THE UNIVERSITY OF CALGARY

# INTEGRATION OF CADASTRAL SURVEY PLANS INTO A COMPUTERIZED LAND INFORMATION SYSTEM

BY

MICHAEL P.A. MEPHAM

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE
DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF SURVEYING ENGINEERING

CALGARY, ALBERTA
DECEMBER, 1988

## ABSTRACT


A primary source of data for Land Information Systems (LISs) is the cadastral survey plan. An automated system for entering, analyzing, and integrating the information shown on these plans is developed in this thesis. Existing systems are evaluated on the basis of their speed, robustness, completeness and simplicity, and are found to have serious limitations.

The design of this new system begins with an enumeration of the types of entities that appear on cadastral plans and their relationships with each other. Rules are developed for the automatic detection and modelling of these relationships. This includes the formation of the polygons defining the lots, blocks, etc. shown on the plan, and the detection of constraint conditions, such as points that must lie on a straight line.

Constrained least squares is presented as a powerful tool that uses not only the observations in the calculations, but also the topological relationships, for the calculation of the positions of the points on the plan. Data errors are detected using a combination of topological rules and statistical analysis. This ensures that the data entered into the LIS are virtually error free.

The proposed system is implemented in a prototype program and the results of testing using real cadastral survey plans are presented. The implications of using this program in an existing production environment are examined and it is expected that significant productivity improvements and cost savings can be realized.

## ACKNOWLEDGEMENTS

The completion of this thesis would not have been possible without the help and support of many people. Dr. G.D. Lodwick was the chairman of my supervisory committee, which also included Dr. E.J. Krakiwsky and Dr. M.R. Coulson. They freely gave of their time and knowledge when I needed assistance or support. Their contributions were invaluable.

My fellow graduate students in the department earned my gratitude by serving as sounding boards for ideas and concepts that were not always correct. Our, frequently long, discussions were always useful.

Financial support for this project came from many sources. I would like to thank the Province of Alberta for assistance through the Professorship in Digital Mapping and Spatial Data Management, the Government of Canada for support through an Energy Mines and Resources research grant, the Natural Sciences and Engineering Research council for assistance in the form of graduate student assistantships and, finally, Walker Newby and Associates for awarding me the Walker Newby graduate award on two occasions.

## TABLE OF CONTENTS

TABLE OF CONTENTS (Cont'd)

CHAPTER PAGE

TABLE OF CONTENTS (Cont'd)

CHAPTER PAGE

TABLE OF CONTENTS (Cont'd)

CHAPTER PAGE

## LIST OF TABLES

## LIST OF FIGURES

# Chapter 1

# INTRODUCTION

Since the Babylonians produced the first cadastral maps almost 6000 years ago, government authorities have recognized that the maintenance of a consistent, complete record of the ownership of the land under their authority is one of their prime responsibilities. Since then, efforts have been made to make this process as efficient as possible. Over the last two decades the use of computers has emerged as a very effective approach to the storage and manipulation of spatially referenced data. The development and implementation of a data entry system for cadastral survey plans, to be used as the first component of a computerized land information system, is described in this thesis.

## 1.1 Objectives

The history of computerizing land records goes back over two decades. In 1966 the American Bar Association sponsored a conference on the modernization of land records at the University of Cincinnati. This was followed by two more conferences in 1968, one in 1972, and one in 1974. The principal result of these conferences was general agreement that the most suitable building block for a modern computerized system of land records is the parcel [Phillips, 1975]. Since the North American Conference on Modernization of Land Data Systems in 1975, there have been hundreds, if not thousands, of papers presented at conferences and published in journals dealing with the whole concept of computerizing cadastral data.

While this was going on, the idea of creating Land Information Systems (LISs) for purposes other than simply cadastral land management was also growing [Bathke, 1979]. In Alberta the two separate approaches have coalesced into a unified approach, whereby the positions of the cadastral survey points form a part of the geographic positioning base for a comprehensive Land Information System [Kennedy, 1986]. Separate component of the system are the cadastral maps describing all of the land holdings in Alberta [Beraha, 1986].

The data source for both the positions of the cadastral survey points and the mapping information is the registered survey plans held by the Land Titles Office. A typical example of such a plan is shown in Figure 1.1. (All information identifying the location of this plan and the surveyor responsible for it has been removed.) In 1986 there were 110,000 such registered plans on file and the number is increasing by 5000 per year [Beraha, 1986]. The design of a system for the entry of such a large volume of data requires careful study and analysis. The system must meet the following objectives:

a) The entry of the data must be fast. If it requires two hours to enter a typical plan, like the one shown in Figure 1.1, then it will require 110 person years to enter the existing plans on file in Alberta and five person years every year thereafter to enter

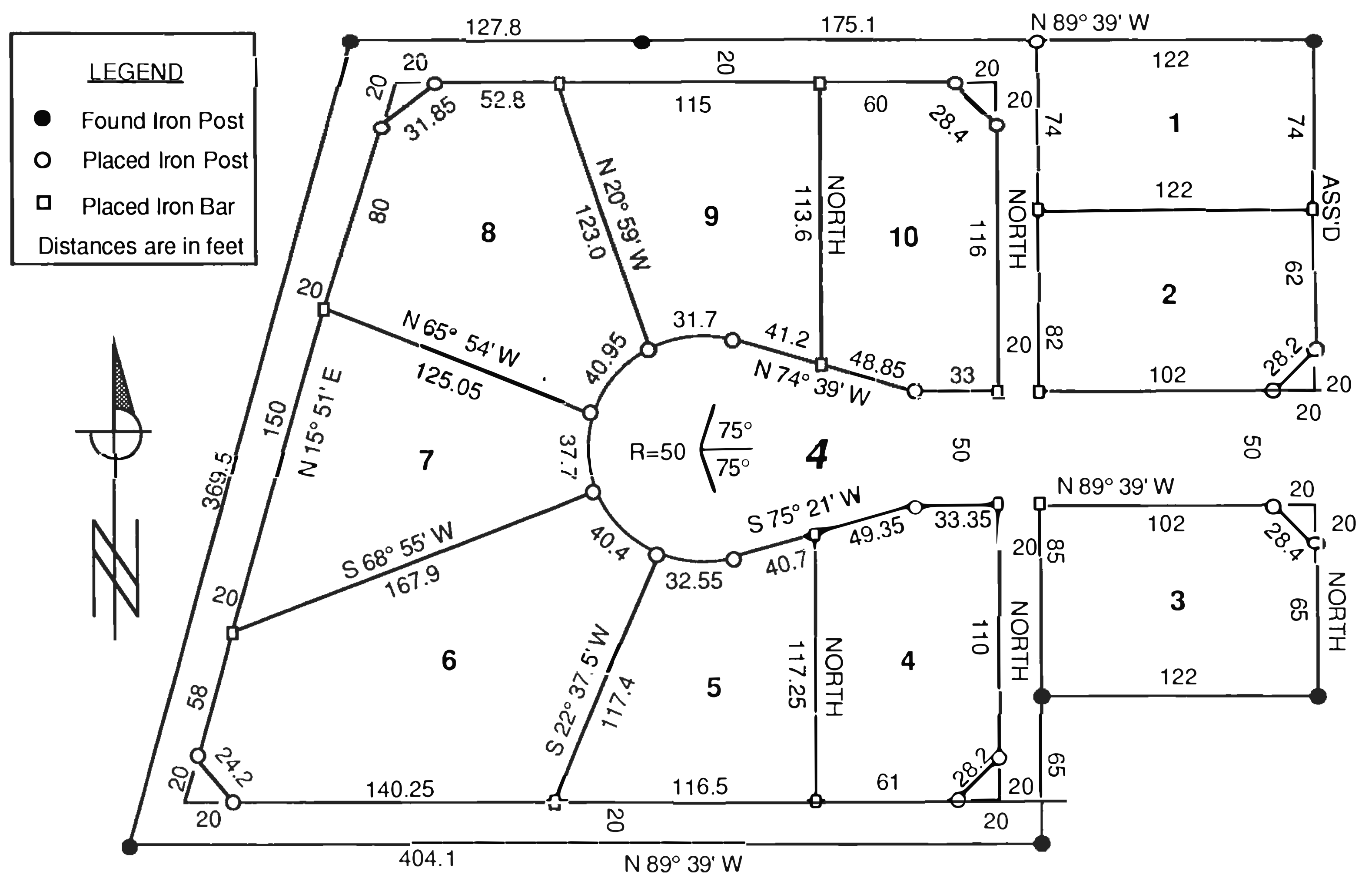



Figure 1.1: A typical cadastral survey plan

the updates. These figures are based on a 40 hour work week and a 50 week work year and as such are somewhat optimistic.

b) The data entry process must be robust. No attempt has been made in the past to ensure that all of the metric data shown on the plans is correct. The plan shown in Figure 1.1 contains an error that was not discovered in the 25 years that elapsed between its registration and the computation of the coordinates of the property corners shown on the plan. In addition, the system must deal gracefully with any data entry errors made by the operator. Program crashes resulting from simple mistakes are not acceptable.

c) All the information shown on the survey plan must be entered if the digital plan is to serve as a substitute for the paper plan. There are three different types of data shown on the plan. They are:

i) The textual data. This includes such things as the date of the survey, the name of the surveyor, and the purpose of the survey. The entry of this information can be simply a matter of typing the information at a keyboard.

ii) The metric data. This includes all the observations shown on the plan. In addition to the observed value, the identity of the points between which the observed value applies must also be entered. The observed values must be maintained at all times as attribute information. After the coordinates of the corners have been computed, the observed and computed values will differ slightly due to the inexact nature of observed values and the high level of redundancy of the information shown on the plans.

iii) The topological information. This includes the connectivity between the points related to the observations as well as other topological conditions that exist on the plan. Included in these other conditions is the description of what lines form the boundaries of the various polygons shown on the plan, and such conditions as a set of points that lie on a straight line between two other points.

d) The system should be simple to operate and should operate on simple hardware. If the program is simple to operate then there will be fewer operator errors and the training time for new operators will be reduced. The use of simple hardware for the operation of the program implies a lower capital cost for its operation, making its use widely available. For example, with a low capital cost the program would be available to private land surveyors who could then perform plan entry on a contract basis for government agencies.

With these objectives in mind the remainder of this chapter presents an examination of the two most popular data entry approaches in use today with a view to how well they meet these objectives. Also given is a preview of the approach used in the system that has been designed and implemented as a result of this research.

Specific details of the design and implementation follow in subsequent chapters.

## 1.2 Existing Approaches

The entry of cadastral survey plans into computer readable form is not a new activity within the realm of Land Information Systems. It has long been realized that such a computer map is an essential component of these systems [e.g. Wahl, 1975]. In this section two approaches to the computerization of survey plans are presented and evaluated.

### 1.2.1 Digitizing

The idea of using a digitizing tablet to convert a legal survey plan from paper form to digital form seems an obvious solution. The use of this approach was described by Cremont [1980] as being suitable if "maps and documents with satisfactory quality are available". In Alberta there is no doubt that the maps and documents are available and generally of a high quality.

The basic approach with a digitizing table is to digitize all points and lines shown on the plan without regard to the underlying topology. The result is a, so called, spaghetti file consisting of random points and lines which, when plotted on a screen, provide only a graphic replication of the original plan. The text shown on the original plan is also entered but only as a series of text strings, each of which is to be placed at a given position with a given rotation and displayed with a particular font and size. The text is not associated, in the file, with any particular lines or points.

The accuracy of the coordinates of the digitized points is not very good. Petersohn and Vanderhoe [1982] reported mean positional differences of up to 7.8 metres in a comparison of cadastral survey plan digitizing and field surveys, for the purpose of obtaining coordinates of property corners. Some of the discrepancy is due to errors on the plans being digitized and parts are due to the digitizing process itself. The plans that were used were less than 17 years old and considered to be of good quality. Errors of this magnitude are clearly unacceptable.

The digitizing process results in what appears to be a digital form of a cadastral survey plan, in that the plan can be displayed on a graphics screen or plotting table and when viewed it transmits all of the same information to the user that the original paper plan did. However, the information is obtained by the viewer himself making a visual analysis of the shapes and patterns formed by the lines. The user can, for example, easily identify which lot is lot number five, by finding the area on the plan that has a "5" located within it. There is nothing in the data that enables a computer program to identify the limits of lot five automatically. This inability to apply intelligent processing to the data severely restricts its future use for anything other than reproducing copies of the original survey plan.

A second disadvantage to the use of digitizing is the inability to check the dimensions shown on the plan. The effect of this is to perpetuate any errors shown on the plan in the digital copy of the plan. This failing, in conjunction with the positional errors described earlier, further reduces the utility of the data, as well as user confidence in the entire data base, once the users encounter the errors themselves when they try to use the data.

Finally there is the pure tedium of the digitizing process itself. Any person having to perform such a task is going to make errors. These may be errors of omission (missing data), incorrect attributes (typographical errors), or simply a failure to properly join two lines without under or over shooting. Such problems are discussed in detail in Burrough [1986] and Mulaku [1987]. The result is that any digitized plan requires extensive manual checking and editing to ensure that it is as correct as the original plan.

In summary, of the objectives identified in the preceding section, digitizing only satisfies objective (d). Digitizing is a relatively simple process to learn and perform and the hardware required is widely available at a reasonable cost. While digitizing is reasonably fast, as required by objective (a), the checking and editing process adds considerable time to the process causing it to fail objective (a). As far as the objectives of robustness and completeness, objectives (b) and (c), are concerned the digitizing approach to the computerization of cadastral survey plans fails miserably.

### 1.2.2 Coordinate Geometry

In an effort to improve the positional accuracy of the digitized plans, and to help detect any errors in the observations shown on the plan, many agencies have adopted a coordinate geometry approach to the calculation of the coordinates of the property corners. This is the approach currently used by the Province of Alberta [Mepham et al., in prep.] as well as other provinces and municipalities [e.g. Province of Manitoba, 1986].

Using this coordinate calculation method, the operator forms and calculates a series of closed traverses through the plan being computerized. He checks the closures on the traverses in an attempt to locate any errors in the data shown on the plan. Any traverse that fails to close within tolerances is considered to have had one or more erroneous observations used in it. Once the bad observation has been identified, it can either be ignored for all future calculations or the plan can be sent back to the surveyor for correction (the usual practice).

This results in many different sets of coordinates for each point on the plan, since each point is used in more than one traverse and thus has its coordinates computed several times from different data. The various sets of coordinates must be resolved to a single set. This may be done by either taking the mean of all the sets for each point or choosing the set from the "best" traverse.

Once the coordinates of the points have been computed then the appropriate connections between points are identified. This process is similar to that used in the straight digitizing approach with the exception that all the lines must start and end at one of the coordinated points. These connections can be entered using either a digitizing tablet or screen plot of the points and pointing to the points to be joined, or by entering a list of pairs of points to be connected. In either case it is a tedious process and, once again, results in a spaghetti file of connections.

The use of traverse closure specifications for the detection of blunders in the observations is not a statistically valid operation [Blachut et al., 1979]. In order to properly determine whether a traverse closes or not, the standard deviations of the observations should be propagated into a covariance matrix for the coordinates of the end point and these coordinates tested for closure in the context of their accuracy [Mepham, 1983b]. This is rarely done. The result is that the detection of errors is an empirical operation with no statistical basis and may vary in its application from operator to operator.

This blunder detection process depends on the operator forming a sufficient number of unique traverses to check each and every observation shown on the plan. Mepham [1983b] shows that the minimum number of traverses required is

$$NT = NL - NP + 1, \qquad 1.1$$

where: NT is the number of traverses,
NL is the number of lot lines on the plan, and
NP is the number of points on the plan.

In the simple plan shown in Figure 1.1 there are 65 lines and 48 points. Applying Equation 1.1 to this plan shows that a minimum of 18 traverses must be computed to check all the data shown on the plan. It is unreasonable to expect an operator to manually form the minimum number of closed traverses due to the complexity of the process. The result is that, in general, more than the minimum number of traverses will be formed and computed, extending the time required to complete the coordination of a plan.

In summary, the coordinate geometry solution to the problem of converting cadastral survey plans to digital form is more robust than the digitizing approach but is more time consuming. It is more complicated to use as a result of the requirement for the manual formation of the traverses. Both methods result in a similar digital product in terms of content but the coordinate geometry approach has better positional accuracy and will result in fewer erroneous observations being stored in the data base.

