KINEMATICAL ANALYSIS OF BUBBLE CHAMBER PICTURES

B. Ronne,
Track Chamber Division, CERN

I. INTRODUCTION

There are many different systems used in the treatment of bubble chamber pictures. Every big laboratory has its own system. In all of them the scanning and measurement has to be followed by geometrical reconstruction and kinematical calculations. Big computers are always used in the last two steps.

We will not try to make any comparison between different systems but only describe in some detail how the analysis is performed at CERN, especially by the group working with a heavy liquid chamber as detector.

Many examples in later Chapters will be taken from one special experiment concerning the properties of negative cascade-hyperons, and we will therefore concentrate somewhat on describing this experiment.

II. TREATMENT OF BUBBLE CHAMBER PICTURES

The three corresponding rolls with different views of the same events are examined visually and all events fulfilling special criteria are written down. The criteria of Ξ^- are mainly:

1) that a Λ (or V^0) comes from a kink of a negative track,
2) that the transverse momentum of the decay particles are about the same and oppositely directed,
3) that the transverse momentum is not obviously much greater than what is possible. The maximum transverse momentum ($\sim$ 140 MeV/c) is equal to the momentum of the particles in the Ξ^- rest system.

An event fulfilling the above criteria and where further a K^+ or a K^0 is associated with the production point of the Ξ^-, is noted as a signed Ξ^-. Those events without the K signature are noted as unsigned. In Fig. 4 we see one event which fulfils the criteria of signed Ξ^-.

Photographs are taken of all interesting events and short descriptions of the events are given on scanning sheets.

Before an event can be measured it has to be carefully prepared. The preparation of events is needed because the treatment of measured data is performed by a computer, and this must know for instance which tracks stop and which belong to a special interaction.

In our case the preparation includes labelling of vertices, tracks, stopping points, decay points and intermediate points of tracks. As this experiment is performed in heavy liquid, we often have a particle which is scattered a few degrees at one point. If one assumes that the particle loses only a small amount of energy in this single scattering, it may be worth while to measure the track after the interaction point also, to receive a more accurate value for the momentum of the particle. This is especially true if the particles stops in the chamber. Multiple scattering may further have as a result, a track which looks partly straight and partly curved. The geometry programme, which is described below, can only accept a track which has the same curvature along its entire length. Other tracks have to be divided in several pieces, with the help of intermediate points and each piece is treated separately by the geometry programme.

The measurements are made on a digitized projector, which gives the output coordinates (precision a few microns on the film) on a tape. This can be treated by the following chain of programmes which are described below.

REAP → THRESH → JOIN → GRIND

III. GEOMETRICAL RECONSTRUCTION

The data from the measuring machines are first sent through a programme named REAP. Some types of measuring errors and errors in labelling of tracks and vertices are now detected. REAP further prepares the data to be in a suitable order for the following programmes.

The geometrical reconstruction of tracks is made in THRESH. A detailed description of THRESH is given by Cnops (THRESH Manual) and Moorhead (CERN 60-33). The input tape from REAP contains measured coordinates of fiducial marks and of points along the tracks and also information which makes it possible to identify the event and recognize different types of tracks and vertices. This is needed later in GRIND.

Before the reconstruction of a track can start, one has to provide THRESH with some general information needed for all events in an experiment, e.g. coordinates of cameras and fiducial marks, refractive index of the media between the cameras and the back glass of the chamber, path length in each media, tolerances for the measured coordinates of fiducial marks, tolerances on the error of a reconstructed point etc.

A reference system is fixed. The z direction is chosen to be parallel to the uniform magnetic field. The cameras are positioned around the z axis. The plane $z = 0$ is defined to be in the back of the front glass. The apparent position (x_o, y_o) in the $z = 0$ plane of the fiducial marks, as seen from the cameras, can be calculated from the general information above. The coordinates (x_m, y_m) measured on the film can further be transformed through intermediate media to the plane $z = 0$. This transformation is described by the following equations:

$$\begin{cases} x_o = \alpha_1 + \alpha_2 x_m + \alpha_3 y_m \\ y_o = \alpha_4 + \alpha_5 x_m + \alpha_6 y_m \end{cases} \qquad (1)$$

The six constants α_i can be evaluated by the method of least squares (Chapter IV.2), if at least four fiducial marks have been measured. A set of α values is calculated for each view.

The number of measured fiducial marks needed can be reduced if one includes some additional conditions, i.e. that the coordinate systems are orthogonal and that the magnification in the x and y directions is the same.

The coordinates of all measured bubbles can now be transformed to the reference plane z = 0.

A light ray, which passes from a point inside the chamber to a point on the film, may in the chamber be described by the following two equations:

$$\begin{cases} x = F_x z + G_x \\ y = F_y z + G_y \end{cases} \qquad (2)$$

where F_x, F_y, G_x and G_y can be calculated (Moorhead, CERN 60-33). The line Eq. (2) is called a reconstruction line and can be derived for each measured point.

The position inside the chamber of a bubble, measured in at least two views, is identical with the intersection point of the corresponding reconstruction lines, where the intersection point is defined to be that point which is closest to the two lines. It is then easy to find the true coordinates of such bubbles.

A track is, however, measured at about 5-10 arbitrarily chosen points, and the measured points are normally not the same for different views. At the first reconstruction of a track, one uses the following method.

The two best views, according to the following criteria, are picked out. The first view is chosen to be the one on which the track is most nearly seen as an orthogonal projection. We then calculate the angle between the tangent to the track at the initial point and the line joining the camera corresponding to the first view with each one of the other cameras. The second view is chosen to be the one for which the corresponding camera gives the greatest angle. All points in space corresponding to the measurements in the first view are reconstructed. Before one can find the points in space, one must, however, have the reconstruction lines from exactly the same points in the second view. These lines are not available, but it is possible to interpolate between reconstruction lines of near positioned points until one finds

a missing line. This will intersect the corresponding line from the first view in the true space position of the bubble.

After all points in one view have been reconstructed, one determines a first approximation to a helix, which can describe the motion of the particle inside the chamber. The helix is described by the following equations:

$$\begin{cases} x' = \rho(\cos\Theta - 1) \\ y' = \rho \sin\Theta \\ z' = \rho\,\Theta\,\mathrm{tg}\,\alpha \end{cases} \qquad (3)$$

The origin is at the starting point of the track, the z direction, as before, along the magnetic field and the y direction along the tangent of the projection of the track in the z = 0 plane. ρ is the projected curvature in the same plane. α is the dip angle of the track and Θ is the angle between the x and x′ directions.

The parameters of the helix derived above are used as starting values in a final least squares fit, where the reconstruction lines for all measurements in all views are included. The last fit is an iterative process, and the parameters converge normally (in about 97% of the tracks) to the best fit solution.

For each track THRESH gives curvature, dip angle and azimuth angle.. It is preferable to use curvature instead of momentum because the first variable is more normally distributed than the second one, and the method of least squares is based on the assumption that the variables are normally distributed.

The angles of a track are given for one or both endpoints, while the momentum is given for a meanpoint. At a later stage, when one has assumed a mass for the particle, one can make an energy-loss correction and thus find the momentum at other points of the track.

THRESH also gives errors in the derived quantities. The uncertainty of track variables is mainly due to uncertainties of the measuring device and contributions from Coulomb scattering.

The uncertainty of coordinates from the measuring machine is assumed to be known. This constant error is propagated through all the space reconstruction and in the best fit solution it gives the corresponding uncertainties of the derived quantities.

The uncertainties due to Coulomb scattering depend on the mass of the particle. THRESH is mass independent and the multiple scattering errors are therefore included at a later stage.

All parts of a track which has been measured in several pieces are treated as separate tracks in THRESH. One has to combine the information from these pieces to one track. This is done in JOIN.

IV. KINEMATICAL CALCULATIONS

1) Constraint equations

A signed Ξ^- may look as in Fig. 4.

The possible hypotheses for the interpretation of an event have to be tested by kinematic calculations, which are made in GRIND. A V^o may, for example, be due to the decay of a K^o, Λ or $\bar{\Lambda}$, or it may be a two-prong star from the interaction of a neutral particle. Often there is more than one possible origin for the neutral particle and all hypotheses have then to be tested for each origin. The V^o's in Fig. 4 may, for example, have either A or R as origin. A V^- may be due to the decay of a Ξ^-, Σ^-, K^-, π^- or μ^-, or it may be a scattering of a negative particle.

The basis of testing different hypotheses is that the four momentum and energy equations have to be fulfilled at the decay of a particle. If we have less than four unknown quantities in these equations the system is overdetermined, and the possibility to satisfy all equations give a test of how good the hypothesis is. The test is in general performed by the method of least squares, which is described below. All measured quantities are fitted to give the best solution and the probability that the hypothesis is correct is also given.

For the Λ decay we can write the four energy and momentum equations in the following way:

$$\begin{cases} f_1 = P_\Lambda \cos\lambda_\Lambda \cos\varphi_\Lambda - P_p \cos\lambda_p \cos\varphi_p - P_\pi \cos\lambda_\pi \cos\varphi_\pi = 0 \\ f_2 = P_\Lambda \cos\lambda_\Lambda \sin\varphi_\Lambda - P_p \cos\lambda_p \sin\varphi_p - P_\pi \cos\lambda_\pi \sin\varphi_\pi = 0 \\ f_3 = P_\Lambda \sin\lambda_\Lambda - P_p \sin\lambda_p - P_\pi \sin\lambda_\pi = 0 \\ f_4 = \sqrt{P_\Lambda^2 + M_\Lambda^2} - \sqrt{P_p^2 + M_p^2} - \sqrt{P_\pi^2 + M_\pi^2} = 0 \end{cases} \qquad (4)$$

where P_i, λ_i, and φ_i are the momentum, dip angle and azimuth angle of particle i.

If we assume that all three parameters of the proton and the pion are known and further that the lambda comes from a given vertex, we have only one unknown quantity, P_Λ. In this case the event is subject to three constraints, and we talk about a 3 C-fit.

We may treat the mass of the decaying particle as unknown and have then a 2 C-fit.

Assuming the production point of the V^0 to be unknown, there are three missing variables (P_Λ, λ_Λ and φ_Λ) in the equations above, and we have still a 1 C-fit to find out if the Λ hypothesis for the V^0 is correct.

For the Ξ^- decay we have equations corresponding to Eq. (4) if we take the fitted lambda as measured. In the best case we have no unknown quantities and thus a 4 C-fit. Mostly, however, the Ξ^- track is straight and therefore the momentum of the Ξ^- is missing. We have then a 3 C-fit with the Ξ^- mass known and a 2 C-fit if the mass is calculated.

In the examples above we have only tracks from one vertex, which are subject to a fit. These types of fits are called 1-vertex fits. Once all hypotheses have been tested by 1-vertex fits, one may go a step further and, for example, combine the four equations from the lambda decay and the four from the Ξ^- decay to a 2-vertices fit with eight equations, which in the best case gives a 7 C-fit (P_Λ unknown).

2) The method of least squares

A description of different ways of using the method of least squares in kinematical analysis of bubble chamber events is given by R. Böck (CERN 60-30 and 61-29) and Berge et al. (UCRL-9097).

We denote the measured variables by m_i, i=1,2 ... I, and the unknowns by x_j, j=1,2 ... J. The superscript 0 refers to unfitted quantities. The constraint equations are written $f_k(x,m)$, k=1,2 ... K. The error matrix of the measured variables is called G_m^{-1}, where the element $(G_m^{-1})_{ii}$ is the variance of the element m_i and $(G_m^{-1})_{ij}$ the covariance of elements m_i and m_j. The inverse, G_m, of an error matrix is called a weighting matrix.

The method of least squares is based on the assumption of normally distributed variables and states that the best set of variables is that for which the function

$$\chi^2 = \sum_{j=1}^{J} \sum_{i=1}^{I} (m_i - m_i^0)\ (G_m)_{ij}\ (m_j - m_j^0) \qquad (5)$$

has a minimum, and where the variables fulfil the equations:

$$f_k(x,m) = 0 \quad k = 1,2 \ldots K\ . \qquad (6)$$

Assuming uncorrelated variables $[(G_m^{-1})_{ij} = 0$ for $i \neq j]$ we find that Eq. (5) is reduced to the well-known expression

$$\chi^2 = \sum_{i} \frac{(m_i - m_i^0)^2}{(G_m^{-1})_{ii}} \qquad (5')$$

i.e. one has to minimize the sum of the weighted squares of the deviations between fitted and measured quantities.

The least squares solution may be found in one of the following two ways.

2.1) The elimination method

The J unknowns x_j are eliminated from the constraint equations, and we get a new set of K- J equations

$$f_k(m) = 0 \qquad k = 1,2 \ldots K-J \qquad (7)$$

There are two possibilities to go on.

One may use the equation (7) to eliminate K- J of the measured variables from the function $\chi^2(m)$. The equations

$$\frac{d\chi^2(m)}{dm_i} = 0 \qquad i = 1,2 \ldots I-K+J \tag{8}$$

may then be solved to give the solution for the remaining I- K+ J variables and so on.

The other and better method is to introduce the Lagrangian multipliers α_k, k = 1,2 ... K-J. The problem is then reduced to finding the minimum value of

$$\chi^2(m,\alpha) = \sum_{j=1}^{I}\sum_{i=1}^{I} (m_i - m_i^0)\,(G_m)_{ij}\,(m_j - m_j^0) + 2\sum_{k=1}^{K-J}\alpha_k f_k(m). \tag{9}$$

In the first step we assume that the measured variables are uncorrelated. The minimum of χ^2 is then found by solving the following system of equations:

$$\begin{cases} \dfrac{d\chi^2}{dm_i} = 2\left\{(m_i - m_i^0)\,(G_m)_{ii} + \displaystyle\sum_{k=1}^{K-J}\alpha_k f_{ki}(m)\right\} = 0 & i=1,2 \ldots I \quad (10) \\[2ex] \dfrac{d\chi^2}{d\alpha_k} = 2\,f_k(m) = 0 & k=1,2 \ldots K-J \quad (11) \end{cases}$$

where

$$f_{ki}(m) = \frac{df_k(m)}{dm_i}\,.$$

Equations (11) are just the constraint conditions.

As these are in general not linear, one has to repeat the calculations below to get a still better solution until some given criteria are satisfied. A superscript ν means that this value has been derived in the ν-th iteration.

