A RATIONAL APPROACH TO AUTOMATED CARTOGRAPHY

by

I.D.G. MACLEOD, J.G. FRYER & D.R. SMITH

A joint publication:

Technical Report EP–RR 36 of
The Australian National University
Department of Engineering Physics — Head: Professor S. Kaneff

and

Technical Report 16 of
The Department of Minerals and Energy — Minister: The Honorable R.F.X. Connor, M.P.
Secretary: Sir Lenox Hewitt, O.B.E.

Division of National Mapping — Director: B.P. Lambert, O.B.E.

Published by the Division of National Mapping
CANBERRA, AUSTRALIA
JUNE 1973

A RATIONAL APPROACH TO AUTOMATED CARTOGRAPHY

1. INTRODUCTION

The Division of National Mapping of the Department of Minerals and Energy, and the Department of Engineering Physics in the Research School of Physical Sciences of the Australian National University (ANU), have been working jointly on two projects which are relatively inexpensive and easy to implement, yet which have important implications in the field of automated cartography. These two projects are the use of a simple digitizer to record data for a thematic mapping data bank and a scanning technique for changing map projections.

One aim in these projects has been to minimise investment in specialised hardware; some is extremely expensive, and it easily gets out of date. Inexpensive and simple hardware, when suitably programmed, can achieve valuable economies in map production. One hardware item developed was a "string" digitizer which has been used as an interactive device for recording and editing data in the thematic mapping data banks. A second item developed was a photo-scanning head which fits in the pen carriage of a conventional X-Y drum plotter. It can scan maps which can then be redrawn on a different projection and at a different scale, saving much time.

2. A THEMATIC MAPPING DATA BASE

2.1 Introduction

A thematic mapping data base has been developed to aid in the production of thematic maps which use census data collections as raw material. With a new census being taken every five years, this type of thematic map should be produced fairly quickly, so that the data portrayed are still valid when the maps are issued. Unfortunately, a shortage of cartographers, together with extensive changes in the boundaries of census collection areas between each census, lead to a lengthy map production cycle, especially with the current largely manual methods.

Census data are already available in computer readable form. Thus, if a file containing geographical boundaries of the collection areas is created, compilations of population dot distribution maps, for example, can be plotted automatically. This can then be used by a draftsman to produce his "fair drawn" copy, ready for reproduction. As the compilation phase is a major component of current methods, it is clear that significant savings in time, manpower, and cost can be achieved with a suitable automatic system. By eliminating manual collation and transcription of statistics and by automating computations of (for example) population densities, an overall improvement in accuracy could also be expected.

The census collection areas are organized in hierarchical fashion from the level of the basic collectors' districts (CDs) up to the State level. In 1971 there were about 22,000 CDs, which usually consist of about 100 families or 500 people and can vary in size from one tall building in Sydney to an area approaching one quarter of West Australia. Each collector requires a map showing the boundaries of his collection area and these maps are aggregated onto 3000 or so sheets which form the permanent geographical record of the data collection. A unique four-number code is assigned to each CD: the State; the Local Government Area (LGA); the Local Government Area Part (LGAP); and the CD - for example, 03-172-02-12. An LGAP is a region consisting of about ten to twenty adjacent CDs, which is the responsibility of an Area Manager or Supervisor during the census. LGAs with small populations usually contain only one LGAP.

Mapping takes place both before and after a census. Pre-census mapping, known as field mapping, is concerned with the design of the CDs and the production of maps for field collectors and area supervisors and for record purposes. Post-census mapping displays the information obtained by the census. Apart from the publication of a series of maps showing the data collection boundaries in urban areas immediately after the census, population distribution, population density and other statistical maps are prepared. For this purpose, base maps showing LGA and LGAP boundaries within Australia must be prepared and kept up to date. These are commonly used as base maps for many other thematic maps which the Division of National Mapping produces for client departments and for the Atlas of Australian Resources.

The boundaries of individual CDs are most readily digitized from existing field maps. A typical urban field map and the inexpensive "string" digitizer used in this project are shown in Fig. 2.1 The boundaries of each area could be digitized separately, but as most boundaries are shared by two CDs this approach doubles the amount of digitizing and leads to problems with supposedly adjacent CDs over-lapping or being separated by a thin strip of "no man's land". To avoid these problems the system developed regards the CD boundaries as a sequence of "boundary segments", separating adjacent CDs, which are digitized only once. This saving in digitization is, however, offset by the need to key in more encoding information when the boundary is broken into segments.

There are several areas of difficulty associated with the digitizing, most of which are a consequence of the sheer volume of data which has to be acquired. The 22,000 CD boundaries usually consist of four or more segments. The simplest method of encoding the boundary segments is to record a sequence of co-ordinate pairs of points along the segments, with the points assumed to be joined by straight lines. For city areas, where the boundary segments are usually composed of straight line sections, this method is very suitable. Outside city areas many segments follow natural features such as rivers or coastlines and these usually require a large number of individual boundary points for satisfactory description. Some of the requirements to be considered are:

.1 In view of the time and expense incurred in digitizing the field maps, security of the data acquired must be assured.

.2 As the volume of material to be digitized and the number of people operating digitizers grows, the book-keeping task of ensuring that each segment is digitized once and only once becomes more important.

.3 Errors in digitizing and/or encoding data must be detected and corrected as early as possible, before either the cause of later inconsistencies is hard or find or uncorrected errors are perpetuated in production maps

.4 Inaccuracies in the acquired data must be expected because of inaccuracies in:

  .1 the digitizer itself
  .2 the placement of the cursor by the operator
  .3 the field map being digitized (as a result of distortions of the paper map sheet and inaccuracies in surveys and drawing).

With scale variations in the field maps of from 1:5,000 up to 1:2,000,000, adjacent map sheets can sometimes differ in scale by a factor of ten or more e.g., a town insert in a rural area. Relative to the larger scale map, substantial errors between supposedly coincident points digitized from such maps are often encountered.

2.2 Advantages of On-Line Digitizing

In the light of the above problems, on-line digitizing offers several advantages when compared to the more usual off-line methods:

.1 Detailed instructions may be given to the operator (preferably by means of a keyboard display) so that digitizing is simplified and the amount of operator training is minimised. Encoding of data is easier and less error-prone if a question and answer type approach is used, together with free format input of numeric codes and a facility for defining abbreviations for frequently required codes.

.2 Book-keeping functions can be performed automatically so that any attempt to re-digitize an existing boundary segment (possibly digitized by another operator) can be indicated. Missing segments within an LGA or LGAP can also be detected readily with a verifying plot.

.3 The most important feature of on-line digitizing is the potential it offers for early detection of errors, while the map sheet being digitized is still in position on the digitizer, rather than much later. Apart from the usual checks of range and of alpha versus numeric data type for keyed-in codes, several types of consistency checks can be performed on both the digitized values and keyed-in data. Furthermore, a check plot out of each area digitized can be obtained for manual verification before moving to a new area.

.4 By using a co-ordinate transformation based on known control points on the map sheet, map set-up can be simplified and linear (or curvilinear) distortions in the map sheet and the digitizer can be reduced. In an on-line approach an additional known checkpoint can be used to verify the operation of the equipment and the accuracy of the transformation and to indicate if any error has been made in the set-up procedure (or perhaps in marking the control points). In this case the set up can be repeated or the map sheet put aside for checking. In an off-line environment such errors may go undiscovered until a lot of digitizing effort has been wasted.

.5 Editing of boundary segment description codes and of individual co-ordinate pairs is much simpler, particularly if current values are to be inspected or extra points are to be inserted.

Given the obvious advantages of an on-line approach, the question arises of how best to provide such a facility within a typical drawing office. On-line equipment required by the digitizer operator includes the digitizer, a keyboard display, and a plotter (preferably both hard copy and vector-graphics). All these devices can be interfaced without too much difficulty to a remote time-sharing computer via several pairs of telephone lines.

Another approach would be to use a local mini-computer to handle the required computations and file storage and there are several reasons why a local system with stand alone capability may be preferable to a remote time-sharing system:

.1 With the long "connect" times involved with digitizing, computing costs for the remote machine may be somewhat greater than those for a local mini, particularly if a continuing operation is envisaged or several digitizers are to be serviced.

.2 Draftsmen using the system as an aid to map production can see the computer as a component of this system. They will thus have a better understanding (and therefore acceptance) of the overall operation than if the computer is something mysterious at the far end of a 'phone line tended by men in white coats.

.3 The response time following operator requests will usually be shorter with a local dedicated machine, and the speed of plotting on a vector-graphics display will not be limited by the data rate over 'phone lines.

.4 Availability, access, data security and backup, recovery of freshly digitized data following a system malfunction, etc. all tend to be more easily controlled with a dedicated machine.

A realistic minicomputer system complete with digitizer, 30-inch drum plotter and vector-graphics facilities costs about $70,000. A disadvantage of this system is that having acquired the geographical data of interest, only a limited amount of data base enquiry is possible. If simultaneous access by many users is required, a large time-sharing system is necessary.

2.3 The Present System

2.3.1 Encoding of Boundaries

The Department of Engineering Physics' PDP-15 computer system includes 32K words of 18 bit core memory, two 256K work fixed-head disks, DECTAPES, analogue to digital and digital to analogue converters, a VTØ5 keyboard display, a PEP-400 video graphics terminal, and the string digitizer shown in Fig. 2.1. To facilitate the production of maps in various projections and scales, all points on boundary segments are recorded in decimal degrees of latitude and longitude. As all longitudes in Australia are greater than 100°E, longitudes are encoded with respect to 100°E rather than 0°. The 36-bit representation for floating point numbers in the PDP-15 has more than 7 digits of precision so that the range of Australian latitudes and longitudes can be represented with an accuracy of better than 10^{-5} of a degree or about 1 metre on the ground. This accuracy is more than adequate for most purposes.

Referring to Fig. 2.2, each boundary segment is given a code consisting of the CD codes on the left and right of the segment in the direction of digitizing. More than one segment can divide the same two areas so that an additional code (part N of M parts) is required to uniquely identify a boundary segment - for example, the two segments dividing CDs A1 and C3 in Fig. 2.2. Each segment is stored in the computer as a "head" section followed by a string of co-ordinate pairs. The header consists of:

.1 The codes of the two adjacent CDs
.2 The scale of field map from which digitized
.3 The number of co-ordinate pairs (must be $\geqslant 2$)
.4 A "reconciliation" code showing if the end point co-ordinates have been adjusted to match the end points of adjoining segments
.5 Part N of M parts (usually part 1 of 1).

