# Lawrence Berkeley National Laboratory
**Recent Work**

**Title**
FILTERuA TOPOLOGICAL PATTERN-SEPARATION COMPUTER PROGRAM

**Permalink**
https://escholarship.org/uc/item/6t5848gj

**Author**
Innes, Daphne J.

**Publication Date**
1960-11-01

# UNIVERSITY OF CALIFORNIA

## *Ernest O. Lawrence*
## *Radiation Laboratory*

## BERKELEY, CALIFORNIA

# DISCLAIMER

UCRL-9461

## UNIVERSITY OF CALIFORNIA

Lawrence Radiation Laboratory
Berkeley, California

Contract No. W-7405-eng-48

# FILTER--A TOPOLOGICAL PATTERN-SEPARATION COMPUTER PROGRAM

Daphne J. Innes

November 1960

# FILTER--A TOPOLOGICAL PATTERN-SEPARATION COMPUTER PROGRAM

Daphne J. Innes

Lawrence Radiation Laboratory
University of California
Berkeley, California

November 1960

## ABSTRACT

The advent of high-energy particle accelerators and liquid bubble chamber detectors has added the demands of high-speed data reduction to the many problems of modern nuclear physics research. For example, one 6-month experiment on the University of California 72-inch hydrogen bubble chamber yields photographic records of millions of nuclear events. This paper discusses one of the new measuring and topological identification devices which has been developed to analyze these great volumes of research data.

Dr. Bruce McCormick has proposed a scanning technique which allows rapid recognition, separation, and measurement of the photographic records of star-type nuclear events. A device known as the Spiral Reader measures background and star-type event features impartially, discriminating against nonradial patterns by the geometry of its rotating scanning element. The event measurements are separated from the background measurements by an IBM 704 computer under the direction of a program called FILTER. The separated nuclear event measurements are subsequently reconstructed in space for physics analysis.

FILTER exploits the observation that if a segment of a circular arc is rotated about a point on that arc, intercepts occur at regular intervals along a radius to the point at constant angular intervals of the rotation azimuth. The Spiral Reader, by placing the burden of event discrimination on a high-speed digital computer, minimizes the need for either special analysis equipment or a human operator to make the topological separation. Simulation, calibration, and cathode-ray-tube display routines have been included in the filter system.

This paper describes the computer program FILTER, which separates the topological star-type event configurations from undesired background features.

# FILTER--A TOPOLOGICAL PATTERN-SEPARATION
# COMPUTER PROGRAM

Daphne J. Innes

Lawrence Radiation Laboratory
University of California
Berkeley, California

November 1960


Many modern high-energy nuclear physics experiments are performed by observing the interactions of elementary particles in a liquid bubble chamber. Tracks of tiny bubbles in the liquid define the paths of ionized particles, much as do the fog droplets in the familiar Wilson cloud chamber. Stereo photographs of the chamber preserve these tracks so that the nuclear interactions may be analyzed. A typical 72-inch hydrogen bubble chamber experiment at the Lawrence Radiation Laboratory produces six stereo triad photographs every minute. As almost every triad includes a nuclear interaction, several million events are available for analysis each year.

A stereo triad of the 72-inch chamber comprises three views, each 33 by 125 mm. An average bubble image is 40 microns in diameter on the film. The track images are opaque against a clear background, the film being a negative of the dark-field-illuminated chamber. Figure 1 is a typical photograph of the chamber.

Human operators search the films for interesting nuclear interactions, designating those track combinations which are to be measured. Rectangular-coordinate point measurements along the participating event tracks define the track locations with respect to reference marks in the chamber. The measuring technician must select the event tracks from the numerous background features as he directs the operation of the currently used semi-automatic measuring projector.[1] An IBM 704 computer program, PANG, makes a spatial geometrical reconstruction of the event from the measured two-dimensional data.[2,3]

The topologically reconstructed event is kinematically fitted to physics hypotheses by a second computer program, KICK.[4] The physicist makes experimental conclusions and evaluations from these analyses.