Assume that we have passed the ν-th iterative step and have to go on at least one step more. From Eq. (10) we have:

$$m_i^{\nu+1} = m_i^0 - \sum_{k=1}^{K-J} (G_m)_{ii}^{-1} f_{ki}^{\nu}(m)\ \alpha_k^{\nu+1} \qquad i=1,2 \ldots I. \tag{12}$$

The constraint equations can be expanded in the following way:

$$f_k^{\nu}(m) + \sum_{i=1}^{I} f_{ki}^{\nu}(m)\ (m_i^{\nu+1} - m_i^{\nu}) = 0 \qquad k=1,2 \ldots K-J. \tag{13}$$

We can now eliminate $m_i^{\nu+1}$ from equations (12) and (13) and get:

$$R_j^{\nu} - \sum_{k=1}^{K-J} S_{jk}^{\nu}\ \alpha_k^{\nu+1} = 0 \qquad j=1,2 \ldots K-J \tag{14}$$

where

$$R_j^{\nu} = f_j^{\nu}(m) + \sum_{i=1}^{I} f_{ji}^{\nu}(m)\ (m_i^0 - m_i^{\nu}) \tag{15}$$

$$S_{jk}^{\nu} = \sum_{i=1}^{I} f_{ki}^{\nu}(m)\ (G_m)_{ii}^{-1}\ f_{ji}^{\nu}(m)\ . \tag{16}$$

Eq. (14) is a system of K- J linear equations from which the unknown variables, $\alpha_k^{\nu+1}$, k=1,2 ... K- J, can be solved.

The new set of $m_i^{\nu+1}$ values is then obtained from relation Eq. (12), and the iteration process can go on until some stopping criteria are fulfilled. The iteration process is started with the values $m_i = m_i^0$ and $\alpha_k = 0$.

The calculations are more simplified when matrix notations are used. We re-write the equations above in this way and extend them to be valid also for correlated measurements (G_m^{-1} is no longer a diagonal matrix).

$$\chi^2 = (m - m^o)^T G_m (m - m^o) + 2\alpha^T f \tag{9'}$$

where T means a transposed matrix

$$\frac{d\chi^2}{dm} = 2\left\{(m - m^o)^T G_m + \alpha^T f_m\right\} = 0 \tag{10'}$$

where $f_m = df(m)/dm$ is a $(K - J) \times I$ matrix

$$m^{\nu+1} = m^o - G_m^{-1} f_m^{\nu T} \alpha^{\nu+1} \tag{12'}$$

$$f^\nu + f_m^\nu (m^{\nu+1} - m^\nu) = 0 \tag{13'}$$

$$R - S\,\alpha^{\nu+1} = 0 \tag{14'}$$

with

$$R = f^\nu + f_m^\nu (m^o - m^\nu) \tag{15'}$$

$$S = f_m^\nu\, G_m^{-1}\, f_m^{\nu T} \tag{16'}$$

and we get

$$\alpha^{\nu+1} = S^{-1} R \,. \tag{17}$$

If Eq. (12') is introduced into Eq. (9') we find

$$\chi^2 = (\alpha^{\nu+1})^T R \tag{18}$$

2.2) A generalized method

The elimination method has the disadvantage that the constraint equations have to be modified in many ways, depending on which variables are missing. One can, however, make the fit in a more general way, so that the same system of constraint equations can be used in all cases.

We have to minimize the expression:

$$\chi^2(m,x,\alpha) = (m-m^0)^T G_m (m-m^0) + 2\alpha^T f(x,m) \tag{19}$$

which means that the following system of equations has to be solved:

$$\begin{cases} \dfrac{d\chi^2}{dm} = 2\left\{(m-m^0)^T G_m + \alpha^T f_m\right\} = 0 & (20) \\ \dfrac{d\chi^2}{dx} = 2\alpha^T f_x = 0 & (21) \\ \dfrac{d\chi^2}{d\alpha} = 2f(x,m) = 0 & (22) \end{cases}$$

where $f_x = df(x,m)/dx$.

The constraint equations can now be expanded in the following way

$$f^\nu + f_x^\nu (x^{\nu+1} - x^\nu) + f_m^\nu (m^{\nu+1} - m^\nu) = 0 \ . \tag{23}$$

Exactly in the same way as above Eqs. (12′ - 17) we obtain:

$$\alpha^{\nu+1} = S^{-1} \left[R + f_x^\nu (x^{\nu+1} - x^\nu)\right] \tag{24}$$

with R and S defined by Eq. (15′) and 16′).

We introduce this vector into Eq. (21) and find:

$$x^{\nu+1} = x^\nu - (f_x^{\nu\,T} S^{-1} f_x^\nu)^{-1} f_x^{\nu\,T} S^{-1} R \ . \tag{25}$$

The problem is then solved. From the new set of $x^{\nu+1}$ values we calculate $\alpha^{\nu+1}$, Eq. (24), and finally $m^{\nu+1}$, Eq. (12′).

As above, Eq. (18), we have

$$\chi^2 = (\alpha^{\nu+1})^T \left[R + f_x (x^{\nu+1} - x^\nu)\right] . \tag{26}$$

The iteration process has to go on until some test indicates that a received set of approximations is good enough. One may, for example, require that the difference in χ^2 between two consecutive iterations is smaller than a given value, or still better that the constraint equations and the derivates Eq. (20) and (21) are zero within some given limits.

2.3) Calculation of errors

From Eqs. (12'), (15'), (24) and (25) above, we see that $m^{\nu+1}$ and $x^{\nu+1}$ can be expressed as explicit functions of m^o. Let us linearize these equations and write:

$$m^{\nu+1} = g(m^o) \tag{27}$$

$$x^{\nu+1} = h(m^o) \ . \tag{28}$$

The error matrices are then obtained from the formulae:

$$G_{m^{\nu+1}}^{-1} = \frac{dg}{dm^o} \, G_m^{-1} \left(\frac{dg}{dm^o}\right)^T \tag{29}$$

$$G_{x^{\nu+1}}^{-1} = \frac{dh}{dm^o} \, G_m^{-1} \left(\frac{dh}{dm^o}\right)^T \tag{30}$$

and the correlation between measured and unmeasured variables from:

$$C_{(mx)^{\nu+1}} = \frac{dg}{dm^o} \, G_m^{-1} \left(\frac{dh}{dm^o}\right)^T \ . \tag{31}$$

The two derivatives needed above, dg/dm^o and dh/dm^o, are obtained in the following way:

$$\begin{aligned}
\frac{dg}{dm^o} &= 1 - G_m^{-1} \, f_m^T \, \frac{d\alpha}{dm^o} \\
&= 1 - G_m^{-1} \, f_m^T \, S^{-1} \left\{ \frac{dR}{dm^o} + f_x \, \frac{d(x^{\nu+1} - x^{\nu})}{dm^o} \right\} \\
&= 1 - G_m^{-1} \, f_m^T \, S^{-1} \left\{ \frac{dR}{dm^o} - f_x (f_x^T \, S^{-1} f_x)^{-1} \, f_x^T \, S^{-1} \, \frac{dR}{dm^o} \right\} \\
&= 1 - G_m^{-1} \, f_m^T \, S^{-1} \left\{ f_m - f_x (f_x^T \, S^{-1} \, f_x)^{-1} \, f_x^T \, S^{-1} \, f_m \right\}
\end{aligned} \tag{32}$$

$$\frac{dh}{dm^o} = (f_x^T \, S^{-1} \, f_x)^{-1} \, f_x^T \, S^{-1} \, f_m \tag{33}$$

After having introduced these expression into Eq. (29 - 31) and simplified the results we obtain:

$$G_m^{-1}{}_{\nu+1} = G_m^{-1} - G_m^{-1} f_m^T S^{-1} f_m G_m^{-1} + G_m^{-1} f_m^T S^{-1} f_x (f_x^T S^{-1} f_x)^{-1} f_x^T S^{-1} f_m G_m^{-1} \quad (34)$$

$$G_x^{-1}{}_{\nu+1} = (f_x^T S^{-1} f_x)^{-1} \quad (35)$$

$$C_{(mx)^{\nu+1}} = -G_m^{-1} f_m^T S^{-1} f_x (f_x^T S^{-1} f_x)^{-1} . \quad (36)$$

One can see from Eq. (34) that the errors of the measured quantities are reduced in the least squares solution and that there are correlations between fitted variables even when the measured quantities are uncorrelated. Some examples are given in Chapter IV.4 and .5.

3) The χ^2 test of a hypothesis

After an event has passed through the kinematical fitting programme, one has to decide if the tested hypothesis is correct or not. All measured variables are forced to fulfil the constraint equations. The magnitude of χ^2 is dependent on how big the differences between the fitted and the measured values are. One can convert the χ^2 value and the number of degrees of freedom ($n = K - J$ in our notations above) into a probability by using the theoretical χ^2 distribution. The latter is given by the formula:

$$f_n(\chi^2)\, d\chi^2 = \frac{(\chi^2)^{n/2-1}}{2^{n/2}\Gamma(n/2)} e^{-\chi^2/2}\, d\chi^2 \quad (37)$$

where Γ is the Gamma function.

A derivation of the χ^2 distribution is given by D. Hudson (Lectures on elementary statistics and probability, CERN 63-29).

Examples of χ^2 distributions for three different numbers of degrees of freedom are shown in Fig. 1.

The mean of a χ^2 distribution is at $\chi^2 = n$, and we have further

$$\int_0^\infty f_n(\chi^2)\, d\chi^2 = 1 \,. \tag{38}$$

The probability that the χ^2 is greater than or equal to a given value χ_0^2 is

$$P_n(\chi^2 \geq \chi_0^2) = \int_{\chi_0^2}^\infty f_n(\chi^2)\, d\chi^2 \,. \tag{39}$$

These probabilities can be found in most statistical tables. One example is given in Chapter V.3, (Table 3).

Assume that the least square method above, gives as result the value χ_0^2 when testing a hypothesis. We can conclude that the hypothesis gives a possible interpretation of the event if the derived probability (P_n) is greater than 5%. If the probability is between 5% and 1%, we find the hypothesis to be doubtful, and if P_n is less than 1% we conclude that this hypothesis is very improbable.

The two percentage limits above (5% and 1%) may be compared with the corresponding limits (2.0 and 2.6 standard deviations) for a normally distributed variable.

Before we can use the probability test we have to check that the χ^2 distribution obtained from the fit procedure agrees with the theoretically expected one. If the experimental χ^2 distribution has a mean value, which is too low (high) we may suspect that the errors of the input parameters are systematically too big (small).

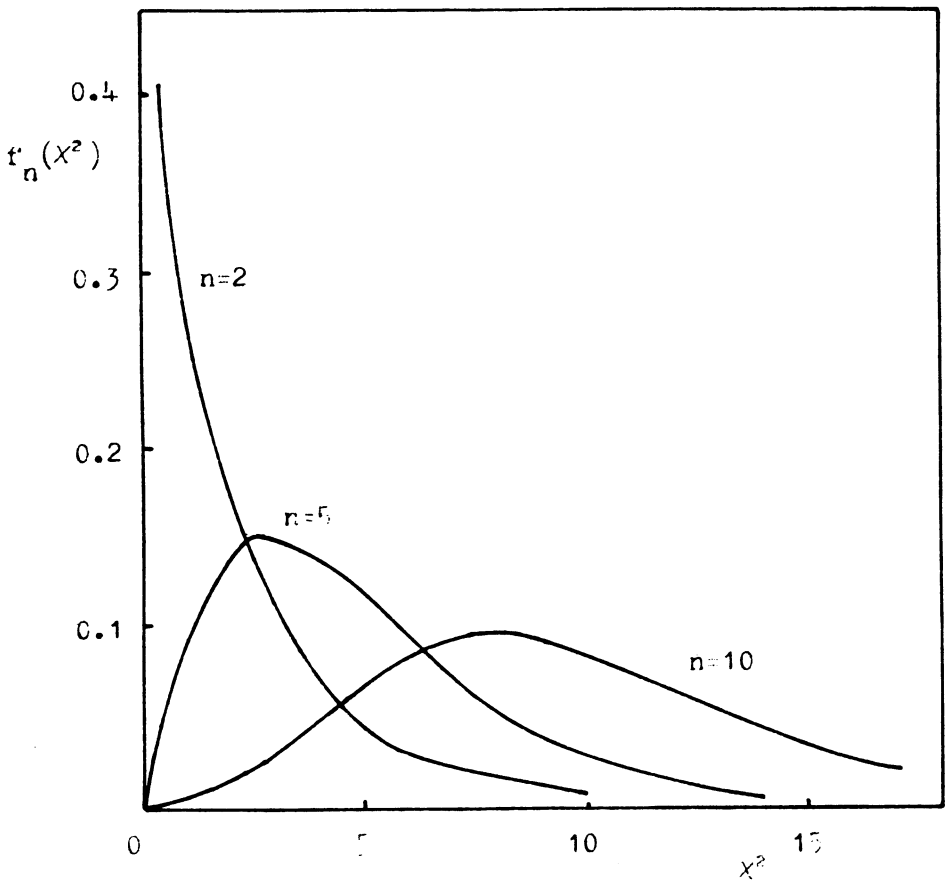


Fig. 1

χ^2 distributions for three different degrees of freedom

It is also possible to calibrate the errors assigned to each variable m_i. We know that

$$S(m_i) = \frac{m_i - m_i^o}{\sigma(m_i - m_i^o)} \qquad (40)$$

should be normally distributed with unit standard deviation and mean value zero. We can then correct the errors until $S(m_i)$ has the correct width. If the $S(m_i)$ distribution is very asymmetric, we conclude that the variable m_i is not normally distributed or that we have a systematic error in the measurement of m_i.

In Fig. 2 we give the experimental and theoretical χ^2 distribution for the 1-vertex Ξ^--fit with M_{Ξ^-} unknown, (see Chapter IV.1). The agreement between the two distributions seems to be good. A method to test the agreement by using probability paper is described in Appendix II.

4) Numerical example; fit of a Λ decay

We illustrate the practical use of the generalized method of least squares by testing the hypothesis that the event below (Fig. 3) is a decaying lambda

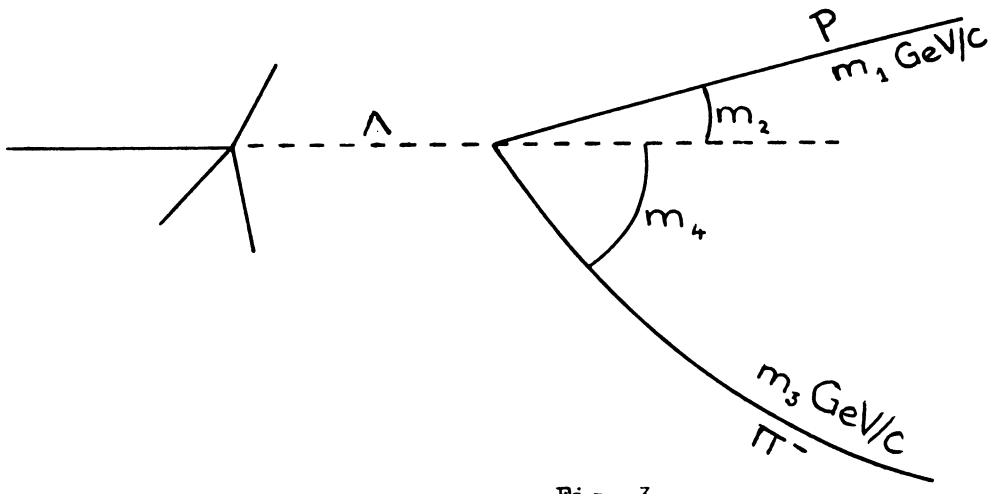


Fig. 3.

