Title
DAPR - DIGITAL AUTOMATIC PATTERN RECOGNITION FOR BUBBLE CHAMBERS

Permalink
https://escholarship.org/uc/item/5w80b92g

Author
White, Howard S.

Publication Date
1967-05-29

UCRL-17589

University of California

Ernest O. Lawrence Radiation Laboratory

DAPR - DIGITAL AUTOMATIC PATTERN RECOGNITION FOR BUBBLE CHAMBERS

Howard S. White

May 29, 1967

Berkeley, California

DISCLAIMER

Presented at the
Symposium on Automatic Photointerpretation
Washington, D. C., May 31 - June 2, 1967

Co-sponsored by
Office of Naval Research
University of Maryland
Pattern Recognition Society

UCRL-17589
Preprint

UNIVERSITY OF CALIFORNIA

Lawrence Radiation Laboratory
Berkeley, California

AEC Contract No. W-7405-eng-48

DAPR - DIGITAL AUTOMATIC PATTERN RECOGNITION FOR BUBBLE CHAMBERS

Howard S. White

May 29, 1967

DAPR - DIGITAL AUTOMATIC PATTERN RECOGNITION FOR BUBBLE CHAMBER

Howard S. White

Lawrence Radiation Laboratory
University of California
Berkeley, California

29 May 1967

The bubble chamber is one of the most versatile tools for high energy physics research. It is used to detect the passage of particles and to measure their trajectory, momentum and velocity. Nuclear events are produced by the interaction of beam particles from the accelerator with nuclei of the chamber medium.

Charged particles which pass through the chamber during the sensitive period of its expansion cycle cause formation of strings of bubbles along their trajectories, with bubble density determined by the particle velocity. A magnetic field is present to allow particle momentum to be measured by the radius of its trajectory. Generally, three cameras photograph the bubble chamber during the sensitive period of its expansion cycle, so that by intercomparison of views a space reconstruction of particle tracks can be made.

Bubble chambers currently used are typically two meters in length, and can accept 15 to 20 beam tracks per accelerator pulse. When filled with hydrogen, they yield about one useful event per expansion. Typical experiments take from 100,000 to 1,000,000 picture sets (triads) during 10 to 100 days of chamber operation. Figure 1a shows the LRL 72" Hydrogen Bubble Chamber, housed in its magnet and cryogenic equipment. A view of the chamber being lifted into place is shown in Figure 1b.

Figure 2 contains the three views of a representative bubble chamber triad, and illustrates the typical conditions of unwanted background which confront one who would try to apply pattern recognition in this field.

In past years these pictures have been analyzed by a tedious process which includes manually scanning each triad for events of interest, sketching the events to guide their subsequent measurement, performing the measurement with a manually controlled precision microscope, numerically analyzing the results, and gathering the data into a statistical summarization. Since the last two steps of this process require a much greater time to do manually, automation began with analysis and summarization. With conventional measuring equipment, a severe limitation to event processing rate is imposed by the fifteen minutes required to measure an event.

FSD Hardware

In 1960, it became clear that a Hough-Powell Device (HPD) could adequately digitize a bubble chamber picture with the required resolution and accuracy.[1.] This device employes a mechanically generated flying spot to produce a raster scan of the film image. The spot is generated by the inersection of a moving slit with a fixed slit, illuminated by a mercury vapor arc lamp. The spot is imaged both upon the film to be measured, and upon a precision grating from which the spot position can be accurately known. The other dimension of the raster is generated by moving a mechanical stage, carrying the film along a direction orthogonal to the direction of spot motion.

The optical layout of the Berkeley Flying Spot Digitizer (FSD) is illustrated in Figure 3. It may be noticed that there are no parts in the optical system which move during the change between normal and orthogonal scan modes, so that a permanent calibration of the instrument may be made.

The actual FSD hardware is shown in Figure 4. The HPD in the foreground is used to measure film from a variety of bubble chambers. About half of the electronics equipment is visible in the background. The entire machine was constructed in the Berkeley shops.

In comparison with a CRT spot generator, the HPD has better resolution, greater light flux, higher accuracy, and allows faster spot motion for a given signal to noise ratio. However, because of the inertial characteristics of the mechanical stage, an entirely serial scan of the picture must be made. This introduces no delay since all the area must be scanned, but it does impose a constraint against "second looks" at any picture area.

The Berkeley FSD is attached to an IBM 7094 II computer by means of a Direct Data Connection. Our 7094 is a rather large system having 65,536 words of core memory, five data channels, a disc module and 16 tape drives, and is operated in a multiprogrammed mode.

An executive program, TRIST, responds to the needs of the FSD in such a way that the processing of information from it is given absolute priority, while all central processor (CPU) time not needed for this use is devoted to numerical analysis of events which have been measured at a previous time.[2.]

This hardware configuration of computer and FSD has been used in the routine measurement of bubble chamber data since 1963.[3.] The HAZE system requires manual scanning and rough digitizing of tracks comprising events to be measured, followed by the automatic precision measurement. HAZE thus breaks through the restrictive barrier of slow event measurement, but stimulates the desire to have entirely automatic processing of bubble chamber events. The hardware system has been made thoroughly reliable during the processing of more than 600,000 events by the HAZE system, which now operates at a rate of about 300,000 events per year.[4]

The DAPR System

Even before completion of the FSD hardware, development of a procedure for entirely automatic processing of bubble chamber pictures was begun.[5] In the same spirit of cooperation which marked the collaborative development of FSD hardware at Berkeley, Brookhaven and CERN[6] the first extensive automatic scanning programs which were developed by Pasta, Marr and Rabinowitz, were generously shared with us at Berkeley.[7] From this basis, further development produced in 1965 a set of programs which were prototypes for the production system.

The prototype programs demonstrated that several necessary constraints could be satisfactorily met. It is necessary that an automatic system be able to follow tracks without becoming confused by their crossing, or by the large amount of optical noise present in the pictures. The system must recognize patterns of tracks as desired events, and must find these in all configurations at least as well as human scanners. But just as important is the requirement that the system operate at no greater cost than the cheapest manual system. This requirement imposes a minimum speed of operation upon the system. We therefore asked the question: "Could an automatic system which operates at a cost per event that is no greater than a manual system, produce bubble chamber events with an efficiency and measurement quality that is at least as good as the manual system?" Tests of both speed and quality were very favorably met by the prototype programs.

The DAPR System (Digital Automatic Pattern Recognition), is based upon these prototype programs. Because of its size, and in order to take best advantage of the multiprogrammed computer, it is organized into five principal programs operating in three phases. These phases correspond to the three levels of priority which must be given to their operation. The relationship of the programs to each other and the means by which they interchange data is indicated in Figures 5 and 6.

DAPR Measurement and Abstraction

Highest (A-level) priority is given to the program (511) which controls the FSD and receives data from it. This program and the executive monitor are always resident in one of the core modules. Data from the image being digitized arrive continuously at the computer; about 60,000 words are transmitted for each image in a time of about nine seconds. Accumulation of 2,000 words of digitizings causes immediate transfer of CPU control to the A-level program, and these words are reduced within the time in which the next 2,000 words are accumulated. A real-time constraint is thus imposed upon the program, since failure to complete the reduction of one buffer during the time required to fill the other will result in an irretrievable loss of data. All excess CPU capacity is used by the background program.

Reduction of the digitizings is achieved by deciding for each point in turn whether it should be associated with a track already being followed, whether it may be associated with other points from recent scan lines to initialize the following of a new track, or whether it should be placed in the residue buffer. Initialization and early following associations are based upon linear extrapolations, while well initialized tracks are followed by circular extrapolations. Tracks must maintain a minimum density of digitizations or following will stop. A minimum of twelve digitizings is required to complete initialization in the present version of the program. The result of one sweep (normal or orthogonal mode scan of a single image) is a set of lists, each containing up to 18 precision points representing the locus of each track segment, together with a measure of bubble density along that segment.

With a tandem FSD, sweeps of the two film images alternate. Since retrace and film positioning operations occupy almost as large a fraction of the cycle as does the measurement phase, we plan the addition of a second HPD to the existing hardware, to form a tandem machine.

Three sweeps are typically required to digitize the whole area in both normal and orthogonal scanning mode. Reduced data from each in turn is written onto the disc for communication to the next phase of DAPR.

About once each five minutes, a batch of sufficient size to warrant further processing is built up on the disc for each FSD unit. When this happens, the executive program retires the current background program to the disc, and initiates a cycle of intermediate priority (B-level) DAPR programs. The first of these is brought into core memory, and thus replaces the background program as the user of CPU time not needed by the A-level program controlling the FSD.

The first B-level program (512) acquires the data from the three sweeps of an image, and after converting the track measurements to standard coordinates, joins track segments within sweeps, and links the resultant segments between sweeps. This is a more global pattern recognition than is performed by the

A-level program, but the algorithims still rely upon the fact that tracks can be reasonably well represented as arcs of circles. The 512 program also recognizes fiducials, which have been followed as tracks, and replaces their point lists by the coordinate pair defining each fiducial location. The unwanted fiducials, reference markings for other measurement systems, and certain features of the illumination optics which are in reality fixed noise signals are deleted. The output written to disc is thus a greatly reduced set of point lists, now corresponding to each followed track in the image. Further images are processed in a similar manner until the entire batch has been handled. The executive monitor then retires this program to disc, and brings the next program into core memory.

The second B-level program (513) finds track verticies which correspond to particle interactions and associates the corresponding track images from the separate views. A vertex is defined as the common intersection of two or more circles representing tracks, having the intersection near the end of at least one of the tracks. Each view in turn is searched for verticies. Tracks belonging to a vertex are associated in each view. When all three views of a frame have been searched, the found verticies are compared in terms of the chamber optical configuration. Vertices and their component tracks are associated between views to achieve recognition of geometric configurations in space. Provision is made for association ambiguities to be transmitted to a later and more physics-oriented program for resolution. Results are written onto tape, which therefor contains a digital abstraction of all significant information from the film, together with the lists of recognized charged clusters.

The third B-level program (514) performs no pattern recognition, but is an essential part of the artificial intelligence which must be posessed by any automatic system. It has the function of monitoring system performance and reviewing the data according to criteria of reasonableness. Throughout the operation of the earlier programs, various flags and error codes have been stored to indicate abnormalities in hardware, film, program or operator action. Because most of the errors are of random nature affecting only occasional sweeps, it is usually profitable to repeat a sweep measurement when an error is detected. The A-level program makes this decision, and stores codes describing the initial reason for rejection. The performance monitoring program gathers these bits of qualitative information and collects them into a statistical analysis of system performance during the entire run.

Urgent problems are brought to the attention of the operators before a large amount of time has been used. The program also maintains a journal of processing history for individual frames to allow production control.