Items (.3) to (.5) are added to the header automatically. Fig. 2.3 shows the structure of a typical boundary segment.

2.3.2 <u>Digitizing Procedure</u>

An accuracy of about 0.1 mm of the thematic map's production scale is desirable to conform with current specifications. The current production scale is usually much smaller than that of the field maps used for digitizing so that a digitizer accuracy of ± 0.5 mm is adequate. The inexpensive "string" digitizer shown in Fig. 2.1 was developed in the Department of Engineering Physics for mapping applications and is accurate to better than ± 0.5 mm over an area approximately 45 cm square. Two important features of this device are that it can be placed on top of the material to be digitzed and that it employs a hand-held "pen" which is more convenient than the usual type of cursor. See Annex A for further details.

The top 5 cm of the format area is used as a "function box" to input control commands such as "delete last point". A pressure-actuated switch in the pen tip allows co-ordinate pairs along boundary segments to be selected simply by pressing the pen tip on the map at the desired points. The keyboard display is programmed to emit a short "beep" as confirmation that each point has been accepted.

The input procedure has been designed around a single LGAP which usually covers a map area about the size of the digitizer's working area. An average LGAP contains about fifty boundaries and an operator can digitise this number efficiently, in about one hour, without accidentally omitting any boundaries or having to resort to more than one time consuming check plot (the input program allows the operator to obtain visual verification of his input at any stage via a plotter or graphic display). Sufficient core storage is available in the PDP-15 for an LGAP. A combination of disk and core storage will be used in a production system being developed by the Division of National Mapping, which will have several digitizers operating concurrently.

A generalised flow chart of the input procedure is shown in Fig. 2.4, while the function boxes employed are shown in Fig. 2.5 A useful feature of the input procedure is that abbreviations are employed for the codes of LGAP being digitized (which is given the abbreviation "A" by default) and adjacent LGAPs - See Fig. 2.2. The entry of the CD codes for a given boundary segment is then simply a matter of typing codes such as "A1,A2" for a segment within the LGAP (between CDs 1 and 2) or "B1,A1" for a segment on the LGAP boundary. The use of abbreviations speeds the entry of such codes, reduces the number of errors, and improves the legibility of the plotted output.

2.3.3 <u>Adjustment of Junction Points</u>

One of the consequences of digitizing boundary data from a series of maps at various scales is that lines which should meet at a point fail to do so. At the time of data editing, the operator can call into core memory from disk all boundaries inside a given region and compute a weighted mean of the end point co-ordinates. The end points will then join neatly when plotted out and calculations of areas, etc., can proceed. In searching for groups of end points which should be joined, the criterion adopted is that if an "error circle" drawn around an end point overlaps the error circle of another point then they are grouped - see Fig. 2.6. The diameter of the error circles is a function of the scale of maps from which the points were digitized. An exhaustive search is made of each group to see if a closed cycle of CD codes can be formed around a set of end points. Such a cycle occurs only when all relevant end points are within the group. The cycle 1-2-3-4-1 can be formed around the end point depicted in Fig. 2.6. When a cycle is found, a weighted mean of the end point co-ordinates is computed (with co-ordinates digitized at large scale being given large weights and vice versa), and this value is assigned to all end points in the cycle.

At this point a number of consistency checks are made. For example, for CDs which lie wholly within other CDs and have only one boundary segment, the end points of the segment must coincide and there must be more than three points (including the end points) in that segment. Failure to "reconcile" end points which are supposedly coincident is often a result of incorrect entry of CD codes, thus preventing a cycle from being formed - reversed CD codes (equivalent to wrong direction of digitizing along a segment) are detected and listed for correction. Figs. 2.8 and 2.9 show plots of the information digitized from the LGAP shown in Fig. 2.2, both before and after reconciliation of end points.

2.3.4 Assembly of CDs from Boundary Segments

Following successful reconciliation of the end points of a CDs' boundary segments, the individual segments are used to make up CD boundaries, LGAP boundary segments, LGAP boundaries, etc. up to the level of State boundaries. As CD boundaries are completed, CD description blocks containing such information as area in hectares, latitude and longitude extremities, and centre point of each CD are automatically generated. LGAP and LGA description blocks are formed in a similar manner. Each block has a number of words set aside to be used for specific data such as population of the area defined. Further consistency checks are performed as LGAP and LGA blocks are formed. For example, the area of an LGAP computed from its boundaries should equal the total area of its component CDs.

All items are entered in a disk data structure whose organization is basically hierarchical but which also includes index tables to facilitate access to individual items, in particular to new CD boundary segments before they have been merged into the overall structure. A master index table assigns regions of the disk to secondary index tables for CD, LGAP, LGA, and State boundaries and for LGA's. The rest of the disk is used for storage of the boundary segments themselves and other data. An example of the disk storage structure is shown in Fig. 2.7. The index for CD boundaries contains the CD codes on both sides of each boundary segment and a pointer to the disk location of the header plus co-ordinate pairs. The other indices contain similar information.

To assist in development of the system, all CD boundaries in the Northern Territory were digitized. The number of boundaries was much smaller than in States such as Victoria or New South Wales, but most of the problems likely to be met were found in the Northern Territory. Typical of these problems were

(i) CDs which are defined only vaguely in a geographical sense (e.g. all persons resident in the Andamooka opal fields)
(ii) CDs within CDs within CDs
(iii) disjoint CDs
(iv) CDs which include an offshore area of ocean, where this area is not precisely bounded.

2.4 Discussion

Once census boundaries have been digitized and collected together to form Collectors' Districts and Local Government Areas etc., thus forming a geographic data base, it is a relatively simple matter to marry this base with other computer files containing various statistics. This amalgamated file can then be used to produce thematic maps. Before examining how this can be done, it is best to study the current manual methods of producing thematic maps in order to contrast the manpower savings when using the system developed.

After the census collection has been completed the statistics are recorded via a mark sense card reader onto suitable media such as disk and magnetic tape. The statistics are aggregated, manipulated, extracted, and output to a line printer in a suitable format. In the meantime, draftsmen have to make a pencil plot of every collector's district boundary on a map base at a suitable scale and projection, using the various mechanical means currently available to them.

To produce, say, a population distribution map, the draftsman then has to embark on the laborious task of extracting from the computer print-out, the population figures for every collector's district and has to position a dot, varying in size with the population, at the appropriate place on the prepared base map.

Compare now, a computer oriented system. With simple plotting software, all boundaries can be automatically plotted at any desired scale and map projection, and in most cases dots can be plotted in their correct position, varying in size with the population. Note that a facility exists for manually changing CDs' "centre-points" to the population centres in sparse CDs (the centre-points are normally set to the mean of the CDs extremities).

In developing this system, we have not attempted to replace the draftsman with a machine, rather we have tried to release him from painstaking and monotonous tasks and to shift his work load to where it rightfully belongs - in creating the final fair drawn copy, ready for reproduction. Maps can then be produced more quickly and cheaply, and more importantly at the time when statistics are current. In many cases, the present system produces records which are primarily of historical value. When the automated system is implemented it is hoped that maps will be available to the analyst and planner shortly after the census has been taken.

2.4.2 Other Applications

The main aim of developing this system has been to solve the particular problem outlined above and those of producing other thematic maps associated with the census. During the development phase however, it has become clear that many other maps for which the Division of National Mapping is responsible can be produced in a similar manner. The Resources Atlas is a typical example. This collection of maps shows various themes such as forests, wheat, sheep and beef, and in each instance uses Local Government Areas for their basis of statistical collection. A file of LGA boundaries can easily be extracted from the disk data structure and amalgamated with these statistical data to produce map compilations ready for fair drawing.

Another project completed by the Division has also proven that the technique can be very simply adapted to suit any geographical area system which is hierarchically numbered. The project was to produce, for inclusion in the Commonwealth Year Book, new figures for the areas and length of coastline of each State. Data was to be recorded from a series of 1:250,000 scale maps covering Australia. Each map in the series has a unique number, being part of an international map numbering system. By treating each map as an LGAP, the land area as CD No. 1, the sea area as CD No. 0 and islands as a disjoint CD No. 2, the project fitted neatly into the census system. Simple algorithms to compute boundary lengths were added to the software library and production was able to commence immediately. Naturally, a permanent record of the digitized coastline and State boundaries has been kept so that it can be used for re-plotting in future. An equal latitude/longitude plotout of the digitized coastline is shown in Fig. 2.10. As an indication of the level of detail, Fig. 2.11 shows Tasmania plotted on a larger scale - over 100,000 points were recorded around the Australian coastline. Incidentally, it is interesting to note that although the new and existing areas agreed to within 0.05%, the coastline length grew from approximately 12,000 miles to 22,877 miles. The reason for this increase is that the present measurement was made at a finer level of detail than previously.

Another envisaged benefit which has not yet been fully investigated is that of interrogating the data base to provide answers to specific questions. A frequent question put to the Bureau of Census and Statistics is "How many people lie within a radius of x miles from point y", or a more recent example - "How many people reside under the flight path of the Concorde on its supersonic journey between Darwin and Sydney". Answers to this type of question are currently extracted manually by plotting the area required onto relevant maps, noting the CDs falling within it, and then looking up the computer printouts to add up the numbers.

Fairly simple area-intersection algorithms could be developed to extract this information from the geographical data base once it has been established. Use would be made of the recorded extremities of States, LGAs, LGAPs, and finally CDs to quickly eliminate areas outside the enquiry region. A more sophisticated test of inclusion would be applied to areas which had an extremity inside the enquiry region's extremities, until a list of included CDs was generated. Treatment of CDs intersected by the enquiry region would depend on the type of enquiry but the simplest technique is to include only those CDs which have more than half their area inside. Automatic extraction and aggregation of the relevant statistics is straightforward once a list of included CDs is formed.

When detailed enquiries such as "How many pre-school age children live within a given region" can be answered readily, then much greater use can be made of census data for planning purposes.

