

最可几法在带电粒子径迹重建中的应用*

陈廷杨

(南京大学物理系 南京 210008)

1992年12月28日收到

摘 要

本文介绍了用最可几法拟合具有非高斯型测量误差的空间径迹,与常用的最小二乘法拟合相比较,分辨率大大提高,因之也提高了识别真假径迹的可靠性。

关键词 最可几法,最小二乘法,直接搜索法。

1 引 言

在高能物理实验的离线分析中,带电粒子的径迹重建是一个很关键的问题。径迹重建要解决两个问题,首先是从一系列测量得到的空间点中找出径迹并定出径迹的参数如截距、斜率或方向余弦等,其次是给出每条径迹的品质因素。在径迹重建中,通常用最小二乘法来对可能的径迹进行拟合,从中给出径迹的参数和品质因素 χ^2 。由于在实验数据的采集过程中探测器的效率不可能为100%,以及噪声和相互感应等的存在,因此在径迹重建中常常会出现一些假径迹,如何除去这些假径迹而又不牺牲真径迹是一个重要的问题。

如果有 M 个测量点 $x_i(i=1,2,\dots,M)$,从中要拟合出与 $N(<M)$ 个线性参数 $b_n(n=1,2,\dots,N)$ 有关的函数 $\bar{x}(z,b)$ ^[1],

$$\bar{x}(z,b) = \sum_{n=1}^N b_n f_n(z), \quad (1)$$

这里 z 为自变量; $b=(b_1,b_2,\dots,b_N)$ 为一 N 维矢量; $f_n(n=1,2,\dots,N)$ 为 N 个已知的函数。如果 $x_i(i=1,2,\dots,M)$ 是相互独立的,最小二乘法由下式来决定 b_n :

$$\chi^2 = \sum_{i=1}^M \left[x_i - \sum_{n=1}^N b_n f_n(z_i) \right]^2 / \sigma_i^2 = \min, \quad (2)$$

这里的 σ_i 是 x_i 的标准偏差。严格说来,用最小二乘法对径迹进行拟合时要求测量误差必须是高斯型^[1],在高能物理实验中,有些探测器的测量误差分布却并非高斯型,如硅条探测器,多丝正比室以及垫块室(Pad chamber)等等,它们的误差分布近似矩形(见图1),

* 国家自然科学基金资助。

在这种情况下虽然用最小二乘法仍能得出 b_n 与 χ^2 , 但从式(2)看出这时得到的只是一种近似, 如果探测器的误差为图 1 形式的分布,

(2)式中 x_i 为第 i 条丝(多丝室)的位置或第 i 条(硅条探测器)的中心位置, s_i 为半丝距。实际上粒子在 $x_i - s_i$ 到 $x_i + s_i$ 之间任何一点上通过的几率是相同的, 但从式(2)可以

看出 χ^2 随 x_i 与待拟合曲线 $\sum_{n=1}^N b_n f_n(z_i)$ 的

差距成平方增加, 即使其差距小于 s_i 也是如此, 这与实际情况(即误差分布近似于矩形分布)有很大出入。在最小二乘法中, 拟合的结果是迫使被拟合的径迹向 x_i 靠近, 而实际分布几率在 $(x_i - s_i, x_i + s_i)$ 之间是均匀的。

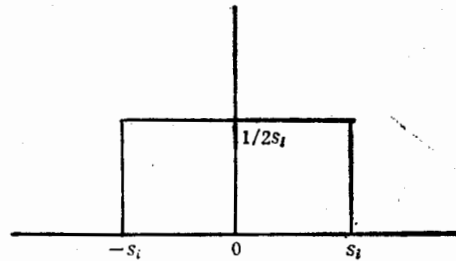


图 1. 多丝正比室或硅条探测器的理想误差分布

从理论上讲, 最重要而又最普遍的方法应是最可几法^[2]。假设有一个带电粒子通过探测器, 为了简单起见, 首先考虑粒子径迹在 $x-z$ 平面上的投影, 该径迹由(1)式定义, 用 $b = (b_1, b_2, \dots, b_N)$ 表示 N 维参数。设径迹在第 i 个探测器上产生讯号 x_i 的几率密度为

$$p_i(x_i, z_i, b),$$

这里 z_i 为该探测器的 z 坐标。如果有 M 个探测器, 则粒子产生 $x_i (i = 1, 2, \dots, M)$ 讯号的总几率为

$$L(b) = \prod_{i=1}^M p_i(x_i, z_i, b), \quad (3)$$

如果能找到一组 b , 使 $L(b)$ 达到最大值, 这就是最可几法。很显然如果所有的 $p_i(x_i, z_i, b)$ 都是高斯型:

$$P_i(x_i, z_i, b) = \frac{1}{\sqrt{2\pi}\sigma_i} \exp \left[-\frac{1}{2} \left(\frac{x_i - \bar{x}(z_i, b)}{\sigma_i} \right)^2 \right], \quad (4)$$