## 1.3 Proposed Solution

Neither of the two methods examined above meet the objective of completeness. Both methods yield results that can best be described as "digital reprography" since no analysis or decision making can be performed using the data without viewing it graphically. In addition, neither approach models the basic unit of the cadastral plan, the lot. A cadastral plan is not a set of points and lines but rather is a graphical model of a set of lots and roads that are bounded by points and lines. The proper modelling of the plan requires that the complete topology shown on the paper plan be properly modelled as well. In order to provide the position of a lot, for example, one must know what lines bound that lot and what points delimit the lines. This aspect of the problem is covered in detail in Chapter 2 of this thesis.

In addition to the points, lines, and polygons shown on a cadastral plan that must be entered, there are many special topological conditions that must also be identified and entered. Among these is the situation of a series of points that must lie on a straight line or along a circular curve. The requirement that the operator find all of these conditions and enter them has the dual disadvantages of making the data entry process more complicated and thus slower, and of increasing the chances for errors of omission. For these reasons rule based systems for the automatic detection of these conditions were developed and implemented as part of the design and implementation of the data entry program. These systems are described in Chapter 3. For similar reasons the formation of all the polygons that form the basic units of the plan is highly automated. The process for this is presented in Chapter 4.

Robustness was considered at all stages of the design and implementation of the data entry program. Errors and omissions made by the operator need to be detected and, where possible, corrected. An area where robustness is especially important is the calculation of the coordinates of the points and the detection of any errors in the data. For this reason the method of least squares was chosen for the coordinate calculations. The topological conditions that exist on the plan are modelled mathematically and incorporated into the least squares adjustment as constraint equations, improving both the accuracy and reliability of the solution. The results of the adjustment are rigorously analysed using statistical tools to ensure that there are no errors in the data. This process of coordinate computation and analysis is presented in Chapter 5.

If the proposed data entry process is to meet the objective of completeness the challenge is also to make it fast and simple (objectives (a) and (d)). As will be seen in the following chapters the data entry process requires repeated identification of the survey points shown on the plan. As an example, the operator has to provide the identifiers of the points at the ends of observed lines, points that are the start and end of curves, etc. One of the fastest and most natural methods of identifying objects is to point at them [Newman and Sproull, 1979]. The implementation of this pointing approach and other steps to simplify the operation of the program are presented in Chapter 6.

## 1.4 Summary

This chapter has outlined the current situation concerning digital data entry of cadastral survey plans into LISs. It has also identified important objectives that must be met by an effective solution to this problem. The cadastral data entry system described in the following chapters meets all four objectives described at the beginning of this chapter. It has been implemented and tested and the results are presented in this thesis. The system works and should meet the data entry needs for cadastral survey plans for many years to come.

# Chapter 2

## ENTITY TYPES AND CLASSIFICATION

In this chapter the entity types used to describe the positional and topological information shown on a legal survey plan are described. In addition, a classification scheme for these entities is given that allows the formation of rules to be used in the examination and coordination of a legal survey plan.

### 2.1 The Three Basic Entity Types and Their Relationships

There are three basic types of spatial entities. These are points, line segments, and polygons. By using just these three entity types, it is possible to model the planimetry of all features on the surface of the earth, whether natural or man-made. Points are the most elementary of these entities. They are used to model dimensionless features (e.g. survey control points) or to model small features with a well defined size, shape, and orientation (e.g. power poles or manhole covers). In the latter case the coordinate information stored for the point is used only to locate the feature and would typically be the position of its centre.

The second of the basic entities is the line segment. These are used to model one dimensional (linear) features. The best example of a linear feature is a legal boundary. Line segments can be used to model other apparently linear features, such as roads and rivers, by storing the centre line of the feature but this is, in fact, a generalization of the feature. While the feature can be mapped using a line at some scales (e.g. 1:50000) it should be drawn as an areal feature at other scales (e.g. 1:1000). Railway tracks can be modelled as linear features since they have a standard width (gauge) but the associated right-of-way cannot as it tends to have a variable width. The definition of a line used in this thesis is the simplest definition possible. A line segment starts and ends at defined points and is either straight or circular with a known radius and orientation. Only two points are required to define these line segments.

The third basic entity is the polygon. Polygons are used to define the extent of areal features, such as legal parcels, lakes, or forest cover. Polygons may be defined in many ways [Lodwick and Feuchtwanger, 1987] but the most natural approach, and the one used in this thesis, is to define the polygon by listing the line segments that form its boundaries. Conventionally, polygons are described in a clockwise manner. For this reason the direction, forward or reverse, that each line segment is used in defining the limits of the polygon must be recorded.

Figure 2.1 is a simple Entity-Relationship Model (ERM) diagram showing the relationships between points, line segments, and polygons. (For a description of ERM diagrams see Chen [1977] or Feuchtwanger and Poiker [1987].) The simple relationships shown in this diagram illustrate the conceptual simplicity (and elegance) of using the three entity types described above for modelling the features of the earth. All the positional information is stored in a single

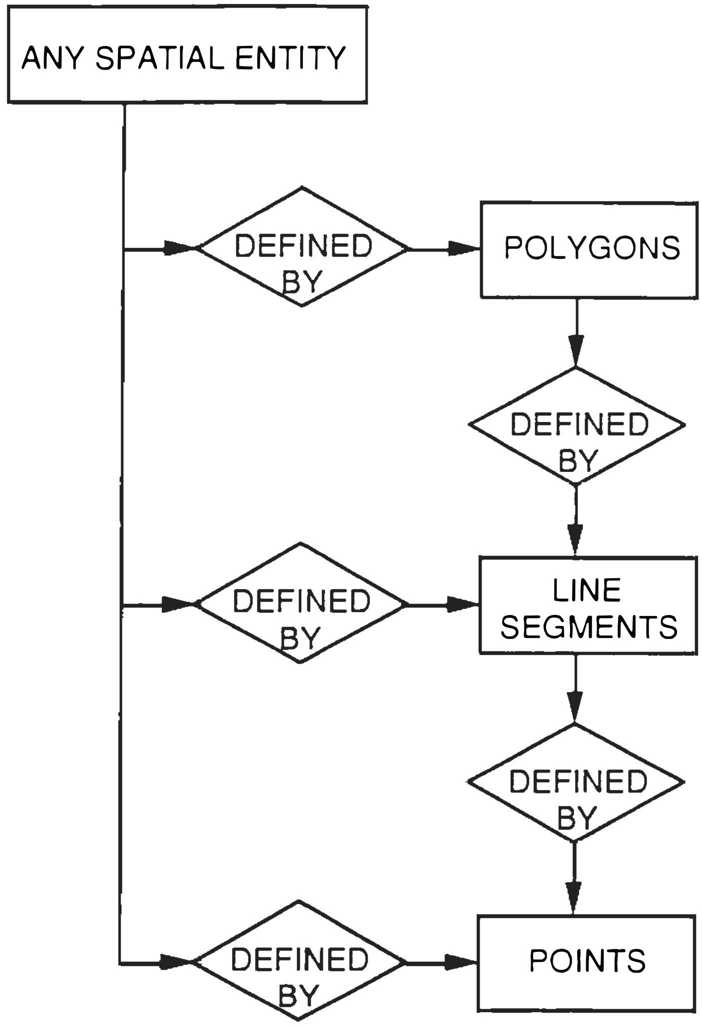



Figure 2.1: Entity-relationship model diagram for basic entity types

location, the point file, without losing the positions on the line segment and polygon entities. Similarly, all the connectivity information is stored in a single location, the line segment file. This storage of similar types of information in a single location eliminates redundancy [Lodwick et al., 1986]. There is no need to store the coordinates of a point in two separate locations just because it is used in the definition of two separate entities.

## 2.2 Additional Entities Defined in Terms of the Basic Entities

While it is possible to describe any feature using the three basic entities described above the task becomes simpler if three additional entity types are defined. The ERM diagram for the extended entities is shown in Figure 2.2 and each of the extended entities is described in the following paragraphs.

The polygon entity type can be extended to include both simple polygons and super polygons. A super polygon is a polygon that is made up of several other polygons, each of which may be a simple polygon or another super polygon. By recursively replacing each super polygon with the definitions of the polygons that form it, a list of only simple polygons will eventually be obtained, which, when looked at together, define the area covered by the super polygon.

An example of a super polygon is the block polygons on a legal survey plan. Each block polygon is made up of several lot and lane polygons, which are all simple polygons. Similarly, the plan polygon is a super polygon that is made up of all the block and road polygons. In this case the block polygons are super polygons while the roads are simple polygons. Since the description of super polygons allows for recursive definition, care must be taken to ensure that the definition is finite. For example, if A, B, and C are each super polygons then one cannot have the situation where A includes B, which includes C, which includes A. A second condition placed on super polygons in this work is that the defining polygons must all be adjacent with no overlaps or

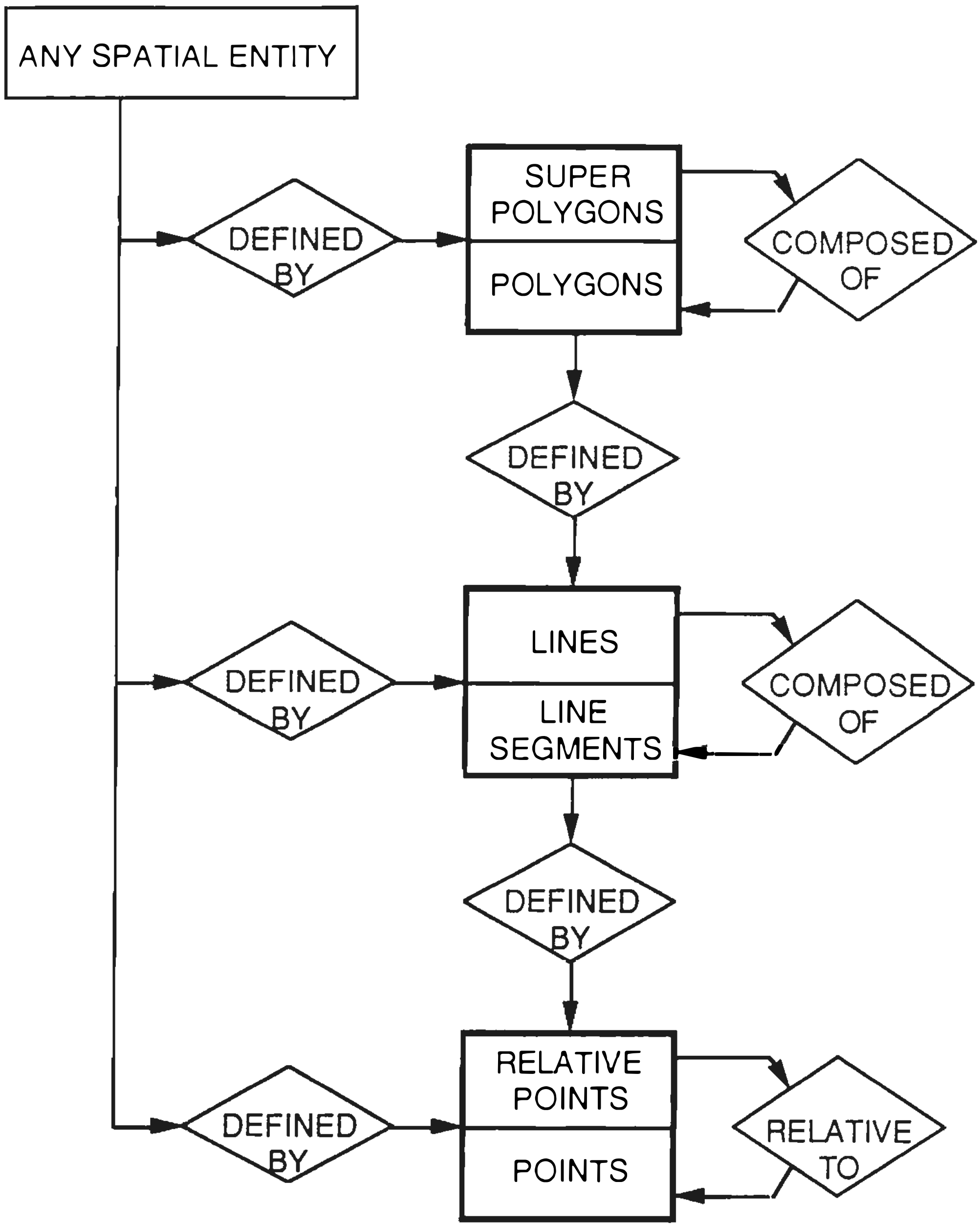



Figure 2.2: Entity-relationship model diagram for extended entity types

holes. In general, this condition is not absolutely essential but it does simplify the processing and manipulation of the data.

In a similar manner to the definition of super polygons, lines are composed of several line segments. An example of this is the line forming a block face, which is composed of several lot faces each defined by a line segment. Lines may be composed of both line segments and other lines. Once again, circular definitions must be avoided. In the definition of a line used in this thesis, lines may not have any branches in them and must be continuous. Lines may or may not have observed values as attributes assigned to them. As an example of this, it is common on legal survey plans to have an azimuth on a block face. If the block face is modelled as a line composed of several line segments, this implies that each of the line segments also has that same azimuth.

The third extended entity type is the relative point. This is a point whose position is defined relative to another point, rather than in terms of absolute coordinates. There are many ways to define the position of one point relative to another. One of the most useful in dealing with legal survey plans is to define a point as lying on a straight line between two points at a certain percentage of the distance from the initial point. By storing the positions of lot corners in this manner, using for reference the points at the ends of the block face, the positions of the lot corners will automatically follow any changes to the positions of the block corners using the method of proportioning. This is the accepted method of "re-positioning" a lot corner from the block corners. In other words, such a storage method automatically implements current standard practice into the data base.

### 2.3 Point Classification

There can be many different types of points shown on a legal survey plan. In this section an attempt is made to enumerate the types of points and to describe the implication of each point type on the entry of the plan into a data base. The classification scheme that is used and discussed here is shown in Figure 2.3.

The initial breakdown of the points is into the categories of marked and unmarked [Mepham, 1987a]. This breakdown is used to differentiate between points that either are, have been, or will be permanently marked on the ground and those that are not, never were, and never will be. Three categories of unmarked points have been identified. The temporary classification is used for points that have been occupied and measured but not permanently marked, such as intermediate traverse points. The theoretical classification is used for points whose coordinates are required but the point has never been actually occupied or measured to. The centre of a curve is an example of a theoretical point. The third category of unmarked points is simple lot corners where the Surveys Act does not require a post and no bar has been or will be placed.