3. AUTOMATION OF MAP PROJECTION CHANGE

3.1 Introduction

A new projection of an existing small scale map is sometimes required, in which case it is usually easier to transform an existing map to the new projection than to compile afresh from raw data. The transformation distorts the original map non-linearly. The usual method of achieving the required transformation is to overlay the original with a graticule of lines of latitude and longitude and the individual "tiles" so formed are linearly distorted by a graphical process of iterative adjustment until they fit a corresponding graticule drawn in the new projection. The tile size has to be chosen so that the errors introduced by linear distortion of each tile are insignificant.

The manual technique described is slow and requires great care if the precision of the original map is to be maintained. The mathematical formulae involved in changing projection are well known. An automated system has been developed which uses a digital computer to transform detail from one projection to another by scanning the input material, storing the locations of black points, and plotting these points in the new projection. The plotted output is used to prepare a guide image suitable for scribing. A projection change task being undertaken by the Division of National Mapping was used as an example in the development of this system, which has been implemented on the Department of Engineering Physics PDP-15 computer installation.

This task involved conversion and assembly of a set of international Map of the World (IMW) series maps, on the Lambert Conformal projection at a scale of 1:1,000,000 into a four sheet map of Australia on the Simple Conic projection at a scale of 1:2,500,000. Each IMW map covers 4^{o} in latitude by 6^{o} in longitude and is photographically reduced to 1:2,500,000 before processing. The IMW map information consists of four overlays: contours; drainage net and coastlines; roads; and railways with town centres. The distortion between the two projections ranges up to 8% in longitude but less than 1% in latitude. Figure 3.1 gives an example of the two projections of the same area. In this figure the Lambert Conformal Projection, which was scanned, is shown in blue, and the Simple Conic Projection output is shown in red.

3.2 Scanning

A picture scanner has been constructed by replacing the pen assembly of a CALCOMP drum plotter with a scanning head - see Annex B and Figure 3.2. The plotter drum is only 25 cm wide and the IMW overlays have to be processed in two $4^{o} \times 3^{o}$ sections of about 19 cm by 13 cm. One of these sections is aligned with respect to previously plotted positioning crosses and is fastened to the plotter drum. The scanning head is moved across the overlay section using the Lambert Conformal projection equations to drive the head along meridians of longitude. The grey levels of points along each scan line are examined and the locations of black points are stored on disk in terms of the scan line number and the position in that line. The spacing between points along the scan line is 0.125 mm. The spacing between successive scan lines can be varied according to the nature of the overlay being scanned - a spacing of 0.25 mm may be required for the contour overlays but 0.35 mm is usually adequate for the road and railway overlays. With a scan line spacing of 0.25 mm, a 4^{o} x 3^{o} section takes about fifty minutes to scan including the time required to write the locations of the black points on disk.

3.3 Plotting

The scanning head is replaced by a 0.1 mm pen, and a strip of acetate film is fastened to the plotter drum. The simplest method of plotting is to read the locations of black points from the disk, in the order stored during scanning, to convert the line and point numbers into latitudes and longitudes and thence via the Simple Conic equations into plotter co-ordinates. The points are plotted out individually. This simple method of plotting takes up to $1\frac{1}{4}$ hours for a 4^{o} x 3^{o} section and the continual pen lowering and raising leads to an excessive ink flow.

These problems have been overcome by writing routines which attempt to plot the points along lines in the original overlay, leaving the pen down between adjacent points. These routines plot connected sets of points by searching for adjacent black points as each point is plotted. There is space for only about ninety scan lines, each 4^{o} deep, in the core memory of the PDP-15 computer, even when the scanned elements are represented by single bits. Plotting is therefore performed in several strips across the map.

A matrix of 1-bit elements is mapped onto the 18-bit computer words with those elements corresponding to black points set to 1. This matrix includes a 2 element wide border of zeros to simplify subsequent processing. Starting at one corner of the matrix a search is made in the form of an expanding square annulus until a black point is encountered. Any black neighbours of the point are noted and the bits corresponding to the starting point and its neighbours are set to zero before a search is made to see which of the neighbours has the most new black neighbours. This point is recorded for use as the starting point in the next cycle of the process and the current starting point and its recorded black neighbours are then plotted with the next starting point being plotted last. When two of the current neighbours are not adjacent, lifting of the pen is avoided by moving between them via the starting point.

The plotting of connected points is terminated and the pen lifted when none of the current points being plotted has any black neighbours. Using this technique, all black points in a connected set can usually be plotted without lifting the pen. Exceptions occur when the line in the original overlay is thicker than 3 points wide. This causes odd points on the edge of the line to be plotted individually at the end of the continuous string. When all strips have been plotted, the output phase is completed by superimposing a graticule on the plotted points, to assist in relating overlays to each other.

A dramatic improvement in output time results from plotting connected sets of points - the time for plotting a typical 4^{o} x 3^{o} overlay is reduced to about twenty minutes. Including set up time and data entry, the overall time required to process a 4^{o} x 3^{o} overlay section is about $1\frac{1}{2}$ hours, which compares favourably with the time required for manual techniques, especially when the limited amount of manual intervention is considered. Additional photo-detectors could be added to the scanning head so that several adjacent lines could be scanned with each traverse of the overlay. The scanning time will then be reduced from fifty minutes to perhaps only ten minutes, thus nearly doubling the throughput.

4. CONCLUSION

These two projects demonstrate that by an appropriate division of labour between man and computer, substantial improvements over conventional manual methods can be achieved. The investment in computer hardware is not great and the total program developed time for both projects has so far been less than 18 man-months.

The computer-oriented system developed recognise that the computer is better at repetitive operations, data storage, retrieval and calculation, whilst the draftsman is superior in aesthetics and judgment. The automated system can easily be introduced into a drawing office as they parallel the existing manual methods.

As a result of the successful experiments into automated cartography techniques at the Department of Engineering Physics, the Division of National Mapping is buying a mini-computer system to establish a thematic mapping data bank. Delivery is expected in mid 1974.

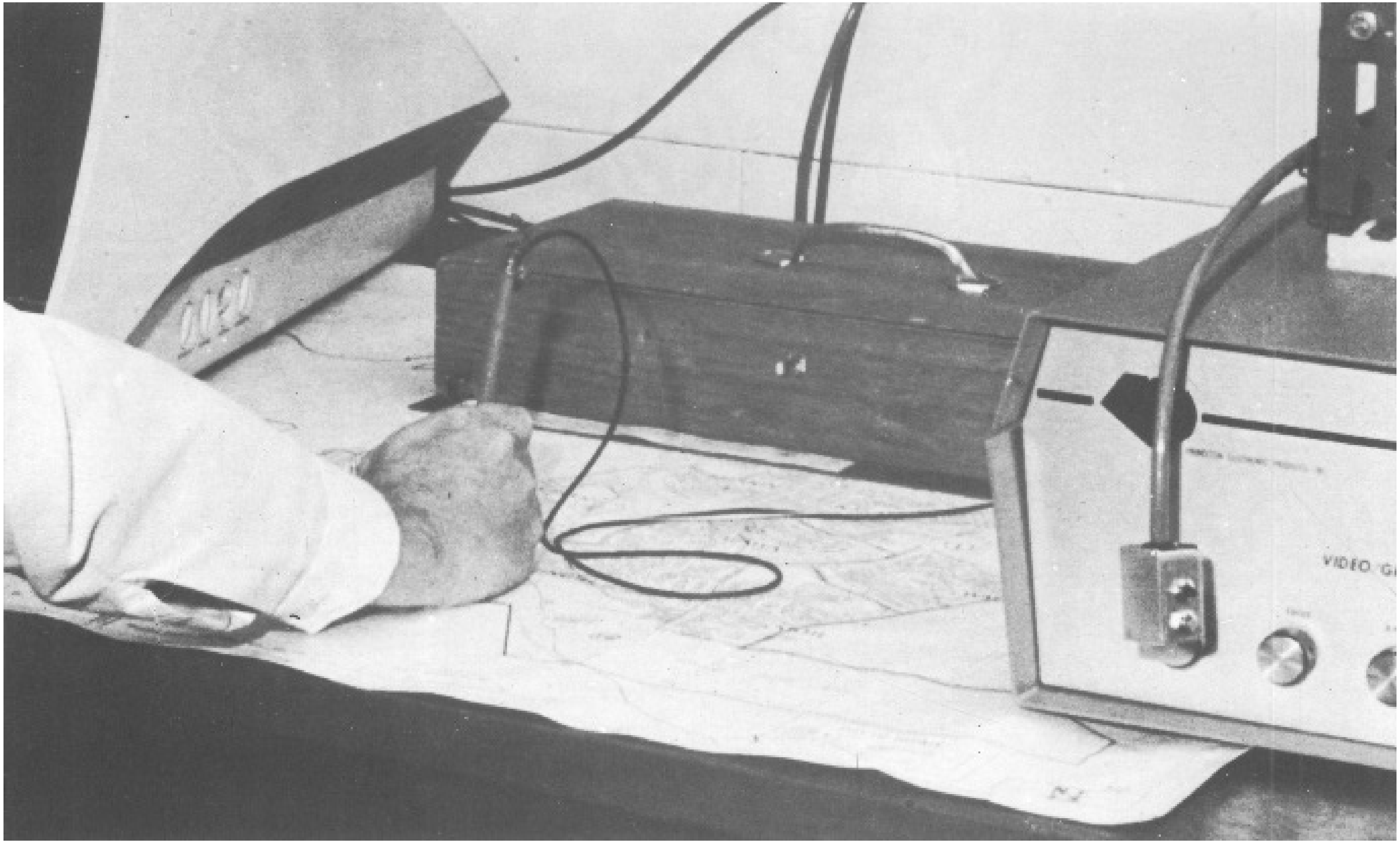


Figure 2.1.: Digitising a census map.