One of the most numerous topological event types is the single-vertex interaction (Fig. 2). The currently available measuring techniques, being completely saturated by processing the rarer interactions, must neglect the important research information of these events. Dr. Bruce McCormick has proposed a scanning method especially suited for measuring such star-type events. A magnified image of the event is projected on a rotating disk. The event vertex image and the center of the rotating disk are superimposed. This disk is opaque except for one segmented radial slit. The slit is about one-third as wide as the image of a typical track projected upon it. Each of the radial segments scans an annular path ten average track widths long.
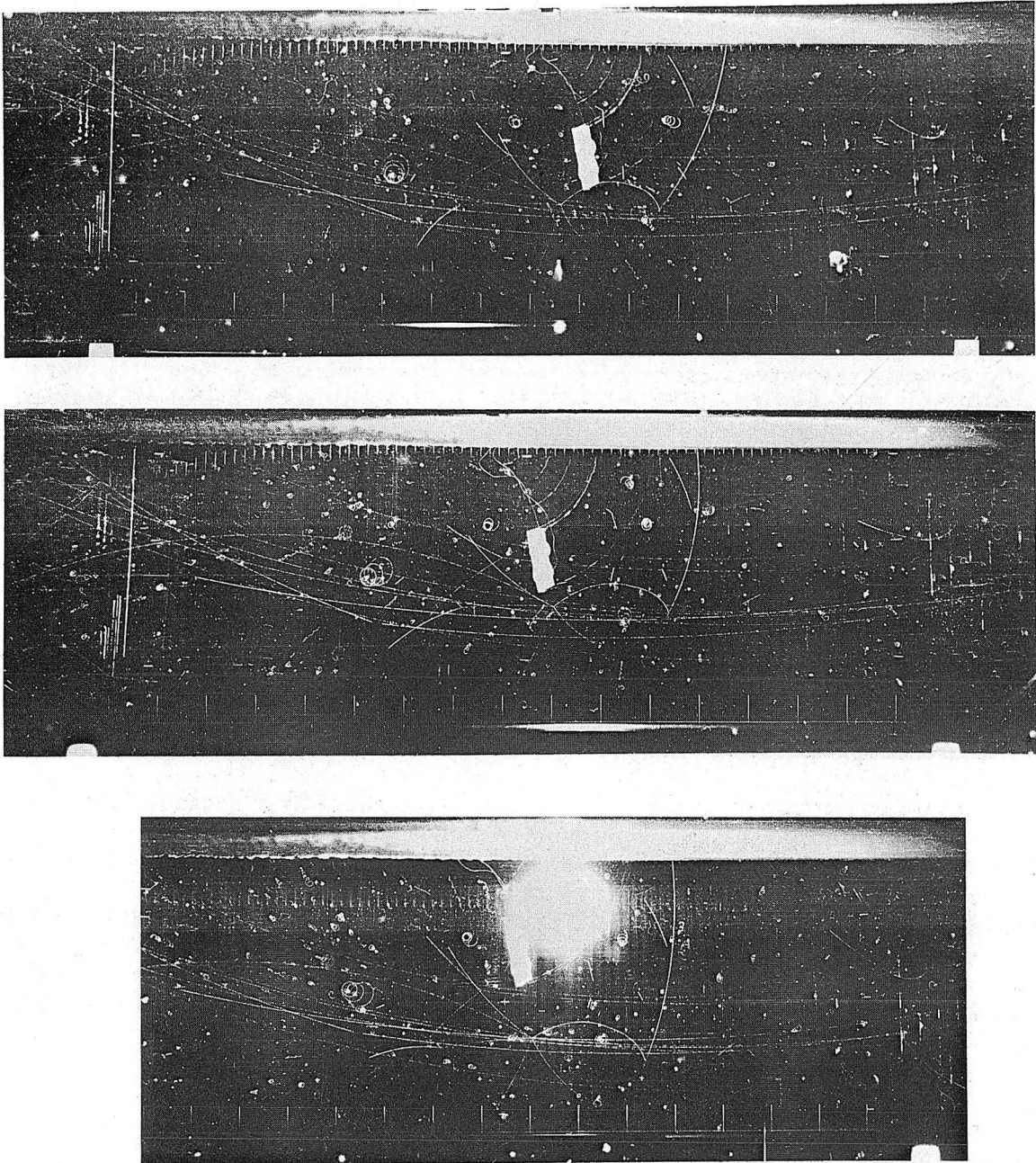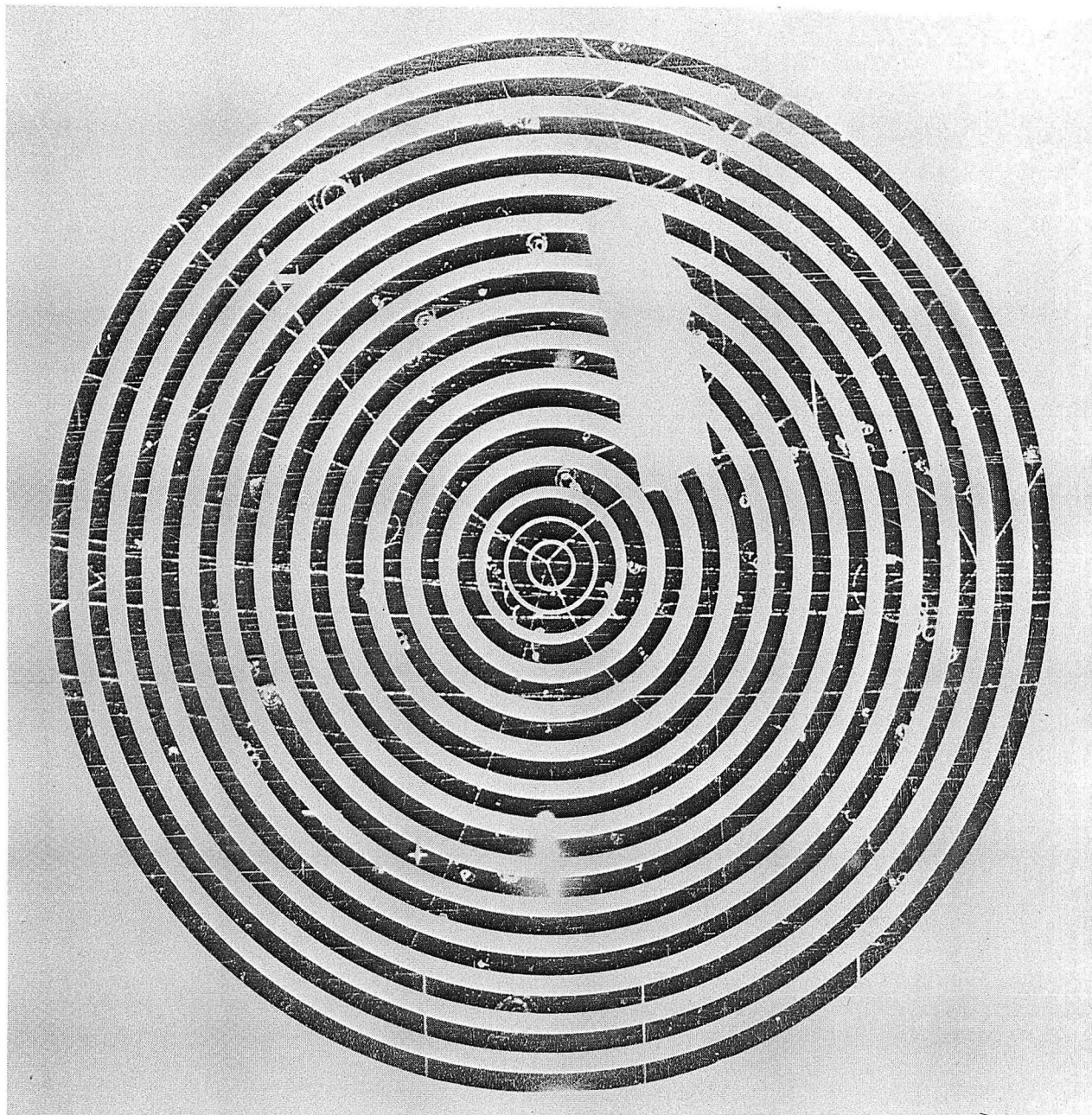
ZN-2608

