ORGANISATION EUROPÉENNE POUR LA RECHERCHE NUCLÉAIRE
CERN
EUROPEAN ORGANIZATION FOR NUCLEAR RESEARCH

PROGRAMMING FOR FLYING SPOT DEVICES

A conference held at
The Centro Nazionale Analisi Fotogrammi, I.N.F.N., Bologna,
on 7-9 October, 1964.

(organized jointly by the Centro Nazionale Analisi Fotogrammi,
I.N.F.N., Bologna and the Data Handling Division, CERN, Geneva).

PROCEEDINGS
edited by
W.G. Moorhead
B.W. Powell

GENEVA
1965
Propriété littéraire et scientifique réservée pour tous les pays du monde. Ce document ne peut être reproduit ou traduit en tout ou en partie sans l’autorisation écrite du Directeur général du CERN, titulaire du droit d’auteur. Dans les cas appropriés, et s’il s’agit d’utiliser le document à des fins non commerciales, cette autorisation sera volontiers accordée.

Le CERN ne revendique pas la propriété des inventions brevetables et dessins ou modèles susceptibles de dépôt qui pourraient être décrits dans le présent document ; ceux-ci peuvent être librement utilisés par les instituts de recherche, les industriels et autres intéressés. Cependant, le CERN se réserve le droit de s’opposer à toute revendication qu’un usager pourrait faire de la propriété scientifique ou industrielle de toute invention et tout dessin ou modèle décrits dans le présent document.

© Copyright CERN, Genève, 1965

Literary and scientific copyrights reserved in all countries of the world. This report, or any part of it, may not be reprinted or translated without written permission of the copyright holder, the Director-General of CERN. However, permission will be freely granted for appropriate non-commercial use. If any patentable invention or registrable design is described in the report, CERN makes no claim to property rights in it but offers it for the free use of research institutions, manufacturers and others. CERN, however, may oppose any attempt by a user to claim any proprietary or patent rights in such inventions or designs as may be described in the present document.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preface</td>
<td>v</td>
</tr>
<tr>
<td>Acknowledgements</td>
<td>vii</td>
</tr>
<tr>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>Introductory remarks</td>
<td>3</td>
</tr>
<tr>
<td>Experience with the SPASS system</td>
<td>7</td>
</tr>
<tr>
<td>DISCUSSION</td>
<td>13</td>
</tr>
<tr>
<td>List processing techniques in the automatic analysis of spark chamber data</td>
<td>15</td>
</tr>
<tr>
<td>DISCUSSION</td>
<td>23</td>
</tr>
<tr>
<td>The automatic analysis of 200,000 spark chamber pictures using the CERN HPD</td>
<td>25</td>
</tr>
<tr>
<td>A program for calibrating the flying spot cathode ray tube for CHLOE</td>
<td>35</td>
</tr>
<tr>
<td>DISCUSSION</td>
<td>41</td>
</tr>
<tr>
<td>Automatic analysis of photographic data in biology</td>
<td>43</td>
</tr>
<tr>
<td>DISCUSSION</td>
<td>55</td>
</tr>
<tr>
<td>Spark chamber photograph analysis using LUCIOLE</td>
<td>57</td>
</tr>
<tr>
<td>DISCUSSION</td>
<td>69</td>
</tr>
<tr>
<td>Program developments in the CERN FSD spark system</td>
<td>71</td>
</tr>
<tr>
<td>PEPR status report</td>
<td>81</td>
</tr>
<tr>
<td>DISCUSSION</td>
<td>89</td>
</tr>
<tr>
<td>The CERN picture number finding routine</td>
<td>91</td>
</tr>
<tr>
<td>DISCUSSION</td>
<td>Page</td>
</tr>
<tr>
<td>----------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>The status of HPD programs at Brookhaven National Laboratory</td>
<td>95</td>
</tr>
<tr>
<td>K. Abrams</td>
<td>97</td>
</tr>
<tr>
<td>J. Cockrill</td>
<td></td>
</tr>
<tr>
<td>P. Connolly</td>
<td></td>
</tr>
<tr>
<td>D. Crennell</td>
<td></td>
</tr>
<tr>
<td>F.V.C. Hough</td>
<td></td>
</tr>
<tr>
<td>L. Plotter</td>
<td></td>
</tr>
<tr>
<td>D. Stonehill</td>
<td></td>
</tr>
<tr>
<td>R. Strand</td>
<td></td>
</tr>
<tr>
<td>W. Thompson</td>
<td></td>
</tr>
<tr>
<td>DISCUSSION</td>
<td>113</td>
</tr>
<tr>
<td>CERN version of filter</td>
<td>115</td>
</tr>
<tr>
<td>W. Krischer</td>
<td></td>
</tr>
<tr>
<td>W.G. Moorhead</td>
<td></td>
</tr>
<tr>
<td>DISCUSSION</td>
<td>121</td>
</tr>
<tr>
<td>Experience with the CERN HAZE program</td>
<td>123</td>
</tr>
<tr>
<td>J.W. Howie</td>
<td></td>
</tr>
<tr>
<td>DISCUSSION</td>
<td>137</td>
</tr>
<tr>
<td>The post filter routine of CERN HAZE</td>
<td>139</td>
</tr>
<tr>
<td>H. Hannerfeld</td>
<td></td>
</tr>
<tr>
<td>J. Seyerlein</td>
<td></td>
</tr>
<tr>
<td>DISCUSSION</td>
<td>143</td>
</tr>
<tr>
<td>The FORTRAN diagnostic routine used with CERN HAZE</td>
<td>145</td>
</tr>
<tr>
<td>M.L. Luvisetto</td>
<td></td>
</tr>
<tr>
<td>Preliminary report on an &quot;Integral&quot; filter technique for the HPD</td>
<td>153</td>
</tr>
<tr>
<td>N. Webre</td>
<td></td>
</tr>
<tr>
<td>D. Burd</td>
<td></td>
</tr>
<tr>
<td>L. Gardner</td>
<td></td>
</tr>
<tr>
<td>D. Tycko</td>
<td></td>
</tr>
<tr>
<td>B. Zaslavsky</td>
<td></td>
</tr>
<tr>
<td>DISCUSSION</td>
<td>171</td>
</tr>
<tr>
<td>HPD programming at the Rutherford Laboratory</td>
<td>173</td>
</tr>
<tr>
<td>J.W. Burren</td>
<td></td>
</tr>
<tr>
<td>J.W. Gardner</td>
<td></td>
</tr>
<tr>
<td>M.J. Mitchell</td>
<td></td>
</tr>
<tr>
<td>M.C. Turnill</td>
<td></td>
</tr>
</tbody>
</table>
Results obtained with an off-line HPD system

M. Bloch
M. Schiff
C. de la Vaissière

205

DISCUSSION

Control of the HPD by a small computer

Ph. Leblond

207

DISCUSSION

Preliminary results on HPD ionization measurements at CERN

L. Michejda

217

DISCUSSION

HPD programming at the University of Pennsylvania

J.D. Engman

219

DISCUSSION

Current status of the digital automatic scanning programs

R.B. Marr

227

DISCUSSION

Concluding remarks

L. Kowarski

229

List of participants

235

237

257

259

267
A few explanatory remarks are necessary about the preparation of these proceedings. The contributions have been copied from the type-scripts provided by the authors and only the more obvious errors have been corrected. For the figures, we have done our best with the material provided and in some instance redrawn them when reproduction from the original was not desirable. We have also thought it desirable to contract and re-word much of the discussion. Since in the majority of cases speakers have not had any opportunity to check the final version of the text, all inaccuracies should be attributed to us and not to the speakers themselves.

For the report by Deutsch on the status of FEPR, it was agreed to reproduce the text of the paper given by Taft at the 11th International Conference on High Energy Physics held in Dubna in 1964 rather than to use a text based on what was actually said at Bologna.

W.G. Moorhead
B.W. Powell
Acknowledgements

It was a pleasure to collaborate with M. Masetti and the members of the I.N.F.N. group at Bologna in organising this meeting. Due to them the meeting took place informally and with a minimum of difficulties.

We would like to thank the Scientific Secretaries for their work in compiling the text and in checking the stencils and figures.

We are indebted to Mme G. Andréossi for the organising and the execution of all the secretarial work associated with the proceedings. We would also like to thank E. Bissa and D. Boileau for the tape recording of the proceedings and Mlle R. Willibald for her help with the figures.

W.G. Moorhead
B.W. Powell
INTRODUCTION

E. CLEMENTEL
Centro di Calcolo Nucleare, Bologna

On behalf of C.N.E.N. and the Computer Centre of Bologna, I have the privilege and pleasure of welcoming you here to this meeting on "Programming for Flying Spot Devices".

All of you know better than I do, that the development of bubble chamber and spark chamber experiments has introduced to high-energy physics, flying spot devices and high speed computers as a powerful and necessary means of examining data to aid in the conducting of experiments and understanding and analysis of results.

Apart from some technical aspects, which, I understand, have already been discussed during the informal meeting of the last few days, it is also well known that the optimum use of spot devices and computers, from the point of view of speed and computing capacity depends very much on the programming.

Several groups in different laboratories have made a great effort in this direction during the last few years and I see that the representatives of these laboratories are now attending this meeting to exchange and discuss their different experiences, results and philosophies.

Since the very beginning of the founding of the National Centre for Analysis of Photographs here in Bologna we have enjoyed the help and collaboration of several laboratories, especially CERN. I take the opportunity of this welcoming talk to thank all the people who have helped us in overcoming all the difficulties which the establishment of a new national centre implies.
INTRODUCTORY REMARKS

C. FRANZINETTI
CERN, Geneva

Flying spot digitizers have become an essential part of the equipment of high energy physics. This is due essentially to two factors: on the one hand there is the increasing demand for bubble chamber experiments, each involving the scanning of several hundred thousand pictures; on the other hand there is the continuous development of more powerful and faster computers which make that demand acceptable in principle. The connection between these two factors is made by automatic devices for fast and highly precise scanning of the pictures, the information recorded on a picture being translated into a language understandable to computers.

The origin of the demand is to be found in the present stage of high energy nuclear physics. We are exploring the internal structure of matter and we do so by studying the behaviour of particles in their interaction with other particles and the characteristics of the final products of high energy collisions. Many of the properties of these particles are known. Most probably progress will take place in experiments which select comparatively rare events and can detect the detailed structure of the observed distributions. In fact, experiments in which events are to be selected having a relative abundance of $10^{-6}$ with respect to the main competing process have been discussed and are being planned. Determination of moments down to a fraction of a percent are also being discussed.

All these requirements call for faster scanning and higher resolution devices. The use of manpower for scanning and measuring can not match in efficiency (apart from cost and time required) an automatic machine of any of the types already in existence. Already in Europe we are witnessing an increasing interest in automatic measuring devices. A questionnaire
circulated last year in Europe (CERN, DD/CO/FN17) showed that at least six flying spot digitizers were being planned in various European countries, and may be in operation by the end of 1965; and that the needs for computers and automatic instruments for measurement and analysis were likely to progress on a basis of an annual doubling. Surely this progress will meet saturation some day, but probably we are not yet near that moment. The limitation may come from the number of available pictures. In fact, the demand from European laboratories in 1966 is expected to reach 6,000,000 pictures, a figure which is far above the expected output of European laboratories, until new accelerators come into operation. Moreover, larger recording devices have been planned as you all know. At Brookhaven there is a project for a hydrogen bubble chamber of 16 feet diameter and 14 feet depth. At CERN similar projects are also being discussed. This sets new problems for you. Most probably, these large chambers will be viewed by many cameras, each recording a section of the sensitive volume. The reconstruction of an event and its analysis will probably require the simultaneous scanning of a large number of pictures. Big computers with large memories will be required. Fast scanning will become essential.

I dare to say that we shall be in a better position to ask for financial support for more ambitious projects, because the cost of both accelerators and detectors will be so high that also the instrumentation for the analysis will have to be increased in proportion. But the scope of your work is wider than bubble chamber picture scanning, even if that is the main concern now. It extends to spark chambers and to counter experiments with on-line computers. Spark chamber pictures are much easier to analyse. They convey less information than a bubble chamber picture. So both the scanning and the interpretation of the events is simpler.

Spark chambers in their present form will remain a useful instrument only so long as the event to be analysed is geometrically simple, for example elastic scattering, pair creation by γ's, charge exchange scattering. Already high energy γ's set a problem for spark chambers which is better
solved in many cases by using a total absorption Čerenkov counter, or a Čerenkov in combination with a few spark counters.

However, spark chambers are being developed into more sophisticated instruments. We have already seen in the last few years pictures of "track following spark chambers" and it is not improbable that a number of such instruments will come into operation in the near future. Then spark chamber pictures will present the same sort of problems as those of bubble chambers.

Filmless spark chambers can be regarded as counters with high spatial resolution. No picture is taken, and thus no automatic scanning is involved. Both filmless spark chambers and counters are being used with on-line computers, in the United States and also in Europe. The analysis has to take place simultaneously with the recording of the events or later, using the information memorized on magnetic tape. Mention should also be made of the automation in other branches of science. Biologists are beginning to use automatic methods of analysis and automatic methods of counting. At this conference we shall hear a communication on the automatic classification of chromosomes. We know that similar methods have been employed for the analysis and correlation of nerve pulses. Complicated animal functions, such as the mechanism of vision are now being studied in a much deeper way than would have been possible a few years ago before the use of automatic analysis. New fields of research are thus being developed, which would have otherwise been closed to scientists.

In the course of this meeting we shall be hearing a number of very interesting reports on the work in progress on these subjects in European and American laboratories. We would like to thank all the contributors and organisers.
EXPERIENCE WITH THE SPASS SYSTEM

M. DEUTSCH
Massachusetts Institute of Technology, Cambridge

(presented by M. Deutsch)

Most of this audience is probably familiar with the basic features of the SPASS system. It was originally designed to be the poor man's answer to the problem of scanning relatively simple spark chamber pictures. It was never meant to be developed into a general scanning system for bubble chambers or very large and complex spark chambers. In the event, the system turned out to be more versatile than we had expected. It has now operated successfully for about two years. In that time we have successfully analysed—in the sense that results have been published—three experiments and we are working on the fourth.

The main message that I can give you on the basis of this experience is this: possession of a scanning apparatus which is functioning in the technical sense is only a small part of the problem. Let me remind you that SPASS uses the standard display oscilloscope of a PDP-1 computer to interrogate the film. This procedure has the virtue of requiring essentially no special hardware. It provides completely random access to every point on the film but at the cost of slow operation: about 50 μsec per point interrogated. We were therefore forced into a tightly programmed track-following system designed to scan only a small fraction of the picture area. For example, we search only a certain number of gaps for possible track origins, follow each only far enough to determine its relevant characteristics and, whenever possible, abandon a picture as soon as we are certain that it contains no valid event. In this manner we have kept the average scanning and measuring
time at about one second per frame. The time for a valid event may be considerably longer, up to four or five seconds per frame for a difficult situation. The first experiment to which SPASS was applied was ideally suited for this procedure. In three of the four chambers involved only a single track could occur in a valid event and this track had to enter the chamber in a known direction and, for two of the chambers, in a predictable region. We were therefore able to reject immediately any event in which a view did not yield exactly one successful track originating in the proper zone of one of the first two gaps. Since there was a total of 42 gaps in the chambers, the average time saving over a complete search was very large.

As the complexity of the pictures increases, the advantages of the tight feedback tend to diminish. If the number of tracks may be large and if they may originate in a large part of the chamber volume, one does not lose much time by searching the entire chamber. For example, in the case of a gamma ray shower chamber, the average scanning time would probably not increase by more than a factor of two or three if we simply collected all spark co-ordinates.

After having talked this long about the question of scanning speed, I should like to say that this parameter is probably of less importance than we usually think. Even now, more than two years after the first successful operation, the time actually spent in the routine scanning and measuring of the pictures is still small for a new experiment compared with the time spent on developing and trying the program, investigating the film, making mistakes, and doing things over again.

In a tight-feedback system, using a small computer, this development occurs very largely on-line, in contact with the film. We sit in front of the oscilloscope, making changes in calibration or strategy with the on-line typewriter, scanning the same film, the same frame or even the same track many times over again. I believe that this procedure is probably quite efficient in terms of the intellectual effort involved.
but we are in any case forced to use it since the information contained in the output of the scanning program is so limited.

In a system such as the FSD in which most of the information on the film can be transferred to magnetic tape, this effort occurs to a large extent in more usual program development methods and may not appear quite so obviously as part of the "processing time". Nevertheless, I think it unlikely that 1966 will see many installations processing $10^6$ events of various kinds, as has been predicted here, unless processing means only the transfer of the information from film to magnetic tape. Perhaps it would be safer to predict that in two years each working installation will be capable of processing two or three experiments per year, regardless of the number of pictures, perhaps up to nearly $10^6$, with much time still being spent on program development.

The second question usually asked after that of processing rate, refers to the yield or the rejection rate. Here again I believe that our experience may have relevance for all other devices, perhaps more for spark chamber than for bubble chamber pictures.

When we process an entirely new experiment, the rejection of valid pictures is initially likely to be 30%, more or less. We learn very quickly the main causes of these failures, e.g. crossing tracks, multiple sparks, missing fiducials, or whatever special new problems occur in the particular experiment. Within a week or two we make the necessary changes in the program and the yield rises from 70% to, let us say, 85%. This is of course very encouraging. We go through a second round of refinement and get to a yield of 89%. But after this there seems to occur a saturation which makes it extremely hard to raise the yield above some value between 90 and 95%. The reason for this is a multiplicity of minor errors each of which causes the loss of only an occasional event, just as in the case of a human scanner. It will therefore require careful investigation of many events and a great deal of ingenuity to remove this last hard core of errors. I believe that it would be a mistake to do this in a concentrated effort. Every week we recognize and
remedy some little flaw in the program and in due time we shall have encountered most of them. The initial yield for a new experiment will also rise with accumulating experience. In the meantime, I believe, we should avoid automatic scanning of experiments in which a precise knowledge of the absolute yield is essential and ascertain, as far as possible that no relevant bias is introduced by the rejections.

I should now like to say a few words about the impact of the availability of automatic scanning on the production of actual physics. At least until now it has been true that the principal results of a given experiment could be obtained as fast by hand as by machine, but I think that this comparison tends to obscure the two most important contributions of the automatic method.

1. It is generally true that only a small, selected part of the data needs to be evaluated to reach the most important conclusions. It might be said that frequently ten percent of the data yield ninety percent of the results. Obviously a development which greatly reduces the effort required for a large volume of data will affect primarily the "less important" ninety percent. In fact, while the hand-scanning effort is linear with the number of pictures, the automatic scanning requires primarily a development effort which is almost independent of the number of pictures. As with any new technique, its greatest strength does not show itself in cases which can also be done well by older methods but in those which would be prohibitive or not worthwhile.

2. Despite what I have just said, it is possible to obtain preliminary results extremely rapidly, almost "on-line" by the automatic method. While it may take several weeks or even longer to attain a sufficiently low rejection rate for complete processing, reasonable foresight has permitted us to evaluate the performance of our experiments, even during the early part of an accelerator run with a feedback time of less than 24 hours and with sufficiently good yield - certainly
better than 50% - to verify the proper functioning of the apparatus and to determine the number of pictures required. This rapid feedback is of great value in spark chamber experiments. It may be less important for bubble chambers.

What I have just said about "reasonable foresight" brings me to my last point: in the preparation of an experiment the method of data processing must receive as careful consideration as the method of data acquisition. The full advantages of the automatic method cannot be realized - especially point 2, above - unless the experimenter is fully committed to the method. Data acquisition should not begin until the scanning program has been tested on preliminary pictures. I appreciate the pressure of accelerator run deadlines but data may well be obtained faster if one does not consider oneself ready for a run until all parts of the experimental procedure have been tested. We have been able to follow this procedure quite successfully in our own experiments. Our very pleasant and fruitful collaboration in the Brookhaven-Maryland experiment on $K^0$ decay suffered from the fact that hand measurements were carried out at the same time and the physicists most thoroughly committed to the experiment were not in residence at M.I.T.. This reduced the pressure so much that it took well over a year to complete the experiment with complete technical success.

None of these remarks is meant to discourage the processing of the large numbers of "less interesting" pictures remaining from experiments which were originally not planned for automatic processing.

In conclusion, I should like to make some brief remarks which may not have much bearing on the large programs for bubble chamber processing. SPASS was developed as a small-scale effort, to permit us to use visual techniques in a normal university environment, in which an experiment is the affair of a professor, an assistant, a graduate student and a technician. It has worked in that manner, although it
has taken up a little too much of the professor. Although the system is surely capable of technical refinement, which it will receive as the need arises, I do not think that it should be used as a starting point for a much more elaborate system. This attitude has created a staff problem: almost any young engineer or technically minded physicist competent to make a significant contribution to our project would naturally rather work on PEPR. Although this difficulty is a special feature of SPASS, or perhaps of its inventor, it may be generally true that when a device reaches the point where its major difficulties are overcome but some problems remain, a physicist will have to continue to think about it, even though it was built by an engineer.

Added notes:

1. Immediately after the conclusion of the Bologna meeting, $2 \times 10^5$ events remaining from the Brookhaven-Maryland experiment were processed in two weeks by one physicist (not the author) with some help of a graduate student.

2. We have found that the addition of two new instructions to the PDP-1 computer at negligible cost (less than $\$200$) permits us to execute search scans along arbitrary lines at the same speed previously possible only for orthogonal straight lines.
DISCUSSION

FRANZINETTI: With regard to the six million pictures which I mentioned, it should be realised that this is the total number, out of which a more limited number will be analysed for the final results. This reduction in the preliminary analysis was taken into account in deciding whether CERN could meet these requirements.
LIST PROCESSING TECHNIQUES IN THE AUTOMATIC
ANALYSIS OF SPARK CHAMBER DATA

R.K. CLARK
Argonne National Laboratory, Argonne

(presented by R.J. Royston)

The problem which we will discuss in this paper is common to
the processing of data taken from any type of spark chamber, namely, given
a collection of points in Euclidean 3-space which represent spark locations,
how to reconstruct the physical event. We will propose a solution to this
problem which has been realized in the 3600 FORTRAN program, LINK. A sam-
pole of 50 real events, measured on the SCAMPS measuring tables, has been
chosen for test purposes. At the present time approximately 90% of these
events are processed without error. Running time is a function of the
topological complexity of the event and the number of vertices. It avera-
ges 5 seconds per event and has varied from 0.1 seconds for a single
track event with 50 sparks to 30 seconds for a very complex event with
280 sparks. LINK will eventually be a part of the automatic spark
chamber data processing system, AIRWICK, which is being developed at
Argonne.

In this paper we will use some graph-theoretic and list-proces-
sing terminology whose meanings we hope will be intuitively clear. The
references give more detailed information¹,²).

*)

Work performed under the auspices of the U.S. Atomic Energy Commission.
Any event in a spark chamber can be described by a forest of trees (a graph whose connected components are circuit-free) where:

1) The vertices of the graph are given by the points at which the paths of the particles have been recorded (the locations of sparks).

2) The edges of the graph are the unobserved portions of the trajectories between two vertices. An edge can be interpreted as a line in space joining two vertices (sparks).

To solve the reconstruction problem, we must reconstruct the descriptive graph on its known set of vertices by determining the proper edges.

The graph is described by a multi-word list structure which is contained in a FORTRAN array of M x W words, where M is the maximum number of sparks (the maximum number of points allowed in the set of vertices) and W is the number of words allocated to each multi-word list item. One such item will be associated with each spark. In LINK, $M = 300$ and $W = 17$. The seventeen words in each multi-word list item are used as shown in Fig. 1.

The local degree of each vertex, i.e. the number of sparks to which a given spark can be linked, must not exceed seven. In addition to yielding topological information about the graph, the local degree can act as an upper bound on sequencing through a list item to obtain pointers. The only restrictions on the complexity of our graph (and hence on the topology of the event) are imposed by the amount of storage available to this FORTRAN array. In figure 2 we see some of the topological information given by the local degree. If 0, we have an isolated spark; if 1, a track end-point; if 2, a spark interior to a track or the vertex of a "Y"; and if $N \geq 3$, the vertex of an $(N-1)$-pronged "Y" or an $N$-pronged "Y".
A link or edge is constructed between two sparks, I and J, by:

1) increasing the local degrees of each spark by one;
2) inserting J in the first available pointer space of spark I;
3) inserting I in the first available pointer space of spark J.

This bi-directionality enables us to determine all sparks connected to any given spark. Figure 3 shows a possible configuration and its representation in the multi-word list structure.

The basic criterion for establishing links between two sparks is the Euclidean distance between them. This criterion is not sufficient to make many decisions, and we add or remove links on the basis of:

1) the Euclidean distance between sparks;
2) the geometry of the spark chamber;
3) the current stage in the reconstruction process;
4) the number of vertices in the graph;
5) the topological complexity of the descriptive graph;
6) angles formed between adjacent edges;
7) the results from linear and helical extrapolation.

Figure 4 outlines the flow in the current version of LINK. Figure 5 gives the flow in the version which is under development. We will now give a more detailed description of the subroutines.

LINKSTB(I,J) establishes a link between sparks I and J.
ERASE(I,J) removes the link between sparks I and J.
FINDDIST(I,J) determines the distance between sparks I and J.
CANGLER(I,J,K) determines the angle at spark J between the lines IJ and JK.
SUBTREE(I) identifies all sparks in the same branch as spark I.
TREESEARCH(I) identifies all sparks in the same connected component as spark I.
The six subroutines given above are basic building blocks for most of the subroutines described below.

SORT is dependent on the configuration of the chambers, and assigns a gap and chamber number to each spark.

PRELOOK (1) sorts sparks by gap and chamber number, (2) determines the event centroid, (3) determines the basic dimension of the searching cone, (4) determines the linking distance for each chamber, and (5) determines convex hulls for the event and for that portion of the event which is in each chamber. These hulls must be rectangular parallelepipeds whose sides are parallel to the co-ordinate planes.

INTRAGAP links a spark to any spark within the same gap whose distance is within 70% of the linking distance for the chamber.

INTERGAP links a spark whose (gap, chamber) = (i, j) to any spark whose (gap, chamber) = (i + 1, j) and whose distance is within 140% of the linking distance for the chamber.

CLEAN1 executes a smoothing algorithm, erasing links established by INTRAGAP AND INTERGAP when certain conditions of complexity are met. Specifically these are (1) if the local degree is greater than 2, the longest link is removed, and (2) if the local degree is equal to 2, both links are allowed to stand if the two sparks are in the same gap. If these two sparks are linked, the link between them is removed. If the two sparks are in different gaps and one spark is in the same gap as the original spark, this link is removed. If the two sparks are in different gaps and neither is in the same gap as the original spark, both links are allowed to stand.

PASS1, PASS2, and PASSN are concerned with linear extrapolation along a directed line segment. Sparks which are end-points always form one end (called end 2) of the directed segment. PASS1 allows extrapolation in a direction away from the event centroid. PASS2 allows extrapolation in a direction toward the event centroid. PASSN allows extrapolation in any direction. All three subroutines use the
subroutines FINDOUT and EXTRAP. FINDOUT determines which spark is to be used for end 1 of the directed segment. EXTRAP searches along the line determined by this segment, in the direction from end 1 to end 2, collecting all sparks interior to the search cone, and, if possible, linking to one of them. This cone is shown in Fig. 6.

CLEAN4 investigates the angles formed by all edges at points of local degree 2 or more. If the angle is less than 90° and one of the two line segments is longer than 1.4 x L, where L is the linking distance for the chamber, this link is broken. If both links are less than this distance, both are broken.

CLEAN5 removes "hairs," i.e., all links between sparks I and J where the local degree of I is 1 and the local degree of J is ≥ 3. If a link exists between sparks I and J and if an isolated spark K is such that

\[ 90° \leq \text{ANGLE}(I,K,J) \]

the link is broken and links are established between sparks I and K and between sparks K and J.

CLEANUP identifies and sequences through all connected components of the graph.

GRADES investigates the connected component for the existence of circuits. If M is the set of indices corresponding to the sparks in the component, we calculate the circuit-rank.

\[ \text{C}(M) = (\text{number of edges}) - (\text{number of vertices}) + 1 \]

The number of vertices is the cardinality of M. The number of edges is given by

\[ \frac{1}{2} \sum_{I \in M} \text{(local degree of I)} \]
If \( C(M) = 0 \), the component is circuit free. If \( C(M) \geq 1 \), the component contains at least one circuit. These circuits can be removed by using the following algorithm:

If \( G \) is a connected graph with \( n \) vertices and if a length \( \ell(E) \) is defined for each edge, a tree \( T \) may be selected such that \( T \) is a subgraph of \( G \), \( T \) and \( G \) have the same vertex set and \( T \) has the minimum total length:

\[
L = \sum_{E \in T} \ell(E)
\]

\( T \) is selected in the following manner:

1. Select \( E_1 \) such that \( E_1 \) is minimum \( E_j \) of the graph.
2. Select \( E_2 \) such that \( E_2 \) is the minimum remaining \( E_j \) of the graph and \( \{E_1, E_2\} \) contains no circuits.

\[ ............. \]

\( \text{(n-1) Select } E_{n-1} \text{ such that } E_{n-1} \text{ is the minimum remaining } E_j \text{ of the graph and } \{E_1, E_2, \ldots, E_{n-1}\} \text{ contains no circuits.} \]

The tree \( T \) is called the minimal connector tree.\(^3\)

When we have obtained a circuit-free connected component, we make dictionary entries identifying (1) all sparks of local degree 1 (end-points) and (2) all sparks which represent physical vertices. In addition we identify all track segments as indicated in Fig. 7.
References


Figure captions

Fig. 1  Contents of multi-word list item pertaining to a spark.

Fig. 2  Topological information given by the Local Degree.

Fig. 3  A possible configuration and its representation in the multi-word list structure.

Fig. 4  Flow in current version of LINK.

Fig. 5  Flow in version of LINK under development.

Fig. 6  Search cone.

Fig. 7  Identification of track segments.
<table>
<thead>
<tr>
<th>Word</th>
<th>Data Mode</th>
<th>Information Contained in Word</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Floating</td>
<td>X-coordinate of spark *</td>
</tr>
<tr>
<td>2</td>
<td>Floating</td>
<td>Y-coordinate of spark *</td>
</tr>
<tr>
<td>3</td>
<td>Floating</td>
<td>Z-coordinate of spark *</td>
</tr>
<tr>
<td>4</td>
<td>Fixed</td>
<td>Gap Number **</td>
</tr>
<tr>
<td>5</td>
<td>Fixed</td>
<td>Chamber Number **</td>
</tr>
<tr>
<td>6</td>
<td>Fixed</td>
<td>Local Degree of Vertex ***</td>
</tr>
<tr>
<td>7</td>
<td>Fixed</td>
<td>Pointer to 1st connecting spark ****</td>
</tr>
<tr>
<td>8</td>
<td>Fixed</td>
<td>Pointer to 2nd connecting spark ****</td>
</tr>
<tr>
<td>9</td>
<td>Fixed</td>
<td>Pointer to 3rd connecting spark ****</td>
</tr>
<tr>
<td>10</td>
<td>Fixed</td>
<td>Pointer to 4th connecting spark ****</td>
</tr>
<tr>
<td>11</td>
<td>Fixed</td>
<td>Pointer to 5th connecting spark ****</td>
</tr>
<tr>
<td>12</td>
<td>Fixed</td>
<td>Pointer to 6th connecting spark ****</td>
</tr>
<tr>
<td>13</td>
<td>Fixed</td>
<td>Pointer to 7th connecting spark ****</td>
</tr>
<tr>
<td>14</td>
<td>Floating</td>
<td>Distance to closest spark</td>
</tr>
<tr>
<td>15</td>
<td>Fixed</td>
<td>Pointer to closest spark *****</td>
</tr>
<tr>
<td>16</td>
<td>Floating</td>
<td>Distance to 2nd closest spark</td>
</tr>
<tr>
<td>17</td>
<td>Fixed</td>
<td>Pointer to 2nd closest spark ****</td>
</tr>
</tbody>
</table>

* Coordinate information is provided as input to LINK.

** This information is furnished by the subroutine SORT and depends on the configuration of the chamber(s).

*** This value will be N (0 ≤ N ≤ 7). It indicates the number of connected sparks. If non-zero, only the first N pointers will be used. All others will be set to zero.

**** If the contents of the word are zero, all remaining pointers will be set to zero. If non-zero, the value I lies in the range 1 ≤ ISM and is the index of the multi-word list item pertaining to the connected spark.

***** The contents of these words will always contain a number I (1 ≤ ISM) which is the index of the multi-word list item pertaining to the appropriate spark.
Local Degree = 0
Isolated Spark

Local Degree = 1
Endpoint in a Track

Local Degree = 2
Spark in a Track or Vertex of a "V"

Fig. 2
| Spark | 1  | 2  | 3  | 4  | 5  | 6  | 7  | 8  | 9  | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 |
|-------|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
|       | -2.01 | 30.53 | 14.92 | 1  | 2  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 1.12 | 2  | 2.14 | 3  |
| 2     | -2.18 | 29.51 | 15.36 | 2  | 2  | 2  | 3  | 6  | 0  | 0  | 0  | 0  | 0  | 1.04 | 3  | 1.12 | 1  |
| 3     | -2.08 | 28.53 | 15.70 | 3  | 2  | 2  | 2  | 4  | 0  | 0  | 0  | 0  | 0  | 1.03 | 4  | 1.04 | 2  |
| 4     | -2.39 | 27.55 | 15.71 | 4  | 2  | 2  | 3  | 5  | 0  | 0  | 0  | 0  | 0  | 0.98 | 5  | 1.03 | 3  |
| 5     | -2.39 | 26.57 | 15.73 | 5  | 2  | 2  | 2  | 4  | 0  | 0  | 0  | 0  | 0  | 0.98 | 4  | 1.14 | 7  |
| 6     | 0.73  | 26.57 | 14.60 | 5  | 2  | 2  | 2  | 8  | 2  | 0  | 0  | 0  | 0  | 1.08 | 8  | 3.31 | 5  |
| 7     | -2.66 | 25.58 | 16.24 | 6  | 2  | 2  | 5  | 9  | 0  | 0  | 0  | 0  | 0  | 0.98 | 9  | 1.14 | 5  |
| 8     | 1.16  | 25.58 | 14.45 | 6  | 2  | 4  | 6  | 11 | 12 | 14 | 0  | 0  | 0  | 1.08 | 6  | 2.74 | 11 |
| 9     | -2.67 | 24.60 | 16.25 | 7  | 2  | 2  | 7  | 10 | 0  | 0  | 0  | 0  | 0  | 0.98 | 7  | 1.08 | 10 |
| 10    | -3.00 | 23.62 | 16.55 | 8  | 2  | 2  | 9  | 13 | 0  | 0  | 0  | 0  | 0  | 0.99 | 13 | 1.08 | 9  |
| 11    | 3.03  | 23.62 | 14.08 | 8  | 2  | 2  | 2  | 15 | 8  | 0  | 0  | 0  | 0  | 1.41 | 14 | 1.69 | 15 |
| 12    | 8.91  | 23.62 | 11.57 | 8  | 2  | 1  | 8  | 0  | 0  | 0  | 0  | 0  | 0  | 5.13 | 15 | 6.38 | 11 |
| 13    | -3.05 | 22.64 | 16.66 | 9  | 2  | 2  | 2  | 10 | 0  | 0  | 0  | 0  | 0  | 0.99 | 10 | 2.02 | 9  |
| 14    | 2.20  | 22.64 | 13.49 | 9  | 2  | 2  | 1  | 8  | 0  | 0  | 0  | 0  | 0  | 1.41 | 11 | 2.14 | 15 |
| 15    | 4.33  | 22.64 | 13.66 | 9  | 2  | 2  | 1  | 11 | 0  | 0  | 0  | 0  | 0  | 1.69 | 11 | 2.14 | 14 |

Fig. 3
<table>
<thead>
<tr>
<th>Step</th>
<th>Subroutine</th>
<th>Duties</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>SORT</td>
<td>On basis of chamber configuration, assigns gap and chamber numbers.</td>
</tr>
<tr>
<td>2</td>
<td>PRELOOK</td>
<td>Investigates spark dispersion determines parameters.</td>
</tr>
<tr>
<td>3</td>
<td>INTRAGAP</td>
<td>Attempts linking within gaps.</td>
</tr>
<tr>
<td>4</td>
<td>INTERGAP</td>
<td>Attempts linking between adjacent gaps.</td>
</tr>
<tr>
<td>5</td>
<td>CLEAN1</td>
<td>Removes complex linkages.</td>
</tr>
<tr>
<td>6</td>
<td>PASS1</td>
<td>Linear extrapolation away from event centroid.</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td>Links established?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Yes - go to Step 6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>No - go to Step 8</td>
</tr>
<tr>
<td>8</td>
<td>CLEAN2</td>
<td>Removes links forming small angles.</td>
</tr>
<tr>
<td>9</td>
<td>PASS2</td>
<td>Linear extrapolation toward event centroid.</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td>Links established?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Yes - go to Step 9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>No - go to Step 11</td>
</tr>
<tr>
<td>11</td>
<td>CLEAN2</td>
<td>Removes links forming small angles.</td>
</tr>
<tr>
<td>12</td>
<td>PASSN</td>
<td>Arbitrary linear extrapolation.</td>
</tr>
<tr>
<td>13</td>
<td></td>
<td>Links established?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Yes - go to Step 12</td>
</tr>
<tr>
<td></td>
<td></td>
<td>No - go to Step 14</td>
</tr>
<tr>
<td>14</td>
<td>CLEAN5</td>
<td>Removes hairs; inserts sparks.</td>
</tr>
<tr>
<td>15</td>
<td>CLEANUP</td>
<td>Identifies connected components.</td>
</tr>
<tr>
<td>16</td>
<td>GRAPHES</td>
<td>Investigates topology of connected components; if component is not circuit-free, selects minimum spanning subtree; makes dictionary entries for track endpoints and physical vertices.</td>
</tr>
<tr>
<td>17</td>
<td></td>
<td>Have all connected components been investigated?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Yes - go to Step 18</td>
</tr>
<tr>
<td></td>
<td></td>
<td>No - go to Step 16</td>
</tr>
<tr>
<td>18</td>
<td>GEOMETRY AND KINEMATICS PROGRAMS.</td>
<td></td>
</tr>
<tr>
<td>Step</td>
<td>Subroutine</td>
<td>Duties</td>
</tr>
<tr>
<td>------</td>
<td>------------</td>
<td>----------------------------------------------------------------------</td>
</tr>
<tr>
<td>1</td>
<td>SORT</td>
<td>Same as before.</td>
</tr>
<tr>
<td>2</td>
<td>PRELOOK</td>
<td>Same as before.</td>
</tr>
<tr>
<td>3</td>
<td>INTRAGAP</td>
<td>Same as before.</td>
</tr>
<tr>
<td>4</td>
<td>INTERGAP</td>
<td>Same as before.</td>
</tr>
<tr>
<td>5</td>
<td>CONTROL</td>
<td>Determines topological and geometric complexity of graph, removes incorrectly established links; determines next step (6, 8, 10, 12, or 14)</td>
</tr>
<tr>
<td>6</td>
<td>PASS1</td>
<td>Linear extrapolation away from event centroid.</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td>Go to Step 5.</td>
</tr>
<tr>
<td>8</td>
<td>PASS2</td>
<td>Linear extrapolation toward event centroid.</td>
</tr>
<tr>
<td>9</td>
<td></td>
<td>Go to Step 5.</td>
</tr>
<tr>
<td>10</td>
<td>PASS3</td>
<td>Arbitrary linear extrapolation.</td>
</tr>
<tr>
<td>11</td>
<td></td>
<td>Go to Step 5.</td>
</tr>
<tr>
<td>12</td>
<td>PASS4</td>
<td>Arbitrary helical extrapolation</td>
</tr>
<tr>
<td>13</td>
<td></td>
<td>Go to Step 5.</td>
</tr>
<tr>
<td>14</td>
<td>CLEANUP</td>
<td>Same as before.</td>
</tr>
<tr>
<td>15</td>
<td>GRAPHS</td>
<td>Same as before.</td>
</tr>
<tr>
<td>16</td>
<td></td>
<td>Have all connected components been investigated?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Yes - go to Step 17</td>
</tr>
<tr>
<td></td>
<td></td>
<td>No - go to Step 15</td>
</tr>
<tr>
<td>17</td>
<td>GEOMETRY AND KINEMATICS PROGRAMS.</td>
<td></td>
</tr>
</tbody>
</table>

Fig. 5
D is the search cone dimension determined by PRELOOK.