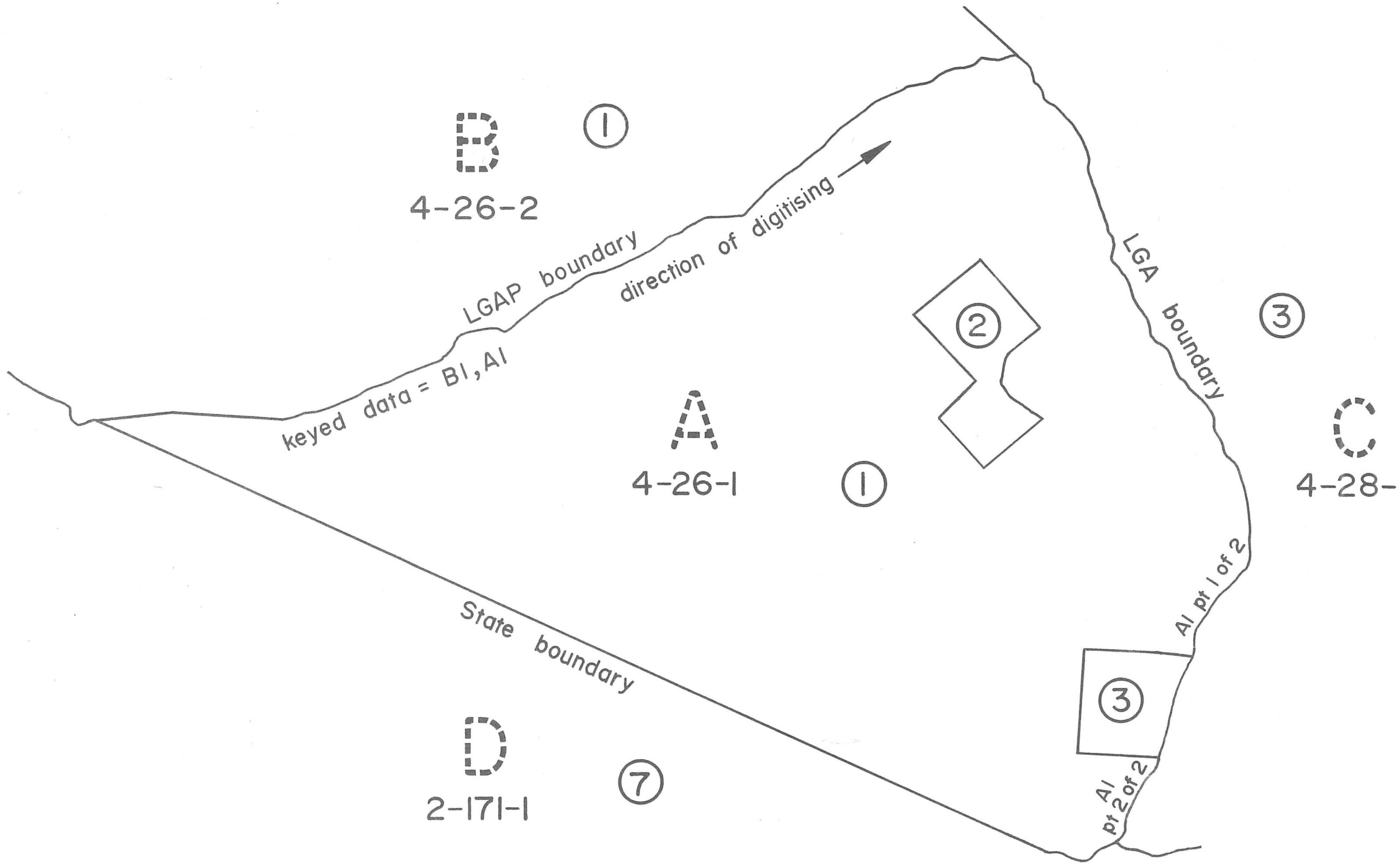


Figure 2.2 : A typical LGAP

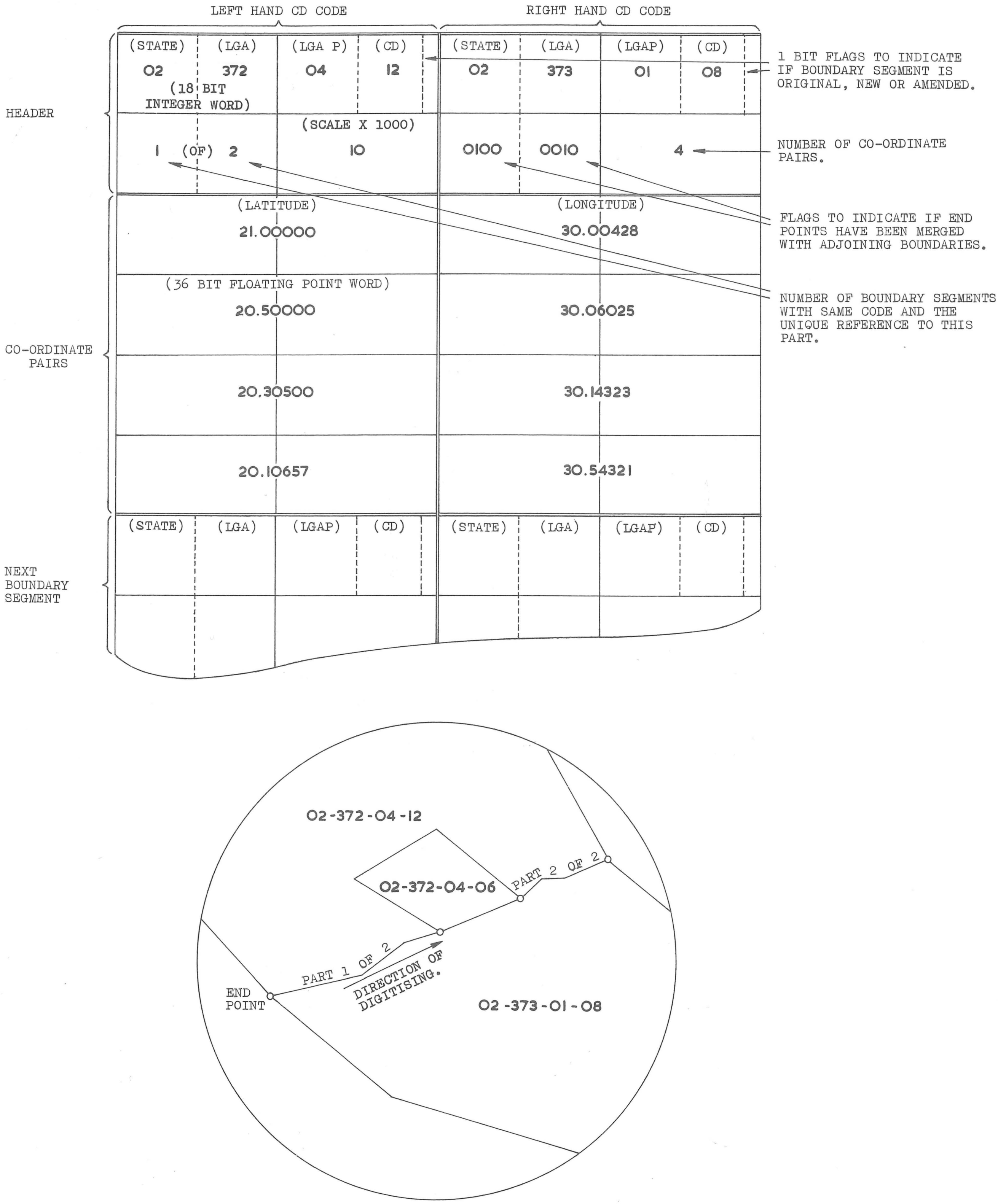


Figure 2.3.: Structure of a boundary segment.

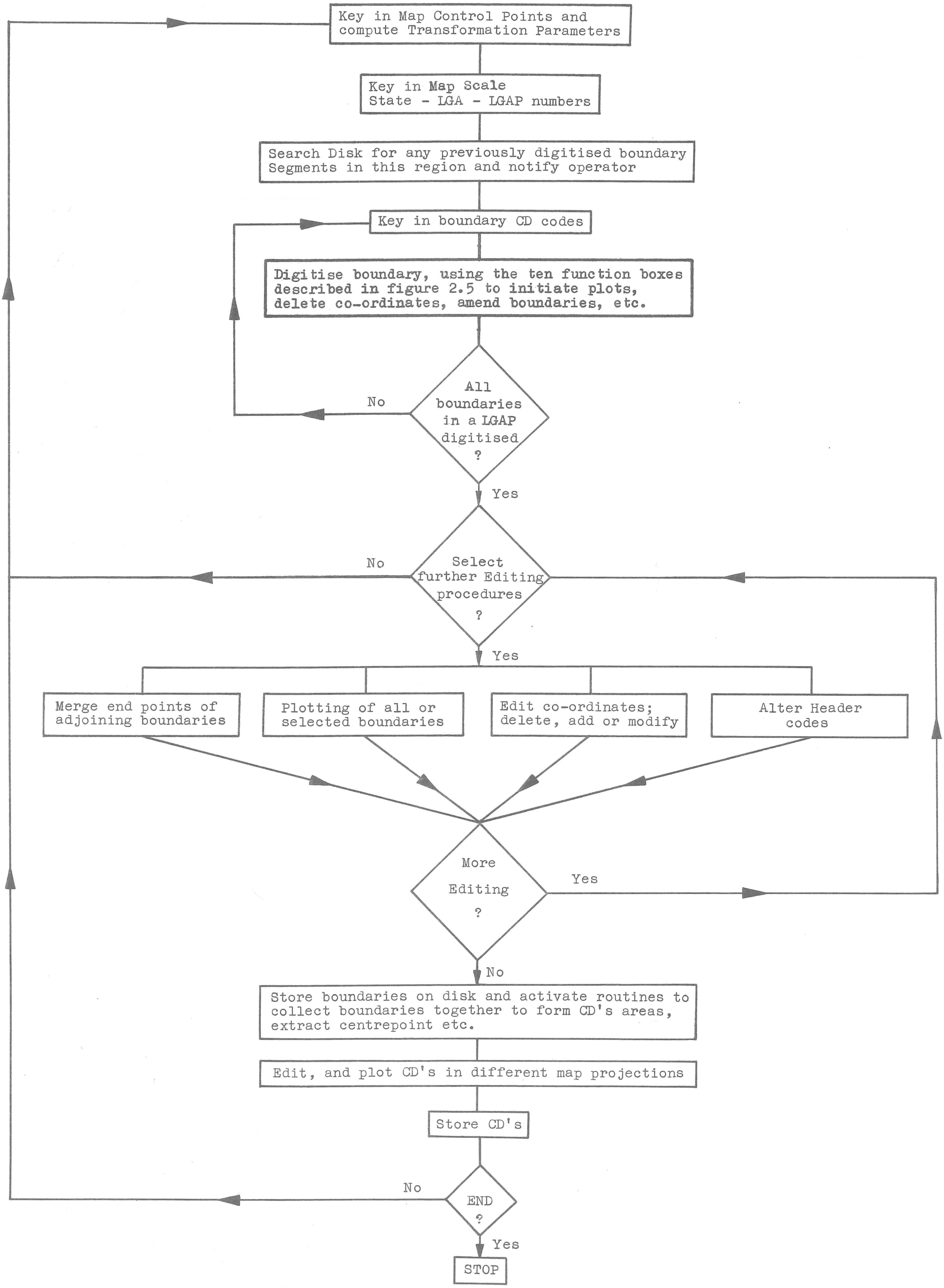


Figure 2.4 : Generalised flow chart for census mapping data entry.