Fig. 1.  A typical 72-inch hydrogen bubble chamber stereo triad.
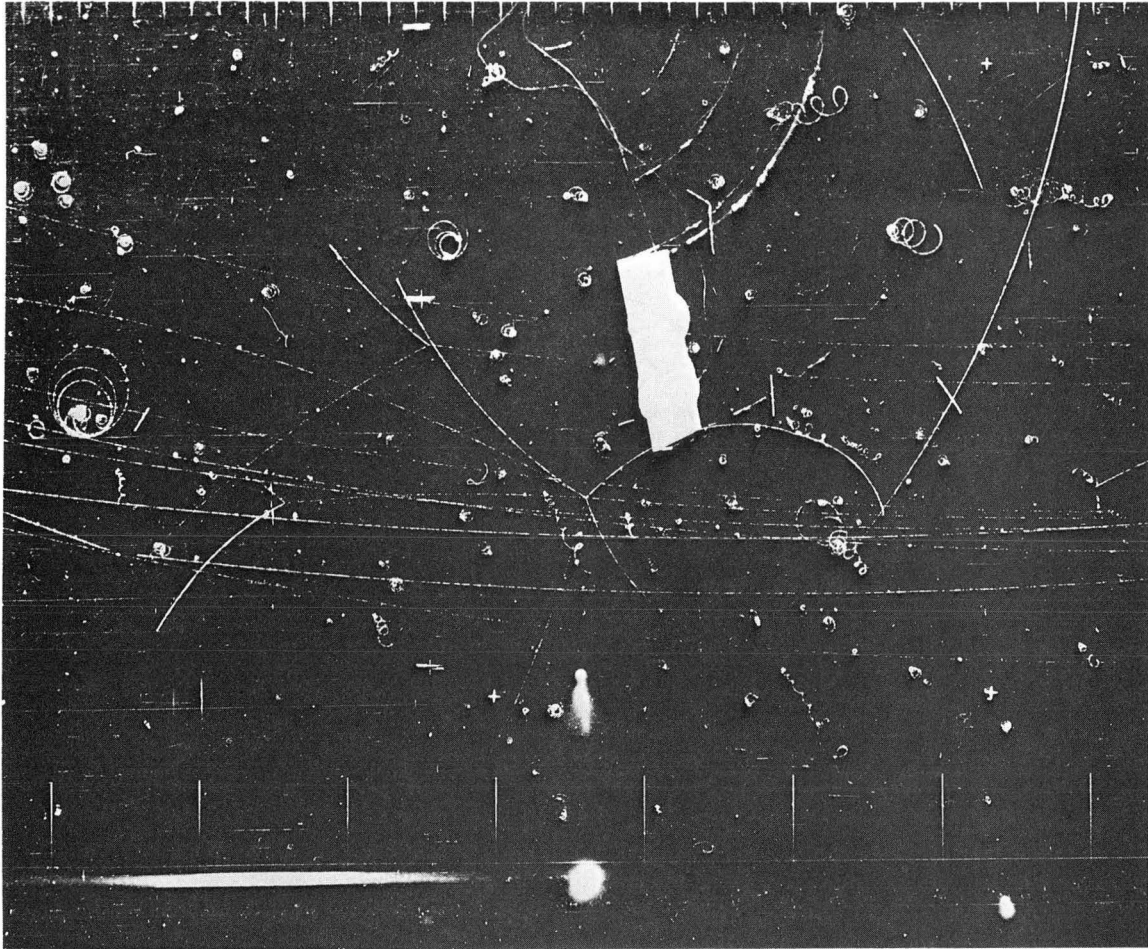
ZN-2607

Fig. 2.  An enlargement of the single-vertex event appearing in
        Fig. 1.

Fig. 3. Annuli seen by Spiral Reader scanning disc.

Figure 3 indicates the annular areas scanned around a vertex. This scanning technique discriminates geometrically against tracks which do not emanate from the event vertex. A photomultiplier collects light passing through the scanning slit. As the disk rotates over the image, the electrical signal from this photomultiplier displays the film-density variations. Radial tracks appear as pulses. A rotary analog-to-digital shaft encoder attached to the scanning disk permits azimuthal measurement of the track or feature pulses.

Track pulses which exceed a preset amplitude are squared by a discriminator. The angles at which the leading and trailing edges of these squared pulses occur are the azimuthal pulse-pair measurements of the tracks. A magnetic-core buffer memory temporarily stores these azimuths along with the relative pulse amplitude number and the slit radius. These measurements--a rectangular coordinate vertex location and event descriptive information--are recorded on IBM compatible magnetic tape. This tape is the input data for the event-separation program, FILTER.

Bubble chamber photographs are cluttered by noninteracting beam tracks, frost on the optical surfaces, gas bubbles, electron spirals, and other irrelevant features. The Spiral Reader mades only a partial attempt to discriminate against these noise features by favoring radially disposed tracks. The burden of separating the event measurements from the background information is placed upon a digital computer program rather than upon human judgment or special complicated equipment.