1. FVAL(I) for PASS1 and PASS2

\[ FVAL(I) = \frac{(E \times T(I))}{\cos^2 A} \]
where
\[ E = \text{distance between spark I and spark 2} \]
\[ A = \text{angle made with axis of search cone by line between spark I and spark 2} \]
\[ T(I) = \begin{cases} 1.0 & \text{if local degree of I is 0} \\ 1.1 & \text{if local degree of I is 1} \\ 3.0 & \text{if local degree of I is 3} \\ 4.0 & \text{if local degree of I is 2} \\ M+1 & \text{if local degree of I is } M \geq 4 \end{cases} \]

2. FVAL(I) for PASSN

\[ FVAL(I) = \frac{E}{\cos^4 A} \]

3. FVAL(I) is determined for each spark I found in the search cone and a link is made between spark J and spark 2 provided

\[ FVAL(J) = \min_{\text{I is SearchCone}} \{ FVAL(I) \} \leq 5.85 \times D \]

Fig. 6
DICTIONARY

Endpoints: 1, 13, 14, 21

Physical Vertices: 7, 18, 27

Tracks: (a) 1-2-3-4-5-6-7
          (b) 7-8-9-10-11-12-13
          (c) 7-22-23-24-25-26-27
          (d) 27-28-29-30-31-32-33-18
          (e) 18-17-16-15-14
          (f) 18-19-20-21

Fig. 7
DISCUSSION

TYCKO: You do spatial reconstruction first. Why is this preferable?

ROYSTON: First, we have virtually no ambiguities because we can make use of spark width and also we have 180° stereo. Secondly, there is a more positive advantage. If tracks cross in two dimensions, it is hard to say whether they intersect in space. Normally, it is possible to distinguish non-intersecting tracks more easily in three dimensions.
THE AUTOMATIC ANALYSIS OF 200,000 SPARK CHAMBER PICTURES USING THE CERN HPL

P.M. BLACKALL, B.W. POWELL, P. ZANELLA
CERN, Geneva

(presented by P.M. Blackall)

INTRODUCTION

This paper describes in general terms the methods used to automatically scan and measure the photographs from a spark chamber experiment. The pictures were scanned using the CERN Mark I HPD, operating on-line to the IBM 7090 computer. The program processed each event up to the stage of completing the geometrical reconstruction in space concurrently with the measuring process.

In addition to an important gain in speed over the hand analysis, it is also shown that the automatic measurements were more precise and more complete than the hand measurements. The reliability of the automatic scanning procedure is also discussed and results of a comparison with hand scanning are given.

1. SPARK CHAMBER EXPERIMENT

1.1 Description of the experiment

The spark chamber experiment was run at the CERN PS in January 1963 by Caldwell et al.\textsuperscript{1}) The aim of the experiment was to measure the elastic scattering differential cross-section for \( \pi^- - p \), \( \pi^+ - p \) and \( p - p \) scattering at angles between 1.25\(^\circ\) and 7.45\(^\circ\) in the laboratory system at incident momenta of 8.5, 12.4 and 18.4 GeV/c. In addition it was planned to investigate the behaviour of the diffraction peak for \( \pi^\pm - p \) scattering at high energies for a possible shrinking as predicted by Regge pole theory.
A schematic diagram of the experimental arrangement is shown in Fig. 1. An incident beam of pions or protons is deflected through 65 mrad in the first bending magnet and is then focussed onto a 20 cm long liquid hydrogen target. Beam particles that do not interact in the target are deflected through 60 mrad in the second bending magnet and then pass through the anti-coincidence counter behind spark chamber S08. For an elastic scattering interaction a triple coincidence occurs between the beam counter telescope, one of the scattered particle counters and the recoil particle counter. When such a coincidence occurs the tracks of the incident particle, scattered particle and recoil particle are recorded in the nine spark chambers S01 - S09. The nine spark chambers are photographed in 90° stereo and by means of a mirror-lens system all 18 views are arranged on a single 24 x 36 mm film frame. This represents a demagnification of about 60. Some 150,000 pictures were taken in the above triple coincidence triggering mode and in addition a further 100,000 pictures were taken with $\pi^-$ at 8.5 and 12.4 GeV/c in which the requirement of coincidence of the recoil particle was removed. These latter pictures were taken for observation of inelastic events in which a $\pi^-$ is produced with a momentum only slightly smaller than that of an elastically scattered pion at the same scattering angle.

1.2 Frame format and characteristics of the scan

Figure 2a) shows a typical frame and figure 2b) shows a schematic representation of the chamber view positions on the frame. Each view has its own fiducials to eliminate errors produced by movement in the mirror-lens system. The three large fiducial crosses are included for use by the scanning program to position the frame within the HPD co-ordinate system.

The pictures were scanned with a spot size of approximately 15 microns diameter and a scan line separation of 60 microns. The scan lines were 35 mm long and typically provided about 5,000 co-ordinates per frame. To detect as many as possible of the fainter spark images it was found necessary to tolerate a relatively high background rate of spurious
digitisings. The accuracy of the individual digitisings was approximately ± 5 microns rms.

2. **THE SCANNING PROGRAM**

2.1 **Basic subroutines**

The IBM 7090 scanning program, which together with its working space, occupies some 30,000 words of core storage, uses a set of basic subroutines which are essentially independent of the experiment and frame format. They perform the following functions.

1. Control the HPD and the flow of data from the digitiser to the computer.
2. Recognise a fiducial cross.
3. Find the sparks in a chamber view.
4. Find linear tracks from the sparks within a view.
5. Correlate the tracks found in two stereoscopic views of a chamber or of a group of chambers.

Detailed descriptions of these subroutines and their organisation have been given in a previous paper\textsuperscript{2}.

Two 7,000-word storage buffers are allocated to contain the digitisings from two frames. Whilst the contents of one buffer are being processed by the scanning program the other buffer is being filled with digitisings from the next frame. Each time the digitising of a new frame is begun the roles of the two storage buffers are interchanged. In this way the measurement process is limited only by the speed of the measuring device.

2.2 **General flow**

The sequence of operations carried out for each frame is determined both from the scanning requirements of the experiment and the format of the frame. If at any point in the sequence the necessary scanning conditions are not met, the frame is rejected and an entry is made in the appropriate rejection code.
For elastic scattering events the sequence is listed below:

1. Locate the fiducial crosses and position the frame in the HFD co-ordinate system.

2. Find the tracks of the incident particle in both views of chambers 1 and 2.

3. Check that there are not two incident particles in both chambers 1 and 2.

4. Find the tracks of the incident particle in both views of chambers 3 and 4.

5. Check that there are not two incident particles in more than one of the four chambers 1 to 4.

6. Check if the tracks in both views of chambers 1 and 2 correspond to the same incident particle.

7. Check if the tracks in both views of chamber 3 and 4 correspond to the same incident particle.

8. Extrapolate the tracks in chambers 1, 2 and chambers 3, 4 to the central plane of the first bending magnet and check that they intersect within 1 cm in space in the horizontal plane and 1.5 cm in the vertical plane.

9. Find the track of a recoil proton in both views of chamber 9, correlate the tracks in the two views if there are more than one and check that the proton comes from the target volume.

10. Find the tracks of scattered particles in both views of chambers 8, 7, 6 and 5.

11. Select the tracks in both views of chambers 7 and 8 which correspond to the same scattered particle.

12. Select the tracks in both views of chambers 5 and 6 which correspond to the same scattered particle.

13. Extrapolate the tracks in chambers 5 and 6 and chambers 7 and 8 to the central plane of the second bending magnet and check that they intersect within 1 cm in space in the horizontal plane and 1.5 cm in the vertical plane.

14. Correlate the selected trajectories from chambers 5 to 8 in the two stereoscopic views.

15. Check that the incident particle track, the scattered particle track and the recoil proton track extrapolate to a scattering vertex within the target volume.

16. Output onto magnetic tape the spatial co-ordinates and directions of the tracks in chambers 1 to 9 together with the approximate position of the scattering vertex.
2.3 Use of chamber fiducials

For each roll of film (approximately 3,000 frames) the relative positions of the chamber fiducials with respect to the fiducial crosses were measured for the first frame on a digitised projector (IEP). This method was necessary because the fiducials were about 150 microns long on the film corresponding to only 3 digitisings per line and thus accurately defined search areas were necessary to avoid incorrect identification due to background. On scanning the film with the HPD these measurements were then corrected for the first frame and subsequently dynamically corrected for each frame by recognition of the chamber fiducials. These fiducials are used to reconstruct tracks in space. All recognised tracks are reconstructed before linking them with tracks in neighbouring chambers to eliminate the variations in chamber demagnifications. For each track reconstructed, one point - the intersection of the track with a given fiducial plane - and the track direction are calculated by linearly interpolating between the two nearest chamber fiducials.

3. RESULTS

From January to April 1st, 1964 a total of 200,000 frames were scanned and measured automatically with the HPD. 100,000 of the 150,000 frames taken in the elastic scattering triggering mode and all of the 100,000 inelastic events were processed. The recognition of interesting events and their spatial reconstruction were done simultaneously with the measurement process at a rate of 1,200 frames per hour. The average computing time per frame was 1.5 sec so that the processing speed was limited by the mechanical speed of the HPD. In comparison with hand measurements, for the elastic events in which two out of five frames contain events to be measured the increase in speed is a factor of 80 whereas for the inelastic events in which all events must be measured before a selection can be made a speed factor of 200 is attained.

A detailed comparison of results from 10,000 frames was made for identification reliability and for freedom from systematic errors before the scanning system was put into production.
The efficiency of the system in both recognizing events and in correctly classifying the type of rejection was concluded to be between 92 and 95 per cent. This figure compares very closely with the efficiency obtained from the manual measurements. The remaining 5 to 8 per cent were due either to a track image being too faint to be digitized by the HPD or to a track containing an insufficient number of sparks for recognition (< 50% chamber sparking efficiency). Good agreement was found between the results of the 50,000 hand measured frames and those of the 100,000 machine measured events.

The overall precision of the HPD for this experiment is about 6 microns per track which corresponds to 0.4 mm in space. This is approximately a factor of 2 better than the hand measurements which were done on digitized scanning tables built specifically for the analysis of spark chamber pictures. The increased precision is due to the intrinsic precision of the HPD, the large number of scans made per track segment (each spark is scanned two or three times in its length) and the recognition of chamber fiducials for every frame. In the hand measurements a complete set of fiducial measurements were made only once every 100 events. Figures 3 and 4 show the difference between measurements obtained with the HPD (above) and hand measurements (below) for \( \pi^- + p \) elastic scattering at 12.4 GeV/c. The kinematical analysis of both the HPD measurements and the hand measurements were made by the IBM 7090 program SCRAP\(^3\).

In figure 3 the quantity plotted is the difference between the computed momentum and the measured momentum of the scattered pion. The zero has been displaced for convenience in presentation. The intervals of the histogram are 50 MeV/c.

In figure 4 the quantity plotted is the sum of the squares of the residuals after the geometrical fitting of the vertex. The residuals are measured at a fixed plane in each of the six chambers used to define the three tracks of the event.
In addition to the improvements in speed and in precision of measurement the automatic scanning system demonstrated further advantages.

1. The simplification in book-keeping due to no remeasurements which was appreciated particularly by the physicists.

2. By giving on-line summaries of event and rejection distributions during the scanning operation a check could be made on the correct functioning of the device and of the program.

ACKNOWLEDGEMENTS

We would like to thank D. Harting, G. Giacomelli and L. Monari for their collaboration in the examination of the results and in determining the method of analysis.

We would also like to express our thanks to B. Evershed, G. Durupthy, F. Marciano, G. Pacteau, L. Sohet and R. Zurbuchen for their help in keeping the HPD operational during the period of analysis. We are also indebted to Mrs. B. Powell and to Miss M. Rey for their help in preparing the input data for the program and to those computer operators who volunteered to work overtime so that we could complete the analysis sooner.
References


3. B. Zacharov, Nuclear Instruments and Methods, to be published.
Figure captions

Fig. 1  Schematic diagram of $\pi^+ - p$ elastic scattering experiment.

Fig. 2 (a) Typical frame showing the positions of chamber views.  
(b) Schematic diagram of frame format.

Fig. 3  Histograms of the difference between the computed momentum and the measured momentum of the scattered pion - above HPD measurements, below hand measurements.

Fig. 4  Histogram of the sum of the squares of the residuals of the space points for chambers 3, 4, 5, 6 and 9 after the geometrical fitting of the scattering vertex - above HPD measurements, below hand measurements.
Schematic diagram of the $\pi^\pm - p$ elastic scattering experiment at 8.12 and 18 GeV/c

- SC1 - SC9: spark chambers
- C: coincidence counters
- A: anticoincidence counters

Fig. 1
Schematic of Chamber/View Positions

8/1       8/2
2/1  2/2  4/1  4/2  6/1
3/1  3/2  5/1  5/2
7/1  7/2

9/1 normal view
9/2 stereo view

Fig. 2a

Fig. 2b
A PROGRAM FOR CALIBRATING THE FLYING SPOT CATHODE RAY TUBE FOR CHLOE

R.J. ROYSTON, J. BUTLER and P. PENNOCK
Argonne National Laboratory, Argonne

(presented by R.J. Royston)

SUMMARY

The flying spot cathode ray tube in CHLOE is calibrated by scanning a suitable photograph which is also measured on an accurate conventional measuring table. The type of photograph suitable for this application differs from that which is used for calibrating a conventional table and the considerations which lead to its choice are discussed. The CHLOE output is then processed by a program which identifies and measures the position of the calibration marks. These are then compared with their true measured positions and a correction function is computed by means of a fitting procedure.

* * *

In calibrating a manual measuring table it is both customary and convenient to use a film on which a rectangular grid has been drawn with great precision. One can, for example, then run the position marker along the grid lines and observe the changes in scales attached to the x and y encoders of the table.

*) Work performed under the auspices of the U.S. Atomic Energy Commission.
In an automatic CRT digitizer such as CHLOE it is not clear that such an accurate grid is necessary, or even that it is desirable to use a grid at all. Data coming from an automatic digitizer has to be processed by a computer, and it is well known that writing programs to process photographs containing large numbers of intersecting lines can be quite difficult! Moreover, attempting to bring an on-line operator into the system by means of a slave scope is also liable to involve fairly extensive programming.

The task of locating and separating the calibration points from one another is made simpler if the picture to be scanned is broken into separate, disjointed pieces. Then no matter whether one uses a histogramming technique or a binning technique, the digitizations for each calibration mark can be separated from one another without any difficulty.

In our first calibration we used a set of X's (see Fig. 1). They were drawn by a draughtsman simply using a ruler. We had the drawing photographed and printed on 35 mm film. The film was then placed in the CHLOE scanner and scanned with a raster of 4096 x 4096 points, the digitizations being written on a magnetic tape. It should be remembered that the digitizations come in pairs, one when the scanning spot first encounters the arm of the cross and one when it first encounters the clear film surrounding it again.

The magnetic tape was taken to the CDC 3600 where we had a program which regarded each pair of digitizations as a line segment. These line segments were "binned" by starting a new "bin" (or array) every time a line segment was found which did not overlap with the last line segment in one of the bins already started. If the segment did overlap with the last segment in some existing bin, and the two segments occurred in successive scan lines, it is placed in that bin and it now becomes the last segment. This basically simple algorithm was modified in various ways to take care of all the complications that could arise. Excessively long line segments were discarded altogether; one scan was allowed to be missed without closing off the bin; if one segment fitted in two bins, the two bins were combined into one; if two segments fitted into the same bin the test segment for subsequent scan
lines was made to include both segments. When the binning was complete, all bins with less than 20 line segments in them were thrown away and two roads, at 45° and 135°, were constructed through the centroid of the system. Two straight lines were fitted to the points that lay within these roads. In order to prevent anomalies, if one end of a line segment fell in the road, both ends were used in the fit.

The intersection of these two lines was then found and this was taken to be the position of the calibration point which was output on a punched card. In order to avoid a lot of unnecessary pattern recognition work we also drew what CHLOE saw on a CALCOMP plotter via magnetic tape (see Fig. 2). In this picture we included the intersection point and numbered all the bins. It will be noticed that one or two items not corresponding to useful information were processed. We simply threw away the cards corresponding to these.

The calibration picture was also measured directly. In order to calculate the distortion using these two sets of measurements we performed a least squares fit of one set of measurements to the other, with a correction applied to the measurements from CHLOE. The correction was in its general form and the parameters in it were varied to produce the best fit.

Since a device like CHLOE has several focussing and correcting coils and at least one lens, it is not likely to have one simple optical axis. So it is hard to justify the use of the standard pincushion formula. We therefore took the attitude that some ad hoc formula would be in order provided that it worked; and in fact we found that quadratic formulae of the type

\[
x = a_1 + b_1 x + c_1 y + d_1 x^2 + e_1 xy + f_1 y^2
\]

\[
y = a_2 + b_2 x + c_2 y + d_2 x^2 + e_2 xy + f_2 y^2
\]

gave a very good fit indeed, leading to a root mean square deviation of the fitted values from the measured values of about 2 units when the correction was about 30 units.
Having simplified the specifications of the grid somewhat from what one usually associates with a calibration picture the question now arises, as to whether further simplifications could be made. These 16 circles (Fig. 3) represent an approach to such simplification. If they are produced in high contrast white on black as they are shown here, they give a lot of latitude to the CRT brightness and photomultiplier bias level settings without distorting the shape of what one sees or introducing background. The centre of a circle is its centroid, which is pretty simple to compute, and in fact we were able to process this film through CHLOE with a cut down version of J. Butler's EIM program. The direct measurement of the film is simplified by actually marking the centre of each of the circles. The results obtained with 16 circles were as accurate as those obtained with the larger number of crosses (see Fig. 4).
Figure captions

Fig. 1 Photograph of the calibration pattern.
Fig. 2 CHLOE-output by Calcomp plot for Fig. 1.
Fig. 3 Calibration pattern with circles.
Fig. 4 CHLOE-output for Fig. 3.
DISCUSSION

BAZIN: Why is the distortion 1% here whereas it is 15% at M.I.T.?

ROYSTON: We have an analogue circuit, which calculates the distortion and changes the deflection current by the appropriate amount to square the distortion off again.

DEUTSCH: I think the answer is - the Argonne value is after correction.

POWELL: How often do you think it is necessary to do this calibration?

ROYSTON: I hope during production, it is no more than once a week.
AUTOMATIC ANALYSIS OF PHOTOGRAPHIC DATA IN BIOLOGY *)

J.W. BUTLER, MARGARET K. BUTLER and AGNES STROUD
Argonne National Laboratory, Argonne, Illinois

(presented by J.W. Butler)

I. FOREWORD

The project to be described is being carried out as a joint effort of the Applied Mathematics and Biological and Medical Research Division of Argonne National Laboratory. Mrs. Stroud has been responsible for the biological laboratory procedures, while Mrs. Butler has written most of the difficult portions of the computer programs.

The need for automatic analysis of chromosome data was originally brought to our attention by Austin M. Brues in February 1962, and work on the problem was begun in earnest early in 1963 when the CHLOE film measuring system reached a satisfactory state of development.

II. INTRODUCTION

Recent advances in cytogenetic techniques have caused the subject to acquire considerable interest in connection with radiation biology and also to emerge as a potentially important clinical tool for recognition and identification of congenital diseases. More particularly for clinical application, however, use of the method has been hampered by the tedious and time-consuming nature of the laboratory procedures. For this reason, it seemed to us that it would be worthwhile to make use of the rapid information processing capabilities of a digital computer to reduce the time extension of the process and thus render it more appropriate for clinical and research application. In addition, an important feature of a digital procedure is the

*) Work performed under the auspices of the U.S. Atomic Energy Commission.
production, in a reasonable length of time, of statistically significant amounts of numerical data, making possible the discovery of new relationships on a rational inductive basis.

With the exception of the chromosomes indicating the male sex, chromosomes of most species occur generally in pairs. The characteristics of the chromosome population vary with species but are consistent from cell to cell in a given species. In the operation being discussed, photographs are obtained showing all the chromosomes possessed by individual cells; these photographs form the input data for the classification procedure. A typical such photograph is shown in Fig. 1. Such photographs can only be obtained during the metaphase stage of mitotic division, since it is only in this stage that the separate chromosomes are visually resolvable.

The chromosomes naturally take up a more or less random arrangement in the photographs, and the objective of the present computer programs is to automatically make the necessary measurements and pair each chromosome with its companion on the basis of shape similarity.

III. OVERALL PROCESS FLOW

The sequence of operations in the existing procedure is shown in Fig. 2. The first three boxes in the chain refer to the laboratory techniques which are described in the next section. "Mechanical Preparation" refers to the actual handling of the materials and preparation of the microscope slides. "Recognition" relates to the operation, at present carried out by humans, of recognizing the particular cells to be photographed on the basis of various criteria of desirability. At the moment, this function presents the chief obstacle to complete automation; however, it is not especially time-consuming and is not a serious bottleneck. The photographic part of the process is already practically automatic and could easily be eliminated, if desired, by working directly with the prepared cell material. The next two blocks are the principal subject of the present discussion, but I will first give a brief description of the biological laboratory procedures.
The chromosomes of individual cells are made visible by the tissue culture method. This requires the maintenance and propagation of cells away from the host organ in a nutrient medium which enables the cells to divide under favourable conditions. Treatment of the culture with the drug colchicine causes the cell division process to be inhibited in metaphase, with the consequence that cells moving into this stage are accumulated. The spindle attachments to the visible chromosomes at the centromeres are also destroyed by the drug, making it possible to spread the chromosomes\textsuperscript{1) with the addition of a hypotonic solution. An alcohol and acetic acid fixative is used to kill the cells and a drop of this solution containing cells is spread on a glass slide and is allowed to dry in the air. For photomicrography an aceto-orcein stain is used to stain the chromosomes a purplish red. A 35 mm Zeiss Ikon camera attached to a Zeiss phase microscope (100X oil objective) records the chromosome images on film strips for input to the CHLOE machine.

V. DESCRIPTION OF SCANNING EQUIPMENT

The measurement function and part of the data processing are performed by the CHLOE film measuring system, designed and built in the Applied Mathematics Division at Argonne by an engineering group led by Donald Hodges. The basic function of the equipment is to serve as a powerful and flexible means of digitizing photographic information and performing necessary computations with the numerical data obtained from the photographs.

A functional block diagram of the system is shown in Fig. 3. The two main components are a digital computer and an optical scanner, the latter being under control of the computer. The computer is the commercially available ASI-210\textsuperscript{2}, while the cathode ray tube (CRT) scanner was built in the Applied Mathematics Division of ANL. The spot of light from the CRT is projected onto the film, and the light transmitted through the film is viewed by a photomultiplier by means of which a decision is made regarding the density of the film at the point in question. The light spot is driven by two counting registers so as to scan a rectangular area on the film, the extent of this area in both directions being determined by the computer program. The spot does not actually move, but appears in one place for one \(\mu s\), is blanked
out for three μsec, and reappears in the adjacent location for one μsec. When the photomultiplier unit detects a sufficiently large change in the transmitted light from one point to the next, the contents of the counters are sent into the computer memory as the co-ordinates of the point.

A system of this sort has a certain quality of universality, in that it is capable, within limits, of doing anything that we are clever enough to instruct it to do. Consequently, in order to apply the machine to a particular experimental situation, a program must be written to appropriately control the operation of the scanner and to deal with the information obtained from the film.

VI. COMPUTER PROGRAMS

The overall structure of the computer programs is shown in Fig. 4. The first two indicated functions - the measurement and the shape construction - are controlled by the ASI-210 computer which forms a part of the CHLOB system; the remaining operations, with the exception of the picture reconstitution, are performed by the IBM 704 program. Reconstitution of the pictures from the digital information is done with the home-built computer GEORGE, which is equipped with a cathode-ray tube output device. Input for this step is a magnetic tape which is written out by the IBM 704 at an early stage in the processing of the data.

The first section of the ASI-210 program is a routine to set up a repetitive display on the monitor CRT to allow the operator to adjust the beam current on the scanning tube. When this display assumes the proper appearance, control is transferred by operator action to the next section of the program, at which time the background density of the film is computed, the photomultiplier discriminator is set at the indicated level, and control branches to the next part of the program which controls the actual measuring process.

The measuring program causes the entire area of the film to be scanned and detects and labels any connected shapes which may be present in the photograph.
Figure 5 shows some rather miscellaneous shapes such as might be found on a photograph. The area external to the closed shapes should be thought of as "blacker" than the interior zones. The co-ordinate origin is placed conventionally in the lower left-hand corner with x increasing to the right. It may be recalled that the CHLOE scanning unit only sends co-ordinates into the computer when there is a density change of specified magnitude on the film. This means that only the edge co-ordinates of the shapes are measured - this was, of course, done deliberately to reduce the peak data rate - so that the information in the computer memory may be regarded as a collection of horizontal line segments.

After a buffer area in the ASI-210 core memory is filled with line segments, the computer examines the horizontal position of each segment relative to those on the next y level below (or the next N levels if the operator so desires). Segments whose projections onto the x axis have a non-zero intersection are then caused to acquire the same label, indicating that they belong to the same connected shape. Since the line segments are horizontal, the right hand y co-ordinate is superfluous and these memory locations are used for labelling. Line segments which the computer is unable to associate with a shape are labelled with a negative number.

The labelled segments are then written out on a magnetic tape for use as input to the IBM 704 computer, which has been programmed to perform the remainder of the analysis.

The IBM 704 program first causes a magnetic tape to be written containing the paired line segments in the proper format for input to the GEORGE CRT plotting device. Regarding the closed shapes as regions of constant density, it then computes the first ten absolute xy moments (through the third order) of each shape and attempts to form initial clusters of shapes for possible assembly into chromosomes. The reason for this last step is that chromosomes often appear in the photograph as several nearby pieces, and these must be put together before processing can proceed. This is done essentially by approximating each shape by an ellipse and associating them into the same initial cluster if the corresponding ellipses overlap. The existing program is able to deal with an initial cluster of up to six elements.
The method used in the program for classifying chromosome shapes is based on the use of moment invariants as described by Hu. These are algebraic combinations of absolute moments which are invariant under various sub-groups of the full affine group in the plane. The particular invariants used here may be referred to as Euclidean moment invariants, since they are chosen to be invariant under the extended Euclidean group, that is, the group in the plane consisting of translations and proper and improper rotations. We have used the first 7 of these invariants, which are formed from absolute moments through the third order. Each chromosome may then be associated with one point in a 7-dimensional space, which may be made into a metric space by introduction of a suitable metric. Moment invariants have also been used for recognition of printed letters by Alt.

After an initial cluster is formed, the program tests all possible sub-clusters to see which of them "look" like chromosomes. This is done by merging trial sub-clusters (adding their absolute moments), computing the 7 moment invariants for the merged clusters, and accepting the largest clusters whose representative points lie inside of a certain polygonal region in the moment space. Successful clusters are then stored in the computer memory in the form of 9 quantities for each cluster, these quantities being the co-ordinates of the centroid and the 7 invariants.

The moment space is now converted into a metric space by definition of the distance

$$D = \frac{|J - J'| + |J_1 - J'_1| + \ldots + |J_6 - J'_6|}{|J_0| + |J'_0| + |J_1| + |J'_1| + \ldots + |J_6| + |J'_6|},$$

where $J_0 \ldots J_6$ are the 7 moment invariants and the primed and unprimed symbols refer to the two chromosomes between which the distance is being measured. This distance function is scale invariant and also has the obvious property $0 \leq D \leq 1$. In computation of the invariants, the formulas given by Hu were modified in such a way as to cause each invariant to have the same physical dimensions.
An array of numbers is then formed whose elements are the distances between each pair of chromosomes in the 7-dimensional metric space. The pair with the smallest distance is then selected out as a valid pair, followed by the pair with the next smallest distance, etc., until all the chromosomes are represented. The monotonically increasing sequence of distance measures obtained in this process is independent of the scale of the photograph and the arrangement of chromosomes in the spread and is therefore an invariant property of the karyotype.

Notice that there has been a subtle change in the use of the word karyotype. Instead of the usual meaning of the term as referring to the photograph of the chromosomes arranged in neat rows and collected into appropriate groups, the mathematical viewpoint leads to an abstract definition of karyotype as the set of measurable properties of a metaphase spread which are invariant under change of scale and under rearrangement of the individual chromosomes. Indeed, we can go further and define an ideal karyotype by averaging these properties over an infinite "ensemble" of cells of the same kind. Figure 1 shows a spread from a pig kidney tissue culture which has been maintained by Mrs. Stroud for a number of years, while Fig. 6 shows the conventional method of exhibiting the karyotype in which the pairing of the chromosomes and the group structure are clearly visible.

Getting back to the numerical output of the program, what might we use these numbers for? One possibility might be this:

Consider, for example, a human tissue culture spread with 46 chromosomes. The output from the computer program is then a monotonically increasing sequence of 23 distance measures, which can be associated with a point in a 23 dimensional measurement space. To make the situation easier to visualize, let us cut the number of dimensions down to 2, so that we can draw a figure on paper, (see Fig. 7).
Thus we know that the measurement points must lie inside the truncated cone. Now, a large number of measurements on a normal cell population might cluster as shown in Fig. 8. If some of the chromosomes fail to pair properly, some of the co-ordinates will increase in value, leading to a shift of the cloud of measured points. If there is a significant difference between these two clouds of points, known statistical techniques can be applied to compute a separating hyper-plane which serves to classify any new items entering the system with a certain confidence level in the statistical sense. The ideal members of the two classes are then approximated by the mean points of the respective clouds.

The result of this process is then an objective classification of cell populations into normal and abnormal. The statistical techniques can be extended without difficulty to produce more classes if desired. The success or failure of this classification procedure depends on the adequacy of the model used to reduce the data.

The structure of the process up to this point is:

![Diagram](attachment:image)

Almost any experimental situation can be decomposed in this way. In particular, experiments in high energy physics carried out with spark chambers fall naturally into an identical scheme. The only real difference in the character of the model.

Figure 9 shows a spark picture taken by the Argonne group at CERN in Geneva in 1962. The model in this physical situation, in simple terms, tells us that the sparks are produced along the paths of particles moving in accordance with certain differential equations, an hypothesis which enables the reduction of
the raw data to only a few numbers. These numbers are then used to classify the photographs according to reaction event types.

In the biological situation which we are talking about, the model is furnished by the combinatorial structure of the karyotype, meaning the property of chromosomes to occur in pairs and separate further into groups as you can see in Fig. 6. I believe it can now be seen that the two situations are remarkably similar, and that the same strategy can be used in dealing with either experimental situation.

VII. CURRENT ACTIVITY AND RESULTS

Figure 10 shows the reconstitution of the digital information obtained by scanning the pig kidney spread (Fig. 1). These have not been retouched; they are shown directly as they came from the GEORGE cathode ray tube output.

The outlines of the chromosome shapes, again constructed from the digital information, is shown in Fig. 11, together with connecting line segments illustrating the pairing obtained from the computer program. If this slide were carefully compared with the previous karyotype illustration, one would see that about 65% of the chromosomes have been paired in a reasonable way. Actually, one thing that caused the pairing to go off in this example is that the film was inadvertently not centred properly and about 4 of the chromosomes were lost off the edge of the picture.

Figure 12 shows a page of the printed output from the IBM 704 program corresponding to the pairings exhibited on the last slide. Figs. 13 and 14 represent more or less typical output from the CRT plotting device.

The structure of the computer programs for producing karyotypes (pairing) is now stable and has been so for 6 or 8 months. We are now engaged in the tedious process of optimizing the performance of the system by adjusting the various parameters initially built in for this purpose. The programs have now been brought to production status, which means that the whole computer system can be operated by computer operators on a routine basis.
The previously mentioned pig kidney culture is being used as a control for optimization purposes. So far we have looked at several hundred spreads from this culture, running them through the system repeatedly during the process of parameter adjustment.

In the operation as presently carried out, there are four main stages which must be optimized for maximum performance:

1) Photographic techniques
2) Data acquisition from the film
3) Recognition of chromosomes by the computer programs
4) Optimization of karyotype production (pairing).

Stage 2 concerns the optimization of operating procedures and parameter values in such a way that the digital information supplied to the computer is a faithful representation of the information which it is desired to extract from the photograph. Actually, one of the most difficult and subtle points connected with the design and operation of a film scanning machine is the proper adjustment of light discriminator levels so as to recover the information from the photograph. For instance, if this level is badly set, the machine may ingest a large quantity of data which is merely related to the "graininess" of the film and has no relevance to the desired application. On the other hand, one may get essentially no information at all under certain conditions.

A little reflection on this problem leads to the realization that the machine cannot be expected to perform this operation "blindly" but must be instructed in some way as to the character of the information it is to obtain.

We have chosen to give the machine what appears to be the least information possible in this regard, namely, we instruct the machine to so adjust its operating condition that the total number of co-ordinates produced by a scan is between an upper and lower bound, those bounds being input parameters to the program. Surprisingly enough, this small amount of information enables the machine to do quite a respectable job if the upper and lower bound parameters are adroitly chosen.
Due to a certain design limitation of the scanning machine, however, this procedure must be supplemented by the previously mentioned manual procedure for adjustment of the beam current in the cathode ray tube light source.

I have recently tried this technique on a few spark chamber pictures and it appears to yield acceptable results with these pictures also.

In the optimization process for the chromosome project, we still have a little way to go on stage 3, and there are strong indications that the pairing process will proceed very well when the recognition of individual chromosomes is brought to an acceptable degree of efficiency.

Acknowledgements

I would like to thank B. Kroupa, R. Wehman, and G. Wittrus for keeping the various items of special equipment operating during the early phases of the effort. The Digital Operations Group is also to be commended for efficient operation of the several computing machines used in the project.

References


Figure captions

Fig. 1  Photograph of chromosomes for input to CHLOE.
Fig. 2  Schematic diagram of sequence of operations.
Fig. 3  Functional block diagram of the CHLOE system.
Fig. 4  Block diagram of the computer programs.
Fig. 5  Examples of different shapes found on the film.
Fig. 6  Example of karyotype with the pairing of chromosomes.
Figs. 7 - 8 Clustering of cell population for 2-dimensional case.
Fig. 9  A spark picture taken by the Argonne group at CERN.
Fig. 10 Digitally reconstituted photograph of chromosomes
        from a pig's kidney.
Fig. 11 Outlines of chromosome shape from digital information
        indicating the pairing obtained from the computer program.
Fig. 12 Numerical output from the IBM 704 program.
Figs. 13 - 14 Typical output from the CRT plotter.
### CHROMOSOME PAIRING OPTIONS
- **PLOT=1**
- **MSL=128**

### SERIES A FRAME COUNT
- **max=0**
- **SDX=9**
- **SDY=9**
- **E=0.0000906**
- **H=2**
- **N=1**
- **MSL=0.25000000**
- **UPH=2.25714**
- **CML=128**

### 70% PARAMETERS
- **Amax=0.00002394**
- **AMIN=0.00000000**
- **AMINC=3.00724**
- **AMERG=0.00000000**

### DEPRE
- **0.00000000**
- **0.00000000**
- **0.00000000**
- **0.00000000**

### NUMBER OF CHROMOSOMES
- **17**

<table>
<thead>
<tr>
<th>XG</th>
<th>YG</th>
<th>JG</th>
<th>JI</th>
<th>J2</th>
<th>J3</th>
<th>J4</th>
<th>JC</th>
<th>J6</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.923</td>
<td>0.571</td>
<td>1.166E-02</td>
<td>3.75E-06</td>
<td>2.249E-10</td>
<td>6.076E-12</td>
<td>1.356E-11</td>
<td>1.199E-20</td>
<td>1.878E-15</td>
</tr>
<tr>
<td>0.923</td>
<td>0.659</td>
<td>1.166E-02</td>
<td>3.69E-06</td>
<td>2.231E-10</td>
<td>1.905E-12</td>
<td>2.672E-11</td>
<td>1.772E-20</td>
<td>1.591E-15</td>
</tr>
</tbody>
</table>
| **SEPARATION DISTANCE OF PAIR 0.0000**

| 0.956 | 0.795 | 1.556E-02 | 1.657E-05 | 3.689E-10 | 6.574E-12 | 4.647E-10 | 1.763E-17 | 3.838E-13 |
| 0.987 | 0.599 | 1.556E-02 | 1.658E-05 | 4.619E-10 | 1.025E-09 | 9.071E-10 | 6.631E-17 | 7.966E-13 |
| **SEPARATION DISTANCE OF PAIR 0.0000**

| 0.952 | 0.726 | 1.166E-02 | 4.835E-06 | 2.904E-10 | 7.11E-09 | 2.636E-09 | 6.472E-16 | 7.092E-13 |
| 0.945 | 0.650 | 1.166E-02 | 4.836E-06 | 2.789E-10 | 6.269E-10 | 3.19E-10 | 1.198E-17 | 7.802E-14 |
| **SEPARATION DISTANCE OF PAIR 0.0000**

| 0.981 | 0.864 | 1.166E-02 | 7.69E-06 | 3.129E-10 | 7.71E-09 | 7.325E-09 | 4.719E-15 | 4.029E-12 |
| 0.954 | 0.771 | 1.166E-02 | 7.326E-06 | 3.083E-10 | 2.416E-09 | 2.203E-09 | 4.357E-16 | 1.183E-12 |
| **SEPARATION DISTANCE OF PAIR 0.0000**

| 0.953 | 0.810 | 1.556E-02 | 5.857E-05 | 5.606E-09 | 5.298E-09 | 1.947E-09 | 1.514E-16 | 1.287E-12 |
| 0.916 | 0.746 | 1.556E-02 | 7.467E-06 | 5.606E-09 | 9.444E-12 | 6.701E-12 | 4.053E-17 | 3.124E-13 |
| **SEPARATION DISTANCE OF PAIR 0.0000**

| 0.949 | 0.761 | 1.361E-02 | 1.059E-05 | 3.40E-10 | 2.80E-09 | 2.666E-09 | 5.420E-16 | 1.896E-12 |
| 0.976 | 0.650 | 1.550E-02 | 1.948E-04 | 5.892E-08 | 2.185E-08 | 4.678E-08 | 6.173E-14 | 5.293E-10 |
| **SEPARATION DISTANCE OF PAIR 0.0000**

| 0.924 | 0.629 | 3.845E-03 | 4.737E-06 | 2.09E-07 | 2.093E-07 | 1.140E-11 | 2.576E-10 |
| 0.970 | 0.610 | 5.798E-03 | 4.504E-04 | 7.676E-08 | 9.814E-07 | 1.611E-06 | 3.111E-10 | 1.203E-07 |
| **SEPARATION DISTANCE OF PAIR 0.0237**

| 0.949 | 0.764 | 9.67E-03 | 4.105E-06 | 5.165E-11 | 1.513E-08 | 1.273E-08 | 1.825E-14 | 5.38E-11 |
| 0.926 | 0.713 | 1.358E-02 | 2.623E-03 | 4.625E-07 | 4.235E-04 | 4.119E-04 | 1.250E-05 | 7.940E-05 |
| **SEPARATION DISTANCE OF PAIR 0.2753**

| 0.965 | 0.777 | 1.558E-02 | 3.979E-12 | 1.261E-05 | 2.796E-02 | 2.793E-02 | 5.747E-02 | 8.18E-02 |

**Fig. 12**
DISCUSSION

TURNILL: Are the specimens only two dimensional or is it possible to get overlapping of the chromosomes?

