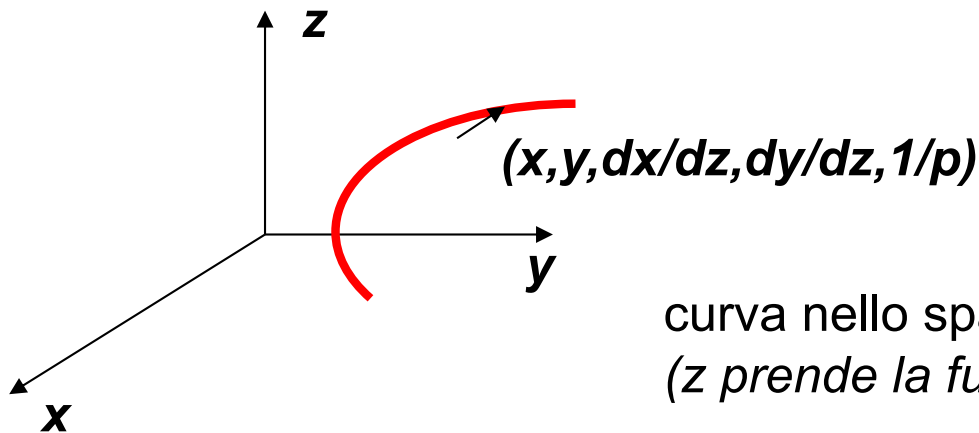
C \Rightarrow Fit del Vertice

Il FIT delle tracce tratta della STIMA dei parametri delle tracce.

Il FILTRO tratta dell'analisi (*lineare*) dei sistemi dinamici stocastici.

equivalenza se:

Si identifica il vettore di stato del sistema dinamico con un vettore x a 5 parametri, che descrive la traccia in ciascun punto della sua traiettoria spaziale.



curva nello spazio, parametrizzato in z
(z prende la funzione del tempo, o dell'indice k)

Filtro di KALMAN (espressione esplicita)

Equazione del sistema: $\mathbf{x}_k = \Phi_{k,k-1} \cdot \mathbf{x}_{k-1} + \omega_k$

Equazione della misura: $\mathbf{y}_k = \mathbf{H}_k \cdot \mathbf{x}_k + \varepsilon_k$

PREDIZIONE

$$\hat{\mathbf{x}}_k(P) = \Phi_{k,k-1} \cdot \hat{\mathbf{x}}_{k-1}$$

$$P_k(P) = \Phi_{k,k-1} \cdot P_{k-1} \cdot \Phi_{k,k-1}^T + Q_k$$

$$r_k(P) = y_k - H_k \cdot x_k(P)$$

$$R_k(P) = V_k + H_k \cdot P_k(P) \cdot H_k^T$$

definizioni :

$$Q_k = cov\{\omega_k\}$$

$$V_k \equiv G_k^{-1} = cov\{\varepsilon_k\}$$

$$P_k = cov\{x_k\}$$

residui :

$$r_k = y_k - H_k \cdot x_k$$

$$R_k = cov\{r_k\}$$

Le equazioni si trovano calcolando $E[\mathbf{x}_k]$, cioè la media, e poi lo scarto quadratico $E[(E[\mathbf{x}_k] - \mathbf{x}_k)^2]$

Notiamo che $P_k(p) = \Phi_{k,k-1} \cdot P_{k-1} \cdot \Phi_{k,k-1}^T$

corrisponde alla classica legge

di propagazione degli errori: $\sigma_{x_i x_j} = \sum_{lm} \frac{\partial x_i}{\partial x_l} \frac{\partial x_j}{\partial x_m} \sigma_{x_l x_m}$ con $\Phi_{ij} = \frac{\partial x_i}{\partial x_j}$

FILTRO
$$\mathbf{x}_k(F) = \mathbf{x}_k(P) + \mathbf{K}_k \cdot [\mathbf{y}_k - \mathbf{H}_k \cdot \mathbf{x}_k(P)]$$

dove \mathbf{K}_k è la matrice di **GUADAGNO** di KALMAN:

$$\begin{aligned} \mathbf{K}_k &= \mathbf{P}_k(P) \cdot \mathbf{H}_k^T \cdot \left(\mathbf{V}_k + \mathbf{H}_k \cdot \mathbf{P}_k(P) \cdot \mathbf{H}_k^T \right)^{-1} \\ &= \mathbf{P}_k(P) \cdot \mathbf{H}_k^T \cdot \mathbf{G}_k \end{aligned}$$

$$\begin{aligned} \mathbf{P}_k(F) &= (\mathbf{I} - \mathbf{K}_k \cdot \mathbf{H}_k) \cdot \mathbf{P}_{k-1}(P) \\ \mathbf{r}_k(F) &= \mathbf{y}_k - \mathbf{H}_k \cdot \mathbf{x}_k(F) = (\mathbf{I} - \mathbf{H}_k \cdot \mathbf{K}_k) \cdot \mathbf{r}_k(P) \\ \mathbf{R}_k(F) &= (\mathbf{I} - \mathbf{H}_k \cdot \mathbf{K}_k) \cdot \mathbf{V}_k = \mathbf{V}_k - \mathbf{H}_k \cdot \mathbf{P}_k(F) \cdot \mathbf{H}_k^T \end{aligned}$$