An effective event-separation program must detect all valid tracks in each view. It is permissible for an occasional extraneous track to be mistaken as valid in one of the three views. An erroneous labeling of a background feature can be corrected by comparison with the other two views.

In addition to removing noise measurements from the data, the event-separation program must satisfy several topological conditions:
1.  Tracks passing through obscuring features are not be be lost.
2.  Tracks that branch into two or more tracks are to be separated and identified.
3.  Tracks that cross over one another are to be individually defined.
4.  That data are missing from an occasional gap in the track image is not to stop track separation.
5.  The program is to tolerate data distortions from small misalignments of the scanning-disk center and the event-image vertex.

The radial scanning technique affords advantages other than the obvious discrimination against nonradial tracks. As the rotating scanning disk is centered over the event vertex, all tracks in the event are in principle scanned by the first three inner slits. Actually many event tracks appear in only two of the first three slits because of centering errors and track gaps. The tracks near the vertex appear as straight lines. Therefore all tracking of event tracks may be initiated by finding all pulse combinations in the first three slits which form radial lines. As the area scanned by the first three slits is small, only a few track pulses must be considered for the track-initiation part of the program. Typically, ten to twenty pulses are found in the third-slit data. Very little time is spent by the computer while making an exhaustive search of the data to form straight-line groups. Some of these

BLOCK   DIAGRAM   OF   FILTER   PROGRAM
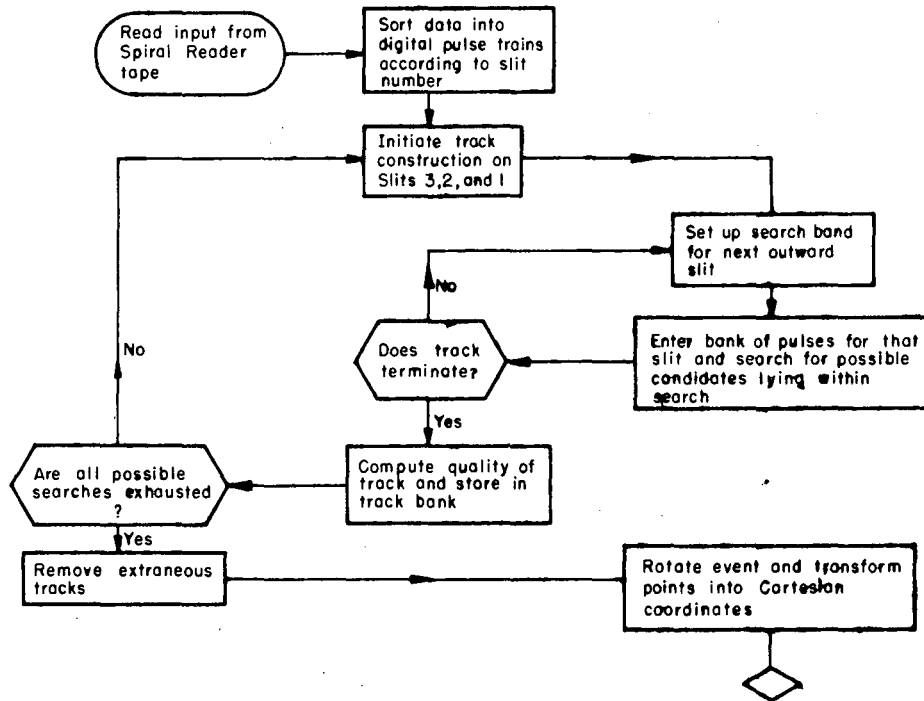


Fig.  4.   Block diagram of FILTER Program.

initial straight-line combinations do not represent tracks but are formed by dirt and isolated bubble images. Subsequent program operations remove these anomalies.

No feature that obscures less than one-fifth of the area of a scanning slit is measured. This discimination often avoids recognizing small particles of dirt and individual bubbles. Tracks which cross a scanning slit at an angle greater than 30 deg are usually not measured. Unfortunately, short-radius tracks are lost by the outer-disk slits.

The FILTER program flow diagram is shown in Fig. 4. The polar coordinate measurements made by the Spiral Reader are read from magnetic tape into the computer memory. The Gray coded azimuth numbers are converted into binary numbers. Possible malfunctions of the Spiral Reader azimuth digitizing system are detected by verifying that the successive angle measurements from each slit are in ascending order. The measurements must be arranged in sequence from the innermost to the outermost slit.

The azimuth data from each slit are separated into leading and trailing edge pulse-pair combinations, these pulses being stored in memory sections called slit banks. The relative pulse-height designating numbers are also stored in the slit banks to be used in a subsequent track ambiguity resolving operation. The pulse-pair words in the slit banks can be regarded as forming a digital pulse train in which each word defines one pulse.

After data storage and verification, the program searches Slits 1, 2, and 3 for pulses that form straight radial lines. The straight-line fit used to initiate event tracks is not sufficiently good for track interpolation beyond the third slit. We have found that, for the most curved tracks that can be seen by the Spiral Scanning Disk, a linear spiral approximates the actual curve with sufficient accuracy to predict a track pulse azimuth to within 50 microns of its true position as far as two slits beyond the last established pulse on the track.