BUTLER: This varies according to species. There is also a certain amount of overlap which is a function of the amount of spreading in the preparation. It will generally increase the values of the variants of anything one is trying to find, so that one must take more pictures than otherwise.

TURNILL: What is your rejection rate?

BUTLER: It is not quite fair to use that term in this case. At the moment we are probably analyzing 10% to 15% of the pictures correctly and this is gradually rising as the parameters are optimized properly.

BAZIN: You drew an analogy between your work and spark chamber work. You characterize chromosomes with seven parameters. Could you tell us how many parameters would characterize a spark chamber picture?

BUTLER: The analogy I referred to is that in the systems. From a systems point of view the activities involved in getting the answer are more or less similar.

DEUTSCH: Could you give us an example of what the parameters represent?

BUTLER: The first parameter is simply the area. The next two are related to the eccentricity of the object, for example its length relative to its width. One can introduce an infinite set of numbers of this type which do not depend on the orientation or location of the object.

LEBLOND: Do you ever get two areas that have the same parameters and do not look alike?
BUTLER: I have not really looked at them all.

ROYSTON: I just wanted to comment on the number of parameters which characterize a spark chamber. Three parameters will characterize a spark, two for its position and one for its width. For fiducials we use two parameters - the x and y coordinate of the centre. So there may be some analogy there.

FRANZINETTI: The comparison should be made not between one artificial set of electroplating but between one chromosome and one spark but between a chromosome and a whole track. In this problem the points might not be independent and the whole set of parameters fact the comparison with a bubble chamber track is closer. In the latter case one would of course have to deal with an infinite number of invariants.

I also wonder if you have attempted any kind of automatic recognition on the basis of learning of patterns as is being investigated in some laboratories.

BUTLER: This is something similar to what we are doing in the optimization process, except that we change the parameters by hand. We put a large amount of data through the system and if it does not look right when it comes out we punish the machine by changing a parameter, until it begins to give the right answers.
SPARK CHAMBER PHOTOGRAPH ANALYSIS USING LUCIOLE

A.E. HEAD, J.A. WILSON
CERN, Geneva

(presented by A.E. Head)

1. LUCIOLE

1.1 The device

LUCIOLE is a device which generates a flying spot of light to scan 35 mm spark chamber film, and to feed digitized data on-line to a large computer, at present the IBM 7090.

The light spot is generated by a high precision cathode ray tube and has a diameter which does not exceed 35 microns on the screen.

The spot is demagnified in the ratio 1:0.45 onto the film by a lens system, giving a spot diameter of 17 microns, and a scanning area of approximately 24 mm x 24 mm on the film. A condensing lens is placed behind the film in order to distribute the light passing through the film uniformly over the surface of the photo-multiplier. In order to achieve maximum stability and reproducibility of the scan, a raster of parallel lines is used which is continuously rewritten in the same manner.

On the film negative, items of useful information (i.e. sparks, fiducial marks etc.) appear as black images and the output from the photo-multiplier is therefore inverted before being sampled by a discriminator, pulses greater than that of the discriminator being taken as possible useful information. The wave is then "squared" before being sampled by the digital circuit, which gives a coded form of data containing both width and positional information of the pulses. This system of digitizing, known as Sampled Analogue Digitizing enables the co-ordinates of closely spaced tracks to be packed into a single 36 bit word so that
it is not necessary to have an intermediate buffer store between the
digitizer and the computer.

The least count in the slow scan direction is 32.5 microns and
in the fast scan direction it is 14 microns. The resolution of the
scanner corresponds approximately to one spot diameter of 35 microns.

The TV-like scan of the film takes 1.510 sec, which includes
350 msec fly-back time for the spot, during which the film is automatically
advanced to the next frame. The frame-to-frame advance time is 250 msec
and the film transport gives a positional accuracy of ± 100 microns in
the direction of motion (as opposed to ± 1 mm for normal camera mechanisms).
In its present configuration LUCIOLE will handle only 35 mm perforated
film; this may be modified in later models. The frame currently in posi-
tion is continuously scanned by the CRT but data is only transmitted to
the computer if it has been requested. If data has not been requested,
then the film is not advanced.

1.2 Correction of systematic errors

In achieving stability and reproducibility of scan, various
systematic errors are introduced into LUCIOLE digitizings. The more im-
portant errors are as follows:

a) Overall errors due to
   i) optical distortion by the lens system,
   ii) pin-cushion distortion by the CRT.

b) Errors in the slow scan direction (i.e. X) due to the
   fact that
   i) scan lines are not parallel,
   ii) scan line separation is not constant.

c) Errors in the fast scan direction (i.e. Y) due to the
   fact that
   i) spot velocity is not constant across a scan line.
In the slow scan direction (the X-direction), the co-ordinate is obtained from a count of the scan lines, while in the fast scan direction (Y) the co-ordinate is obtained essentially by measuring the time of passage of the spot.

To correct for these errors over the total screen area, a calibration grid of crossed, straight lines, at 45° to the scan line direction, is scanned every hour.

The intersection points on the grid are determined in the LUCIOLE co-ordinate system. The points have also been measured with a precision microscope to an accuracy of 3 microns. From these two sets of measurements the errors in both X and Y in the LUCIOLE co-ordinate system are determined using a least squares fit over all the intersection points. Figure 1 shows typical error distributions in X and Y. The unit in X is in least squares (i.e. 32.5 microns) the unit in Y is in half-least counts (i.e. 7 microns).

2. The EXPERIMENT

2.1 Set-up of the experiment

The first spark chamber experiment selected for data analysis by LUCIOLE is a hyperon missing mass experiment run at the CERN PS in December 1963 by the Lundby group. The reactions studied are those arising from an incident π⁻ beam hitting a liquid hydrogen target.

The momentum of the incident beam was varied from 1.1 GeV/c to 2.75 GeV/c.

The reactions considered are

\[ \pi^- + p \rightarrow Y^- (\Sigma^-, \Xi^-, \ldots) + K^+ \]
\[ M^- (\pi^-, \rho^-, \ldots) + p \]

Figure 2 shows a schematic diagram of the experimental arrangement and optics.
An incident beam of negative π-mesons is focussed onto the liquid hydrogen target and then passes through the hole in chamber 3. The meson produced passes through chambers 3 and 2, is deflected by the magnet, and continues through chamber 1 and the Čerenkov counters to the K-decay counter.

Decay products are recorded in chambers 3, 4 and 5. The system is triggered by the counter telescope and the $K^+$ - decay counter. All 5 chambers are photographed in $90^\circ$ stereo and the optical system arranges the views on a single 24 x 24 mm frame. The demagnification factor is approximately 40. A total of 100,000 pictures were taken of which approximately 80% contain interesting events.

2.2 Typical frame

A typical frame is shown in figure 3. For each chamber the two stereo views are adjacent to one another. There are four fiducial crosses per view, each approximately 300 microns square. The quality and size of these fiducials have given considerable difficulty in locating them, as they are crossed by at most seven scan lines.

The data box, at the top of the frame, displays a pattern of lights representing binary digits. The pattern gives (1) the binary frame number and its complement; (2) the decay time and its complement (in the first 6 positions). The last light is present in both arrays and is used as a location marker on the frame.

3. The PROGRAM

3.1 The program system

The complete program system is subdivided into four sections:

1. determination of the systematic errors;
2. the location of all fiducials on the first frame;
3. the scanning and geometrical reconstruction of events;
4. The kinematical and subsequent statistical analysis of events.
3.2 Determination of systematic errors

It has been outlined before that the errors at the intersection points of the calibration grid are determined by a least squares fit to a set of precision microscope measurements. The errors on this grid are then transformed to errors on a uniformly spaced rectangular mesh and stored in a packed tabular form giving ease and speed of access when interpolating between the nearest four mesh points.

3.3 Location of all fiducials on the first frame

In the absence of a unique distinguishing mark on the frame, two pieces of information are required, namely, the relative positions of the chambers and data box on the film, and the absolute position (in LUCIOLE co-ordinates) of some distinguishing mark, e.g. a fiducial cross.

It has been noted that each chamber has four fiducial crosses per view. If the positions of all these fiducials are known (for example by finding the first frame of the roll with LUCIOLE or by measuring on a HEP), then a measurement of 3 of them, by LUCIOLE, is sufficient to determine the remainder. It was decided to accept 3 fiducials out of a possible 4, where each fiducial was known to lie within a small region on the frame.

This method was adequate when clean films were used, but if the film had previously been scanned it was quite useless since the search regions had more background (scratches etc.) than useful digitizings; the useful digitizings being from fiducials only 300 microns long.

In order to locate the fiducials at all, the histogram technique is employed, with the background levels set to minima, that is, accepting any pulse greater than unity in a bin and accepting an additive bin count of 3 for an arm of a fiducial. By these means fiducials as small as 100 microns, arm to arm, have been located and positioned within a least count in both scan directions.
To obtain a better chance of locating the fiducials it is desirable to reduce the search areas. Fortunately this can be done by locating the double-pulse on the data box, using a fairly large search area, and once it is found, predicting the search areas for the fiducials.

This method has led to a very high degree of efficiency, about one frame per thousand being discarded because the double pulse cannot be located, and, to date, none discarded because the set of 3 fiducials out of 8 cannot be found. The relative positions of the chambers, one to another, is constant from frame to frame on a roll of film, and also for considerable periods during the experiment.

Only four IEP calibration measurements should be required for the complete 100,000 pictures. The program, at present, has the option of either locating at least 2 fiducials per view or using the IEP calibration measurements.

3.4 The scanning and geometric reconstruction of events
The general procedure in this program is as follows:
a) locate double pulse in data box;
b) locate three fiducials;
c) predict the positions of the remaining 37 fiducials;
d) position each chamber, both views;
e) search in gaps for sparks;
f) associate sparks into tracks;
g) correlate track pairs in the two views of a chamber;
h) calculate equation of track in LUCIOLE co-ordinates;
i) correct track for parallax;
j) convert to space co-ordinates;
k) output results on magnetic tape.

The packed format of the LUCIOLE digitizations is not particularly suited to the IBM 7090 order code. So the LUCIOLE digitizations corresponding to regions of interest on the frame are converted to an HPD-type data
format. This has two advantages; it facilitates programming and enables existing HPD spark chamber routines to be used.

The routines for locating fiducials, locating sparks and associating sparks into tracks are HPD spark chamber routines with only slight modifications. When searching for sparks, it is desirable to have the chamber plates parallel to the scan lines, but this is not always the case. However, on the present experiment the problem was not serious, and for chambers slightly out of alignment it was sufficient to take an average line for the centre of the gap and then search in a conservative region each side of this datum. When the sparks are found which associate to form a track, each spark is corrected for LUCIOLE errors and the track equation is fitted by least squares.

To obtain the necessary matching of tracks in the two views of a chamber, each track in view 1 is compared in turn with each one in view 2, and the number of times that sparks do not correlate is found. This number is taken as an indication of the efficiency of the correlation. If there are more than 5 such mis-matches, a further investigation is made. There are 10 gaps in all chambers. So if the number of mis-matches exceeds 5, the correlation factor is set infinite and that pairing is not considered further.

For the remaining cases, if the number of matching sparks equals or exceeds the number of mis-matches the pairing is taken as a possibility.

A matrix is thereby constructed consisting of all possible correlations of all tracks. This matrix is then scanned for its minimum element; if this is unique it specifies a track pair and the appropriate row and column are deleted and the process is repeated.

Events containing multiple possibilities for tracks are discarded if no unique track exists in that chamber, but, to date, no event has been discarded for this reason.
The method of correlation works well for this particular experiment, but it clearly relies very strongly on the chambers being inefficient.

3.5 Kinematical statistical analysis of events

The equation for the missing mass is of the form:

\[ M_y^2 = \left( \sqrt{P_{\pi}^2 + M_{\pi}^2} - \sqrt{P_{\pi}^2 + M_{\Sigma}^2 + M_{\pi}^2} \right)^2 - P_{\pi}^2 - P^2 + 2P_{\pi} P \cos \theta \]

\( P \) and \( M \) are respectively the momentum and mass of the \( K^+ \) or \( p \), and \( \theta \) is the production angle.

The particular reaction to be considered for an event is determined by the deflection angle \( \alpha \) in the magnet.

If \( \alpha \geq 0.12 \) radians then the reaction considered is:

\[ \pi^- + p \rightarrow \Sigma^- + K^+ \]

If \( \alpha < 0.12 \) radians then the reaction considered is:

\[ \pi^- + p \rightarrow \pi^- + p \]

\( \alpha \), the production angle, is the angle between the incident beam and the product recorded through chambers 3 and 2. \( \alpha \) is the deflection of the product as it passes through the magnet. Knowing the field of the magnet, and the direction of the particle entering and leaving the magnet, the momentum \( P \), of this particle is calculated.

4. PRELIMINARY RESULTS

Figures 4a and 4b show histograms for the mass of the \( \Sigma^- \) as obtained from measurements using a digitized projector (Fig. 4a) and from LUCIOLE (Fig. 4b).

The resonance is at 1196 MeV; the interval is 5 MeV.

The measurements have been made on the same roll of film, although only part of the roll was measured on the digitized projector. The LUCIOLE results do not include the corrections for the systematic errors and they have used a IEF calibration measurement to give the relative positions of all fiducials.
These preliminary results, based on a small sample, indicate less precision in the LUCIOLE measurements than in the digitized projector measurements. The inclusion of the corrections for the systematic errors will indicate whether the complete system is verified or if further detailed analysis is required before production is begun.
References


Figure captions

Fig. 1  Distribution of errors in X and Y.
Fig. 2  Schematic diagram of experimental set-up.
Fig. 3  A typical frame.
Fig. 4  Histogram for the mass of Σ
 a) from digitized projector
 b) from LUCIOLE.
### Fig. 4a

<table>
<thead>
<tr>
<th>SIGN</th>
<th>CONTENTS</th>
<th>GROUP</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>11 11 11 XXXXXXXXX</td>
<td>0</td>
</tr>
</tbody>
</table>

### Fig. 4b

<table>
<thead>
<tr>
<th>SIGN</th>
<th>CONTENTS</th>
<th>GROUP</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>11 11 XXXXXXXXX</td>
<td>0</td>
</tr>
</tbody>
</table>

---

**Histogram of Missing Mass and Diagram of Deflection and Production Angles**

**Block 2**

- **Mode:** 94-441
- **Histogram:** 1
- **Missing Mass:**
- **Test:** 0 2 3 4 5 6

**Data Table:**

- **Channel:** 1234567890
- **Continuous Data:** 1234567890
- **Lower Range:** 0011111111
- **Upper Range:** 1111111111
- **Contents:** All Channels: 0, 0.00 Not Including Underflow, 133.00 Max Overflow, 0, 572 Entries in All.

---

**Diagram:**

- **Fig. 4a:**
- **Fig. 4b:**

---

**Notes:**

- **Page:** 3
- **Date:** 7/09/64
- **Dimensions:** 595.3x841.9

---

**Image:**

- **Image 0x-14 to 616x842**
DISCUSSION

DEUTSCH: I am struck by the fact that the program is almost identical in its structure to the one we use when we search for information on the scope rather than in the computer memory. For example we also find it extremely useful to orient oneself on the film first using some gross object which is bound to be found.

However, we find it better in the conversion of CRT co-ordinates to space co-ordinates to go directly without the intermediate step of using film co-ordinates. There are sufficient fiducials for fitting a reasonable polynomial which can be used for transformation directly to spark chamber space. Is there a specific reason why you do not do this? Perhaps you want to know about the instrument anyway?

POWELL: There has been great difficulty in finding sufficient fiducials due to the fact that the experiments have not been designed primarily with automatic measurement in mind. Apart from this I see no particular reason for not doing what you suggest. It is simply that we have not arrived at a full appreciation of how to do it.

ANDERS: I have a comment on the distortions which may not be evident from the table. In fact, the distortion is nowhere greater than 1% of the scan length. Also the distortions remain constant within one least count per hour.

In answer to Deutsch's question, we used a special grating for the calibration because the film was so poor in quality. Also we want to have the possibility of measuring other experiments.

LORD: How does the computing time compare with the measuring time?

HEAD: It takes 1.5 sec to scan a frame. We are hoping to take less time than this to process the event so that we can use full buffering i.e. computing one event while reading in the next. At present we take longer than this because we use several debugging routines, and also we output intermediate results on magnetic tape.
PROGRAM DEVELOPMENTS IN THE CERN F3D SPARK SYSTEM

P. M. BLACKALL, P. ZANELLA
CERN, Geneva

(presented by P. Zanella)

ABSTRACT

This paper describes the current program developments at CERN for the automatic analysis of a second spark chamber experiment using the HPD flying spot digitizer.

1. THE EXPERIMENT

1.1 Description of the experimental arrangement

A preliminary report on the experiment has been presented at the International Conference on High-Energy Physics (Dubna 1964) by P. Astbury et al. 1 Figure 1 shows a schematic diagram of the experimental arrangement.

A high-energy negative beam from the CERN PS traverses a beam telescope with beam defining counters. Two threshold Cerenkov counters are used to separate $\pi^-$, $K^-$ and $\bar{p}$ in the beam. Four small spark chambers and a bending magnet define the path and the momentum of the incident particle. The hydrogen target is surrounded by a lead scintillator sandwich anti-coincidence counter. The secondaries are detected in a large magnet spark chamber which has a useful volume of $60 \times 67 \times 170 \text{ cm}^3$ in a field of 10.7 kG and which contains 72 gaps.

Figure 2 shows a schematic drawing of the magnet, the magnetic spark chamber and the optical system. Two stereo views of the spark chamber are taken with two lenses on 70 mm unperforated film. The stereo
angle is $12^\circ 30'$; the demagnification factor at beam level is about 26.

Reactions of the following types are being studied:

(1) $K^- + p \rightarrow \bar{K}^0 + n$ at 9.5 GeV/c

(2) $\pi^- + p \rightarrow K^0 + K^0 + n$ at 9.5 GeV/c

Figure 3 shows two $K^0_1$ meson decays from the reaction of the second type. About 800 events of this type have been photographed in February 1964 together with 500 events of type 1. They have been measured on IEP equipment and have been analysed by the REAP - THRESH - GRIND chain of programs.

The preliminary results indicate that the system is satisfactory. So, it is planned to take a large number of pictures starting in December 1964.

Justifications for automatic measurement with the HPD are the desire for a greater accuracy and the pressing need for a shorter analysis time. It takes at present about 40 minutes to measure an event by hand and 30% of the measurements have to be repeated.

1.2 The frame format

Figure 4 shows another double neutral event and a typical frame format. The length of the frame is 130 mm. On the left hand side a set of lights indicate the picture number and the triggering mode in a binary code for automatic recognition. The image of the large magnetic spark chamber surrounded by fiducial crosses occupies the central part of the frame. It consists actually of six chambers each containing twelve 8 mm gaps separated by 12 mm thick plates. On the right hand side are placed the two $90^\circ$ stereoscopic views of the four beam chambers. Since the HPD is not yet equipped with a 70 mm film transport, each roll has been divided down the middle into two 35 mm films, each containing one view. Because each view contains all the necessary features such as Brenner marks and data box, there is no difficulty in analysing each such 35 mm film strip separately.
2. **THE ANALYSIS OF THE DATA**

2.1 **Pre-scanning**

Due to the small events-to-pictures ratio (about 1/15, 1/40 for events of type 1, 2 respectively) and the relatively large HPD measuring time (approximately 40 sec for the 2 views), pre-scanning is economic for this experiment (average time taken is 10 sec per picture).

The only information needed by the computer is the number of the picture containing a good event. Further desirable information may easily be provided, such as the type of the event (one or two neutral V's), the number of the gap where an apex appears to be and the number of background beam tracks traversing the magnetic chamber.

2.2 **The automatic measurement**

The computer reads from magnetic tape the scanning information for each event to be analysed and instructs the HPD to advance to the appropriate frame. Measurement is then started and the frame number immediately recognised to make sure that the correct picture is being processed. In order for the flying spot to cross a spark at least three times, a scan line separation of 90 microns has been chosen for this experiment. A circular buffer, at present of 14,000 words, is set up in the IBM 7090 memory to receive the digitisations from the on-line HPD. (About 15 - 20,000 per view.) Data reduction is performed simultaneously with the measurement. The program continuously checks that the data are not overwritten before processing occurs.

2.3 **Data reduction**

Fiducials and sparks are found, while measurement is taking place, by the subroutines already used for the analysis of the previous experiment\(^2\), \(^3\). A slight improvement was made to the spark finding routine in order to reduce the background due to scratches and other objects on the film. In fact, as suggested by the geometry of this chamber, a sufficiently long object can be distinguished from a spark by its presence also in the thick plate region.
The reconstruction of the linear tracks in the beam chambers is performed by the existing routines described in references 2 and 3.

2.4 Pattern recognition in the magnet spark chamber

Association of sparks into tracks and recognition of neutral V's occurs while the HPD advances the film to the next event.

First of all, displacement of sparks from tracks due to magnetic and electric fields are removed or, at least, minimised by applying systematic corrections to the co-ordinates of the spark images. In particular, the magnetic field staggering is removed by adding or subtracting an average staggering correction term to the spark co-ordinate (parallel to the plates). The electric field staggering is reduced by displacing the other co-ordinate (x), nearer to the cathode.

The sparks found in the chamber image are scanned backwards (from gap 72 to gap 1) by the track following program. At each gap, every spark is either assigned to a previously established track or to a new "starting-trial". The threshold for an established track is at present fixed at 4 sparks.

Coarse predictions from linear extrapolations are used in the initialisation stage. More accurate predictions from a parabolic least squares fit are used for the track following. Due to the regularity of the x-co-ordinates, each prediction is calculated as a simple linear combination of the x-co-ordinates of the sparks found in the 'n' preceding gaps. Missing sparks are replaced by their corresponding predictions. Errors due to residual staggerings are minimised by using an even number for 'n'. A maximum of n = 6 is used and it has proved to be quite adequate. Whenever a very precise prediction is needed, for instance when two or more tracks are crossing or when two tracks are approaching an apex or when a track has several consecutive missing sparks, then an ellipse is fitted to all the sparks of the known portion of the track.

In fact, due to the optics of the experiment, the track images in the magnet spark chamber are best approximated by ellipses. The ratio,
between the major and minor axes, and the axis orientation of the ellipse are known with sufficient accuracy. Such an ellipse may be represented by an equation of the type:

$$x^2 + Ry^2 + ax + by + c = 0$$

The three parameters $a$, $b$, $c$ are determined by forcing the curve through three given points on the track. These are chosen as widely separated as possible on the parabolic segment which best fits in the least squares sense all the sparks previously associated with the track.

Since all the tracks in a view are followed simultaneously, information on their relative behaviour is instantaneously available to the program, which can therefore anticipate the track crossing regions and the apices and can then decide the suitable prediction mode. The uncertainties associated with each kind of prediction have been determined experimentally and have been used to define the acceptance regions for sparks in succeeding gaps.

Apices are recognised in the gaps indicated by the scanner wherever two tracks having curvatures of opposite signs are closer than a preset maximum separation. Solution of ambiguities, if any, is postponed to the end of the track-following (gap 1). For ambiguities which cannot be solved at this level of the program the decision is left to THRESH or GRIND which will be used for the geometrical reconstruction and the kinematical analysis of the events, respectively.

Figure 5 shows one ambiguous case which cannot be resolved at this level of the program.

2.5 Correlation of the two views

The part of the program which deals with this is not yet tested. Correlation of neutral V's is based on the comparison of apex positions. Ambiguities are solved by matching spark patterns. The arms of the V's are correlated by comparing the signs of the curvatures.
2.6 How to handle the rejections

Although it is expected that a reasonably good efficiency will be achieved with the system, no doubt there will be a small percentage of non-recognised events.

The idea is to avoid measuring these pictures on IEP's and to make use instead of the HPD measurements available in the computer memory. To this end a set of subroutines is being written which will store the spark array of the rejected events on magnetic tape and will provide an off-line plot, on a Calcomp plotter, of the recognised tracks. (It takes about 5 min/plot.) Looking at such plots and knowing the reason for rejecting the event, one can provide the program with supplementary information to obtain the correct recognition. Doing so, remeasurements are avoided and off-line reprocessing should bring in a good fraction of the lost events.

3. PRELIMINARY RESULTS

3.1 Experience with the system

The main differences between this and the previous spark chamber experiment analysed on the HPD are:

1) The number of digitizings is increased by a factor of 3 - 4.
2) Tracks are non-linear and are complicated by magnetic and electric field staggerings.
3) Events are more complex.
4) Films are pre-scanned.

Problems arising from the first three points seem to have been satisfactorily solved, but human intervention at point 4) is proving to be quite unreliable. It is certainly desirable to change the logic of the event recognition to give less weight to the scanning information, if any weight at all. In particular we shall try to suppress the indication about the position of the apices.
3.2 Performance of the program

So far only a small sample of the double neutral V events have been measured by the HPD. A total of 31 events have been submitted to the program which recognised all of them correctly. The recognition was limited to one view of the magnetic spark chamber. Figures 6 and 7 show some other typical double neutral-V events correctly recognised by the program. We consider these first results as quite encouraging and expect that the production program will be ready at the beginning of December 1964.

ACKNOWLEDGMENTS

We would like to thank J.C. Lassalle and P. Nielsen for their help in the programming and L. van Loon for his valuable help with the HPD.
References


3. P.M. Blackall, B.W. Powell and P. Zanella, "The automatic analysis of 200,000 spark chamber pictures using the CERN HPD", paper presented at this Conference.
Figure captions

Fig. 1  Schematic diagram of the experimental set-up.
Fig. 2  Schematic drawing of the magnet, of the magnet spark chamber and of the optical system.
Figs. 3 - 4 - 5 - 6 and 7 show typical photographs of two $K^0$ mesons decaying in the magnet spark chamber.
$S_{1,2,3,4,5,6}$, BEAM DEFINING COUNTERS
$R_{1,2}$, ROUND ANTICOINCIDENCE COUNTERS
$F_{1,2,3}$, FLAT ANTICOINCIDENCE COUNTERS
$\bar{C}_{1,2}$, CERENKOV COUNTERS
FOR SELECTION OF $\pi^-$, $K^-$, AND $\bar{\nu}$

BEAM MOMENTUM DEFINING MAGNET

BEAM DIRECTION SPARK CHAMBERS

$\pi^-$ TRIGGER = $S_{1,2,3,4,5,6}$, $C_1$, $C_2$, $R_{1,2}$, $\bar{F}_{1,2,3}$
$K^-$ TRIGGER = $S_{1,2,3,4,5,6}$, $C_1$, $C_2$, $R_{1,2}$, $F_{1,2,3}$
$\bar{\nu}$ TRIGGER = $S_{1,2,3,4,5,6}$, $\bar{C}_1$, $\bar{C}_2$, $R_{1,2}$, $F_{1,2,3}$

CERN MAGNET SPARK CHAMBER
TEST RUN
LAY-OUT OCTOBER, 1963

Fig. 1
PEPR STATUS REPORT

M. DEUTSCH

Massachusetts Institute of Technology, Cambridge

No formal paper on PEPR was presented. However, M. Deutsch described PEPR informally. In order to make the subsequent discussion more intelligible, we reprint here a paper on PEPR which was presented at the 11th International Conference on High Energy Physics held at Dubna, 1964. For a more detailed reference, see Methods in Computational Physics V (Academic Press, 1965).

A PRECISION ENCODING AND PATTERN RECOGNITION SYSTEM (PEPR)*

I. PLESS, L. ROSENSON, P. BASTIEN, B. WADSWORTH, T. WATTS, R. YAMAMOTO

Massachusetts Institute of Technology, Cambridge

M.H. ALSTON, A.H. ROSENFELD, F.T. SOLMITZ

Lawrence Radiation Laboratory, Berkeley

H.D. TAFT

Yale University, New Haven

(presented by H.D. Taft)

INTRODUCTION

PEPR is a computer controlled device to automatically measure and "pattern recognize" bubble chamber events on film. The precision element is a magnetically focussed and deflected cathode ray tube. The system has two logically separate functions. The first is to locate and recognize a track element which is defined as the image on film of a short (~ 2 cm) section of bubble chamber track. The results of this

* ) Work performed under the auspices of the U.S. Atomic Energy Commission.
recognition are the x, y and θ co-ordinates of the track element on the film. This is called the PR (Pattern Recognition) mode of the system. The second function is to measure the x, y co-ordinates of the center of this track element with an accuracy of one part in 40,000. This is known as the PE (Precision Encoding) mode of the system.

OPERATION OF THE HARDWARE

1. The PR mode

The pattern recognition of a track element is accomplished by an analogue technique. It is based on the fact that an efficient mask for identifying a line is another line. In particular, a line of light which can be oriented under computer control, both in angle and position, is formed on the face of the cathode ray tube. This line can then be swept either along the x axis or along the y axis. When this line falls upon a track in the film plane, the height and width of the output pulse of a viewing photomultiplier depends on the relative orientation of the line of light and the track element.

An analogue circuit called TED (Track Element Detector) measures the full width at half height of the output pulse. If this is narrow enough to be considered a track element, TED produces an information pulse. This pulse can then be considered to be the recognition of a track element. The time of the arrival of the pulse indicates the x and y co-ordinate of the track element and the known angle of the flying line on the CRT constitutes a measurement of the angle. The angular resolution of the system in both PR and PE modes is about one degree. The least count in the PE mode is 25 microns and the reproducibility is a fraction of that.

The basic PR logic is as follows. A PDP-1 general purpose computer with 32,000 18 bit words acts as the logic control unit. It communicates with a special purpose computer called the controller. By program control, the PDP-1 chooses a 2 mm by 2 mm square on the film
plane. It also chooses the angle interval (± one degree) and position interval (± one track width) inside the square it wishes to investigate. The controller accepts this information, sets up the required voltages and transmits them to the angle and position amplifiers which are connected to the precision CRT. The signals seen by the photomultiplier are then sent to TED which in turn sends its decisions to the controller. The controller then converts the applied voltages at this instant through its internal logic into co-ordinate and angle information which is then sent back to the PDP-1.

2. The PE mode

The precision encoding is accomplished by effectively measuring the deflection currents at the time of recognition. These measurements are made with a reproducibility of approximately one part in 40,000. Figure 1 shows a block diagram of the complete PEPR System.

The precision encoding sweep has one tenth the length of the PR sweep. However, the basic control logic is the same. In the PE mode, the least count is 2 microns and the reproducibility also about 2 microns. Figure 2 shows a plot of measurement data versus time in the PE mode. This figure indicates that the total mechanical and electrical system drifts of the order of one micron per five minutes.

Figure 3 shows the result of measuring a straight line and plotting the deviation from a straight line as a function of distance along the line. It is seen from this figure that the error due to equipment fluctuations is less than 2 microns. Finally, Fig. 4 shows a PEPR reproduction of a bubble chamber picture scanned by the "area scan" program. All track elements in the original film were located and stored in the memory of the computer. After the scan a special program generated a short line segment at the position and angle of each element on a display scope, thus producing the figure.
PEPR Programs

The basic ingredients such as calibration, track element recognition, area scanning, track following and vertex location which make up the scanning and measuring scheme have been designed and programmed, and preliminary attempts to combine all of these have resulted in the successful recognition of interactions specified only as to their approximate location on the film. We shall give here a brief description of the logical function of the four major sections of the program which together form a complete pattern recognition and measurement system.

1. The black box

To achieve the greatest possible simplicity and generality in the computer programs it is necessary that all of the programs operate in an absolute and undistorted co-ordinate system in the film plane. Since the required corrections to the hardware are too complicated to be applied by analogue circuitry, a program has been written to perform all calibration functions as well as to code and decode data and perform all transfer of information in and out of the PEPR controller. This "black box" program is thus a software interface which converts the hardware to as ideal an instrument as is possible.

When scanning with a spot, at least four calibration routines are required in the "black box". A straightforward co-ordinate transformation has been constructed to calibrate the main deflection system. Using all terms through cubic, it has been possible to obtain an r.m.s. deviation of approximately 3.5 microns for this function over a four inch diameter area on the scope face. Pin-cushion and other distortions must be removed from the high-frequency sweep length as well as from the direction of the sweep. Further, the point on the sweep at which the spot coincides with the main deflection co-ordinates must be determined for each mode of operation. Additional corrections must be programmed
when scanning with a flying line segment. The correction to the angle of the line segment is a function both of position and angle. Finally, any small misalignment of the diquadrupoles which distort the beam into a line produces a pinwheel effect when the line is rotated, and this must also be calibrated.

2. Element recognition

The raw data from the FEPR hardware consists of strings of hits, i.e. co-ordinates along the sweep axis at which track signals are obtained. The problem of correlating all the hits within a given scan cell and producing sets of single co-ordinates and angles, called "track elements", is a pattern recognition problem on the most fundamental level.

Since the sweep at any given angle may generate several hits, a logical scheme has been programmed which matches hits on consecutive angles in such a way as to generate a minimum number of incomplete sequences. The groups of correlated hits thus produced are the track elements which form the basis for the area scanning and track following program to follow.

3. Area scan

Although some systems for the complete analysis of bubble chamber photographs may be worked out using track following routines alone, an area scan is required if neutral decays or recoils are to be located. The object of the scan is to search a specified area as completely and rapidly as possible and to produce an output as possible in which at least one element on every track segment in the area is identified. In order to simplify the logic of connecting track elements the area scan is broken up into two $90^\circ$ scans which together cover the entire area and angular range.

As the scan advances, a 20 word track bank is set up for every new element which is found which cannot be associated with a track bank already formed. At the end of the scan the banks contain the beginning and
end co-ordinates and angles of each track segment as well as associated information. Because provision must be made to accommodate a large number of tracks simultaneously the prediction and correlation techniques used must be less sophisticated than those used in the track following routine but are sufficiently powerful to produce an output meeting the specifications stated above.

Because of the interrupt feature, the area scan program has been designed so that processing of the data proceeds simultaneously with the scan itself. In this way very little of the time during which the scan is covering blank area is wasted and most of the time necessary for area scanning with the PDP-1 computer is spent in actual computation and data processing.

4. Track following and vertex recognition

Starting with the track banks produced by the area scan, a track following routine has been written which by means of successive extrapolation and correlation, is able to follow a track across a 35 mm picture in about 1/2 second. This routine is also able to extrapolate across gaps which are due to the presence of large areas of confusion on the film, and is able to follow tracks through the interface between the two modes (90° apart) of the area scan. The extrapolation used is circular, and due to its higher intrinsic accuracy, co-ordinate information only is used in making this extrapolation. Naturally angle information is used in matching to new data since signal to noise is greatly improved by this information. In order to locate the end point of the track segments with greater accuracy the flying line segment may be reduced in length, and routines are under development which are designed to achieve an accuracy of a single bubble on these end point locations.

A minimum guidance system involving all of the programs described above is currently being set up. In this system the approximate locations and the number of prongs of a given interaction are specified by a pre-scan.
The area scan is first used to locate the track segments. "Through" tracks are discarded and those which begin or end in the area of interest are then track followed and the end points are located. The vertex locating program is then applied and, if successful, may be followed by high precision measurements on tracks leading into and out of the vertex. Finally a gap length distribution is obtained by scanning across the track every 25 microns.

A primitive version of the minimum guidance program, called "PEPR III", has been constructed. Its performance is illustrated in Figure 4. It should be stressed that this minimum guidance program does not differ in any essential way from a complete area scan and can be extended to handle bubble chamber photographs with no pre-scanning if the resulting loss of speed is not important or if faster computers are attached to the hardware. Other approaches to the problem which are appropriate for particular experiments will be developed in the future.

ACKNOWLEDGEMENTS

The development of the PEPR system has been the result of the combined efforts of many people. We should like to acknowledge in particular the invaluable assistance of Mr. Lars Monrad-Krohn, Mr. Tor Lingjaerde, Dr. Eugenio Sartori, Mr. Jack Sharp and Mr. Ray Kenyon, all of whom have contributed in an essential way to the project.
Figure captions

Fig. 1 Block diagram of the PEPR system.

Fig. 2 Reproducibility of the PE mode of PEPR.

Fig. 3 Measurement in the PE mode of 1000 points on a straight line. The vertical scale is magnified 500-fold with respect to the horizontal one. The smooth sagitta of about 50 microns is systematic pin-cushion distortion, later corrected by the "black box" program; the "point scatter" of 2 microns can also be decreased by measuring each point several times and averaging. The width of the line being measured was 25 microns.

Fig. 4 Performance of PEPR III on a bubble chamber photograph: (a) actual photograph of one view; (b) display of a "TV scan" (using a spot) around a grid region associated with the event at the left; (c) display of the area scan (using 2 mm line segment) in the grid region; (d) end points found by the area scan (104 end points in all); (e) tracks selected by "vertex finder" program after track following; (f) the upper display shows the elements of the first outgoing track. The base line merely joins the first and the last elements found. The x-axis units are 512 main deflection counts; the y-axis units are one main deflection count. The lower display shows the difference between track elements as actually found and as predicted by the track follower routine. The x-axis is the same as above; the y-axis units are one main deflection count. Note that the "point scatter" is only a few microns.
Fig. 1
Fig. 2

Fig. 3

Measurement Deviation
Units: 25 microns

Measurement Distance
Units: 1 cm.
DISCUSSION

MARR: Could you give us some idea of the overall processing speed anticipated?

DEUTSCH: About 10 sec, as an order of magnitude. It is possible that one will do better, as we are at least half computer limited.

MARR: How were the pre-scanning measurements prepared? On the scan table?

DEUTSCH: I don't know. I don't think this will cause any great difficulty.

TURNILL: How do you find a "V" using a line? As I understand it you follow the beam tracks.

DEUTSCH: In the example which I showed, we were not following beam tracks. What we were doing was to follow all tracks which cross a rectangle placed about a point measured in the pre-scanning. This is of course logically equivalent to following all beam tracks if one moves the boundaries to the end of the film.

EDMONDS: The 25 μ accuracy of measurements, which you mention, presumably refers to the surface of the tube. However, a photograph is normally rectangular, for example 35 mm × 135 mm while the CRT is normally round or square. How does FEPR deal with this?

DEUTSCH: The precision of the device is, of course, given in terms of the CRT face. It is clear that if the optics is set up so that the scanning is very inefficient, then we will correspondingly lose. If the event does not take up the whole length of the picture, one might, by pre-scanning, decide to project only a part of the frame.

Otherwise, the fact has to be faced that the resolution given is for a 1/1 magnification, corresponding roughly to a 75 mm film.
EDMONDS: You have said that computers are becoming cheaper. Some people take the view that it will be more economical in the long run to use a simple TV raster scanning device and do the recognition of line segments in the computer, rather than have a special purpose device like PEPR. What are the current views on this?

DEUTSCH: This is, of course, the subject of a continuing debate. Approximately half of the cost of PEPR is for hardware which it is necessary to build anyway. What you are asking is whether it is preferable to spend $250,000 on a small computer, as we do, or to use the money to buy time on a larger computer.

I think that it will be much easier to achieve pattern recognition with a device, like PEPR, which is guided. Then, once the problem is solved with one device, it should be possible to program for any device. The economic question then becomes relevant.

KOWARSKI: This question should really be discussed in the concluding remarks which have been promised to you; and I do not propose to make my concluding remarks now. It is exactly what I called, last year, the tricky hardware versus tricky software question.

I do not think it is exactly an economic question. The reason why there might be some interest in using software alone, is not that hardware is so expensive but that hardware is considered not to be stable enough and to be difficult to make stable enough.