DELETE LAST XY	DELETE BOUNDARY COORDINATE STRING	BOUNDARY CONTINUES ON ANOTHER MAP	END OF BOUNDARY, NEXT BOUNDARY HAS SAME END POINT	END OF BOUNDARY, NEXT BOUNDARY STARTS ELSEWHERE	SET UP MAP FOR DIGITISING	BOUNDARY LEAVES MAP AND RETURNS, NEXT POINT IS RETURN POINT	END OF LGAP	PRINT BOUNDARIES	DELETE BOUNDARY COORDINATE STRING AND CD CODES

Figure 2.5 : Function box commands

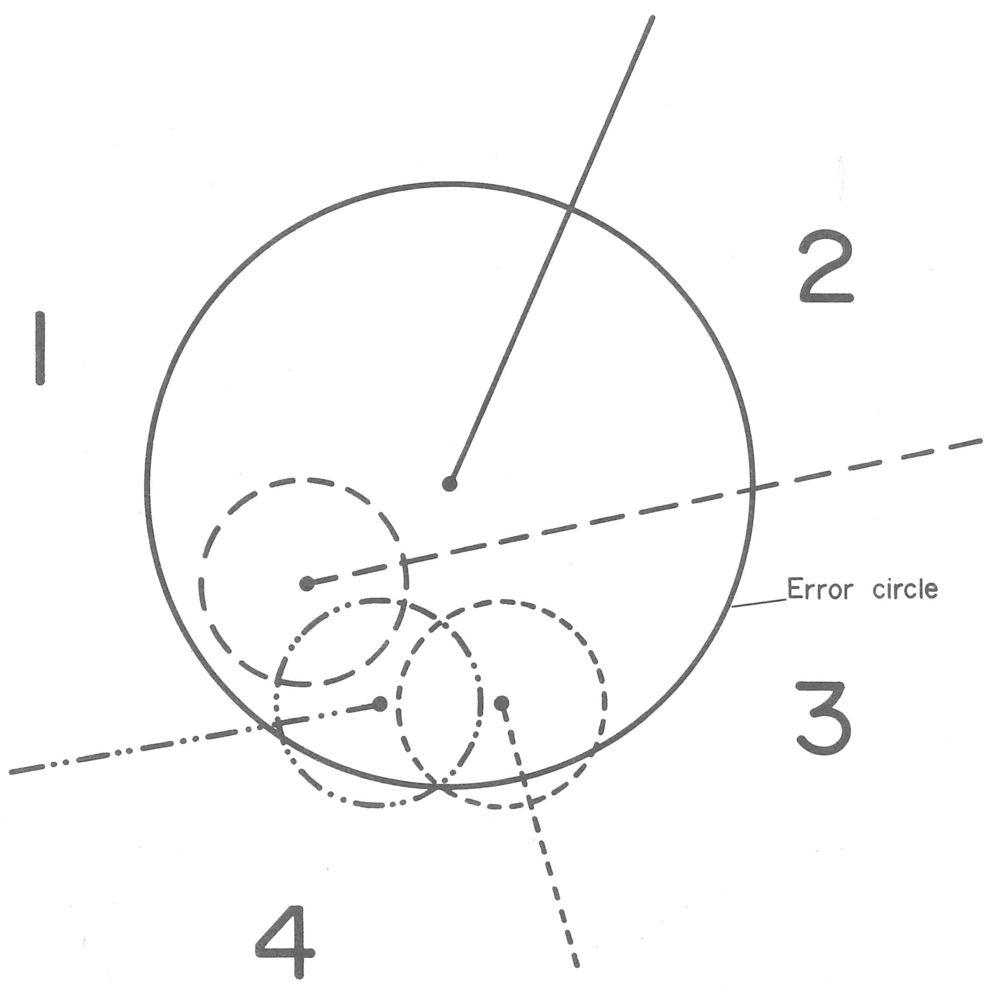


Figure 2.6 : End point merging

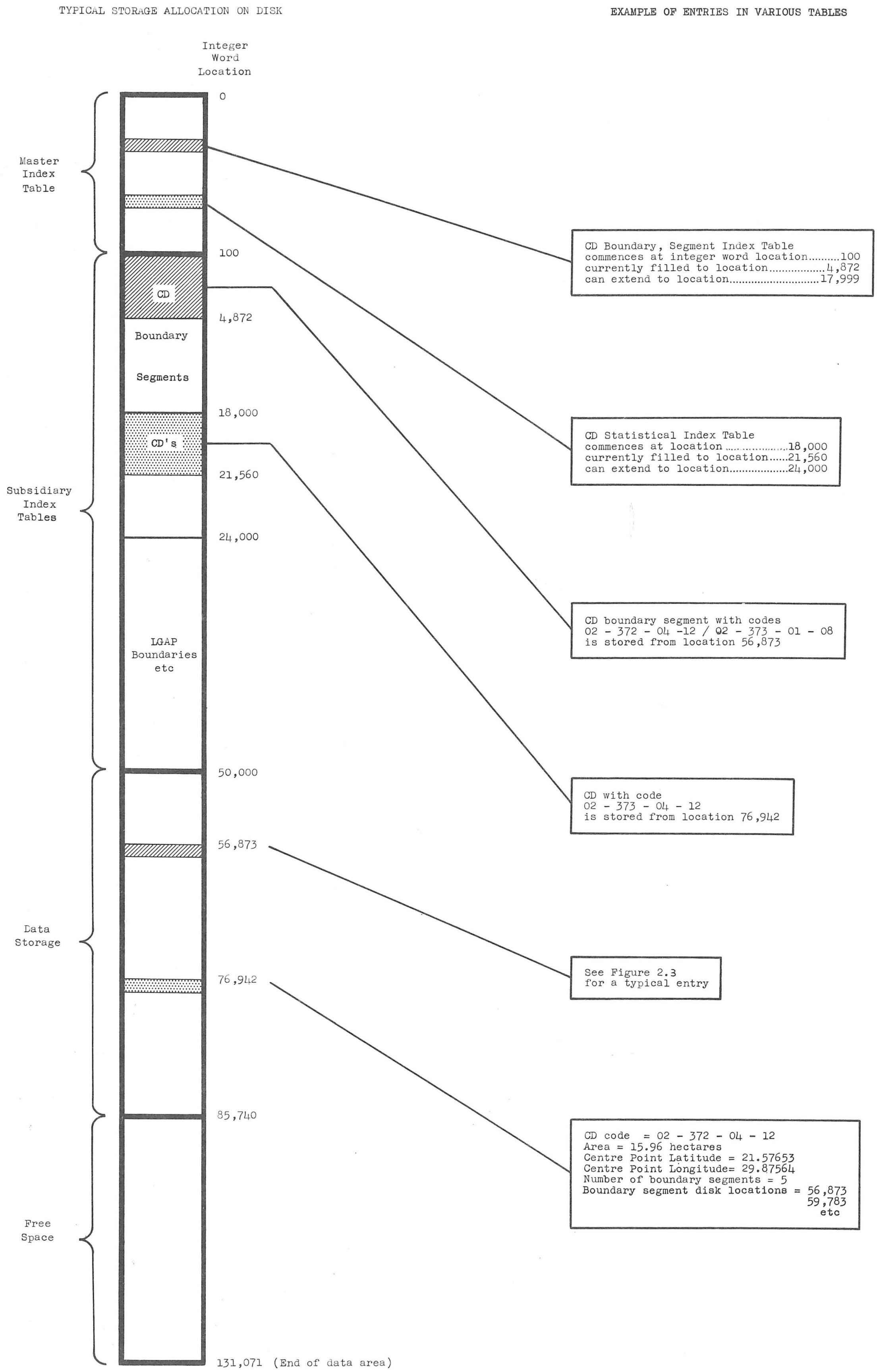
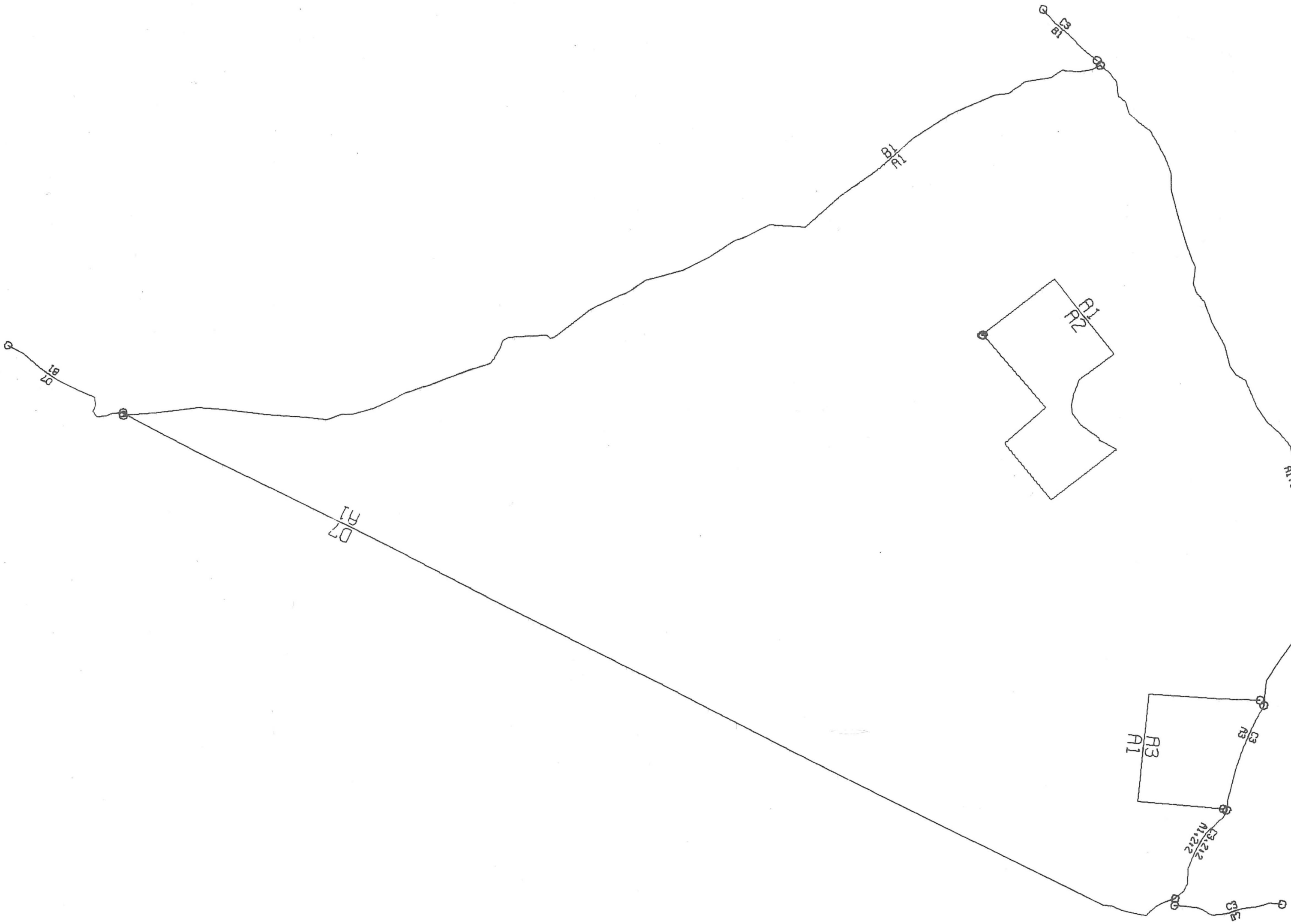