When the batch of data has been written onto tape, the executive monitor retires to disc the last of the B-level program stack, and restores the C-level program back to operation. Since only four core-loads are exchanged each five minutes, system overhead is below five percent.

The result of A- and B-level program operation is the generation of a data abstract tape (DAT) containing a representation of each track by its precision coordinates and bubble density parameters measured in each view, together with a list of tracks associated with each spatial vertex. The spatial verticies will be identified with various nuclear interactions in the following scanning program. A good approximation to each particle's memorandum vector is also contained on the DAT. This allows the digital scanning program to use more sensitive criteria than those used by human scanners who must work with qualitative criteria. It is perhaps of interest that the DAT is about three times more compact, by volume, than the original film.

Because each film image must be entirely scanned, almost all of DAPR's time goes into the abstraction process. Using a tandem FSD, the production system will operate at a rate of 180 triads per hour, and will use about half of the 7094's CPU capacity during this time.

DAPR Scanning

Scanning is done by a C-level program (516) whenever a sufficient quantity of film has been abstracted. Since no biasing physics assumptions have been introduced into the data during the abstraction process, a library of DATs will be built, from which many experiments can be quickly performed. Because no further access to the film is required, scanning takes place at 10,000 triads per hour on the computer, as contrasted with the 60 triads per hour scanning rate achieved by a good person.

Scanning consists of recognizing events which meet certain selection criteria, labeling both the verticies and the tracks in accordance with the assumed identities, and editing the measurements contained on the DAT to the format required for further processing. The data now are equivalent to those from other measurement systems. As is true for manual scanning, not only the track configurations, but also the bubble densities yield highly significant clues to the identification of particles participating in events. Neutral particles, which do not leave bubble chamber tracks, are introduced into the assumptions at this point. Neutral connecting tracks between charged clusters can be verified from momentum conservation at the verticies. Any of the criteria used by human scanners can be applied by this program to data contained on the DAT.

Experience with DAPR

We are just now approaching production with the DAPR system. The A- and B-level programs have been operated on a few dozen representative frames. Our experience with them indicates that very satisfactory data abstract tapes can be obtained, and that the system will operate at the planned speeds. Debugging of later programs is still in progress. Comparison of data for individual tracks indicates that DAPR measurement quality will probably exceed that of HAZE, the best current standard.

At this point in the evaluation of results, great attention is given to each individual frame processed. For this purpose, it is very convenient to study the point lists as they are progressively treated by each phase of DAPR. A diagnostic display program (515), which can drive a CalComp plotter, is used to produce a graphical display of these track lists, and to mark the verticies found by the program.

Figures 7 through 16 show the data of a representative frame at several steps from digitization to final form on the DAT. The frame chosen is the one shown in Figure 2. Thresholds for track following were established from other frames having considerably darker tracks. It is thought that the strobe light was fainterthan usual on this exposure, and that this accounts for the greater than usual amount of unfollowed track. Minor threshold adjustment will cure this, and optimal settings are now being found. Comparison of Figures 2 and 16 show that a faithful and unconfused abstraction of data has been made. In Figure 16 the events marked by DAPR correspond to those that can be seen in Figure 2.

While its performance in an actual physics experiment is the ultimate standard by which one can measure the success of DAPR, we have tested all of the elements required by a successful system. From extensive HAZE experience, we know that the hardware is reliable and capable of adequately digitizing the pictures. A detailed study of program diagnostics shows that the digitizings are properly assigned to the track lists. Bubble densities are measured in the same manner as with HAZE, which is known to be satisfactory. A study of the track linking process demonstrates that segments are not confused between tracks. Vertex finding and track association procedures have been demonstrated to yield the same results that are obtained manually. System performance has been timed on enough frames to reliably verify program execution speeds and the meeting of real-time constraints. We are eager to finish final debugging and to begin production.

Where from here?

DAPR will continue to be extended to new experiments, chambers and kinds of conditions. Experience gained during early production will doubtless point toward desirable modifications. The problem of short tracks in the presence of much track-like noise will be further explored.

Bubble chambers are changing, too. Today's two meter chamber is small beside the six meter chambers now being designed. These will be photographed by arrays of cameras which have front lens elements projecting like fisheyes into the chamber medium. The photographs produced will be severely distorted, and each view will cover only part of the chamber. Manual processing of these pictures appears impractical but the insertion of a simple transformation in DAPR's program 512 restores the data back into present form.

As technology continues to improve, it is likely that direct electronic sensing of the bubble chamber images will become possible. Limitations now imposed by resolution, accuracy and bit rates will gradually disappear, and the techniques learned with today's hardware will provide the firmest basis for future developments.

ACKNOWLEDGEMENTS

The development of the Flying Spot Digitizer at Berkeley has been the work of many people. The mechanical hardware was designed and built under the leadership of Jack Franck. The electronic design was initiated by Jerome Russell, and implemented by a group working under Jack Salvador and Gene Binnall. Programming developments toward DAPR derived great benefit from the work of John Pasta, Robert Marr and George Rabinowitz at Brookhaven. Initial programming for the prototype programs was performed by Charles Dickens and Mary Downton. The final DAPR system has been developed by a number of programmers working in the Berkeley Data Handling Group, among whom, Dennis Hall, Joan Tyson, Frank Windorski, Joyce Crawford, Joan Franz and Nan Jontulovic deserve particular mention.

The continued interest, encouragement and firm support given to these projects by LRL Director Edwin McMillan has made it all possible.

REFERENCES

1. P. V. C. Hough and B. W. Powell, A Method for Faster Analysis of Bubble Chamber Photographs. Proceedings of an International Conference on Instrumentation for High Energy Physics, 1960. Berkeley (Interscience Publishers, N. Y., 1961) p. 243.
2. C. Osborne, R. Larson, T. Oliver, J. Stekler, and H. S. White, TRIST - A Tri-level Multiprogramming Executive Involving Real-Time Calculations and an Asynchronous On-line Data Source. Lawrence Radiation Laboratory Report UCRL - 16199, June 1966.
3. D. Hall, FSD Operating Experiences at Berkeley. Proceedings of a Conference on Programming for HPD and Other Flying Spot Devices, Paris, 1963. CERN Reports 63-34, October, 1963.
4. H. S. White, Status of the LRL Flying Spot Digitizer, Proceedings of the 1966 International Conference on Instrumentation for High Energy Physics, Stanford. (Clearinghouse for Federal Scientific and Technical Information, NBS, Springfield, Va.)
5. H. S. White, An Automatic Scanning System for Bubble Chambers. Report of the July, 1962 Informal Meeting on Track Data Processing, Geneva. CERN Report 62-37, page 33.
6. H. S. White, T. Aronstein, C. Osborne, N. Webre, W. G. Moorhead, Preliminary Operating Experience with Hough-Powell Device Programs. Proceedings of the 1962 Conference on Instrumentation for High Energy Physics, Geneva. Nuclear Instruments and Methods 20 (1963) page 393.
7. R. Marr, J. Pasta, and G. Rabinowitz, A Preliminary Report on Digital Pattern Recognition of BNL. Brookhaven National Laboratory Report 6866, February, 1963.

FIGURE CAPTIONS

(When a figure contains more than one part, the labeling letters run from left to right in alphabetical sequence).

Fig. 1a The Berkeley 72" Hydrogen Bubble Chamber, surrounded by its magnet and cryogenic equipment. The actual chamber is near the center of the magnet, and is photographed by a camera that looks downward from the floor level of the upper platform.

Fig. 1b The Berkeley 72" Hydrogen Bubble Chamber being lifted into place in its magnet.

Fig. 2 Three views of the chamber during the sensitive time following an accelerator beam pulse. (a), (b), and (c) are, respectively, views 1, 2, and 3. Intercomparison of the loci of tracks referenced to the fiducial marks, allows spatial reconstruction of the tracks and from these parameters describing the nuclear interactions are computed. Note the flare which is always present in view 1, and the striations due to the illumination optics, especially in views 2 and 3.

Fig. 3 A schematic drawing showing the optical arrangement of the Berkeley Flying Spot Digitizer.

Fig. 4 A view of the Berkeley Flying Spot Digitizer, showing the Hough-Powell Device in the foreground, and about half of the electronics in the background. This equipment is attached online to an IBM 7094 II.

Fig. 5 A block diagram of the A- and B-level DAPR programs which prepare the Data Abstract Tape.

Fig. 6 A block diagram of the C-level programs which analyze events "scanned" from the Data Abstract Tape. The DAPR scanner is shown as program 516 on the left of the diagram.

Fig. 7 (a) The actual bubble chamber photograph (cf. Figure 2b).

(b) A plot of the digitizings produced by the FSD operating in the normal scanning mode. Note that solid lines running in a direction parallel to the short edge of the picture are not digitized in this mode. However all tracks running parallel to the long direction are faithfully reproduced.

Fig. 8 (a) The normal mode digitizings . Digitizings are assigned by program 511 either to the track buffers or to the residue buffer.

(b) At the end of a normal sweep, the track buffers contain a set of average points formed from all digitizings assigned to the track. The display program connects these average points with lines.

(c) The residue buffer contains all digitizings not associated with tracks. Tracks visible here had digitizings that were more sparse than allowed by track following thresholds.

Fig. 9 (a) A part of the normal mode digitizings shown in Figure 8a.

(b) The corresponding area, taken from Figure 8c, showing residue points. Inspection shows that digitizings near but not on tracks are not assigned to the tracks, but are correctly put into the residue.

Fig. 10 (a) The actual bubble chamber photograph.

(b) A plot of the digitizings produced by the FSD operating in the orthogonal scanning mode. Note that solid lines running in a direction parallel to the long edge of the picture are not digitized in this mode. However, good digitizings of lines parallel to the short edge are obtained. The two solid lines are due to an artificial digitizing which is used to check the FSD hardware.

Fig. 11 (a) The orthogonal mode digitizings.

(b) Tracks followed in the orthogonal mode.

(c) Residue points in the orthogonal mode. Note that the large group of lines resulting from the chamber illumination noise have been left in the residue.

Fig. 12 (a) The normal mode tracks as they stand after program 511 has written them on the disc.

(b) The orthogonal mode tracks after program 511.

(c) Program 512 brings together tracks from all sweeps. These are shown before joining or linking has taken place.

Fig. 13 (a) Tracks from all sweeps before joining process.

(b) Tracks after partial joining. An example of joining is the leftmost beam track, which was followed in two segments (No. 31 and No. 67) and which has now been joined together.