BUTLER: I think that the inclusion of a small computer may be justified simply on the basis of data flow through the system. The increased capacity is possibly due to the presence of another arithmetic unit working at the same time.
THE CERN PICTURE NUMBER PINGING ROUTINE

M. ROSENBLUM*)
CERN, Geneva

(presented by P. Seyboth**)
CERN, Geneva

Purpose of the routine

The picture number finding routine, a FAP subroutine of the CERN HAZE program, was written for locating and reading the picture number on bubble chamber pictures taken at the Saclay 81 cm chamber. The Brenner marks, the picture number and the HPD fiducials are contained in the data box at the beginning of the photograph. When advancing the film, the HPD stops on the Brenner marks, but the positioning accuracy is only of the order of millimetres. Therefore, the program first searches for the bits and then predicts the fiducial positions from their location, to avoid excessively large gating regions.

Format of the picture number

The picture number consists of two columns of equidistant, horizontal strips of lights (spacing 600 μ, length 500 μ approximately). The first column contains the frame number, the second the roll number. There are 25 bits in each column, each bank of five representing a digit in BCD code with one parity bit. The top group is always illuminated to make it easier to locate the bits.

Method of the routine

a) Location of bits

For all the digitizings in a scan line the distances from a line below to the left of the bits are calculated. After having been redu

*) On leave from Brookhaven National Laboratory, Upton, L.I., N.Y.

**) On leave from Max-Planck-Institut für Physik und Astrophysik, München
modulo the bit spacing, a histogram is made with these values. If when stepping through the digitizations in the scan line, a pulse appears in the histogram, it is assumed that the bits were detected. Using the stage co-ordinate of the scan line the HPD fiducial positions are calculated. Using the mean distance (reduced modulo the bit spacing and accumulated in the histogram) of the pulse from the reference line, the program updates the predicted bit positions. If no pulse is found the histogram is zeroed and the next scan line is processed.

b) Decoding the bits

Starting with the scan line on which the bit pattern was found, the digitizations of six scan lines are used to tally a second histogram which has one bin for every bit in the column. When a digitizing lies within ~ 30μ of the predicted position of a bit, a count of one is added to the appropriate bin. The digitizing is ignored otherwise. When there are at least 3 hits in a bin the corresponding bit is assumed to be on.

The pattern of each bank of five bits may be a legitimate BCD digit with parity bit, may be made into one by a single correction, or might contain at least two errors. The program takes appropriate action by executing the instruction in a table at the place corresponding to the 5-bit binary number. In the first case the parity bit is removed and the BCD digit is stored, in the third case an error word is set to indicate failure. In the second case a correction is attempted by inserting a bit which is known to be weak. If the pattern does pass the parity test now, it is considered correct. Otherwise the original pattern is matched against the expected one and if there is a discrepancy of only one bit, the latter is taken to be correct. A different error word is set when a correction was made.
When the program finds a wrong picture number, it gives an appropriate film move order to the HPD and starts again; if it has to do this too often, or when the bit pattern was unreadable, it stops and asks for operator intervention. If the bits were not found an order is given to the HPD to re-position the frame and the program tries again.
DISCUSSION

BLOCH: Since the program was written, the number box of this particular bubble chamber, which is the 81 cm hydrogen bubble chamber from CERN and Saclay, has been modified according to Rosenblum's ideas.

SEYBOTH: Even so, the other method has proved to be very reliable.

DEUTSCH: The obvious thing to do is to take proper pictures. It is usually true that you only have to watch two or three bits, because the numbers do not change at random; they change consecutively. When we had trouble with the numbers we usually ignored all but the least significant four bits.

STRAND: Do you ever have the situation where you predict the position of the road fiducial and fail to find it?

SEYBOTH: We have not had that yet.
THE STATUS OF HPD PROGRAMS AT
BROOKHAVEN NATIONAL LABORATORY

K. ABRAMS, J. COCKRILL, P. CONNOLLY, D. CRENNELL, F. HOUGH,
L. POTTER, D. STONEHILL, R. STRAND and W. THOMPSON

Brookhaven National Laboratory, Upton, L.I., N.Y.

(presented by R. Strand)

INTRODUCTION

This paper describes the Brookhaven National Laboratory, Hough-
Powell Digitizer system for rapid processing of bubble chamber pictures.
The HPD hardware and most of the software have been developed by members
of the BNL Bubble Chamber Group. The software runs on the IBM 7094 and
CDC 924 computers of the Applied Mathematics Department (AMD). This
paper emphasizes software and experience that was gained from the first
production running of the system.

The heart of the HPD system is the flying spot digitizer (FSD)
which traces a master scan over bubble chamber photographs. Bubble positions,
accurate to a few microns are transmitted to the memory of an IBM 7094 com-
puter. With the help of previously rough-digitized three-point roads, the
computer is able to locate the tracks within the set of all bubble positions.
Average points are obtained which simulate the output of the conventional
manual spot digitizers. In addition, bubble density information is obtained
for each track. The rest of the analysis, scanning, geometry and kinematics,
use the Brookhaven BCG data processing programs.

This paper is organized into seven sections in order of increasing
amount of detailed description. Readers with HPD experience should be able
to understand the HPD system by referring to the eight figures at the end of
the text.
Section I contains a general description of the whole HPD system as it is applied to an experiment. Picture quality requirements, event processing rates and event rejection percentages are reported. In the next three parts of this paper the HPD system is divided into three parts. Each part is discussed in more detail. The road making operation is discussed in Section II, the FSD measurements in Section III, and the kinematics and event selection in Section IV. Section V contains a description of the FSD control program called HAZE. The use of FSD measured bubble density to compute the best event interpretation(s) is presented in Section VI. In Section VII future plans for the HPD system are presented.

I. PRODUCTION EXPERIENCE WITH THE 35 MM HPD SYSTEM

The first large scale experiment to be processed with the HPD system is a study of $3.5 \text{ GeV/c } \pi^- + p$ reactions in the BCG 20 inch hydrogen chamber. The picture quality of this chamber is well suited to the elementary track finding methods used in the HAZE program, which controls the FSD. HAZE is able to cope with only a limited number of beam tracks. About twelve tracks per picture were taken. A fast beam pulse was used to provide uniform track age in the chamber. The bubble density was set to about twelve bubbles per centimetre in space. The HPD uses X-shaped fiducial marks on the chamber windows. The bubble chamber arc voltage was adjusted until these fiducial marks had the appearance of stopping proton tracks. The track contrast was brought to a suitable level for the FSD by adjusting the bubble chamber piston stroke or the arc delay. A binary data box is applied to each new photograph. The FSD film transport reads the binary frame number from this data box, and also stops the film with reference to a binary clock track on this data box. It was necessary to monitor the data box quality during the run. Also some of the pictures were tested for overall quality on the FSD.

A block diagram of the current 35 mm HPD system is shown in Fig. 1. The OFFMIF operation has not yet been used for production, and the first production runs have been made without its advantages of saving 7094 main frame time.
In Fig. 1 we see that the film is first scanned and predigitized. Three points are taken on each track in each of the three views to construct 400 micron wide circular roads which guide the HAZE program by eliminating most of the picture from consideration. We refer to this process as maximum guidance.

FILTER, the pattern recognizing subroutine in the HAZE program, attempts to find the correct track within the road. Crossing tracks, close parallel tracks, and noise in the road all contribute to the difficulty of track recognition. It was soon discovered that track failures were too frequent. Only about 80% of the events could be correctly measured with only the computer FILTER. At this time human guidance was brought to bear upon the pattern recognition problem. The rejected road contents were displayed on the IBM 7094 digital CRT. An operator filtered the road by telling the program where the track was located through the use of the IBM 7094 console keys. We called this process manual filter.

Manual filter permitted us to correctly measure 90% of the events. More important though, it helped us find bugs in the system. We were able to measure an average of 80 events per hour with manual filter on-line.

Next the manual filter operation was moved off-line from the IBM 7094 by using a magnetic drum on which to store the rejected roads. Special hardware permits an operator to display the rejected roads on a CRT. The desired co-ordinates can be marked by touching a light pen to the surface of the CRT. This operation is known as "off-line manual filter" or OFFMIF. HAZE is now able to measure an average of 100 events per hour and HAZE plus OFFMIF provide correct measurement of about 95% of the events.

Next in Fig. 1 POG and CLOUDY perform conventional geometry and kinematics calculations. Finally two versions of the Brookhaven BCG Suffolk County FAIR program are used to select the best fits for each event and to display physical parameters of event collections. The fit selection program is possible because of the ability of the FSD to measure bubble density.
About 10,000 events in 50,000 pictures have been measured with the HPD system. The logistics of this production are presented in Table I. The kinematic analysis is done in two passes through the computer. CLOUDY 208 prepares the lists of mass assignments and CLOUDY 209 constrains the events. Clearly most of the time is spent making the roads. This time could be perhaps cut in half by a "minimum guidance" scheme under which only vector points would be predigitized. It is expected that the next generation computer programs will make this possible. The FSD measuring time, which is most costly, can be reduced by optimizing the HAZE program and improving the FSD film transport and the binary data box on the film. Current stage speeds would limit the maximum attainable production rate to about 200 events per hour.

II. BNL 35 MM HPD SYSTEM PRIOR TO HAZE

We now examine the HPD system in more detail. First before the film is read by the FSD we are concerned with the preparation of the rough digitized roads. A block diagram for this part of the system appears in Fig. 2. We found that the road making operation is more complex than we had expected. In order to improve the road making efficiency we found it necessary to check the quality of the predigitizing with an extra computer program, MISTY. Our goal was to mount the film on the scan table only once and require all the events to be correctly predigitized.

We have used a three-roll unit of about 900 events in about 4,500 total pictures. We prepare checked roads for all of the 900 events before going on to the next three-roll unit. The three-roll unit was selected to optimize both tape storage and the amount of film that can be processed during a four hour run of the FSD. As a redundancy check the scan technician writes down the picture number as well as entering it in the picture number switches. This enables our computing aide to compare a scan list with the events as they are checked by the MISTY program, as shown in Fig. 3. The scanning and predigitizing rate of about 10 events per hour would be approximately doubled if no predigitizing were performed or if the number of measured points were
greatly reduced. At present three points on each track and two fiducial
marks are predigitized in three views. The set of events which meets the
requirements of MISTY and which can be fixed up by the data processing side
is about 95% of the total input. Our data processing side provides rapid
feed back of rejected or missing events to the scan technicians. In most
cases these events are predigitized again before the film has been removed
from the scan table. As a precaution the last few rejected events are
measured twice before the film is removed. In practice we find that the
combination of the MISTY program and our data processing side, effectively
trains our scan technicians. They simply improve their techniques until
their rejection rates are satisfactory.

We recognized the need to check for format errors in the MISTY
program. For example, each event has a format code (i.e. 102 for a two-
prong event) that specifies the correct number of tracks for the event.
Digitizer errors are also detected by requiring approximately equal point dis-
tribution over a moderate turning angle. A standard fiducial mark separa-
tion is required of each view. We will add to the MISTY program three more
checks. First we will require agreement of the vertex points. Secondly
we will make a test stereo reconstruction of each track in order to check
track matching errors between views. Thirdly we will reject those events
for which two outgoing tracks lie wholly within one road. We have not
solved this problem but we have developed a method to handle the case of
two beam tracks in the same road. We select arbitrarily one of the tracks.
Tests showed that this method has no effect on our measurement of the beam
track because of our editing method for beam tracks. After the three-roll
unit has passed the checks in MISTY it is used to update the list of events
for the experiment which is kept on a bookkeeping tape. The data also is
passed through the MIST program which separates it by view and writes it onto
three scan tapes for use in the HAZE program. MIST requires the events to
be in order and it prepares useful summary listings for each run. MIST also
accepts a table of instructions for the mass assignments for each event
topology. Mass assignments for kinematic analysis are generated for all
events from this table.
At this point in the analysis the three rolls of film are taken from the scanning lab. to the FSD lab. and the three scan tapes are stored in a rack ready for FSD production.

III. BNL 35 mm HPD SYSTEM FROM THE FSD THROUGH GEOMETRY

The HAZE program controls the 35 mm FSD as shown in Fig. 3. We will return to the HAZE program in more detail in section V. For the moment we should note that HAZE produces film geometry points and film bubble density measurements for about 90% of the track views that were pre-digitized. If the track is short, close to another track, or in general complicated with other tracks or with noise it may not be found by HAZE. In this case all of the points within the road are saved for human inspection as was mentioned in section II. A failed road's contents are displayed on a CRT. With a light pen, a technician can mark those co-ordinates which belong to the track. This operation is under control of the program OFFMIF which stands for off-line manual filter, as shown in Fig. 3. The associated magnetic drum is connected to one data channel of the BNL 7094 computer. The OFFMIF program is called into operation by the technician's pushing an interrupt button for the data channel. At the end of the next Brookhaven AMD monitor job OFFMIF is loaded into the 7094 memory and control is passed to it. If the drum is empty OFFMIF fills it with 32 failed-roads' contents and returns to the monitor system. The technician is able to look at the roads in order and fix or reject each track with the light pen. When the 32 tracks have been examined, the interrupt button is again pushed. OFFMIF is called into core and it reads the 32 corrected roads from the drum and loads the next 32 failed-roads. The corrected roads are merged into the normal HAZE output record. This operation takes about half an hour per drum load.

The drum CRT performs two functions of great importance. In the first place it permits us to achieve a production system before the pattern recognition problems of maximum guidance are satisfactorily solved. Secondly, it tells us how to solve our difficult pattern recognition problems in the HAZE program and it enables us to monitor the effect of HAZE program changes.
HAZE and OFFMIF run with one view at a time. When the three views of our unit of three rolls of film have been measured the measurement quality is checked by a test geometry reconstruction. After OFFMIF about 5% of the events fail to be reconstructed because a track was not properly measured in one essential view. To recover these events, a measuring technician measures the required track in its view along with necessary fiducial marks. These data are merged with the main FSD track data record in the program MTR which stands for multiple track re-run. More than one track-view can be measured and so merged. This program runs on the Brookhaven AMD CDC 324 computer. Four hours of FSD production produces enough work for one manual digitizer shift. OFFMIF or MTR may one or both be omitted after HAZE since all three output formats are identical. We have adopted these methods in order to have 100% event transmission without passing the film through the FSD a second time.

The final FOG run for geometry information terminates the measurement phase of the system. At this time a calculation of the cosine of the angle between the flying spot velocity vector and the track direction is performed in each view. These data are stored in the FOG output records. Each track's bubble density must be corrected by this factor. Since bubble density measurements are in an experimental stage, the raw data along with these cosines are preserved for later study in the Suffolk County FAIR program. More details of the bubble density methods will be presented later in section VI. The output from the FOG runs is merged into one logical FOG library for input to the kinematic analysis programs.

IV. **BNL 35 MM HPD SYSTEM FROM KINEMATICS THROUGH ANALYSIS**

The CLOUDY kinematic analysis is performed upon a FOG library which may include any number of events in numerical order. The CLOUDY output library contains the kinematic analysis for each desired mass set as shown in Fig. 4. In our experiment frequently the kinematic analysis favours two or more of the mass sets for each event. A typical example is the substitution of a positive pion for a real proton track followed by a successful kinematic
fit with a missing neutron. Indeed all bubble chamber experimentalists have at one time or another resolved mass assignment ambiguities by looking at track bubble density on a scan table. We have written a fit selection program which uses the FSD bubble density information and computes the resolution of ambiguities as a physicist might by looking at the scan table. The fit selection version of the Suffolk County FAIR program has been rather controversial, but we have attempted to select the fits step by step as a physicist would select them. FSD bubble density measurements are compared with the measured track momenta and their masses in the usual way in order to be able to select the best mass assignments for an event. The mass fits are then deployed onto their respective tapes as shown in Fig. 5. The bubble density methods are discussed further in section VI. Suffolk County FAIR is used to produce histograms and other plots of the data for a particular reaction. For each requested distribution three plots are prepared based upon the unambiguous sample, the ambiguous sample and the total weighted sample.

V. MAXIMUM GUIDANCE OF THE FSD WITH HAZE

A block diagram of the major features of the HAZE program is shown in Fig. 5. We will outline the HAZE program here. Readers who are not interested in the details of HAZE will be able to skip the next three subsections which will describe our HAZE program in more detail.

At the start of HAZE in Fig. 6, the road information is read from the scan tape. Next HAZE asks the FSD for the frame number which was read from the scan tape and for a scan of the picture. HAZE then waits for the FSD to fill the first 2000 word buffer. While waiting, a routine tests the format of the FSD data and keeps records of mistakes or prints a message when the data becomes impossible.

When the first 2000 word buffer is full, HAZE begins real time competition with the FSD to process one 2000 word buffer before the other is filled. During this real time period the road fiducials and glass fiducials are found as they occur in the 2000 word buffers. Also in real time the
subroutine GATE fills the roads with 20 points and then calls on FILTER to histogram the points and finds the tracks. When the 2000 word buffer is exhausted HAZE returns to the pause loop and the FSD data format test.

At the end of the scan the GATE plus FILTER combination handles the last part of the picture. Then HAZE checks the results of the scan. The so-called servo-subroutines adjust the road laying parameters so that the tracks are centred within their roads on the average. The tracks for this event are checked for bad master points or intolerably long gaps. Rejected tracks are flagged and their road contents saved for visual inspection and manual pattern recognition as was already discussed. If an orthogonal scan is required HAZE asks the FSD for an orthogonal scan. Tracks which make 45 degree angles or greater with the long axis require an orthogonal scan. Otherwise the event is written as one record on the output tape and control returns for the next picture.

1. Pre-Real Time HAZE

In more detail we consider that part of HAZE before real time as shown in Fig. 6. In this part of the paper we assume that our readers are program experts so that the Fig. 6 will tend to speak for itself. On the average HAZE seems to spend about 40% of its real time in the pause loop waiting for a 2000 word buffer to be filled. The rest of the total real time is divided approximately into 50% for GATE and 10% in FILTER.

2. Real Time HAZE

In Fig. 7 the real time part of HAZE shows two reasons for halting a scan and asking for a re-scan. If the program failed to keep up with the FSD, a slower re-scan is requested. The stage is asked to move at half speed and the co-ordinate flow circuits are asked to transmit only alternate passes of the flying spot or "one-half density". In this way the peak data rates are halved while the scan line separation is held sufficiently constant. Should the road fiducial remain unfound, the film is moved in and out one frame and the picture is rescanned. This occurs if the film fails to stop properly due to substandard binary data box quality or hardware failure.
Fiducials are sought by the H-10 subroutine which searches a histogram of digitization taken along the 45 degree slope of the fiducial arms. The fiducial arms must be straight for this procedure to work reliably. In Fig. 7 we also see that FILTER calls subroutine DIAG to save all of the gated points for possible later use by a manual filter technique.

FILTER collects bubble density information from the histogram pulse in each twenty point byte. In real time no correction is made for the angle between the flying spot and the track. The average angle based on the chord is determined in the geometry program. A chord to arc length correction will be made in the geometry program. We have found the edge of a photograph to give useful geometry data but erroneous bubble density data since the bubble image size is shrinking to zero. We now take bubble density information only within a smaller area. At the merge of two tracks we do not take bubble density data for that byte. No bubble density data is taken if there is a long gap in the string of track points. The bubble density data is an accumulation of the total number of scans, hits and gaps. Each digitization is a hit and a gap is composed of one or more contiguous misses.

Our FILTER routine initializes a track in its road with some difficulty if there are many tracks in the byte. For this reason we scan the pictures against the beam direction so as to allow track initialization under the more favourable conditions away from the vertex.

3. Post Real Time HAZE

After the scan HAZE checks the result as shown in Fig. 8, a parabolic fit on the film checks for bad points located transverse to the track by more than 30 microns. After the track passes the transverse test a check is made that no gap of more than 30% of the track is devoid of master points. If the track passes this longitudinal test it is accepted.

Rejected tracks may be inspected and repaired if possible by on-line manual filter or by off-line manual filter as mentioned in section I.
In order to make the roads fit well, two separate routines serve the Y-magnification and the rotation angle and X-magnification respectively as shown in Fig. 8. In both routines the rough scan table co-ordinates are compared with the precise FSD co-ordinates in order to cause the first to agree with the second on the average. These servo-subroutines have solved our road-fit problems. The film is stopped by the FSD when the correct binary data box is read. A search box for the road fiducial is placed in a position determined by the average position where the road fiducial is found. This is done by the RFOO subroutine shown in Fig. 8. These servo-subroutines are good examples of how production problems can be solved by software instead of by hardware.

VI. COMPUTED FIT SELECTIONS

We have outlined the fit selection process in section I. We will now describe this important use of FSD bubble density measurements in greater detail.

After some hardware and software debugging we have been able to measure bubble density with the FSD to a 15% standard deviation. This result was obtained by selecting a sample of 4-constraint events that had unambiguous kinematic fit. This distribution of the percentage of difference between the measured and expected bubble densities is approximately normal and its standard deviation is about 15%.

In each view the measured bubble density is corrected for the number of spurious misses per hit. Our current data indicates that there are three spurious misses per one hundred hits.

The next part of the bubble density calculations concerns itself with checking the consistency of the bubble density measurement in the three views. To begin with, those tracks which dip by more than 60° are omitted from the analysis. From the mass and measured momentum the predicted chamber bubble density for each track is computed from the inverse velocity squared law.
This predicted bubble density is then carried to the film in each view. The ratio of the predicted bubble density to the measured bubble density is obtained for each of the remaining views. The ratios are required to agree so as to indicate the correct answer by an agreement of two of the three.

The mass set is then tested by the minimizing chi-squared of a function of the misses predicted and measured while adjusting one parameter which is the bubble density of the beam averaged over the three views.

In our terminology an event which has a positive missing mass is completely acceptable. Those which have negative missing mass are considered to be no-fits.

The scan technicians are able to label a track with a special comment card. The fit selection program checks these comments which may specify as mass (i.e. \( \pi - \mu - e \)) or which may indicate that the beam track is overlaid on one view and will have two times minimum bubble density.

The program allows a bias in favour of 4-c fits over 1-c fits. The probability for the bubble density fit is multiplied by the probability for the kinematic fit to get an overall probability for the event. These probabilities are compared with thresholds and the best fits are selected.

Currently we find that human event selection agrees with the computed selection for fits. However, the computed selection yields twice as many no-fits as the human selection. We find by inspection that about 5% of our two prongs and 15% of our four-prongs are no-fits.

VII. SOME FUTURE PLANS FOR USE OF BROOKHAVEN BCG FSD’s.

The 70 mm FSD (ENL MARK II) is being turned on for the 30 and 80 in. chambers. One experiment is ready in each chamber. The 70 mm machine has been transmitting coordinates for a few months and it has been run under limited control of the HAZE program. We expect to be providing useful track measurements during the winter of 1965.
We will remain as IBM 7094 users after the next Brookhaven AMD general purpose computer is installed some time in 1966. Our FSD production time will increase from 30 hours to 60 hours per week during 1965. Eventually the Brookhaven DCG expects to use most of the IBM 7094 time. A large fraction of this will be used for our two FSD's.

ACKNOWLEDGEMENTS

This research has benefitted much from the support and encouragement of Dr. Ralph Shutt and Dr. Alan Thorndike of the Brookhaven Bubble Chamber Group. We have had much help from the members of the Brookhaven Applied Mathematics Department. They have helped write some of our programs and have operated the computer so as to make it accessible for debugging both our on-line hardware and our real time software.
TABLE I

Logistics of the analysis of 10,000 π⁻ + p events with the HPD system

<table>
<thead>
<tr>
<th>Operation</th>
<th>No. of Events (in units of 1,000)</th>
<th>Time to Process Events (in hours)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scan and predigitize</td>
<td>10</td>
<td>1,000</td>
</tr>
<tr>
<td>35 mm FSD</td>
<td>10</td>
<td>130</td>
</tr>
<tr>
<td>FOG</td>
<td>9</td>
<td>4.5</td>
</tr>
<tr>
<td>Two-prong events:</td>
<td>7</td>
<td>8.5</td>
</tr>
<tr>
<td>(CLOUDY 208)</td>
<td></td>
<td>2.5</td>
</tr>
<tr>
<td>(CLOUDY 209)</td>
<td></td>
<td>4.5</td>
</tr>
<tr>
<td>(S.C. FAIR)</td>
<td></td>
<td>1.5</td>
</tr>
<tr>
<td>Four-prong events:</td>
<td>2</td>
<td>7.5</td>
</tr>
<tr>
<td>(CLOUDY 208)</td>
<td></td>
<td>1.5</td>
</tr>
<tr>
<td>(CLOUDY 209)</td>
<td></td>
<td>5.0</td>
</tr>
<tr>
<td>(S.C. FAIR)</td>
<td></td>
<td>1.0</td>
</tr>
</tbody>
</table>
Figure captions

Fig. 1  BNL 35 mm HPD system.
Fig. 2  BNL 35 mm HPD system prior to HAZE.
Fig. 3  BNL 35 mm HPD system from HAZE through FOG.
Fig. 4  BNL 35 mm HPD system after FOG.
Fig. 5  HAZE.
Fig. 6  HAZE (pre-real time).
Fig. 7  HAZE (real time).
Fig. 8  HAZE (post-real time).
START HAZE
READ CONTROL CARDS
RETURN FOR NEW PICTURE
END ROLL FILM
READ PRE-DIG. TAPE
PREPARE FOR ALL FSD SCANS
ASK FSD FOR A NORMAL SCAN
CHANNEL D FILLS 2000 WORD BUFFER AND TRAPS 7094
WAIT FOR CH D TO FILL 2000 WORD BUFFER
SET CHANNEL D TRAP
FSD DATA
CHECK FSD DATA
TRAP
SEE CHART HAZE-2

Fig. 6
DISCUSSION

ROYSTON: What is the difference between the road fiducials and the chamber fiducials?

STRAND: The road fiducials are encountered before the picture in the HPD scan. They are predigitized on the scan table and are used to calculate the transformation from the scan table co-ordinate system to the HPD system. The chamber fiducials are used for reconstruction of the event.

BLOCH: Do you think that a least count of 60 microns, on the film, for the predigitizer is satisfactory?

STRAND: Our new predigitizers will have a 30 micron least count so that the filter program can use smaller roads. This will increase the efficiency of the filter program.

BLOCH: For your ionization measurements, how do you deal with tangent tracks?

STRAND: $\delta$-rays may cause spurious gaps by shading the track. For a perfectly black track the program allows for 3 misses in 100 hits. Parallel beam tracks are noted while predigitizing and are specially treated. Merging tracks are detected in the filter program and for the region where the tracks are very close the bubble density information is not used.

BENOT: Is a 40 micron scan line separation satisfactory for bubble density measurements?

STRAND: It is a compromise between scan speed and accuracy of bubble density measurement, the FSD bubble width is about 40 microns, so this is satisfactory.
The CERN version of FILTER is a routine of about 2000 FAP instructions. The purpose of FILTER is:

a) to compute the road parameters used by GATE, i.e. at the beginning of each slice the starting y-value, starting tangent, tangent increment per scan line, bin-width and road-width;

b) to improve the road parameters, using information already gathered along the road;

c) to separate interesting tracks from background and other tracks;

d) to compute master points for our spatial reconstruction program.

For a certain fixed number of scan lines, called a "slice", a track is considered to be a parabola, i.e. to have a constant tangent increment. The increment of tangent per scan line we call DDTAN. The slice length, a program parameter, is at present set to 32 scan lines. For highly curved tracks, and for short tracks we use a half slice mode, i.e. at present 16 scan lines.

GATE collects all points lying in the roads as defined by FILTER, and builds up the histograms $\Sigma l$, $\Sigma x$ and $\Sigma y$, where $\Sigma l$ means the point


**) On leave from Max-Planck-Institut für Physik und Astrophysik, München.
count per bin, \( \Delta x \) the distance from the beginning of the slice and \( \Delta y \) the positive \( y \) distance from the road edge. At the end of each slice \( \text{FILTER} \) is called.

\( \text{FILTER} \) searches at first the point-count histogram for pulses. It uses three thresholds, a background, an area and a width constant. All three parameters are functions of how many scan lines constituted the slice, which bin-width was used, etc. If a pulse (of \( p \) bins-width) is found we get the abscissa \( x \) of the master point in the form

\[
\bar{x} = x_{\text{beg}} + \frac{\sum \Delta x}{\sum 1},
\]

where \( x_{\text{beg}} \) is the abscissa at the beginning of the slice. The expression

\[
\frac{\sum \Delta y}{\sum 1}
\]

represents the mean \( y \) distance from the parabolic road edge. From the \( y \) value of the upper road edge at the slice end, the tangent (at the slice end) and the \( \text{DDTAN} \), the equation of the parabola \( y = y_p(x) \) used in that slice is known. The ordinate of our master point is then

\[
\bar{y} = y_p(\bar{x}) - \frac{\sum \Delta y}{\sum 1}
\]

At the beginning of each track \( \text{FILTER} \) starts off with the road constructed by fitting a circle through the three rough digitizings. When at least two master points have been found "track following" can be started.

Depending on the magnitude of \( \text{DDTAN} \), we use either a circle or a parabola for extrapolation. We postulate that the extrapolated curve passes through the point \( (x_2', y_2') \) with that tangent \( y_2' \), with which this master point has been found, and that it has at the point \( (x_1, y_1) \) and the tangent \( y_1' \), the tangent of slice 1, see Fig. 1.
The tangent used in slice \( 3, y'_{3f} \), is then the weighted mean of

1) the tangent obtained by extrapolation, and

2) the tangent of the circle through the rough digitizing.

The weighting factor on the latter tangent decreases like \( \frac{1}{n} \) (where \( n \) is number of found points) and is zero after 5 points have been found.

The roads are always centred around the predicted \( y \) value, using as a correction the difference between the last found \( y \) value and the centre of the road. With this new tangent we compute the new DDTAN, extrapolate linearly the old one and use their weighted sum as the DDTAN for the next slice.

If a point \( (x_{3f}, y_{3f}) \) has been found in the next slice

we correct at first \( y'_{3p} \) - let us call this \( \tilde{y}'_{3p} \) taking into account the difference \( x_{3f} - x_{3p} \) and using DDTAN. We now fit a parabola through the points \( (x_2, y_2) \) and \( (x_{3f}, y_{3f}) \) using \( y_2' \) and get the following simple formula for the tangent at \( (x_{3f}, y_{3f}) \).

\[
y'_{3f} = 2 \frac{y_{3f} - y_2}{x_{3f} - x_2} - y_2'
\]

The next extrapolation is then based on the corrected tangent \( y'_{3c} \), the weighted sum of \( y'_{3p} \) and \( y'_{3f} \), where the weight on \( y'_{3f} \) decreases with the number of master points found.
When looking at the flow charts of FILTER (Fig. 2 to Fig. 6), the following remarks might be helpful:

a) For our histograms we use three different bin-widths: 16, 8 and 4 least counts. We always start off with a bin-width of 16 least counts. The bin-width is then a function of the goodness of the pulses.

b) We always start with a road-width of 400 least counts. After two points have been found, we start track-following as explained above, and use a road-width of 240 least counts, because the resolution of our HPD is about 100 least counts, we cannot use smaller road-widths, if we wish to notice diverging tracks. If we have found a second track in our original road, we narrow the roads of both tracks so that they do not include the other track.

c) We always keep two histograms. Sometimes we combine them to a "double slice" histogram. The double slice mode serves two purposes; it is used for very sparse tracks and to split two diverging tracks. If we find no pulse or two pulses in the histogram, we set a flag and call GATE to process the next slice. When FILTER is called the next time for this track and if this flag is set, we combine the two histograms and look for pulses in this accumulated histogram. If we find any pulse(s), we search the old and the new histogram (with lower background constants) for pulses matching the double slice pulse(s) and compute the average point(s) if we find two double slice pulses, it is assumed that we have two diverging tracks and we follow both.

d) For the computation of ionization we keep, for each slice in which a point was found, the cosine, the number of hits and the bin-width used, all packed in one word.

e) Eventually we expect a FILTER rejection rate of less than 0.5 per hundred.
Figure captions

Figs. 2 - 6  Flowcharts of CERN FILTER Routine.
Fig. 2

Fig. 3
DISCUSSION

STRAND: Does the track following scheme use only the two previous master points, or the complete previous history of the track on which to base its prediction?

KRISCHER: The extrapolation uses only the two previous master points and tangents, but these tangents are corrected using information on the track up to that point.

BURBERRY: Do you have experience with highly curved tracks, i.e. stopping tracks?

KRISCHER: The program works satisfactorily on these tracks.
EXPERIENCE WITH THE CERN HAZE PROGRAM

J.M. HOWIE
CERN, Geneva

In this paper I present some of the results we have obtained at CERN from the measurements of bubble chamber photographs on the HPD. The people who have contributed to the various aspects of the analysis of the results are: A. Accensi, Mme V. Alles-Borelli, B. French, A. Frisk, H. Hannerfeld, W. Krischer, A. Moyer, L. Michejda, W.G. Moorhead, B.W. Powell, P. Seyboth and J. Seyerlein. Before presenting any numerical results I will first indicate what the main developments of the CERN HAZE program have been since last year's conference in Paris.

a) A new version of FILTER has been written, which now employs the basic philosophy of track following.

b) A routine has been written to deal separately with the beginning part of the track - the idea being to clear up any bad master points, which FILTER might have produced in this region. This routine is also designed to deal with short tracks and cases where FILTER produces too few master points. Elsewhere in this paper this routine will be referred to as the HH routine, the initials being those of the author of the routine, Harold Hannerfeld.

c) A diagnostic program has been written which produces a print-out of what is happening in FILTER, and it also indicates where all the digitizations are relative to the road edge.

The part of HAZE which deals with the abnormal scan was not properly tested until 8 weeks ago, when we had the first run of the HPD for bubble chamber pictures in the Abnormal Scan Mode. The bugs produced were easily dispensed with, and we have had a fully working program for the last 6 weeks. During these 6 weeks we have measured just over 500 events of the
5.7 GeV/c anti-proton experiment.

The events measured were nearly all 4-prong events. The beam is $\bar{p}$ at a momentum of 5.7 GeV/c and the chamber used was the Saclay 80 cm hydrogen bubble chamber. The interactions studied are anti-proton plus proton combining to produce a combination of protons, anti-protons and positive and negative pions, sometimes together with one or two neutral particles i.e.:

$$\bar{p} + p \rightarrow (\pi^+ \pi^- p p) + (\pi^0 n \bar{n})$$

These are various aspects of the results which we need to look at, such as success of the HFD measurements, comparison of results obtained from IEP and HFD measurements, etc, and the results have therefore been divided into three basic sections.

1. Taking only events measured on the HFD, we have made a breakdown of all the events and also all the tracks into either a success category or various failure categories.

2. A sample of events (68) have been taken which were measured both on the HFD and the IEP, and various comparisons between the results have been made.

3. A small sample of events were measured a few times on both the IEP and the HFD. From an analysis of the results for these repeated measurements, one can investigate the consistency of the HFD and IEP measurements, and also estimate the measurement errors in both cases.

1. Analysis Sequence of HFD Events

The path of an event which is to be measured on the HFD is given in Fig. 1. First the event is measured roughly on the MILADY scan-table; the scan-cards produced are then read by the program MIST, which checks the cards for any obvious inconsistencies. The output of MIST is a magnetic
tape (the scan-tape), which is used to tell the HAZE program which frames to select for measuring on the HPD and also whereabouts the event is within the frame. The track and fiducial measurements found by HAZE are then passed, via magnetic tape, into the program THRESH for the geometrical reconstruction of the tracks in space, and from there the event is passed into the program GRIND for the kinematic analysis. What happens beyond GRIND does not concern us in this paper.

Of the 500 events measured on the HPD, 157 have been analysed in detail after obtaining the THRESH results. Apart from the printed geometry results, we have also had the help of two very useful diagnostic outputs; the HAZE diagnostic which has already been mentioned and a plot of the measurements on the CALCOMP plotter of all tracks which either failed to converge in THRESH or had too large an error on the fitted radius (greater than 8%).

a) Analysis of tracks

A breakdown of the fate of all the tracks in the 157 events is made in Table 1. It has been found necessary to separate the tracks into two categories, beam and non-beam, because the results for beam tracks are much poorer than for the other tracks. The main reason for this difference is that there are many beam tracks on the film, usually very close together, and it is extremely difficult for FILTER to separate them. A further difficulty is introduced by the fact that, for this experiment, the operating conditions of the chamber were deliberately set to produce a fairly low ionization level (in order to make the job of measuring the ionization with the HPD easier), and thus many of the tracks, including beam tracks, tend to be quite weak.

The first column in Table 1 gives the total number of tracks and the second column gives the number of tracks which were reconstructed successfully. The criteria for success was (1) complete convergence in the least squares fitting procedure in THRESH and (2) the errors on the geometrical parameters all had to be small (less than 5%). A track which did not meet
these success criteria was counted to have failed. The remaining columns in Table 1 distinguish between the various reasons for a track failing.

S  The track failed because of an error on the scan tape. This column therefore includes all errors due to the scan table, the scan table operator and the MIST program.

F  The track failed because the FILTER routine in HAZE gave some wrong points.

H  The track failed because of faults in the HH routine. This column includes two types of errors, (1) where the HH routine found wrong points, even though all the FILTER points were good and (2) where the HH routine failed to remove a bad first point on the track, though it was clear from the digitizings available, that it should have done so easily.

W  The track has too few digitizings. This causes two possible results, either (1) FILTER finds no points because the track is indistinguishable from the background, or (2) another track, which is strongly digitized lies within the same search area (the road), and FILTER follows this wrong track. This latter case has only been found to occur with beam tracks. It is worth adding here that in the majority of the "W" cases, it was impossible from looking at the digitizings within the road, to tell even by eye that the correct track existed.

T  The track failed because of an error in THRESH.

N  The track failed, but it was impossible to give the reason why; i.e. the measurements looked perfect and there were no ambiguities in FILTER, but the geometry results were poor.

B  The track failed because of bad tape. These few cases exist because for these events HAZE was reading the digitizings from tape (i.e. off-line), and sometimes a bad patch of tape would cause a read redundancy, which meant that one view would be lost
for an event. The tracks included in this column are those which absolutely needed the missing view for their reconstruction.

**Classification of Tracks**

<table>
<thead>
<tr>
<th></th>
<th>TOTAL NUMBER</th>
<th>OK</th>
<th>S</th>
<th>F</th>
<th>H</th>
<th>W</th>
<th>T</th>
<th>N</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>BEAM</td>
<td>157</td>
<td>106</td>
<td>9</td>
<td>20</td>
<td>2</td>
<td>12</td>
<td>6</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>NON-BEAM</td>
<td>638</td>
<td>557</td>
<td>45</td>
<td>16</td>
<td>10</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>3</td>
</tr>
</tbody>
</table>

**Table 1**

From the figures in Table 1 we arrive at the following conclusions. (All percentages quoted are a percentage of the total number of tracks measured.)

**Success rate**

BEAMS = 67.5%  
NON-BEAMS = 87%

**HAZE errors**

BEAM  
FILTER accounted for 13%  
HH  "  " 1%  
TOO WEAK  "  " 7.5%  
21.5%

NON  
FILTER accounted for 2.5%  
BEAM  
HH  "  " 1.5%  
TOO WEAK  "  " 0.5%  
4.5%

Notes (1) The errors caused by the HH routine are certainly bugs in the program, and it is hoped that they will all be cleared up soon.

(2) Of the 16 non-beam tracks which failed for reasons of FILTER, it is known definitely that 4 of these were caused entirely by a program mistake, which caused FILTER to reject the track,
even though the points were perfectly alright. Therefore to assess the success rate of the FILTER routine as it is at the present time, we should really subtract out these four cases from the failure list. The revised figure for non-beam tracks thus becomes

Reject Rate due to FILTER = 1.9%

This figure corresponds to about 0.7% per view.