Incremento del χ^2 : $\chi_+^2 = \mathbf{r}_k^T(F) \cdot \mathbf{R}_k^{-1}(F) \cdot \mathbf{r}_k(F)$

Espressione equivalente per la covarianza $\mathbf{P}_k(F)$: $\mathbf{P}_k(F) = \left[\mathbf{P}_k^{-1}(P) + \mathbf{H}_k^T \cdot \mathbf{G}_k \cdot \mathbf{H}_k \right]^{-1}$

Cioè, in termini di matrice d'informazione, $\mathbf{I}_f = \mathbf{P}^{-1}$, $\mathbf{I}_f(F) = \mathbf{I}_f(P) + \mathbf{I}_f(\text{misura})$
 il FILTRO corrisponde all'aggiunta dell'informazione dovuta alla misura !

SMOOTHING *(interpolazione)*

Date l misure, $\mathbf{x}_k(\mathbf{S})$ con $k < l$, è dato da:

$$\mathbf{x}_k(\mathbf{S}) = \mathbf{x}_k(\mathbf{F}) + \mathbf{A}_k \cdot (\mathbf{x}_{k+1}(\mathbf{S}) - \mathbf{x}_{k+1}(\mathbf{P}))$$

\mathbf{A}_k : **smoother gain** matrix:

$$\mathbf{A}_k = \mathbf{P}_k(\mathbf{F}) \cdot \Phi_{k+1,k}^T \cdot \mathbf{P}_{k+1}^{-1}(\mathbf{P})$$

Vale:

$$\begin{aligned} \mathbf{P}_k(\mathbf{S}) &= \mathbf{P}_k(\mathbf{F}) + \mathbf{A}_k \cdot [\mathbf{P}_{k+1}(\mathbf{S}) - \mathbf{P}_{k+1}(\mathbf{P})] \mathbf{A}_k^T \\ \mathbf{r}_k(\mathbf{S}) &= \mathbf{r}_k(\mathbf{F}) - \mathbf{H}_k \cdot (\mathbf{x}_k(\mathbf{S}) - \mathbf{x}_k(\mathbf{F})) = \mathbf{y}_k - \mathbf{H}_k \cdot \mathbf{x}_k(\mathbf{S}) \\ \mathbf{R}_k(\mathbf{S}) &= \mathbf{R}_k(\mathbf{F}) - \mathbf{H}_k \cdot \mathbf{A}_k \cdot [\mathbf{P}_{k+1}(\mathbf{S}) - \mathbf{P}_{k+1}(\mathbf{P})] \cdot \mathbf{A}_k^T \cdot \mathbf{H}_k^T \end{aligned}$$

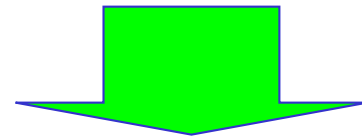
Nota: la matrice di covarianza interpolata $\mathbf{P}_k(\mathbf{S})$ è “più piccola”, cioè contiene “più informazione”

della matrice di covarianza filtrata $\mathbf{P}_k(\mathbf{F})$
(informazione di **tutte** le misure)

In generale:

$$\mathbf{x}(z_k) \equiv \mathbf{x}_k = \mathbf{f}(\mathbf{x}_{k-1}) + \mathbf{g}(\mathbf{x}_{k-1}) \cdot \boldsymbol{\omega}_k \quad \text{evoluzione del vettore di stato da } \mathbf{k-1} \text{ a } \mathbf{k}$$

$$\mathbf{y}_k = \mathbf{h}_k(\mathbf{x}_k) + \boldsymbol{\varepsilon}_k \quad \text{misure in } \mathbf{k}$$



LINEARITA'

$$\mathbf{x}_k = \boldsymbol{\Phi}_{k,k-1} \cdot \mathbf{x}_{k-1} + \boldsymbol{\omega}_k \quad \text{rumore bianco}$$

$$\mathbf{y}_k = \mathbf{H}_k \cdot \mathbf{x}_k + \boldsymbol{\varepsilon}_k \quad \text{rumore bianco}$$

Note:

- essendoci campo magnetico, il sistema è a priori NON-lineare
- $\boldsymbol{\varepsilon}_k$ è l'errore di misura, solitamente gaussiano
- $\boldsymbol{\omega}_k$ è il disturbo casuale, NON-lineare ! (esso è dovuto alla diffusione multipla subita dalla particella carica attraversando il mezzo gassoso, secondo lo scattering elastico di Coulomb)
- la NON-linearità aumenta se, inoltre, la particella perde energia (varia l'impulso) attraversando il mezzo, essendo in prima approssimazione:

$$\frac{dE}{dx} \propto \frac{Z}{A} \cdot \rho_{\text{densità}}$$

Scattering Multiplo

Per un attraversamento L , lo spread in posizione è: $\sigma(x) = \frac{L}{\sqrt{3}} \cdot \sigma(\theta)$

con spread in angolo: $\sigma(\theta) = c \cdot \sqrt{L} \cdot \frac{1}{p}$
 p ← impulso