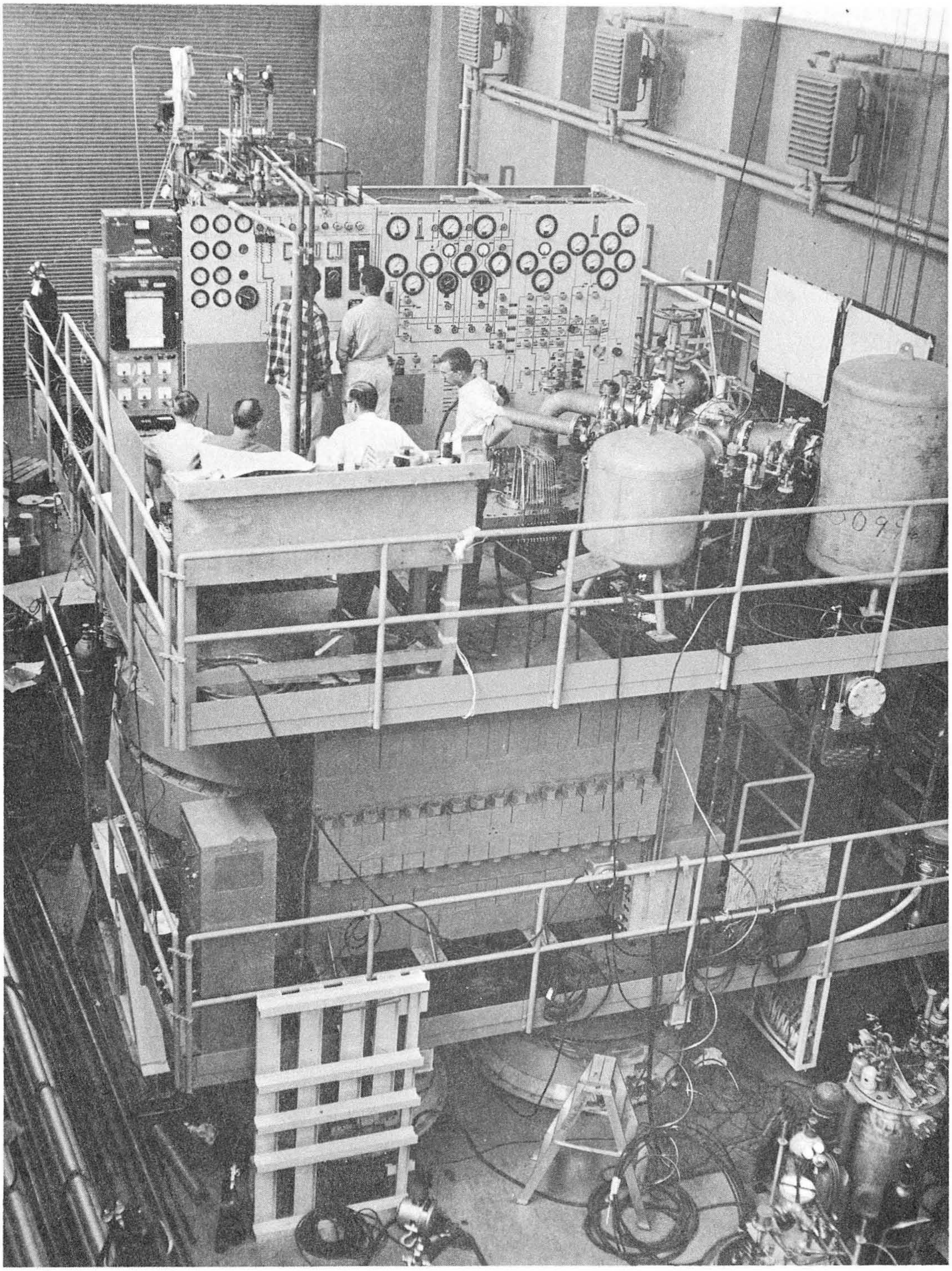
(c) Tracks after joining has been completed. The fiducials have been replaced with their point of intersection, and other fixed noise has been deleted.

Fig. 14 (a) The original photograph of view 2.

(b) The tracks as they are written onto the Data Abstract Tape. Programs 511 and 512 have operated on the data at this step.

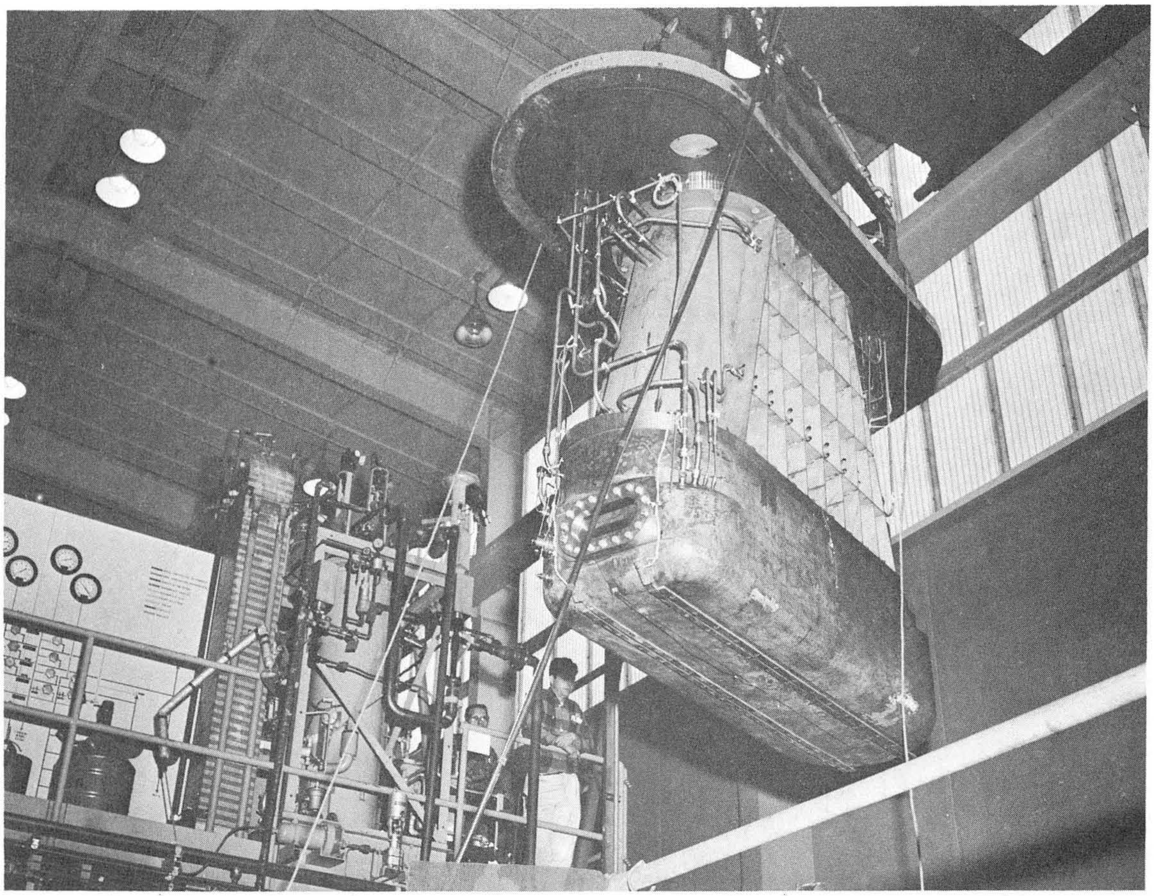
Fig. 15 Tracks from the three views of the chamber, as they stand at the completion of program 511 and 512. These may be compared with the original photographs in Figure 2.

Fig. 16 Tracks from the three views, with verticies found in individual views by program 513 shown as small circles. Large circles mark the verticies which have been associated in at least two of the views. These are candidates for analysis when selected by scanning criteria.