Scan errors

Reject Rate for all tracks = 7%
Errors due to the MILADY hardware = 3%
Errors due to the MILADY operator = 3%
Errors due to the MIST program = 1%

Note: The main error made by the scan table operators is mixing up two tracks on one of the views: for example tracks 4 and 5 will be measured correctly on views 1 and 2, but on the third view, track 5 is measured as track 4 and vice-versa.

b) Analysis of events

In Table 2 a breakdown of the events is made similar to that which we made for the tracks in Table 1. The first column in Table 2 gives the total number of events and the second column gives the number of events which were successfully reconstructed. An event was considered satisfactory if each of the individual tracks belonging to the event was well reconstructed, and also the vertex was reconstructed without large errors. In the version of THRESH used for the HPD, the vertex is reconstructed by extrapolating back along the tracks belonging to the vertex and finding the best "point of intersection" of the tracks. The third column in Table 2 contains the number of events which were successful except for the beam track; i.e. all tracks except the beam track were well reconstructed, and the vertex had small errors. Now in the particular experiment which we are measuring on the HPD, the properties of the beam tracks are known.
extremely well, and hence it is not necessary that the beam be well measured, for the event to produce a useful kinematic result. In the GRIND program, the parameters of the beam track may be input into the program as data and these values will overwrite the values found by THRESH. Therefore one may expect the events given in column 3 to be as useful for the kinematic analysis as those in column 2. An event is considered to have failed if one of the non-beam tracks failed in THRESH, and these events are allocated to one of the remaining columns in Table 2. The meaning of these columns is the same as for Table 1, with the exception that all the HAZE failures are now collected together into one column headed HAZE.

Classification of Events

<table>
<thead>
<tr>
<th>TOTAL NUMBER</th>
<th>OK</th>
<th>OK without beam</th>
<th>S</th>
<th>HAZE</th>
<th>T</th>
<th>N</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>157</td>
<td>85</td>
<td>33</td>
<td>16</td>
<td>16</td>
<td>2</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

Table 2

The numbers in Table 2 provide us with the following success and failure rates.

Success rate (with the beam included) = 54%
Success rate (with the beam not counted) = 75%
Events rejected because of HAZE errors = 10%
Events rejected because of SCAN errors = 10%

2. Comparison between IEP and HPD results

A sample of 68 four-prong events were measured on both the IEP and the HPD.
a) Badly measured tracks

In GRIND various tests are made on the results coming from THRESH and error flags are set if the geometry results are poor. A generally accepted criteria is that if the GRIND error number for a track is greater or equal to 100 then the track is considered to be "badly measured" and the event will probably be sent back for remeasuring. Without giving any details (these can be found in the GRIND Manual\textsuperscript{1}), this means roughly that a "badly measured" track is one which either did not converge in THRESH or else the error on the radius was too large. A comparison between the failure rates for non-beam tracks yield the following figures.

Total number of tracks considered = 272
Percentage badly measured on IEP = 11%
Percentage badly measured on HPD = 9%

Only 4\% of the tracks were badly measured on both the IEP and the HPD.

b) Errors from THRESH

THRESH makes a least squares helix fit for each track and it is useful to compare the size of the errors found for the track parameters. These parameters are the radius $\rho$, the dip angle $\lambda$ and the azimuthal angle $\phi$. We have normalized the THRESH errors by dividing them by the errors used in GRIND. The GRIND errors are true external errors. They indicate what the expected measurement errors are and they are quite independent of any fitting process. The formulae used to compute these external errors can be found in the GRIND manual\textsuperscript{1}.
Comparison of THRESH errors

<table>
<thead>
<tr>
<th></th>
<th>HPD</th>
<th>IEP</th>
<th>IEP/HPD</th>
</tr>
</thead>
<tbody>
<tr>
<td>MEAN of $\Delta(1/\rho)_T/\Delta(1/\rho)_G$</td>
<td>0.43</td>
<td>0.56</td>
<td>0.77</td>
</tr>
<tr>
<td>MEAN of $\Delta\lambda_T/\Delta\lambda_G$</td>
<td>0.90</td>
<td>1.14</td>
<td>0.79</td>
</tr>
<tr>
<td>MEAN of $\Delta\phi_T/\Delta\phi_G$</td>
<td>0.51</td>
<td>0.70</td>
<td>0.73</td>
</tr>
<tr>
<td>MEAN of $\Delta P/P$</td>
<td>0.85%</td>
<td>1.04%</td>
<td>0.82</td>
</tr>
</tbody>
</table>

Table 3

The suffix T means THRESH and the suffix G means GRIND.

In Table 3 the mean values of the normalised THRESH errors are given for the three geometrical parameters for both the IEP and the HPD measurements. Only non-beam tracks, which were not "badly measured" were used to compute these means. The fourth row of Table 3 contains a similar assessment of the average measurement error found on the momentum $P$; i.e. the mean of $\Delta(1/\rho)_T/\frac{1}{\rho} \approx \frac{\Delta P}{P}$ is computed for both the IEP and HPD measurements. The figures in column 3 are the ratio of the figures in the first two columns and they give a positive indication that the HPD measurements are indeed more accurate than those of the IEP.
c) Successful events and values of $\chi^2$

We define a "good event" as an event in which all non-beam tracks are well measured.

Total number of events measured = 68
Number of "good events" from IEP measurements = 48 = 70%
Number of "good events" from HPD measurements = 51 = 75%

The number of these good events which were common to both the IEP and the HPD was 40. The kinematic hypotheses tried in GRIND for this experiment were either one-constraint fits (with one neutral particle) or four-constraint fits (with no neutral). Considering only the 40 common "good events", it was found that every four-constraint fit found with the IEP measurements corresponded to a four-constraint fit with the HPD, and the same was also true for one-constraint fits and the cases which produced NO Fits, except for just one event which gave a one-constraint fit for the IEP and a NO Fit for the HPD.

It was easy to confirm that the four-constraint fits were identical in every case, but because of many ambiguities in the cases of one-constraint fits, together with a shortage of time, we have not yet confirmed that the same is true for all the one-constraint fits.

In Fig. 2 a comparison is made of the $\chi^2$ values found for all the four-constraint fit events measured in both IEP and HPD. There are 14 events altogether, but in fact they do not all belong to our sample of 68 events. Some other results did exist for four-constraint fit events and to make the sample for Fig. 2 as large as possible we have included all available events. For each event a point is plotted on the graph in Fig. 2, the x co-ordinate being the $\chi^2$ value found with the HPD measurements and the y co-ordinate being the $\chi^2$ value found with the IEP measurements. The average value of $\chi^2$ is 4.5 for the IEP measurements and 3.7 for the HPD measurements, but the scatter diagram in Fig. 2 is much more informative than these averages.
3. Repeated Measurements

In order to determine the relative consistencies of the IEP and HPD measurements, repeated measurements of the same event were made on both the IEP and the HPD. The sample size was six events and each event was measured six times, making 36 events in all. (A new scan tape was made each time for the HPD events.) Unfortunately not all the events were reconstructed successfully and the final tally of good events was rather less than the number we had at the beginning.

One of the events measured on the HPD, consistently on all measurements, had wrong points on two of its tracks due to the HH routine, and it was finally decided to reject this event altogether from the analysis. The remaining 30 events were not entirely free of operator errors, and for this reason two events were lost from both the IEP and the HPD samples. After obtaining the GRIND results for the 28 good events left, three tracks from both the IEP and the HPD events were found to fail the "well measured" criteria and these tracks have also been excluded from our analysis. Though this reduction in our sample size obviously reduces the reliability of any conclusions we might make, the sample is still large enough to provide useful results. Nevertheless it is agreed that a bigger sample would be more desirable and it is intended to repeat the investigations in the near future on a larger scale.

a) Comparison of external errors

Our first investigation was to try to assess the external measurement errors on the geometrical parameters $\frac{1}{\rho}$, $\lambda$ and $\phi$. In section 2 we made a comparison of the internal errors given by THRESH. For both the IEP and HPD events the following computations were made.

1) For each track we compute the mean of $\frac{1}{\rho}$, $\lambda$ and $\phi$ from the results obtained from the six different measurements.
2) Now using these means we compute the values of \( \frac{1}{\bar{\rho}} - \frac{1}{\rho_1} \) for each track. \( \Delta(\frac{1}{\rho_1}) \) is the GRIND external error and we use it here again as a normalising factor.

3) The distribution of \( \left( \frac{1}{\bar{\rho}} - \frac{1}{\rho_1} \right) / \Delta(\frac{1}{\rho_3}) \) for all tracks of all events is considered as a whole and the standard deviation is calculated.

4) Similar distributions are made for \( \frac{\bar{\lambda} - \lambda_1}{\Delta \lambda_1} \) and \( \frac{\bar{\phi} - \phi_1}{\Delta \phi_1} \).

The distribution of \( \left( \frac{1}{\bar{\rho}} - \frac{1}{\rho_1} \right) / \Delta(\frac{1}{\rho_3}) \) for both the IEP and the HPD events is plotted in histogram form in Fig. 3. In Table 4, the standard deviations of all the distributions for the HPD and the IEP are given in columns 1 and 2 respectively. Since the errors have been normalised with respect to the GRIND error, one would expect the standard deviation for the IEP events to be about 1. In fact it has been known for some time that the GRIND errors for \( \frac{1}{\rho} \) and \( \phi \) are slightly overestimated and the error for \( \lambda \) is underestimated, and this probably accounts for the deviations from unity in column 2. The measurement errors for the HPD events are seen to be clearly smaller than those for IEP. In column 3 the ratio of the standard deviations \( \sigma_{\text{HPD}} / \sigma_{\text{IEP}} \) is given.

**Comparison of external errors**

<table>
<thead>
<tr>
<th></th>
<th>HPD</th>
<th>IEP</th>
<th>( \sigma_{\text{HPD}} / \sigma_{\text{IEP}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard deviation of ( \frac{1}{\rho} )</td>
<td>0.68</td>
<td>0.99</td>
<td>0.69</td>
</tr>
<tr>
<td>Standard deviation of ( \lambda )</td>
<td>0.92</td>
<td>1.15</td>
<td>0.80</td>
</tr>
<tr>
<td>Standard deviation of ( \phi )</td>
<td>0.64</td>
<td>0.87</td>
<td>0.74</td>
</tr>
</tbody>
</table>

Table 4
b) Comparison of $\chi^2$ for four-constraint fits

Of the five events we used for the repeated measurements, three of them were four-constraint fit events. For each of these events, the histogram of the values of $\chi^2$ found with each of the measurements is plotted in Fig. 4. These histograms show that the HPD results are extremely consistent.

Measurement speeds

Finally I will give some brief details on the measurement times.

a) Scan table

The average rate of measuring is 6 events/hour (i.e. 10 minutes per event). This time includes:

1) the scanning of the pictures for suitable events;
2) skipping over frames without events;
3) the measuring of the events.

The time actually spent measuring an event is about 5 minutes.

b) HPD

During an actual hour's production run, a rate of 54 events/hour was achieved. The average time for completely measuring an event is about 22 seconds. For the events we are measuring, an abnormal scan is needed in about 40% of the cases. One reason for this rather slow measurement rate is that the HPD Mark I, which we have at CERN, was designed to work with the IBM 709 computer, and we are not able to take advantage of the increased computing speed, which the 7090 now gives us.

Reference

1. GRIND MANUAL, Section C, EXTER, CERN 1962.
Figure captions

Fig. 1 Analysis sequence of HPD events.
Fig. 2 Comparison of $\chi^2$.
Fig. 3 Distribution of $\left( \frac{1}{\bar{p}} - \frac{1}{p} \right)/\Delta(\frac{1}{\bar{p}})$.
Fig. 4 $\chi^2$ distributions for repeated measurements.
Fig. 1

Fig. 2
DISCUSSION

SOOP: Comparing the values of $\frac{A}{\rho}$ for HPD and IEP results is not very conclusive if multiple scattering is not taken into account particularly for low energy tracks. You should compare the same points measured repeatedly on HPD and IEP.

HOWIE: Multiple scattering is not taken into account but the tracks are mostly of high energy.

STRAND: Do you have a check on fiducial separation in MIST?

POWELL: Not at present. The program is being rewritten to include additional checks.

BURREN: When were tracks counted as successful?

HOWIE: They had to be measured correctly in all 3 views to be considered successful.

STRAND: At Brookhaven we compare HPD and IEP measurements by suppressing one track in the fit and looking at the missing mass. Longitudinal and transverse momentum balances are also compared.
THE POST FILTER ROUTINE OF CERN HAZE

H. HANERFELD
CERN, Geneva *

J. SEYERLEIN
Max-Planck-Institut für Physik und Astrophysik, München

(presented by J. Seyerlein)

The overall purpose of this routine is to increase the accuracy of tracks reconstructed by THRESH from HAZE output and also to enable THRESH to calculate those tracks for which FILTER found only one or two average points. The latter case is mainly associated with short tracks.

For all tracks the first two average points found by FILTER, if they exist, are replaced by new average points. The replacement is made for the two reasons: 1) FILTER uses a rather large bin width for the first two slices. These wide bins may contain noise, thus the average points may be inaccurate. 2) Because of new tracks opening in the first two slices it may be, that the first or the second point found by FILTER lies on a neighbouring track. For all tracks the program attempts to find the first point of the track, that means the first digitizing which seems to belong to the track. For all tracks on which FILTER found less than seven average points, the routine attempts to increase the total number of average points.

*)

On leave from Lawrence Radiation Laboratory, Berkeley
The working medium of the routine is a number of digitizings which have been kept by GATE under control of FILTER. Until FILTER found at least two average points, all digitizings in the Milady-Road are stored in an intermediate buffer. So the routine finds in this buffer at least the digitizings in the first two FILTER-slices of each track.

The routine is divided in five sections, which are called PHASE 1, PHASE 2, TWO POINT ROUTINE, ONE POINT ROUTINE and EXIT ROUTINE. The course of action is determined by the number of average points found by FILTER.

PHASE 1

In PHASE 1 the program tries to find three new average points by refiltering the digitizings which have been kept by GATE using a smaller road than FILTER. It is important that the road is smaller and that the road is based on a circle through three average points found by FILTER. The routine tries to avoid using the first two average points found by FILTER. However, when the number of average points is four, only the first average point is ignored and when the number of average points is three, the program is forced to use these three average points.

Refiltering the digitizings, the program does not use a fixed number of scan lines per slice but computes the slice length based on the number of digitizings in the Milady-Road. The road is divided in \((20_8)\) histogram bins. Looking for a pulse in the histogram the program uses background constants which are based on the number of digitizings which are used to build up the histogram.

When the program does not succeed in finding three average points, provided the number of average points found by filter is greater than four, it re-enters PHASE 1 using the newly found average points and as many of the old points as necessary for determining a circle. Upon repeated failure in PHASE 1 the program transfers directly to the exit without having altered the average points found by FILTER.
PHASE 2

The three new average points found by PHASE 1 are used to compute a new cyclic road to refilter all the digitizings in the intermediate buffer. The road width is half that of PHASE 1. PHASE 2 tries to provide THRESH with at least eight average points.

ONE POINT AND TWO POINT ROUTINES

The two point routine and the one point routine try to find a second and/or third average point to enter PHASE 1.

In the two point routine all digitizings between the two average points are refiltered with a straight road.

In the one point routine the program computes a straight line through the one average point and through each digitizing on both sides of the average point. For each straight line the program counts the number of digitizings within a road of ± 6 least counts from this line. Then the program takes on both sides of the average point that line with the largest number of digitizings falling in its road and computes an average point using all digitizings in those roads on each side. With these two average points and the one point found by filter the program enters PHASE 1.

EXIT ROUTINE

The purpose of the exit routine is the reordering of the old average points by FILTER and the new average points from the present program. A further purpose is, that the EXIT ROUTINE looks for the first digitizing which seems to belong to the track. It chooses the first digitizing in the intermediate storage which falls either in any of the bins from which the last average point in PHASE 2 has been computed or falls in the two bins in the neighbourhood of that region.
DISCUSSION

ACCENSI: Is there any check for agreement between the curvature found on the track or the curvature based on the rough digitizings, and the curvature given by the points of this routine?

SEYERLEIN: No.
THE FORTRAN DIAGNOSTIC ROUTINE USED WITH CERN HAZE

M.L. IUVISETTO
CERN, Geneva

The GATE routine of CERN HAZE divides the picture into slices (picture slices), each one containing a constant number of scan lines. When a track is encountered a track slice is initiated at the beginning of this track. For each slice GATE finds all the digitizings in the road and makes histograms using these points. To set up the histograms, the road is divided into strips parallel to the road edge (bins). For each bin, the following quantities are stored:

1. The total number of points in the bin = $E_1$

2. $EAX = \frac{\sum X}{N}$, the sum of the X co-ordinates of the points in the bin relative to the X value at the beginning of the slice.

3. $EAW = \frac{\sum W}{N}$, the sum of the W co-ordinate increments with respect to the road edge of the points in the bin.

Then the FILTER processes these histograms predicting the road for the next slice and finding the average points, beginning in the coarse mode.

FILTER is described in another paper. The main characteristics which concern the Diagnostic routine are as follows. As soon as FILTER finds two average points, the mode is changed from coarse to medium and if another point is found, the mode is changed to fine. If in one slice a pulse is found, but its area is lower than a fixed background, then on the next slice previous and present histograms are associated (double slice mode)

*)

On leave from the Centro Nazionale Analisi Fotogrammi, Bologna

9758/ga
to find average points for the present and previous slices. If there are
two pulses in the histogram, FILTER opens a new road for the extra track
(called a subsidiary one) and follows both tracks to decide which is the
good one, on the basis of the Milady circle. At the beginning of each track,
all the digitizings inside the road are stored to find the very first point
(MWFEPT routine), and the program continues to store digitizings until the
first two average points have been found. These digitizings are stored in
a 1000 word buffer labelled SL28.

The track information resulting from the GATE and FILTER operations
is stored in a block called BUFB (about 60 words × 16 tracks). Each time
a slice is filtered FILTER calls MONIT, a routine which writes onto tape
the contents of BUFB for that track. This tape is the input to the FORTRAN
Diagnostic routine which reorders the data to give an ordered output for
checking FILTER operation. The tape consists of a sequence of variable
length records, each type of record being named and distinguished from the
others by means of the contents of its first word, which must be one of the
following octal numbers:

11000, 12000, 13000, 13400, 14000

The sequence of records for one event in one view is the following:

11000 Record (10 words)

This record contains the information about the event in general,
i.e. Experiment No., Roll No., Frame No., View, Subexperiment No., Event No.,
etc.

12000 Record (9 words per track – maximum 20 tracks)

This record contains the general information about the tracks
in that event, such as:
total number of tracks, Milady co-ordinates, circle parameters from the
Milady co-ordinates, etc.
A sequence of 13000 Records (100 words each)

Each 13000 Record contains a BUFB plus the 21 histograms. These records are written on tape in the order in which tracks are filtered.

12400 Record (4 words)

This is a dummy record to indicate the end of the sequence of 13000 records.

At this point another 12000 record or a 14000 record may follow. In the first case it means that the event either has been analysed in two scans (one normal and one abnormal or two normal ones in the case of overflow) or that some operation fault has occurred and that the HAZE program has been restarted. In the second case the event is finished and the 14000 record contains the SL2G block of digitizations from the beginning of the tracks.

The diagnostic program needs, besides the described data tape, some control cards which provide a means to select events from it. There are the following possibilities:

1. process the complete tape;
2. process single events;
3. process events by groups, (i.e. from Frame m-View X to Frame n-View Y) with the only assumption that frame numbers in the same view are in increasing order.

The diagnostic routine reads all the information for one view of one event into core and then processes the 13000 records. It has been proved that the events fit well in core and only rare cases give rise to core overflow. For these special cases a restart procedure is provided and the event is analysed in two steps.

An example of the output from the diagnostic routine is shown in Fig. 1.

Each page has the heading FILTER DIAGNOSTIC, followed by the number of tracks in this scan and the type of scan. Then comes the Milady information transformed into the HPD 1.6 μ least count system.
Vertex number and Track number.

XYBB = co-ordinate at the beginning of the track.
XYBB = co-ordinate at the end of the track.

LAMB, LAMB2, FORCB = circle parameters given by:

\[
\text{LAMB} = \frac{X_2 - a}{4c} ; \quad \text{LAMB2} = \frac{X_2 - b}{4c} ; \quad \text{FORCB} = 4c
\]

where \((a, b)\) = circle centre; \(c\) = circle radius; \((X_2, Y_2)\) = co-ordinates of the second Milady point.

XYPQB and WPQB are the \((X_2', Y_2')\) co-ordinates.

Then follows the specification whether the track is a principal or a subsidiary one. If the track is not a principal one, the previous information is missing and there is the number for this subsidiary (subsidiary No 1, No 2, etc.). After this, the information from BUFB follows, and this is:

SL = slice number (octal)
X32TB = X value at the beginning of the slice
P/C = information on the track-following extrapolation (P for parabola, C for circle)
TANGMB = tangent computed from the Milady points (mean value)
TANCUB = new predicted average value calculated by FILTER and corrected on the basis of the average points found
TANB = tangent at the beginning of the slice extrapolated backwards from the middle
DDTANB = increment of the tangent per scan line, considered constant over one slice
XOB, YOB = X and Y value at mid-slice on the left hand edge of the Milady road
WPRED = predicted value for the left road edge, Y value
WRR = usually contains the value for the right road edge, (but in this particular run it contains other information)
RDW = Road width
ST = status, indicating if in coarse (ST = 4), medium (ST = 3) or fine (ST = 2) mode
AVX, AVY = average point from present slice
AVX1 } = average point from previous slice
AVY1
ERWDB = error word, which is different from zero if one of these error conditions is met: 1) too many gaps for what track,
2) points found too far from Milady circle
SLAGPB = slice and gap counter

HISTOGRAMS

The histograms are plotted using the bin number versus the slice number, the number of bins per histogram depending on the mode and being computed as:

\[ \text{NBINS} = \text{RDW shifted right by ST places} \]

The pulse used to compute the average points is given by the numbers in brackets.

At the end of an event in one view there is the output related to the SL2G block. The digitizings belonging to each track are printed in the following order: vertex and track number, scan line number, Y co-ordinate, all Y co-ordinates for that scan line belonging to that track, and so on till the next track.

Looking now at the output shown in Fig. 1, we notice that AVX for slice number 1 and 2 is negative, as flag for the HEBPT routine. In slice 1 and 2 two average points are found, so the road width is reduced from 400\_l.c. to 240\_l.c. and the mode is changed from coarse to medium. Another point is found in slice 3, then FILTER goes to fine mode for slice 4 but in this slice the pulse is too wide for fine mode and the program goes to medium mode for slice 5. No pulse is found and we have to go back to coarse mode again for slice 6. In this slice there is a pulse, but its
area is lower than the threshold, and we have to keep in coarse and use
the double slice procedure. Associating the histograms for both slice 5
and 7, we find a big pulse and it is possible to compute two average points,
so for slice 7 we have two different values, one for (AVX, AVY), and the
other for (AVX1, AVY1). We keep in coarse mode (to find another point) be-
fore going to medium and fine. In slice 10 there is a pulse but still the
area is too small and we go to double slice mode. We find two points in
slice 11 and one point in slice 12, so we go to medium mode for slice 13
and, as we still find a point, we pass to fine for slice 14. No point is
found and we go back to medium in slice 15. One point is found, so we go to
fine again for slice 16, and, as we find a point, we keep in fine mode.

From the histogram we may also see if the track is centred or not
in the road. We see that already at the beginning the points are out of
centre (bad Milady measurements), in fact the pulse should be in bin No. 10,
instead of bin No. 12, and WRR should be 200, not 245. We also see that
the centring improves as we go on processing the track and, at the end, the
track is in fact only 1 l.c. out of centre: WRR = 117 instead of 120.

At present this program has been used only for HAZE off-line runs
because it is very likely that the processing will not keep up with the
input of digitizing from the HPD on-line if many records have to be writ-
ten on tape, as well. This could be avoided by packing the data into larger
records.

The time required to process one event in one view is approximately
20 seconds.
Figure caption

Fig. 1 Example of a filter diagnostic.
### Normal Scan 4 Tracks

**Vertex 1**

```plaintext
XYBB = 66141  LAM1B = -0.0956  FORCB = 102202  XYPB = 153550  WPQ = 27275
XYEB = 114462  LAM2B = 1.9977
```

### Principal Track

<table>
<thead>
<tr>
<th>SL</th>
<th>X32TB</th>
<th>P/C</th>
<th>TANGMB</th>
<th>TANCDB</th>
<th>TANB</th>
<th>DDTANB</th>
<th>XDB</th>
<th>YDB</th>
<th>WPRD</th>
<th>WRR</th>
<th>RDW</th>
<th>ST</th>
<th>AVX</th>
<th>AVY</th>
<th>AVX1</th>
<th>AVY1</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>66121</td>
<td>0.324555</td>
<td>0.324555</td>
<td>-0.340459</td>
<td>-0.996E-03</td>
<td>66760</td>
<td>24666</td>
<td>0</td>
<td>0</td>
<td>400</td>
<td>04</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>67601</td>
<td>0.292786</td>
<td>0.292786</td>
<td>-0.308720</td>
<td>-0.996E-03</td>
<td>70424</td>
<td>25256</td>
<td>0</td>
<td>0</td>
<td>245</td>
<td>04</td>
<td>0</td>
<td>466753</td>
<td>25136</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>71272</td>
<td>0.260212</td>
<td>0.258230</td>
<td>-0.274494</td>
<td>-0.996E-03</td>
<td>72125</td>
<td>25624</td>
<td>0</td>
<td>237</td>
<td>400</td>
<td>04</td>
<td>0</td>
<td>470420</td>
<td>25520</td>
<td>0</td>
<td>466753</td>
</tr>
<tr>
<td>3</td>
<td>72371</td>
<td>0.229287</td>
<td>0.230510</td>
<td>-0.264441</td>
<td>-0.996E-03</td>
<td>73563</td>
<td>26126</td>
<td>0</td>
<td>126</td>
<td>240</td>
<td>03</td>
<td>0</td>
<td>72107</td>
<td>26060</td>
<td>0</td>
<td>25520</td>
</tr>
<tr>
<td>4</td>
<td>74403</td>
<td>0.198621</td>
<td>0.197282</td>
<td>-0.212965</td>
<td>-0.952E-03</td>
<td>75222</td>
<td>26401</td>
<td>0</td>
<td>111</td>
<td>240</td>
<td>02</td>
<td>0</td>
<td>73544</td>
<td>26365</td>
<td>72107</td>
<td>26060</td>
</tr>
<tr>
<td>5</td>
<td>76042</td>
<td>0.167965</td>
<td>0.167068</td>
<td>-0.182036</td>
<td>-0.937E-03</td>
<td>76665</td>
<td>26624</td>
<td>0</td>
<td>111</td>
<td>240</td>
<td>03</td>
<td>0</td>
<td>73544</td>
<td>26365</td>
<td>0</td>
<td>26060</td>
</tr>
<tr>
<td>6</td>
<td>77527</td>
<td>0.136725</td>
<td>0.136608</td>
<td>-0.151594</td>
<td>-0.937E-03</td>
<td>100353</td>
<td>27022</td>
<td>0</td>
<td>111</td>
<td>240</td>
<td>04</td>
<td>0</td>
<td>0</td>
<td>73544</td>
<td>26365</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>101222</td>
<td>0.105180</td>
<td>0.112559</td>
<td>-0.127582</td>
<td>-0.910E-03</td>
<td>102057</td>
<td>27167</td>
<td>0</td>
<td>137</td>
<td>240</td>
<td>04</td>
<td>0</td>
<td>500454</td>
<td>27306</td>
<td>477007</td>
<td>27111</td>
</tr>
<tr>
<td>8</td>
<td>102662</td>
<td>0.075438</td>
<td>0.083441</td>
<td>-0.097991</td>
<td>-0.910E-03</td>
<td>103506</td>
<td>27276</td>
<td>0</td>
<td>137</td>
<td>240</td>
<td>04</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>11</td>
<td>104343</td>
<td>0.044759</td>
<td>0.053442</td>
<td>-0.068473</td>
<td>-0.912E-03</td>
<td>109170</td>
<td>27357</td>
<td>0</td>
<td>112</td>
<td>240</td>
<td>04</td>
<td>0</td>
<td>503512</td>
<td>27556</td>
<td>502014</td>
<td>27435</td>
</tr>
<tr>
<td>12</td>
<td>106020</td>
<td>0.014219</td>
<td>0.020314</td>
<td>-0.035644</td>
<td>-0.930E-03</td>
<td>106650</td>
<td>27407</td>
<td>0</td>
<td>113</td>
<td>240</td>
<td>04</td>
<td>0</td>
<td>505452</td>
<td>27652</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>13</td>
<td>107470</td>
<td>0.015929</td>
<td>-0.016363</td>
<td>-0.000192</td>
<td>-0.930E-03</td>
<td>110286</td>
<td>27557</td>
<td>0</td>
<td>120</td>
<td>240</td>
<td>04</td>
<td>0</td>
<td>0</td>
<td>105452</td>
<td>27652</td>
<td>0</td>
</tr>
<tr>
<td>14</td>
<td>111146</td>
<td>-0.046283</td>
<td>-0.048502</td>
<td>-0.032044</td>
<td>-0.103E-02</td>
<td>111771</td>
<td>27356</td>
<td>0</td>
<td>120</td>
<td>240</td>
<td>02</td>
<td>0</td>
<td>0</td>
<td>106542</td>
<td>27652</td>
<td>0</td>
</tr>
<tr>
<td>15</td>
<td>112604</td>
<td>-0.076120</td>
<td>-0.079271</td>
<td>-0.061762</td>
<td>-0.106E-02</td>
<td>113425</td>
<td>27276</td>
<td>0</td>
<td>132</td>
<td>240</td>
<td>03</td>
<td>0</td>
<td>111710</td>
<td>27654</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>16</td>
<td>114247</td>
<td>-0.013170</td>
<td>-0.109126</td>
<td>-0.091598</td>
<td>-0.106E-02</td>
<td>115070</td>
<td>27165</td>
<td>0</td>
<td>117</td>
<td>240</td>
<td>02</td>
<td>0</td>
<td>113557</td>
<td>27563</td>
<td>111710</td>
<td>27654</td>
</tr>
<tr>
<td>17</td>
<td>115711</td>
<td>-0.106317</td>
<td>-0.109126</td>
<td>-0.126179</td>
<td>-0.106E-02</td>
<td>115070</td>
<td>27165</td>
<td>0</td>
<td>117</td>
<td>240</td>
<td>02</td>
<td>0</td>
<td>0</td>
<td>113557</td>
<td>27563</td>
<td>0</td>
</tr>
</tbody>
</table>
```

**ERWCB = 000000000000  SLAGPB = 100001200017**

### Histogram

```
47 46 45 44 43 42 41 40 37 36 35 34 33 32 31 30 27 26 25 24 23 22 21 20 17 16 15 14 13 12 11 10 7 6 5 4 3 2 1
```