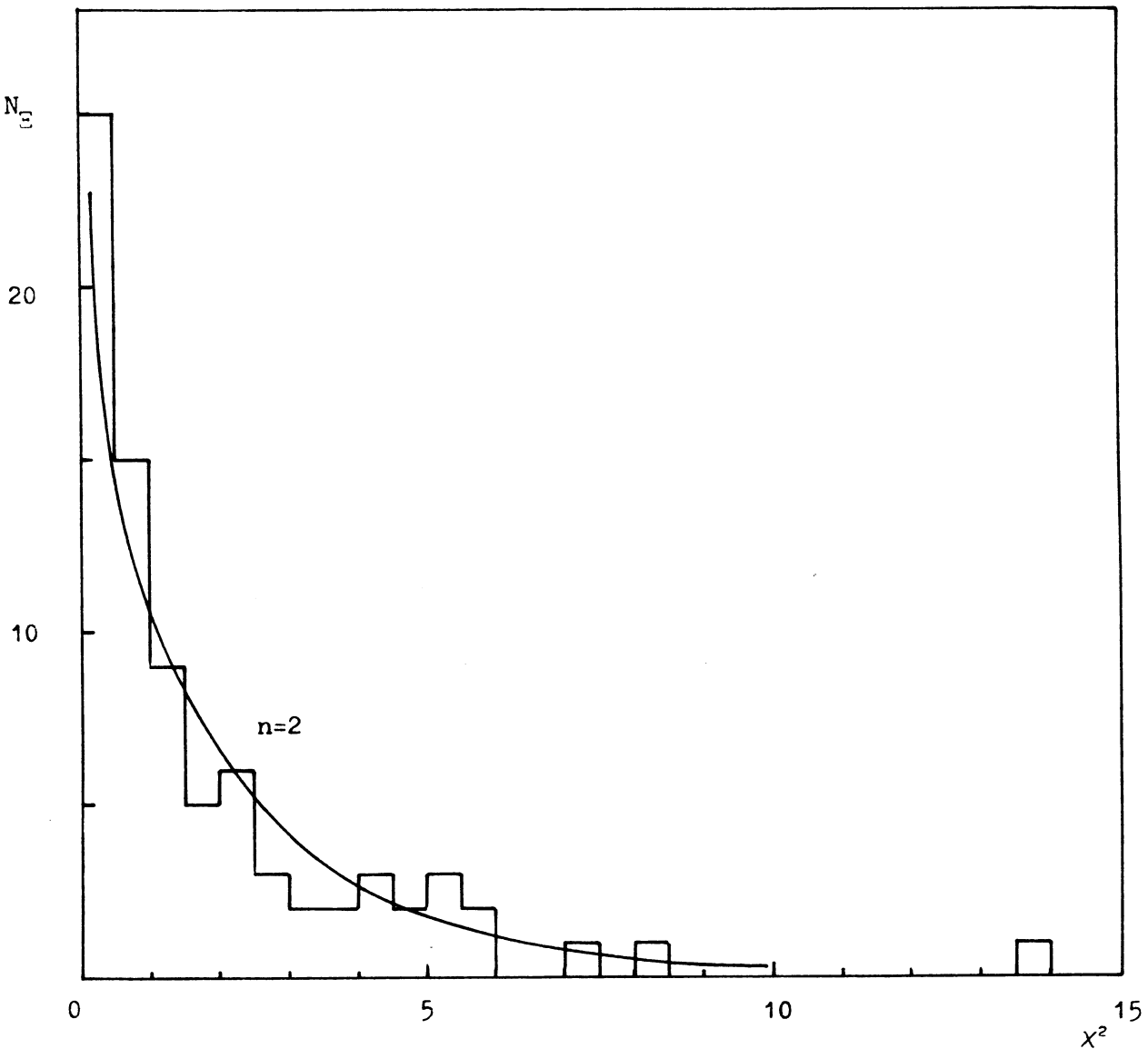


Fig. 2

Experimental and theoretical X^2 distribution for a 2 C-fit.

To make the calculations somewhat easier, we assume that the three tracks (Λ,P and π^-) are coplanar and further that the measured quantities are uncorrelated. The notation of the measured quantities (m_i) with errors (σ_i) is explained in the figure.

The numerical values are

$$\begin{aligned} m_1^0 &= 900 \pm 100 \text{ MeV/c} \\ m_2^0 &= 5^\circ.7 \pm 1^\circ.0 \\ m_3^0 &= 100 \pm 20 \text{ MeV/c} \\ m_4^0 &= 57^\circ.3 \pm 1^\circ.0 \ . \end{aligned}$$

The constraint equations can be written

$$\begin{aligned} f_1 &\equiv x - m_1 \cos m_2 - m_3 \cos m_4 = 0 \\ f_2 &\equiv m_1 - \sin m_2 - m_3 \sin m_4 = 0 \\ f_3 &\equiv \sqrt{x^2 + M_\Lambda^2} - \sqrt{m_1^2 + M_p^2} - \sqrt{m_3^2 + M_\pi^2} = 0 \end{aligned} \qquad (41)$$

where the only unknown quantity (x) is the momentum of the lambda.

We have three equations and one unknown, and thus a 2 C-fit.

The error matrix is

$$G_m^{-1} = \begin{bmatrix} \sigma_1^2 & 0 & 0 & 0 \\ 0 & \sigma_2^2 & 0 & 0 \\ 0 & 0 & \sigma_3^2 & 0 \\ 0 & 0 & 0 & \sigma_4^2 \end{bmatrix}$$

and the weighting matrix

$$G_m = \begin{bmatrix} \frac{1}{\sigma_1^2} & 0 & 0 & 0 \\ 0 & \frac{1}{\sigma_2^2} & 0 & 0 \\ 0 & 0 & \frac{1}{\sigma_3^2} & 0 \\ 0 & 0 & 0 & \frac{1}{\sigma_4^2} \end{bmatrix}$$

We have to minimize the expression

$$\chi^2(m,x,\alpha) = (m - m^0)^T G_m (m - m^0) + 2\alpha^T f(x,m) \qquad (19)$$

which can be written

$$\chi^2 = (m_1 - m_1^0 ; m_2 - m_2^0 ; m_3 - m_3^0 ; m_4 - m_4^0) \begin{bmatrix} \frac{1}{\sigma_1^2} & 0 & 0 & 0 \\ 0 & \frac{1}{\sigma_2^2} & 0 & 0 \\ 0 & 0 & \frac{1}{\sigma_3^2} & 0 \\ 0 & 0 & 0 & \frac{1}{\sigma_4^2} \end{bmatrix} \begin{bmatrix} m_1 - m_1^0 \\ m_2 - m_2^0 \\ m_3 - m_3^0 \\ m_4 - m_4^0 \end{bmatrix} +$$

$$+ 2(\alpha_1 ; \alpha_2 ; \alpha_3) \begin{bmatrix} f_1 \\ f_2 \\ f_3 \end{bmatrix} =$$

$$= \sum_{i=1}^{4} \frac{(m_i - m_i^0)^2}{\sigma_i^2} + 2 \sum_{j=1}^{3} \alpha_j f_j \, .$$

This is reduced to the first well-known term when the constraint equations are fulfilled.

The following system of equations has to be solved.

$$\begin{cases} \dfrac{d\chi^2}{dm_i} = 2 \left\{ \dfrac{m_i - m_i^0}{\sigma_i^2} + \displaystyle\sum_{j=1}^{3} \alpha_j \frac{df_j}{dm_i} \right\} = 0 \qquad i = 1,2,3,4 & (42) \\ \dfrac{d\chi^2}{dx} = 2 \displaystyle\sum_{j=1}^{3} \alpha_j \frac{df_j}{dx} = 0 & (43) \\ \dfrac{d\chi^2}{d\alpha_j} = 2 f_j = 0 \qquad j = 1,2,3 & (44) \end{cases}$$

We have thus eight equations and eight unknowns (α_1 α_2 α_3 m_1 m_2 m_3 m_4 and x).

The constraint equations can be expanded in the following way after we have passed the ν-th iterative step

$$f_j^\nu + \sum_{i=1}^{4} f_{ji}^\nu (m_i^{\nu+1} - m_i^\nu) + f_{x_j}^\nu (x^{\nu+1} - x^\nu) = 0 \qquad j = 1,2,3. \tag{45}$$

We have

$$\left\{ f_{ji}^\nu \right\} = \left\{ \frac{\partial f_j^\nu}{\partial m_i^\nu} \right\} \equiv \begin{bmatrix} -\cos m_2^\nu & + m_1^\nu \sin m_2^\nu & -\cos m_4^\nu & + m_3^\nu \sin m_4^\nu \\ +\sin m_2^\nu & + m_1^\nu \cos m_2^\nu & -\sin m_4^\nu & - m_3^\nu \cos m_4^\nu \\ -\dfrac{m_1^\nu}{\sqrt{m_1^{\nu 2} + M_p^2}} & 0 & -\dfrac{m_3^\nu}{\sqrt{m_3^{\nu 2} + M_\pi^2}} & 0 \end{bmatrix} \tag{46}$$

and

$$\left\{ f_{x_j}^\nu \right\} = \left\{ \frac{\partial f_j^\nu}{\partial x^\nu} \right\} \equiv \begin{bmatrix} 1 \\ 0 \\ \dfrac{x^\nu}{\sqrt{x^{\nu 2} + M_\Lambda^2}} \end{bmatrix} \tag{47}$$

We can re-write the constraint equations in a more suitable form.

$$\sum_{i=1}^{4} f_{ji}^\nu (m_i^{\nu+1} - m_i^0) + R_j + f_{x_j}^\nu (x^{\nu+1} - x^\nu) = 0 \qquad j = 1,2,3 \tag{48}$$

where

$$R_j = f_j^\nu - \sum_{i=1}^{4} f_{ji}^\nu (m_i^\nu - m_i^0) \tag{49}$$

Equation (42) can be written

$$m_i^{\nu+1} - m_i^0 = -\sigma_i^2 \sum_{j=1}^{3} \alpha_j^{\nu+1} f_{ji}^\nu \qquad i = 1,2,3,4 \ . \tag{42'}$$

We can now eliminate $m_i^{\nu+1}$ between Eqs (42') and (48).

$$\sum_{i=1}^{4} \sigma_i^2 f_{ji}^{\nu} \sum_{k=1}^{3} \alpha_k^{\nu+1} f_{ki}^{\nu} = R_j + f_{x_j}^{\nu} (x^{\nu+1} - x^{\nu}) \qquad j = 1,2,3. \tag{50}$$

After re-arranging the terms we have

$$\sum_{k=1}^{3} \alpha_k^{\nu+1} \sum_{i=1}^{4} f_{ji}^{\nu} f_{ki}^{\nu} \sigma_i^2 = R_j + f_{x_j}^{\nu} (x^{\nu+1} - x^{\nu}) \qquad j = 1,2,3 \tag{51}$$

or

$$\begin{cases} \alpha_1^{\nu+1} \sum_{i=1}^{4} f_{1i}^{\nu 2} \sigma_i^2 + \alpha_2^{\nu+1} \sum_{i=1}^{4} f_{1i}^{\nu} f_{2i}^{\nu} \sigma_i^2 + \alpha_3^{\nu+1} \sum_{i=1}^{4} f_{1i}^{\nu} f_{3i}^{\nu} \sigma_i^2 = R_1 + f_{x_1} (x^{\nu+1} - x^{\nu}) \\ \alpha_1^{\nu+1} \sum_{i=1}^{4} f_{2i}^{\nu} f_{1i}^{\nu} \sigma_i^2 + \alpha_2^{\nu+1} \sum_{i=1}^{4} f_{2i}^{\nu 2} \sigma_i^2 + \alpha_3^{\nu+1} \sum_{i=1}^{4} f_{2i}^{\nu} f_{3i}^{\nu} \sigma_i^2 = R_2 + f_{x_2}^{\nu} (x^{\nu+1} - x^{\nu}) \\ \alpha_1^{\nu+1} \sum_{i=1}^{4} f_{3i}^{\nu} f_{1i}^{\nu} \sigma_i^2 + \alpha_2^{\nu+1} \sum_{i=1}^{4} f_{3i}^{\nu} f_{2i}^{\nu} \sigma_i^2 + \alpha_3^{\nu+1} \sum_{i=1}^{4} f_{3i}^{\nu 2} \sigma_i^2 = R_3 + f_{x_3}^{\nu} (x^{\nu+1} - x^{\nu}) \end{cases} \tag{51'}$$

The three unknowns $(\alpha_i^{\nu+1})$ can be solved from this linear system of equations.