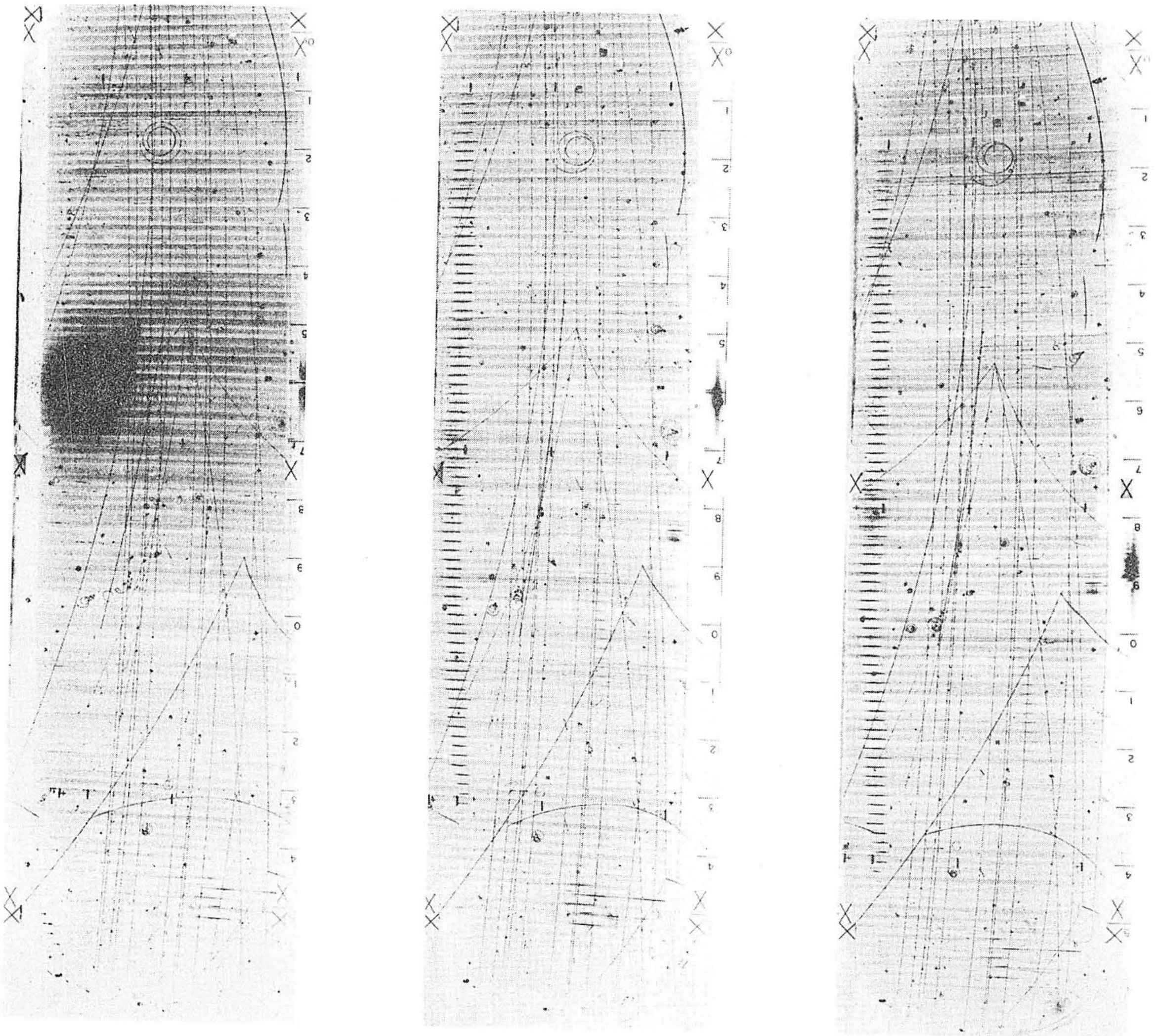
Bub Ch-625

Fig. 1a



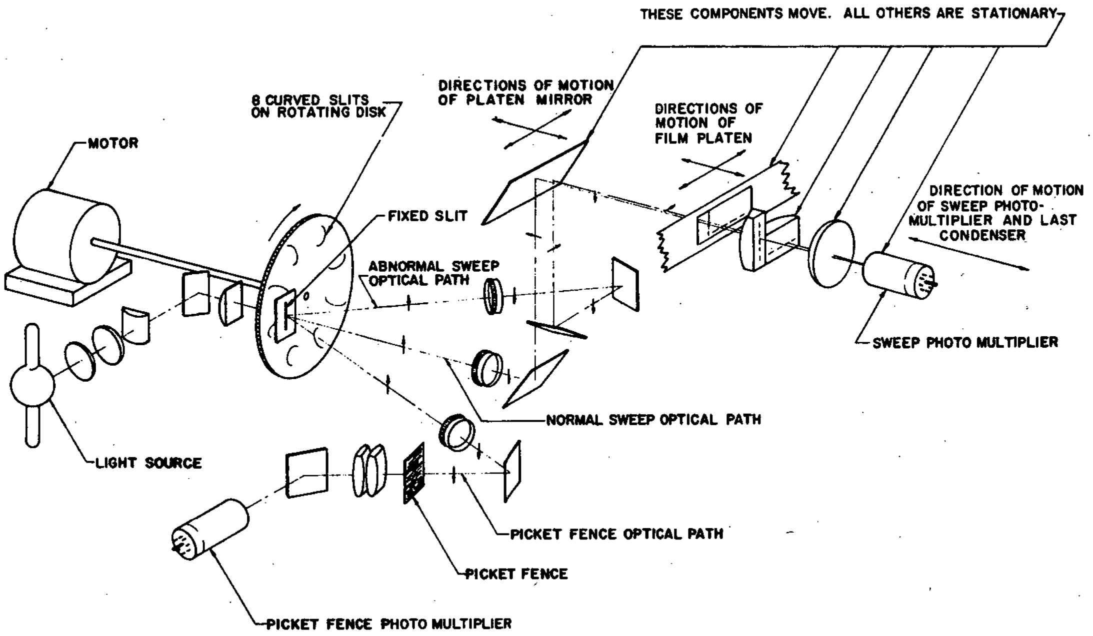
Bub Ch-687

Fig. 1b



XBB 675-3003

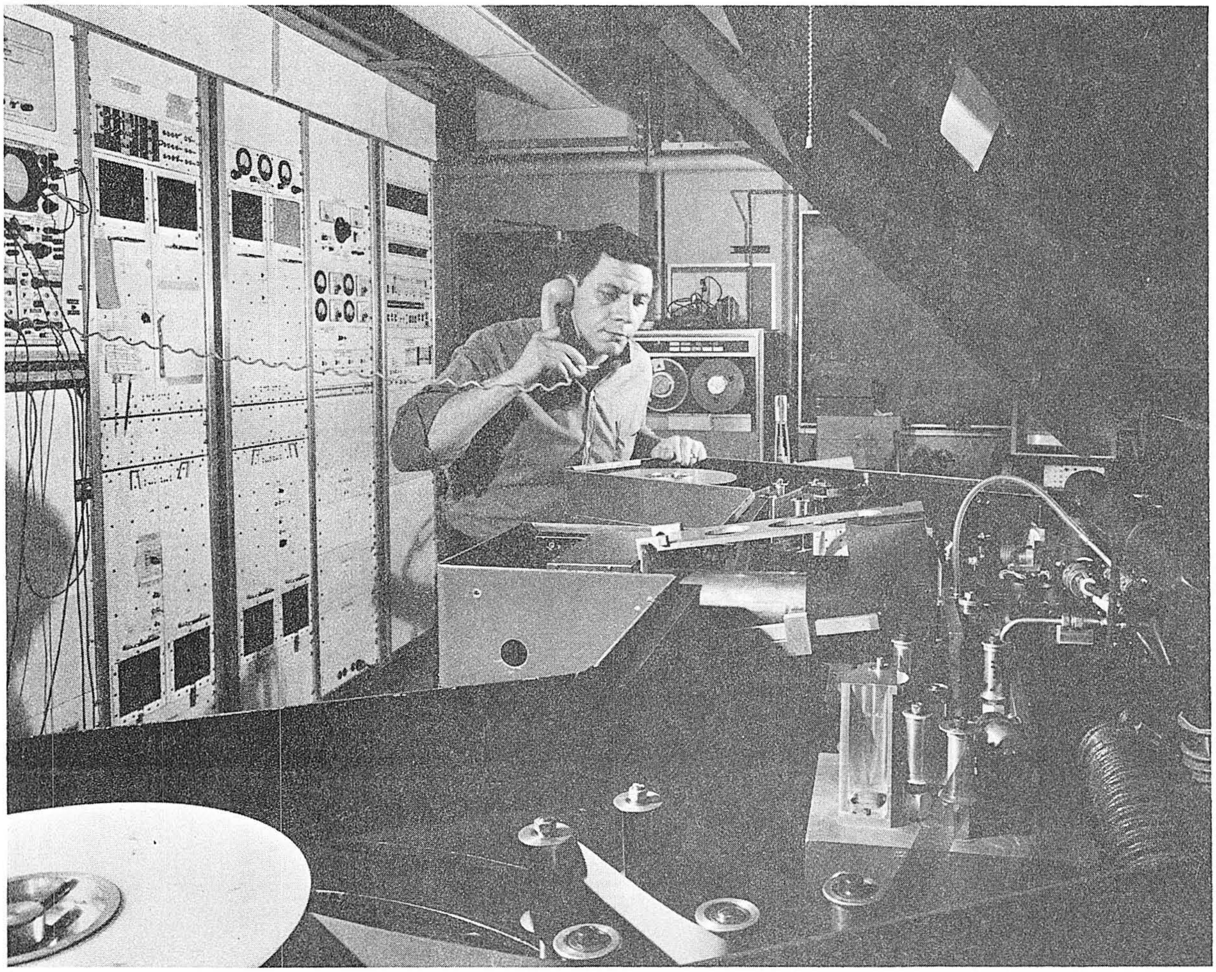
Fig. 2



FLYING SPOT DIGITIZER
OPTICAL SCHEMATIC