```
<table>
<thead>
<tr>
<th>1</th>
<th>3</th>
<th>3</th>
<th>(21)</th>
<th>2</th>
<th>3</th>
<th>1</th>
<th>4</th>
<th>4</th>
<th>2</th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
<td>(6)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>10</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>11</td>
<td>3</td>
<td>(14)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>12</td>
<td>3</td>
<td>(12)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>13</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>14</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>15</td>
<td>(4)</td>
<td>7</td>
<td>4</td>
<td>(12)</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>16</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>17</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>
```
PRELIMINARY REPORT ON AN "INTEGRAL"
FILTER TECHNIQUE FOR THE HPD.

N. WEBRE, D. BURD, L. GARDER, D. TYCKO, B. ZASLAVSKY
Columbia University, New York

(presented by N. Webre)

We have been interested in adopting a new filter method for our
HAZE program since the rejection rates of Filter II at Brookhaven and LRL
seemed to be too high for a production program. The most promising method
was the track-following technique\(^1\) now being developed at CERN. This, how-
ever, seemed to us to have several inherent difficulties:

a) difficulty of initialization in confused regions;

b) susceptibility to local phenomena (though very much less than
   Filter II), especially in the early stages of following;

c) non-recoverability: the inability to pick up the correct track
   if the wrong one is being followed.

We thought that a filter method which used all the information in
the road might not have these problems. Furthermore, if the contents of
the roads could be condensed into some small number of words, they could
all be saved and more powerful techniques than can be used in real time might
be applied at the end of the scan.

The problem of recognizing a circle with a computer appears to be
much more difficult than that of recognizing a straight line. Straight
line recognition has been demonstrated by Brookhaven's and LRL's experience
with the H-10 fiducial finding routine\(^1\). This suggested the possibility
of reducing the circular case to the linear case, which, in turn, suggested
conformal mappings. The following is the description of a method which
uses such a mapping.
CONDENSING THE ROAD DATA

This is done in two steps; first, picking the digitizations in the road (Gating), and second, reducing the digitizations to "master points" and slopes.

Gating is done by the CERN method\(^1\), \(^2\), but with the road edge always being calculated by using the rough digitized road, and the road width being held constant. For each point \((X', Y')\) found in the road, the following quantities are calculated:

\[1, X', Y', (X')^2, (X' \cdot Y')\]

where \(X'\) is the stage co-ordinate and \(Y'\) is the spot co-ordinate. For simplicity, we shall restrict our discussion to one scan mode.

Using \(\Delta Y'\), the distance of the point from the edge of the road, a histogram bin number is calculated for the point. For each bin the quantities

\[\Sigma 1, \Sigma X', \Sigma Y', \Sigma (X')^2, \Sigma (X' \cdot Y')\]

are calculated where the sums are taken over all points in the bin.

Ionization is being omitted for this discussion, but is easily included in the program.

After the scan has progressed \(N\) scan lines, the histogram of \(E1\) is searched to find pulses. A floating threshold is used

\[T = k_1 \cdot q + k_2\]

where \(k_1\) is a constant \(\geq 1\), \(q\) is the average number of points per bin, and \(k_2\) is an integer \(\geq 2\). A slice is an \(N\) scan line section of a road (see Fig. 1). The quantities \((E1)_p, (\Sigma X')_p, (\Sigma Y')_p, (\Sigma (X')^2)_p, (\Sigma X'^{\ast} \cdot Y')_p\) where \(p\) is the pulse number, are computed for each pulse. The points of each pulse are used to determine the slope of a least squares straight line.
\[ b' = \frac{(x_1) \cdot (x') \cdot (x') - (x')^2}{(x_1) \cdot (x') - (x')^2} \]

where \( Y = a' + b'X \). In addition, the co-ordinates of the "master" point are taken as the average co-ordinates of the points in the pulse \( X_m = \frac{(x')}{(x_1)} \)

\[ Y'_m = \frac{(y')}{(x_1)} \]

The above quantities are transformed to the square orthogonal system:

\[ b' = \frac{b'}{K} \left( 1 - b' \cos \phi \right) \]

\[ X'' = K \cdot X'_m + Y'_m \cos \phi \]

\[ Y'' = Y'_m \]

where \( K = \frac{X \text{ least count}}{Y \text{ least count}} \)

and \( \phi \) - the angle between the stage and scan line directions (see Fig. 2).

For the entire road we collect the set of master points and tangents \( (x'_i, y'_i, b'_i) \) \( i = 1, 2, ..., n \)

where \( n \) is the number of master points in the road. It is interesting to note that the information in the roads is thus reduced to small line segments similar to those produced by a flying line scanner such as PEPR 5, 6).

THE TRANSFORMATION AND FILTERING

Given this set of points and tangents, the object is to recognize all the tracks present in the road. Suppose we now consider the set of
master points \((X''_1, Y''_1)\) to be in the complex plane such that

\[
Z''_1 = X''_1 + iY''_1
\]

Make the transformation

\[
Z_i = C(Z''_1 + D)
\]

where \(C\) and \(D\) are complex numbers. This is equivalent to a translation, and rotation (see Fig. 3).

\((X''_B, Y''_B)\) and \((X''_E, Y''_E)\) are the beginning and end rough digitizings of the road. \(C\) and \(D\) are computed such that the \(X\) axis falls on the chord between \((X''_B, Y''_B)\) and \((X''_E, Y''_E)\) and the \(Y\) axis passes through the midpoint of the chord.

If \(b''_1 = \tan \theta''_1\), the transformed tangent

\[
b_1 = \tan \theta_1 = \tan (\theta''_1 - \alpha) = \frac{b''_1 - \tan \alpha}{1 + b''_1 \tan \alpha}
\]

where

\(\theta''\) is the angle in the \(Z''\) plane

\(\theta\) is the angle in the \(Z\) plane

\(\alpha\) is the angle of rotation between the \(Z''\) and \(Z\) planes

Transforming all the master points and tangents \((X''_1, Y''_1, b''_1)\), we now have the set

\[(X_i, Y_i, b_i) \quad i = 1, 2, \ldots, n\]

Suppose we pick a master point in the \(Z\) plane near the \(Y\) axis, say \(Z_\ell\), and translate the origin to it. (In the final program, we intend to make the \(Y\) axis pass through the middle rough digitizing and pick master points which are close to it. The scanners will be instructed to put the middle rough digitizing in a clear region of the track, thereby insuring a master point on the track at that place in the road. In this way, the scanner will be able to directly guide the Filter process.) Now consider
the transformation

\[ W_i = \frac{1}{(Z_i - z_\ell)} = \frac{1}{(X_i - x_\ell)} - i(Y_i - y_\ell) = U_i + iV_i \]

This mapping has the property that circles in the \( z \) plane will map into circles in the \( W \) plane except those which go through the origin in the \( Z \) plane. They will transform into straight lines. The equation of a circle through \( z_\ell \), the translated origin in the \( Z \) plane, is

\[(X - x_\ell)^2 + (Y - y_\ell)^2 + \lambda_1(X - x_\ell) + \lambda_2(Y - y_\ell) = 0\]

Applying the above transformation with

\[ U = \frac{X - x_\ell}{(X - x_\ell)^2 + (Y - y_\ell)^2} \quad V = \frac{Y - y_\ell}{(X - x_\ell)^2 + (Y - y_\ell)^2} \]

we get

\[ 1 + \lambda_1 U + \lambda_2 V = 0 \]

With this mapping, the tangents transform as follows:

\[ \left( \frac{dV}{dU} \right)_{i \neq \ell} = \frac{(X_{i \ell}^2 - Y_{i \ell}^2)b_{i} - 2X_{i \ell}Y_{i \ell}b_{i}}{-(X_{i \ell}^2 - Y_{i \ell}^2) - 2X_{i \ell}Y_{i \ell}b_{i}} \]

\[ \left( \frac{dV}{dU} \right)_{\ell \ell} = \frac{b_{i}}{2X_{i \ell}Y_{i \ell} + b_{i}} \]

where

\[ X_{i \ell} = (X_i - X_\ell), \quad Y_{i \ell} = (Y_i - Y_\ell) \]
If the point $Z'_c$ were on a track, all those master points of the same track in the Z plane should be collinear in the W plane. Points of any other track not passing through the origin should appear as points on a circle. Since a straight line has the property that its slope is constant, histogramming the tangents $\left(\frac{d\bar{V}}{dU}\right)_i$ should produce a peak at the angle of the straight line in the W plane. If we pick the points in this tangent histogram peak, they should in general be the points on a track. If the number of points which are in the pulse but not on the track is small, then they can be discarded by doing a circular least squares fit and throwing out the points with highest deviations as is done by the spatial reconstruction programs $^3$.

ERROR PROPAGATION

The points and tangents $(X, Y, b)$ have measurement errors $(\varepsilon_X, \varepsilon_Y, \varepsilon_b)$. It is interesting to investigate the behaviour of these errors under the transform from the Z plane to the W plane.

$$f (X, Y, b) = \frac{d\bar{V}}{dU} = \frac{(X^2 - Y^2)b - 2XY}{-(X^2 - Y^2) - 2XYb}$$

Assuming that $X$, $Y$, and $b$ are independently measured quantities

$$df = \frac{\partial f}{\partial X} \cdot dX + \frac{\partial f}{\partial Y} \cdot dY + \frac{\partial f}{\partial b} \cdot db$$

$$= \frac{\partial f}{\partial X} \cdot \varepsilon_X + \frac{\partial f}{\partial Y} \cdot \varepsilon_Y + \frac{\partial f}{\partial b} \cdot \varepsilon_b$$

$$A (X, Y, b) = \frac{\partial f}{\partial X} = \frac{-2Y^3 + 2b^2Y^3 + 2X^2Yb^2 + 2X^2Y}{(X^2 - Y^2 + 2XYb)^2}$$

$$B (X, Y, b) = \frac{\partial f}{\partial Y} = \frac{2X^3 + 2X^2b^2 + 2XYb^2 + 2XY^2}{(X^2 - Y^2 + 2XYb)^2}$$

9758/ga
\( C(x, y, b) = \frac{\partial \Phi}{\partial b} = -\frac{x^4 + y^3 + 2xy^2}{(x^2 - y^2 + 2xyb)^2} \)

Some plots of these quantities for circles of various radii are shown in Figs. 4, 5 and 6.

These expressions are approximately correct for noise points, points on other circles, etc. It can be used to investigate the behaviour of the entire road.

**TEST AND RESULTS**

The data we are using are the roads rejected from the Brookhaven HAZE program which uses Filter II. The film is from the BNL 20 inches bubble chamber exposed to a 3.5 GeV/c \( \pi^- \) beam. When a road is rejected, they write the entire contents of the road onto a tape. Due to a bug in their system some unrejected roads were also written. We read these tapes, make master points and tangents by dividing the road into 32 scan line sections, apply the transformations to the tangent for several choices of origins, and compute the tangent histogram in the \( W \) plane for each origin. The plots shown in Figs. 7 and 8 are the contents of the road, the \( Z \) space, and the tangent histogram.

We then did by hand what we eventually expect the computer to do—pick the points on the track by examining the tangent histogram. We applied the following rules:

1. find the highest pulse using a fixed window to search the tangent histogram, i.e. sliding a window of a fixed number of bins along the tangent histogram and looking for the strongest pulse. For these tests, we used a window width of 6 bins, where 1 bin = 0.0025.

2. If there are two or more pulses of the same height, choose the one closest to \( \left( \frac{dV}{dU} \right)_\ell \), the tangent at the origin.
3. Accept the pulse if \( P > \frac{S}{2} \)

where \( P \) is the number of points in the pulse, and \( S \) is the total number of slices in the road. Assume no track is present if no pulse meets this requirement.

4. If two master points belonging to the same slice fall into the same \( \frac{dV}{dU} \) pulse reject both points.

We ran 75 randomly selected roads from rejected Brookhaven events. Of these, 19 were beam tracks. We obtained the following results:

- number of roads - 75
- number not attempted - 5
- number attempted - 70
- number successful - 68
- percentage - 97%

Where "number not attempted" includes empty roads, roads with tracks less than two slices, and roads rejected because of tape check errors, and where "number successful" are roads in which at least the correct track was found with no more than one incorrect point. For some cases, mostly parallel beam tracks, other tracks besides the correct one were found.

Both cases that were rejected were due to the same error—no master point could be found near the centre, so there was no point on which to place the origin. We think this can be remedied by the scanner guidance of the origin which was mentioned before.

We also investigated the efficiency of the method, i.e. how many of the points on the track would the program pick, and how many incorrect points would it pick.
Number of points on tracks - 342
number of points picked - 797
efficiency - 95%
number of incorrect points - 6

No track had more than 1 incorrect point. Of the 6, 2 were questionable, 1 was on parallel beam tracks 60μ apart, 2 were on parallel beam tracks 30-40μ apart.

In no case was a track constructed out of random noise master points.

We ran with the following program parameters:

road width = 400μ
one bin width in ΔY' = 16μ
slice length (N) = 32 scan lines
threshold parameter (k1) = 2
threshold parameter (k2) = 3

In Figs. 7 - 21, we have shown some cases which demonstrate some of the properties of the system. For each case, the following displays are printed:

a) the actual bubble chamber photograph;
b) the contents of the road. The ordinate and the abscissa are ΔY' (ΔW) and the scan line number, respectively;
c) the road condensed to master points in the Z plane. The units are millimetres on the film. The circled points are the ones picked by applying the previously mentioned rules 1 through 4 to the tangent histogram;
d) the histogram of the tangents in the W plane;
e) c and d are repeated if another track in the road is shown.

Case 1 (see Figs. 7 - 8)

This road contains a strong track with a bit of noise (a typical case).
Case 2 (see Figs. 9 and 10)

This shows a three slice track and the resulting tangent histogram (a short track).

Case 3 (see Figs. 11 and 12)

This illustrates the excellent performance of the fixed number of scan lines slice and the floating threshold in discriminating against noise. The lone point at -0.050 in the tangent histogram is the point on the track in the last slice. This often occurs when the last slice contains less than the maximum number of scan lines.

Case 4 (see Figs. 13, 14 and 15)

This shows the effect of putting the origin first on and then off the track. The misplaced origin is approximately 90μ from the track.

Case 5 (see Figs. 16 and 17)

This shows tracks less than 50μ apart almost parallel for a large portion of the road and then crossing at an angle of less than 1°. The vertex is on the right.

Case 6 (see Figs. 18, 19, 20 and 21)

This demonstrates the effect of putting the origin on each of three almost parallel beam tracks.

CONCLUSIONS AND FURTHER TESTS

The first striking thing about the data was that the master points and tangents made with the fixed length slice and the floating threshold gave an excellent representation of what was happening in the road. It also gave a very effective pre-filtering of the data. This is significant if one thinks of gating and making master points on a small computer, and processing the roads later on another machine.
Since the master points give a true representation of the road, there would be no need to remeasure the road if it turned out to be difficult to filter.

The tests seemed to indicate that if a master point to use as the origin could be found on a track, then the track could be recognized. The need for scanner guidance of the origin and a more determined search for a master point seems to be necessary. These tests included some very difficult cases, e.g. case 5 was the most difficult case of a low angle crossing, and not shown is a difficult example of parallel beam tracks 30 - 40μ apart and of length 23 mm on film.

The fact that in no case was more than one incorrect point picked as part of the track indicates that a least squares circular fit to the data and the rejecting of the point of highest deviation would have a high probability of discarding the bad point. This will be checked during our next phase of testing.

The efficiency of the method in picking the available track points in general eliminates the need for "gathering" the missed points. In a few cases which were the main contributions to the few percent of the master points lost, a technique to gather the residual good points has been proposed. It is to fit the points in the $\frac{dV}{dU}$ pulse with a least squares circle, throw away the worst measurement if it has greater than some maximum deviation, refit, and then gather all points which are within some small tolerance of the fitted curve. This will be tested also.

Of course, some development of logic is needed too, but this does not appear to be very difficult or complicated. The program will be constructed to find all the tracks in any road. Given all of the possible tracks, the choice of the correct one could be made by inquiring which remains in the road for the entire length, or by a simple vertex calculation. We have not mentioned before that the end points of the
tracks will be calculated by extrapolating the fitted curve to the end of the road.

The next phase of testing will contain the following procedures:

1. programming the search for the pulse in the $\frac{dV}{du}$ histogram;
2. least squares fitting of circles to the points in the pulse and a study of the deviations;
3. refining the CRT plots to facilitate the study of these results.

If these tests turn out favorably, the next step will be a production version of the program written in FORTRAN and employing the logic discussed above.

There are three major questions still remaining:

a) is the real time portion of the program fast enough to keep ahead of the HPD?
b) will the system handle non-circular tracks?
c) short tracks?

To check (a) some test will be run with the BNL Mark I machine as soon as the program has been debugged using tape as input.

We have passed some non-circular tracks measured on projection microscopes through the transformation. Their characteristic is that the centre of the $\frac{dV}{du}$ histogram moves away from zero and the width of the pulse is greater. It seems that there may be some correlation between the pulse position and its width, but the data is inconclusive because we have so few cases (b) and it is not HPD data. We will study these further as soon as more data is available. We might note that even if the transform method fails, the set $(X, Y, b)$ still contains the information necessary for some sort of track-following techniques.

Short tracks (less than 3 slices) cannot be handled. We will try reducing the slice length to 16 scan lines or even less (only for these
tracks) to see if they can still be done by the method.

PROGRAMMING AND TIME ESTIMATES

The Gating program must, of course, run in real time. Since the Filter program must have all the information in the road before it can operate, it must wait until the end of the HPD scan. It seems that the logical place for it to run is during the stage retrace of the HPD.

We have made time estimates based on the mechanics of the operations. We think that less than 0.15 sec will be needed to do a very large, noisy road of about 80 master points, with smaller roads being done in a proportionally smaller time. We have not actually measured the real running time of the prototype program yet.

The storage need for saving the road information (master points, tangents, and ionization) is probably not more than 2,000 words for the maximum case of six roads per scan.

Taking Filter out of real time seems to have several advantages:

1. very difficult roads do not compound the logic of the real time program. All gating has to do is pick more points in the road and calculate new master points. This increases the time, of course, but in a very predictable and linear fashion. The unpredictable timing of complex logic is now out of real time where we are not hurt so badly if it expands for certain cases. The worst that could happen is that the HPD is momentarily delayed;

2. several different methods could be tried for difficult cases;

3. it may be feasible to use a small computer to reduce the HPD data to the \((X, Y, b)\) sets, write these onto tape, and use the large computers for filtering, track reconstruction, etc. On the average, the amount of data that would be written onto tape would be slightly more than three times that which is now output after filtering.
4. It makes it possible to write the Filter program in FORTRAN so that it is easily manipulated and changed even during production. It is also more easily understandable to other users.

ACKNOWLEDGEMENTS

We wish to thank Dr. P. Hough, Dr. R. Strand, Mr. W. Thompson, and the members of the BNL - HPD Project for their helpful discussions and for supplying the data for these tests. We wish also to thank Dr. E. Fowler of Duke University for the use of his film. Our thanks also to the National Aeronautics and Space Administration for the use of their Stromberg-Carlson 4020 plotter and to the staff of the Columbia University Computing Center.

FURTHER RESULTS

Since the first writing of this paper the program had undergone some changes and additions, and there has been more experimenting with the data. The program now automatically searches to find a pulse in the tangent distribution (see pp. 159-160 Nos 1, 2, 3 and 4). The tangents are no longer histogrammed, but instead are ordered in increasing order. Moving the "sliding window" along now consists of testing each tangent in succession to see how many other tangents are within a specified distance (i.e., the window width) from it. This makes possible a more precise determination of the pulse position.

When a track is recognized, the points assumed to be on it are fitted to a least squares circle (see "LAL Memo No. 220", Frank Solmitz). The deviation from the circle of every point in the road is computed and the standard deviation of the fitted points is calculated.

We ran 60 tracks in this manner with the following results:

1. for the tracks for which the method picked the apparently correct points, 77% of the fits had $\sigma < 2.0\mu$, 93% had $\sigma < 3.0\mu$, and 100% had $\sigma < 4.0\mu$.  

9758/ga
2. The three cases which included 1 incorrect point all had 
   \( \sigma > 10.0 \mu \)

**PROGRAM ORGANIZATION**

Columbia University's HAZE program was developed from Brookhaven's HAZE, which was in turn developed from the original HAZE written at LRL. The major organizational difference is that the latter two use the FOG-CLOUDY-FAIR System, while the Columbia HAZE operates under the FORTRAN II Monitor System. The program is fully subroutinized, i.e., relocatable FAP and FORTRAN subroutines. This gives us the advantages of having to reassemble only the subroutines being modified and of having an automatic recording of the changes. We also gain the ability to assemble and run at the same time. We have used FORTRAN to rewrite several of the subroutines (e.g., the scan tape reading routine, the transformation from scan table to HPD co-ordinates, Filter, etc.). In general, they are routines which do not operate during the measurement period.

All I/O with the exception of the HPD control and digitizations, is done by FORTRAN and is not buffered.
References


Figure captions

Fig. 1  A schematic representation of the jth slice.

Fig. 2  The definition of $X'$, $X''$, $Y'$ and $\varnothing$.

Fig. 3  The translation and rotation achieved by the transformation.

Figs. 4, 5, 6  Plots showing the behaviour of $A$, $B$ and $C$ for different values of $R$.

Figs. 7, 8  The digitisations obtained from the primary of the interaction shown in 7a are plotted in 7b. In Fig. 8a is shown the result of the transformation and 8b the tangent histogram for the same track.

Figs. 9, 10  The same information as for the previous case is given for the short track indicated in Fig. 9a.

Figs. 11, 12  The same information is plotted for the incident track of the interaction shown in Fig. 11a.

Figs. 13, 14, 15  These figures show the effect of displacing the origin by about 90 microns.

Figs. 16, 17  Show a case of two tracks about 50 microns apart for much of their length.

Figs. 18, 19, 20  These demonstrate the effect of putting the origin on each of three closely spaced beam tracks.
$j^{th}$ SLICE

Fig. 1

SPOT MOTION

Fig. 2

$Z^\prime$ PLANE

Fig. 3
FIG. 4
WEBRE'S 1/Z TRANSFORMATION

HISTOGRAM OF TANGENTS IN W PLANE

FIG. 8
WEBRE'S 1/Z TRANSFORMATION

HISTOGRAM OF TANGENTS
IN W PLANE

FIG. 10
WEBRE'S 1/Z TRANSFORMATION

HISTOGRAM OF TANGENTS IN W PLANE

FIG. 12
WEBER'S 1/Z TRANSFORMATION

HISTOGRAM OF TANGENTS IN W PLANE

FIG. 14
WEBRE'S 1/Z TRANSFORMATION

HISTOGRAM OF TANGENTS
IN W PLANE

FIG. 15
Fig. 16a

Fig. 16b
Webre's $1/2$ Transformation

Histogram of Tangents in W Plane

Fig. 17
WEBRE'S 1/Z TRANSFORMATION

HISTOGRAM OF TANGENTS
IN W PLANE

FIG. 19
WEBRE'S 1/Z TRANSFORMATION

HISTOGRAM OF TANGENTS
IN W PLANE
WEBRE'S 1/Z TRANSFORMATION

HISTOGRAM OF TANGENTS
IN W PLANE

FIG. 21
DISCUSSION

KRISCHER: Is your transformation in fact equivalent to the following? Using the centre co-ordinates \((a, b)\) of the circle passed through the rough digitized points, the tangent value \(y'_1\) for a master point \((x'_i, y'_i)\) could be calculated from \(y'_i = \frac{x'_i - a}{y'_i - b}\). These values \(y'_i\) would then be compared with the values \(y'_i\) found by the program, and master points which gave differences deviating too far from a mean difference could be discarded.

TYCKO: I think the difference is that we do not assume the track to be a particular circle. We only use the fact that a circle will transform into a straight line. Normally the tracks are quite circular. Non-circular tracks can be handled by treating them in sections which are more nearly circular.

KRISCHER: As your method seems to depend on finding pulses in the histograms, why don't you use the tangent values which you have already gathered to update the tangent and road for the next slice?

WEBRE: Before you can update the road you must decide where the track is and I do not wish to make that decision until I have accumulated all the information along the track.

TURNILL: Don't you tend to get bad average points in confused regions because you are using a wide road?

WEBRE: There is a width criteria applied to the pulses which should eliminate crossing regions. Looking at the distribution of the master points around the least squares circle the points appear to be very good.
SEYBOTH: Have you thought of applying your method to a system of minimum guidance?

WEBRE: No, but we think it is a good solution of the present problem. However, we think our next step will be pattern recognition.

MARR: Webre's method might prove useful in solving the problem of locating kinks in track segments that are reconstructed by our pattern recognition program.

MOORHEAD: The difficulty appears not to be that of finding master points from pulses in coarse histograms but that of finding master tangents.

WEBRE: Occasionally tangents turn out to be bad but we do not depend on any one tangent, all we depend upon is that in general they are good.
HPD PROGRAMMING AT THE RUTHERFORD LABORATORY

J.W. BURREN, J.W. GARDNER, M.J. MITCHELL
Rutherford Laboratory, Chilton

M.C. TURNILL
Imperial College, London

(presented by J.W. Burren)

Over the past year the main effort of the Rutherford Laboratory group has been taken up in reprogramming the track following methods for the ORION computer. We have kept basically to the methods described in (1) but have taken the opportunity of improving many of the programming methods. The program is written in FORTRAN apart from the Gating subroutine and some of the "service" routines which are in machine code. First to serve as a reminder we give a brief outline of the method.

1. TRACK FOLLOWING METHOD

We use the method of "gating" a road for \( N \) scan-lines and then "filtering". During the gating phase the points in the road are histogrammed relative to the road-edge. The road-edge at \( X \) is given by

\[
Y_R = Y_o + (X - X_0) (\tan \theta + \frac{n}{2} \Delta \tan \theta)
\]  

(1)

where \((X_0, Y_0)\) is the position of the road-edge at the beginning of the slice, \(\tan \theta\) is the slope at the beginning of the slice, \(\Delta \tan \theta\) is the change in slope per scan-line and \(n\) is the number of the scan-lines in the slice. We use a histogram bin width of either 16, 8 or 4 least counts and form for each bin 21 the number of hits in the bin, \(S_X\) the sum of \(X\) co-ordinates, and \(S_Y\) the sum of \(Y\) co-ordinates relative to the road edge.

In addition we have a facility to store the co-ordinates of all the points in the road for the slice.
After N scan-lines the histogram is examined by the "FILTER" part of the program. The filter subroutine has two functions, firstly to find an average ("master") point for the current slice and secondly to determine values for the gating parameters $I_0$, $\tan \Theta$, $\Delta \tan \Theta$ and the bin size for the next slice. We scan the hit-count histogram for bumps above a threshold and have criteria of width and area to define a "good" bump. For a good bump we calculate an average point using

$$X = \frac{\Sigma X}{N} + X_c(Y)$$

$$Y = \frac{\Sigma Y}{N} + Y_R\left(\frac{XX}{N}\right)$$

(2)

where $N$ is the number of points in the bump, $\Sigma X$ and $\Sigma Y$ are summed for all the bins in the bump, $Y_R(X)$ is the road-edge at $X$ and $X_c$ is the $X$ correction for the pitch of the scan-lines.

By comparing the position of this point with its position predicted from the previous slice, we can correct the $\tan \Theta$ and $\Delta \tan \Theta$ for the slice. The corrections are

$$\Delta \tan \Theta = W_1 \frac{(Y_n - Y_p)}{(X_n - X_{n-1})}$$

$$\Delta \Delta \tan \Theta = W_2 \frac{(Y_n - Y_p)}{(X_n - X_{n-1})} \frac{dX}{(X_n - X_{n-1})}$$

(3)

Where $W_1$ and $W_2$ are weighting factors which we make dependent on the number of points already found on the track, $(X_n, Y_n)$ are the co-ordinates of the average point found for the slice and $X_{n-1}$ is the $X$ co-ordinate of the previous point, and $dX$ is the scan-line separation. On the assumption that the curve is a circle we can extrapolate $\tan \Theta$ and $\Delta \tan \Theta$ to the next slice and can also calculate a predicted $Y$ position for the point in the next slice.
We adjust the road edge so that the predicted Y position is at the centre of the road. Finally we use the width of the bump to determine the bin-size for the next slice.

We deal with the case of multiple bumps in the road by setting up a new road for each bump and then following separately with each road. This is of course necessary if we wish to use narrow bin sizes for histogramming.

2. PROGRAMMING IMPROVEMENTS

In order to have variable slice lengths for tracks, we use the concept of a "picture" slice suggested last year by Moorhead. That is we divide the picture into slices of 48 scan-lines and we associate a control word with each scan-line. The control word gives a list of all the roads that are to be filtered (or started) on the particular scan-line. The word is divided into eight 6-bit sections (ORION is a 48 bit machine) and thus we can filter up to 8 tracks on a scan-line. In general there will be no filtering required and so the very quick test of the control word against zero is all that is required every scan-line. When a track is filtered we count on the number of scan-lines of the next slice modulo 48 and insert the road number into the appropriate control word (we shift the word up by 6 bits and insert into the bottom 6 bits).

We also use the picture slice to control the starting of roads and the ordering of roads in Y. Every 24 scan-lines we look to see if any tracks are due to start in the next 24 scan-lines and if so we insert their numbers, together with a flag to indicate that they are starting into the appropriate control word. The current roads are reordered so that they are in order of increasing Y co-ordinate of their road edges at the middle of the next 24 scan-lines.

To save time in the gating subroutine we have separate blocks of identical coding for each of the current roads. Our gating routine consists of 11 blocks; one block which searches for the X co-ordinate
of the scan-line and controls the filtering and starting of tracks, 6 identical blocks each of which gates one track-road, 2 identical blocks which gate fiducials and 2 blocks which can store all the points in the fiducial road for subsequent histogramming and filtering (these latter blocks are used if the program is following 4 or more tracks when the fiducial is due to be measured). Every scan-line the gating routine loops through the X search block and as many of the other blocks as are currently being used. At the start of the picture the X search block loops on itself, then as tracks are started (via the control words) road and fiducial blocks are brought into the loop. The threading of the loop from block to block can easily be reordered if for instance tracks cross one another. As tracks end, roads and blocks are taken out of the loop until finally the X search block is again looping on itself. On the ORION computer this method of organising the gating gives a considerable increase in speed.

3. RESULTS OBTAINED WITH THE PROGRAM

We have tested the program on a sample of 20 events measured on the CERN HPD together with rough digitisings of the events from the CERN "Milady". The events are a mixture of $K^-$ at 3.5 GeV/c and $K^+$ at 3.0 GeV/c from a run in the 80 cm hydrogen chamber. The events are mainly 2 prong with one V but we have other topologies. The total number of tracks is about 100, sample which is of course too small for reliable statistics on rejection rates. However, we have found that our difficulties have been confined to two sources. Firstly highly curved tracks with changing radius (tracks with radii of 10 - 15 cm and less in the chamber) are sometimes difficult to follow. We hope by some adjustment of our constants and by bringing the minimum slice length allowed down from 20 to say 16 scan-lines to be able to deal with most of these tracks.

Our second difficulty comes from the problem of continuity of tracks. If we find a point for a slice we define a continuous region of the road for the next slice in which we expect to find the next point on the track. We widen this continuous region for the following slice if we fail to find a point.
We have to deal with the case when we find two or more good points in the continuous region and our first solution was to set up a new road for each new point and to follow these roads independently. This simple solution works surprisingly well but leaves two difficulties. First the case when we have two beam tracks in the road which are merging or crossing. In this case we set up too many roads. Secondly we have found that we sometimes set up a second road in the first slice because another track is leaving the vertex at a small angle to the track of interest. Thereafter the road for the correct track follows quite normally, but the incorrect road fails to find a point in the second and sometimes third slices and then because the continuous region has widened out it finds a point on the correct track and follows this. The net result is that we have two sets of points for the track which are identical apart from their first points. We have solved this difficulty with the following scheme. When we find multiple-bumps in the continuous region or when the road has split in the previous slice, we examine each bump to see if it is nearer to some other road which is being followed for the same track. If it is nearer then we ignore it for the road in question. This procedure has eliminated the continuity difficulties at least for our current sample of events.

Our most encouraging results have been obtained using our Geometry program on the output for these events. The Geometry program does a helix-fit by projecting a helix in space (with slowing down correction) back through the optical system onto the three film planes and then minimising the normal distances of the points from the curves. For each track we print an R.M.S. error of these distances summed over the three views. For hand measurements this error has an average value of about 4 microns for the very best measurements but in general averages about 6 - 7 microns. From the HFD, again within our limited sample, we find an error of about 2 microns. However what is perhaps more significant is that the spread of errors is much smaller for HFD than for hand measurements, nearly all measurement errors being less than 6 microns. We have found that larger errors can usually be traced to mistakes in the filtering. In fact, the geometry program seems likely to prove a very useful diagnostic tool.
4. TIDYING OPERATIONS

At the end of the scan we carry out some tidying operations to deal with short tracks (tracks covered by less than 50 scan-lines) and to decide, in the case of more than one track passing through the rough digitising, which track to take.

a) Short Tracks

We use a method of point following from scan-line to scan-line similar to the methods used in the pattern recognition programs. The method appears to be quite successful but our statistics are very limited.

b) Elimination of Tracks

We try to decide between ambiguous tracks by a close examination of the vertex area. For the co-ordinates of a vertex, we have the rough digitising \((x_R, y_R)\) which has an error \((\Delta x_R, \Delta y_R)\), the co-ordinates of the last point (digitising) on each track, and finally we find the point which is "nearest" to all the tracks at the vertex together with errors. By nearest we mean the point for which the sum of the squares of the normal distances to the tracks at the vertex is as minimum. By assessing the consistency of this information we try to eliminate the unwanted track. We also use the information to calculate the vertex co-ordinates. First results indicate that the method can be made quite powerful.

CONCLUSION

Results obtained with the program from a small sample of data have been encouraging. We hope in the very near future to run the program with very much larger samples of data.
References


DISCUSSION

MARR: Have you considered scanning upstream instead of downstream?

BURREREN: Yes, we shall probably do this on our own experiments.

WEBRE: Is the polishing routine used because the first point was calculated using the crude histogram?

BURREREN: We need a good first point and tangent at the beginning of the track for a reliable vertex determination. The first point may either be missed completely because of confusion at the vertex or may be bad because we use a coarse histogram at the beginning. This is why we use a polishing routine, however it might be unnecessary if you scan upstream.
RESULTS OBTAINED WITH AN OFF-LINE HPD SYSTEM

M. BLOCH, M. SCHIFF
Laboratoire de Physique Nucléaire, Collège de France, Paris

C. de la VAISSIERE
Institut du Radium, Paris

(presented by M. Schiff)

Introduction
The design of our system has been described elsewhere¹). The main part of our software has recently been completed and is now being tested with some photographs digitized with the CERN HPD. After a brief reminder of the principles of the system, we shall describe the off-line program. We shall then discuss our present results.

I. THE OFF-LINE SYSTEM

1) Outline

The first operation is that of digitized scanning: roads are digitized in the usual fashion. The second operation is the precision measurement with the HPD digitizer; under the control of a small computer (CDC 160A), a bubble chamber picture is transferred onto a magnetic tape; the computer time available during this on-line operation is used to perform a very crude gating of the digitizations²). Finally, the magnetic tapes are treated by a CDC 3600, to produce an output acceptable to a geometry program.

The data reduction is thus decoupled from the on-line operation of the HPD. Each of these two operations can then proceed at its natural speed. We expect our Mark I HPD to digitize about 50 triads per hour; on the other hand, our present processing speed with a CDC 3600 corresponds to a rate of about 500 triads per hour.

9758/p/mn
2) **Advantages of being off-line**

The practical and economic advantages are obvious: we will be able to use our digitizer in a fashion and at hours suited to a man machine system because the computer used to operate the HPD is relatively cheap.

Our method of data reduction itself is favourably influenced. Since bottlenecks were no longer to be feared we could afford to write our program entirely in FORTRAN; also, the program can use as much time as necessary at a difficult point. A second and even more important consideration is the following: the program has an easy access to all the data. This is useful, both in the search for fiducials and in the search for tracks (see below).

3) **Possible drawbacks of a very small computer**

We do not anticipate any difficulty in the control and transfer functions of the CDC 160A. The only difficulty could arise from a need to perform checks in real time. Three kinds of check might be needed: check the co-ordinates given by the digitizer; check that the pictures have been properly counted; check that the frame to be measured is properly centered on the stage. All these checks can be performed by an on-line program designed to decode the BCD frame numbers that appear on each picture. Periodic systematic checks will also be needed. There is no difficulty of principle, but the coding is not trivial.

II. **THE OFF-LINE PROGRAM**

The treatment can be divided into three main phases. During the first phase, the data are read from a magnetic tape and simultaneously sorted and stored in appropriate blocks of memory. During the second phase, the exact position of fiducials is found. During the third phase, the track images are reconstructed. The program has been coded in FORTRAN for the CDC 3600.
1) **Input and storing of the data**

Three magnetic tapes (one per view) are mounted on three tape units. For a given event, the 3 views are treated in sequence. The results found for each view are stored and merged in core to meet the format requirements of a geometry program (THRESH).

The first record of each file (picture) contains the information coming from the rough digitizer (roads); the other records contain co-ordinates from the HFD digitizer. These are read into alternate buffers; while one buffer is being filled, the other one is being decoded.

Each word in a buffer contains two digitizings. Each digitizing is examined and stored into the X or into the W block, depending on its nature. In a third block, one stores the address of the last W for each scan line; this allows one to have an easy access to the content of any scan line.

2) **The search for fiducials**

For each fiducial, the search proceeds within an appropriate window. There are 3 modes of operation for the search. The first mode is used when no fiducial has yet been found: the window is large. After one fiducial has been found (any of those available), the position of the others can be predicted with a fair accuracy: the search region for each fiducial, and parameters of the search subroutine are changed accordingly. Finally, when 2 fiducials have been found, the position of the others is in principle exactly known; a third mode is nevertheless used, to make a consistency check between all fiducials.

The 3 modes of search only differ by the numerical value of the parameters involved. For any mode, the computation performs the following functions:

- extract a window from the picture; then for each arm:
  - make a histogram of distance to the assumed position of arm
  - make a least squares linear fit to the points in the peak
  - choose the best fitting points and make a second fit
  - check direction found, value of fit and number of points found.

The position of the fiducial is given by the intersection of the 2 arms.

9758/p/mn
3) The search for tracks:

The content of the program is outlined in flow charts 1, 2, and 3, describing respectively the treatment of an event in one view, the treatment of one track and the treatment of one segment of track.

a) Treatment of one picture (flow chart 1)

Labels for tracks and vertices, as well as the event type, are determined by the program. The information that has to be digitized at scanning time is thus kept to a minimum: co-ordinate pairs for 2 fiducials and for 3 points per track. Only in 2 cases does a scanner have to push a special button: when a track seems to be stopping or, when it is so long and so curved that the road is defined by 6 points instead of the usual 3.

After having determined the topology of the event, the program determines the transformation coefficients that will allow one to express the rough digitizations in the HPD co-ordinate system.

The program then considers each vertex in turn, starting with the primary vertex. The tracks coming out of the vertex are treated and checked, one after the other (see below). If a track is satisfactory, its summary curve nearest to the vertex is saved for future computation of the vertex point. The intersection of "good" tracks is computed by least squares: the sum of squares of the distance of the vertex to the tracks is minimized.

The exact position of the vertex is of interest in itself. Also, it can provide redundancy checks for the tracks; finally, it is used to rescue tracks that have not been located by the normal procedure. A most common case will be that of two beam tracks inside one road. Using an approximate value of the curvature, and the known position of the vertex, one makes an angular search for an element of track.
b) Treatment of one track (flow chart 2)

It can be divided into three phases:

α) a road is computed and used to find a set of points

β) a search is made for track elements within the road; the
elements found are represented by "summary curves"

γ) in case of success: check that the curves found do coincide
with a track, count bubbles and compute summary points.

a) Road and gate

A road is defined by 3 rough points $P_1$, $P_2$ and $P_3$. The 2
extreme points define a "natural" co-ordinate system for the
track: $P_1 P_3$ is on the x axis and the y axis bisects $P_1 P_3$.
The road curve is defined in the natural co-ordinate system as
that parabola which goes through $P_1$, $P_2$ and $P_3$. The co-ordinates
of the gated points are expressed in the natural system. For
each scan line the ordinate of the road curve is subtracted from
the ordinate of the gated points. This removes the average
curvature from the track.

β) Search for track elements within the road (second page of flow
chart 2)

Before we start describing our filtering procedure, let us point
out three of its distinctive features. The first one is that the
recognition of a track element is done by a least squares parabolic
fit, rather than by a histogram. The second is that the size of
the element has a wide range; between 30 scan lines and the full
length of the track. Finally, the accurate location of the track
is usually found through successive approximations; a table is
used to record the information obtained on each portion of the track;
this information is constantly kept up to date and is used to reduce
the width of the search regions. For instance, if an element of
track has been definitely located, it can be used to extrapolate a
path within which the next element should be looked for,
The road is divided into segments. In each segment of road one searches for a track element (the search for an element in a segment is described in section c below). The search for elements proceeds as follows.

For the first trial, the road is considered as a single element. If an element has been located with sufficient accuracy, the program proceeds to part Y. Otherwise, the road is divided into 2 segments for a second trial. If necessary, the road can be divided into 4 or 8 segments for a third or fourth trial.

Once an element of track has been located with adequate accuracy (good fit → element fully satisfactory), it is used to extrapolate a path within the road, for the adjacent segment; at the same time, the smaller elements that it may contain are ipso facto recorded as being fully satisfactory.

If an element has been located, but with insufficient accuracy (mediocre fit → element partially satisfactory), it is used to define a path for the smaller elements that it contains (interpolation).

To treat a segment of road, one attempts successively interpolation, extrapolation or the starting procedure. In that latter case, the width of the path is taken equal to the width of the road.

Y) Checks and summary points

Once all elements are satisfactory, for a given trial, one checks and exploits the results.

Elements of tracks have now been summarized by arcs of parabola. For conventional geometry programs, tracks are defined by measured points; the HPD program must therefore provide summary points. These are obtained by computing the intersections of the track with a set of straight lines perpendicular to it.
The summary curves are also used to obtain a histogram of the distances of all the digitizations of the track to the computed trajectory. This histogram serves to check the summary curves (see next chapter, section C2) and to compute the apparent ionisation: a record is made of hits and misses along each scan line.

c) The search for one element within a segment of road (see flow chart 3)

This search is performed by a subroutine called GATSAMFI (contraction for GATE, SAMPLE and FIT). A segment of road is defined by its first and last scan line. Within the road, a path is defined by a parabola and by a width. If the number of sample points found in the path is too small, control returns to the calling program. This allows one to try successively interpolation and extrapolation: if those do not work, the starting procedure is used, i.e. the path is made as wide as the initial road.

a) First phase: GATE and SAMPLE within a path

The segment is divided into zones (say 10 zones with 4 scan lines each). In each zone, the program searches for a sample point located inside the path. When GATSAMFI is called for the first time, the segment is as long as the track and the path is as wide as the road; in this case, sampling plays an important role. It proceeds according to the following recipe: if a scan line contains a single point within the path, take that point as a sample point, if there are two points, take alternately the top and bottom ones; if there is no point in the path or more than 2, try the next scan line in the zone.

β) Second phase: fitting the sample points

A parabola is fitted to the sample points. A choice is then made of the best points: for instance one can choose the 7 points that lie closest to the computed parabola. A parabola is again fitted to those 7 points.
The value of that second fit and the parameters of the parabola are transmitted to the calling program. The magnitude of the fit is used to decide whether a segment is fully OK, partially OK or not OK.

III. THE STATUS OF THE PROGRAM AND ITS PERFORMANCE

The program described above has been run as a single package on real data with a CDC 3600.

The input part, the search for fiducials, and the logistics to treat tracks are working fairly well; the latter includes the normal filtering procedure, but not a special procedure (tracks reconstructed from the vertex) which is not yet debugged.

The interface with a geometry program is not yet complete. Finally, some parts not mentioned in the description of the program still remain to be written (essentially a provision for the orthogonal scan and a provision for short tracks).

In order to understand the behaviour of the program and study its performance, we have used a display of the content of roads, a display of histograms and a time analysis program. We shall now describe some of the results of this study.

A - Input of HFD data from magnetic tapes

Data supplied by the CERN HFD group were converted to our format to provide an input for the program.

The saving obtained by packing and by pregating the HFD digitizings was studied. Also, detailed measurements were made of the times involved in the various phases of the input. Results obtained with the present version of the program can be summarized as follows.