把它代入(3)式中, 很容易看出最可几法与最小二乘法一致:

$$L(b) = (2\pi)^{-\frac{M}{2}} \frac{1}{\prod_{i=1}^M \sigma_i} \exp \left[-\frac{1}{2} \sum_{i=1}^M \frac{1}{\sigma_i^2} (x_i - \bar{x}(z_i, b))^2 \right],$$

这里 σ_i 为 x_i 的标准偏差。要使 $L(b)$ 达到最大值的充要条件是

$$\sum_{i=1}^M \frac{1}{\sigma_i^2} (x_i - \bar{x}(z_i, b))^2 = \min.$$

2 最可几法在径迹拟合中的应用

要求得(3)式的最大值一般来说是很复杂的, 如果每一 $P_i(x_i, z_i, b)$ 都大于零, 则可对(3)式取对数得

$$l(b) = \ln L(b), \quad (5)$$

假设每个 $P_i(x_i, z_i, b)$ 都是可微的, 则可用通常的求极值方法来求解:

$$\frac{\partial l(b)}{\partial b_n} = \frac{1}{L(b)} \frac{\partial L(b)}{\partial b_n} = S_n = 0, \quad n = 1, 2, \dots, N. \quad (6)$$

图1中硅条探测器和多丝室的测量误差从理论上讲应为

$$P_i(x_i, z_i, b) = \begin{cases} \frac{1}{2s_i}, & |x_i - \bar{x}(z_i, b)| \leq s_i \\ 0 & |x_i - \bar{x}(z_i, b)| > s_i \end{cases}, \quad (7)$$

这里 $\bar{x}(z, b)$ 为待拟合径迹的方程, 在这种情况下, (5)式的对数在

$$x_i = \bar{x}(z_i, b) \pm s_i \quad (8)$$

时无意义。在实际情况下 $P_i(x_i, z_i, b)$ 在 $|x_i - \bar{x}(z_i, b)| > s_i$ 时并不为零, 这是因为径迹与探测器往往会有一定的交角, 因此在相邻近区间会产生讯号, 这些讯号有一定的几率被采集到数据中, 另外由于多次散射等原因, 使 $P_i(x_i, z_i, b)$ 在 $\bar{x}(z_i, b)$ 满足(8)式时并不为零而有一定的延伸。用下式^[3]作为近似可以消除(5)式中的奇点:

$$P_i(x_i, z_i, b) = \begin{cases} \frac{A_i}{2s_i}, & |x_i - \bar{x}(z_i, b)| \leq s_i \\ \frac{A_i}{2s_i} \exp\left[-\frac{1}{2} \left(\frac{x_i - \bar{x}(z_i, b)}{\sigma_i}\right)^2\right], & |x_i - \bar{x}(z_i, b)| > s_i \end{cases} \quad (9)$$

如图2实线所示, 这里 A_i 是归一化常数, σ_i 为其标准偏差。作了这种替代后, (5)式可改写成

$$l(b) = \text{const} - \frac{1}{2} F(b),$$

$$F(b) = \sum_{i=1}^M \frac{1}{\sigma_i^2} [x_i - \bar{x}(z_i, b)]^2, \quad (10)$$

这里 Σ' 表示求和是对高斯型误差项进行的, 如果是(9)式的分布, 则只对 $|x_i - \bar{x}(z_i, b)| > s_i$ 进行, 这时最可几法等价于

$$\chi^2 = F(b) = \sum_{i=1}^M \frac{1}{\sigma_i^2} [x_i - \bar{x}(z_i, b)]^2 = \min. \quad (11)$$