Figure 2.7 : Disk data structure.



A=4-26-1 B=4-26-2 C=4-28-1 D=2-171-1

Figure 2.8 : Plot of un-reconciled boundaries

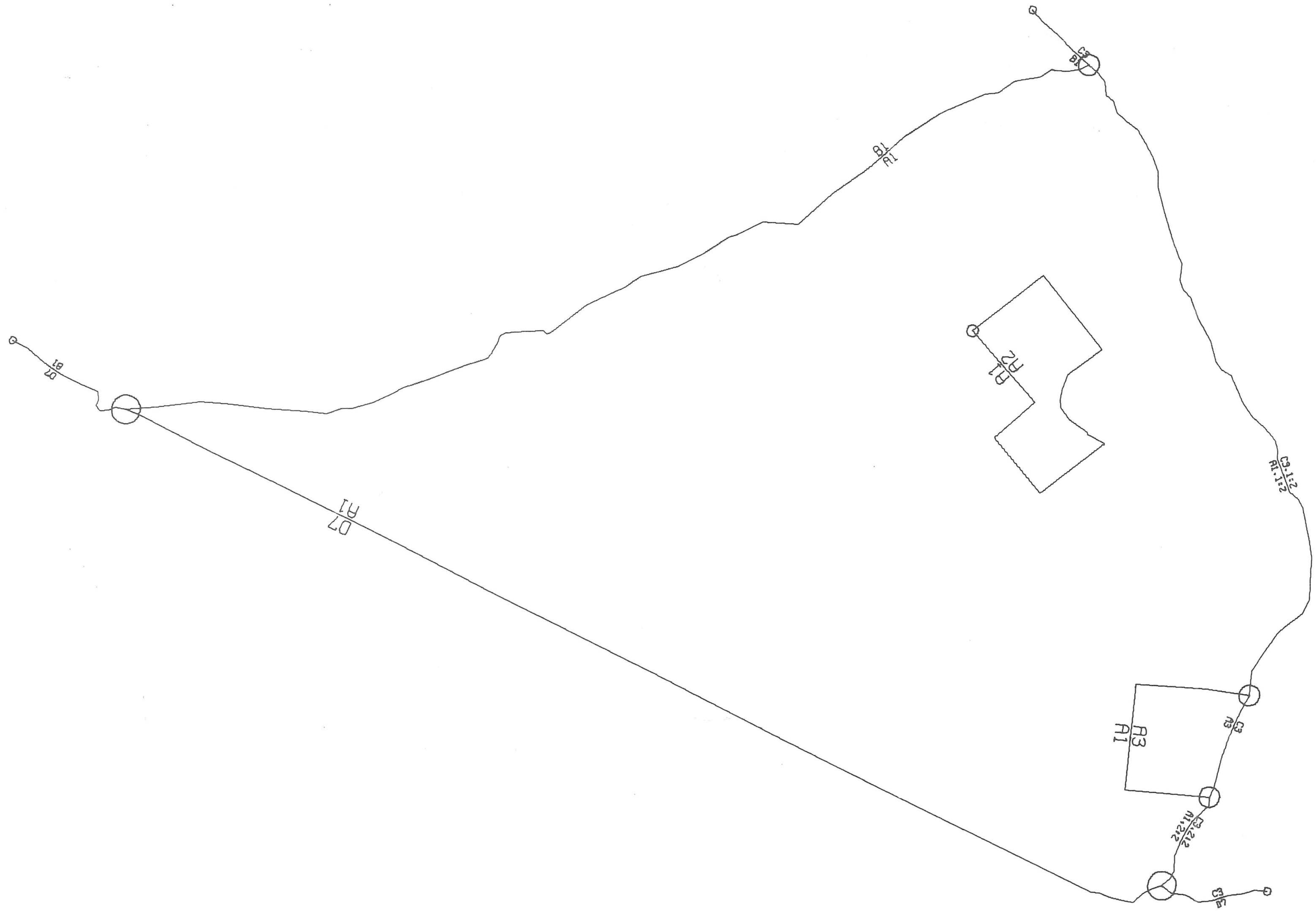


Figure 2.9 : Plot of reconciled boundaries

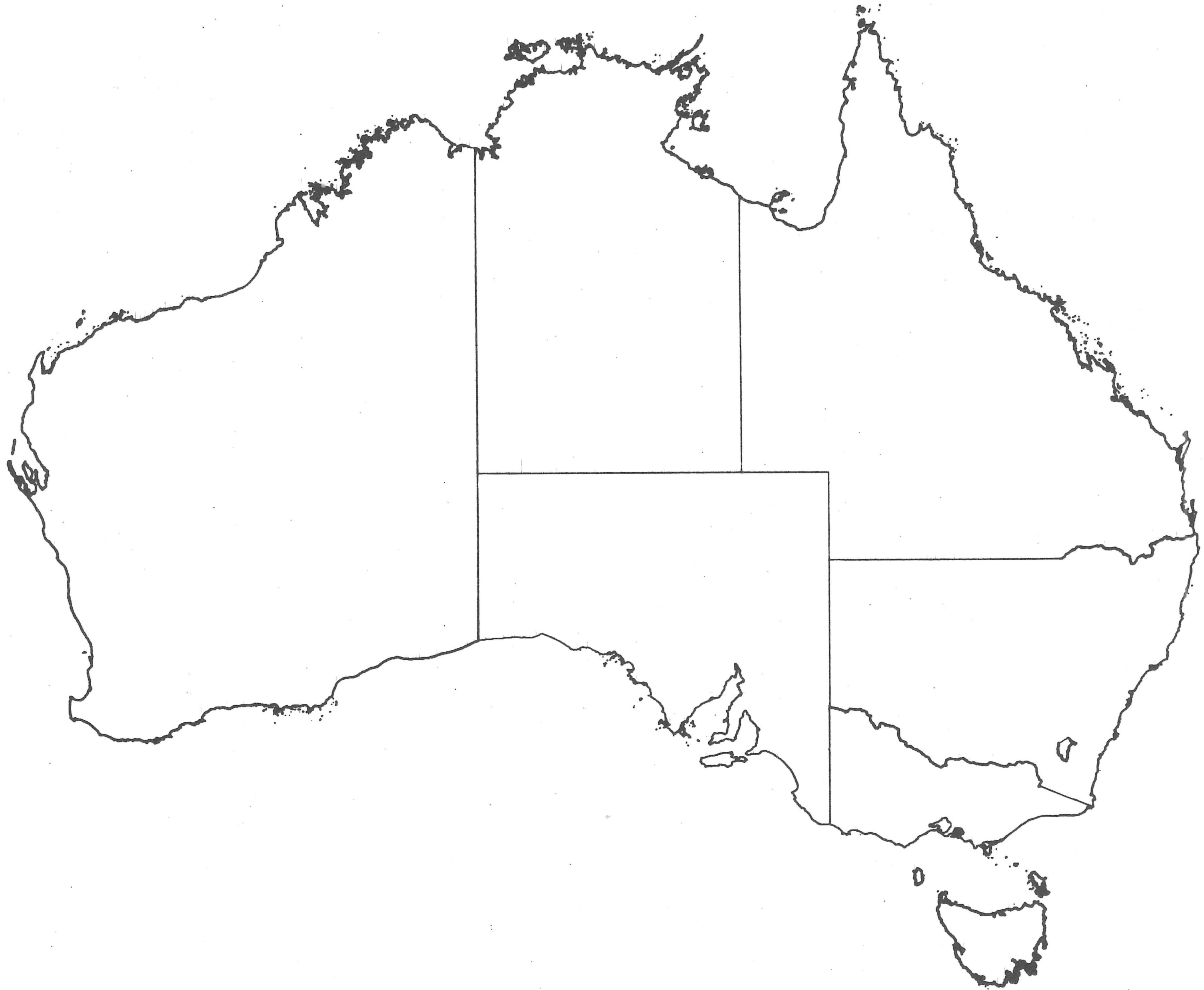


Figure 2.10 : Digitised coastline of Australia

Figure 2.11 : Digitised coastline of Tasmania

FIGURE 3.1 COMPARISON OF LAMBERT CONFORMAL AND SIMPLE CONIC MAP PROJECTION

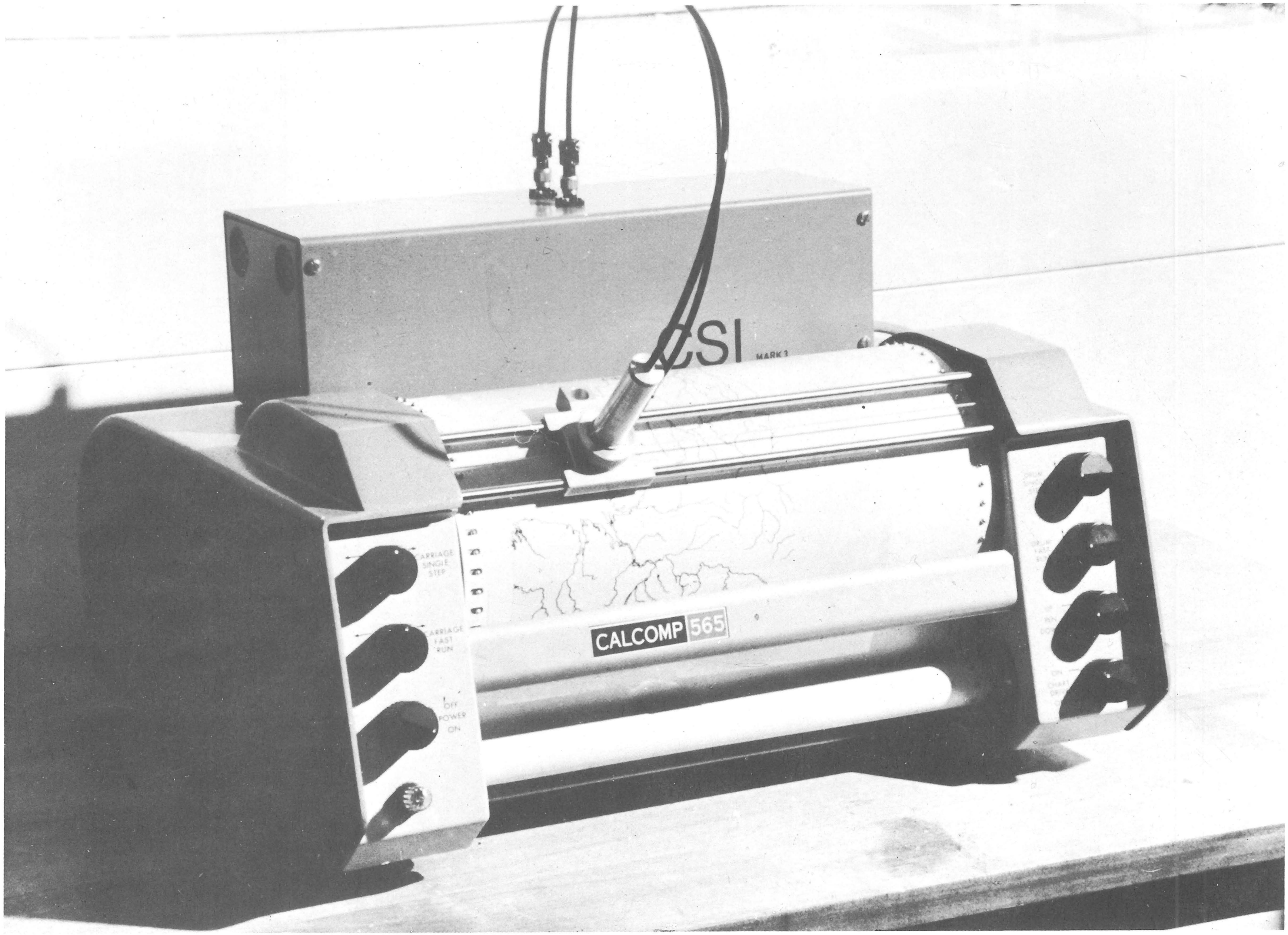


Figure 3.2 : Scanning a stream pattern.

An Inexpensive "String" Digitizer

Figure 2.1 shows a simple and inexpensive digitizer which has been made by connecting a cursor or pen by two inextensible strings through fixed eyelets to take-up drums. Changes in string length are converted into voltage changes by mounting the drums on the shafts of potentiometers - typically 10 turn, 0.1% linearity. The strings are lightly tensioned by winding elastic cord around separate sections of the drums. This constant force has little effect on the operator's ability to follow curves. Inertial forces can be more serious, but these have been kept small by reducing the drum weight to a minimum.

The configuration of the strings, take-up drums and elastic cords is shown in Figure A.1; the electrical circuit is diagrammed in Figure A.2. The take-up drums and eyelets are at the bottom of the digitizer box which can be moved around over the material to be digitized. This feature is very useful with large maps (see Figure 2.1) which might otherwise have to be cut up to fit on a digitizing table.

A spring-loaded shaft passing down the centre of the pen opens a switch whenever the pen is pressed down. Fine fibreglass cord with a braided nylon sheath is used for the strings; after stretching slightly when first used it maintains its length well. Backlash is reduced by mounting the take-up drums on roller bearings.