1) Our format is 3 times denser than the usual one.

2) By pregating the picture, we chop off about 75% of its w's; this leads to an additional gain in input time.
3) The tape handling associated with the operation of our HPD digitizer is expected to be as follows:

- number of pictures on one tape 400
- time to read and decode 3 tapes 16 minutes

4) At present, the input is limited by the decoding of the digitizings (about 180 μsec per word), rather than by the read-in time (100 μsec per word). By making the decoding time match the reading time, we could reduce the overall time by about 30%. This would probably require rewriting the decoding subroutine in machine language.

B - The search for fiducials

1) Scope

In each picture, we search for the 6 internal fiducials of one of the windows. We do not use external fiducials. With an early version of the program (IBM version), we located fiducials on all of the four triads used, with a large redundancy for internal checks.

2) Accuracy

Using our present improved version (CDC version), we made a study of the performances of the various modes of searching for a fiducial. Out of 21 fiducials, 20 were located in the coarse mode and all were located in the final mode.

Figure 1 illustrates the fact that the final mode is more accurate than the coarse mode, at least in the presence of background. In order to know to what accuracy internal fiducials can be located, we computed and plotted histograms like those shown in Fig. 2. One can see that fiducials can be located to within a few microns.

3) Speed

Finally, a time analysis was performed. To accurately locate 6 fiducials, we use about 400 msec:

- 80 msec to locate the first one in the coarse mode
- 40 msec to locate the second one in the second mode
- 260 msec for the 6 fiducials in the final mode.
In short, the search for fiducials is safe and accurate and it is fast enough to be used as it is.

C - The search for tracks

1) Scope of the normal filtering procedure

Two kinds of "difficult" cases were encountered, which do not fall within the scope of the normal procedure. These are discussed below.

- Parallel beam tracks

The case where a road contains several parallel tracks has been specifically provided for by the special procedure that starts from the vertex. This procedure has worked, but it is not completely debugged yet: it was called for in 2 cases; it worked in one and failed in the other.

- S shape within the road

Because of conical projection and/or energy loss, tracks going through a large angle can appear S shaped inside a circular road. There were 2 such tracks in our sample for which the amplitude of the S was about 400 microns (on the film); we now believe that this very large amplitude was due mostly to an error in the computation of the road *). The fact still remains that in its present state, our filtering procedure does poorly on "S shaped" tracks.

- "Normal" tracks

All the other tracks available (23 tracks) were properly handled. The accuracy of the summary points computed by the program is discussed in the next section.

*) Note added in proof. This was indeed found to be the case.
2) **Accuracy of the summary points**

a) **Means used to study it**

For each track, the program computes summary curves, and from these it computes summary points. There are 3 ways in which the results can be judged.

The most direct and fool-proof check on the filtering is that obtained by a visual display of the content of the road. The summary points are plotted, together with all the digitizings of the road; the resolution of the display is 1% of the width of the road. The displays are obtained by off-line printing (see Fig. 3). Although they are invaluable for studying the program, they cannot be used extensively, for practical and economic reasons.

A more compact display can be obtained by plotting the histogram of the distances between all points within the road and their respective summary curves. In fact, only the central part of this histogram is used (see Fig. 4). For the time being, we simply look at these histograms.

The acceptance criteria presently used are based on a parabolic fit to sample points. Criteria based on fits are easy to program; but we have found them to be sometimes deceptive. We now think that it may be worthwhile to complement them by tests performed on histograms.

b) **Results found**

Using the display of summary points, we have found that the program provides an accuracy close to that permitted by the natural width of the track.

In order to be able to judge objectively how modifications in the program or in the parameters affect the results, we have started keeping a record of the fit values and of the width of the histogram, for each track. An example of such a record is summarized below in the form of two histograms (one least count is about 1.6 microns).
The width of the histogram is defined for each track by the strip of the histogram containing just above 50% of the bubbles of the track. The half widths thus defined are shown here.

In 19 tracks the histogram was off centre by less than 1 least count. In 4 tracks it was off centre by 1 to 2 least counts. In no case was it off centre by more than 2 least counts.
3) **Time used to treat tracks**

The results of the time analysis are summarized below, for 5 track events, on one view.

<table>
<thead>
<tr>
<th>Input</th>
<th>Time (msec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>FIDUCIALS</td>
<td>400</td>
</tr>
<tr>
<td>GATE</td>
<td>500</td>
</tr>
<tr>
<td>PARABOLIC FIT</td>
<td>155</td>
</tr>
<tr>
<td>CHOOSE</td>
<td>120</td>
</tr>
<tr>
<td>GATSAMTI</td>
<td>165</td>
</tr>
<tr>
<td>HISTOGRAM, BUBBLE COUNT</td>
<td>240</td>
</tr>
<tr>
<td>VERTEX</td>
<td>20</td>
</tr>
<tr>
<td>All the rest</td>
<td>110</td>
</tr>
</tbody>
</table>

**Total**: 2.2 seconds

Before concluding this chapter on results, we wish to emphasize that all parts of the program are in FORTRAN and that little effort has yet been made to save time. We expect that additions and refinements to the program will be more than compensated for by an effort to speed up a few of the most time consuming subroutines. This effort should be greatly facilitated by the use of our time analysis program.

**IV. CONCLUSIONS**

We have found that magnetic tapes provide a means of entering HPD data into a large computer which is both more practical and more economical than the conventional method.

Our data reduction is done off real time. This has permitted us to use a program that is 100% in FORTRAN. Our program takes advantage of the absence of bottle necks in the time flow and of an easy access to the whole picture.

*) In the absence of pregating, or when circular roads are used, this time becomes 1.0 sec.
2) Our procedure to find fiducials does not rely on external fiducials and can start anywhere in the picture. It has been found to be safe, accurate and sufficiently fast.

3) The procedure to locate tracks within a road works in three steps. First, one attempts to reduce the effective width of the road by making a parabolic fit to sample points, evenly distributed along the track. In the second step, the road is divided into segments inside which one searches for track elements; as soon as one element has been located, the continuity between elements is used to help the search for other elements. Finally, hits and misses are counted in the usual way, using histograms.

This procedure takes care of the bulk of the tracks in an efficient way. Using visual displays, we have found it to be almost as accurate as the jitter of the digitizer will allow.

4) From accurate time measurements performed on the program, we conclude that our data reduction, including the input, will proceed at a rate of about 500 triads per hour.

5) Our program has been run as a single package, on real events, with a CDC 3600. We believe that it is adequate to measure events on bubble chamber photographs which have a good contrast and where the beam tracks do not swamp the secondary tracks. To qualify this statement, however, we must add that several additions and refinements still remain to be made before this goal is achieved.

V. ACKNOWLEDGEMENTS

The input part of the program has been skilfully written by Mr. P. Margerand. We have received considerable help in the writing, the debugging and the running of the program from Mr. Ton That. We are very indebted to the CERN HFPD group, without whose assistance this work would have been impossible.
References


2. P. Leblond, see following paper.
Figure captions

Fig. 1  The plots demonstrate the behaviour of the fiducial finding routine in the coarse mode (above) and in the final mode (below).

Fig. 2  The histograms show the accuracy of the points found by the two modes of the fiducial finding procedure. The first, based on 520 points from 20 different fiducials, corresponds to the coarse mode. The second shows the accepted 320 points after the final mode. The abscissa is in 1.6 micron counts.

Fig. 3  The elimination of the background by extrapolation is demonstrated here where the total width of the plot corresponds to 1000 microns.

Fig. 4  A plot of the distances between all points within the road and the corresponding summary curve. From left to right the track has been treated as a whole, as two segments and as four segments respectively. The bin width corresponds to 32 microns on the film.
Fig. 2

Fig. 3
APPENDIX

Analysis of input into the CDC 3600. Use of the millisecond clock.

a) Pregating and packing

In an early version of the pregating program (IBM version), we found that pregating reduces the number of W's by a factor of about 4. Another gain is achieved by the packing of 2 digitizings into a 48 bit word. Finally we only use one X per scan line, instead of 8 X's and one full grating count.

b) The conversion from CERN format to our format

For the road measurements we used the card output of the CERN MIST program; the HPD digitizings were a copy of tapes obtained with the CERN HPD. The road cards were used to compute a mask for each picture.

The pregating and the packing were achieved by a CDC 160 A program. The tapes produced in this way were used as input to the CDC 3600 program.

c) The time analysis subroutine

A special subroutine was written (and the program modified), to analyse the time flow of the program. Each time the program enters or leaves a subroutine, a call is made to the millisecond clock. A "time account" is kept up to date while the program is being executed. At the end, the time account is decoded and printed. In this way, an accurate picture of the time flow is obtained, with little perturbation to the program under study.

d) Timing of the input

One example of time measurements performed on real pictures is shown below (view 1 of the first event; the number of scan lines was 1300 and the number of W's before and after pregating was 8000 and 2000).
<table>
<thead>
<tr>
<th></th>
<th>Without pregating</th>
<th>With pregating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Read in and decode road points. Start reading of first record of HPD digitizations</td>
<td>80</td>
<td>80</td>
</tr>
<tr>
<td>Read HPD digitizations and 10 interrecord gaps*</td>
<td>70</td>
<td>70</td>
</tr>
<tr>
<td>Decode and store X and W's</td>
<td>860</td>
<td>310</td>
</tr>
<tr>
<td><strong>TOTAL TIME SPENT</strong></td>
<td>1,010</td>
<td>460</td>
</tr>
</tbody>
</table>

*) This is the time not spent in simultaneity; when simultaneity was temporarily suppressed this time was 600 msec, for the unmasked picture.

e) **Tape handling to be expected**

In order to make meaningful statements let us define a fictitious "standard" picture:

- Average number of scan lines: 2000 lines
- Average effective number of tracks: 20 tracks
- Average percentage of hits: 50%

Using the results given above, and assuming that for a "standard" picture 75% of the W's will be removed by pregating, one can predict what the tape handling is going to be. The results are given in the text.
FLOW CHART NO. 1

TREATING ONE EVENT ON ONE VIEW

TOPOLOGY
associate tracks into vertices
count tracks and vertices

ROTATION
PARAMETERS
Scan table → HPD co-ordinates

Loop over vertices

Loop over tracks

TREAT
CHECK
TRACK

OK ?

Store for vertex

last track?

Find VERTEX with good tracks
Rescue bad tracks

Last vertex?

Next picture
FLOW CHART NO. 2  page 1 of 2

TREATING ONE TRACK

Compute ROAD in HPD and in track co-ordinate system

GATE points inside the road Express them in track co-
ordinate system

Try first overall fit to about 20 points -- If successful reduce width of road

Loop over trials

Loop OVER SEGMENTS OF TRACK
See next page

All segments OK for this trial?

CHECK summary curves by histogram
BUBBLE count
Compute SUMMARY POINTS

Last trial?  →  Next track
FLOW CHART NO. 2  page 2 of 2

TREATING SEGMENTS OF ONE TRACK

Loop over segments of track

Is segment already fully OK?

Partially OK? — Is previous segment fully OK?

Define path by INTERPOLATION

Define path by EXTRAPOLATION

path = road

TREAT SEGMENT INSIDE PATH

Is segment now OK?

Store parameters of summary curve for segment and for segments it contains

Last segment?
TREATING ONE SEGMENT: GAT SAM FI

INPUT

- Loop over sampling zones
- Loop over scan lines in one zone
- Position of path
  \[ IY = RAO \times X^2 + RAI \times X + RA2 \]
  - Found sample point?
  - Last line of zone?
  - Last Zone?
- More than J5 points?

RETURN

FIT sample points to a parabola

CHOOSE J5 beat points

FIT them to a parabola

OUTPUT

- Summary curve PA0, PA1, PA2
- Fit value
DISCUSSION

SEYBOTH: How do you get your first points on the track?

SCHIFF: We first try to locate the track using about 20 sample points. The track is divided into zones and within each zone, in a rather random way, a point is chosen and a parabola fitted through these points. Since within the road, the average curvature has been removed, one expects that the fit would be a straight line, but in practice it is usually more like a circle, or even an S shape. In many cases though, the fit will be reasonably good, 20 - 30 µ as I said, and this is how you most often get started. If you cannot get started on the whole track, then you try on half the track, and here we have used Burren's suggestion of starting away from the vertex. Usually it is the vertex region that is messy, so you begin as far away from the vertex as possible.

EDMONDS: Do you consider it to be a practical proposition to store all the digitizings on tape or is the pre-gating in the small computer a necessity?

SCHIFF: The pre-gating can equally well be done in the big computer during the input of the digitizings - clearly this is not as economical as doing it in the small computer, but it is quite feasible. With our format we expect to get between 10,000 and 20,000 words of digitizings per picture.

STRAND: I am not sure I understand your use of the MILADY. Do you have a switch that independently gives you the topology, or is it left to the program to discover the topology?

SCHIFF: We do have a switch on our console, but we do not intend to use it. The main reason for this is that most of the events which we shall be measuring initially will be 4-prong events, and so one might just as well not use this information. Of course, there are some internal checks in the program, such as the number of points on each of the three views must be the same.
STRAND: We have been measuring a mixture of events and we did find it useful to have this piece of redundancy. Do you have provision for entering what we call "comments"?

SCHIFF: In principle yes, though this part is not programmed yet.

STRAND: I have a short observation to make relating to the earlier discussion. It seems to me that you will need to have your pictures taken very carefully, with a lot of contrast, so that you don't have to digitize noise in the picture. I think this will be very important. Assuming this is true and you have 20 tracks per picture, how many events do you expect to get onto a spool of tape, say 4-prong events?

SCHIFF: With pre-gating about 300 - 400 events on 3 magnetic tapes; that is one tape per view.

STRAND: If your 3 tapes are mounted on the 3600 at the same time, then, since you have 64 K of storage at your disposal, you have the interesting possibility of being able to do the geometrical reconstruction at the same time.

SCHIFF: We do load the 3 input tapes together, so that we can do the merging for the geometry program in core. I agree that in principle we could also do the reconstruction then; it is just a matter of more programming.

TYCKO: It may encourage you to know that Columbia is also planning to keep scan table operations down to a minimum. We shall reconstruct topologies from the rough digitizings, the track labels will be simply automatically generated sequential numbers and the frame numbers will be read automatically.
CONTROL OF THE HPD BY A SMALL COMPUTER

Ph. LEBLOND
Collège de France, Paris

(presented by Ph. Leblond)

The Collège de France system operates in two stages\(^1, 2\). the first stage operates essentially as a film digitizer with a magnetic tape output; the second stage is a program for a large computer which uses this magnetic tape output to reconstruct the events.

The first stage incorporates a small computer which controls the measurement and performs a certain amount of coarse gating on the co-ordinates before recording them. This gating involves only the simplest logical operations and reduces the amount of data by a factor of about four.

While this paper deals mainly with the actual HPD control program, it will be necessary to describe beforehand the mechanical and logical structure of the system, and the handling of the preliminary rough digitizing step.

I. DESCRIPTION OF THE SYSTEM

1. Arrangement of the components

The Collège de France system includes the following components (Fig. 1).

a) A Sogenique mechanical flying spot digitizer and its associated electrical equipment.
b) A CDC 160A, computer with 6192, 12 bit words of memory, two input-output channels, two external interrupt lines and various peripheral devices. The two input-output channels are capable of handling data in 12 bit words at a rate of 39 to 50 thousand words per second. One of the two channels is completely buffered. The peripheral equipment includes: four IBM compatible magnetic tape units, a card reader, a card punch, a printer, a type-writer and an X-Y plotter. This system is more elaborate than required to handle the operation of the HPD since the computer is used for other purposes as well.

c) The interface between the computer and the HPD, consisting of two separate connections:

- A 128 word buffer memory, to smooth the irregular output of data from the HPD and adapt it to the input rate of the computer.

- A control interface to operate the HPD according to commands received from the computer in the form of 12 bit words.

It should be noted that the magnetic tape units are connected to the buffered channel and the interface connections are made to the unbuffered channel.

2. Premasurement testing program

Referring to Fig. 2 the processing chain of events is seen to include a scanning and manual rough digitizing or premasurement step, that produces a set of cards for each selected event. These cards are processed by a specialized premasurement program that tests the cards to insure that they record measurable events. Cards containing errors are rejected for remeasurement of the event. The program also transforms the co-ordinates to Cartesian form and uses this premasurement
data to create the mask that will govern the pregating. This mask is a binary representation of a coarse $18 \times 32$ element grid that is applied to the photograph. Those elements which contain a portion of a premeasured track are assigned a value one. The other elements are given a value zero.

The output of the testing program is a magnetic tape which contains the information on the three views of all the events selected from a film roll. This tape serves as input to the HPD control program.

II. THE HPD CONTROL PROGRAM

A) General description

There are two parts to the control program: the first, or initialization phase, advances the film to select the proper event and sets up the measuring conditions. The second or measuring phase sets the film stage in motion and sends the co-ordinates to the computer for recording on magnetic tape.

1. The initialization phase

In the initialization phase, the event number is obtained from the input tape. A command is then sent to the film transport system to advance the proper number of frames. Then the mask, corresponding to the view being measured, is read in and stored in the memory so as to be ready at the processing time. Once the event has been properly positioned, commands are sent to the HPD to select the actual measuring conditions: speed and direction of the measuring stage, fraction of the scan lines to be accepted.

2. The measuring phase

In the measuring phase, each co-ordinate is examined as it arrives from the HPD. The co-ordinate is then accepted or rejected depending on the state of the corresponding element of the mask. Control of the position of the measuring stage is effected by the computer
which stops the transfer of co-ordinates as soon as the end of the photograph has been reached. At that time if there is no perpendicular scan, a stage return is ordered and processing of the next event commences.

B) Control of the Sogenique equipment

Control of the HPD is achieved by the 160A in the following manner: first the computer selects the control interface in one of three modes. The film mode, the measurement mode or the status request. Then in the first two cases the computer sends a command in the form of a 12 bit word.

1. The film mode is used to control the movement of film transport. The 12 bit command consists of 11 bits indicating the number of frames to be moved, and the twelfth bit indicating the direction of movement.

2. The measurement mode controls the actual measuring conditions and the movement of the film stage. Table 1 lists the various commands available to the 160A. Each function can operate independently, which provides the system with a great flexibility of operation. As an example, 011, 111, 110, 100 would be the command for a normal measurement at maximum stage speed, all scan lines accepted; 010, 110, 000, 100 is the command for a stage return.

3. The status request mode. When a status request is sent, the control interface responds with a 12 bit word through the buffer memory. For the moment only the two first bits are used, the first indicating a proper operation of the system, the second indicating an alarm condition.
c) **Transfer of co-ordinates to magnetic tape**

This is done during the second phase by the measurement routine.

1. **Simulation**

As the Sogenique equipment has not yet arrived, the routine has not been run under actual measuring conditions. However a simulation program was written to take the place of the HPD, using CERN HPD tapes as the input. These CERN tapes contain all the measured co-ordinates of several events from a roll of film. They are read in by the simulation program which feeds them to the measurement routine in the correct format.

2. **Processing times**

The simulation program has made it possible to debug the measurement routine and to evaluate the times required to process a co-ordinate. These figures are found to be: 65 μsec to examine, and when necessary reject a co-ordinate, 100 μsec to accept a co-ordinate. If one third of the co-ordinates is accepted the program can process 6600 co-ordinates per second. With the model I HPD (10 msec/scan line) this corresponds to a maximum average of 40 hits per scan line which is amply sufficient for most cases.

3. **Protection against buffer memory overflow**

There is always the possibility of a picture being so dense as to saturate the buffer memory. To prevent loss of information under such conditions, a safety valve has been built into the system. When the capacity of the buffer memory reaches the three quarter mark, an interrupt signal is automatically sent to the computer. This interrupts the normal program sequence and transfers control to another routine which by-passes the preeating section and accepts the co-ordinates at the maximum rate of the computer, that is 16500 co-ordinates per second. Once the buffer memory has been emptied, control reverts to the normal
program.

D) The problem of controlling output quality

Since the 160A does not itself yield the full reconstruction of an event, it is difficult to make an immediate check on the quality of the measurement. This may be done when the large computer program has processed the HPD output. However, this is a real disadvantage since remeasurement becomes then both inconvenient and costly.

It is proposed to minimize this difficulty by the following procedures:

1. Periodic checks with standard photographs and/or test patterns.

2. During the measurement of a photograph, verification of the position of known fiducial marks in the picture, in our case the position of the binary lights on the NIRNS numbering system.

CONCLUSION

Sufficient tests have now been performed to show the feasibility of efficiently controlling an HPD by a small computer. Although this method increases the amount of manual data handling it has the advantage of simplifying the overall HPD system, both for the electronic equipment and the programs involved.

The speed of the 160A limits to about 3000 RPM the maximum speed of rotation of the disk in the Sogenique apparatus. It would be simple to run at a faster rate by replacing the central processing unit by a faster one. There exists on the market a number of small relatively inexpensive computers that would be suitable for such a purpose.
References


2. M. Bloch, M. Schiff and C. de la Vaissière, "Results obtained with an off-line HPD system", preceding paper.
<table>
<thead>
<tr>
<th>BIT POSITION</th>
<th>COMMAND</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>NOT USED</td>
</tr>
<tr>
<td>2</td>
<td>STOP-START</td>
</tr>
<tr>
<td>3</td>
<td>FORWARD-BACKWARD</td>
</tr>
<tr>
<td>4</td>
<td>STAGE SPEED 11 = MAX</td>
</tr>
<tr>
<td>5</td>
<td>00 = MIN</td>
</tr>
<tr>
<td>6</td>
<td>SCAN LINE 00 = 1 LINE IN 8</td>
</tr>
<tr>
<td>7</td>
<td>DENSITY 11 = ALL LINES ACCEPTED</td>
</tr>
<tr>
<td>8</td>
<td>TRANSMIT W</td>
</tr>
<tr>
<td>9</td>
<td>MOVE X STAGE (=0), MOVE Y STAGE (=1)</td>
</tr>
<tr>
<td>10</td>
<td>TRANSMIT X (=0) OR Y (=1) CO-ORDINATES</td>
</tr>
<tr>
<td>11</td>
<td>ENABLE X OR Y PERRANITI REGISTER CLEAR ON STAGE RETURN</td>
</tr>
<tr>
<td>12</td>
<td>NOT USED</td>
</tr>
</tbody>
</table>

**Table 1. COMMANDS TO CONTROL INTERFACE**

- MEASUREMENT MODE
Figure captions

Fig. 1  System layout.

Fig. 2  Processing of an event.
PRELIMINARY STAGE

SCANNING

ROUGH DIGITIZING (PREMEASUREMENT)

PREMEASUREMENT

TESTING PROGRAM
→ MASK CREATION

FIRST STAGE

HPD CONTROL PROGRAM

SECOND STAGE

HAZE

Fig. 2
SOOP: Do you use the x - y plotter for anything in particular?

LEBLOND: We use the Calcomp plotter for testing purposes only since this is a lengthy procedure. When there is a lot of plotting to be done we prefer to use a 1401 program at Saclay, which prints the output on a 1403 printer equipped with a special draughting chain. The resolution is 40 points per inch, and the result is very good.

EDMONDS: How soon will you be in operation after the delivery of your machine from Sogenique?

LEBLOND: The programs have already been written and tested with simulation routines. Consequently they should work once the rest of the equipment is in operating condition. I think, that after a short debugging period of a month or so we should be able to digitize a photograph properly.

STRAND: Do you plan to monitor the number of digitizations with the computer?

LEBLOND: We have not yet considered this possibility. Also, I believe, there will be an automatic gain control included in the electronics.

BUTLER: We use the method of controlling the number of digitizations to set the photomultiplier sensitivity level on our chromosome scanning program. I have tried this method also on some rather "dirty" spark chamber pictures, with what I judged to be acceptable results. I think it should work even better with the HPD.

EDMONDS: What computer would you choose if you were buying one now.
LEBLOND: The decision to buy a CDC 160A was made eighteen months ago when the computer market offered a much smaller choice than today. There are available now a number of small computers suitable for the job. In any case, when choosing a computer, I would insist on the following properties: low cost CPU, memory cycle about 2 μsec, word length of at least 18 bits, IBM compatible tapes, ease of adaption of "home made" peripheral equipment i.e. nonsynchronous input-output.
The procedure for bubble density evaluation and the application of ionization information in identification of bubble chamber events, which are being developed at CERN, follow closely those proposed by Strand\textsuperscript{1}, 2).

At present a scan line separation $s = 66 \mu$ is used, and for the pictures from the Saclay $81$ cm HBC with which we are working, an effective scan width\textsuperscript{3}) was found to be $a = 40 \mu$. Therefore, it is a case without scan overlap, where any bubble can be digitized at the most once. For that case Strand has given the following formula\textsuperscript{1}).

\begin{equation}
kn = \ln \frac{T}{M} \quad \text{(1)}
\end{equation}

where $T$ is a total number of scan lines, $M$-number of misses, $k$-density of bubbles in the film and $a$-effective scan width. This formula was derived for straight tracks which are orthogonal to the scan line direction.

We had to generalize this formula for arbitrary tracks. One notices that for inclined tracks a projected density $k_{\text{proj}} = K/\cos \alpha$ is being measured ($\alpha$ is the angle between the tangent to a track and the normal to the scan direction). So for curved tracks, the observed number

\text{---)

\text{*) On leave from the Institute for Nuclear Research, Warsaw}
of misses will change along the track according to

$$M = T \exp(-ka/\cos \alpha), \quad (2)$$

where $\exp(-ka/\cos \alpha)$ is the probability for a miss. In the derivation of the general formula for $ka$ we have assumed that $\cos \alpha$ is constant within every "slice" (a part of the track usually 32 scan lines long, i.e. about 2 mm in the film), and following reference 1 we have written the likelihood function:

$$a(ka) = \prod_{i} (\exp(-ka/\cos \alpha_i))^{M_i} (1-\exp(-ka/\cos \alpha_i))^{H_i}, \quad (3)$$

where index $i$ numbers the slices and $H_i$ is the number of hits in the $i^{th}$ slice ($H_i + M_i = T_i$). Its maximum is given by the value of $ka$ which solves the following equation:

$$\sum_{i} \frac{M_i}{\cos \alpha_i} = \sum_{i} \frac{H_i}{\cos \alpha_i} \frac{1}{\exp(ka/\cos \alpha_i) - 1}, \quad (4)$$

Let us consider the programs at CERN which deal with ionization information.

In HAZE ionization as seen in the film is evaluated, as well as its external and internal errors*). We start by applying an empirical correction for hits which were lost during the digitization and filtering processes (only slices for which average points were found are used). This correction was found to be $\sim 30\%$. Then we reject slices for which the measured values of ionization differ by too much from the mean value.

*) This part has been written in FAP by W. Krischer and P. Seyboth.
This is done as follows: we first correct the numbers of misses for differences between actual and mean values of \( \cos \alpha \):

\[
M'_1 = M_1 \left( \frac{\Sigma T_i / \cos \alpha_i}{\Sigma M_i / \cos \alpha_i} \right) \cos \alpha_i - 1
\]

(5)

then from the \( M^* \) distribution we compute \( \bar{M}^* \) and its internal error

\[
\sigma^{\text{INT}}(\bar{M}^*) = \sqrt{\frac{\Sigma (M_i^* - \bar{M}^*)^2}{N(N-1)}}
\]

(6)

where \( N \) is number of slices. All slices for which \( |M_i^* - \bar{M}^*| \) is greater than \( 2 \sigma^{\text{INT}}(\bar{M}^*) \sqrt{\frac{1}{N}} \) are rejected, and the new mean \( \bar{M}^* \) is computed. The external error is calculated according to the following formula:

\[
\sigma^{\text{EXT}}(\bar{M}^*) = \sqrt{\frac{\bar{M}^*(T - \bar{M}^*) + A^2 (T - \bar{M}^*)^2}{T \cdot N}}
\]

(7)

where the first statistical term is obtained from the binomial distribution, and the second is an empirical correction which accounts for all other sources of error (\( A = 0.06 \) was found to give agreement between internal and external errors). Ionization is computed by solving the equation (4), its asymmetric errors \( \sigma^+_+(k) \), \( \sigma^-_-(k) \) being the results of the propagation of \( \sigma^{\text{EXT}}(\bar{M}^*) \) and \( \sigma^{\text{INT}}(\bar{M}^*) \) through the approximate formula

\[
kn = \cos \alpha \ln \frac{T}{M^*}
\]

(8)

In the geometrical reconstruction program THRESH, the bubble density in space is computed as a weighted mean of the corrected bubble densities measured in the 3 views. There are 2 geometrical corrections.
which are applied. The first one takes account of the fact that the light-rays are generally not perpendicular to the track. If we define the angle \( \theta \) to be the angle between the light-ray from the camera to the middle of the track and the tangent to the track at this point, then the correction is

\[
(kn)^\text{corr} = kn \cdot \sin \theta \quad .
\]  

(9)

It is computed for each view separately. The second correction takes into account differences in demagnification. If the demagnification in the middle of the chamber is \( D_0 \), then at an arbitrary point it can be written as

\[
D = D_0 (1 + \nu) \quad .
\]  

(10)

The differences in \( D \) influence both the apparent bubble density and the effective scan width \( a \) (as the size of the bubble image in the film depends on \( D \)). These influences are partially compensated and we use

\[
(kn)^\text{corr} = kn(1 - \nu/2) \quad .
\]  

(11)

Finally, the formula for the weights which are used in the computation of the mean value is

\[
P = \frac{1}{\text{Max}(\sigma^2_{\text{INT}}(kn), \sigma^2_{\text{EXT}}(kn))} \quad .
\]  

(12)

The corrected mean values, their errors and geometrical corrections are passed through the kinematical program GRIND, and in a sub-routine JUDGE of the following program BAKE, they are used to identify the events.

We assume that the expected ionizations will be proportional to \( 1/\beta^2 \) and that the value of \( a \) is the same for all tracks.
Therefore we have

\[ I_{\text{calc}}^i = F \left( 1 + \frac{N^2}{p^2} \right) \]

(13)

and

\[ I_{\text{meas}}^i = (ka)^i = ak_i \]

(14)

For all hypotheses not rejected by GRIND we compute the minimum value of the following \( \chi^2 \)

\[ \chi^2_{\text{ION}} = \sum_i \frac{(I_{\text{calc}}^i - I_{\text{meas}}^i)^2}{\sigma_{\text{EXT}}^2(I_{\text{meas}}^i)} \]

(15)

hence determining the factor \( F \). The corresponding probabilities (number of degrees of freedom being equal to number of tracks minus 1) together with the kinematical probabilities determined in GRIND will be used to check hypotheses put forward by GRIND and resolve most of the existing ambiguities.

The programs were tested using 15 four-prong interactions of 5.7 GeV/c antiprotons in the Saclay 81 cm HBC. They were practically uniquely identified in GRIND, as most of them are 4-contraints events \( p + \bar{p} \rightarrow p + \bar{p} + \pi^+ + \pi^- \).

In Fig. 1 the distributions of the lower external and internal errors \( \sigma_{\text{EXT}}(ka) \) and \( \sigma_{\text{INT}}(ka) \) divided by ionization \( (ka) \) itself are compared. Figure 2 shows a plot of the measured ionization (after fitting) vs. momentum. The upper curve is for nucleons, lower for pions. Experimental values for pions are shown as points, and for nucleons as circles. All circles in the region \( p > 5.7 \) GeV/c correspond to beam tracks. Ionization 5.0 which seems to be the upper limit of the method is marked.
Finally the ionization $\chi^2$ distribution (Eq. 15) is presented in Fig. 3 (blank area). For comparison the distribution for all other (and so wrong) hypotheses which are considered in the CERN antiproton HPD experiment is shown as normalized, shaded area. The mean $\chi^2_{\text{ION}}$ value for correctly chosen hypotheses is 4.7 (4 degrees of freedom).
References


Figure captions

Fig. 1  A comparison of external and internal errors.

Fig. 2  A plot of the measured ionization (after fitting) versus momentum.

Fig. 3  The ionization $\chi^2$ distribution (clear area) compared with all other wrong hypothesis (shaded area).
Fig. 1

Fig. 2
Fig. 3
DISCUSSION

STRAND: Do you really apply to your "hits" a correction factor of 30%? Our correction factor is only about 3%.

MICHEJDA: Yes, it is correct.

STRAND: One would expect that to improve. I want to ask you about the $\chi^2$ - plot you showed us. It is somewhat disturbing to see the wrong hypothesis getting a low ionization $\chi^2$. If you get a low ionization $\chi^2$ for a wrong kinematical hypothesis will the high kinematical $\chi^2$ rule out that hypothesis?

MICHEJDA: Yes, that is correct.

WEBRE: Just a comment on the 30% and 3% difference observed between CERN and Brookhaven. I think it is because the master points at Brookhaven are calculated using a crude histogram, which, while it may not clip off the points that are on the track, may include some that are not. The finer histogram used at CERN tends to exclude not only the background but also some points which belong to the track.
HPD PROGRAMMING AT THE UNIVERSITY OF PENNSYLVANIA

J. D. ENGMAN
Pennsylvania University, Philadelphia

(presented by J. D. Engman)

The Pennsylvania HPD is as yet incomplete and therefore most of the programming work which I am going to describe is still in the thinking or planning stage. The stage assembly is due to arrive later this fall, from Sogenique, and we hope to have our machine running shortly after the first of the year.

Our HPD will be directly coupled to an IBM 7040 Computer. This machine is also the primary machine in the University's Computer Center. It has the DDC and 6 tape units on Channel C, a 1301 disk, and 6 tape units on Channel B, and a typewriter and 1622 reader/punch on Channel A. Although our HPD could efficiently use a computer several times faster than the 7040, the 7040 was thought to be the machine best fitted for the initial work load expected.

The 7040 is presently being used about 80 hours a week performing general University computing, including debugging of several large physics programs. When the HPD becomes operational, the 7040 will become saturated within 6 to 12 months, with HPD use taking about 60 hours a week. To increase our measuring capacity, we are considering another computer, possibly a separate machine, such as a small configuration PDP6 to do HPD work, while the University Computer Center is looking at larger systems in the IBM 360-60, Univac 1108, GE635, or even CDC 6600 class.
Since the 7040 is quite similar to the 7090 series machines, (one of the reasons for our getting it), conversion of present 7090-94
HPD programs was simplified. Our first job was conversion of the basic
BNL HAZE program, as set up for the BNL 20 inch chamber format, to the
7040. This job has been completed and the program has been debugged
using tapes of HPD digitizings from BNL. Apart from changing 7090
instructions which do not appear on the 7040, the major jobs were
rearranging input and output formats to conform with University of
Pennsylvania practice, writing new subroutines to communicate with the
HPD, and changing input, output and other monitor communication sections
to work with 7040 IBSYS. We also removed all the CRT subroutines, Manual
Filter and several other BNL diagnostic and manual intervention sub-
routines. We are presently using TRED-KICK for our reconstruction and
kinematic programs. We expect soon to convert to an event library
system developed at Pennsylvania. We are also considering using the
THRESH-GRIND system for future experiments.

MIST changes again were those necessitated by changing from the
7090 to the 7040. We used J. Friedman's BNL version of the program as
a basis. The vertex and track checking sections have been changed to
agree with procedures developed at the University of Pennsylvania. Our
scan tables use the Mangiafico biradial measuring heads, and they punch
2 radial co-ordinates plus an X, Y or no overpunch to indicate in which
of 3 slices the point was measured. The output format has been changed
to a binary record so as to decrease the amount of processor time required
in HAZE.

We are at present writing an HPD diagnostic program, based on
the CERN PAN3 but extended to evaluate digitizings from special test film
strips. (See Fig. 1.) The different weight lines should enable us to
calibrate photomultiplier gain. The slightly divergent lines should tell
us the resolution of the track center circuits. By measuring the fiducials
in both scans we can check orthogonality of both scan modes. In addition
we hope to be able to use the diagnostic program to calculate the HPD machine constants necessary for co-ordinate transformation in HAZE or similar programs. The diagnostic program will also be expanded to verify that HPD orders such as move film, move stage, etc., are properly executed.

The next programming job is modifying HAZE to measure the film which will be produced at the 15 inch Pennsylvania-Princeton bubble chamber, located in Princeton, N.J. Since this film has three views on one piece of 70 mm film, all three views will be scanned at one film setting (see Fig. 2). There will be a NIMNS type data box with a $5 \times 60$ array, which will be encoded with frame number, magnet current and other variables, such as hodoscope output. Major modifications to the HAZE dispatcher are contemplated to attempt to get some degree of overlap between stage and film motion and processing. (See Fig. 3.) This film format is very well suited to future programming developments which would use the information derived from finding an event in one view to find the event in the other two views. This would mean that one could hopefully prescan only one view. Likewise, if one were to use track following and vertex finding programs such as those being developed by Marr and Rabinowitz at BNL and Dickens and Dowton at LRL, one could conceivably use the co-ordinates of an event found in one view to perform some sort of "pregating" to reduce the processing time required to find the event in the other two views.

We also plan to modify our present HAZE to process 70 mm film taken at the BNL 80 inch bubble chamber. The changes contemplated are primarily modifications of the program constants rather than any major logical change to the program.

The spark chamber experimental groups at Pennsylvania have shown a lot of interest in using the FSD to analyze their photographs. We have supplied several of these groups with information and equipment to put HPD data boxes, fiducials and strobe marks on their film. We have not
yet started to do any programming to analyze spark chamber photographs although I expect we will start some of this work by December 1.

To facilitate programming for the HPD by other groups, we have written a series of macros (the 7040 assembler, MAP, will process macros) which will set up the proper calling sequences for the HPD commands.

A short review of the order codes built into the Pennsylvania HPD might be of interest (see Fig. 4). We have an ordinary move film (MF) order which will move film backward or forward to a target frame number or forward by an increment. Our move stage order (MS) is likewise quite conventional, moving the stage to an initial X, W co-ordinate before beginning the scan. The set mode (SM) order, in addition to setting forward or reverse, normal or abnormal (orthogonal) mode and the stopping address, also has several other options. One can set the HPD to produce either an EOF or an EOR when it reaches Xf (or Wf); one can set the HPD to produce an EOF signal when the stage reaches Xf (or Wf) (Xfa = 1) while the stage continues to scan and another SM order would be necessary to stop it (Xfa = 0); and one can set the HPD to produce an interrupt signal at the end of every 1, 2, 4 or 8 scan lines. (N scan lines = 1.) This is in contrast to our usual use of the interrupt line to signal an error condition at the HPD. The annunciator output order (AO) merely turns on a light on the HPD console to indicate end of film, program finished, computer buffer overflow, etc..

The W stage co-ordinate is sent to the computer as the first word in a new buffer, after a read select signal has been received by the HPD, and is not sent on each scan line.

We are also considering the use of a video gating scheme which would permit us to specify up to 8 roads. We would also be able to vary the road width as the scan progressed across the picture using the "N scan lines" or "Xfa" options in the SM order.
We have investigated using this video gating scheme, in conjunction with a track following program of the Marr, Rabinowitz/Dickens Downton variety.

We are planning to analyze the photographs from a 2 GeV $\rho$ resonance study, on the BNL 20 inch chamber, 200,000 pictures, as the first production run on our HPD.

We at Pennsylvania wish to acknowledge the advice and assistance of the HPD groups which are already in operation. Of particular help to me have been Mr. W. Thompson at BNL and Mr. H. White at LRL. I hope that once we get into operation we can begin to repay some of the help we have received.
Figure captions

Fig. 1 The proposed test pattern for use with the HPD.

Fig. 2 The Pennsylvania-Princeton bubble chamber film format.

Fig. 3 The revised flow chart for HAZE when used with film from the Pennsylvania-Princeton bubble chamber.

Fig. 4 The word formats planned for communicating with the HPD.
LINE THICKNESS VARIED BETWEEN 10µ AND 200µ.

Fig. 1

Fig. 2
### SENSE OUTPUT

<table>
<thead>
<tr>
<th>MF</th>
<th>MOVE FILM</th>
<th>1</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>MS</td>
<td>MOVE STAGE</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>SM</td>
<td>SET MODE</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>RO</td>
<td>ANNUNCIATOR</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>NG</td>
<td>NO GATE</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>VG</td>
<td>VIDEO GATE</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

### OUTPUT DATA BUS

<table>
<thead>
<tr>
<th>Frame Number</th>
<th>10^8</th>
<th>10^7</th>
<th>10^6</th>
<th>10^5</th>
<th>10^4</th>
<th>10^3</th>
<th>10^2</th>
<th>10^1</th>
<th>10^0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frame</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- X - Starting Address
- W - Starting Address
- X, S, or W - Stopping Address

### SENSE INPUT

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
DISCUSSION

SCHIFF: I did not understand, what you said about variable roads.

ENGMAN: This idea which was put forward several years ago by Miura at CERN, is for the computer to specify values \( W_1 \) and \( W_2 \) and only \( W \)-values along the scan line which fell within these limits would be digitized. Thus you would only pick up points within the road and then you could vary the width and location of this road as you move along the picture.

MOORHEAD: How do you update the road information and how often do you feed it back to the hardware?

ENGMAN: This is why we have the possibility to interrupt after a specified number (1, 2, 4 or 8) of scan lines, update the road and continue across the picture. We have another possibility of updating our road in our Set Mode operation. In this we have a provision to give an interrupt signal when the stage reaches a preset value.

BLOCH: How do you digitize with the 52 mm scan of the Sogenique a 70 mm film?

ENGMAN: We would digitize one view at a time because the necessary information for one view is contained in a strip of about 51.5 mm.

HART: Where are you going to put your fiducials?

ENGMAN: They will be right on the bolt circle holding the chamber window. We have bolted metal "ears" right on the chamber, which will have the fiducials on them.

HART: It may be difficult to illuminate them in that position. At Brookhaven we preferred to mount them directly on the cameras since
that seemed a very stable and convenient arrangement.

ENGMAN: We considered both possibilities and went to this scheme because we thought it would be more stable.

DEUTSCH: Is it not important to have fiducials rigidly located with respect to the tracks if you want to measure with respect to the fiducials?

HART: It does not matter where you put the fiducials as long as they are in a plane and you know their location well enough relative to the optic axis of the lens - this of course has to be the case in aerial photography. Fiducials on the camera do not tell you about distortions in the chamber, that you will have to gather from having had fiducials at one time in the chamber. However, once you have done that, if the window moves between runs - as it does in the 80" chamber at Brookhaven - it does not really matter. You have calibrated all the space in front of the camera once you have a set of fiducials stable with respect to the camera.
CURRENT STATUS OF THE DIGITAL AUTOMATIC
SCANNING PROGRAMS

R.E. MARR
Brookhaven National Laboratory, Upton, N.Y.

1. INTRODUCTION

In this talk, I shall attempt to give a reasonably up-to-date
picture of work being carried out at both Brookhaven and Berkeley
directed toward a software realization of automatic scanning.

First of all, I must make it clear that, as far as the Berkeley
programs are concerned, I am speaking strictly as an interested observer.
Since no one from the Berkeley group could attend this meeting, they
have supplied me with some slides and briefed me on certain aspects of
their work with the understanding that I might give a joint presentation.
I shall endeavour to make this presentation a balanced one, although my
greater familiarity with the Brookhaven programs will inevitably result
in somewhat more space being devoted to it.

Let me begin with a brief overall view of where things stand
at present. Both systems start off with a track reconstruction phase
which is designed to operate in real time during the FSD scan, and which
has as its object the reconstruction, in as complete a form as possible,
of all particle tracks. The second, track-editing phase (called track
linkage at Berkeley) then checks for spurious or incomplete track segments
and takes appropriate corrective action. It is at this stage in the pro-
cessing that the data from normal and orthogonal scans can be merged.

*) Work performed under the auspices of the United States Atomic
Energy Commission.
The final output of track-editing is a list of legitimate, complete tracks for each view, each track being represented by the precision co-ordinates of a number of average points along its length. A third, vertex recognition, phase may then begin rudimentary "scanning" operations, resulting in the listing of all possible vertex candidates recognizable, in a single view.

These various phases currently exist as separate IBM 7094 programs, interfaced via magnetic tape, and we are continuing to test their performance on a variety of sample data — (at Brookhaven, from the 20" chamber, and at Berkeley from the 72" chamber). The first two phases are at similar, fairly advanced stages of development at both BNL and LAL, with modifications continually being made as more operating experience is obtained.

Succeeding stages of processing, involving the merging of data from different views and the recognition of preselected types of events remain in an area not yet given serious attention by either group, as far as I know, although certain ideas have emerged concerning possible systems configurations. I have a few remarks to make concerning this general area, which I shall reserve for the end of this talk.

2. TRACK RECONSTRUCTION FROM FSD DATA

2.1 General Organization of Program