UCLRL BERKELEY

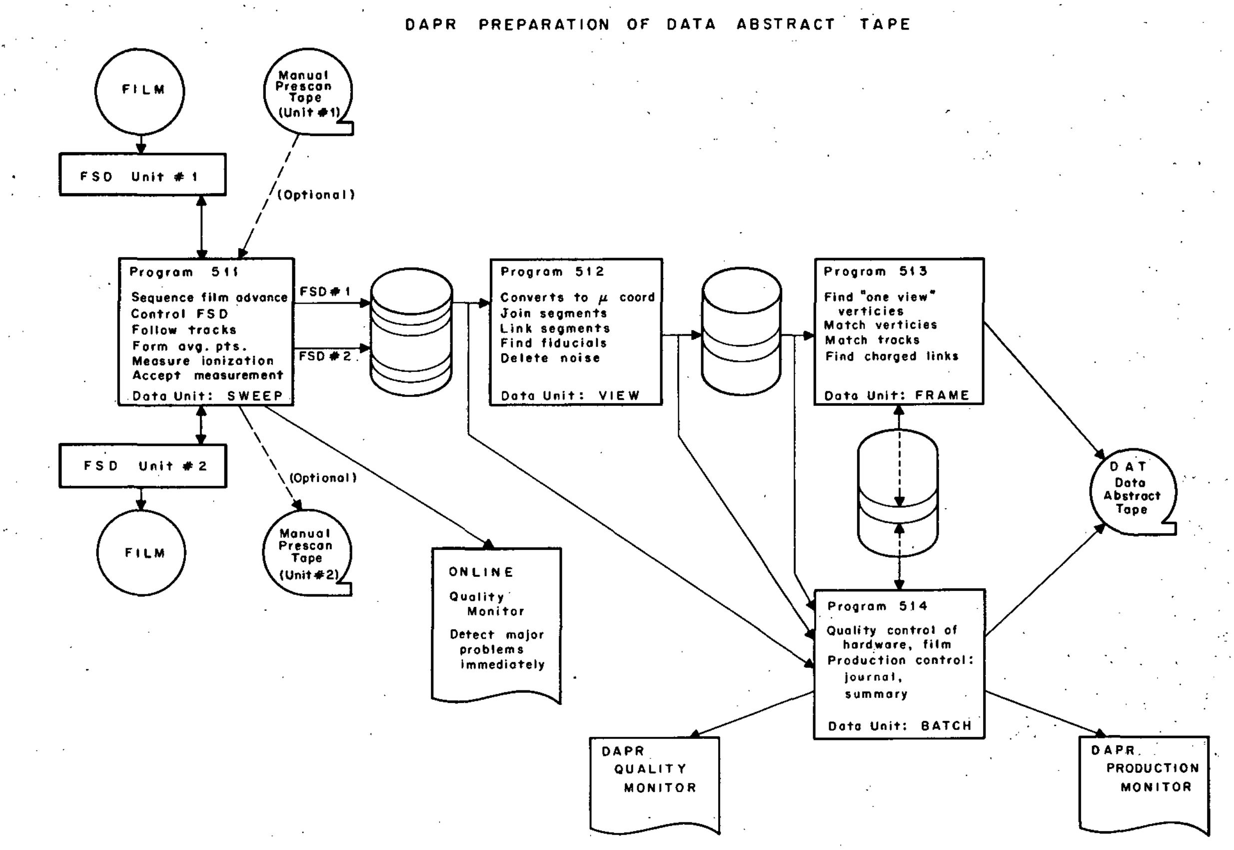
XBL 675-4022

Fig. 3



GPR-3127

Fig. 4



XBL675-3139

Fig. 5

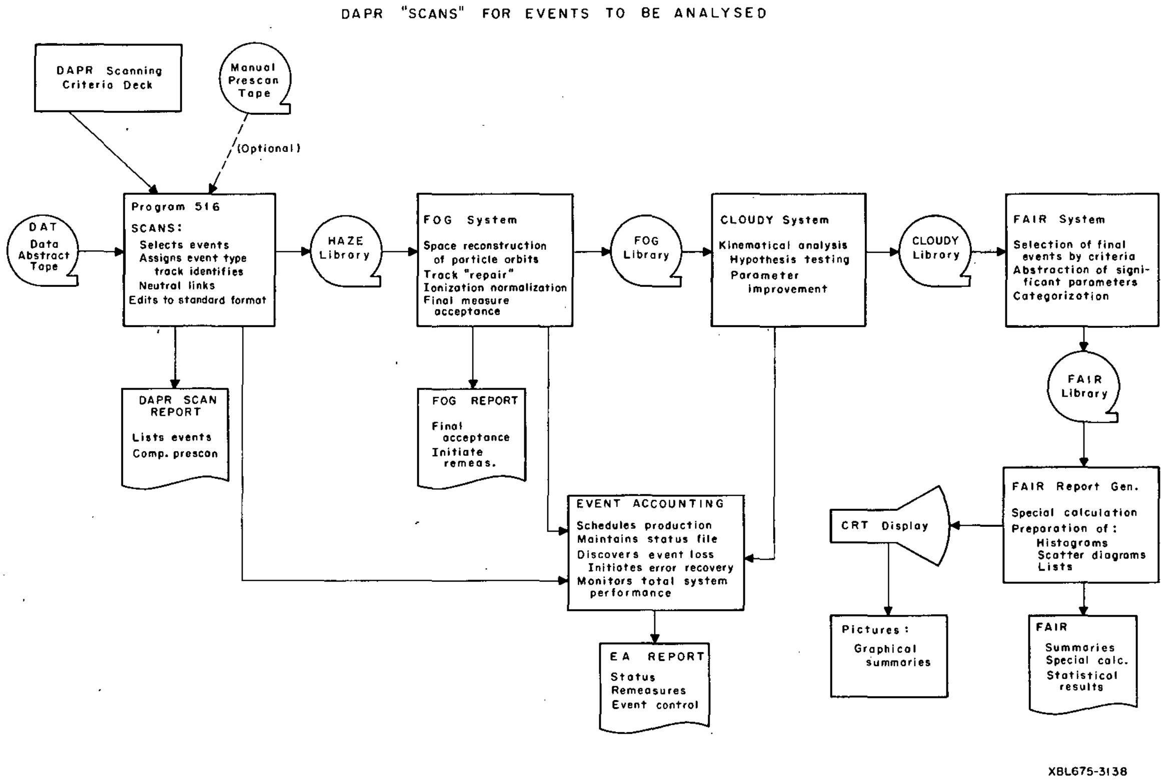
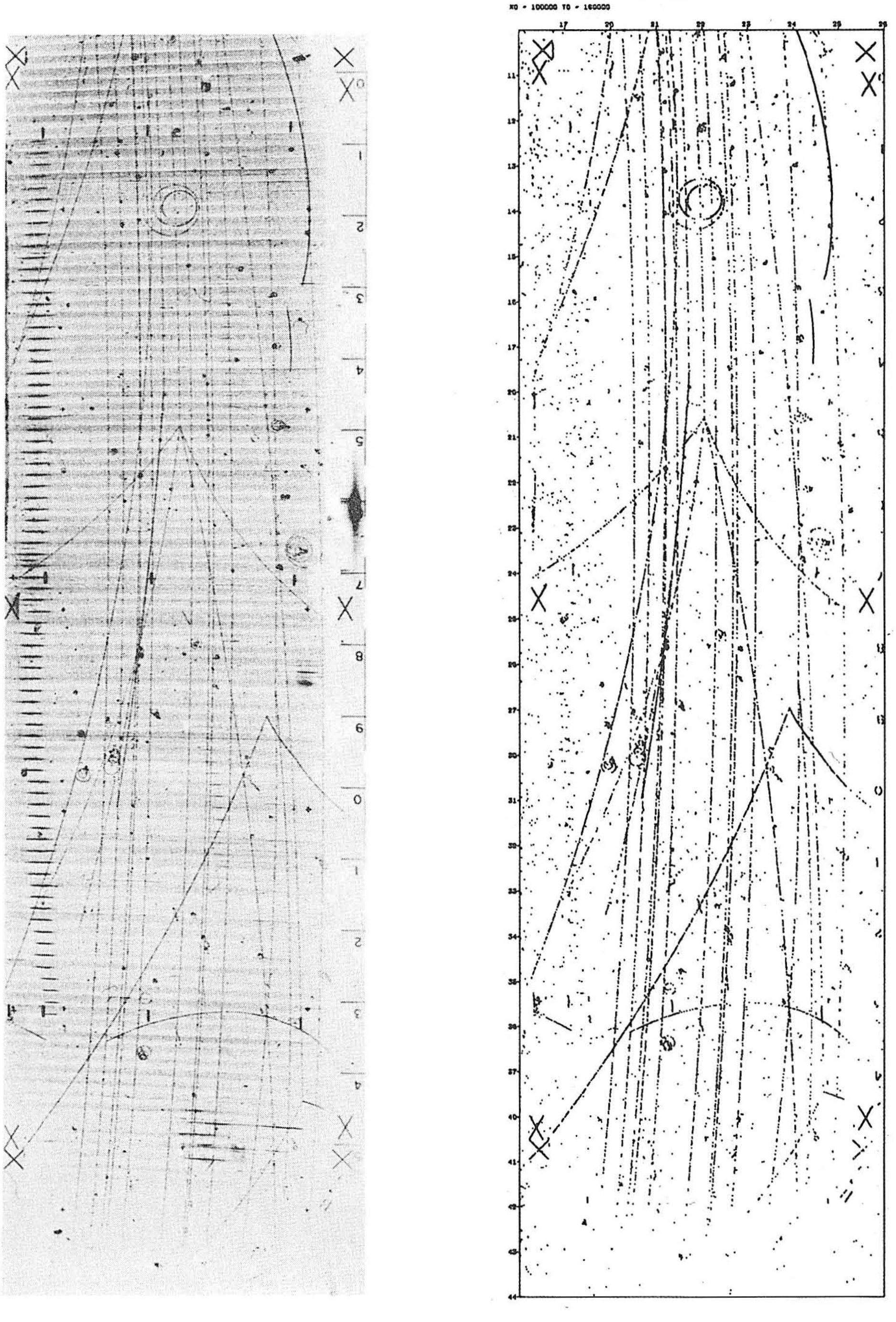
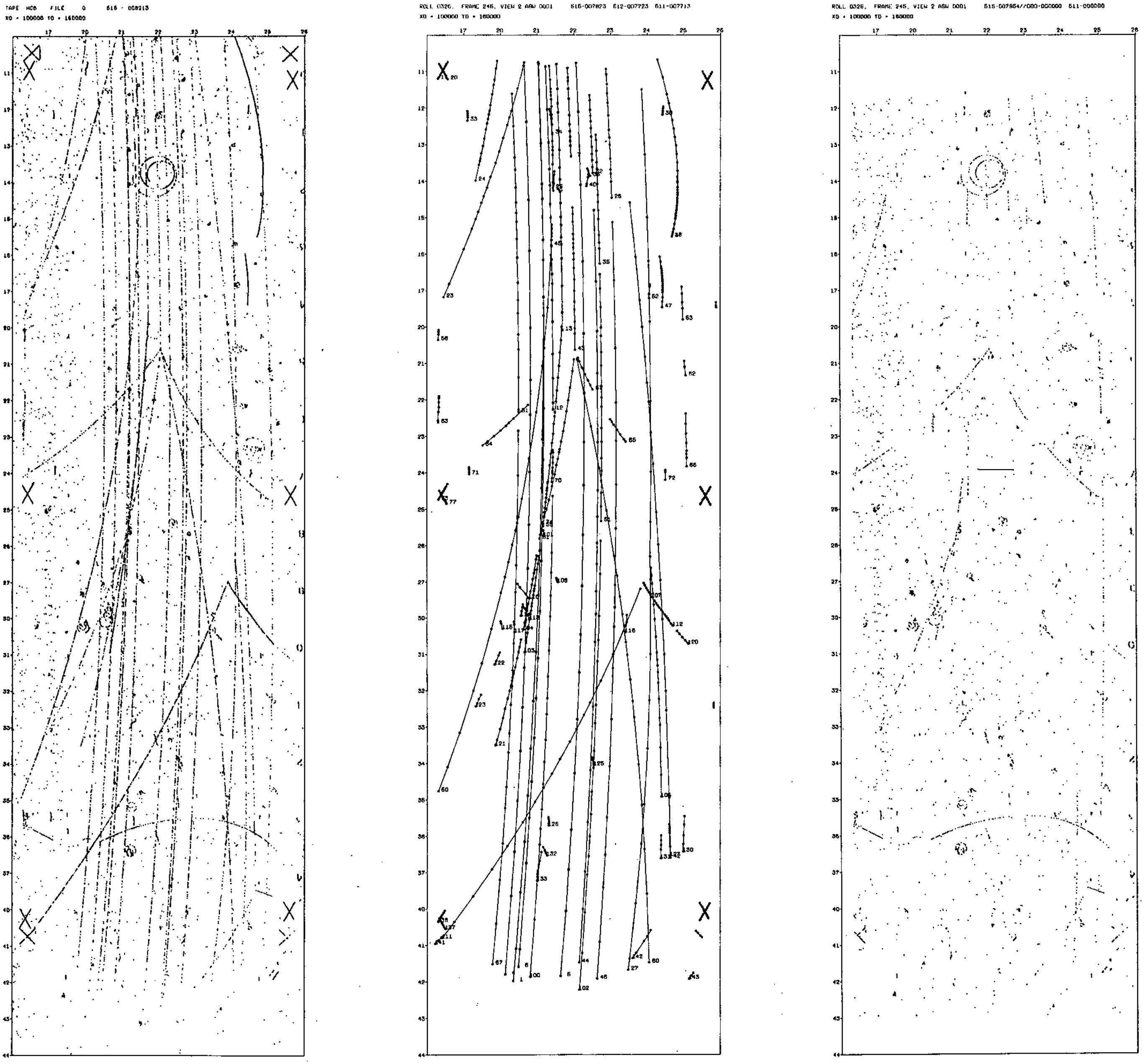