Tracks are reconstructed from the measured data slit by slit from the vertex outward. An azimuth is predicted for the next outward slit track pulse by extrapolating a least-squares fit of the already established track pulses. A search zone 100 microns wide is established about the predicted azimuth. The slit bank is entered and a search made for track pulses lying completely or partly within the search zone. The search is effected by comparing the predicted azimuth range with the smallest-angle pulse pair in the bank. If these azimuth ranges do not overlap and the predicted azimuth is greater than this smallest-angle pulse, a comparison of the predicted azimuth with the largest pulse-pair angle is made. If again the azimuth zones do not overlap, but the predicted angle lies within the range of the smallest- and largest-angle pulse pairs, the predicted azimuth is compared to the azimuth of a pulse pair at the middle of the list of pulse pairs for that slit. If again there is no coincidence of azimuths, the number of pulse pairs to be searched has been reduced to one-half, because the predicted azimuth is either greater than or less than that of the middle pulse pair. This comparison and subsequent halving of the azimuth ranges is continued until either a pulse pair is found containing the predicted azimuth zone or the search is abandoned. The maximum number of
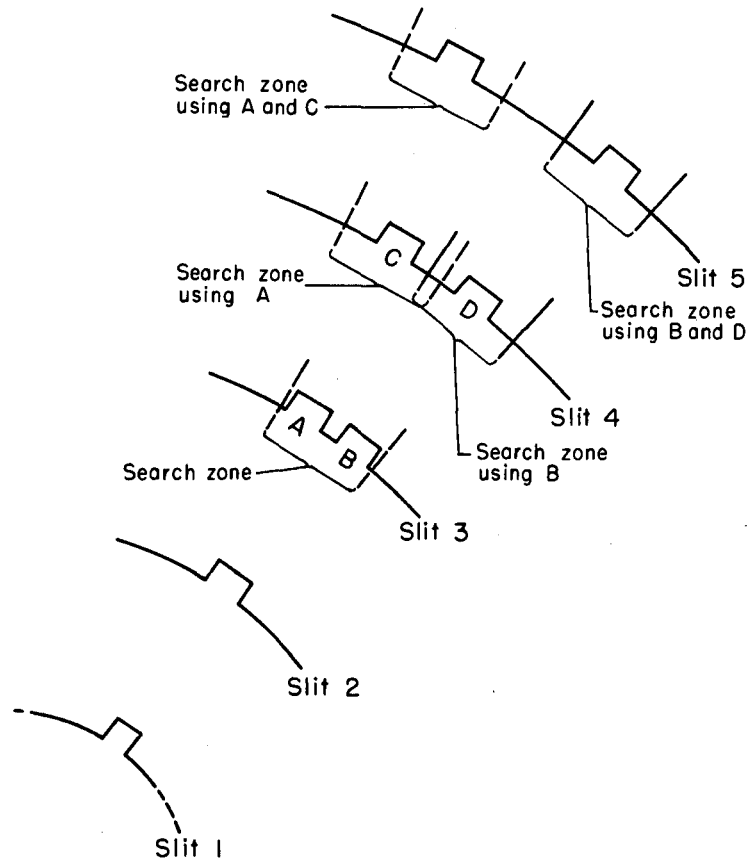
Fig. 5.  Branching track pulse trains from first five slits.

searches made in one slit bank is equal to the logarithm to the base 2 of the number of pulse-pairs in the bank. Even a slit-bank containing 100 pulse pairs, such as has been found on the 15th slit of the Spiral Reader while measuring noisy film, needs only seven or eight searches to define a track.

Since tracks may branch or cross over one another, it is probable that more than one track pulse will be discovered within a search zone. The pulse pair with the lowest azimuth is tentatively selected as the valid track pulse, the higher-azimuth pulses being stored for later investigation.

Figure 5 is a schematic of the pulse trains of the innermost five slits in the neighborhood of the branching track shown in Fig. 2. Only one pulse appears in each of the search zones for Slits 1 and 2, but there are two pulses, A and B, within the Slit 3 search zone. The search zone for the Slit 4 pulses is predicted by extrapolating the pulses on Slits 1 and 2 and 3A. Pulse C is found in the Slit 4 search zone.

The search zone for Slit 5 is now predicted by fitting the pulses 1-2-3A-4C to a linear spiral. A pulse is discovered within the Slit 5 search zone and the tracking is continued to the outermost slits.

When the track of this example terminates, the program returns to Slit 3 to test the alternative track established by the presence of pulse pair 3B. A new search zone is predicted in Slit 4 by the linear spiral fit to the points 1-2-3B. In this way, the program finds the two prongs of a branching track. This method of tracking, in addition to detecting branching tracks, can also generate false tracks – pseudo-prongs.

It is interesting to consider the effects of increasing the search zone from two track widths to four track widths (see Fig. 6). As before, pulses A and B are found in Slit 3, but when 1-2-3A is extrapolated, the new search zone on Slit 4 includes C and D. Similarly, the search zone predicted by fitting 1-2-3B also includes C and D. We find that the search zones in Slit 5 are predicted by four pseudo-prongs: 1-2-3A-4C; 1-2-3A-4D; 1-2-3B-4C; 1-2-3B-4D. Clearly, if the size of the search zone is not limited, the number of pseudo-prongs generated will rapidly become excessive. The zones must be sufficiently wide to avoid losing true tracks. A search zone 100 microns wide appears to be a good compromise for the 72-inch chamber film.