Detailed descriptions of both the Brookhaven and Berkeley versions of the track reconstruction program are available in other reports (ref. 1, 2, 3 and 4). Hence, my remarks concerning the structure of these programs will be in the nature of a review.

The basic design, which evolved out of earlier prototype schemes at Brookhaven, is intended to meet the need for high efficiency in dealing with the large volume of FSD data, while still retaining some flexibility in the choice of criteria for "recognizing" tracks.
The principal features of this design are:

1) reliance on a single *serial* pass over the input data, during which the reconstruction of the various tracks is carried out;

2) the use of iterated blocks of coding, or *track banks* to facilitate the simultaneous reconstruction of many tracks. Each track bank acts primarily as a storage area for all the parameters relevant to a single partially reconstructed track and contains, in addition, executable instructions for the more efficient handling of data pertaining to the track;

3) a combination of certain fast multi-gating techniques for the disposition of individual input digitizations. These techniques are:

   a) for beam tracks a *pseudo-register* scheme, in which the various track acceptance regions, or "roads", exist as pointers to the associated track banks, inserted at the appropriate position in a long block of memory, which then acts as a one-dimensional "road map" for a single scan line; and

   b) for the remaining tracks, a *directed list* scheme, in which the roads exist as numerical parameters stored within the *track* banks, while the track banks themselves are retained in a list structure ordered according to the geometric position of the roads.

Superimposed on this basic framework is a variety of subroutines which deal with each partially reconstructed track as a unit, initializing new tracks from otherwise unassociated data, setting up each road by extrapolation from the accumulated data on a track, reducing the data for each track by averaging groups of consecutive hits, and deciding when a track or track segment must be terminated. Needless to say, these subroutines, which comprise what we at Brookhaven like to refer to as the
"tracking section" of the program, determine the quality of the reconst-
structed track output, and it is in this area that the greatest need for
continued experimentation exists.

2.2 Beam Track Following

At Brookhaven the beam track map is currently realized by a 4000
word block of memory with each word corresponding to a basic cell 3 least
counts (16 μm on BNL HPD) wide. Each road entered in the map occupies
a string of three consecutive words, corresponding to a total road width
of 48 μm with the Brookhaven HPD. At Berkeley a 3072 word map is used
and the basic cell is 16 least counts (16μm on LRL HPD) wide. Five
words are used for each road so that the total road width with the
Berkeley HPD is 80 μm.

Both programs exploit the normally good collimation of the beam
tracks by setting up a transformation which carried each beam track into
a nearly vertical one, thus minimizing the need for updating the map.
At Brookhaven this is accomplished via a sequence of linear shear trans-
formations, the coefficient of each transformation being determined
locally from the average slope of the beam track ensemble. No assumptions
concerning the overall "profile" of the beam are used and a rudimentary
servoing procedure, based on the net amount of road-shifting in the map,
is used for periodically updating the beam slope parameter. In the
Berkeley program, note has been taken of the fact that the curvature
and entry angle of the beam tracks tend to be stable quantities through-
out the course of a bubble chamber run, and a table of shear parameters
versus scan line number, precomputed on the basis of externally given
information is used. Both of these methods have proved extremely effec-
tive. It seems clear, however, that the Berkeley method makes use of
more information, and may be expected to perform more reliably in the
long run; we hope to incorporate it into future versions of the
Brookhaven program, retaining our present method as an option to be
used when the necessary external information is not immediately avail-
able. It might be thought that the Berkeley method would result in
considerably higher efficiency in the beam track following procedure
since the entire transformation is precomputed in advance. The servoing procedure used for updating the beam slope parameter at Brookhaven actually involves rather little time, however, and I suspect that the difference in efficiencies would be found to be negligible. Many milliseconds are, however, gained in the Berkeley procedure for initializing the beam track ensemble. Because the shear transformation is known in advance a reliable histogramming technique can be used near the top of the picture for this purpose. At Brookhaven all tracks are initialized on essentially the same footing with a slope histogramming procedure used at some point near the top of the picture for beam track recognition.

A further difference between the two programs occurs in the manner in which crossing beam tracks are handled. In the Brookhaven program, when two roads collide in the beam track map, both of the interfering tracks are summarily terminated. While this may seem like a somewhat drastic step we have justified it on the grounds that the extensions of the terminated tracks will appear as separate track segments in the final output and can later be linked up as part of the next, track-editing, phase. The Berkeley program attempts to follow both beam tracks through the region of intersection by holding the roads together within the map until the reappearance of a double pulse in the distribution of points. This method has apparently worked with some success. As far as I know, no provision has yet been made for handling overlaps involving three or more beam tracks.

The remaining aspects of beam track following are essentially the same in both versions of the program, with some minor variations. The roads are kept centered about their respective tracks in both cases by a "local guidance" procedure, in which the distribution of hits among the cells of the road (three or five as the case may be) is periodically interrogated in order to determine whether or not a shift in the road position is required. In the Brookhaven program each track is, at some
point in its history, tested for membership in the beam track ensemble; this occurs even for those tracks which are initialized after the original recognition of beam tracks. In the Berkeley program all beam tracks labelled as such must appear in the original group recognized near the top of the picture.

2.3 Non-Beam Track Following

Both programs now include a provision for reordering banks within the directed list, so that non-beam tracks can be followed across intersections.

Both now use a variation of the one-dimensional mapping scheme for initializing tracks, followed by linear and circular fitting schemes for their subsequent extrapolation. The extrapolation strategies used in the two programs, however, differ in many details which I will not attempt to describe at this time. This part of the program is in any case in a high state of flux at both laboratories. The principal qualitative difference appears to be that the road widths used at Berkeley are larger than those at Brookhaven by a factor of two or more.

2.4 Output Format

At Brookhaven an average point is computed for each established track approximately every twenty scan lines. These points are accumulated in an output block assigned to each track bank and all appear as part of the track output along with a number of individual digitizations from the two ends of the track. The data for the so-called "provisional" track segments (those which never attain a length of about 2.5 millimeters) is never reduced, up to sixty-four individual digitizations appearing in the track output for each of these segments. This results typically in about 5000 words of reconstructed track output for a single picture, which currently gets somewhat loosely packed into a 9600 word output buffer area. At Berkeley, finding storage space is an even greater problem than at Brookhaven because of the size of the systems programs, and at most twenty average points are saved for each track. These points are
insured a fairly even distribution along the length of the track through the use of a certain strategy for overwriting previously computed average points with new ones. This results in a much more compact form for the final output although there is, of course, some danger of losing essential information in the case of tracks followed around kinks.

In the Brookhaven program, the residual points - that is, those points which are not assigned to an initialized track when they are encountered - are, for the present, being retained as part of the final output available to the next stage of processing. As yet, however, we have not figured out any good way of making use of this residual information and it may be that future versions of program will see the retention of this residual data abandoned.

2.5 Treatment of Orthogonal Scans

One of the minor innovations introduced into both programs in the past year is a provision for processing orthogonal scans. Basically this is simply a question of turning off the beam following section of the program, following all other tracks in the standard manner. At Brookhaven, however, we have found that beam tracks as they present themselves in an orthogonal scan cause a certain problem in the track initialization phase of the program. This is illustrated in Figure 1, which is a line printer display of the type used extensively by us in the past for examining the functioning of the program on a microscopic level. In this display, a small section of the picture as seen by an orthogonal scan is represented. The scale on the plot is 32 microns per print position and the scan line spacing for this particular data was about 50 microns. It will be noted that the beam tracks, which cross the picture more or less horizontally, cause a bunching of the individual digitizings in small regions of a single scan line. This produces the utmost confusion when our usual track initialization procedure is applied. To avoid the accumulation of large numbers of completely spurious segments, therefore, a modification has been introduced
into the routine which handles the unassociated points in the orthogonal scan. This consists essentially of keeping a running account of the local density of points along a scan line, and suppressing track initialization when this density is excessive. This method has worked quite well over the somewhat limited selection of data to which we have applied it to date. As is apparent from Figure 1, virtually all of the beam track digitizations appear only as unassociated points, indicated by asterisks in the plot, yet the program nevertheless has little difficulty initializing legitimate segments, typified in this plot by the cross arms of the fiducial. A major advantage of weeding out most of the spurious segments at this stage is that very little reconstructed track data is output from the orthogonal scan and it becomes quite feasible to fit it in memory along with the data from the normal scan.

2.7 Memory Requirements and Timing (BNL only)

The memory requirements of the Brookhaven track reconstruction program have been summarized in Table I. It is necessary to make a few explanatory and qualifying comments. The memory allocations among the four principal categories shown determine the spatial constraints which must be taken into account if additional programs, such as track editing and vertex finding, are to coexist within the computer memory. Thus, the non-erasable section represents memory definitely unavailable to such additional programs, the erasable category represents memory usable for storage of intermediate results, and the unused section of memory is available for the additional programs themselves. The track banks, although an integral part of the program, are included under erasable storage because they can be regenerated from a single "template" at the beginning of each picture in 30 to 40 milliseconds. The one-dimensional map used in the track initializing schemes uses effectively zero storage since it is realized via the decrements of the same block of words used by the beam track map. The figures given in the left-hand column represent storage allocation as it actually exists in the current version of the Brookhaven program. In the right-hand column I have
### TABLE 1

MEMORY REQUIREMENTS IN BNL TRACK RECONSTRUCTION PROGRAM

<table>
<thead>
<tr>
<th>Non-erasable</th>
<th>Storage needed (units of 1000 words)</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Current version</td>
<td>Future version (estimate)</td>
<td></td>
</tr>
<tr>
<td>1. Program and subroutines (excluding track banks)</td>
<td>3.5</td>
<td>7.0</td>
<td></td>
</tr>
<tr>
<td>Erasable</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Tracks banks (currently 60 at 42 words each)</td>
<td>2.5</td>
<td>1.5(?)</td>
<td></td>
</tr>
<tr>
<td>3. Beam track and track initializing maps</td>
<td>4.1</td>
<td>4.1</td>
<td></td>
</tr>
<tr>
<td>4. Input buffers</td>
<td>4.1</td>
<td>4.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total erasable area</td>
<td>10.7</td>
<td>10.7a)</td>
</tr>
<tr>
<td>Output</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Track output blocks</td>
<td>9.6</td>
<td>6.5</td>
<td></td>
</tr>
<tr>
<td>6. Residue buffers</td>
<td>6.2</td>
<td>(1.0)</td>
<td></td>
</tr>
<tr>
<td>7. Table of stage coordinates</td>
<td>2.0</td>
<td>0.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total output area</td>
<td>17.8</td>
<td>7.0b)</td>
</tr>
<tr>
<td>Unused (in 32.8 K memory)</td>
<td>0.8</td>
<td>8.1</td>
<td></td>
</tr>
</tbody>
</table>

a) Includes item 6 (see text)
b) Excludes item 6 (see text)
attempted to make an educated guess of how it will finally come out as further improvements are incorporated. Item Number 1 is perhaps the most difficult to estimate, except that it is almost certain to increase as more sophisticated track-following routines for handling a variety of special cases are added. I have simply multiplied the present requirements by a factor of two to arrive at the figure shown. The smaller figure shown under Item 2 is based on the observation that the present complement of sixty banks is excessive, and that, with the advent of better initialization schemes, 30 to 40 banks will be found to be sufficient. The biggest saving of space which we anticipate is in the output buffer category. In the present method of allocating storage for track output, the 9600 word output buffer area is used quite inefficiently, and we are currently working on a scheme which will permit considerably denser packing of the output area. As previously implied, Item 6 is an uncertainty. If we were to abandon the idea of retaining the residue as part of the output this buffer could be shortened to a few hundred words (for use by the track initialization routine) and could be regarded as part of the erasable storage area. At the present time we retain the X (or Y) co-ordinate at the beginning of every scan line (Item 7), certainly an unnecessary procedure. By retaining this number only for every fourth scan line, say, many words of storage may be regained. In summary, then, we anticipate that 8100 words of storage will in the future be available within a 32 K memory for additional programs, with an additional 10,700 words available as an intermediate store. Let me re-emphasize that these figures apply strictly to the Brookhaven version of the program and are based on the assumption that certain improvements now contemplated will work out as we hope.

The computer time required to process various 20" bubble chamber pictures through the Brookhaven track reconstruction program has been measured using the core clock on the 7094. In general, the processing speed seems to come out in the neighborhood of 5,5 thousand W - co-ordinates
per second, a figure somewhat disappointing in view of our earlier estimates of 10 thousand points per second. We suspect that a good deal of the time is spent on bookkeeping and linear fitting operations associated with the unduly large number of short spurious segments which appear in the output. We expect shortly to have a more elaborate track initialization scheme in operation and it is our hope that somewhat greater processing speeds as well as an improvement in the quality of the output will result.

2.7 Display Programs

Until recently, one of the principle stumbling blocks to making progress in the improvement of the program has been the absence of a suitable method of displaying, and hence, of evaluating, the output. The line printer displays, such as Figure 1, have proved useful, but rather large quantities of paper are required to get an overall impression of any reasonably sized section of the picture. Moreover, it is difficult to show all the information which one desires in a plot of this type. Several display programs have now been written for the Calcomp 565 pen-and-ink plotter which show promise of alleviating this problem. A rather good overall impression of the performance of the program for example can be gained from plots of the type shown in Figure 2. On the left is an original photograph from the Brookhaven 20" bubble chamber and on the right a representation of the reconstructed track output. Only track segments containing at least 12 individual digitizations are included in the plot. Beam track average points are plotted as squares and non-beam track average points as open circles. The absence of a line joining consecutive average points on a track indicates that the program did not succeed in following the tracks across this section. This particular plot was obtained from an "upstream scan" of the picture, a mode of operation which seems to yield more reliable results in the vicinity of vertices. The scan line spacing was 37 microns. Figure 3 is a similar type of plot produced at Berkeley from 72" bubble chamber data.
A much more useful type of display for diagnostic purposes is shown in Figure 4, which depicts a single reconstructed non-beam track segment. The data, the plot, and the program that produced it are all from Berkeley. To obtain a clearer display, the track has been straightened by fitting a circle to three of its points and plotting the deviation from this circle versus scan line number. The edges of the road actually used by the track reconstruction program are indicated. Individual digitizings are indicated by (χ) if they were assigned to the track and by a dot (.) if they were not. Average points that were saved and appear as part of the output are indicated by a box and the average points that were not saved are indicated by a circle. The tick marks on the plot represent intervals of 25 scan lines in the horizontal direction, and 100 W- least-counts (100 microns) in the vertical direction. The smaller tick marks labelled S along the lower edge of each strip signify the points at which a new slope was recomputed. At the three points labelled C, a new circle fit was carried out. Through the use of this type of plot, which we hope to have also at Brookhaven in the near future, it should be possible to examine the behavior of the program and make necessary improvements in a much more efficient manner than previously.

3. TRACK EDITING

The Brookhaven track editing program begins with a pass through the list of segments output by the track reconstruction phase, during which a map of track endpoints is set up, and segments containing fewer than 12 individual points are discarded. The remaining segments, exclusive of those which pass all the way through the chamber, are then ordered according to length. The longest segment is selected and tested for the presence of kinks by fitting a circle to three average points on the segment and testing the remaining average points for excessive radial deviation from the circle. Unfortunately, we do not yet have a satisfactory algorithm for locating kinks once their presence is suspected.
the only action taken at present is to set an indicator bit in the track heading. It is worth remarking that in our limited experience to date we have not yet encountered any cases of kinked tracks.

Following the tests for kinks a search for extensions to the segment is carried out among all the remaining (shorter) segments. All shorter segments with an endpoint within a certain coarsely defined region about the ends of the track in question, are selected as possible candidates. The radial deviations from the previously computed circle of the endpoints on each of these candidates are then calculated. If for some candidate neither of these deviations is excessive, a new circle is fitted to points taken from both segments and the deviations of all the average points again computed for the purpose of making a more stringent test. When two segments are linked together in this manner, the interior endpoints are deleted and a search for further extensions is carried out. This entire process proceeds iteratively until no further extensions can be found, at which point the linked segments are labelled as a maximally linked track and the longest remaining unlinked segment is selected to begin a new track.

The Berkeley track segment linking program is divided into two stages of which the first uses only the slope and its rate of change near the two endpoints of a segment in searching for extensions. This is a much faster procedure than the circle fitting one and is apparently sufficiently powerful to carry out about two-thirds of the linking. For the rest of the linking, reliance is placed on the second stage which involves circle fitting to the longer segment of the two being matched. This program was found to be about 90% effective on a sample of 76 tracks from a recent run.

An important aspect of the track linkage phase which has not yet been implemented is the merging of track data from the normal and orthogonal scabs. Coding for this will soon be under way at both Brookhaven and Berkeley.
4. VERTEX FINDING

The vertex recognition program which currently exists at Brookhaven represents a preliminary attempt at finding vertices of physical interest on a single bubble chamber view. The program proceeds serially through the list of endpoints which remain after the track editing phase. When an unassociated endpoint is encountered the associated track is assigned to a vertex cluster and a search for additional members of the cluster is initiated. All tracks with endpoints in a relatively large search region about the original endpoint are considered as candidates and approximate intersection points between each of the candidates and the searching track is first computed. For this, the tangents at the two endpoints, as determined from the previously computed circular fits, are used. Candidates are then rejected unless the intersection point is external to both tracks and within one millimeter of the two endpoints. When more than three tracks remain in a vertex cluster, a further test, based on the distance between pairs of intersections, is carried out.

The vertex clusters defined by the program are placed into four mutually disjoint classes as follows: Type 1 consists of those vertices involving a beam track and one or more non-beam tracks. The Type 2 vertices are those involving two or more non-beam tracks and no beam tracks, Types 3 and 4 are reserved for beam and non-beam tracks, respectively, which appear to stop within the chamber. No provision has yet been included for associating strings of vertices belonging to a common physical event.

Figures 5 and 6 are provided as examples of cases in which the program performs particularly well. On the left half of each figure is a section of a bubble chamber photograph from the Brookhaven 20" chamber containing a simple 2-prong event. On the right is a 0-comp display of the Type 1 vertices found by the vertex recognition program. In these plots average points are indicated by squares and the position of the vertex itself by a dot in a circle. Parentheses drawn at right angles to a track indicate linkages carried out by the track editing phases of the program.
These plots are representative of the type of output which we would like to achieve consistently but cannot be taken as typical of the present level of performance. The overall results achieved to date with this program, on a variety of sample data, must be regarded as inconclusive. In general, the program has no difficulty recognizing simple types of vertices, provided the associated tracks appear in reasonably complete form in the output from the earlier phases. The principal difficulty noted is the large number of spurious vertices which appear in the final output. These appear typically in the neighborhood of electron spirals, and in the vicinity of track intersections. Many of these could presumably be eliminated immediately when the data from different views is compared. We feel nevertheless that it is possible to attain considerably better output from a single view than at present and this will be one of our principal concerns in the future.

A provisional program for single view vertex recognition has also been written at Berkeley but has not been actively pursued in recent months, the effort there having been concentrated on the continued improvement of the track reconstruction and track linkage phases.

6. **SYSTEMS CONSIDERATIONS**

It is generally agreed that the output from the first stage of a production system should consist of digital data containing in highly condensed form all the essential information on the original bubble chamber photograph. This information would be stored, presumably on magnetic tape and would comprise a library of "abstract tapes" which would act as the principal data input to succeeding stages of the system, in which scanning criteria selected by the physicist would be introduced. One of the major questions to be decided is whether this abstraction should take place at the level of tracks (end of track editing phase) or at the level of vertices. At Brookhaven we are tentatively planning a vertex abstraction system. This system would work on-line to the HPD and would result in the preparation of a separate vertex abstract tape for each view. In addition
to the normal scan, two overlapping orthogonal scans would ordinarily be performed on every frame. Fiducial finding in each scan, conversion of all co-ordinates to a standard co-ordinate system, track editing, and vertex finding would be carried out during the retrace motions of the stage. Our estimate of the timing requirements for the various parts of the program indicate that this would be a reasonably balanced use of computer and digitizer. We do not anticipate any difficulties fitting the track reconstruction and track editing phases of the program into a single memory load, once certain space saving modifications to the current version have been made, but we are not nearly as confident about including also the vertex recognition phase; preliminary estimates indicate that it will come out very close.

Berkeley plans to run under a three-priority level, multi-programming system known as TRIST, using an IBM 7094, Model 2, with twin 32 K memories. In this system, the real time limited track reconstruction program would operate within one of the memories at the top priority level outputting its results onto tape or disc. Track linking and, possibly, vertex finding would operate on these results at the second priority level, after a short time delay.

I should like to conclude with a few general remarks concerning what might be termed the Interface Problem in automatic scanning systems. For this purpose, let us define the term "Interface" to mean all the unrealized stages of processing between the "pattern recognition" programs which exist currently, on the one hand, and the body of standard spatial reconstruction and kinematical analysis systems on the other. Let us assume for purposes of argument that either a track or vertex abstraction system can ultimately be achieved in which the abstracted information meets minimum standards acceptable to the bubble chamber physicist. There is no reason in principle why this abstracted data cannot be arranged in some universally agreed upon format compatible with commonly used programming languages such as FORTRAN or APL. This is a rather simple concept but it seems to me one that would have many long term advantages. Procedures for selecting
particular types of events could be coded and tried out with relative ease. FORTRAN or ALGOL coded programs could be tailored to the needs of an individual experiment with a relatively small amount of purely programming effort required; such programs could in effect be thrown away at the conclusion of an experiment. Programs found to be successful in more general types of experimental situations could later, in the interests of efficiency, be recoded in machine language. The individual experimentalist might feel encouraged to participate more directly in the automatic scanning effort. This would remove some of the burden from the data processing specialist, who may perhaps not always be cognizant of all the essential aspects of an experiment and who in any case will have his hands full maintaining the abstraction system at a high level of performance. I do not mean to imply by this that there should be no feedback across the interface between the physicist and the data processing expert, especially as the two are often one and the same person. What I am suggesting is that we attempt to arrive at a well-defined line of demarcation between the purely pattern recognition aspects of computer scanning and its physics-dependent parts.

7. ACKNOWLEDGEMENT

Various persons who have contributed to the Brookhaven effort are B. Arbeit, W. Beard, C. Dickens, S. Hollet, G. Rabinowitz, and E. Weneser.

The prototype programs on which the current efforts at BNL and LRL are based were developed in collaboration with both G. Rabinowitz and J.R. Pasta of the University of Illinois, who initiated this project at Brookhaven in 1961.

The Berkeley programs were developed by C. Dickens, and M. Downton, working under H. White. I should like to take the opportunity to thank the group at Berkeley for their cooperation in sending me material for inclusion in this talk.
References


Figure captions

Fig. 1  Line printer display showing data from an orthogonal scan.

Fig. 2  A 20" bubble chamber picture and a display of the reconstructed track output from the Brookhaven program.

Fig. 3  A 72" bubble chamber picture and a display of the output from the Berkeley track reconstruction program.

Fig. 4  Calcomp Plot showing the functioning of the Berkeley program on an individual non-beam track.

Figs. 5 - 6 20" bubble chamber pictures with plots of Type I vertices found by Brookhaven vertex recognition program.
DISCUSSION

TYCKO: Could you say how ionisation information is kept in this program system?

MARR: Ionisation information at Brookhaven is at present available from the way that we accumulate average points. For each track we accumulate hits in groups of 4, 8 or 16 according to the observed density of the track. From information in the heading we know which mode was used and so from the scan-line numbers of succeeding average points we have a measure of the local density along the track. At Berkeley they save at most 20 average points for each track and use an overwriting strategy to keep the average points uniformly distributed. Thus they would lose the density information in this form but it would be relatively easy in either program to keep a running count of the total number of hits and insert it into the track heading.

DEUTSCH: Does the data come in, in such a way that you could use auxiliary storage, like a disc or drum, to make more room for those message banks?

MARR: That is a possibility. Most of our considerations have been limited to the particular computer configuration we expect to have at Brookhaven which does not - so far as I know at the present time - make allowance for a drum on-line all the time.

STRAND: We have discussed informally the possibility of adding extra store. There are even certain things we had in mind for the HAZE program which could use this. I think it will have to be evaluated as we get closer to working out the probable rate of production, but I think it is definitely a possibility.

DEUTSCH: In the first stage of segment reconstruction would it not be convenient if the hardware also gave you tangent information?
MARR: It would be worth considering. However, I think that except for the basic astigmatism of the HPD — and this means you have to use two scans — the information is there in digital form. One of the principle difficulties is simply the limited size of computer memories. If we could keep the whole picture in some sort of auxiliary bulk store we would always be able to go back and re-examine information from any part of the picture. If you had all the digitisations in memory, you would probably still be better off doing a blind serial scan for the first stage of the processing. This is because it is so difficult to get random access to the whole picture when you have 20,000 or 30,000 digitisations to work with. After the track reconstruction phase then of course it would be nice to have completely random access.

LEBOY: With the system as it is at present, what qualitative improvement would you expect if you limited yourself to pictures with not more than five beam tracks?

MARR: Provided the pictures are reasonably clean, with low background and with the beam tracks well separated, everything would go much faster. I think the rejection rate would go way down.

SCHIFF: Would it not be useful to make a start with picture containing only two beam tracks? In that way you would be able to join up with the experimentalist sooner and bridge this "Interface".

MARR: Yes. We have in mind getting 1,000 events with a very low beam intensity just to get some idea of what the programs can do in such a situation. Of course what we want to achieve ultimately is something a little more general than that. There will be many types of experiments that we do not expect will ever be processed suitably by this method, but for experiments involving thousands of events of a similar type and if the type is not too complicated, then I think automatic scanning is possible.
CONCLUDING REMARKS

L. KOWARSKI
Purdue University, Lafayette *)

Let me start with a disclaimer: these remarks will be concerned only with the programming part of the meeting. It is not that I know any more about the programming than about electronics, but during the electronic part I was not present. I even missed the beginning of the programming part, for which I had to seek some help — the helpers will remain anonymous so as not to be saddled with any responsibility for what I am going to say.

Last year I listed all the meetings of the series of which the present one is the sixth and — let us hope — not the last. This sixth conference is the first at which Paul Hough is not present. This is a loss, although we notice that he has a rather long-reaching arm and his influence was felt through some of the Brookhaven speakers. But it is also the first, in this series of conferences, at which Martin Deutsch is with us. He assures me that there is no connection between this absence and this presence, but anyhow that was not my point: I just wanted to interpret Deutsch's presence here as a sign of the increasing interest shown by the physicists in those aspects of our subject which previously were more completely left to specialists in data processing. Obviously, nothing ultimately useful to physics can be done if such specialists do not very closely collaborate with the physicists, not only because physicists are the customers, but also because only they know exactly what is needed, and how to shape the specification of a device. Otherwise the data processing people, left by themselves, would be likely to produce something very efficient and very

*) On leave from CERN
useless. To achieve this close collaboration, the simplest way—and possibly the most effective at least in the first stages—is to combine the two sides under the same skull. The physicist Deutsch has become a specialist in data processing—perhaps more so than he would be willing to admit to himself—and so his presence is a symbol of the coming together of the two crafts which so far have tended to develop separately. In fact, it confirms what we have been saying here to ourselves for many years: that data processing is now a part of physics and of a physicist's daily work.

One interesting result of this evolution is that FORTRAN becomes more and more prominent. This is simply because the physicists refuse to write in any more basic language. As Schiff showed this morning, one of the advantages of the Collège de France system is that the "decoupling", as they call it, between the digital data-taking and the processing of the tape allows them to organize their computer time more flexibly and therefore to use FORTRAN without aggravating the bottlenecks in the computer operation.

Another result of this wholesale entry of new addicts into the programming business is that so many of them are interested in speedy debugging. Their fresh zeal can and has become quite a menace to the habits of computer services in various laboratories. I used to be responsible for running of the computer service in CERN and, if I may indulge in a bit of Schadenfreude, I was quite glad to find how prominent was this problem in many American research centres. In one of the discussions here (I was not present) it was claimed that the only way out is to provide each programmer or physicist with unrestricted access to his own small computer, or at least to install a very efficient time-sharing system. The problem is getting more acute as more and more physicists enter this field; many a laboratory has already been driven, in desperation, buying several small computers which, besides, tend to become not so small at all. In CERN we have decided to take a very big computer, to devise some reasonably efficient scheme for
time-sharing and to hope for the best.

There has been an interesting comment to the effect that as soon as writing and debugging is done by a full-time programmer, the complaint becomes less pronounced. It seems, then, that the ideal collaboration will still involve some division of labour, the physicists being keenly interested and telling the data-processing specialists exactly what they want. At the same time, the day-to-day collaboration will perhaps lead the data processors towards a more active interest in physics.

After these very general remarks, let us turn more specifically to the spark part of the meeting. Here Deutsch's system merits some special attention, because it was the first to work automatically and so far has accumulated the biggest amount of experience of use for physics. During the last year this system definitely proved its claim to be applicable to experiments other than the one for which it was originally established. We know now that it is not only fast and automatic, but flexible as well. We feel no longer, or almost no longer, that experiments should be selected or even tailored to the possibilities offered by the automatic processing device. The way is open for the device to take on all classes of experiments; there will remain of course a considerable amount of adjustment to be provided in the way of taking the pictures, marking the fiducials etc. This point is certainly obvious to everyone present, but not so obvious to all of the physicists. Some of them still are a bit surprised when the developers of automatic data processing devices insist on some small changes in the hitherto accepted ways of photographing or fiducialling; they would expect that whatever they do in this respect should be taken in their stride by the devices. The fact that the photographic aspects of experimentation are now being devised with some regard for future automation of data processing shows again that the degree of collaboration is increasing.

Deutsch's device is the foremost example of a system based on guided scan. As such, it produced quite a triumph for the programmed-spot
idea, although he says himself that, as experiments become more complex, and pictures show more tracks, details and so on, then the raster scan may come into its own. The competition, if we may use this word, remains open.

On the important subject of the rejection rate, I am told that Deutsch has commented on the difficulty of making the device work so reliably that the rejection rate is seriously reduced. From the original 45% or so (it must be less than 50%) it gets narrowed down to one-third, then perhaps to one-quarter and finally to 10% - I say finally, because getting even that far down appears to be quite hard to achieve. In this connection, and as a transition to the spark devices with raster scan, it is worth noting that the most conspicuous success in this class of devices has so far been scored by HPD, as reported by Blackall. The rejection rate was only 8% - and Blackall does not show any signs of an extreme exhaustion after this effort. This may be a faint pointer in a very controversial direction: maybe, after all, raster-scan methods are more efficient as regards rejection rates. Just this brief mention to see how some people get excited.

The success of the CERN HPD, as a fully automatic device for the analysis of spark chamber pictures, also includes (as reported by Zanella) the already discussed capability of adapting itself to different kinds of experiments. But HPD is not wholly representative of the raster-scan class, since it was developed for bubble picture work and is too accurate and too slow for spark pictures with their simpler details and greater numbers. Cathode-ray tubes do go faster at some expense in accuracy, and that is how CERN's Laciale project was started. It got as far as a first attempt to do actual physics - this was reported here. There has been a snag, which however appears to be trivial so that Laciale is in the usual hopeful stage when physics, with no snags, can be expected to be done within very few weeks.

To finish the spark chapter, a short mention of Chloe: this device appears to be somewhat more sophisticated than the sheer raster-scan systems.
At present it is, I think, used for a kind of raster scan over a pre-selected small region, although it could be used otherwise. I might characterize it briefly by giving it the same place among spark devices as that occupied by PEPF in the bubble analysis. Papers presented on it here went rather deep, but so far the device was used mostly in biology because, I suppose, of internal reasons concerning physics at Argonne. It is therefore too early to say now how it compares with other automatic systems for spark analysis.

Going now to the bubbles, the first remark is that—comparing to the last year's papers—there is an impression of a far greater maturity. This means that there has been a recognition of the need for a lot of hard work, a lot of plodding on details. Luckily enough, physicists are getting again more interested, they come in and help, and extravagant expectations are no longer entertained. How much one can expect depends on how far the programmers' ambitions will be allowed to go. The more impatient among the would-be users of HFD and related devices for bubble work are willing to rely on the simplest versions of HAZE and so on; the first generation of these programmes is at work for physics first of all at Berkeley. Their pioneering use of HFD was reported at the last year's meeting; unfortunately we have here no definite recent news from Berkeley. There is a rumour for which I cannot take the responsibility, that nothing substantially new has been achieved there since; we can repeat such outrageous statements safely only because no one from Berkeley is present. Their current HFD—aided physics is done as I remember it, with a fairly high rejection rate and that is how far their initial ambition went.

Brookhaven, on the other hand, seems to be more ambitious. So far they have done perhaps not so much physics as Berkeley has, but they certainly aim deeper in their programming. We heard a detailed and very interesting paper about it. It was very long and yet dense in presentation, which only goes to show what an enormous mass of detail one has to foresee in order to do physics, and how complicated even the first generation can be.
The second generation aims at more ambitious achievements, but its proponents had to delay the application of HPD to physics. Last year I said, meaning no unkindness, that CERN was ambitious and therefore had not, by that time, done any physics with its HPD. There is no unkindness in mentioning a delay which was to be well used and I am glad to hear now that the bubble chamber use of HPD has now started at CERN, and that there is a good hope of finding that this somewhat later entry into the physics business is rewarding, in the sense that second generation programmes are more flexible and fruitful. There are other systems in the same generation with which, so far and for various reasons, no physics has as yet been done; their approach is possibly even deeper. The work done in England continues on the same principles as those being applied at CERN. Two other systems, that of Columbia and that of Collège de France are completely different. Paris stated this morning that there has been no physics because, for one thing, there is no HPD yet. This is certainly a reason and I think it also holds for Columbia: that is how delays in getting machines allows you to play with deeper programmes — à quelque chose malheur est bon.

Marr has just finished presenting a paper which I could almost set apart as a third generation — somewhat parthenogenetic as generations go, because of the two parent lines of the HPD system one has been discarded. There is no human guidance left; I do not know how completely it has been eliminated, because Marr's last words were about an interface which may be another word for humanity's ugly face showing up again in some way, so that the system is not fully automatic after all.

Generally speaking, the elapsed year left me with the impression that no matter how automatic any given system claims to be, the human element comes galloping back and the system begins to incorporate again some sort of human guidance. This is true of PEPFR and was always true of SPASS, which is on its slow way to complete dehumanization but has not got there yet. This is also true of the new Brookhaven gadget for spark picture analysis which I had not mentioned so far. PEPFR, above all, started out by being completely inhuman, and now it seems to rely on quite a bit of human
guidance. And even the spark analysis by HPD which had already achieved complete inhumanity, will be in some types of experiments made easier, as Zanella told us, if there is some human preselection of eventful spark gaps instead of having to go through a totally automatic routine for pointing them out. In all such situations, humans will be employed wherever they are the cheaper solution. As computers become faster and cheaper, dehumanisation will become more and more complete.

I am coming again to my pet opposition of guided scan versus raster scan. In this meeting we have seen successful examples of both. Quite possibly, when PEPR finally does do physics - and we are told that this will happen very soon - it will bring a new and very weighty item to the guided-scan side of the argument. So far, on the whole, it remains true that any trick that hardware can do, software can do too. But hardware will do it, if it really can do it, faster and cheaper. PEPR may be expected to show these two advantages of speed and lower cost. But it remains to be seen if hardware can be kept as free of snags as software can; we have already seen a possible indication to that effect in the low rejection rate reported by Blackall. It also occurred to me in the last moment when composing these remarks, that the Software-Only philosophy - you commit all of your picture to the computer memory and you play your tricks on it once it is there - may prove to be more flexible than the practice of incorporating some of the tricks into the hardware. As everybody knows, you cannot teach an old hardware new tricks, and what may happen is that a tricky-hardware man would like to use his system in some new way. The new trick would then have to be performed in software; this would require again putting a lot of information into the memory, which means that a very big and fast computer will again be needed and one of the main advantages of the present tricky-hardware and programmed-scan systems will be lost. But their other advantages will remain available to the combined system.

And this brings me to my last remark. It seems at present that there is some cross-fertilization between the systems which were originated as rivals,
particularly in bubbles. The three main contenders - SMP, HPD and FFR - as they become more experienced, do not mind learning from each other and in this way some sort of middle term is in the process of being reached. Perhaps - is this a fantasy? - we shall see some day a universal system which will embody elements from all these three devices. It will contain a very tricky piece of hardware, and yet there will be some additional tricks done by software. For the development of this ultimate system all those which have been presented here will have been necessary and there will be a contribution from each. On this hopeful note I would like to conclude these concluding remarks on the sixth meeting - the meeting which has shown that our subject is growing and achieving maturity in a very gratifying way.
LIST OF PARTICIPANTS

Argonne National Laboratory, Argonne

J.W. BUTLER
Mrs M. BUTLER
R.Y. ROYSTON

Bonn University, Bonn

H. DREVERMANN
H.H. NAGEL

Brookhaven National Laboratory, Upton, L.I., N.Y.

E.L. HART
R.B. MARR
R.C. STRAND

Centre d'Etudes Nucléaires, Saclay

B. BENABOU
B. DELER
M. GRANDDIER

M. GOUGON
E. COULAREAU
M. GOLDSWASSER
J.C. MICHAU

Centro di Calcolo Nucleare, Bologna

E. CLEMENTEL

Centro Nazionale Analisi Fotogrammi, Bologna

E. CARITA
M.L. LUVISETTO
M. MASETTI
A. MINGUZZI-RANZI
G.P. SINI
E. TOMMASI
E. VACCARI
D. VENTURI
U. ZANOTTI
Collège de France, Paris

J. CAILLET
C. GUIGNARD
J.P. GRELE
M. BLOCH
Ph. LEBLOND
M. SCHIFF

Columbia University, Irvington-on-Hudson

W. SIPPACH
D.H. TYCKO
N. WEBRE

Genova University, Genova

M. PRIORI-REZZUTO
G. TOMASINI

Imperial College, London

A.R. EDMONDS
S.K. PATIENCE
M.C. TURNILL

Istituto Nazionale di Fisica Nucleare, Bologna

A. CHIARINI
C. LOLLI

Istituto di Fisica, Milano

G. DEGLI ANTONI
S. RATTI

Istituto di Fisica, Padova

M. CRESTI
E. SARTORI

Istituto Nazionale di Fisica Nucleare, Pisa

M. CARRARA

Istituto di Fisica, Roma

F. LEPRI
Istituto Nazionale di Fisica Nucleare, Torino

A. WERBROUCK

Laboratoire de Physique Nucléaire, Orsay

J.C. BIZOT
J. LABERRIGUE
M. RUMPF
M. SAUVAGE
C. de la VAISSIÈRE

Laboratori Nazionale di Frascati, Roma

M. PANDARESE

Massachusetts Institute of Technology, Cambridge

M. DEUTSCH

Max-Planck-Institut für Physik und Astrophysik, München

H. BREITEL
P. FREUND
L. KVASZ
S. LUERS
P. SEYBOTH
J. SEGERLIN

Pennsylvania University, Philadelphia

J.D. ENGMAN
E.L. LEBOV

Princeton-Pennsylvania Accelerator, Princeton

M. BAZIN
J.W. BENOIT
N. STUTZ

Rutherford High Energy Laboratory, Chilton

J.W. BURREN
J.D. FORBES
R.A. LAWES
D. LORD
E.S. ROSS

and

Visitors, Fellows and CERN Staff
The following commercial firms were also represented:

**CISE, Centro Informazioni Studi Esperienze, Milano**

I. DE LOTTO

**Control Data Corporation, Geneva**

R.L. CHEW

**Fabri-Tek Incorporated, Minneapolis**

G.A. TINGLEY

**IBM - Geneva, Geneva**

C.A. HERITIER

**IBM, British Laboratories, Hursley (England)**

M. KINGSTON

**Laboratorio di Ricerche Elettroniche, C. Olivetti, Ivra**

G. PASANO

M. CONTA

**Munzig International, Zürich**

G.A. TINGLEY