Eq. (51) is of course identical with the equations derived in Chapter IV.2.2.

$$\alpha^{\nu+1} = S^{-1} [R + f_x^{\nu} (x^{\nu+1} - x^{\nu})]. \tag{24}$$

We have namely

$$S = f_m^{\nu} \, G_m^{-1} \, f_m^{\nu\,T} =$$

$$= \begin{bmatrix} f_{11}^{\nu} & f_{12}^{\nu} & f_{13}^{\nu} & f_{14}^{\nu} \\ f_{21}^{\nu} & f_{22}^{\nu} & f_{23}^{\nu} & f_{24}^{\nu} \\ f_{31}^{\nu} & f_{32}^{\nu} & f_{33}^{\nu} & f_{34}^{\nu} \end{bmatrix} \begin{bmatrix} \sigma_1^2 & 0 & 0 & 0 \\ 0 & \sigma_2^2 & 0 & 0 \\ 0 & 0 & \sigma_3^2 & 0 \\ 0 & 0 & 0 & \sigma_4^3 \end{bmatrix} \begin{bmatrix} f_{11}^{\nu} & f_{21}^{\nu} & f_{31}^{\nu} \\ f_{12}^{\nu} & f_{22}^{\nu} & f_{32}^{\nu} \\ f_{13}^{\nu} & f_{23}^{\nu} & f_{33}^{\nu} \\ f_{14}^{\nu} & f_{24}^{\nu} & f_{34}^{\nu} \end{bmatrix} =$$

$$= \begin{bmatrix} s_{11} & s_{12} & s_{13} \\ s_{21} & s_{22} & s_{23} \\ s_{31} & s_{32} & s_{33} \end{bmatrix} \tag{52}$$

where

$$s_{jk} = \sum_{i=1}^{4} f_{ji}^{\nu} \, f_{ki}^{\nu} \, \sigma_i^2 \ . \tag{53}$$

The S matrix has to be inverted before we can solve $\alpha_i^{\nu+1}$. We have

$$S^{-1} = \frac{1}{|S|} \begin{bmatrix} S_{11} & S_{21} & S_{31} \\ S_{12} & S_{22} & S_{32} \\ S_{13} & S_{23} & S_{33} \end{bmatrix} \tag{54}$$

where $|S|$ is the determinant of the matrix S and S_{jk} is the co-factor of the element s_{jk}. (The co-factor S_{jk} is equal to $(-1)^{j+k}$ times that minor which is obtained when the j-th row and the k-th column are deleted.)

The solution of Eq. (51) is

$$\alpha_j^{\nu+1} = \frac{1}{|S|} \sum_{k=1}^{3} S_{kj} \left[R_k + f_{x_k}^{\nu} \left(x^{\nu+1} - x^{\nu} \right) \right] \qquad j = 1,2,3 \ . \tag{55}$$

$\alpha_j^{\nu+1}$ is eliminated from Eq. (43)

$$\sum_{j=1}^{3} f_{x_j}^{\nu} \sum_{k=1}^{3} S_{kj} \left[R_k + f_{x_k}(x^{\nu+1} - x^{\nu})\right] = 0 \qquad (43')$$

and $x^{\nu+1} - x^{\nu}$ can be solved

$$x^{\nu+1} - x^{\nu} = - \frac{\sum_{j=1}^{3} f_{x_j}^{\nu} \sum_{k=1}^{3} S_{kj} R_k}{\sum_{j=1}^{3} f_{x_j}^{\nu} \sum_{k=1}^{3} S_{kj} f_{x_k}^{\nu}} \qquad (56)$$

The solution Eq. (56) is of course also obtained from Eq. (25) after some matrix operations.

The χ^2 value can be calculated from Eq. (26)

$$\chi^2 = \sum_{j=1}^{3} \alpha_j^{\nu+1} \left[R_j + f_{x_j}(x^{\nu+1} - x^{\nu})\right] \qquad (26)$$

The numerical calculations are made in the following order:

(i) An approximate value of x is obtained by solving the first constraint equation (41).

(ii) The derivative matrices Eq. (46) and (47) are calculated. We have $\nu = 0$ in the first step.

(iii) The S matrix Eq. (52) is constructed and inverted.

(iv) The correction to the unknown variable is solved Eq. (56).

(v) The α vector Eq. (55) and the correction vector $(m_i^{\nu+1} - m_i^{o})$, Eq. (42') are evaluated.

(vi) χ^2 is calculated, Eq. (26).

(vii) A test is made to see if the derivatives Eqs. (42 - 44) are smaller than some given values. If this is not the case, we start again with (ii) above and now replace x^{ν} and m_i^{ν} with $x^{\nu+1}$ and $m_i^{\nu+1}$.

(viii) When the given stopping conditions are fulfilled we calculate the errors in the fitted quantities, Eqs. (34 - 36).

Results from calculations on the given example are given in Table 1.

Step	m_1 MeV/c	m_2 degrees	m_3 MeV/c	m_4 degrees	x MeV/c	f_1 MeV/c	f_2 MeV/c	f_3 MeV	χ^2
0	900.0	5.73	100.0	57.30	949.5	0.000	5.703	-6.974	-
1	812.7	5.37	89.9	57.16	857.9	0.090	5.321	-0.076	1.17
2	806.7	5.40	90.4	57.18	852.2	0.001	0.004	0.002	1.22
3	806.9	5.40	90.4	57.18	852.3	0.000	0.000	0.000	1.22

Table 1

We find that the constraint equations are well fulfilled already after two steps. The χ^2 value of the fit is lower than the mean value we expect to have (= 2 for a 2 C-fit) and we conclude that the tested event is in good agreement with the given hypothesis.

The error matrix of the fitted quantities derived after step 3 is (σ_i given in GeV/c and radians)

$$G_{m^3}^{-1} = \begin{bmatrix} 0.002360 & -0.000655 & -0.000353 & -0.000137 \\ -0.000655 & 0.000187 & 0.000105 & -0.000001 \\ -0.000353 & 0.000105 & 0.000063 & -0.000033 \\ -0.000137 & -0.000001 & -0.000033 & 0.000295 \end{bmatrix}$$

and the error in x is 46 MeV/c.

Finally we give the best fit solution of the observed lambda decay.

m_1 = 307 ± 49 MeV/c
m_2 = 5°.4 ± 0°.6
m_3 = 90 ± 8 MeV/c
m_4 = 57°.2 ± 1°.0
x = 852 ± 46 MeV/c

There is at CERN one subroutine (BOECK; written by W. Koch), which makes the calculations (iii) - (vi) and (viii) above, for the general case, when we have correlated variables, (see Appendix I).

5) Results from a Ξ^--fit

A typical Ξ^- candidate is shown on Fig. 4. In Fig. 5 we give the results from the geometrical and kinematical programmes.

The notation of points and tracks is explained in Fig. 4. NTR is the number of tracks leaving a point. The nature (NAT) of a point is for example, 2 if the apex of a V^0 decay and 3 if the apex of a V^- decay starts from the point. The nature of a track is 1 if the particle stops in the chamber, 2 if the track combines two interaction vertices and 10 for beam tracks. The CODE (U = unmeasured, W = measured and F = fixed) is given for the momentum (P), dip angle (DIP), azimuth angle (PHI) and mass respectively. AD in front of a quantity means the error of this quantity. A U or an F after a variable means unfitted and fitted quantities respectively, L is the measured length of a track, H the average magnetic field and BUB the expected bubble density of a charged track.

The results concerning points and tracks are obtained from THRESH and the fit results from GRIND.

We will now look at the different fits.

a) K^0 (Vertex NN)

TYP 200 indicate a V^0 and HYP 1 is a K^0. The first fit ① is a K^0-fit of the NN vertex. The mass of K^0 is fixed (=0.4978 GeV). We see that the momenta of the pions are higher than those given in THRESH. The reason for

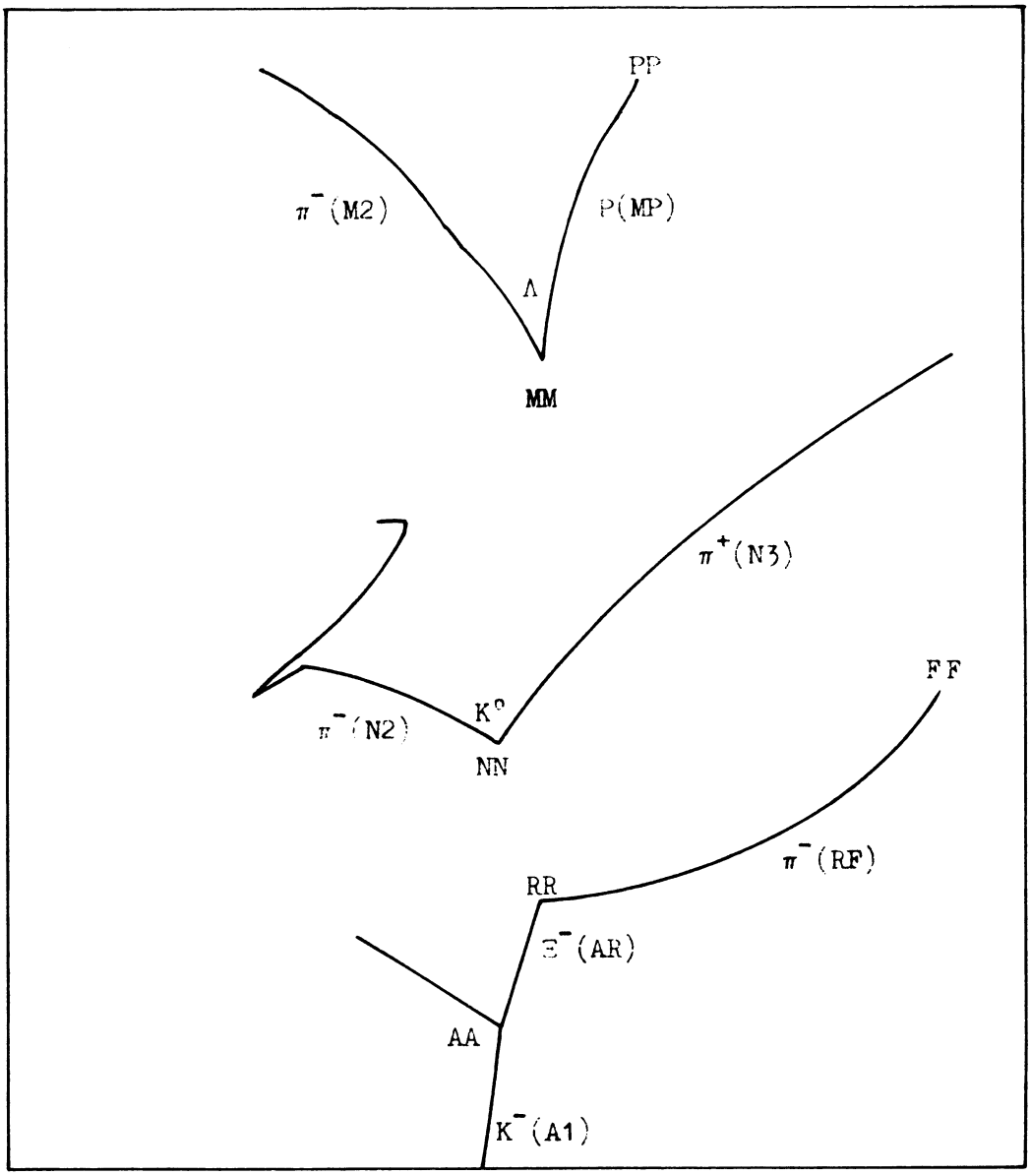


Fig. 4

A Ξ^- candidate. The production process may be

$K^- + N \rightarrow \Xi^- + K^o$.

this is of course that THRESH gives the momentum of a particle at the mean point of the measured track length, while the unfitted momentum in a fit is corrected to be valid at the interaction vertex.

The unfitted errors are greater than THRESH errors because uncertainties due to Coulomb scattering have to be included.

The K^0 is assumed to come from the point AA, (TRACK NA), and we consequently know the direction of this particle. Both pions are completely measured and we have only one unknown in the fit, the momentum of the kaon. This is a 3 C-fit (ND = 3). An unfitted value of the kaon momentum is, however, given. It has been calculated from the constraint equations, and must be available before the least squares fit can start.

At the top we find that the χ^2 of the fit is 1.01, the probability of getting a χ^2 greater than the obtained value is 0.7999 and the number of iteration steps performed to come to the solution is 3.

The K^0 hypothesis is very good for the vertex NN, and the fitted data agree well with the unfitted ones. The fitted errors are generally smaller than the unfitted. This is especially true for the relatively badly measured momenta of the pions (from 67 and 215 MeV/c to 8 and 7 MeV/c respectively).

Fit (2) differs from (1) in that the mass of the neutral particle is assumed to be unknown. We have therefore a 2 C-fit. The obtained probability (0.6048) together with the fitted mass 497.5 ± 63.8 MeV indicate that the K^0 hypothesis is still good.

GRIND has also tried to make a Λ-fit (HYP 2) and an $\bar{\Lambda}$-fit (HYP 3) with the vertex NN, but we see from the summary bank, which follows after the fits in Fig. 5, that an error 15 (ER 15) was found. This type of error indicates that a fit to the given hypothesis is impossible. We may, for example, have some obvious contradiction of a physical law, non-convergence of the iteration process, or corrections to the measured data, which are much greater than what can be expected from statistical fluctuations. The last is what has happened in our case; χ^2 is 26.8 and 41.2 for the Λ and $\bar{\Lambda}$ hypotheses after 3 and 2 steps respectively.

Fig. 5. THRESH and GRIND results for the Ξ^- candidate on Fig. 4.

EVENT 44 467683.1

POINTS

		NTR	NAT	X	Y	Z	DX	DY	DZ
1	A	2	0	6.3861	0.7826	-23.8324	0.0155	0.0149	0.0812
2	R	2	3	7.9113	0.5849	-23.6247	0.0210	0.0199	0.1086
3	M	2	2	17.2402	1.0983	-22.1280	0.0200	0.0177	0.0961
4	N	2	2	10.2983	1.3145	-25.8938	0.0161	0.0150	0.0828
5	P	0	12	27.8238	0.1450	-20.0237	0.0237	0.0191	0.1032
6	F	0	12	17.8081	-8.8863	-27.4560	0.0128	0.0114	0.0624

TRACKS		NAT	CODE	P	DIP	PHI	DP	DDIP	DPHI	L	DL	SAG	PCOSL	H	TESTS
1	AR +-	2	U W W	0.	0.0781	6.1567	0.	0.0506	0.0077	1.59	0.30	0.000	0.	0.	42
2	RF -	1	W W W	0.0650	-0.1949	5.1206	0.0018	0.1176	0.0149	15.07	0.04	2.148	0.064	17.11	0
3	M2 -	0	W W W	0.1747	0.1320	0.2950	0.0129	0.0035	0.0022	8.73	0.30	0.274	0.173	16.93	0
4	N2 -	0	W W W	0.3304	-0.5039	0.9497	0.1683	0.0269	0.0150	4.79	0.30	0.039	0.289	17.20	40
5	MP +	1	W W W	0.4422	0.1768	6.2616	0.0531	0.0037	0.0026	10.91	0.04	-0.167	0.435	16.91	0
6	N3 +	0	W W W	0.3148	-0.2907	5.7404	0.0143	0.0062	0.0029	15.16	0.30	-0.453	0.302	17.25	0
7	A1 -	10	W W W	1.3475	0.0069	3.2228	0.0168	0.0022	0.0012	-57.54	0.30	-0.090	1.348	17.09	0
8	RA +-	2	U W W	0.	-0.0781	3.0151	0.	0.0506	0.0077	-1.59	0.30	-0.000	0.	0.	42

(1)

GOOD FIT NOPT 1 NOTR 3 TYP 200 HYP 1 ER14 NONE CHISQ 1.01 ND 3 PROB 0.7999 STEP 3

TRACK	MASS	CODE	BUB	P U	DIP U	PHI U	DP U	DDP U	DPH U	P F	DIP F	PHI F	DP F	DDP F	DPH F	DM
NA 0	0.4978	U W W F	2.2	0.518	0.481	3.277	0.	0.023	0.005	0.505	0.497	3.277	0.012	0.017	0.005	0.
N3 +	0.1396	W W W F	1.2	0.332	-0.291	5.740	0.067	0.024	0.017	0.336	-0.280	5.743	0.008	0.022	0.017	0.
N2 -	0.1396	W W W F	1.4	0.336	-0.504	0.950	0.215	0.030	0.018	0.315	-0.488	0.948	0.007	0.025	0.018	0.