Figure 7 is employed to describe the effect of dirt or an isolated bubble feature near a true track image. The figure represents the pulse trains on the inner five slits of track 3 in the event shown in Fig. 2. With a search zone of 100 microns, pulse 3A is separated from pulse 3B and the predicted Slit 4 search zone includes no pulse. Neither will any pulses be found in the slit 5 search zone predicted from 1-2-3A; hence the search is abandoned. A noise pulse 4C is so close to 4D that the search zone predicted from 1-2-3B-4C in Slit 5 includes a pulse which actually belongs to the track 1-2-3B-4D. The two track candidates generated by the presence of the noise pulse 4C and the real pulse 4D are stored in track banks for ultimate separation by the program.
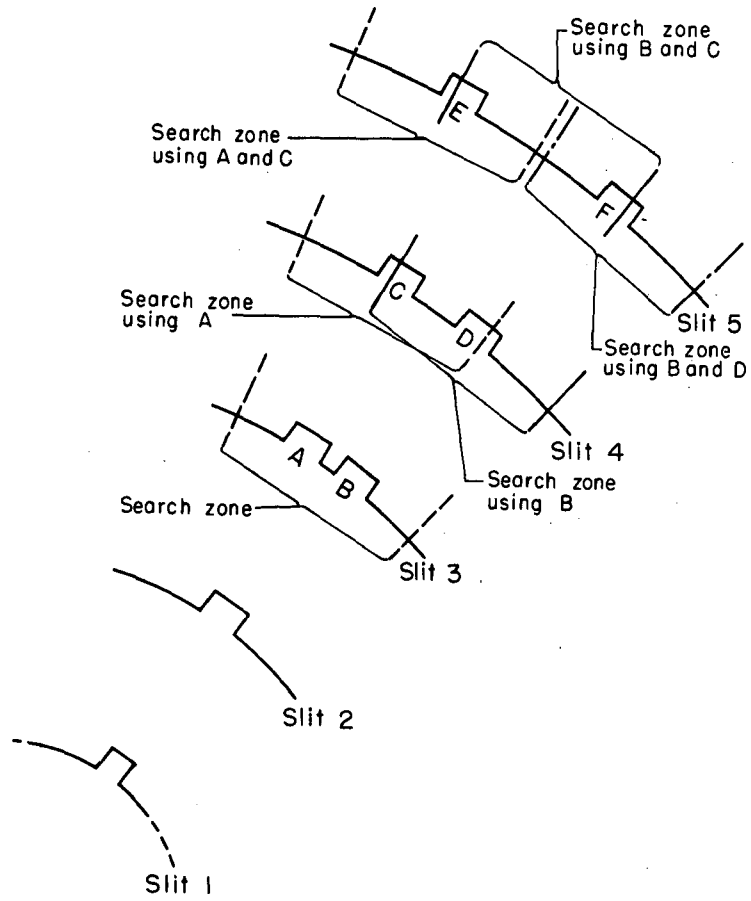
Fig. 6.   Effect of doubling search-zone width.
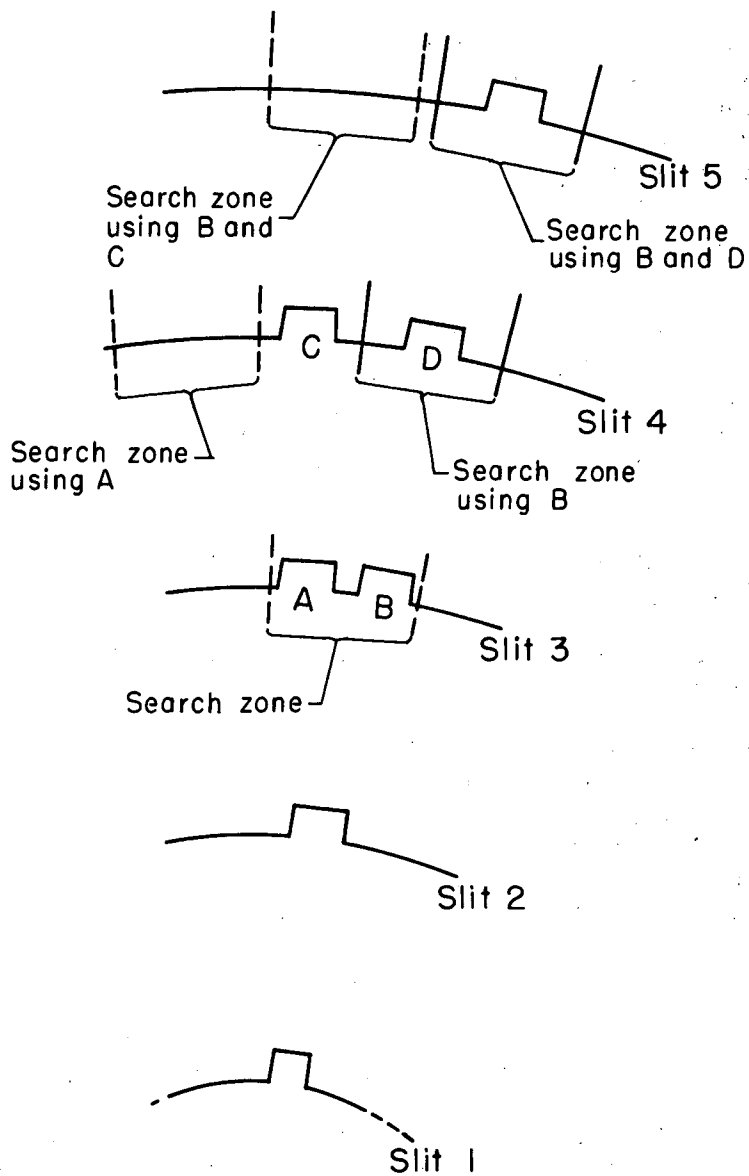
Fig. 7. Effect on tracking procedure of pulse caused by near-by dirt.

The FILTER program is capable of following tracks through large obscuring noise features. Figure 2 shows a thermocouple partially obscuring a track. The thermocouple measurements appear as very wide pulses-- these the program neglects. The extrapolation is continued, through the noise, for as many slits as is necessary. During the extrapolation through gaps or noise, a constant-angular-sector search zone is searched from slit to slit; the search zone is therefore widened as the radius increases.