There are two categories of marked points, survey control and legal. The survey control category is reserved for those Alberta Survey Control points that the plan is tied to. These points are used to

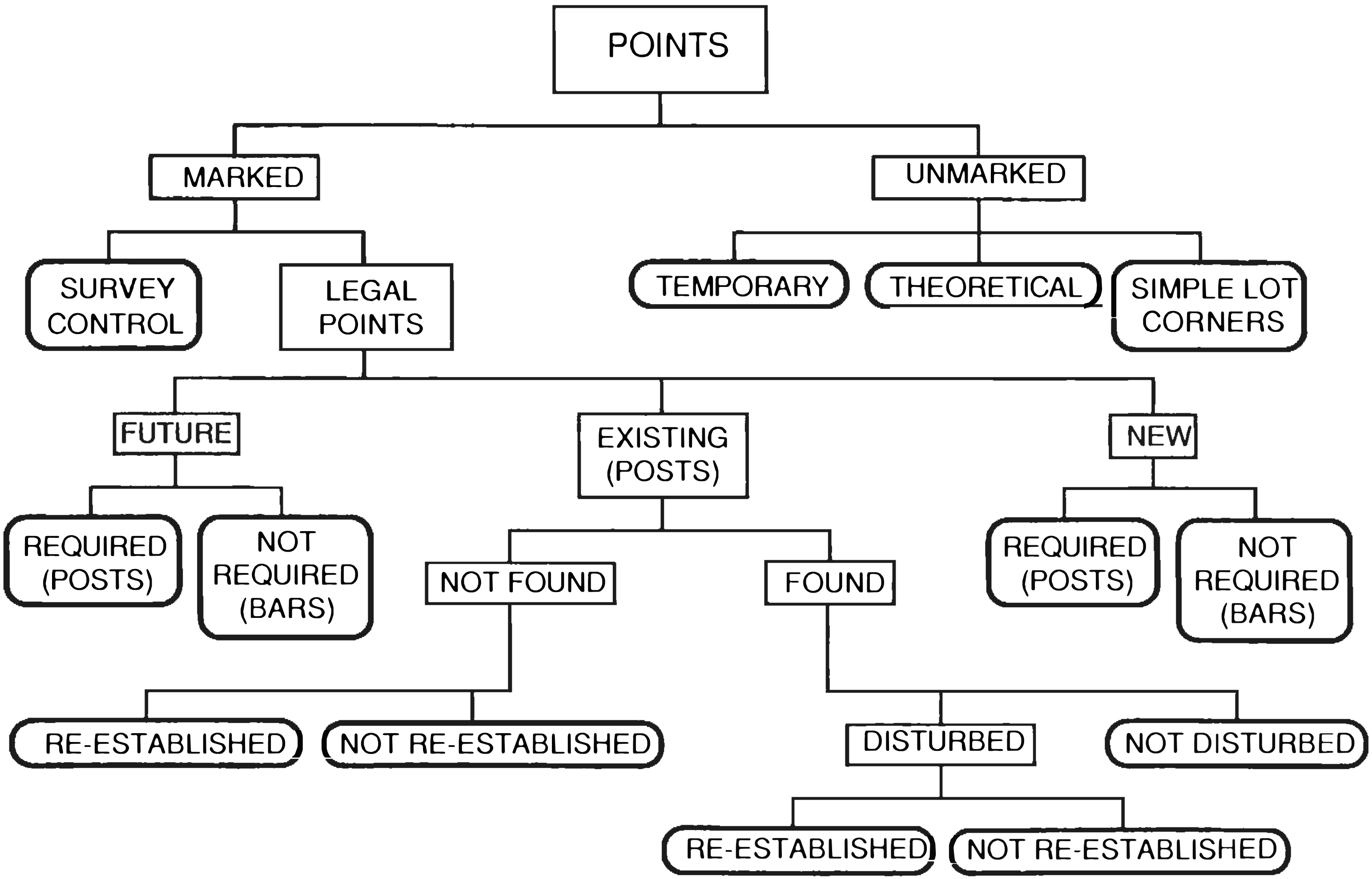



Figure 2.3: Point classification scheme

provide the positional information that is used to compute the coordinates of the points on the legal survey plan. The legal points are those that actually define the boundaries of the lots and blocks of the plan as well as the boundary of the plan itself. The legal category is further divided into three sub-categories. The new points are those that were marked during the survey for the plan. These points may or may not have been required under the Surveys Act and this is recognized by the two classes of new points, required and not required.

The second legal category is the future category. Under section 37(2) of the Surveys Act [Government of Alberta. 1980] the marking of some of the lot and block corners that would normally have to be done at the time of the survey can be delayed for up to one year. The points that fall under this section of the act are classified as future. Again this category can be further subdivided into the required and not required classes.

The final category of legal points is for those that existed before the survey for the plan was carried out. These are points from previous surveys that the new survey was adjacent or close to. The existing points can be classified as to whether or not they were found. The classification of not found is important for two reasons. The first is that under the regulations a tie to the point may have been required. If the point is shown as not found the requirement may be waived. The second reason for identifying these points is that the data base can be kept up to date regarding the presence of the survey marks.

This knowledge is of value to future surveyors working in the area. If the point was not found it may have been re-established. This is recognized by the two classes under the not found category, re-established and not re-established.

The found category under the existing category is subdivided into two categories, disturbed and not disturbed. The disturbed points are further classified as re-established and not re-established. The reasons for this are the same as for the not found categories. Only the positions of points with a classification of not disturbed or re-established can be used in the calculation of the positions of the new and future points. The positions of points in the re-established categories may not be as trustworthy as those in the not disturbed category.

A total of thirteen different point types, identified by the bold boxes in Figure 2.3, results from this classification system. It is thought that this is a complete catalogue of the various point types but, if it is not, the classification system can easily be modified to accommodate other point types.

It should be pointed out here that this classification scheme does not attempt to take the use of the point into account. That is, it does not indicate whether or not the point is a lot corner, block corner or anything else. The use of the point is inferred from the classification of the lines of which it forms an end point. One of the reasons for this is that it simplifies the classification considerably. Using this scheme one point can be a lot corner, a block corner, and a plan corner all at the same time, or it can be none of these. The only point classification that is a function of the point's use is survey control. This could have been avoided by removing the survey control and legal categories from the classification system and putting survey control points under the existing category. This was not done as it was considered that it is important to keep these two types of points separate.

An additional classification of the points that is not shown in Figure 2.3 is the separation of the points into the classifications of on plan and off plan. A point receives one of these two classifications in addition to its classification from Figure 2.3. These two classifications are intended to differentiate between points that are part of the legal fabric of the plan and those that contribute to the positioning of the plan. By definition, all survey control points can be considered as off plan points. In addition, the points that fall into the temporary classification can be considered off plan in all but rare circumstances. All other types of point are classified as on plan points, with the exception of existing points that do not form part of the boundaries of the plan, but were observed for the purpose of positioning the plan with respect to other, existing plans. Note that referring to a point as on plan does not imply that it is within the boundaries of the plan. Curve centre points may, for example, lie outside the boundaries of the plan but they are still considered on plan points. Similarly, survey control points that happen to lie within the boundaries of the plan are still classified off plan.

## 2.4 Line Classification

In contrast to points, there are only five types of line that appear on a survey plan. The classification tree for these lines is shown in Figure 2.4. The remainder of this section is a description of the line classifications.

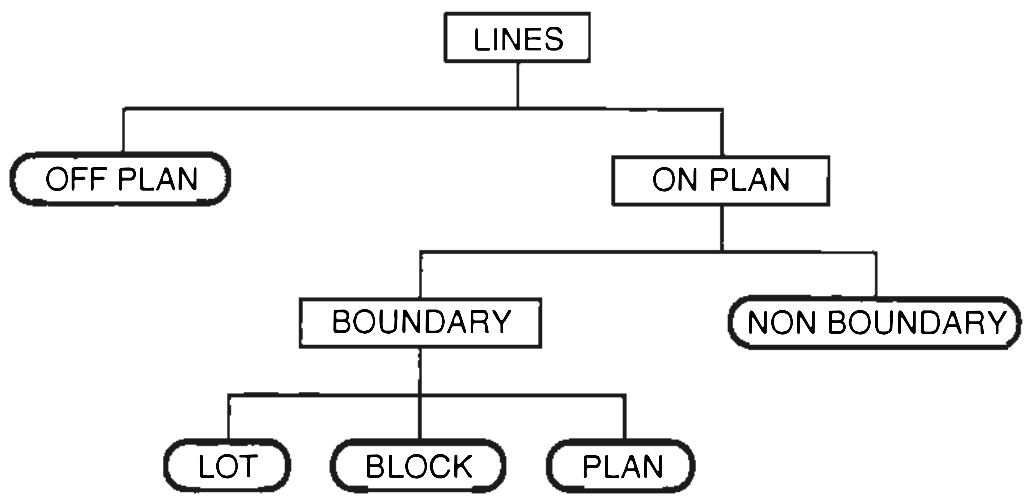



Figure 2.4: Line classification scheme

The initial breakdown of lines is into the on plan and off plan categories. These categories have the same meanings as described in the point classifications described in the previous section. Any line with at least one end point classified as off plan is classified as an off plan line. These lines do not form part of the legal fabric of the plan. The on plan lines are those which have endpoints that are both on plan points. This category of lines is further broken down into two subcategories, boundary and non boundary.

Non boundary lines are those lines which, while they form part of the legal fabric of a plan, do not form part of a boundary of a lot, block, or plan. Examples of non boundary lines include curve radius lines, offset lines, and curve tangent lines. These lines are required for the description of the plan. The boundary lines are just that. They are the lines that actually form the boundaries of the lots, laneways, blocks, roads, and the plan itself. The boundary lines are subdivided into three nonexclusive classes, lot, block, and plan according to the type of boundary of which they form a part. The nonexclusivity of these last three classes results from the fact that a single line can be used as a boundary for more than one part of the plan. It is not unusual for lot, block, and plan boundaries to be coincident at the edges of a plan.

It should be noted that no line may have as its end points either of the two not re-established categories of points. Connection to these points implies occupation of these points, whose classification indicates that they no longer exist. This is clearly impossible. The end points of the lines classified as block or plan boundaries must be classified as required marked legal points as do the end points of lot boundaries that are not identified as being points on line, as described in Section 3.1 of this thesis.

## 2.5 Polygon Classification

The classification of the polygons is very simple as shown in Figure 2.5. The entire plan forms a single polygon at the first level. This super polygon covers all the land that is affected by the registration of the plan. The area within the plan polygon is subdivided into two categories, block and road. The road polygons are simple polygons while the block polygons are super polygons which are further subdivided into the two categories, lot and laneway, both simple polygons.

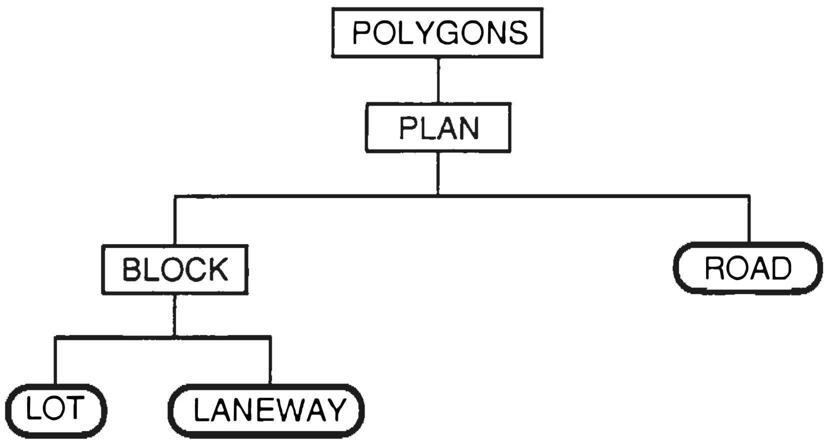



Figure 2.5: Polygon classification scheme

### 2.6 Summary

The classification schemes presented in this chapter have been specifically designed for this research and allow all the points, lines, and polygons shown on a cadastral survey plan to be classified. These classifications take on importance in checking for errors during the data entry process and in the definition of rules to be used in the detection of topological conditions.

# Chapter 3

## TOPOLOGICAL CONSTRAINTS

In this chapter the use of topological constraints on numeric calculations is presented. Employing these constraints makes the use of a least squares adjustment for the calculation of the coordinates and analysis of the data a feasible proposition. A re-examination of Figure 1.1 shows that lots 4 through 10 have no direct connection to any other parts of the plan. Using the approaches described in Section 1.2.2 these lots would probably be positioned by adding a 20 foot line crossing the laneway on either side of the road, as part of a traverse run around the periphery of the road. The positions from this traverse would then be used to position the remaining points in these lots.

The results from this type of approach would be very weak in terms of the accuracy and reliability of the results. The coordinates of the points in the south west corner of lot 6, for example, would probably not be 20 feet from the opposite edge of the laneway. This discrepancy is due to the circuitous route required to run a traverse from the points in lot 6 to the point on the opposite side of the laneway, with no direct connection between the southwest corner points in lot 6 and the points defining the edge of the laneway, only 20 feet away. Topological constraints allow incorporation of the fact that the laneway is 20 feet wide, for example, into the coordinate computation. thus ensuring that the southwest corner points in lot 6 are 20 feet from the opposite edges of the laneway.

The use of topological constraints does more than improve the quality of the coordinates computed for the points on the plan. They also help to maintain the fidelity of the digital representation of the plan. The subdivision designer intended that certain lines be straight or curved, and that sets of lines be parallel. By applying topological constraints at the plan coordination stage, these design elements are maintained. A further advantage of using topological constraints is that they increase the reliability of the coordinate computation. Each constraint equation added to the least-squares adjustment increases the redundancy of the equation system by one.

It should not be the responsibility of the person entering the data to identify all the special conditions that exist on the plan. Doing so is inviting errors of omission. Particularly in large plans, it is doubtful that the operator will consistently find and identify all the special conditions. For this reason, it is important that the detection be performed automatically. A further benefit of automatic detection is that missing data frequently result in the false identification of a special condition. This false identification serves to inform the operator that either data are missing, or there is something abnormal on the plan.

The remainder of this chapter describes the types of constraints that may exist on a legal survey plan and how they are detected. The mathematical modelling of the constraints and their inclusion into the coordinate computation is described in Chapter 5.

### 3.1 Straight Line Conditions

Probably the most frequently occurring condition leading to a topological constraint is that of a series of points lying on a straight line. This is an extremely important condition. The lot corners that lie on the outside of a block must be coincident with the lines defining the boundaries of the block. Without the inclusion of a straight line constraint the individual line segments forming the fronts of the lots may not be coincident with the block boundary line. By imposing the condition that the lot corners must be on the block line, this inconsistency can be eliminated. Figure 3.1 illustrates a portion of a

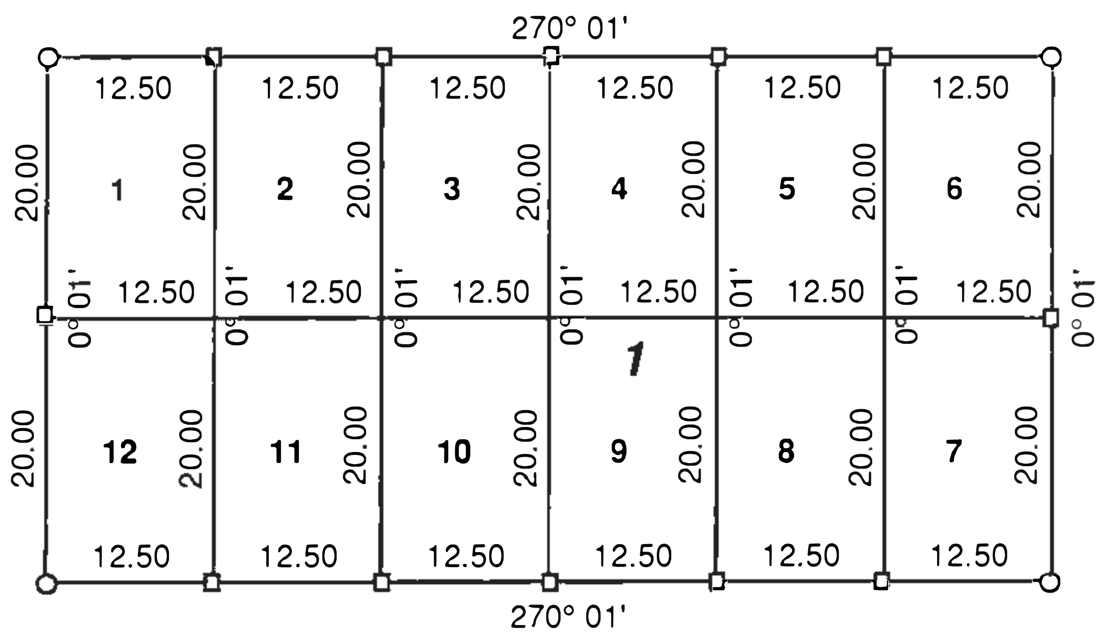



Figure 3.1: Straight line condition

plan where the point on line condition needs to be applied to the lot corners along the top and bottom boundary lines. It should be noted that the point on line condition does not only apply to lot boundaries that are coincident with block boundaries. Frequently the lines within the blocks defining the boundaries of the lots are also straight lines with several lots connected to them. Examples of this are all the vertical lines shown in Figure 3.1, as well as the horizontal line running across the middle of the block.

All of the azimuths shown in Figure 3.1 are given for lines, not line segments. This is the principal clue used in the detection of the straight line conditions. If a line has a given orientation, then all the line segments that form that line must have the same orientation, and thus must all be portions of the same straight line. The process for the determination of what points lie on a common straight line thus becomes:

a) Find a line segment with an observed orientation but no observed distance;

b) Find the shortest path between the end points of the line segment through the sub-network, formed by selecting all of the line segments with observed distances and no orientation, and the link lines;

c) Highlight the chosen line segments on the graphics screen and have the operator verify that these do in fact form a straight line. If the operator rejects the proposed line, there is probably an error in the data entry;

d) If the operator accepts the proposed line then redefine the line segment with the observed orientation as being a line formed by the line segments with the observed distances that form the shortest path.

This process is repeated until all of the straight lines with orientation have been found and their associated points on line identified.