(2)

GOOD FIT NOPT 1 NOTR 3 TYP 200 HYP 1 ER14 NONE CHISQ 1.01 ND 2 PROB 0.6048 STEP 4

TRACK	MASS	CODE	BUB	P U	DIP U	PHI U	DP U	DDP U	DPH U	P F	DIP F	PHI F	DP F	DDP F	DPH F	DM
NA 0	0.4975	U W W U	2.2	0.518	0.481	3.277	0.	0.023	0.005	0.504	0.497	3.277	0.097	0.017	0.005	0.0638
N3 +	0.1396	W W W F	1.2	0.332	-0.291	5.740	0.067	0.024	0.017	0.336	-0.280	5.743	0.065	0.022	0.017	0.
N2 -	0.1396	W W W F	1.4	0.336	-0.504	0.950	0.215	0.030	0.018	0.315	-0.488	0.948	0.059	0.025	0.018	0.

(3)

GOOD FIT NOPT 1 NOTR 3 TYP 200 HYP 0 ER14 NONE CHISQ 0.17 ND 1 PROB 0.6763 STEP 4

TRACK	MASS	CODE	BUB	P U	DIP U	PHI U	DP U	DDP U	DPH U	P F	DIP F	PHI F	DP F	DDP F	DPH F	DM
MO 0	1.1154	U U U F	3.6	0.667	-0.166	3.208	0.246	0.011	0.006	0.694	-0.165	3.218	0.004	0.022	0.016	0.
MP +	0.9382	W W W F	4.8	0.486	0.177	6.261	0.002	0.027	0.019	0.486	0.177	6.261	0.002	0.027	0.019	0.
M2 -	0.1396	W W W F	1.4	0.188	0.132	0.296	0.058	0.037	0.027	0.216	0.132	0.300	0.002	0.037	0.025	0.

(4)

GOOD FIT NOPT 1 NOTR 3 TYP 200 HYP 2 ER14 NONE CHISQ 9.37 ND 3 PROB 0.0248 STEP 4

TRACK	MASS	CODE	BUB	P U	DIP U	PHI U	DP U	DDP U	DPH U	P F	DIP F	PHI F	DP F	DDP F	DPH F	DM
MA 0	1.1154	U W W F	3.6	0.667	-0.156	3.171	0.	0.011	0.002	0.691	-0.158	3.172	0.004	0.010	0.002	0.
MP +	0.9382	W W W F	4.8	0.486	0.177	6.261	0.002	0.027	0.019	0.486	0.168	6.209	0.002	0.018	0.009	0.
M2 -	0.1396	W W W F	1.4	0.188	0.132	0.296	0.058	0.037	0.027	0.214	0.128	0.267	0.003	0.033	0.022	0.

(5)

GOOD FIT NOPT 1 NOTR 3 TYP 200 HYP 2 ER14 NONE CHISQ 2.63 ND 2 PROB 0.2686 STEP 4

TRACK	MASS	CODE	BUB	P U	DIP U	PHI U	DP U	DDP U	DPH U	P F	DIP F	PHI F	DP F	DDP F	DPH F	DM
MA 0	1.0909	U W W U	4.1	0.667	-0.156	3.171	0.	0.011	0.002	0.623	-0.158	3.171	0.023	0.010	0.002	0.0068
MP +	0.9382	W W W F	4.8	0.486	0.177	6.261	0.002	0.027	0.019	0.486	0.166	6.240	0.002	0.015	0.014	0.
M2 -	0.1396	W W W F	2.0	0.188	0.132	0.296	0.058	0.037	0.027	0.143	0.126	0.276	0.024	0.035	0.023	0.

(6)

GOOD FIT NOPT 1 NOTR 3 TYP 200 HYP 2 ER14 NONE CHISQ 2.09 ND 3 PROB 0.5545 STEP 4

TRACK	MASS	CODE	BUB	P U	DIP U	PHI U	DP U	DDP U	DPH U	P F	DIP F	PHI F	DP F	DDP F	DPH F	DM
MR 0	1.1154	U W W F	3.6	0.667	-0.159	3.197	0.	0.015	0.003	0.693	-0.161	3.197	0.004	0.013	0.003	0.
MP +	0.9382	W W W F	4.8	0.486	0.177	6.261	0.002	0.027	0.019	0.486	0.172	6.237	0.002	0.019	0.009	0.
M2 -	0.1396	W W W F	1.4	0.188	0.132	0.296	0.058	0.037	0.027	0.215	0.129	0.285	0.003	0.034	0.022	0.

(7)

GOOD FIT NOPT 1 NOTR 3 TYP 200 HYP 2 ER14 NONE CHISQ 0.28 ND 2 PROB 0.8706 STEP 3

TRACK	MASS	CODE	BUB	P U	DIP U	PHI U	DP U	DDP U	DPH U	P F	DIP F	PHI F	DP F	DDP F	DPH F	DM
MR 0	1.0992	U W W U	3.9	0.667	-0.159	3.197	0.	0.015	0.003	0.650	-0.161	3.197	0.029	0.013	0.003	0.0101
MP +	0.9382	W W W F	4.8	0.486	0.177	6.261	0.002	0.027	0.019	0.486	0.171	6.256	0.002	0.018	0.016	0.
M2 -	0.1396	W W W F	1.7	0.188	0.132	0.296	0.058	0.037	0.027	0.171	0.128	0.290	0.030	0.035	0.023	0.

(8)

GOOD FIT NOPT 1 NOTR 3 TYP 320 HYP 4 ER14 NONE CHISQ 0.18 ND 2 PROB 0.9132 STEP 3

TRACK	MASS.	CODE	BUB	P U	DIP U	PHI U	DP U	DDP U	DPH U	P F	DIP F	PHI F	DP F	DDP F	DPH F	DM
RA -	1.3195	U W W U	4.1	0.754	-0.078	3.015	0.	0.051	0.008	0.753	-0.092	3.015	0.017	0.032	0.007	0.0095
RF -	0.1396	W W W F	1.9	0.152	-0.195	5.129	0.001	0.258	0.166	0.152	-0.275	5.141	0.001	0.172	0.097	0.
RM 0	1.1154	W W W F	3.6	0.693	0.161	0.056	0.004	0.013	0.003	0.693	0.160	0.056	0.004	0.012	0.003	0.

(9)

GOOD FIT NOPT 1 NOTR 3 TYP 320 HYP 4 ER14 NONE CHISQ 0.22 ND 3 PROB 0.9738 STEP 3

TRACK	MASS	CODE	BUB	P U	DIP U	PHI U	DP U	DDP U	DPH U	P F	DIP F	PHI F	DP F	DDP F	DPH F	DM
RA -	1.3215	U W W F	4.1	0.754	-0.078	3.015	0.	0.051	0.008	0.749	-0.087	3.014	0.005	0.020	0.007	0.
RF -	0.1396	W W W F	1.9	0.152	-0.195	5.129	0.001	0.258	0.166	0.152	-0.305	5.121	0.001	0.089	0.028	0.
RM 0	1.1154	W W W F	3.6	0.693	0.161	0.056	0.004	0.013	0.003	0.693	0.160	0.056	0.004	0.012	0.003	0.

(10)

GOOD FIT NOPT 2 NOTR 6 TYP 320 HYP 4 ER14 NONE CHISQ 2.31 ND 6 PROB 0.8891 STEP 4

TRACK	MASS	CODE	BUB	P U	DIP U	PHI U	DP U	DDP U	DPH U	P F	DIP F	PHI F	DP F	DDP F	DPH F	DM
RA -	1.3215	U W W F	4.1	0.729	-0.078	3.015	0.	0.051	0.008	0.749	-0.087	3.014	0.005	0.020	0.007	0.
RF -	0.1396	W W W F	1.9	0.152	-0.195	5.129	0.001	0.258	0.166	0.152	-0.305	5.121	0.001	0.089	0.028	0.
RM 0	1.1154	U W W F	3.6	0.667	0.159	0.055	0.	0.015	0.003	0.693	0.160	0.056	0.004	0.012	0.003	0.
MP +	0.9382	W W W F	4.8	0.486	0.177	6.261	0.002	0.027	0.019	0.486	0.171	6.237	0.002	0.019	0.009	0.
M2 -	0.1396	W W W F	1.4	0.188	0.132	0.296	0.058	0.037	0.027	0.215	0.129	0.285	0.002	0.034	0.022	0.

SUMMARY BANK FOR EVENT 44 467683.1 5

	NOPT	PT	NOTR	TYP	HYP	GDNS	ER14	ER15	CHISQ	ND	PROB	STEPS
1	1	4	3	200	1	77	-	-	0.1006E 01	3	0.7999	3
2	1	4	3	200	2	17	-	FIT 34 NO FIT	0.2679E 02	3	0.0000	3
3	1	4	3	200	3	17	-	FIT 34 NO FIT	0.4123E 02	3	0.0000	2
4	1	3	3	200	2	57	-	-	0.9366E 01	3	0.0248	4
5	1	3	3	200	2	77	-	-	0.2087E 01	3	0.5545	4
6	1	2	3	320	4	77	-	-	0.1816E-00	2	0.9132	3
7	1	2	3	320	4	77	-	-	0.2228E-00	3	0.9738	3
8	2	2	6	320	4	0	-	-	0.2310E 01	6	0.8891	4
9	2	2	6	320	4	0	-	-	0.2310E 01	6	0.8891	4

An error 14 (ER 14), of which we have none in our example, shows that some minor contradiction has been found in the fit, for example, that the momentum of a particle derived from its range and curvature disagree. The fit will in these cases go on, but the contradiction is flagged at the top of the fit results and also in the summary bank.

b) Λ (Vertex MM)

Results from good V^0 fits of vertex MM are given in (3) to (7).

The lambda is assumed to be completely unknown in (3) (HYP 0, CODE U U U F). We have then a 1 C-fit, which gives $\chi^2 = 0.17$ and probability 0.6763. This shows that the lambda hypothesis is possible.

In (4) and (5) we assume that the lambda comes from the primary vertex AA. Fit (4), where the lambda mass is fixed, results in a very high χ^2 and gives a low probability (0.0248). Fit (5) with the mass of the neutral particle unknown gives a high probability, but the derived mass 1090.9 ± 6.8 MeV is 3.6 standard deviations from the lambda mass. We conclude that the hypothesis that the lambda comes from vertex AA is almost disproved by the fit results, or at least shown to be a very improbable solution.

The hypothesis that the lambda comes from the kink RR is a very good solution according to the two fits (6) and (7).

c) Ξ^- 1-vertex fit (Vertex RR)

When a lambda is shown to come from the kink of a negative track, we test the hypothesis: $\Xi^- \rightarrow \Lambda + \pi^-$.

We assume in the 1-vertex fit that the fitted lambda (RM in our notation) is measured and we have a 3 C-fit with the Ξ^- mass known (9), and a 2 C-fit with the mass unknown (8).

The two fits give very high probabilities, and the derived mass, 1319.5 ± 9.5 MeV, is in good agreement with the known Ξ^- mass. We conclude that the Ξ^- hypothesis is a good explanation for this event.

d) Ξ^- 2-vertices fit (Vertices RR and MM)

If a good 1-vertex Ξ^--fit is obtained, we go on by combining the two fits (Ξ^- and Λ decay) in a 2-vertices fit, (10). We then have 8 constraint equations and in our case 2 unknowns (P_{Ξ^-} and P_Λ). The high probability

(0.8891) of the 6 C-fit means that our conclusion above has a good chance of being correct.

A 2-vertices E^--fit is also tried when we do not get a good 1-vertex fit. The 2-vertices fit should in principle be superior, and we have observed a few events where the 2-vertices fit is good, while the 1-vertex fit fails. However, in most cases we obtain both fits, and the fitted data agree extremely well.

V. COMPARISON OF RESULTS FROM AN EXPERIMENT WITH THEORETICAL PREDICTIONS

1) Introduction

Most experiments are performed to test some hypothesis. The observed quantities will of course deviate from the theoretical ones. The purpose of a test is to find out whether the deviation between experimental and theoretical values could be due to random sampling variations, or whether the deviation is too great, from which we may conclude that the hypothesis is probably wrong.

In the following we describe two types of hypotheses which are often tested.

In the first type (Chapter V.2) the hypothesis gives the magnitude of some parameter associated with the population, e.g. the mean value. The test is in general performed in the following way. The sampling distribution of the parameter in question is determined and the measured quantity is compared with its theoretical distribution.

The theoretical frequency distribution is known in the second type of hypothesis. The agreement between the distribution of the observed sample and the hypothesis is in general tested by a χ^2 test of significance (Chapter V.3).

The two types of test will be illustrated by an example taken from an investigation of leptonic decay modes (Physics Letters 6, 186 (1963)). About 90 lambda hyperons decaying to electrons ($\Lambda \rightarrow P + e^- + \bar{\nu}$) were observed.