DAPR "SCANS" FOR EVENTS TO BE ANALYSED
DAPR Scanning Criteria Deck
Manual Prescan Tape
(Optional)
DAT Data Abstract Tape
Program 516
SCANS:
Selects events
Assigns event type
track identifies
Neutral links
Edits to standard format
HAZE Library
FOG System
Space reconstruction of particle orbits
Track "repair"
Ionization normalization
Final measure acceptance
FOG Library
CLOUDY System
Kinematical analysis
Hypothesis testing
Parameter improvement
CLOUDY Library
FAIR System
Selection of final events by criteria
Abstraction of significant parameters
Categorization
FAIR Library
DAPR SCAN REPORT
Lists events
Comp. prescan
FOG REPORT
Final acceptance
Initiate remeas.
EVENT ACCOUNTING
Schedules production
Maintains status file
Discovers event loss
Initiates error recovery
Monitors total system performance
EA REPORT
Status
Remeasures
Event control
CRT Display
Pictures:
Graphical summaries
FAIR Report Gen.
Special calculation
Preparation of:
Histograms
Scatter diagrams
Lists
FAIR
Summaries
Special calc.
Statistical results
XBL675-3138

Fig. 6



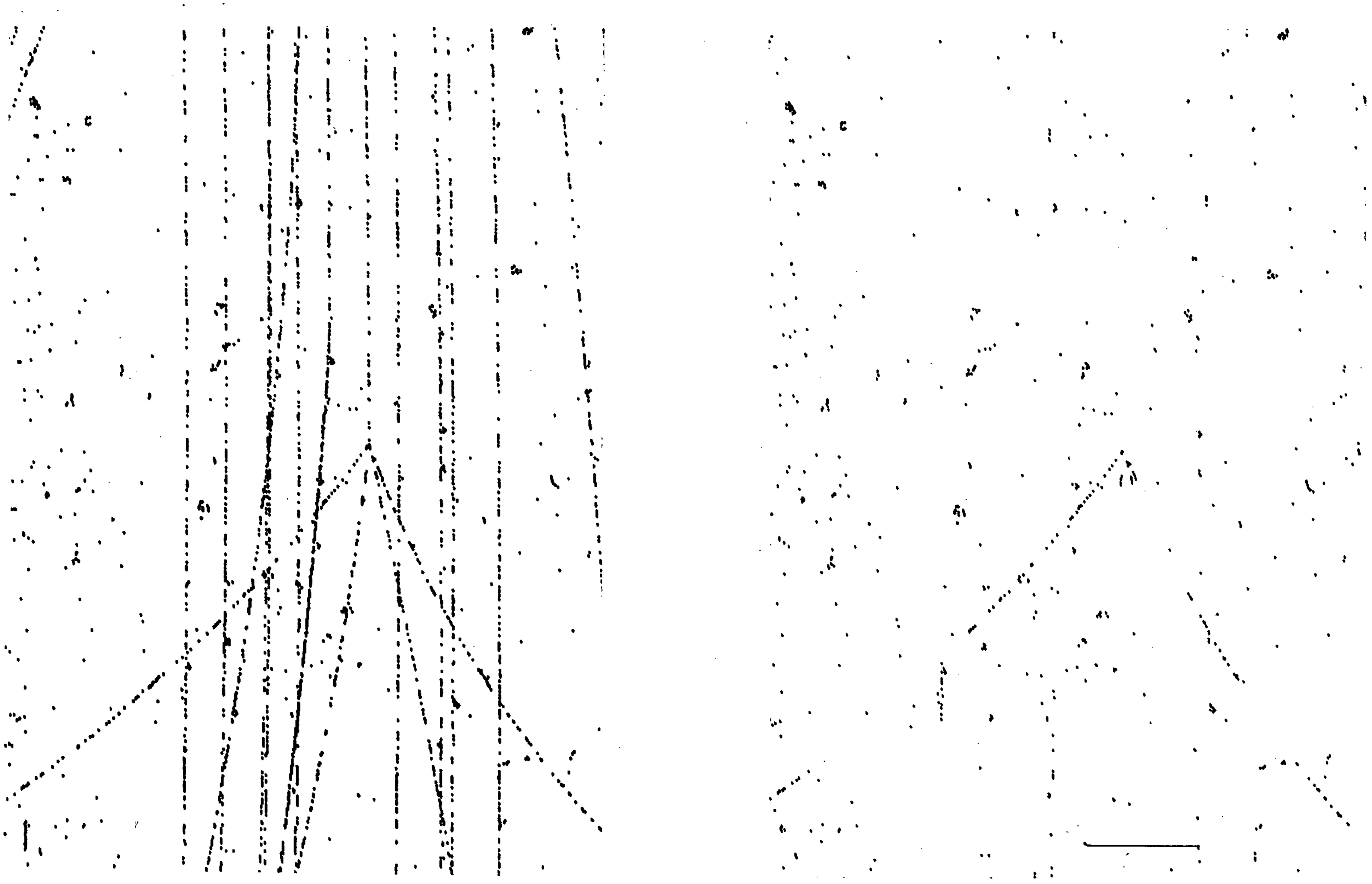