This method will find most of the straight line conditions on a plan but will not find them all. Using Figure 3.1 as an example, this method will detect all the straight lines except one. The line forming the backs of the lots does not have an azimuth shown for it. As a result, it will not be found in step (a) above. The detection of these straight lines (no orientation given) is done at the same time as the detection of the next special condition, that of circular curves.

### 3.2 Circular Curve Conditions

In addition to straight lines, cadastral survey plans also show circular curves. The detection of a series of line segments as all being part of a common circular curve is just as important as the detection of the straight line conditions. Not only is it required to ensure that the block and lot boundaries coincide, but it is required so that the curve distances that were entered earlier can be marked as arc distances. During the observation entry stage (described in Chapter 6) no distinction is made between straight line and arc distances.

The detection of circular curve conditions in the data for a plan starts by finding a connected series of line segments, that have distance observations on them but no orientation, and are not part of a straight line condition. The top line in Figure 3.2 illustrates this situation. Unfortunately the middle line of the figure also fits this criteria despite the fact that it is a straight line, not a curve. This is the situation mentioned in the previous section for the straight line condition detection being incomplete.

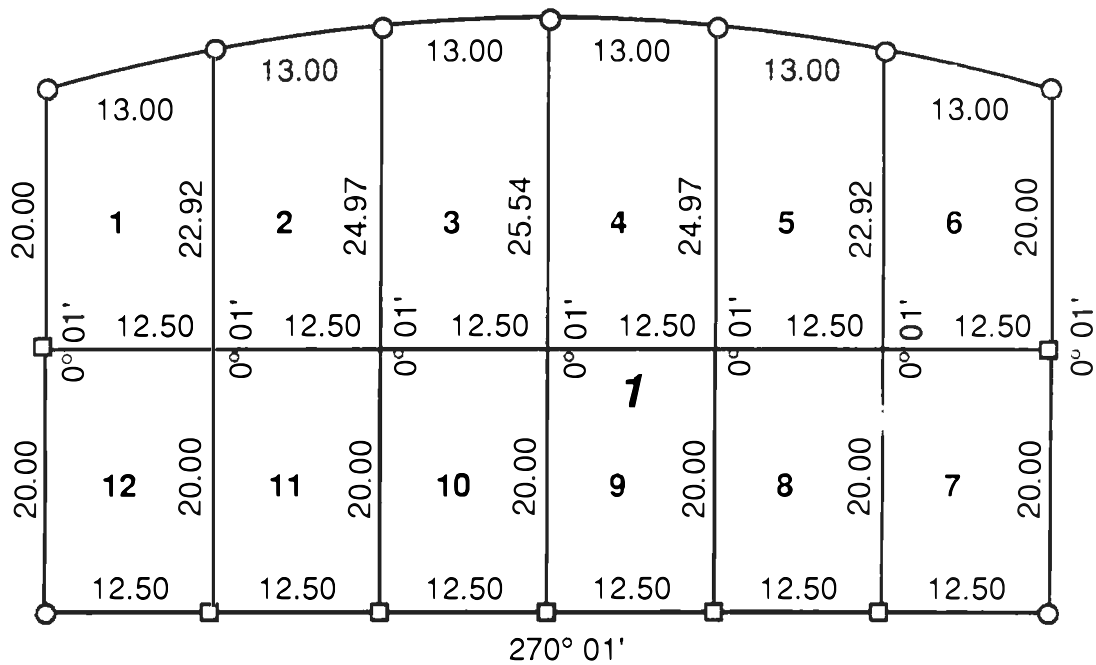



Figure 3.2: Circular curve condition

What is needed now is some way to determine whether the series of line segments is part of a curve or part of a straight line. Section 36 of the Surveys Act states

> "A surveyor shall mark, in all blocks that have curvilinear boundaries, the points of beginning and end of each curve having a constant radius and shall mark all corners of each lot on the curvilinear boundary with iron

posts and where the rear of the lot is bounded by a straight line or lines an iron post shall be planted at each change of direction." [Government of Alberta, 1980].

In other words, all corners along curved lines must be posted. Advantage may be taken of this fact in determining whether or not the line found is a curve or a straight line. This is done by checking the classification of all the points along the line. If all of them are classified as either future required, new required, or existing posts then the line is probably a curve. If any of the points are not posts, then the line cannot be a curve and is either a straight line or there is an error in the data or on the plan. It is possible to have a straight line with all the points posted. For this reason the operator must be allowed to indicate that the line segments form a straight line, not a curve. The opposite is not true. If any of the points are not posted the line segments may not form a curve and, if such is indicated, then the plan must be rejected.

Another situation where this detection process may fail is that of two tangent curves as illustrated in Figure 3.3. There is no way to determine, from the data that have been entered, that the series of line segments represents two tangent curves. The operator has to provide this information to the system. This is accomplished by simply pointing to the start and end points of a single curve within the line segments highlighted on the graphics screen. All the other line segments are then removed from contention.

Once a circular curve has been identified, further information

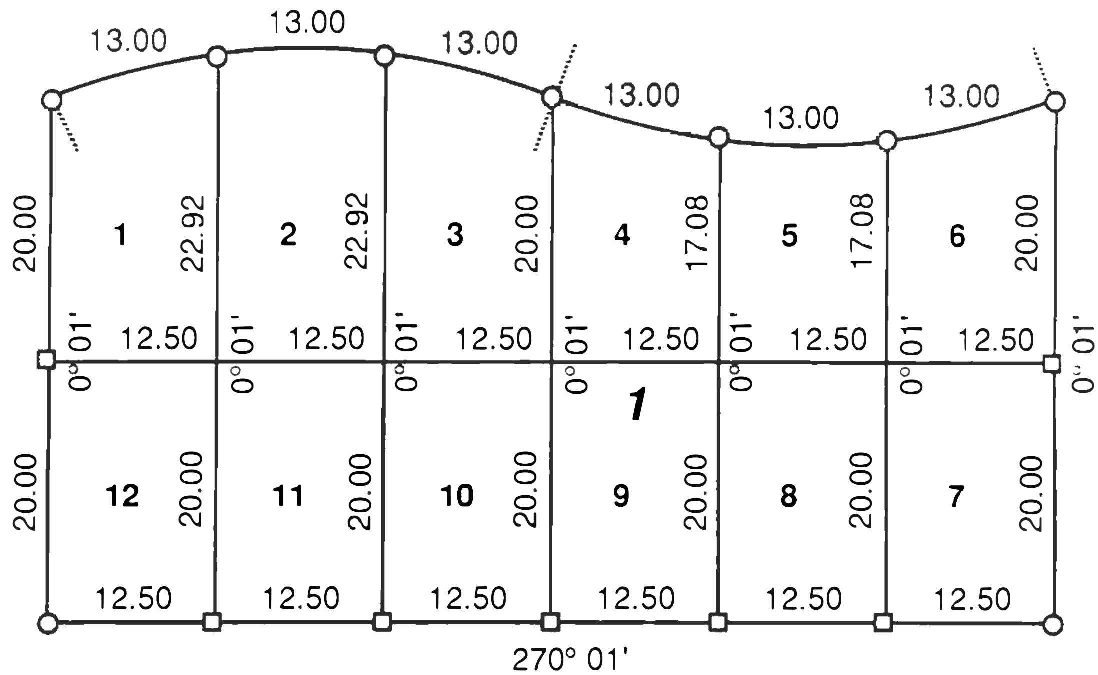



Figure 3.3: Tangential curves condition

must be entered for it to be fully defined. The additional information required is;

i) The centre point,

ii) The radius, and

iii) The delta angle.

For any given curve the radius and/or delta angle may be missing, so provision must be made for them to be entered as "not shown". The centre point must be provided for all circular curves whether one is shown on the plan or not. The centre point is needed during the model formation discussed in Chapter 5. The identification of the centre point is simply a matter of pointing to the centre point on the plan. This pointing will not be confused with the start/end pointing mentioned earlier, because it involves pointing to a point that is not part of the highlighted line segments.

### 3.2.1 Straight Line Editing

The straight line detection technique described in Section 3.1 will always return the complete straight line and nothing else (in the absence of errors in the data). This is not true for lines found during the circular curve detection process. The lines found here may require editing. An illustration of a situation where this may occur is shown in Figure 3.4 which is a detail from Figure 1.1. The line

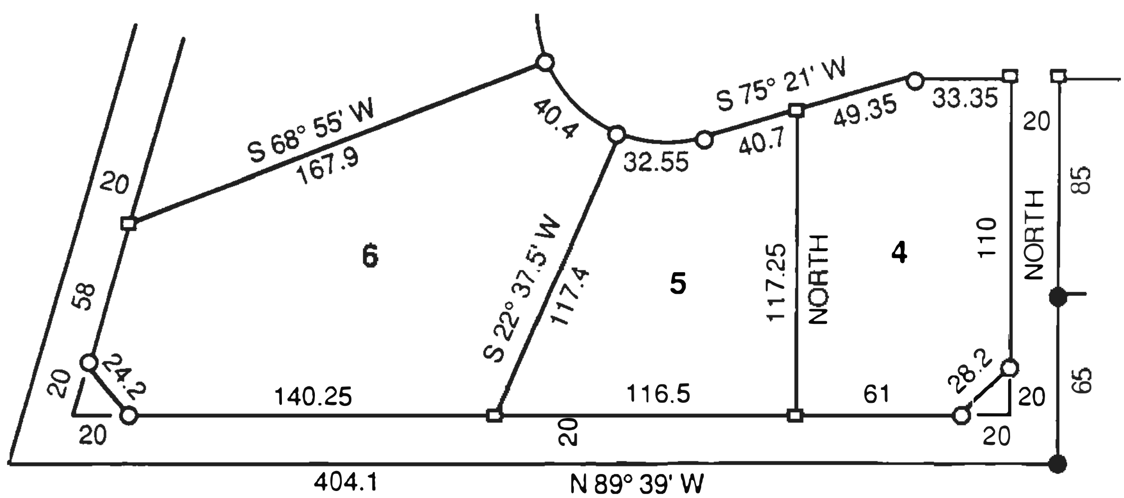



Figure 3.4: Ambiguous straight line condition

segments which form the backs of lots 4, 5, and 6 would be detected as a straight line during the circular curve detection process. It is obvious from the examination of the figure that the straight line should include the two theoretical points defining the corner cuts as its ends. The algorithm, as described above, may not, however, include the 20 foot line segments as part of the line. The "diagonals" of the corner cuts (the 24.2 and 28.2 foot line segments) are, according to the algorithm, just as valid as candidates.

In order to allow the operator to correct any misidentification resulting from the detection algorithm, a line editor has been provided. Editing is accomplished by pointing to points on the plan using the digitizer tablet. If the indicated point is part of the current line definition, then the line segments having it as one of their end points are removed from the line definition. This results in one of two situations. Either the remaining line definition will define one line, if the segments removed were at the end of the line, or it will define two lines, if the segments removed were in the middle of the line. In the latter case, an attempt is made to bridge the gap by finding and inserting a single line segment. If no such line segment exists then the request to remove the line segments connected to the indicated point is ignored. This procedure keeps the line segment properly defined at all times.

If the point that was indicated is not part of the current line definition then it is added to the line definition. This is done by finding the point used in the current definition of the line that is closest to the indicated point. Line segments connecting the indicated point to the closest point, and to one of the points connected to the closest point, are then inserted into the line definition. This implies the removal of a line segment from the definition, unless the new line segment is added at the end of the line.

## 3.3 Parallel Line Conditions

Many of the lines shown on a survey plan are parallel. This applies to both straight and circular lines. This parallelism can be used to advantage in the coordination of the survey plan, as was explained in the introduction to this chapter. The problem is in deciding what lines to force parallel to each other. In a rectangular subdivision it is possible to declare that each line is parallel to half the other lines on the plan. This is obviously somewhat excessive. The question comes down to determining when a parallel line condition needs to be enforced. According to Baker [1987] the lines that must be kept parallel are the opposite sides of rights of way. For this reason, and to avoid imposing an excessive number of conditions on the coordinate computation, the parallel line conditions are restricted to opposite sides of the rights of way only.

The detection of parallel (concentric) curves is very simple. Any two circular curves that form part of the boundaries of a right of way, and also have the same centre point, must be parallel curves. It is the detection of parallel lines that is somewhat more complicated. Examining Figure 1.1 reveals that there are four east-west lines used in defining the boundary of the laneway that surrounds lots 4 through 10. These four lines form two pairs of parallel lines, with the problem being to determine the pairings. The process used to do this is as follows:

a) A list is built of all the straight lines used in the definition of the right of way polygon that have not been marked as parallel to any other lines yet;

b) The first line in the list is checked against all the other lines in the list to see if their extents overlap. (Any lines previously labeled unacceptable in step (c) are skipped as are any lines that have an observed azimuth that differs from the observed azimuth on the first line, if there is one.) The determination of overlapping extents is done by checking to see if the perpendicular at either of the end points intersect the other line between its end points. (This process is described in detail in Section 5.1.) If the lines do overlap, then the perpendicular distances from the end points of each line to the other are computed and the mean of these four values calculated. If this value is less than for the current parallel line candidate, or this is the first overlapping line, then this line is made the candidate;

c) If a candidate for a parallel line to the first line in the list was found then the two lines on the screen are highlighted and a confirmation obtained from the operator that the two lines are indeed parallel. If the operator rejects the selection, then the candidate line is temporarily labelled as unacceptable and step (b) repeated, otherwise the two lines are marked as parallel;

d) The first line and its corresponding parallel line (if one was found) are removed from the list. If there are still two or more lines in the list, any unacceptable labels are deleted from the remaining lines and the process continued at step (b). Otherwise the parallel line determination for this right of way is complete.

This process is repeated for each right of way polygon that exists. It should be noted that, in building the line list, any line segment that is found to be part of a straight line should be replaced by the straight line that it forms a part of. This process requires knowledge of the definition of the polygons that make up the legal survey plan. The process of building the polygons requires knowledge of the straight line conditions but not the parallel line conditions (see Chapter 4). For this reason the processing flow for the definition of the topological constraints is split up. The straight line and circular curve conditions are found before the polygon formation, and the parallel line conditions are found afterwards.

One of the attributes of a right of way is its width. Since the place where the width of a right of way is constant is that portion bounded by parallel lines, the width of the right of way should be entered whenever two lines are accepted as parallel. This applies to both straight and circular lines. The width can be treated as an observation and included in the coordinate computation. In fact, as explained at the beginning of this chapter, this is the manner in which blocks 4 through 10 of the plan shown in Figure 1.1 are positioned.

## 3.4 Summary

The rules developed in this chapter form a critical component of the cadastral survey plan data entry system. The information gained by applying these rules during the data entry process can be stored in the topologically complete Land Information Systems that are currently in development and production. Not all systems available today are capable of storing and processing this information but this should not stop its collection. This information is additional to the spaghetti data that are currently collected and processed by many systems. Their collection does not prevent the use of the point and line data in the existing spaghetti systems. By collecting it now, and storing it for the future, the data will be ready for eventual use in upgraded systems without the additional cost of re-collecting it.

A second reason for collecting these data now is the sensitivity of the rules to missing data. If a distance or azimuth is missed during the data entry process, the results of applying these rules will result in incorrect identification of straight lines and curves. This is a powerful indication to the operator that there is missing data, and the error can be corrected immediately, before the data base has been updated. This one reason alone is enough to justify the small added expense of applying these rules. Finally, there is also the improved accuracy and reliability of the coordinates computed for the points on the plan when this information is included in the calculation process. This will be discussed more fully in Chapter 5.

The topological conditions identified here are not the only ones that are of importance. The identification of the polygons that model the lots, blocks, roads, etc. shown on the cadastral survey plan is of critical importance. This aspect of the topology of a cadastral plan is discussed in the next chapter.

# Chapter 4

## POLYGON FORMATION

Chapters 2 and 3 have described the types of entities that may exist on a cadastral survey plan and the process of finding the topological conditions that exist and will be used in the coordinate computation stage. In this chapter the determination of the polygons that form the plan is described. It must be remembered that the purpose of any legal subdivision of land is to divide the land into parcels that can be dealt with as individual units. Most of the processing of the automated legal survey plan after its entry and acceptance will involve the parcels, not the individual points and lines. These parcels are modelled by the polygons formed using the techniques described in the following sections.