Often track candidates share pulses with other track candidates. Whether this sharing is an indication that real tracks are branching or crossing or whether the tracks have been created by noise must be determined. Any track containing four or more pulses is tentatively assumed valid. A $\chi^2$ test is computed on the geometric fit of the pulse pairs to a linear spiral for each track condidate. An ionization-consistency rating is based upon the fact that the product of the pulse amplitude and pulse width should remain constant along the length of the track. A quality criterion of the track, the weighted sum of these two designators, is stored in the track bank with the pulse measurements.

It is convenient to group the track candidates into subsets. A subset contains track candidates that intersect any other track candidates in that subset. No track candidate in one subset intersects any track candidate in another subset. The track candidate with the lowest-best quality criterion in each subset is selected as a real track. Each other track candidate in the subset, in order of increasing quality function, is then compared to the real track. The assumption is made that is any track candidate has two or more consecutive pulses independent from those of any real track and if it has a quality criterion better than a predetermined limit, it too is a real track. Once a track candidate is classified as real, subsequent prong comparisons are made to it as well as to the original track.

This operation has proved adequate to insure that all real tracks are discovered. Noise features close to real pulses on two or more consecutive slits create pseudo-prongs which are difficult to distinguish from close crossing tracks. More stringent criteria would cause the loss of real tracks. After the searching and testing on an individual-view basis, a stereo reconstruction of the event is attempted. Since the three stereo views are available for comparison, it has usually been possible to resolve these ambiguities by eliminating tracks which do not appear in at least two of the three views.

For each track, measurements from two views are selected which give the best stereo reconstruction. These measurements constitute the input data for the subsequent geometric reconstruction program PANG.[2] One view of a single vertex event is typically separated from the background in 3 to 5 seconds of IBM 704 operation.

The ability of the program to separate an event from the irrelevant features on the film is demonstrated by a comparison of the plot of input data to a plot of the reconstructed event data. Figures 8 and 9 are on-line plots of the data made by an IBM 709 from the event shown in Fig. 2. These on-line plots are made by the IBM cathode-ray oscilloscope as the input data are processed. This part of the FILTER program, although most useful for the equipment and program development, will not necessarily be used for the production program.
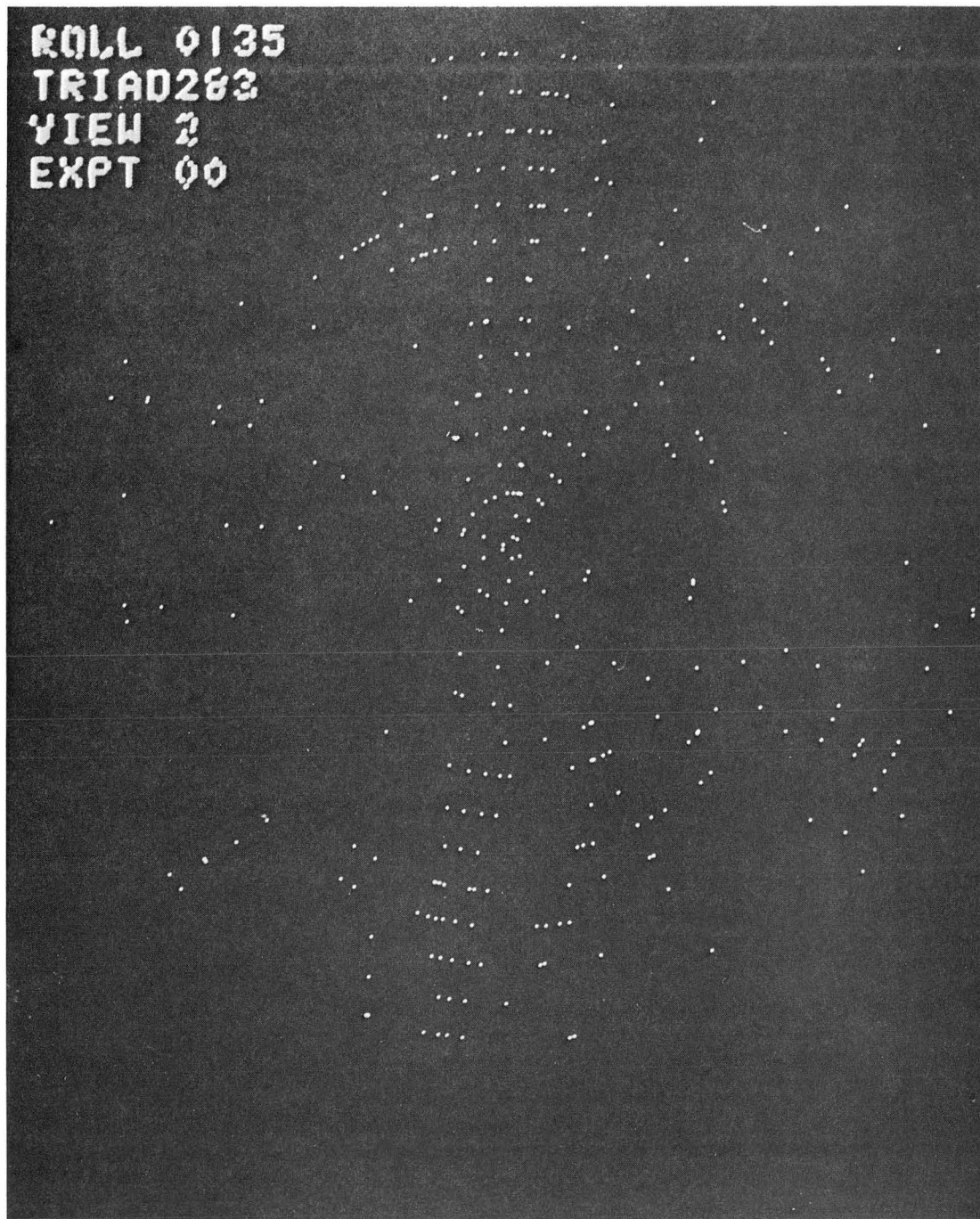
ZN-2610