XBB 675-3004

Fig. 7



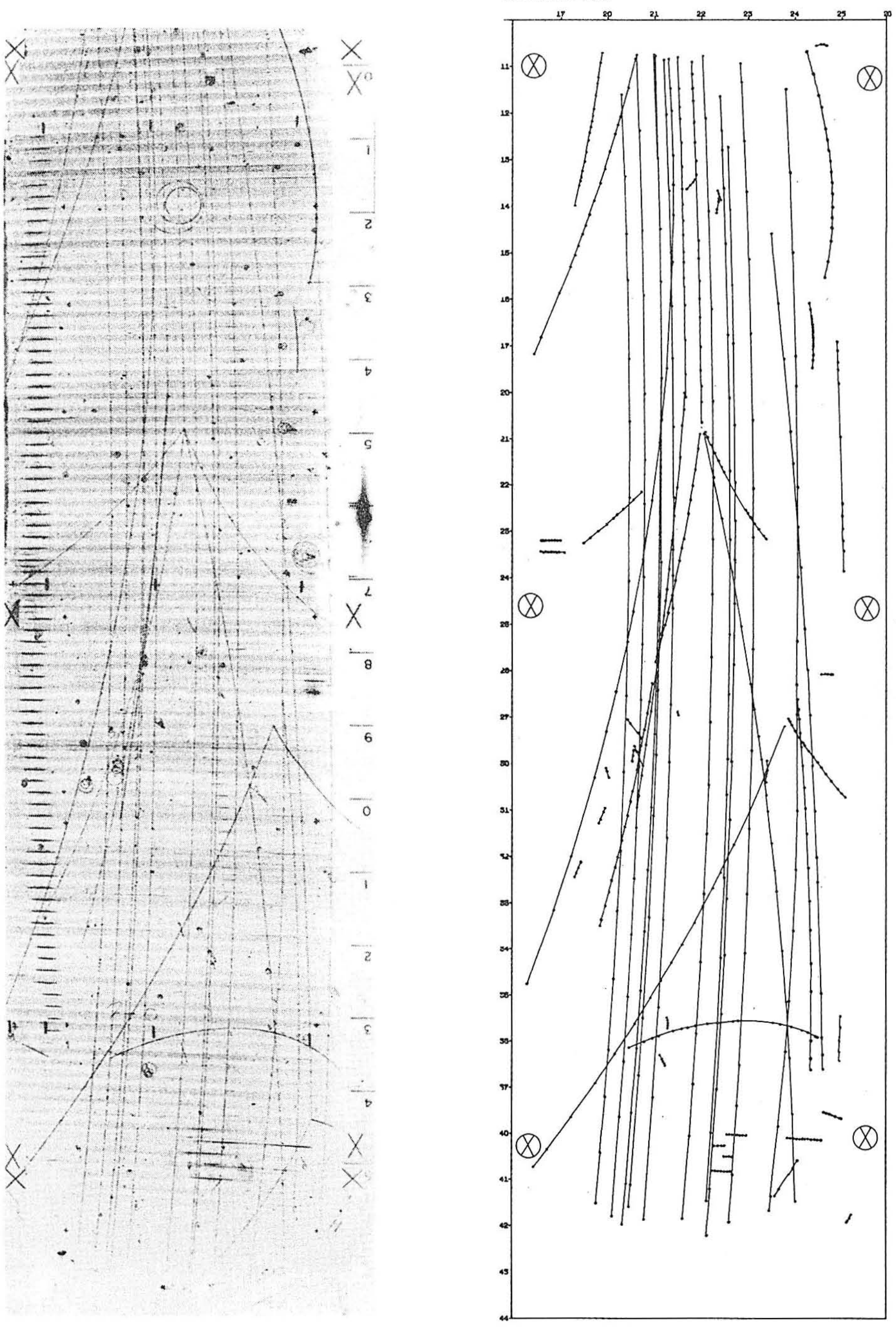
XBL 675-4020

Fig. 8



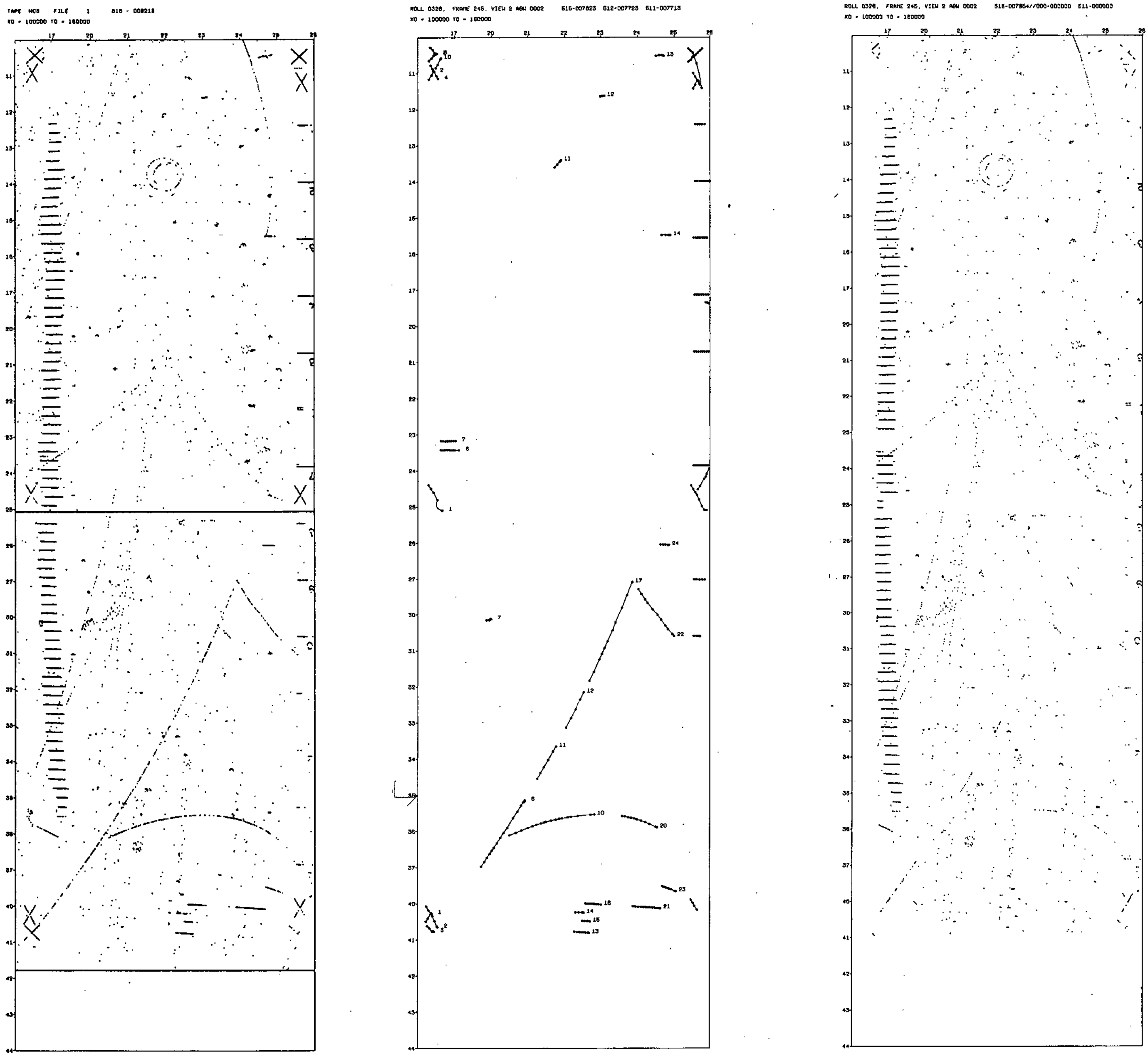
XBL 675-4021

Fig. 9



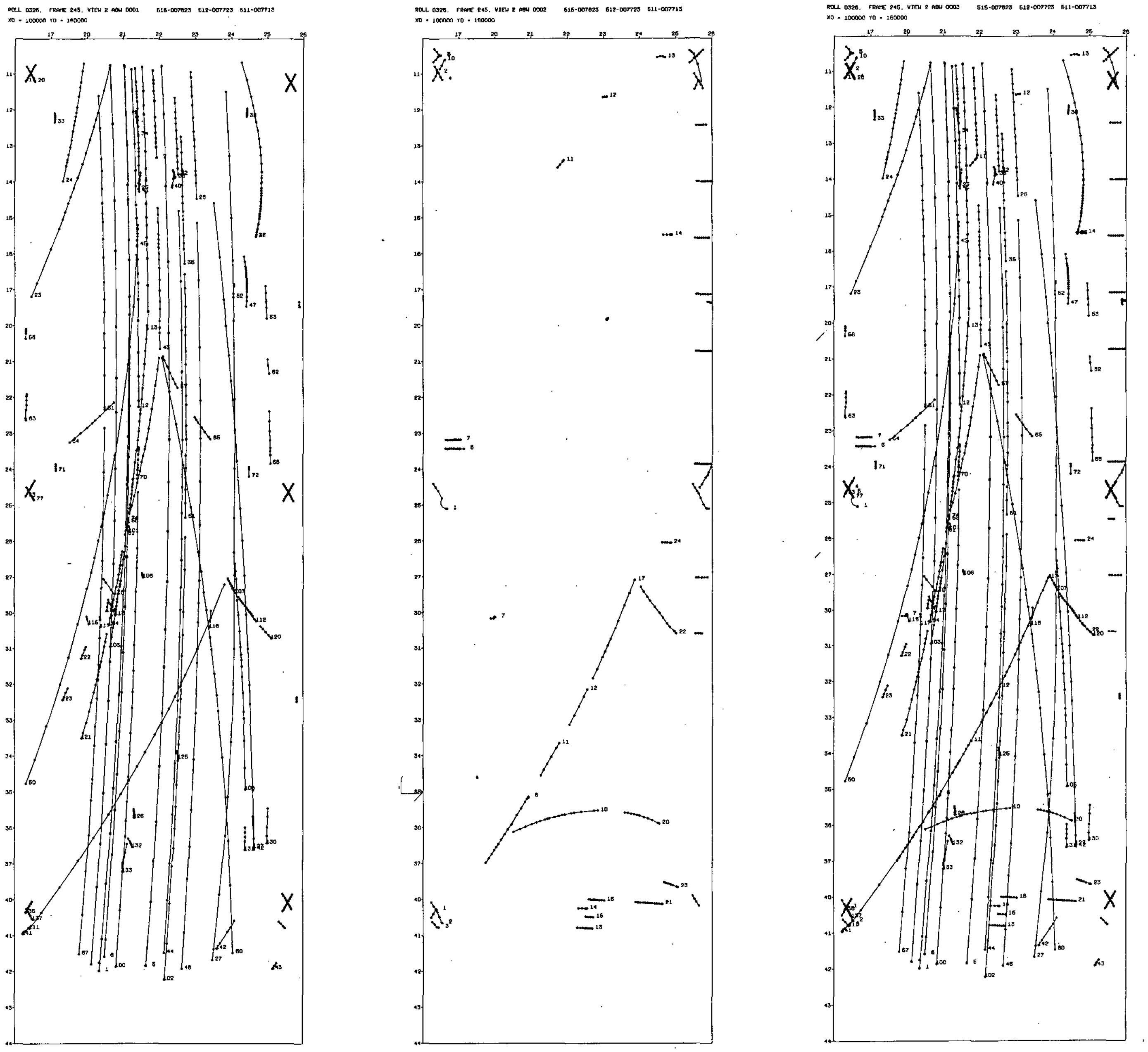
XBB 675-3006

Fig. 10



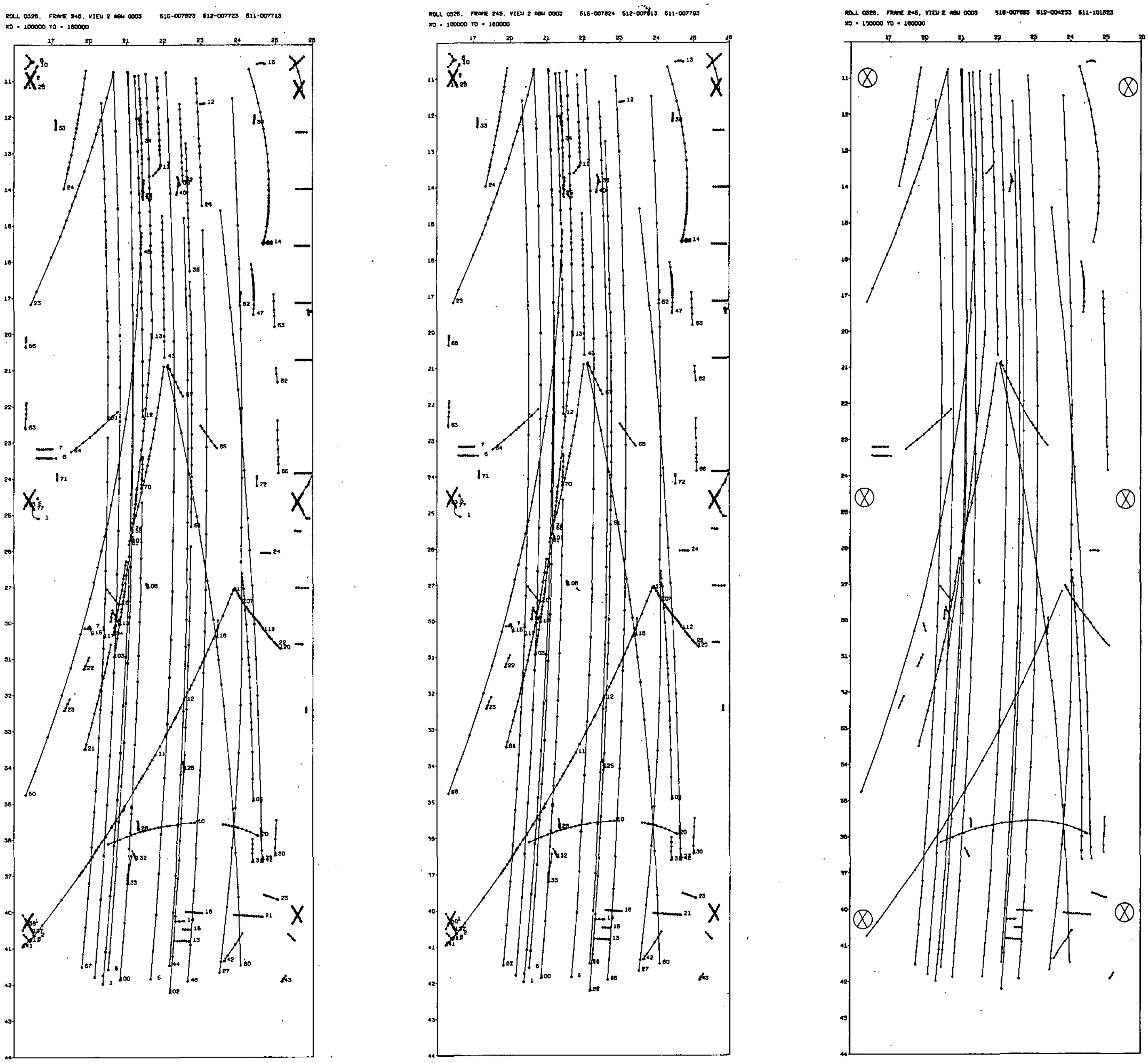
XBL 675-4023

Fig. 11



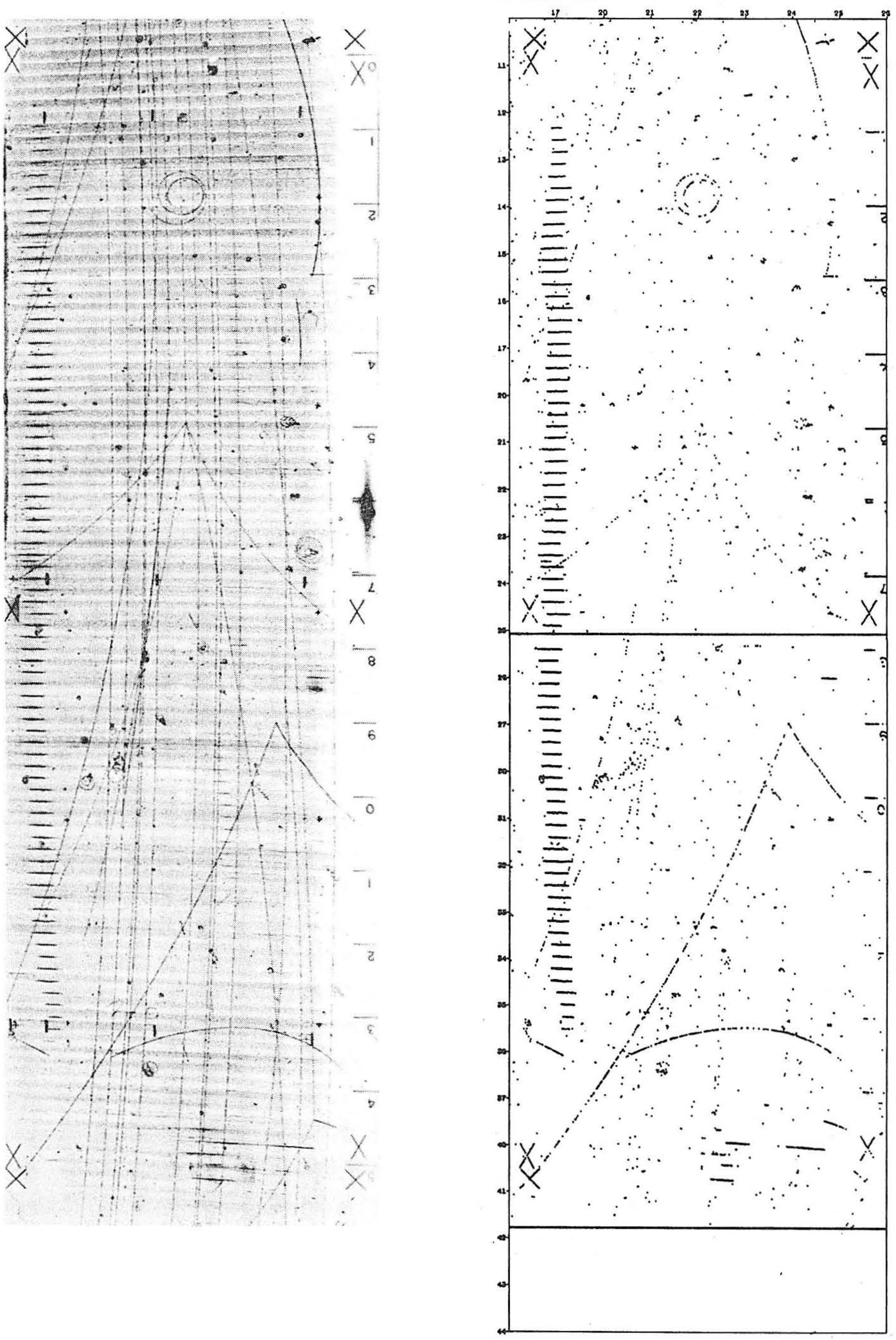
XBL 675-4024

Fig. 12



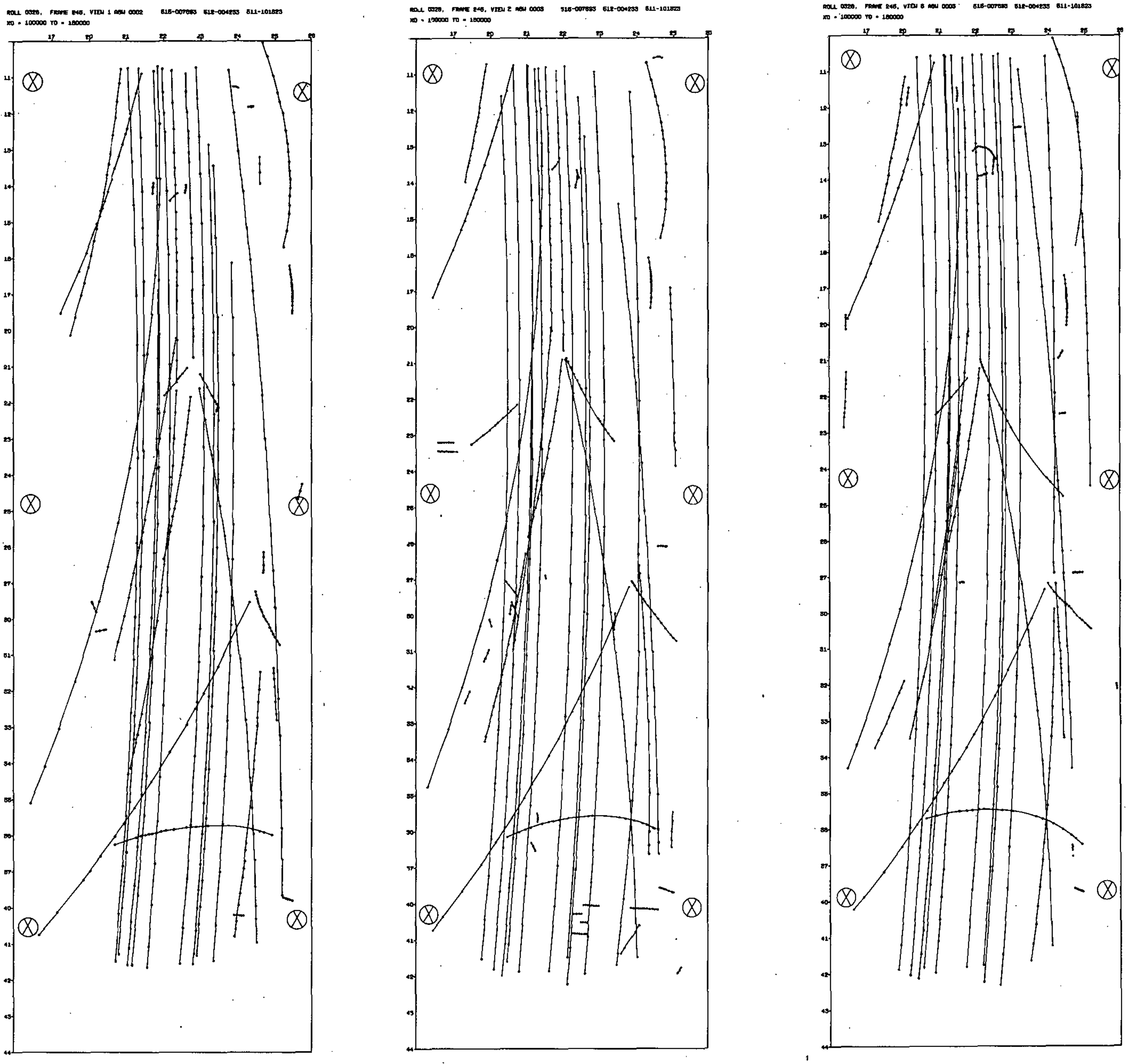
XBL 675-4025

Fig. 13



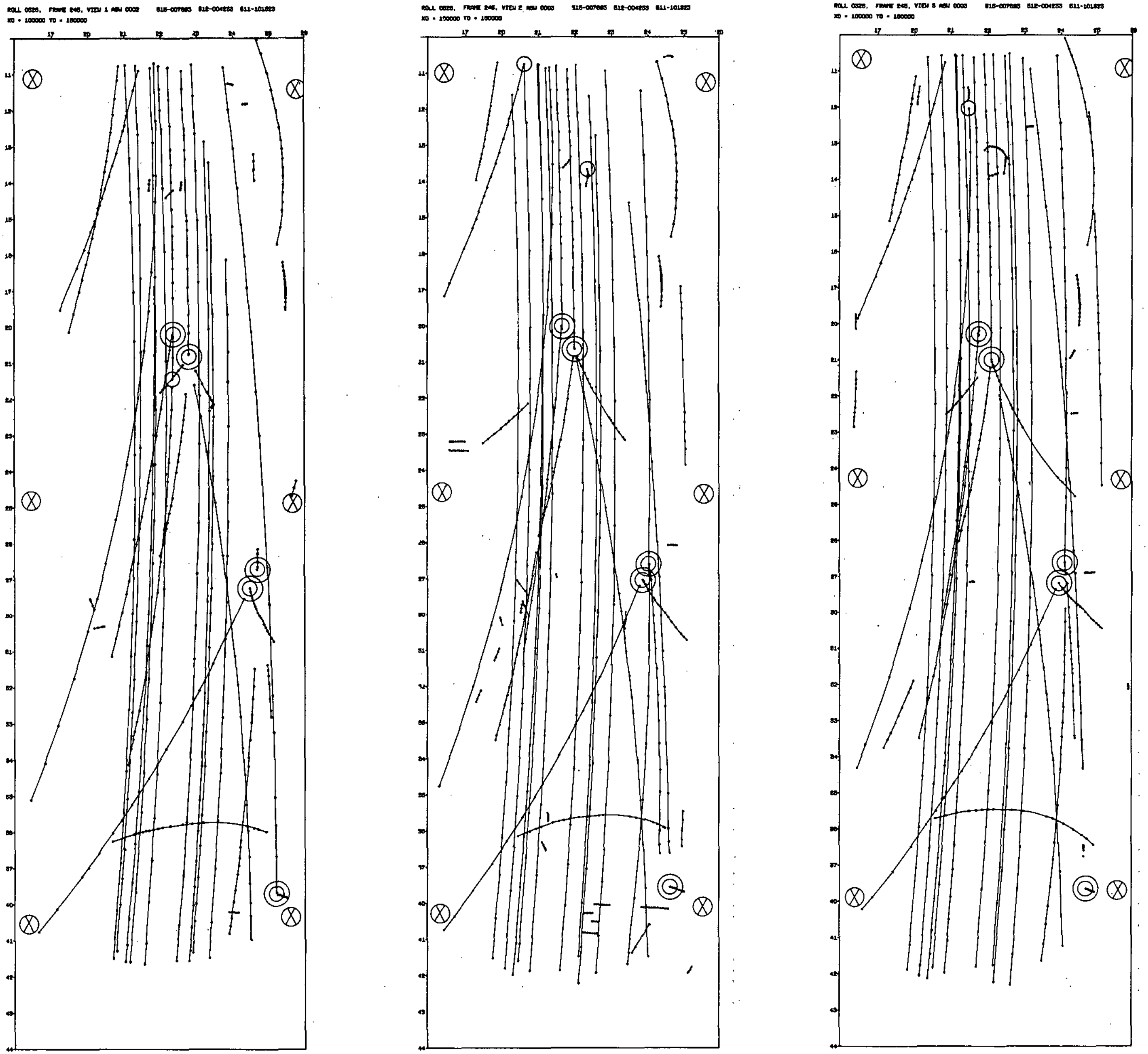
XBB 675-3005

Fig. 14



XBL 675-4029

Fig. 15



XBL 675-4030

Fig. 16