In all the polygon definitions that follow, the definition is strictly in terms of line segments. Any time a line is found as a candidate during polygon formation it is replaced by its constituent line segments. The reason for this is to simplify the making of changes to the polygon definition at a later date. The primary evidence of the location of a property corner is the pin in the ground. If, at some later date, a re-survey of the area covered by the plan or adjacent to the plan, finds that one of the pins is in an incorrect location, then the new position of that point should be registered in order to make the data base better reflect reality. If the point was used to define a line segment that was part of a straight line, for example, then the definition of the line must be changed to exclude the line segment(s) defined by that point. If polygons are defined only in terms of line segments then the update is complete. If not, then all polygons that use the straight line as part of their definition must also be updated.

## 4.1 Simple Polygons

There are three simple polygon types on a legal survey plan. They can all be formed using the same approach and then differentiated by assigning different polygon type codes to them. These three polygon types are the lots, lanes, and roads. Algorithms do exist that will form all the simple polygons that form a plan with no operator intervention. These algorithms are not used in this project for the following reason. Using any algorithm, the process of adding attributes, such as polygon type and identifier, will require that the operator somehow point to, or otherwise indicate, what polygon should receive which attribute.

This requirement that the operator identify the polygons individually is taken advantage of in the algorithm developed as part of this project. The algorithm, shown in Figure 4.1, is as follows:

a) The operator digitizes a spot inside the simple polygon to be formed. The actual location of the spot is immaterial so long as it lies inside the polygon.

b) The boundary type line segment closest to the digitized spot is found and made the first line segment in the polygon definition. (This step is explained in more detail in Section 4.1.1.)

c) The line segment connected to the end of the line segment just added, that forms the minimum counter-clockwise angle to the just added segment, is made the next line segment and added to the polygon definition.

d) Step (c) is repeated until the polygon closes, at which time the polygon formation is complete.

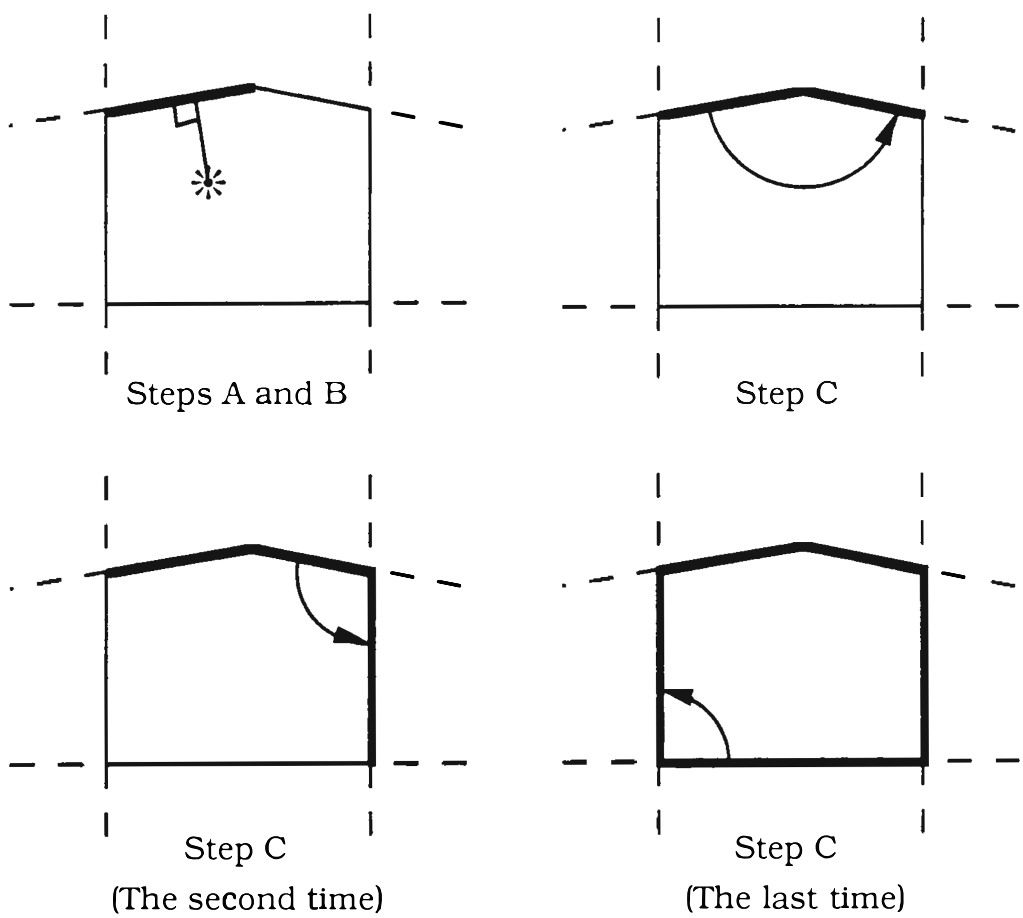


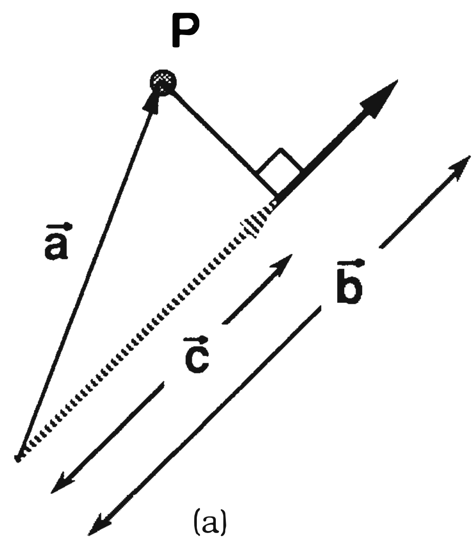

Figure 4.1: Simple polygon formation

e) The operator is asked to accept or reject the polygon just formed. If it is accepted then the attributes for the polygon are requested and recorded. If the operator rejects the polygon then there is an error in the data or on the plan.

This process is repeated for each simple polygon on the plan until all polygons they have been formed and saved. As each polygon is saved, the line segments used to form it are marked to show what polygons the segments have been used in, and the type of polygon. This information will be used in the forming of the block and plan polygons. It should be noted at this point that only two simple polygons may use the same line segment in their definition. For more than two simple polygons to share a single line segment would imply that the polygons are overlapping, a violation of the definition of simple polygons.

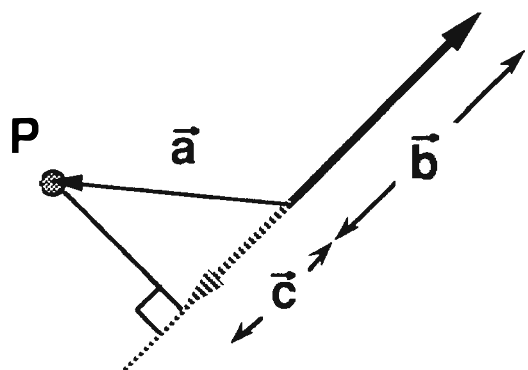
### 4.1.1 Finding the Closest Line to a Point

Determining what line segment is closest to a given point is a relatively complicated process. Before this can be determined, the concept of "line segment closest to a point" must be defined. What is needed is to decide to what location on the line the distance to the point is to be measured. In this case the perpendicular distance from the point to the line segment is used (Figure 4.2a). This raises the second question of what to do with line segments that do not have a perpendicular distance to them, as in Figures 4.2b and 4.2c. In determining lot polygons these lines are not considered as candidates



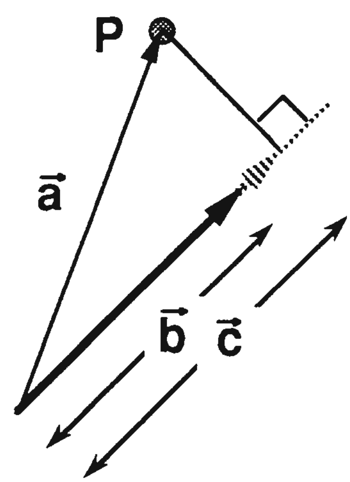

(a)

Perpendicular intersects line segment $\vec{b}$




(b)

Perpendicular intersects behind line segment $\vec{b}$




(c)

Perpendicular intersects ahead of line segment $\vec{b}$

Figure 4.2: Relationships between a point and a line segment

for the starting line. Since most line segments on a plan do not have a perpendicular distance from them to the point, they should be eliminated from consideration as quickly as possible.

To do this two vectors, $\vec{a}$ and $\vec{b}$, are defined as shown in Figure 4.2. The projection of $\vec{a}$ onto $\vec{b}$ is shown as $\vec{c}$. The magnitude of $\vec{c}$, ($|\vec{c}|$), can then be computed as

$$|\vec{c}| = \frac{\vec{a} \cdot \vec{b}}{|\vec{b}|} \qquad 4.1$$

[Leithold, 1972]. If $|\vec{c}| < 0$ then the situation shown in Figure 4.2b exists and there is no perpendicular from the point to the line segment. Similarly, if $|\vec{c}| > |\vec{b}|$ the situation shown in Figure 4.2c exists and there is no perpendicular from the point to the line segment. From this it can be stated that a perpendicular exists from a point to a line segment only if the condition

$$0 < \vec{a} \cdot \vec{b} < |\vec{b}|^2 \qquad 4.2$$

holds. The perpendicular distance, s, from the point to the line is given by the formula

$$s = \sqrt{|\vec{a}|^2 - |\vec{c}|^2} \qquad 4.3$$

which uses all the values computed during the elimination checking. The elimination checking, therefore, involves no extra computations over and above those required for the calculation of the perpendicular distance from the point to the line segment.

The next step in the formation of the lot polygon is to use the end points of the line segment chosen as the first two points in the lot polygon. In order to maintain the clockwise rule of coding polygons [Baxter, 1976], one must know whether the point is to the right or left of the line segment. If it is to the right, then the points are added in the order they define the line segment, otherwise they are added in reverse order. The determination of what side of the line the point lies on is carried out as follows:

a) Extend the two dimensional vectors $\vec{a}$ and $\vec{b}$ to the three dimensional vectors $\vec{a_3}$ and $\vec{b_3}$ with the Z values set to zero. This makes the vectors lie on the XY plane.

b) The cross product $\vec{c_3} = \vec{a_3} \times \vec{b_3}$ will have zeros for its X and Y elements and a non-zero value for its Z component.

c) If the Z component value of $\vec{c_3}$ is positive, the point is to the right of the line; if it is negative, the point is to the left; and if it is zero, the point is on the line.

The reasoning behind this method is as follows, Rees and Sparks [1969] show that

$$\vec{a_3} \times \vec{b_3} = \vec{n} \cdot |\vec{a_3}| \cdot |\vec{b_3}| \cdot \sin\theta \qquad 4.4$$

where: $\theta$ is the angle measured from $\vec{a_3}$ to $\vec{b_3}$, and
$\vec{n}$ is a unit vector perpendicular to $\vec{a_3}$ and $\vec{b_3}$.

Here $\theta$ must lie between 0 degrees and 90 degrees or 270 degrees and 360 degrees since the perpendicular from the point to the line segment intersects with the line segment only for values of $\theta$ in that range. In addition, the vector $\vec{n}$ must have the value [0, 0, 1] since $\vec{a_3}$ and $\vec{b_3}$ lie on the XY plane. The norms of the vectors $\vec{a_3}$ and $\vec{b_3}$ are positive, by definition, so the Z component of $\vec{c_3}$ can be negative only if sin θ is negative. This implies that θ is in the range of 270 degrees to 360 degrees, showing that the point must be on the left side of the line segment, if the cross product has a negative Z element.

## 4.2 Plan and Block Polygons

The plan and block polygons are super polygons composed of other polygons. The block polygons are defined by groups of lot and lane polygons, while the plan polygon is defined by all the block and road polygons. The formation of the block polygons is a simple matter if the block identifiers are entered when the lot and lane polygons are formed. In this way the block polygon is built up at the same time as the lot and lane polygons. The plan polygon is formed after all the block and road polygons have been defined.

The definition of these two types of polygon could be fully automatic without any operator intervention at all. This is not recommended. By forming and displaying the block and plan polygons for the operators acceptance, an extra level of error checking is provided. In particular, this will help to ensure that all of the simple polygons are identified. Any lots, lanes, or roads that were not identified at the simple polygon formation stage will appear as holes or gaps in the block and plan polygons.

## 4.3 Summary

This concludes the identification of the entities and the topological conditions that exist on a cadastral survey plan. The result of the application of the ideas introduced in this and preceding chapters, to the entry of cadastral survey plan data into a Land Information System is a much more complete representation of the plan in digital form than has previously been the norm. The following chapter shows how this process yields not only a more complete modelling of the plan but also a more accurate one.

## Chapter 5

## COORDINATE COMPUTATION AND ERROR DETECTION

In this chapter the methods used for the calculation of the coordinates and the detection of any errors or blunders in the data are described. The method of least squares was chosen for the coordinate computations because it can handle non-homogeneous, inconsistent, overdetermined systems of equations in a simple manner. In addition, the solution arrived at for the coordinates, using the least squares approach, can be statistically analysed as will be seen in Section 5.5. This analysis is the heart of the error detection process.

### 5.1 The Method of Least-Squares

The basic equations underlying and used in the least squares method are well known and so are reviewed here without proof or derivation. The interested reader may find the derivations and statistical basis for the method in any of the many excellent books on the subject, e.g. Hamilton [1964], Mikhail [1976], or Vaniček and Krakiwsky [1982]. The basis of the method of least squares is as follows:

A set of non-linear equations of the form

$$\mathbf{f}(\mathbf{x}) - \mathit{l} = 0 \tag{5.1}$$

where: $\mathbf{f}$ is the vector of equations,

$\mathbf{x}$ is the vector of the unknown parameters (i.e. coordinates), and

$\ell$ is the vector of the observations,

can be approximated using the linear portion of a Taylor's series expansion and re-written as

$$\mathbf{A}\hat{\delta} - \hat{\mathbf{r}} + \mathbf{w} = 0 \qquad 5.2$$

where:

$$\mathbf{A} = \left.\frac{\partial \mathbf{f}}{\partial \mathbf{x}}\right|_{\mathbf{x}^{\circ},\ell} \qquad 5.3$$

$$\hat{\delta} = \hat{\mathbf{x}} - \mathbf{x}^{\circ} \qquad 5.4$$

$$\hat{\mathbf{r}} = \hat{\ell} - \ell \qquad 5.5$$

$$\mathbf{w} = \mathbf{f}(\mathbf{x}^{\circ}) - \ell. \qquad 5.6$$

**A** is referred to as the first design matrix while $\hat{\delta}$ is the vector of corrections to the initial approximations of the solution $\mathbf{x}^{\circ}$. The vector of residuals, $\hat{\mathbf{r}}$, is the estimate of the corrections to the observed values, $\ell$, to give a consistent set of observations, $\hat{\ell}$. The vector of misclosures, **w**, is a measure of the fit between the estimates for the solution, $\hat{\mathbf{x}}$, and the observations, $\ell$. The cap (^) on the various vectors is used to denote that the value is a least squares estimate. The solution for the unknowns is given by:

$$\hat{\mathbf{x}} = \mathbf{x}^{\circ} + \hat{\delta} \qquad 5.7$$

where:

$$\hat{\delta} = -\mathbf{N}^{-1} \cdot \mathbf{u} \qquad 5.8$$

$$\mathbf{N} = \mathbf{A}^T \cdot \mathbf{C}_\ell^{-1} \cdot \mathbf{A} \tag{5.9}$$

$$\mathbf{u} = \mathbf{A}^T \cdot \mathbf{C}_\ell^{-1} \cdot \mathbf{w} \tag{5.10}$$

The $\mathbf{C}_\ell$ matrix is the covariance matrix associated with the observations and allows each observation to be weighted according to its accuracy. Since Equation 5.2 is only a linear approximation of Equation 5.1 the solution must be iterated using the $\hat{\mathbf{x}}$ from the last solution as the $\mathbf{x}^\circ$ for the next solution until the computed $\hat{\delta}$ is less than some, pre-determined, value $\varepsilon$.