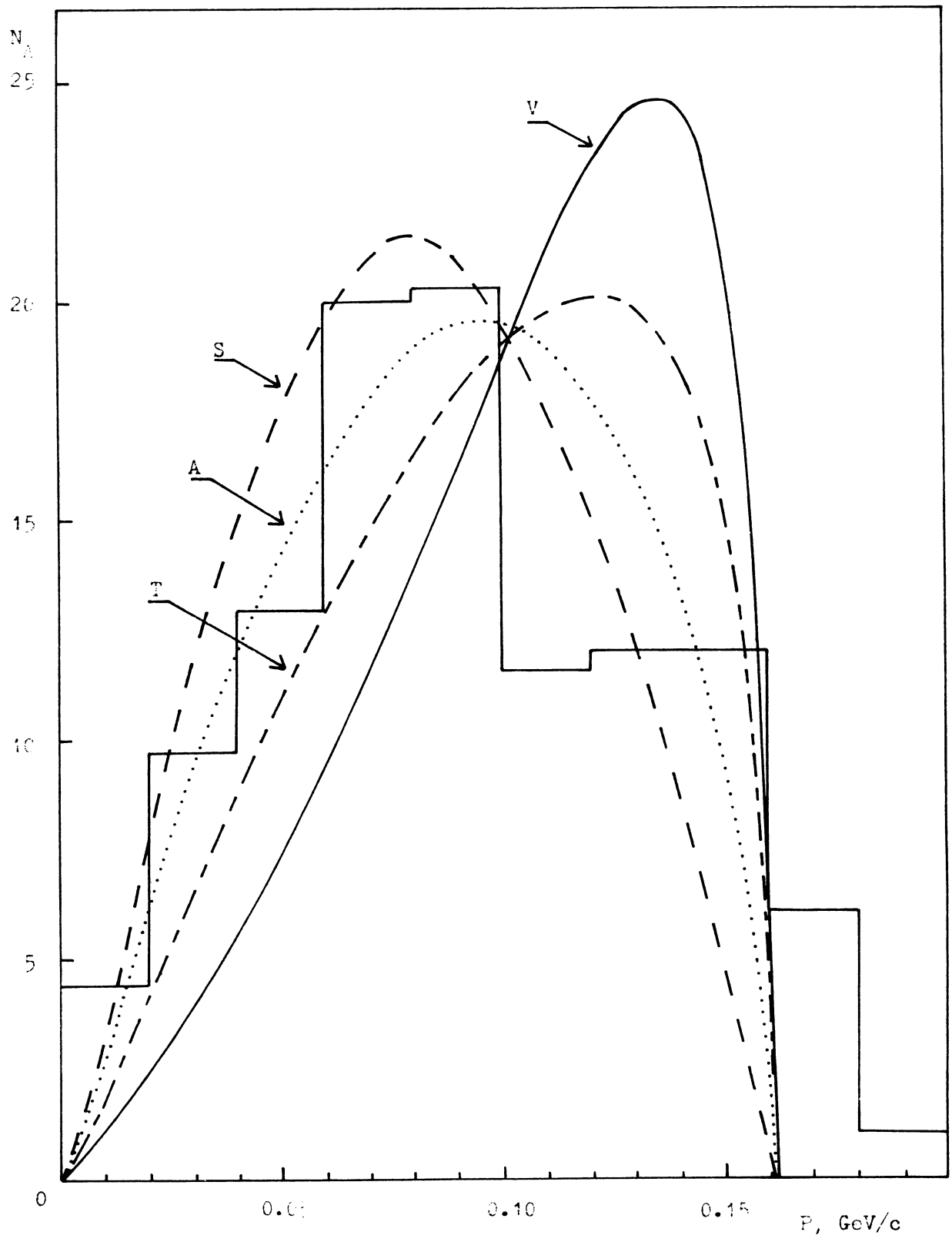


Fig. 6

The four curves show different theoretical transverse momentum spectra of the proton for Lambda hyperons decaying to electrons. The histogram gives the observed distribution. (Physics Letters 6, 185 (1963)).

There are different forms of interaction (V = vector, S = scalar, A = axial vector and T = tensor), which might be responsible for the decay. We give the transverse momentum spectrum of the proton ($P_\perp = P_p \sin \Theta_{\Lambda p}$) derived from the four theories in Fig. 6. We have also given the observed spectrum, which is corrected for detection efficiency. To make the following tests easier, we assume that the histogram contains observed events only (110 in number).

There are some events in Fig. 6 which have transverse momenta above the theoretically maximum possible value. The reason for this is of course uncertainties in the measured variables. One should therefore, in a treatment which is more strict than that below, take these uncertainties into consideration. One possibility is, for example, to modify the theoretical curves in such a way that the uncertainties in the measured variables will be included.

2) Test of a mean value

The mean value ($<P_\perp>$) of the theoretical distributions takes the following values:

V	106	MeV/c
T	98	"
A	89	"
S	81	"

The measured spectrum has the mean value $(P_\perp)_{exp} = 91$ MeV/c.

The theoretical $P_\perp$ distributions are not normal, but the distribution of the mean value of a sample will in most cases approach the normal distribution when the sample size increases. This is, for example, the case, if the population has finite variance (σ_p^2) and mean. The mean value distribution will have the same mean as the population and the variance σ_p^2/n. The standard deviations of the four theoretical distributions are about the same (34 - 38 MeV/c), which means that the mean value distribution will have a standard deviation (σ_{mv}) of about 3.5 MeV/c (n = 110).

The deviation between the theoretical and observed mean values is then:

V	4.3	σ_{mv}
T	2.0	"
A	0.6	"
S	2.9	"

The probability that $\langle P_\perp \rangle_{exp}$ will deviate by more than $\alpha.\sigma_{mv}$ from the mean value of the population is of course:

$$Pr = \frac{2}{\sqrt{2\pi}} \int_{\alpha}^{\infty} e^{-\frac{(P_\perp - \langle P_\perp \rangle)^2}{\sigma^2_{mv}}} dP_\perp \ . \qquad (57)$$

This probability can be found in most statistical tables. In our case we have:

V	<	0.01%
T		4.6%
A		55 %
S		0.4%

The level of significance for a test is arbitrarily fixed at, for example, 5%, 1% or in the more stringent tests 0.1%. Let us fix the significance level at 1%. We then conclude that the pure V and S hypotheses are apparently disproved by the experiment, while the A and T hypotheses are reasonably practical interpretations of the measured mean value.

3) χ^2 test of distributions

This test (ref. D. Hudson, CERN 63-29) is used when one wants to find out how well the whole sample distribution agrees with a known population distribution. It is therefore more general than the test described above.

The distributions are divided in class intervals of the variable ($P_\perp$ in our case). The number (N_i) of observed events in an interval i, is compared with the expected number (N_i^h) from the given hypothesis. The following quantity is calculated:

$$\chi^2 = \sum_{i=1}^{r} \frac{(N_i - N_i^h)^2}{N_i^h} \quad . \tag{58}$$

The distribution of χ^2 will approach the χ^2 distribution given in Chapter IV.3 when the size (n) of the sample increases. One may use the limiting distribution if the size of the sample is so large that each N_i^h is greater than 10.

$P_\perp$ MeV/c	N_i	N_i^h			
		V	T	A	S
0 - 40	14.1	5.7	9.1	12.8	15.7
40 - 80	32.8	20.0	26.6	32.5	39.3
80 - 120	31.9	38.5	38.6	39.3	39.3
> 120	31.2	45.8	35.7	25.4	15.7
χ^2		26.4	6.0	2.8	18.0

Table 2

In our case we divide the distributions into four classes. The expected and observed number of events in each interval is given in Table 2 together with the χ^2 values calculated with the formula above.

The number of degrees of freedom is r- 1 (3 in our case). This follows from the fact that one can only choose r- 1 of the N_i^h values independently when the total number of events in the distribution is fixed.

The probabilities ($P_3\%$) to obtain χ^2 values greater than given values (χ_0^2) are given in Table 3 for 3 degrees of freedom (see Chapter IV.3).

χ_0^2	2.4	3.7	4.6	6.3	7.8	9.8	11.3	16.2
$P_3\%$	50	30	20	10	5	2	1	0.1

Table 3

We conclude that our experimental data are not consistent with hypotheses V and S even if we select the significance level to be as low as 0.1%. The agreement is, however, good between both of the hypotheses A and T and the observed distribution.

The χ^2 test may be used even when some parameters of the population distribution are unknown. In this case one has first to estimate the unknown parameters with the help of the experimental distribution. The number of degrees of freedom will now be r- 1 - a, where a is the number of unknown parameters in the population distribution.

Acknowledgement

I wish to thank Dr. F. Müller for reading this article and making many valuable suggestions.

7863/p/cm

APPENDIX I

DESCRIPTION OF THE SUBROUTINE BOECK

BOECK is a FORTRAN subroutine, which performs needed matrix operations for one step of the best fit calculations.

Input quantities:

NF = Number of constraint equations.
NM = Number of measured variables.
NX = Number of unknown variables.
F(NF) = Constraint equations.
B(NF,NM) = Derivative matrix of measured variables.
A(NF,NX) = Derivative matrix of unknown variables.
GI(NM,NM) = Error matrix of measured variables.

Output quantities:

C(NM) = Corrections of measured quantities; $m^{\nu+1} = m^{0} + C$ (C is put equal to 0 before the first step).
DX(NX) = Corrections of unknown variables; $x^{\nu+1} = x^{\nu} + DX$.
GMFI(NM,NM) = Error matrix of fitted measured variables.
GXI(NX,NX) = Error matrix of unmeasured variables.
B(NM,NX) = Correlation matrix between measured and unmeasured variables.
CHI = χ^2 for the last step.

Some further remarks:

1. The dimensions must be at least as big as those given within the brackets.
2. H(NM,NM), HH(NM,NM), R(NF) and AL(NF) are working matrices.
3. Some of the input matrices are re-defined in later stages of the programme and thus used as working matrices.
4. One must have NM ≥ NF ≥ NX

5. A 2 dimensional array A(1,1) ... A(m,n) is in FORTRAN stored column-wise, i.e. in order A(1,1), A(2,1), ... A(m,1), A(1,2), A(2,2), ... A(m,2), ... A(m,n). It must for this programme be stored row-wise, i.e. in order A(1,1), A(1,2), ... A(1,n), A(2,1), A(2,2), ... A(2,n), ... A(m,n).

6. K is put equal to 0 as long as a further step in the iteration process is needed and equal to 1 when the iteration is finished and the error matrices have to be evaluated.

7. Equations and notations from Böck (CERN 60-30).

8. All matrix operations are performed by a series of subroutines collected under the name MXPACK and described in the GRIND manual. From this we copy the introduction and also descriptions of subroutines used in BOECK.

MXPACK is a summary name for various subroutines written in 709 FAP language for the fast execution of matrix operations. All entry names start with the letter pair MX. All routines assume that matrices are stored row-wise and without gaps (this is important in case of FORTRAN double indexing) in FORTRAN order (i.e. in absolute descending locations).

Description of subroutines:

MXEQU (A,B,I,J) "Matrix Equations"
solves A(I,I) × X(I,J) = B(I,J) for X.
The result (X) is stored into B, A is transformed.
A is assumed to be positive definite. No pivoting is done.

MXUTY (A,I) "Matrix Unity"
writes a (I,I) Unit matrix into A.

MXADD (A,B,C,I,J) "Matrix Addition"
effects A(I,J) + B(I,J) → C(I,J)

MXSUB (A,B,C,I,J) "Matrix Subtraction"
effects A(I,J) - B(I,J) → C(I,J)

MXTRA (A,B,C,I,J) "Matrix Transfer"
effects A(I,J) → C(I,J). The B-address is irrelevant.

MXMTR (A,B,C,I,J) "Matrix Multiplied Transfer"
effects B × A(I,J) → C(I,J)
B is a scalar factor.

MXMPY (A,B,C,I,J,K) "Matrix Multiplication"
effects A(I,J) × B(J,K) → C(I,K)

The following entries use the same sequence of arguments:

MXMPY1 for $A \times B^T \rightarrow C$ (B is a (K,J) matrix)
MXMPY2 for $A^T \times B \rightarrow C$ (A is a (J,I) matrix)
MXMAD for $A \times B + C \rightarrow C$ "Matrix Multiplication and Addition".

```
C     MATRIX OPERATIONS FOR ONE STEP OF LEAST SQUARES CALCULATIONS
      SUBROUTINE BOECK(K)
      DIMENSION F(3),B(12),A(3),GI(16),C(4),DX(1),GMFI(16),GXI(1),
     1H(16),HH(16),R(3),AL(3)
      COMMON NF,NM,NX,F,B,A,GI,C,DX,GMFI,GXI,CHI
      IF(K) 2,1,2
    1 CALL MXMPY(B,C,H,NF,NM,1)
      CALL MXSUB(F,H,R,NF,1)
C               R IS DEFINED IN (15')
      CALL MXMPY1(GI,B,HH,NM,NM,NF)
      CALL MXMPY(B,HH,H,NF,NM,NF)
C               H IS EQUAL TO S, DEFINED IN (16')
      CALL MXUTY(GMFI,NF)
      CALL MXEQU(H,GMFI,NF,NF)
C               GMFI IS THE INVERSION OF S
      CALL MXMPY(GMFI,A,H,NF,NF,NX)
      CALL MXMPY2(A,H,HH,NX,NF,NX)
      CALL MXUTY(GXI,NX)
      CALL MXEQU(HH,GXI,NX,NX)
      CALL MXMPY2(H,R,C,NX,NF,1)
      CALL MXMPY(GXI,C,DX,NX,NX,1)
      CALL MXMTR(DX,-1.,DX,NX,1)
C               DX IS EQUAL TO THE SECOND EXPRESSION IN (25)
      CALL MXMAD(A,DX,R,NF,NX,1)
      CALL MXMPY(GMFI,R,AL,NF,NF,1)
C               AL IS THE SOLUTION (24)
      CALL MXMPY2(AL,R,CHI,1,NF,1)
      CALL MXMPY2(B,AL,H,NM,NF,1)
      CALL MXMPY(GI,H,C,NM,NM,1)
      CALL MXMTR(C,-1.,C,NM,1)
C               C IS EQUAL TO THE SECOND EXPRESSION IN (12')
      RETURN
C               CALCULATION OF ERROR MATRICES (34-36)
    2 CALL MXMPY2(A,GMFI,H,NX,NF,NF)
      CALL MXTRA(H,0,A,NX,NF)
      CALL MXMPY(B,GI,H,NF,NM,NM)
      CALL MXTRA(H,0,B,NF,NM)
      CALL MXMPY(GMFI,B,H,NF,NF,NM)
      CALL MXMPY2(B,H,HH,NM,NF,NM)
      CALL MXSUB(GI,HH,GMFI,NM,NM)
      CALL MXMPY(A,B,H,NX,NF,NM)
      CALL MXTRA(H,0,HH,NX,NM)
      CALL MXMPY2(HH,GXI,B,NM,NX,NX)
      CALL MXMTR(B,-1.,B,NM,NX)
      CALL MXMPY(B,HH,H,NM,NX,NM)
      CALL MXSUB(GMFI,H,HH,NM,NM)
      CALL MXTRA(HH,0,GMFI,NM,NM)
      RETURN
      END
```

APPENDIX II

GRAPHICAL METHODS OF TESTING THE PROBABILITY DISTRIBUTION A SAMPLE IS THOUGHT TO HAVE COME FROM

Contributed by D. Hudson,
Data and Documents Division, CERN

Suppose we have arranged a sample of size n in increasing order of magnitude to get

$$x_1 < x_2 < \ldots < x_n \tag{1}$$

and we wish to test whether it comes from a continuous distribution with probability density function f(x), and cumulative distribution

$$F(x) = \int_{-\infty}^{x} f(t)\, dt \ . \tag{2}$$

In the first instance it is assumed that f(x) is completely known. Later on the technique is extended to the case where a location parameter μ and a scale parameter σ are unknown, and only the general form $f(x; \mu,\sigma)$ is given.