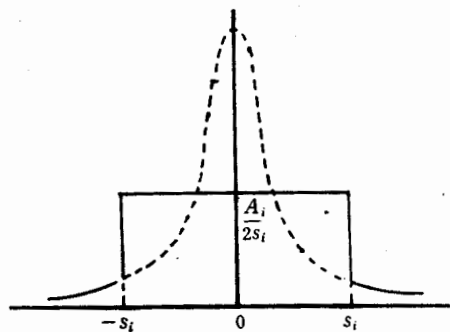


图2 矩形误差分布的理想外推(实线部分)

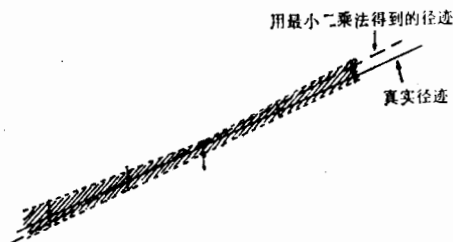


图3 当所有误差都是图1形式分布时最可几法的不确定性(打斜线区域)

从表面上看 χ^2 与最小二乘法中的 χ^2 相似, 实际上则不一样, 在最小二乘法中, 只要 $M \geq N$, $b = (b_1, b_2, \dots, b_N)$ 一般只有唯一的解, 但(11)式则不同, 如果所有点都落在(8)式定义的范围(如图 3 所示), 则 χ^2 在打斜线区间内均相同, 因此有无穷多解, 也就是说在实际情况下, 它们出现的几率是相同的, 因此由图 3 可见, 要使用最可几法必须有“足够”多的丝(或条), 或者加上其它的约束条件, 如径迹必须通过反应点等等。

3 直接搜索法

在一般情况下(11)式的极小值不是用解析方法能求得的, 因为它的导数形式很复杂。下面介绍如何用直接搜索法求(11)式的解。

最简单的直接搜索法为网格法。为简单起见, 假定 $F(b)$ 为二维矢量 (b_1, b_2) 的函数, 也就是说 $N = 2$, 例如直线情况:

$$\bar{x}(z, b) = b_1 z + b_2. \quad (12)$$

如果 b_1 与 b_2 的变化范围已知, 则把 b_1 与 b_2 的变化区间内用 K 个分点各分成 $K + 1$ 个小区间 (b_1 与 b_2 可以分成不同个区间) 这样就有 $(K + 1) \times (K + 1)$ 个网格, 计算 $F(b)$ 在每个网格中心点上的值, 可以得出一个最小值, 相对于这个最小值的 (b_1, b_2) 值就是(11)式的解。但在实际情况中, b_1 与 b_2 的变化范围并不知道, 而只知道各测量值 x_i 的变化范围, 例如该径迹通过的起始点 x_1 和终止点 x_2 的变化范围, 因此如果把(12)式改写成:

$$\begin{aligned} \bar{x} = F(x_1, x_2, z) &= \frac{x_1 - x_2}{z_1 - z_2} z + \frac{z_1 x_2 - z_2 x_1}{z_1 - z_2}, \\ b_1 &= \frac{x_1 - x_2}{z_1 - z_2}, \\ b_2 &= \frac{z_1 x_2 - z_2 x_1}{z_1 - z_2}, \end{aligned} \quad (13)$$

从(7)式知道它们可能的变化范围为:

$$x_{1,2} - s_{1,2} \leq x < x_{1,2} + s_{1,2}.$$

如果在 x_1 与 x_2 的变化范围内分别插入 K 个分点分成 $K + 1$ 个小区间, 每个小区间长度为

$$d_1 = \frac{2s_1}{K + 1}, \quad d_2 = \frac{2s_2}{K + 1}.$$

这样有 $(K + 1) \times (K + 1)$ 个网格, 取网格中心处的 $F(x_1, x_2, z)$ 为 F 的值, 在 $(K + 1) \times (K + 1)$ 个 F 中求出一个最小值, 其对应的 x_1 与 x_2 就是所要求的值, 把它们代入(13)式可得到相应的 b_1 和 b_2 。很显然, 这种网格搜索法可以应用到任何几何形状的径迹。

作为一个应用的例子, 将上述方法试用于 Fermilab 的 E771 实验中, E771 实验是用 800 GeV/c 的质子来打在硅靶上寻找 B 介子, 在这种能量下产生的 B 介子平均衰变距离为 9mm^[4], 为了探测 B 的衰变顶点, 采用硅条探测器作为顶点探测器, 其几何排列如图 4 所示^[4], 其中 L 型为中间 384 条中心间距为 25 μ m 而两边各 152 条中心间距为 50 μ m 的硅条探测器, K 型为中间 384 条中心间距为 50 μ m 而两边各 154 条中心间距为 100 μ m