In addition to computing the coordinates of the points with these equations, their covariance matrix may also be obtained as:

$$\mathbf{C}_{\hat{x}} = \mathbf{N}^{-1} . \tag{5.11}$$

The estimate of the errors in the observations (the residuals $\hat{\mathbf{r}}$), and their covariance matrix, $\mathbf{C}_{\hat{r}}$, are:

$$\hat{\mathbf{r}} = \mathbf{f}(\hat{\mathbf{x}}) - \ell , \tag{5.12}$$

$$\mathbf{C}_{\hat{r}} = \mathbf{C}_\ell - \mathbf{A} \cdot \mathbf{C}_{\hat{x}} \cdot \mathbf{A}^t \tag{5.13}$$

[Mepham, 1983a]. These computations do not need to be iterated and should not be performed until the solution for $\hat{\mathbf{x}}$ has been completed. Equations 5.2 through 5.13 are used for any least squares adjustment of systems described by mathematical models with the form of Equation 5.1. It is the mathematical models which vary from problem to problem.

## 5.2 Mathematical Models

The mathematical models for distance, azimuth, angle, and direction observations are standard models used in surveying. They are given here without derivation or proof. The mathematical model of a horizontal distance observation made between points i and j is:

$$\mathbf{f}(\mathbf{x}) - l = \sqrt{(N_j - N_i)^2 + (E_j - E_i)^2} - s_{i,j} = 0 \qquad 5.14$$

where: $N_i$ and $E_i$ are the coordinates of point i,

$N_j$ and $E_j$ are the coordinates of point j, and

$s_{ij}$ is the observed distance.

The mathematical models for azimuths, angles, and directions measured in the horizontal plane are:

$$\mathbf{f}(\mathbf{x}) - l = tan^{-1}\left(\frac{N_j - N_i}{E_j - E_i}\right) - \alpha_{i,j} = 0 \qquad 5.15$$

$$\mathbf{f}(\mathbf{x}) - l = tan^{-1}\left(\frac{N_k - N_i}{E_k - E_i}\right) - tan^{-1}\left(\frac{N_j - N_i}{E_j - E_i}\right) - \theta_{i,j,k} = 0 \qquad 5.16$$

$$\mathbf{f}(\mathbf{x}) - l = tan^{-1}\left(\frac{N_j - N_i}{E_j - E_i}\right) - \Omega - d_{i,j} = 0 \qquad 5.17$$

where: $\alpha_{ij}$ is the observed azimuth from point i to point j,

$\theta_{ijk}$ is the observed angle at point i backsighting point j and foresighting point k,

$d_{ij}$ is the observed direction from point i to point j, and

$\Omega$ is the orientation unknown associated with the round of directions of which $d_{ij}$ is a member.

## 5.3 The Method of Constrained Least-Squares

In Chapter 3 several topological constraints were identified that need to be applied to the coordinate computation. In this section the extension of the least squares equations to incorporate constraints is presented, as well as the mathematical models used to impose the constraints.

Systems of equations with the form shown in Equation 5.1 may be augmented by a second set of functions with the form

$$\mathbf{f_c}(\mathbf{x}) = 0. \qquad 5.18$$

These functions, known as constraint functions, describe relationships between unknown parameters only, with no observations involved. These relationships are exact, that is, if the function states that two lines must be parallel they will be exactly parallel after the solution. If this requirement of parallelism is contradicted by the observations in the original mathematical model, then the residuals on the observations will reflect the misfit between the constraints and the observations. In addition, care must be taken to ensure that the constraints do not contradict each other. If, for example, a constraint is imposed that point B must lie on a straight line between points A and C and another constraint is imposed forcing point B to be on a circular curve between points A and C, an impossible situation, the system of equations will have no solution.

When constraint functions are included, the equations for solving the system become [Vaniček and Krakiwsky, 1982]:

$$\hat{\delta} = -N^{-1}\left\{ A_c^t \cdot M_c^{-1} \cdot \left( w_c - A_c \cdot N^{-1} \cdot u \right) + u \right\} \qquad 5.19$$

where:

$$A_c = \left. \frac{\partial f_c}{\partial x} \right|_{x^\circ} \qquad 5.20$$

$$M_c = A_c \cdot N^{-1} \cdot A_c^t \text{ ,and} \qquad 5.21$$

$$w_c = f_c(x)|_{x^\circ} \qquad 5.22$$

All of the remaining terms are the same as for the case of unconstrained least squares. The computation of the residuals, $\hat{r}$, and their covariance matrix, $C_{\hat{r}}$, remains the same and the covariance matrix for the unknowns, $C_{\hat{x}}$, becomes:

$$C_{\hat{x}} = N^{-1} - N^{-1} \cdot A_c^t \cdot M_c^{-1} \cdot A_c \cdot N^{-1}. \qquad 5.23$$

In addition, the degree of freedom (or redundancy) of the adjustment, ν, is computed as:

$$\nu = (n + n_c) - u \qquad 5.24$$

where: $n$ is the number of observation equations,

$n_c$ is the number of constraint equations, and

$u$ is the number of unknown parameters.

These solution equations apply the constraints to the unknowns simultaneously with the solution using the observations. Alternate solution methods are available that employ a sequential technique.

The unknowns are first solved for, using the unconstrained least squares approach and then, once the initial solution converges, the constraint equations are used to compute a final update to take the constraints into account. These methods were not chosen here for two reasons. The first is that both methods are equally complex in terms of computations. They both require the same number and size of matrix inversions so there is no advantage in terms of efficiency.

The second reason is that the sequential approach assumes that there are no errors or blunders in the data used in the solution of the unconstrained least squares data. The constraint update is not iterated, on the assumption that the solution from the first step is very close to meeting the constraints. This assumption cannot always be accepted for the cadastral data entry project, since one of the objectives of the project is to find any errors in the data. The plans are not always error free. The errors in the data cannot always be found on the basis of the unconstrained adjustment only and so, the simultaneous solution is the approach taken.

## 5.4 Mathematical Models for the Topological Constraints

In Section 5.1 the mathematical models for the standard surveying observables were presented without any derivation or explanation. In this section the models for the topological constraints detailed in Chapter 4 are presented and fully explained. These models take the form of both Equations 5.1 and 5.18, that is, observation equations and constraint equations.

### 5.4.1 Straight Line Model

The purpose of the straight line model is to express in, mathematical terms, the fact that point C lies on a straight line between points A and B as illustrated in Figure 5.1. If point C is on the straight line then the line segments AC and CB must have the same slope. This can be expressed as:

$$\frac{E_C - E_A}{N_C - N_A} = \frac{E_B - E_C}{N_B - N_C} \tag{5.25}$$

which leads to the mathematical model:

$$\mathbf{f_c}(\mathbf{x}) = \left[(N_B - N_C) \cdot (E_C - E_A)\right] - \left[(N_C - N_A) \cdot (E_B - E_C)\right] = 0 \,. \tag{5.26}$$

This model will work whether point C is between points A and B or is on an extension of the line between points A and B, because the slope of a line remains constant whether the line is defined as going from A to B or from B to A. This is a constraint model as there are no observables involved.

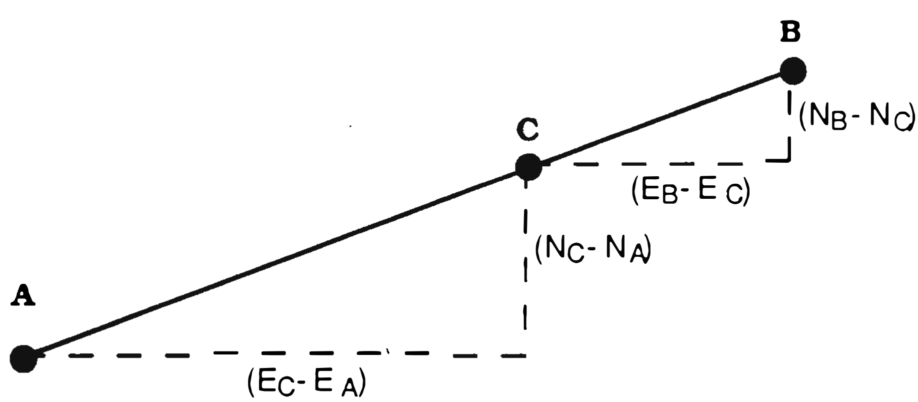



Figure 5.1: Straight line modelling

In practice this model must be applied once for each point that is an end point of a line segment forming a part of a straight line. Care must be taken to ensure that the constraint is applied only once to each point, remembering that, as a rule, each point will be used by two line segments. The condition should also not be applied to the end points of the straight line.

### 5.4.2 Circular Curve Models

The inclusion of a circular curve condition leads to four models that may need to be used. Three of these are observation models and one is a constraint model. Figure 5.2 is used to illustrate the

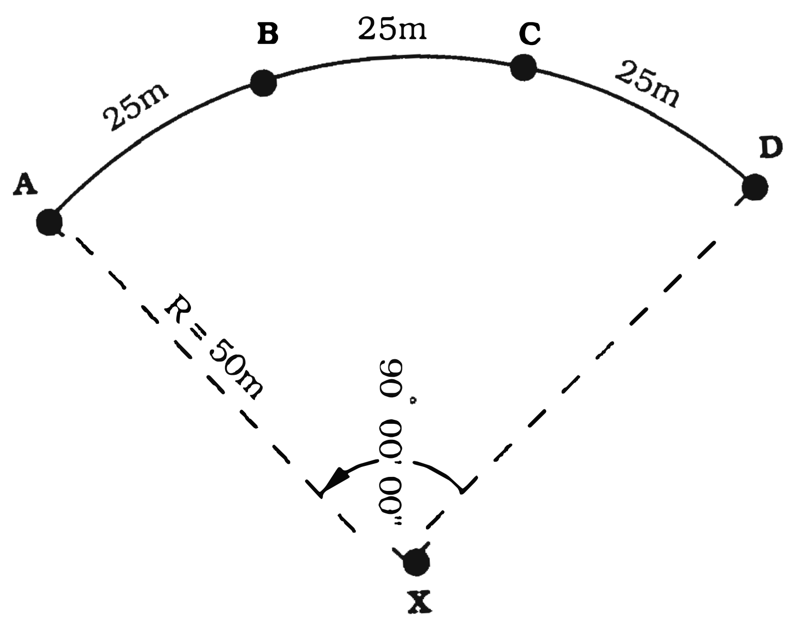



Figure 5.2: Circular curve modelling

formation of these models. In this figure points A, B, C, and D lie on a circular arc with arc lengths between them of 25 metres. The arc is centred at point X and has a radius of 50 metres and a delta angle of

90 degrees. The first condition to model is the constraint that the points lie on the same circular arc. This is done by enforcing the requirement that, if the points are on the same arc, then the distance from each point to the centre of the arc must be the same. This is expressed as:

$$\mathbf{f_c(x)} = \left[(N_X - N_A)^2 + (E_X - E_A)^2\right] - \left[(N_X - N_B)^2 + (E_X - E_B)^2\right] = 0 \ . \qquad 5.27$$

This condition is applied to the end points of each of the arc segments that make up the complete arc. In the situation illustrated in Figure 5.2, it would have to be applied three times.

The intuitive approach to the formation of Equation 5.27 is to simply model the distance between each point on the curve and the centre as being the observed radius. This is not a valid approach. In the formulation of the least-squares equations given earlier in this chapter, there is an assumption made that all of the elements of the observation vector, $\ell$, are linearly independent and that each observation is used once, and only once, in the formation of the functional relationships between the observations and unknowns. If the radius were used in more than one model it would appear to be several different observations. The result would be that each "observed" radius would end up with a different residual, resulting in the points A to D being close to lying on a circular arc, but not exactly. A further problem with this approach arises when the radius is not shown for the circular arc. In this case the approach used in Equation 5.27 will still work, while the intuitive approach will fail completely.

The radius is not ignored, however. Equation 5.27 has the effect that each point on the arc will be exactly the same distance from the centre after the adjustment, but what that distance should be is not specified. If there is a radius shown, then it can simply be included by modelling it as a straight distance (Equation 5.14) between the centre of the arc (point X) and any one of the points on the arc.

The delta angle for the circle also needs to be included in the adjustment. This is done by treating it as an observed angle (Equation 5.16) at the centre of the arc, backsighting and foresighting the end points of the arc. This leaves only the arc distance model to complete the mathematical models for circular arcs. The relationship between the length of the chord, c, the radius of the circle, r, and the arc length, s, is [CRC, 1976]:

$$s = 2r \cdot \sin^{-1}\left(\frac{c}{2r}\right) . \qquad 5.28$$

If the arc length is modelled as the straight line distance between the two points on the arc, and the radius as the straight line distance between the centre of the arc and one of the points on the arc, this gives the final mathematical model for arc distances as:

$$\mathbf{f}(\mathbf{x}) - \ell = 2 \cdot \sqrt{(N_X - N_A)^2 + (E_X - E_A)^2} \times \sin^{-1}\left(\frac{\sqrt{(N_B - N_A)^2 + (E_B - E_A)^2}}{2 \cdot \sqrt{(N_X - N_A)^2 + (E_X - E_A)^2}}\right) - s_{A,B} = 0 . \qquad 5.29$$

Again the observed value of the radius has not been used in the mathematical model for the reasons explained earlier. The choice of which point to use in computing the radius is immaterial, since

Equation 5.27 will be applied to both points, forcing both distances to be the same.

### 5.4.3 Parallel Line Models

The final mathematical models that are needed are those used to describe two parallel lines. The case of parallel curves requires no explicit modelling. The simple fact that two curves use the same centre point in the application of the circular arc conditions will force the curves to be held parallel. However, the situation of two parallel lines does require modelling. The equation used to model the requirement that the line from A to B is parallel to that from C to D is based on the fact that two parallel lines have the same slope. This leads to the model being:

$$\mathbf{f_C}(\mathbf{x}) = \left[(N_B - N_A) \cdot (E_D - E_C)\right] - \left[(N_D - N_C) \cdot (E_B - E_A)\right] = 0 \,. \qquad 5.30$$

This model is very similar to that of Equation 5.26, the straight line condition. In this case, however, there are four points used in the model, while in Equation 5.26 only three points are used, one of them twice.

In addition to two lines being parallel, the separation between them may also be an observable. The model for this is based on the equation for the perpendicular distance from a point to a line, derived in Section 4.1.1, Equations 4.1 to 4.3. If lines AB and CD are parallel straight lines and have a known separation of $\mathbf{w}$, then the distance from point C to line AB must be $\mathbf{w}$. This leads to the mathematical model which is:

$$f(\mathbf{x}) - l = \left[ (N_C - N_A)^2 + (E_C - E_A)^2 - \left( \frac{(N_C - N_A)\cdot(N_B - N_A) + (E_C - E_A)\cdot(E_B - E_A)}{\sqrt{(N_B - N_A)^2 + (E_B - E_A)^2}} \right)^2 \right]^{\frac{1}{2}} - w = 0\ . \qquad 5.31$$

If the two lines are parallel curves with a known separation, then the difference in the radii of the two curves must equal the known separation, w. This may simply be modelled as:

$$f(\mathbf{x}) - l = \sqrt{(N_X - N_A)^2 + (E_X - E_A)^2} - \sqrt{(N_X - N_B)^2 + (E_X - E_B)^2} - w = 0\ . \qquad 5.32$$

## 5.5 Blunder and Error Detection

Once the plan has been adjusted, the results must be examined to check for any errors in the data and the constraints applied to the unknowns. It is this step which, ultimately, verifies that the data shown on the plan and entered by the operator are correct and all fit together. There are several statistical tests that can be applied to the results in the search for errors and these are detailed in the following section. Before this is done, however, the statistical basis and assumptions made in applying the method of least-squares to the data need to be reviewed.

The main assumption made in using the method of least squares is that each of the observations comes from a normally distributed population. This assumption is made for the analysis of the results only. In order to apply the method of least squares, the only requirement is that the observations have a finite standard deviation

[Hamilton, 1964]. The second assumption is that the only errors in the data are stochastic, that is the normal variance due to the sampling process. If there are any non-stochastic errors (blunders) in the data then the solution will be biased away from the true solution and the acceptance of the solution will result in the addition of incorrect information into the cadastral data base.

The third assumption is that the variances of the observations are known. These are required if the covariance matrices computed using Equations 5.13 and 5.23 are to be correct. If the variances are known to within a single scale factor then the solution for the unknowns will be correct but the covariance matrices will be in error by the same scale factor [Vaniček and Krakiwsky, 1982]. If the variances of the observations are not known at all, then the method of least squares cannot be used. Fortunately this situation is extremely rare when dealing with survey data.