La costante c dipende dal mezzo (*lunghezza di radiazione*): $c \propto \frac{1}{\sqrt{\chi_R}}$

con $\chi_R \approx 180 \cdot \frac{A}{Z^2} \text{ (gr} \cdot \text{cm}^{-2}\text{)}$

Per esempio: idrogeno $\chi_R = 61.28 \text{ gr} \cdot \text{cm}^{-2}$

ferro $\chi_R = 13.84 \text{ gr} \cdot \text{cm}^{-2}$

} conclusioni ⇒

Approssimazione LINEARE accettabile per i rivelatori gassosi ...
e per i rivelatori contenenti mezzi "pesanti" ??

Filtro di KALMAN **ESTESO**

Si dimostra che il filtro di KALMAN **lineare**, se applicato a sistemi **NON-lineari**, sviluppati secondo Taylor al 1^o ordine (*linearizzazione*) attorno ad una traiettoria di riferimento, comporta le medesime equazioni, eccetto:

$$\text{(vettore di stato nel punto } \mathbf{P}) \quad \mathbf{x}_k(\mathbf{P}) = \mathbf{f}_{k,k-1}(\mathbf{x}_{k-1})$$

$$\text{con } \Phi_{k,k-1} = \frac{\partial \mathbf{f}_k}{\partial \mathbf{x}_k}$$

Come traiettoria di riferimento si prende ovviamente la traiettoria FILTRATA, oppure quella INTERPOLATA (*in quest'ultimo caso si procede per iterazioni successive*)

APPLICAZIONI

L'algoritmo combinato FILTER-SMOOTHER permette la **STIMA OTTIMALE** dei parametri della traccia in ogni punto di quest'ultima, utilizzando tutta l'**INFORMAZIONE** disponibile.

- E' possibile fare predizioni ottimali in altri rivelatori, per entrambi gli estremi della traccia
- E' possibile combinare segmenti di traccia
- La rivelazione degli **outliers** può essere fatta localmente, utilizzando l'incremento del χ^2
- L'eliminazione di misure "false" è fatta localmente

FIT DEL VERTICE

Utilizzando il vettore di stato $x_k = \begin{pmatrix} x \\ y \\ z \\ p_x \\ p_y \\ p_z \end{pmatrix}$

con l'equazione di sistema $x_k = x_{k-1}$

(non c'è disturbo casuale !)

si può implementare direttamente il filtro di KALMAN!

Rivelatore di μ in ZEUS

8 piani di misura per il calcolo della sagitta delle tracce μ



Metodo Globale (di Wiener) $\chi^2 = \sum_{k=1}^8 (y_k - H_k \cdot x_k) \cdot W \cdot (y_k - H_k \cdot x_k)^T$

matrice dei pesi

Ricerca del χ^2 minimo



ITERAZIONE

(dE/dx : ignorato !!)

$$W = [V_k + Q_k^{M.S.}]^{-1}$$


matrice di covarianza degli errori di misura

matrice di covarianza del Multiple Scattering

Il filtro di Kalman si può applicare (ad esempio !) sui 2 segmenti di traccia, esterni.

⇒ una operazione di FILTERING ed una di SMOOTHING

ma ... attenzione ai valori iniziali della matrice di covarianza
(cioè $\sigma^2_{1/p}$ che è a priori sconosciuto)

Il filtro NON è indipendente dal valore iniziale di $1/p$!  Una stima alternativa dell'impulso (es. calcolo diretto della curvatura) si può aggiungere come informazione indipendente !

Notiamo che la **PREDIZIONE** al 2^o ordine nello sviluppo di Taylor è:

$$\mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_k) + \frac{1}{2} \cdot \mathbf{P}_k \cdot \frac{\partial^2 \mathbf{f}}{\partial \mathbf{x}^2}(\mathbf{x}_k)$$

⇒ equazioni accoppiate per la media e la varianza !

$$\mathbf{P}_{k+1} = \Phi_k \cdot \mathbf{P}_k \cdot \Phi_k^T - \frac{1}{4} \cdot \mathbf{P}_k^2 \cdot \left[\frac{\partial^2 \mathbf{f}}{\partial \mathbf{x}^2} \right]^2 + \mathbf{Q}_{k+1}$$

GUESS: è possibile guadagnarci ??

SI', aiutandosi con le simulazioni Monte Carlo e la verifica sui dati !!

Risultato originale per ZEUS in
 L. Stanco
 Comp.Phys.Comm., 57 (1989), 380.