To convert the string lengths into X and Y co-ordinates, the digitizer was connected to a digital computer via an analogue to digital converter. The co-ordinates are then computed by triangulation. The distance 'a' between the eyelets is known, and the string lengths 'b' and 'c' are obtained from the output voltages. To solve for X and Y, the cosine rule is used as follows:

$$\cos B = \frac{a^2 + c^2 - b^2}{2ac}$$

$$\text{whence } X = c \cos B = \frac{a^2 + c^2 - b^2}{2a}$$

$$\text{and } Y = (c^2 - X^2)^{\frac{1}{2}}$$

The above equations can be solved on most mini-computers (such as a PDP-11) in less than 10 milliseconds, so that to the user, computation appears instantaneous. The equations are reasonably insensitive to small errors in the string lengths, except when the angle A between the strings is outside the range of about 45° to 135°. When Y is small and the angle A is large, Y is ill determined, whereas when Y is large and the angle A is small, X is ill determined. The boundary of the area where $45^\circ < A < 135^\circ$ is shown in Figure A.1. It is convenient to use a 5 cm deep area immediately below the base line, where the accuracy is good in X but poor in Y, for a series of "function boxes". If the pen is placed in a box, this is interpreted by the program as a control function rather than as data.

The digitizer shown in Figure 2.1 has a 45 cm base line between the eyelets. Inside the area of best accuracy shown in Figure A.1, an accuracy of $\pm$ 0.5 mm can be achieved if the pen is held vertically, and if the digitizer parameters (such as the voltage/string length constants) have been determined carefully. This accuracy is adequate for many applications. In operational experience over 18 months of intermittent usage, the basic accuracy has been maintained by recalibrating the digitizer every three months or so, the recalibration taking only a few minutes when a template made for this purpose is used.

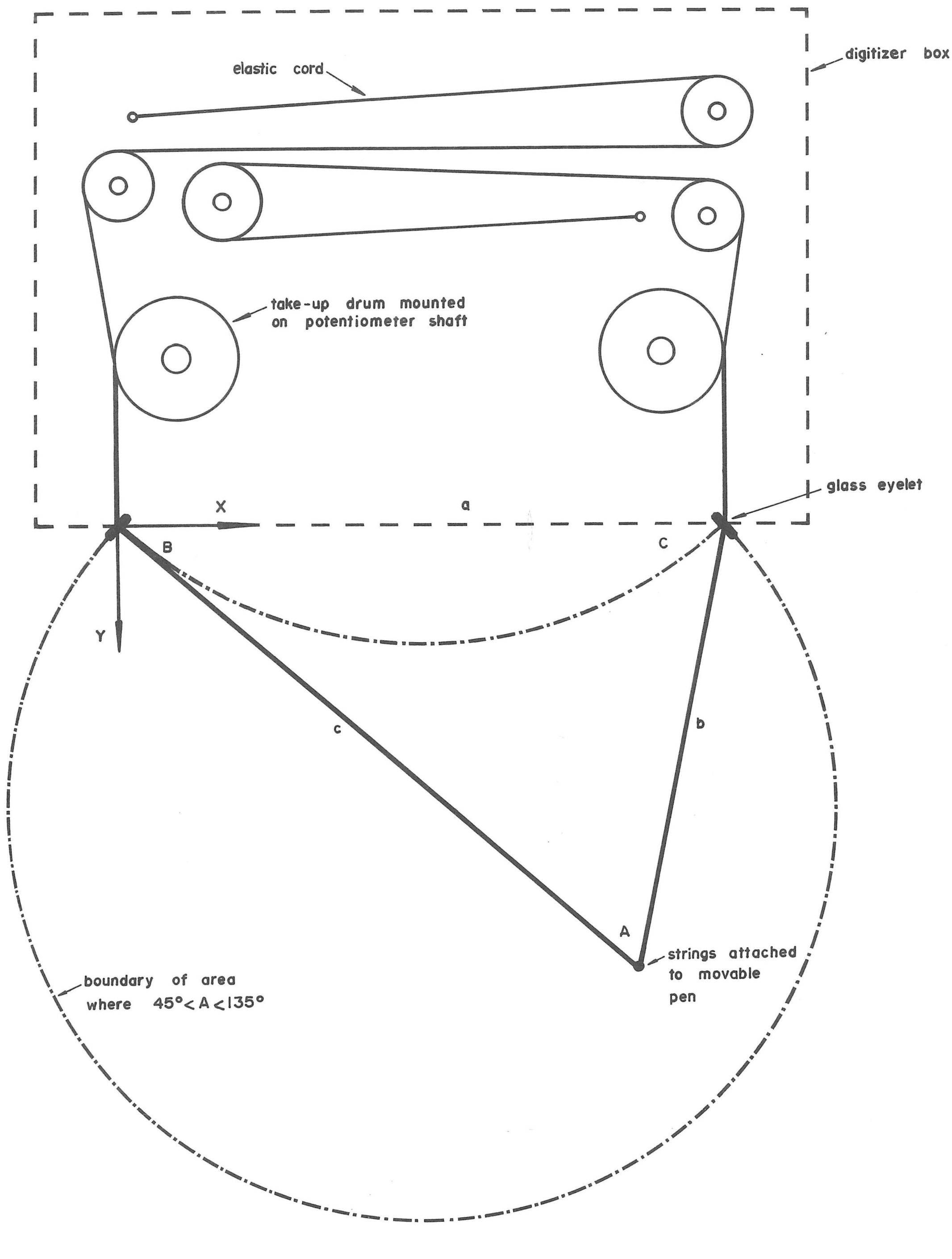


Figure A.1.: Mechanical arrangement of strings and take-up drums.