The first test that is usually applied to the results of a least squares adjustment is the variance factor test. The estimated variance factor, $\hat{\sigma}_0^2$, is given by

$$\hat{\sigma}_0^2 = \frac{\hat{\mathbf{r}}^t \cdot \mathbf{C}_\ell^{-1} \cdot \hat{\mathbf{r}}}{\nu} \qquad 5.33$$

and is the estimate of the scale of the $\mathbf{C}_\ell$ matrix. If the scale of this matrix is known, then this estimated value can be tested to ensure that it is statistically equal to the known a priori value using the test

$$\frac{\sigma_0^2}{\nu} \cdot \chi^2_{\nu,\frac{\alpha}{2}} < \hat{\sigma}_0^2 < \frac{\sigma_0^2}{\nu} \cdot \chi^2_{\nu,1-\left(\frac{\alpha}{2}\right)} \qquad 5.34$$

where: $\sigma_0^2$ is the a priori variance factor, usually one.

$\chi^2$ is the abscissa value from the chi-squared distribution,

$\nu$ is the degrees of freedom from Equation 5.24, and

$\alpha$ is the significance level of the test, usually 0.05 or 0.01%.

This test will fail if [Vaniček and Krakiwsky, 1982]:

a) The observations are not from normally distributed populations, or

b) The $\mathbf{C}_\ell$ matrix is incorrect, or

c) There are blunders in the observations, or

d) The mathematical models used are wrong. Included in this is the application of an incorrect constraint.

This test is analogous to taking a person's temperature. If the temperature is too high or too low then there is probably something wrong, but one cannot tell what is wrong. However, if the temperature is normal one cannot say that the person is healthy. Similarly, if the variance factor test fails then something is wrong and the results cannot be accepted until the problem is found and corrected but, if the test passes, other tests must continue to be applied to the results of the adjustment, in the continuing search for errors.

The most important statistical test that is applied in the detection of errors in the observations is the residual outlier test. Each individual residual in the $\hat{r}$ vector is tested using the equation

$$|\hat{r}_i| = \xi n_{1-\left(\frac{\alpha}{2}\right)} \cdot \hat{\sigma}_{\hat{r}_i}. \qquad 5.35$$

where: $\xi n_{1-\left(\frac{\alpha}{2}\right)}$ is the abscissa value from the standard normal distribution at a $1-\alpha$ confidence level, and

$\hat{\sigma}_{\hat{r}_i}$ is the standard deviation of the estimated residual.

If the residual fails the test then its associated observation is probably an error. It is possible, however, for the residual test to fail when the observation is not an error.

The incorrect application of constraints on the points is one possible cause of these incorrect failures. Examples of incorrect constraints include incorrectly identifying a point as being on a straight line and the use of incorrect coordinates for the known points. Under these circumstances there will usually be many failed residuals on the plan, providing a clue to the cause of the failure.

A second, more common, source of false failures is that one bad observation will affect the adjusted coordinates of several points and thus will cause a failure of the residuals of the observations connected to these points. Mackenzie [1985] has shown that, in the case of several residual failures, the most likely candidate for rejection is the associated observation with the largest standardized total residual.

The total residual is found by dividing the estimated residual $\hat{r}_i$ by its redundancy number $q_i$ where

$$q_i = \left(C_{\hat{r}} \cdot C_l^{-1}\right)_{i,i} \qquad 5.36$$

The redundancy number will always lie in the range 0 to 1 and shows what fraction of an error in an observation will be present in the residual. Low redundancy numbers, typically those less than 0.2, are a warning that errors in the associated observations cannot reliably be detected.

## 5.6 Summary

The least squares approach to the calculation of the positions of the cadastral survey points is very powerful. The inclusion of the topological constraint models developed for this thesis makes it more so. The positions obtained through this process are the best estimates available and reflect the topological nature of the plan. In addition to the positions, the process yields accuracy estimates for the point positions, valuable information that traditionally has been missing from Land Information Systems.

The statistical basis for the least squares method makes statistical analysis of the results possible. Using this approach, the observations shown on a cadastral plan can be rigorously and systematically examined for errors without having to rely on the thoroughness of the operator. With the inclusion of the redundancy number calculation, this analysis is extended to the identification of observations that cannot be checked, valuable information in itself.

This concludes the theoretical background for the development of a cadastral data entry program. The following chapters of this thesis present the implementation of these techniques.

# Chapter 6

# IMPLEMENTATION

This chapter describes the cadastral data entry program (CDEP) and the cadastral adjustment program (CADJ) developed using the ideas presented in the preceding chapters. Primary considerations in the design and coding of the programs were the ease of use and robustness objectives identified in Chapter 1. All input data are validated immediately upon entry and the operator is notified of any problems and potential solutions. The program is completely menu driven, with duplicate menu systems used on the digitizer tablet and the terminal screen, allowing the operator the option of using whatever menu is most convenient.

The first section of this chapter describes the hardware and software available for the implementation, and the overall approach to the programming. This is followed by a description of the menu systems developed for this project, their use, and how they provide for both simplicity and robustness. Following this is a detailed description of the process of entering the data from the cadastral survey plan and the computation and analysis of the data. The chapter concludes with the results of the testing of the programs.

## 6.1 Hardware and Software Considerations

In keeping with the objective of simplicity, the cadastral data entry system should ideally be implemented on a microcomputer.

This would result in the desired low capital cost to any company or government agency wanting to use the system. For the purpose of this thesis, this was not done, for two reasons. The first problem was that of availability. At the time that this project started the Department of Surveying Engineering did not have the necessary microcomputer facilities. The second reason was that it is simpler, in the author's opinion, to develop large programs on larger computers and later transport them to microcomputers, once the ideas and algorithms have been fully tested on the larger systems.

The hardware configuration used for this project is illustrated in Figure 6.1. The computer used is a VAX 11/750 running the VMS operating system. The VAX is configured with a floating point processor, 8 Mb of memory and almost 1 Gb of disk space, providing an excellent program development environment. The device used to display the entered data is a Tektronix 4208 graphics terminal. This is a 16 colour graphics terminal with a 640 by 480 pixel resolution and incorporates a display list processor. The display list processor makes the editing of the graphics image very simple. The primary input device is a Summagraphics ID2000 digitizing tablet with a 36 inch by 48 inch active area. This large surface is required when dealing with large plans. As explained in Section 6.3.2 the tablet is not used for precise digitizing, serving instead as a pointing device. The last hardware component used is a standard ASCII text terminal, in this case a Digital VT220. This terminal is used for the display of messages to the operator and for the menus. Using this second terminal for text reduces the clutter on the graphics terminal.

The only compiler available on the VAX computer at the start of this project was FORTRAN. For this reason the implementation is entirely in FORTRAN. In a microcomputer environment the language of choice would probably be C, as it is more fully developed for microcomputers than FORTRAN. The use of FORTRAN for the development of the prototype cadastral data entry program was not a problem. FORTRAN, as implemented under the VMS operating system, has many extensions to the standard, including facilities for interrupt driven programs and record data structures, that make the design and coding of interactive programs much simpler than with standard FORTRAN.

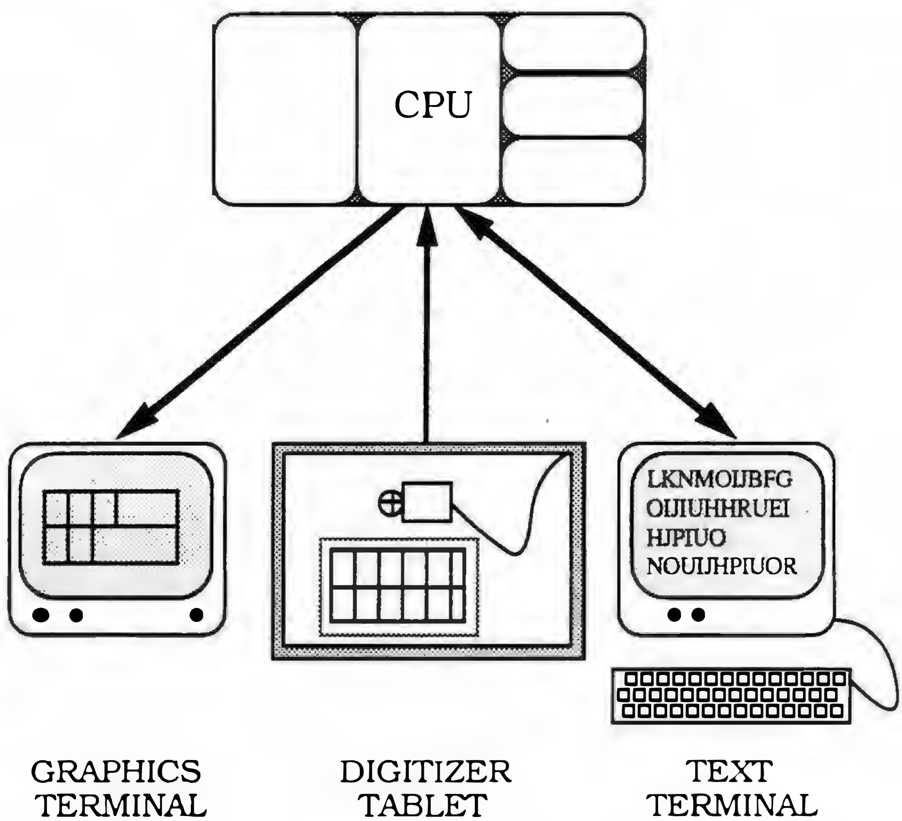



Figure 6.1: Hardware configuration for CDEP

All the graphic display from the program is performed using the GKS graphics standard [Enderle et al., 1987]. The use of this international standard makes the use of other graphics terminals possible and improves the portability of the code. The interface with the digitizer tablet was designed specifically for this project to allow very flexible use of the tablet as a graphic input device. The interface incorporates the use of windows on the digitizer tablet, allowing for the simultaneous use of cadastral survey plans, detail drawings, and menus. Both interfaces have been implemented as stand alone graphics handler (GRH) and digitizer handler (DGH) programs that communicate with CDEP and CADJ via inter-process communication facilities on the VAX. This approach simplifies debugging and reduces the complexity of the programs. Figure 6.2 illustrates the communication paths between the programs.

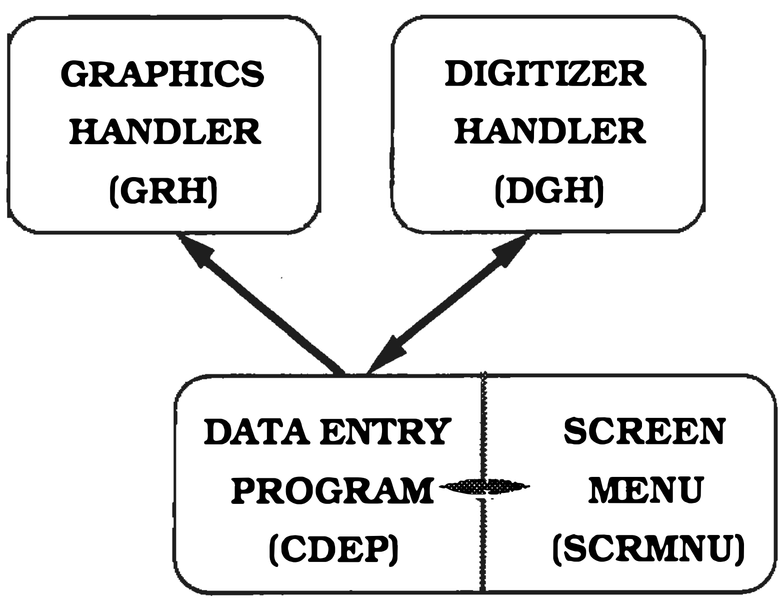



Figure 6.2: Communication paths for CDEP

## 6.2 Menu Systems

As mentioned earlier, CDEP is completely menu driven. The prime advantage of a menu driven system is that the operator does not have to memorize a set of commands in order to perform his job. Using a menu system, only the commands that may validly be chosen at a particular stage in the processing are displayed on the screen. The operator has only to select the command he wants from the list in the menu. This eliminates attempts to perform invalid operations, removing one potential source of errors. In addition, having the full list of commands constantly displayed reminds the operator of all his possible actions.

CDEP uses two complementary menu systems in parallel. Each system has its advantages and disadvantages, but the advantages of one negate the disadvantages in the other. The first menu system is a full screen menu system (SCRMNU), developed by the author as a general purpose menu system for use in this and other projects. The second menu system is part of the DGH program and uses printed menu sheets on the digitizer tablet. These two key components of CDEP are described in the following sections.

### 6.2.1 The Screen Menu System

The SCRMNU subroutine package provides a general purpose full screen menu interface that can be used by any FORTRAN program [Mepham, 1988]. The programmer defines a set of menu pages, each of which may contain an unlimited number of selection lines. Each selection is assigned a code number that is returned to the application program when the associated selection line is chosen by the user. An example of a screen menu display is shown in Figure 6.3.

There are four categories of menu selection line. They are:

a) Immediate action lines. This type of line is used when the command requires no additional data or options. An example of the type of command that falls into this category is "stop";

b) Typed response lines. When this type of menu line is selected the user is prompted to enter from one to ten values. The data type of the responses may be restricted to one of the following:
   i) Text strings,
   ii) Integer numbers,
   iii) Floating point numbers, or
   iv) Angular values.

   The menu system validates all typed input to ensure that no bad data are returned to the application program;

c) Selection response lines. When the user selects this type of menu line he is presented with from two to twenty pre-defined values to choose from. He must select one of the values to complete the menu selection; and

d) Yes/no selection lines. These are used to toggle options on and off. Whenever the user selects a menu line of this type the associated value will toggle between yes and no.

| Menu page CURVE DEFINITION | | RESPONSE |
|---|---|---|
| A | CURVE ID NUMBER | 1 |
| B | END POINT STATION NAMES | 1234-56<br>1234-74 |
| C | CENTER POINT STATION NAME | 1234-321 |
| D | CURVE RADIUS | 50.0 |
| E | RADIUS STANDARD DEVIATION (MM) | 5.0 |
| F | RADIUS BLUNDER WARNING LIMIT (%) | 10 |
| G | DELTA ANGLE KNOWN | <u>**YES**</u> / NO |
| H | DELTA ANGLE | 45-00-00 |
| I | STANDARD DEVIATION (SEC) | 60 |
| J | BLUNDER WARNING LIMIT (DEG) | 10 |
| K | ACCEPT CURVE AS SHOWN | |
| L | REJECT THIS AS A CURVE | |
| M | ABORT CURVE DEFINITION | |

| SELECTIONS |
|---|
| |

| PREV VALUE |
|---|
| |

| SPECL KEYS |
|---|
| HELP/PF2<br>CTRL-W<br>CTRL-Z |

Figure 6.3: Sample screen menu page

The currently active menu selections are displayed on the screen with a single letter (A to Z) beside each line. The user selects a line by pressing the appropriate key on the keyboard. If the menu page contains more lines than can be displayed on the screen the user can use the arrow keys to scroll through the choices. SCRMNU also has a built-in help facility to aid the novice user.

Among the many other features of SCRMNU is an escape mechanism to the operating system. At any time the user may suspend processing of the current program and return to the operating system command level. When he is finished there, he may return back to his application program and continue as if there had been no interruption. Provision is also made to trap any output sent from the application program to the screen, and display it on the screen without interference from the menu text. This permits the use of standard input/output statements in the application program.

The principal disadvantage with the use of SCRMNU, when working with a cadastral plan on the digitizer tablet, is that the operator must continually move between the digitizer and the terminal as he points with the digitizer and selects commands using the terminal. For this reason a digitizer menu system was also implemented for use with CDEP.

### 6.2.2 The Digitizer Menu System

A menu facility is included in the DGH program in order to simplify the use of CDEP. The menu pages to be used with DGH are designed by the programmer according to the needs of the application program. A sample menu page from CDEP is shown in Figure 6.4. As with SCRMNU, each selection box has a numeric code associated with it that will be sent to the application program whenever the user digitizes a spot inside that box. In designing the digitizer menus for CDEP the return codes were selected to match those used with the SCRMNU system. This greatly simplified the programming.

In contrast to the screen menu system, the digitizer menu system makes command selection simple and convenient while the operator is working with a cadastral plan on the digitizer. The drawback to the digitizer menu is that a keyboard is still required for the entry of typed values, such as observations and station identifiers. It is for this reason that the screen menu is used in conjunction with the digitizer menu.