的硅条探测器,在1991年到1992年初的实验中安装了12个 x 、 y 、 u 和 v 方向的硅条探测器,其中 x 和 y 方向各5个, u 和 v 方向各1个, u 与 x 的夹角 θ 为 45° , v 与 x 方向的夹角 θ 为 -45° 。 (u,v) 与 (x,y) 的关系为^[5]

$$\begin{aligned} u &= x \cos \theta + y \sin \theta; \\ v &= x \cos \theta - y \sin \theta. \end{aligned} \quad (14)$$

这里 θ 是夹角的绝对值。当一带电粒子穿过谱仪在这些探测器中给出讯号:

$$x_i (i=1, 2, \dots, n_x), y_j (j=1, 2, \dots, n_y),$$

$$u_k (k=1, 2, \dots, n_u), v_l (l=1, 2, \dots, n_v),$$

这里 n_x 、 n_y 、 n_u 和 n_v 分别为 x 、 y 、 u 和 v

方向探测器上的讯号数。在我们的情况下 n_u 与 n_v 均小于等于1。设一空间径迹在 $x-z$ 和 $y-z$ 平面上投影的斜率和截距各为 α_x 、 β_x 和 α_y 、 β_y ,则相应于(11)式的 χ^2 为:

$$\begin{aligned} \chi^2 &= \sum_{i=1}^{n_x} \frac{1}{\sigma_{x_i}^2} (x_i - \alpha_x z_i - \beta_x)^2 + \sum_{j=1}^{n_y} \frac{1}{\sigma_{y_j}^2} (y_j - \alpha_y z_j - \beta_y)^2 \\ &+ \sum_{k=1}^{n_u} \frac{1}{\sigma_{u_k}^2} [u_k - (\alpha_x \cos \theta_k + \alpha_y \sin \theta_k) z_k - (\beta_x \cos \theta_k + \beta_y \sin \theta_k)]^2 \\ &+ \sum_{l=1}^{n_v} \frac{1}{\sigma_{v_l}^2} [v_l - (\alpha_x \cos \theta_l - \alpha_y \sin \theta_l) z_l - (\beta_x \cos \theta_l - \beta_y \sin \theta_l)]^2, \end{aligned} \quad (15)$$

这里 z_i 、 z_j 、 z_k 和 z_l 分别为 x_i 、 y_j 、 u_k 和 v_l 点的 z 方向坐标, θ_k 和 θ_l 各为 u_k 和 v_l 倾斜角的绝对值,它们可以不一样, σ_{x_i} 、 σ_{y_j} 、 σ_{u_k} 和 σ_{v_l} 分别为 x_i 、 y_j 、 u_k 和 v_l 的标准偏差。 Σ' 与(11)式有相同的定义,在上述的情况下直接搜索法与一维情况下相同,取两对种子点(Seed points) (x_1, x_2) 与 (y_1, y_2) , 并作如下的变换:

$$\begin{aligned} \alpha_x &= \frac{x_1 - x_2}{z_{x_1} - z_{x_2}}; \\ \beta_x &= \frac{z_{x_1} x_2 - z_{x_2} x_1}{z_{x_1} - z_{x_2}}; \\ \alpha_y &= \frac{y_1 - y_2}{z_{y_1} - z_{y_2}}; \\ \beta_y &= \frac{z_{y_1} y_2 - z_{y_2} y_1}{z_{y_1} - z_{y_2}}, \end{aligned} \quad (16)$$

这样(15)式中的 χ^2 就变换为 x_1 、 x_2 、 y_1 与 y_2 的函数,

$$\chi^2 = F(x_1, x_2, y_1, y_2, z) = \min_0 \quad (17)$$

这里种子点的选取应使两点的距离尽量远,因为 z 坐标的测量精度远高于 x 的精度,因此由(16)式可得:

$$(\delta \alpha_x)^2 = \frac{(\delta x_1)^2 + (\delta x_2)^2}{(z_{x_1} - z_{x_2})^2} = \frac{\sigma_{x_1}^2 + \sigma_{x_2}^2}{(z_{x_1} - z_{x_2})^2},$$