Example (i)

The χ^2_2 distribution has density function

$$\begin{aligned} f(x) &= 0 \qquad\quad , \quad -\infty < x < 0 \ , \\ f(x) &= \tfrac{1}{2}\, e^{-x/2} \quad , \quad 0 \leq x < \infty \ . \end{aligned} \tag{3}$$

The cumulative distribution is

$$\begin{aligned} F(x) &= 0 \qquad\quad , \quad -\infty < x < 0 \ , \\ F(x) &= 1 - e^{-x/2} \quad , \quad 0 \leq x < \infty \ . \end{aligned} \tag{4}$$

Example (ii)

The normal distribution has density function

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-(x-\mu)^2/2\sigma^2} \quad , \quad -\infty < x < \infty \tag{5}$$

and so has to be treated by the second method when μ and σ are unknown.

Large samples

When the sample size is large it is possible to draw a histogram and compare the resulting diagram with the expected frequency curve. On page 104 there is a histogram of a sample of size 80 from a χ^2_2 distribution, together with the curve

$$N(x) = 80 \times \tfrac{1}{2}\, e^{-x/2} \qquad (6)$$

Small samples

It can be difficult to decide whether the 'bumps' and 'holes' in a histogram are significant, especially if the sample size is small. A rough guide can be obtained by using the graph of the cumulative distribution $F(x)$. We may interpret $F(x)$ as follows:

Given a number x, the proportion of random variables X in a very large (actually infinite) sample for which $X \leq x$ is $F(x)$. In a finite sample of size n, we can estimate points on the curve $F(x)$ by calculating proportions of the sample less than or equal to x.

The n points $x_1, x_2, \ldots, x_n$ give rise to $(n+1)$ intervals on the x axis, viz.

$$(-\infty, x_1)\ ,\ (x_1, x_2)\ ,\ \ldots\ ,\ (x_{n-1}, x_n)\ ,\ (x_n, \infty) \qquad (7)$$

and it is conventional to say that $1/(n+1)$ of the sample lies in each such interval. Thus $1/(n+1)$ of the sample is $\leq x_1$, so $1/(n+1)$ estimates $F(x_1)$; $2/(n+1)$ of the sample is $< x_2$; ..... ; and $n/(n+1)$ estimates $F(x_n)$. [If we add the points $x_0 = -\infty$ and $x_{n+1} = \infty$ to the sample, then $0/(n+1)$ and $(n+1)/(n+1)$ are the correct values of $F(-\infty)$ and $F(+\infty)$ respectively]. A useful convention therefore is that the graph of $F(x)$ is to be estimated by plotting the n points

$$[x,\ F(x)] = \left[x_i\ ,\ \frac{i}{n+1}\right]\ ,\ i = 1, 2, \ldots, n\ . \qquad (8)$$

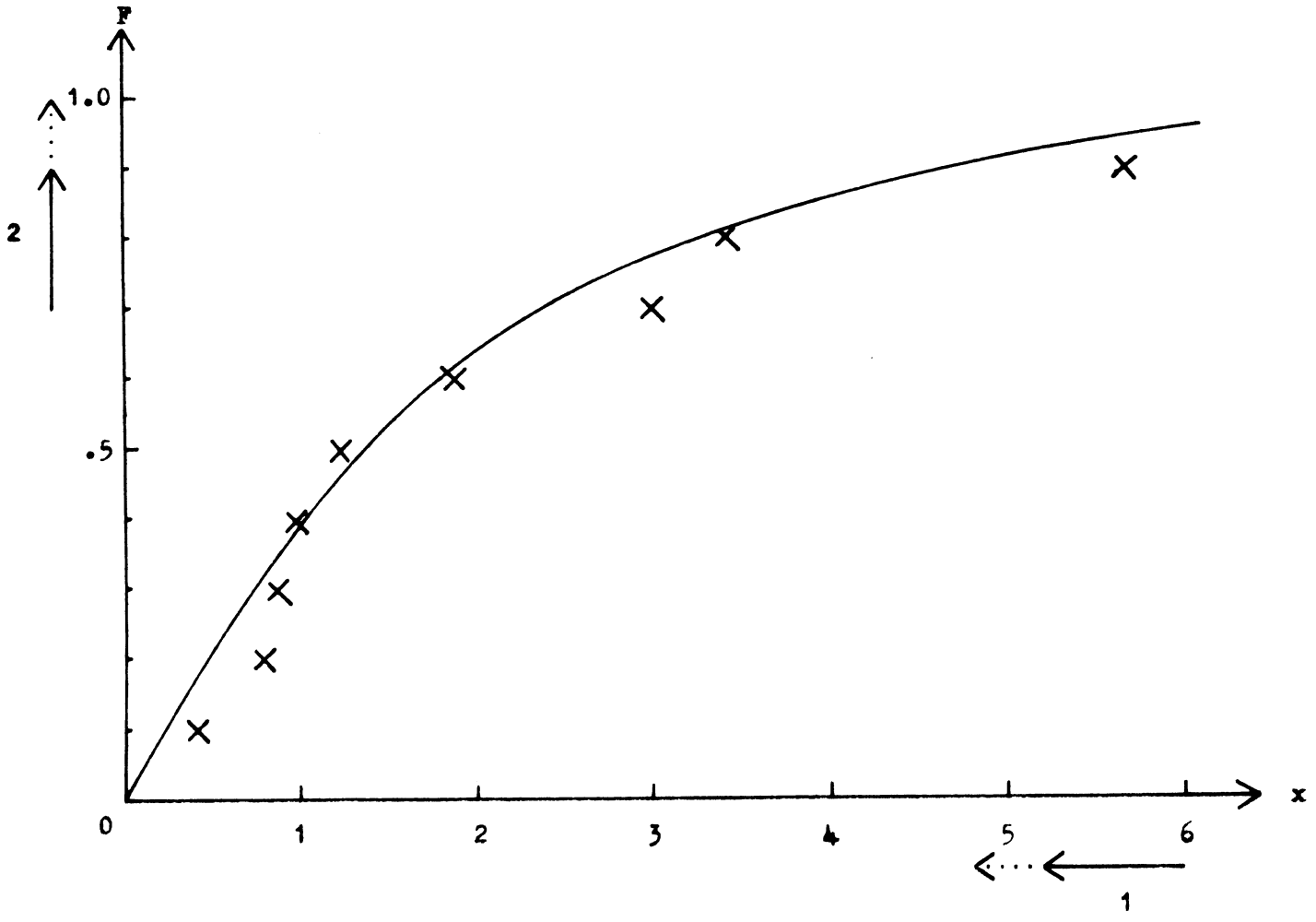
Example

The following is a sample of size nine from the χ^2_2 distribution.

.40, .78, .86, .96, 1.21, 1.86, 2.96, 3.41, 5.68

The points (.40, .1), (.78, .2), etc., are plotted next to the curve

$$F(x) = 1 - e^{-x/2} \tag{9}$$



Probability paper

We next show how to straighten out the curve of F(x) on the above graph so that the plotted points may be compared with a straight line. There are two alternative methods when f(x) is known completely.

(i) We transform the (horizontal) x axis to a new variable z = z(x) in such a way that we 'contract' the x axis most where the graph of F(x) is flat, as indicated by ①←. The transformation is

$$z = \int_{-\infty}^{x} f(t)\, dt \ . \tag{10}$$

The curve Eq. (9) becomes the straight line

$$F(z) = z \ . \tag{11}$$

(ii) We may alternatively transform the (vertical) F axis to a new variable y = y(F) in such a way that we 'stretch' the F axis most where the graph of F(x) is flat, as indicated by ②↑. The transformation is

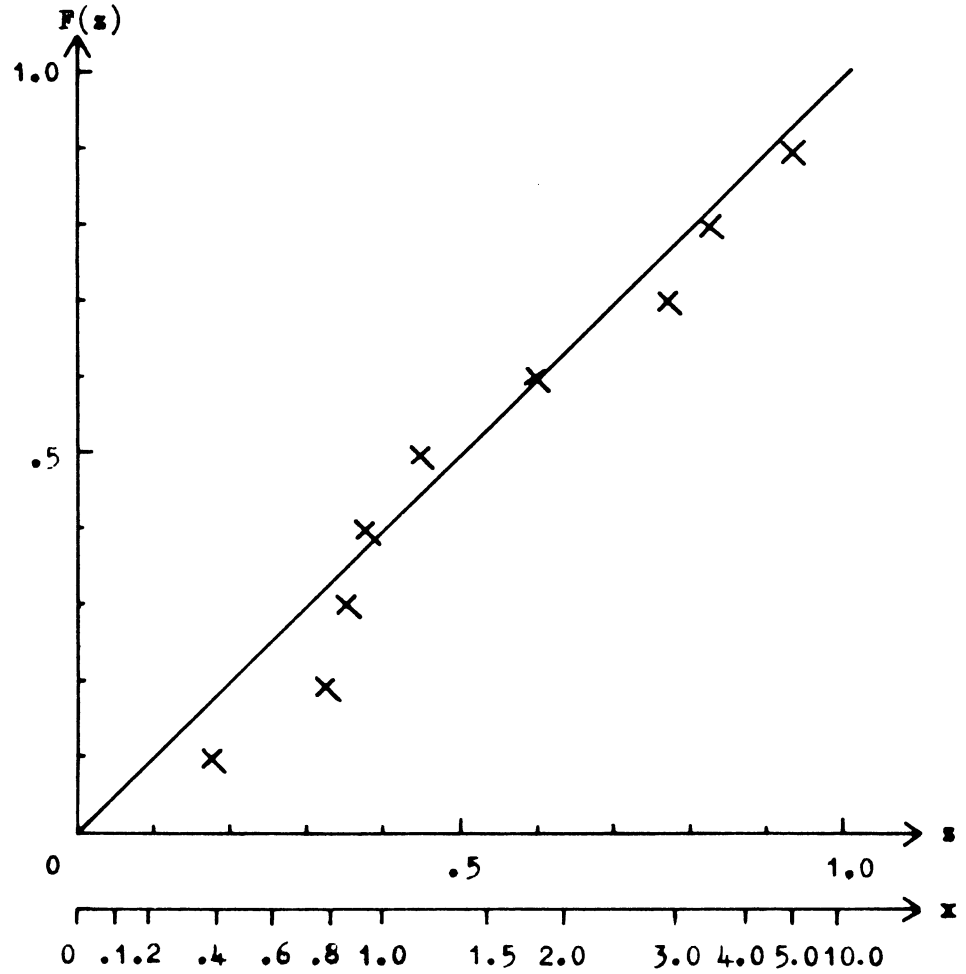
$$F = \int_{-\infty}^{y(F)} f(x)\, dx \ . \tag{12}$$

The curve Eq. (9) becomes the straight line

$$y = x \ . \tag{13}$$

The table required for graph (i) is completed by using 'Table 7: Probability integral of the χ^2_2 distribution and the cumulative sum of the Poisson distribution' in the Biometrika Tables.

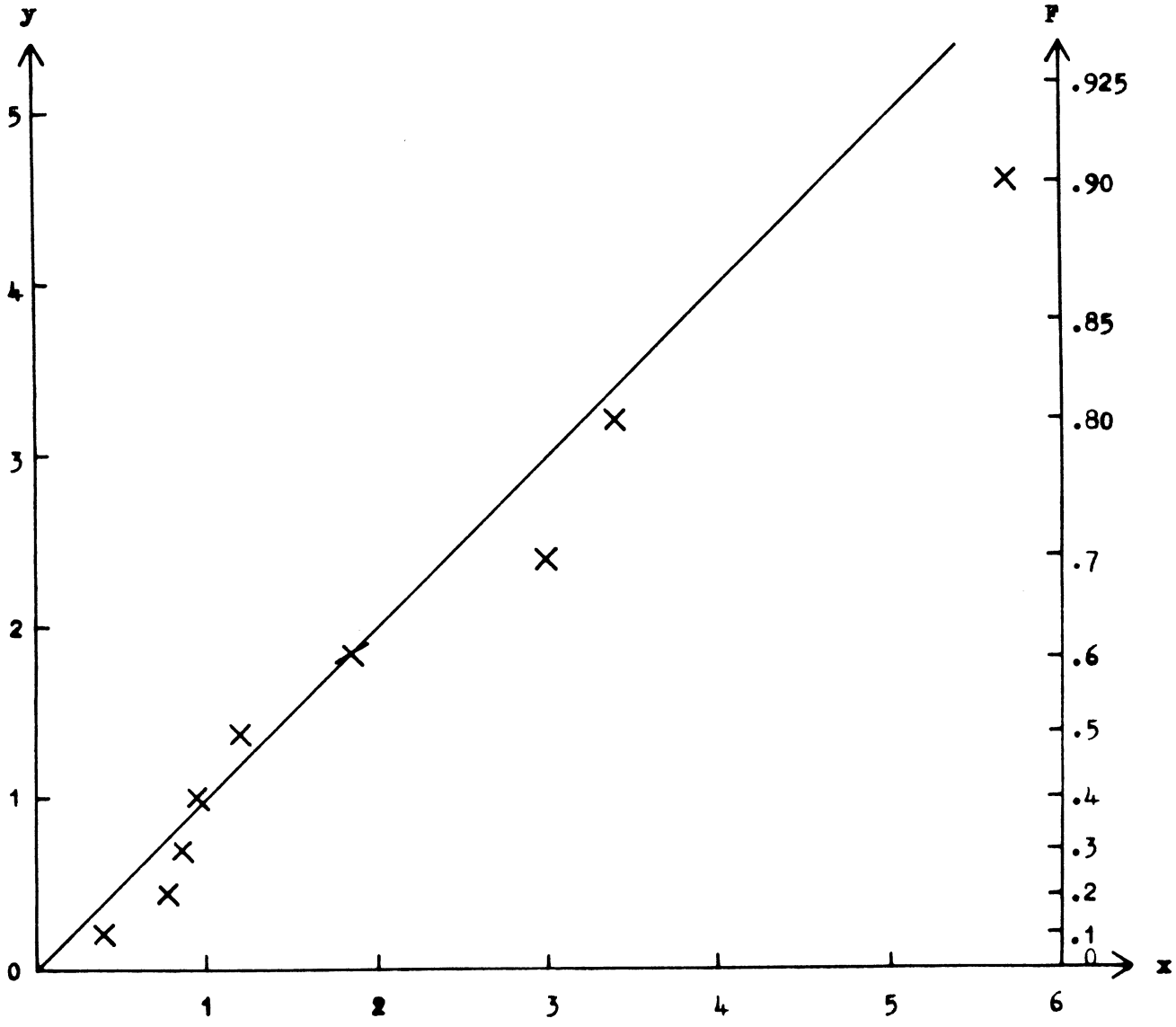
$x = \chi^2_2$	.40	.78	.86	.96	1.21	1.86	2.96	3.41	5.68
z	.18	.32	.35	.38	.45	.60	.77	.82	.94
F(z)	.1	.2	.3	.4	.5	.6	.7	.8	.9



The words 'probability paper' would refer to the above graph paper with only the x and the F scales printed. It would then not be necessary to look up the z values.