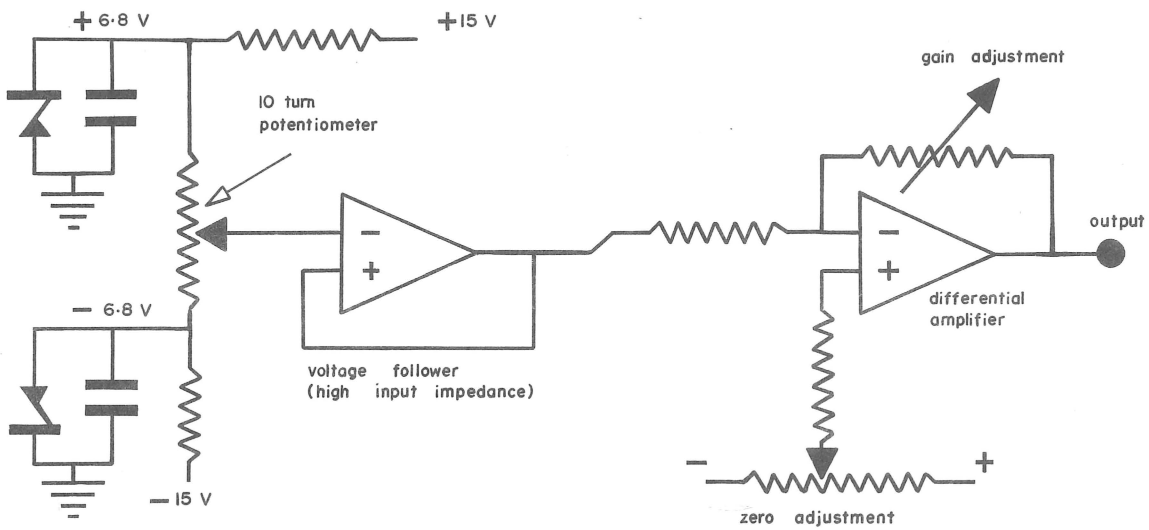


Figure A.2.: Electrical circuit diagram for one string.

A Scanning Head for Incremental Plotters

A CALCOMP 565 drum plotter has been converted into a picture scanner by replacing the normal pen assembly with an inexpensive scanning head. This head consists of a light source which illuminates a small area under the head, a lens which focusses reflected light from this area onto an aperture, and a photo-transistor which generates a signal current proportional to the amount of light passing through the aperture. After amplification this signal current can be digitized using an analogue-to-digital converter.

Figure 3.2 shows the scanning head in position on the plotter. A cross-sectional view of the head and a schematic diagram are given in Figures B.1 and B.2. Two identical photo-transistors are used, only one of which (T1) sees any light. The other photo-transistor (T2) monitors the "dark" current which varies with temperature. The incident light is measured by amplifying the difference between the photo-transistor outputs in a differential amplifier. The light source consists of a 12-volt globe connected to a stabilized power supply, and a fibre optic light pipe which transmits light from the globe to the area under the scanning head. The incident illumination is at an angle of about 45° to the scanned surface to reduce glare from glossy surfaces. The type of light source employed can illuminate the scanned area strongly without the heat and inertia which result if the globe is mounted on the pen carriage. The grey scale resolution is improved with intense illumination, which also allows small appertures to be used. The optical arrangement shown in Figure B.1 gives a magnification close to unity, so that for a plotter step size of 0.125 mm, the smallest useful aperture is about 0.125 mm diameter.

In operation, the input picture is fastened to the plotter drum, and the scanning head is moved across the picture in the desired scanning pattern. The normal plotter software is used except that an analogue-to-digital conversion is initiated at the completion of each plotter step with the digitized grey scale values being stored in a vector. It has been found that the scanning head lags behind the position assumed at the end of each step, thus leading to a "backlash" effect when successive scan lines are in opposite directions. This effect can be eliminated by inserting a delay of about 3 milliseconds at the end of each step before digitizing. This allows the scanning head to reach its rest position, and, as the time required for each plotter step is 3.3 ms, insertion of this delay halves the scanning speed. Experiments have shown however that the backlash is relatively constant at about one plotter step of 0.125 mm, and may thus be allowed for by programming so that scanning at full speed is still practicable.

The performance of the scanner is adequate for many applications. With a 0.125 mm step size and a 25 cm wide drum, the mechanical positioning is accurate to 1 part in 2000. The grey scale resolution is good: when the amplifier is set for a black/white range of 0.5 volts at the output, the combined drift and noise is less than 2 millivolts. The grey scale and positional accuracy compare favourably with flying-spot (CRT) and drum-type scanners costing many times as much. Given that a plotter and analogue-to-digital converter are already available, the additional cost of the scanning head and amplifier is only a few hundred dollars. The main problem with the plotter-based scanner is that its maximum scan rate is limited by the plotter to 300 points per second, which may be too slow for some applications. This shortcoming is offset by the free choice of scanning pattern and the possibility of using the drum surface for both plotting and scanning. For example, positioning marks can be plotted prior to scanning to ensure correct alignment of the input picture.

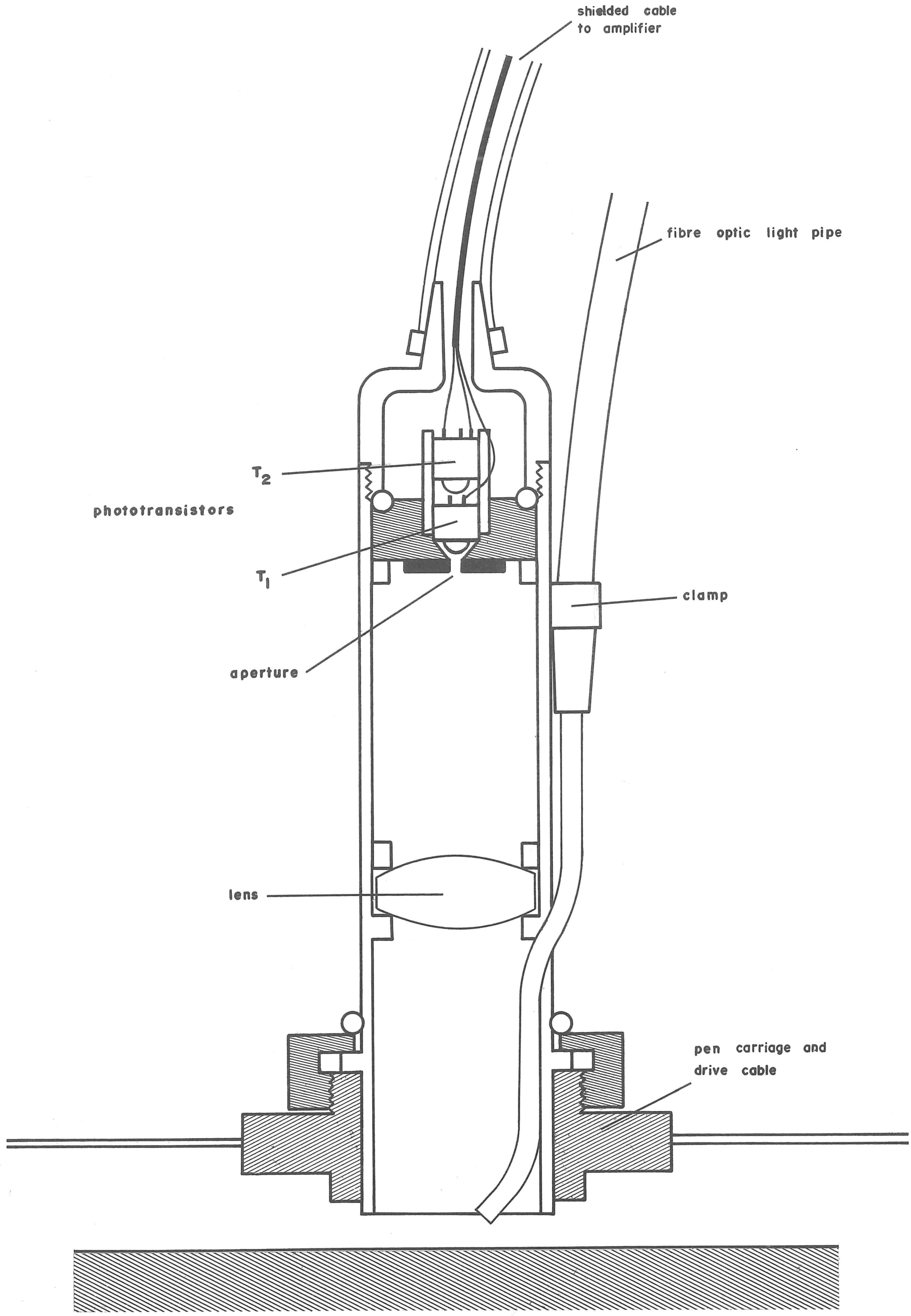


Figure B.1. : Cross sectional view of scanning head.

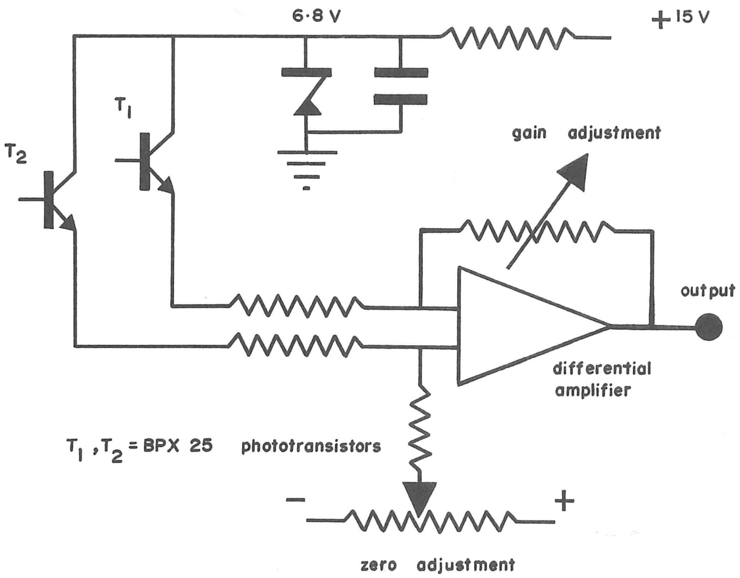


Figure B.2.: Electrical circuit diagram of picture scanner.