σ_{x_1} 与 σ_{x_2} 变化不大,因此 α_x 的误差随 $(z_{x_1} - z_{x_2})^2$ 的增加而减小。因为 (x_1, x_2) 与

(y_1, y_2) 的变化范围已知, 因此很容易对(16)式作网格搜索, 其方法如下: 在 x_1 的允许变化范围内 $(x_1 - s_1, x_1 + s_1)$ 插入 K 个分点, 因此把 $(x_1 - s_1, x_1 + s_1)$ 分成 $K + 1$ 个小区间, 取每个小区间的中点 $x_{1,i} (i = 1, 2, \dots, K + 1)$, 对 x_2, y_1 与 y_2 用同样方法进行分割, 它们的分点数 K 可以各不相同, 通常取相同的 K , 即相同的分点数, 这样每个取样点的误差为

$$\frac{1}{2} d_{1,2} = \frac{s_{1,2}}{K + 1}.$$

从上式看似乎 K 越大越好, 但是当 $\frac{1}{2} d_{1,2}$ 和 z 位置的误差与不同探测器之间 x 位置准直误差数量级相同时, 再增加 K 就不会改善径迹的分辨率. 用上述的网格搜索法可求出一组最佳的 (x_1, x_2, y_1, y_2) 值, 由(16)式可得到 $\alpha_x, \beta_x, \alpha_y$ 和 β_y .

径迹拟合往往与径迹寻找结合在一起, 因为在径迹寻找过程中往往有很多似是而非的径迹, 在这里用 χ^2 代替最小二乘法作为径迹好坏的标准 χ^2 .

为了对上述方法进行检验, 用 PYTHIA 程序产生了 100 个 $B\bar{B}$ 事例, 在事例产生中设探测器的效率为 96% (实际测得的效率为 95% 到 98%^[4]), 噪声水平设平均每个探测器每个事例一个噪声^[4]. 这些反应产生的带电粒子在硅条探测器中产生电离、多次散射等效应, 加上上述的噪声和效率, 在各探测器中得到了空间点分布. 分别用最小二乘法和最可几法来重建径迹, 其差别可在图 5 与 6 中看出(图中纵坐标 N 为事例数), 图 5(a) 与 6(a) 是用最小二乘法求得的径迹与原径迹之间的偏差分布, 图 5(b) 与 6(b) 为用最可几法所得到的结果, 这里取 $K = 4$ 即 x 方向 5×5 , y 方向 5×5 搜索, 为了节省计算机的 CPU 时间, 对 x 方向先进行搜索, 这样在 x 方向进行了 25 次运算, 然后在固定 α_x 与 β_x 后再对(15)式在 y 方向上进行搜索, CPU 时间主要化在 y 方向的搜索上, 因对每一条可能的 $x-z$ 线都要进行搜索, 计算时间似乎要增加 25 倍, 但由于在径迹寻找过程中, 当两个种子点选定后, 由(13)式, 一条尝试性的径迹就建立起来了, 下面的任务就是在这条尝