Fig. 8.  Cathode-ray-tube display of unfiltered data from Fig. 2
event.
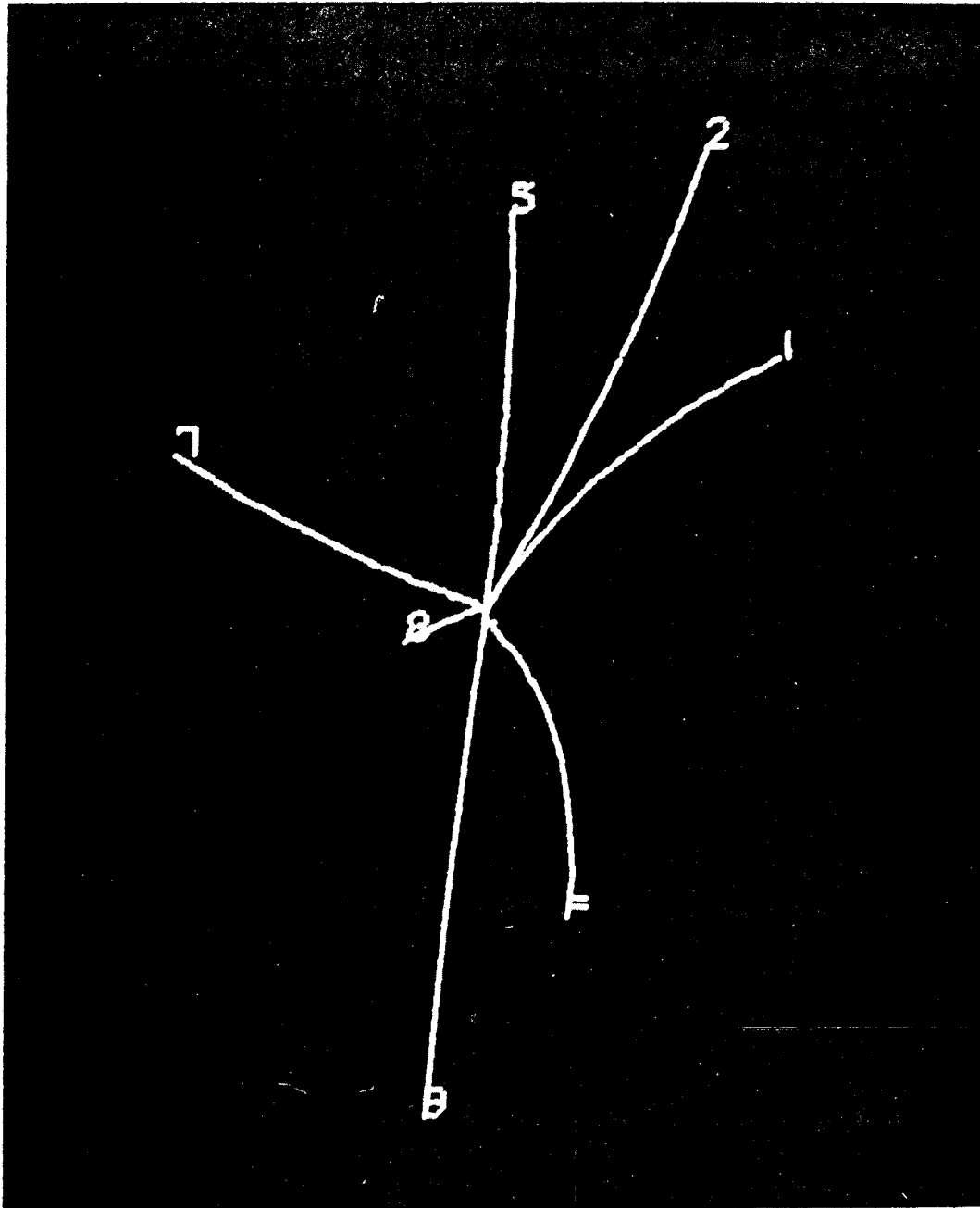
ZN-2611

Fig. 9.   Cathode-ray-tube display of reconstructed event of Fig. 2.

Now that the engineering parameters have been defined by a prototype Spiral Reader and the effectiveness of the FILTER program proved, the Lawrence Radiation Laboratory is constructing a fast scanning and measuring Spiral Reader. This measuring system is expected to process 200,000 single-vertex events each year.[5] It is believed that the tracking techniques developed for FILTER can be employed in the realization of a fully automatic bubble chamber data processing system.

Work on this project has been performed in the Lawrence Radiation Laboratory Hydrogen Bubble Chamber Physics Group under the direction of Dr. Luis W. Alvarez. The Spiral Reader project was originated by Dr. Bruce McCormick, now at the University of Illinois.

-----------------------------------

## References

1.  A. H. Rosenfeld, Digital-Computer Analysis of Data from Hydrogen
    Bubble Chambers at Berkeley, in International Conference on High
    Energy Accelerators and Instrumentation, CERN 1959.

2.  W. Humphrey, A Description of the PANG Program, Alvarez Physics
    Memo 111, 115.

3.  A. Grasselli, B. McCormick, and J. A. G. Russell, A Comprehensive
    Data-Analysis System for the 72-Inch Hydrogen Bubble Chamber,
    submitted to the Atomic Energy Commission Division of Research,
    Nov. 23, 1959.

4.  A. H. Rosenfeld, A Description of the KICK Program, UCRL-9098,
    Feb. 1960.

5.  J. A. G. Russell, The Spiral Reader--Phase Two, Engineering Note
    S. R. 19, March 1960.