The table required for graph (ii) is completed similarly. (An augmented table of percentage points of the χ^2 distribution has been compiled at CERN, using the library tape subroutines referred to in the GRIND manual).

$x = \chi^2_2$	.40	.78	.86	.96	1.21	1.86	2.96	3.41	5.68
F	.1	.2	.3	.4	.5	.6	.7	.8	.9
y(F)	.21	.45	.71	1.02	1.39	1.83	2.41	3.22	4.61



The words 'probability paper' would refer to the above graph paper with only the x and the F scales printed. It would then not be necessary to look up the y values.

Probability paper of type (ii) is used more often in practice. The plotted points will be very bunched at parts of the graph where the original curve of F(x) is steep. Conversely, the data corresponding to 'tail areas' is well spread out. Usually we are particularly interested in these extreme values.

Further example

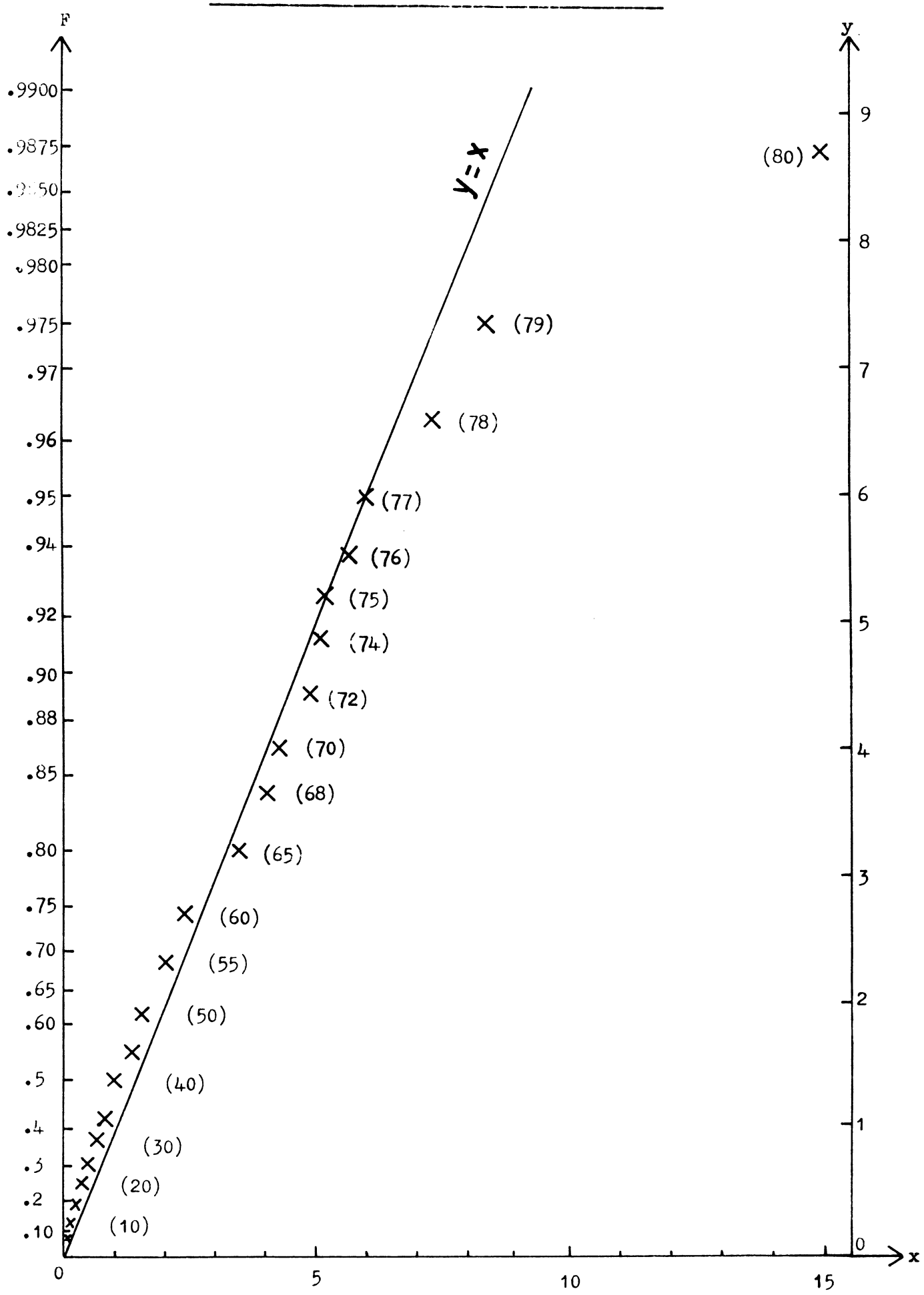
We use probability paper of type (ii) to plot the sample of size 80 on page 104. We wish to do a rough visual test to see whether the sample can reasonably be considered to have come from the X_2^2 distribution. The plotted points are given below. (In order not to crowd the graph, not every possible point is plotted.)

i	5	10	15	20	25	30	35	40	45	50
$x_i = X_2^2$	.08	.18	.27	.37	.49	.64	.82	.98	1.31	1.53
$F_i = i/81$	.06	.12	.18	.25	.31	.37	.43	.49	.56	.62

i	55	60	65	68	70	72	74	75	76
x_i	2.05	2.44	3.46	4.05	4.23	4.95	5.14	5.19	5.65
F_i	.68	.74	.802	.840	.865	.890	.914	.925	.938

i	77	78	79	80
x_i	5.99	7.37	8.32	14.98
F_i	.950	.963	.975	.987

Sample of size 80 from the X_2^2 distribution



Conclusions from the graph

The graph indicates that

(a) on the whole the data follow the χ^2_2 distribution fairly well;

(b) the smaller values (i = 1 to 60) tend to be a bit smaller than expected;

(c) the largest value x_{80} appears to be too large.

Exact test

We can find the exact level of significance of $x_{80} = x_{max}$.

We have $\Pr(x \leq k) = F(k)$

Let $\Pr(x_{max} \leq k) = H(k)$

Then $H(k) = \Pr(\text{All } x_i \leq k)$

$= [F(k)]^{80}$

So $H(14.98) = [F(14.98)]^{80}$

$= (1 - e^{-7.49})^{80}$

$= 1 - 0.044$

$\therefore \Pr(x_{max} \geq 14.98) = 4.4\%$.

The conclusion is that x_{max} is significantly larger than we expect the largest value of χ^2_2 to be in a sample of size 80. The value $x_{80} = 14.98$ may indicate that a different physical hypothesis is required for the relevant event.

Case of unknown parameters

Suppose that the probability density function of x, $f(x; \mu, \sigma)$, contains an unknown location parameter μ and an unknown scale parameter σ such that

$$w = \frac{x - \mu}{\sigma} \tag{14}$$

is a standardized variable with density function $f(w; 0, 1)$ which is completely known.

Let $$F(w) = \int_{-\infty}^{w} f(t;\ 0,\ 1)\ dt \qquad (15)$$

be the standardized cumulative distribution. We can prepare probability paper of type (ii) by making the transformation from the F scale to the y scale

$$F = \int_{-\infty}^{y(F)} f(t;\ 0,\ 1)\ dt\ . \qquad (16)$$

Comparison with Eq. (15) shows that the curve F(w) becomes the straight line

$$y = w\ . \qquad (17)$$

From Eq. (14) we see that if we measure x on the horizontal axis, Eq. (17) becomes the straight line

$$y = \frac{x-\mu}{\sigma}\ . \qquad (18)$$

Thus, when $y = 0$, $x = \mu$;
when $y = 1$, $x = \mu + \sigma$;
when $y = -1$, $x = \mu - \sigma$;

so the parameters can be read off the graph.

Example

A sample of size 7 is available and we wish to do a rough test to see whether they come from a normal distribution. The density function for x is

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}\ .\ e^{-(x-\mu)^2/2\sigma^2}$$

and the standard density function for w is

$$f(w) = \frac{1}{\sqrt{2\pi}}\ e^{-w^2/2}\ .$$

The required transformation from the F to the y scale is

$$F = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{y(F)} e^{-w^2/2}\, dw$$

which is tabulated in the Biometrika Tables.

The data are given below in increasing order of magnitude.

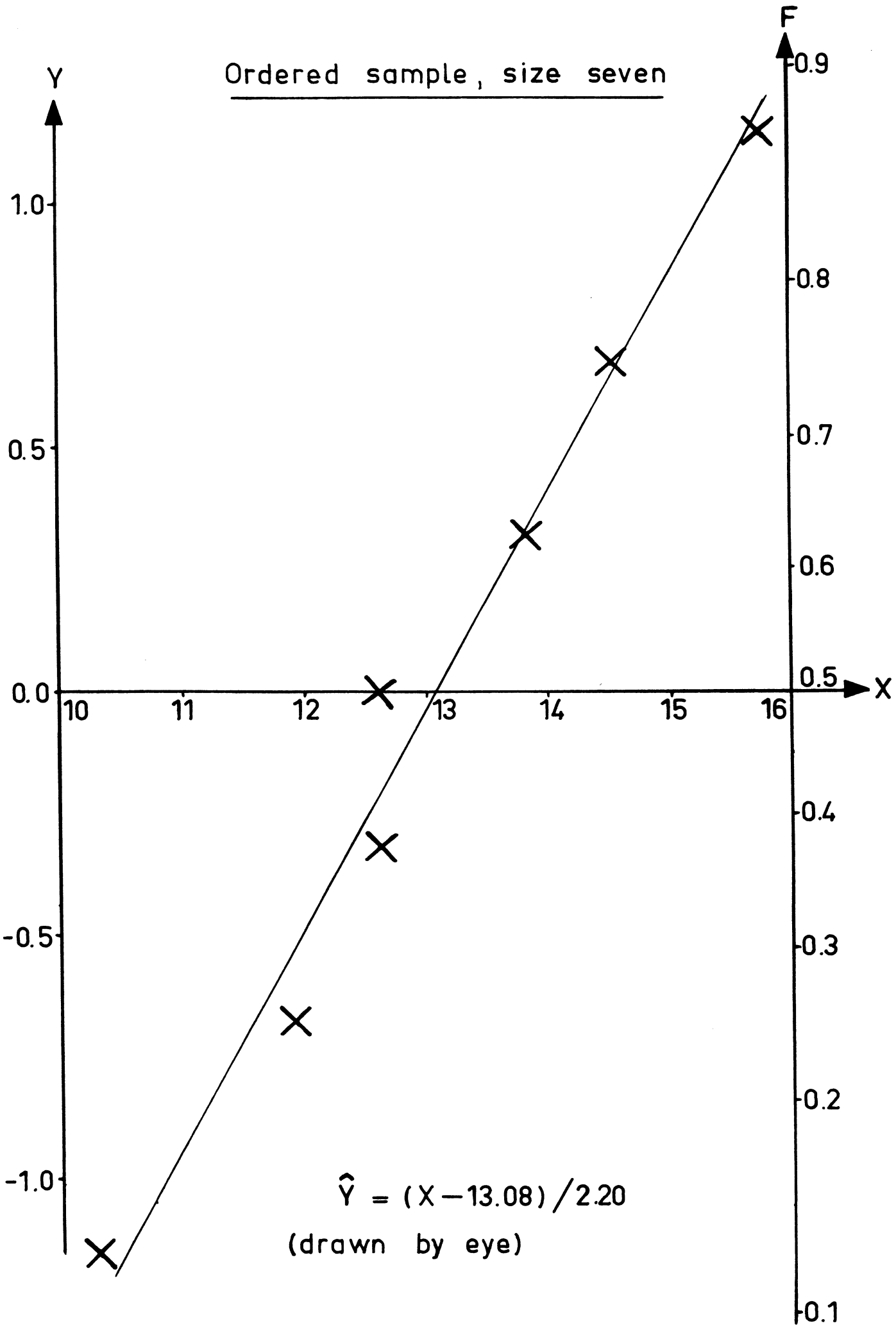
i	1	2	3	4	5	6	7
x_i	10.3	11.9	12.6	12.6	13.8	14.5	15.7
F_i	.125	.250	.375	.500	.625	.750	.875
y_i	-1.150	-.674	-.319	0	.319	.674	1.150

The 'normal probability plot' of the 'order statistics' is shown on page 144. The linearity shown by the plot is a rough assurance that the data are normally distributed. A line drawn by eye through the points gives the estimates

$$\mu = 13.08 \quad , \quad \sigma = 2.2 \ .$$

(The maximum likelihood estimates are

$$\hat{\mu} = 13.06 \quad , \quad \hat{\sigma} = 1.8 \ ).$$


Ordered sample, size seven
Y
X
F
1.0
0.5
0.0
-0.5
-1.0
10
11
12
13
14
15
16
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
$\hat{Y} = (X - 13.08)/2.20$
(drawn by eye)

REFERENCES

1) W. Dixon and F. Massey, Chapter 5.5 'Normal probability paper' and Chapter 13.7 'Binomial probability paper' in "Introduction to statistical analysis" - McGraw-Hill, 1957.

2) H. Chernoff and G. Lieberman, 'Use of normal probability paper' - J.Amer.Statist.Assn. 49, p. 778 (1954).

3) E. Crow, F. Davis and M. Maxfield, Chapter 4 'Tests of distributions as a whole' in "Statistics manual" - Dover, 1960.

4) E. Pearson and H. Hartley, "Biometrika tables for statisticians" - Cambridge U.P., 1958.

5) M. Kendall and A. Stuart, Chapter 14 'Order statistics' in "The advanced theory of statistics, Vol. I" - Charles Griffin, 1958.