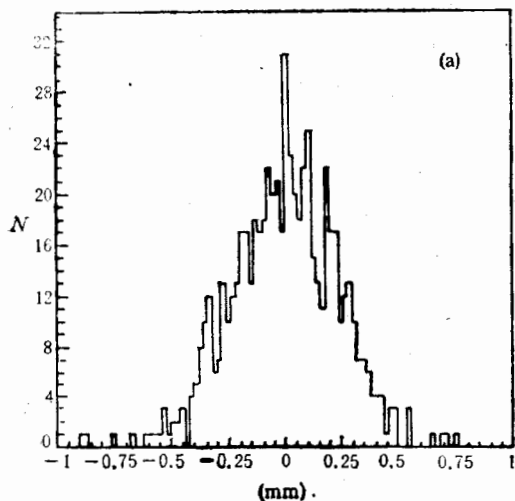


图 5(a) 用最小二乘法拟合得到的 y 截距偏差

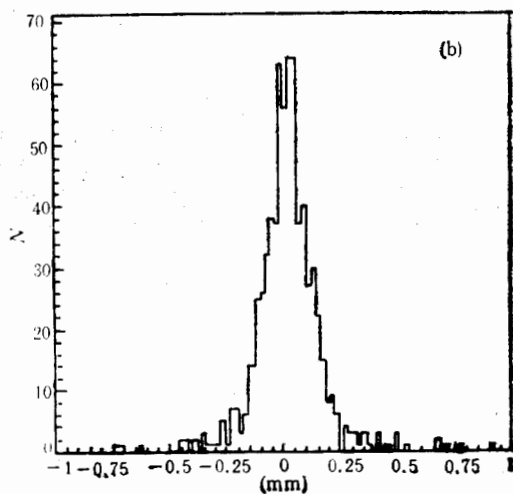
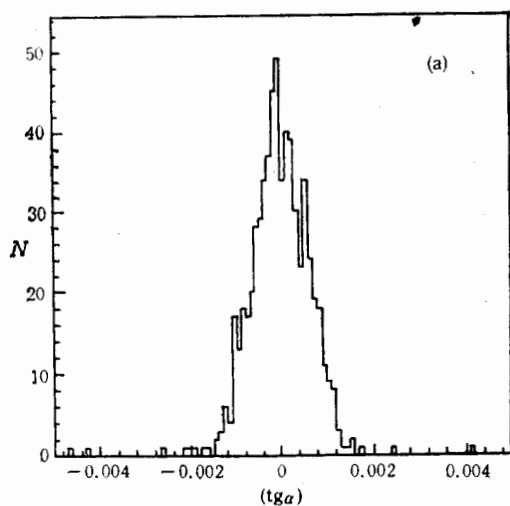
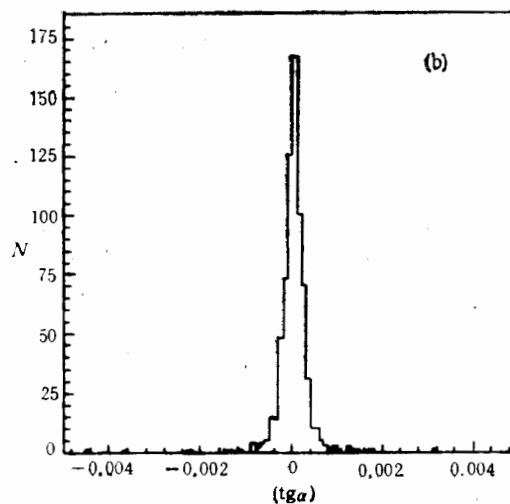


图 5(b) 用最可几法拟合得到的 y 截距偏差

图 6(a) 用最小二乘法拟合得到的 γ 斜率偏差图 6(b) 用最可几法拟合得到的 γ 斜率偏差

试径迹上除种子点外在其它探测器上寻找可能的测量到的点, 由于尝试线的误差和测量点的误差, 因此在尝试线两边建立起一条寻点道 (searching road) 如图 3 打斜线处所示, 用最可几法后, 寻点道由于种子点的误差缩小为最小二乘法的 $\frac{1}{K+1}$, 这样减少了很多不必要的搜索过程, 因此 CPU 时间比最小二乘法增加约 10 倍。即使这样也必须在计算机能力足够的情况下才能应用此法。另外在 E771 实验中用了一个过滤程序, 先把原始数据浓缩, 这样可防止把大量的 CPU 时间用于不必要的搜索过程。

作者感谢美国费米实验室 E771 协作组的同事们的有益讨论。数据的产生用了协作组中加拿大 McGill 大学物理系 John Trischuk 教授等发展起来的蒙特卡罗程序。

参 考 文 献

- [1] Lars Bugge and Jan Myrheim, *Nucl. Instr. and Meth.*, **179**(1981) 365.
- [2] Particle Data Group, *Phys. Rev.*, **D45**(1992)III.36.
- [3] T.Y. Chen and N.G. Yao, *Nucl. Instr. and Meth.*, **A329**(1993)479.
- [4] A.McManus et al, "Performance of a Silicon Microstrip Detector in a Very High Rate Environment", 1991 Nuclear Science Symposium Proceedings.
- [5] T.Y. Chen and N.G. Yao, *Nucl. Instr. and Meth.*, **A290**(1990)390.

Application of maximum likelihood method in reconstruction of charged particle tracks

Chen Tingyang

(Department of Physics, Nanjing University, Nanjing 210008)

Received on December 28, 1992.

Abstract

By use of the maximum likelihood, a new method to fit space tracks with non-Gaussian measurement errors is introduced. Compared with the conventional least squares fit, the resolution has been greatly improved and so does the reliability of distinguishing the false tracks.

Key words Method of maximum likelihood, Method of least squares, Direct search method.