### 6.3 Data Entry

The data entry process involves the entry of all the information shown on the survey plan. This includes the metric, topological, and descriptive information. The metric data include the azimuths, angles, distances, and known positions, while the topological data describe the connectivity of the points on the plan. The descriptive data will include such things as the lot and block numbers, date of survey, surveyors name, purpose of survey, etc. There are be four stages of data entry. These are discussed in the following sections.

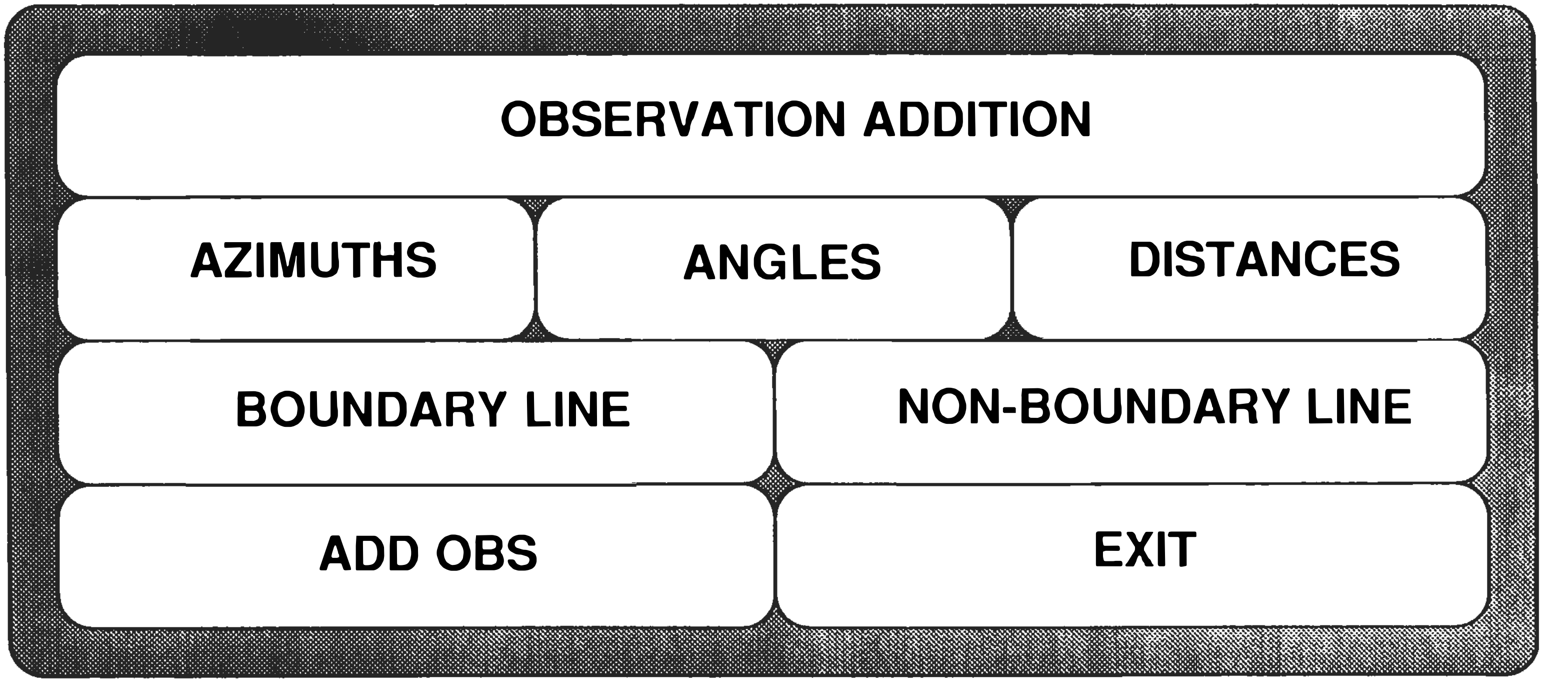



Figure 6.4: Example of a digitizer menu page

### 6.3.1 Initialisation

The initialisation stage of the data entry involves entering the descriptive information about the plan and mounting the plan on the digitizer tablet. The descriptive data to be entered is that information shown in the legend area of the plan (surveyors name, date of survey, purpose of survey, etc.). These have to be entered via the terminal keyboard as they are all textual information. Since the plan inspection process takes place before a plan number is assigned by the Land Titles Office, a provisional plan identifier should be assigned at this time. This will allow future references to the data by the provisional plan number.

The mounting of the plan on the digitizer requires the marking of two registration points on the plan and carefully digitizing these. The registration marks are used to allow re-mounting of the plan at a later date for editing the data. Once the plan is mounted, the north arrow needs to be digitized to provide orientation to the plan. This is to enable the solution of the azimuth ambiguity problem to be discussed in Section 6.3.3. The final step in mounting a plan is to digitize and enter approximate coordinates for from two to twenty points on the plan. These points do not need to be actual legal plan points but may also be grid intersections or any other defined point, for which approximate coordinates are known. This information is used to establish an approximate transformation from digitizer coordinates to real coordinates for all the points that are identified in the next step. These approximate coordinates are used in the least squares adjustment described in Chapter 5.

### 6.3.2 Point Identification

The second stage of the data entry process is to identify to the system, each and every point on the plan. The digitization of the points does not have to be done very carefully, as the digitizer coordinates are not used as metric information. They are used only to identify what points are being selected during the observation entry stage, and for any future editing. It is expected that, by allowing this rough digitizing, the point identification stage will not take as long to perform as it would otherwise.

At the same time as the point is digitized, its attributes are also assigned. The most important of these attributes is the classification of the point (see Section 2.3). Based on the class of the point being entered, other attributes may be either required, optional, or disallowed. For example, if the point is classified as a pre-existing point, then its identifier must be provided. If a point is classified as survey control, then its coordinates must be provided (or obtained from a survey control data base). If it is classified as new, then its coordinates must not be entered as known coordinates, since the coordinates of the new points will be computed using the plan information. These checks are all made by CDEP and the operator informed if they are violated.

### 6.3.3 Line Entry

The line entry stage is the most time consuming step of the data entry operation. Every line on the plan must be entered along with the type and value of any observations on that line. The method for this is to have the operator point to the end points of the line using the digitizer cursor and then enter the observation for that line on the keyboard. Once again the digitizer pointing does not have to be accurate. During line entry any digitization is assumed to be a pointing to an existing station. The station nearest the spot actually digitized is assumed to be the station that is being pointed at. If a point was missed during the point entry stage, it is a simple matter to leave the line entry stage, enter the missing point, and return to the line entry stage without any loss of data.

The actual observed values as shown on the plan are entered using a keyboard. This is a very error prone operation. For this reason procedures to detect errors have been designed and implemented. The first check applied to the entered value is to ensure that it does not contain any illegal characters (e.g. alphabetics in a numeric string). The second check is a blunder check involving the comparison of the entered value with a value computed using the digitized positions of the points. These blunder checks are necessarily coarse since there is no precision in the point digitization. The final check for errors takes place at the coordinate computation phase of the plan entry (Section 5.5).

During the observation entry, one aspect of the line classification scheme presented in Section 2.4 must be taken into account. During the polygon formation process presented in Chapter 4, the classification of a line segment as boundary or non-boundary is required, since only boundary lines may be used in defining the polygons. This information is entered at the same time as the observations are entered. This is the only line classification that needs to be manually entered. All other classifications are determined automatically.

There are four observation types that may appear on a plan. The entry of each of these, and the associated problems that may occur, are discussed in the following subsections. The entry of link lines, lines with no associated observations, is also described.

a) Distances

Distances are expected to be the most frequently occurring observations on the plan. No great problems with their entry are anticipated. Blunders in the entered distances can be detected at the data entry stage. The approximate scale of the plan is known since approximate coordinates of two or more points on the plan have been entered during the initialization stage. Blunders are identified by simply comparing the entered distance value against the value computed from the digitized coordinates. If the difference is greater than a given relative tolerance, the user is warned and given the option of accepting or changing the distance. It should be noted that arc lengths are entered at this stage as if they were straight line distances. These will be changed to arc lengths during the point on curve detection stage of the processing.

b) Azimuths and Bearings

Azimuths are also a frequently occurring type of observation on the plans. Bearings are treated exactly the same as azimuths. Their entry is as simple as for distances and large blunders can also be easily detected. There is one problem peculiar to azimuths, however. When an azimuth is shown on the plan, there is no indication of its orientation. The relative orientation of the two endpoints has to be compared to the orientation of the north arrow, digitized during the initialization stage, and the entered value of the azimuth, to determine which point is the **from** point and which is the **to** point. This is done automatically without any operator intervention or notification. The orientation of the north arrow is also used in checking the entered value of the azimuth for blunders. If the computed azimuth differs from the entered value by more than a given tolerance, the user is given the option of changing or accepting the azimuth.

c) Angles

The entry of angles differs from that of all the other observation types in that an angle involves three points, not two. The three points (occupied, backsight, and foresight) must be correctly identified by the user. Blunder detection is a simple matter for angles as there is no problem with either scale or orientation.

d) Directions

Direction observations do not occur with any great frequency on legal survey plans. The general practice is to show, on the plans, the angles or azimuths computed from the observed directions. Provision has, however, been made for directions. This is more for the sake of completeness than because it is expected that direction observations will become commonplace on legal survey plans. The main problem with the use of directions is associating the orientation unknowns with the rounds of direction observations. As a starting point it is safe to assume that every direction from a given point is in the same round. (Since the use of direction observations is rare in itself, the use of more than one round of directions at the same station is even rarer.) Provision has, however, been made for more than one round of directions at the same station.

The use of the digitized coordinates of the points for a blunder check on the value of the entered directions is done, but only after the second and subsequent directions have been entered at a point. This check is accomplished by examining the angles between two lines with observed directions on them and the angle computed from the digitized coordinates. With this check it is not possible to detect which of the two directions used in the calculation of the angle is in error.

e) Link Lines

There are some lines on legal survey plans that do not have any observation associated with them. They are referred to as "link lines" in this thesis. Link lines may be boundary or non-boundary lines as described earlier. Boundary link lines are used during the polygon formation process, while both types of link lines come into play during the topological constraint detection process. Figure 6.5 shows both types of link lines as they exist in the plan shown in Figure 1.1. The boundary type link lines are used to show where the boundary between the laneway and roadway is, while the non-boundary lines will be used in the straight line formation process. The 20 and 50 foot values shown on the plan are not the lengths of the link lines but are the widths of the right of ways for the laneway and roadway, respectively. These values will be entered later, during the parallel line detection step of the plan entry process.

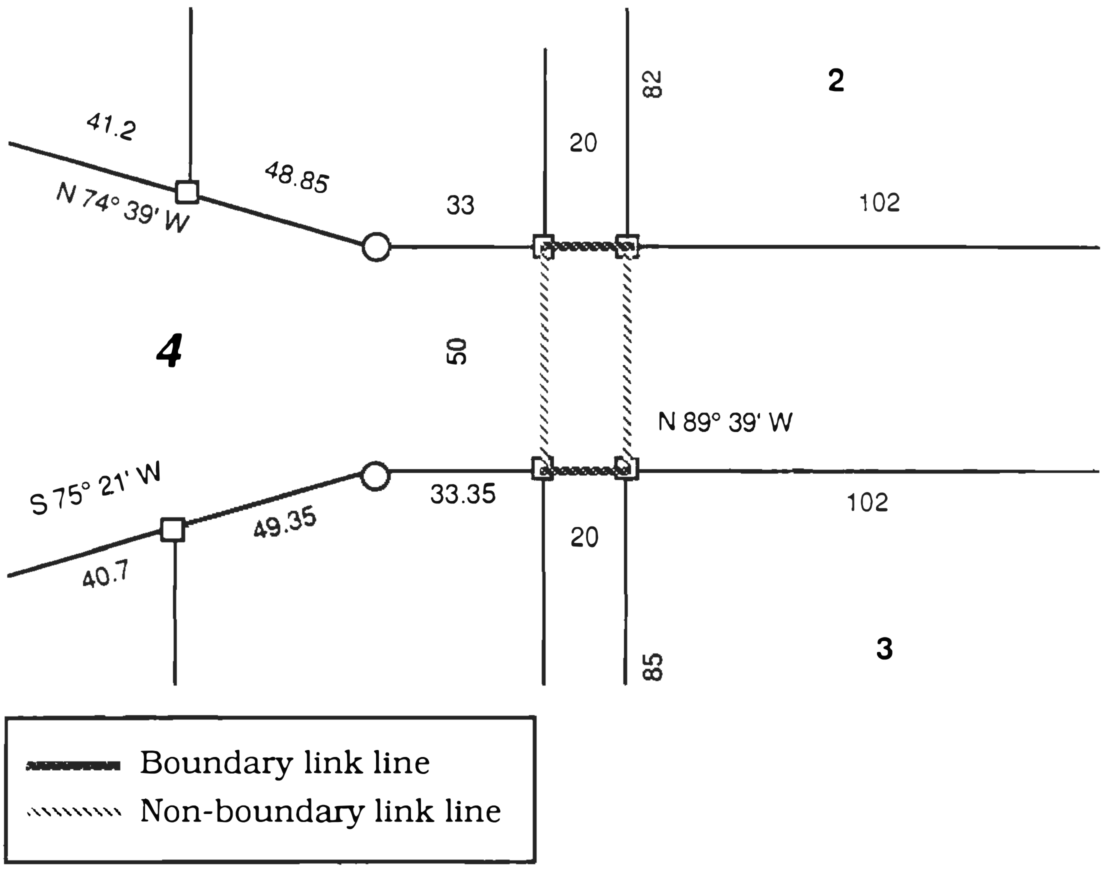



Figure 6.5: Boundary and non-boundary link lines

### 6.3.4 Efficiency Considerations

Steps were taken in designing the observation entry process to simplify and speed up this lengthy process. The reasons for this are twofold. The first is the obvious one that the less time it takes to enter the data the lower the cost in terms of the number of man-hours required. The second, less obvious, reason is that the number of errors made tends to increase with the amount of time spent entering the data.

The first feature for speeding up the observation entry is the use of a stack for the end points of the observations. As each new point is identified it is added to the bottom of the stack, with the contents of the stack all moving up. The point that was at the top of the stack is lost. This results in an efficient point identification when entering the observations along a string of points. The effect is that the previous **to** point will become the **from** point for the current observation, as soon as the current **to** point is identified. (All points, when pointed at, become **to** points by default). If the observation being entered is not connected to the **to** point of the previous observation then it is simply a matter of pointing to both the points involved, in the order **from**, **to**. This simple process reduces the number of pointings made by the operator by up to 50 percent, a significant number.

A second approach to improving the efficiency of the observation entry process is to allow the use of a ditto operation on the observed values. This is done by simply using the last observed value (and type) as the default for the next observed value. The result is that when several observations having the same value are shown on a plan, the value needs to be typed only once by the operator. This use of repeated values is a common occurrence on legal survey plans. Lots are typically all the same size in a subdivision. In combination with the point stack described above, this ditto operation reduces the entry of lot frontages, for example, to simply pointing to the end point of the line segment and then choosing save from the menu, a very fast operation. The combination of these two shortcuts results in a time reduction of up to 80 percent in the entry of the observations on a survey plan, depending on the nature of the plan.

### 6.3.5 Constraint Identification and Polygon Formation

The last step of the data input process is the identification of the straight lines and circular curves, the formation of the polygons, and the detection of the parallel line conditions. This processing takes place in the order:

a) Straight line detection,

b) Circular curve detection,

c) Polygon formation, and

d) Parallel line detection.

The algorithms used are described in Chapters 3 and 4 of this thesis. These processes are highly automated, requiring very little operator intervention.

As each line, curve, polygon, or set of parallel lines is found it is displayed on the graphics screen using blinking lines and points. The operator is then asked to accept or reject the program's selection. If the operator accepts the selection, processing continues normally. A rejection of the selection is an indication of an error in the data, and processing is terminated. The operator may correct the error and then continue with this phase of the data entry.

A third option available to the operator during the curve detection process is the editing option. This involves the use of the line editor described in Section 3.2.1 to correct the definition of straight lines found during the curve detection process. Also during this process, the operator identifies the centre point and enters the radius and delta angle value for the curve as each curve is identified.

During the polygon formation process the operator has to digitize the starting point for each simple polygon as described in Section 4.1 and provide the polygon type and identifier. The super polygons described in Section 4.2 are formed automatically. Both types of polygon are displayed during their formation for the operator's acceptance.

## 6.4 Coordinate Computation and Analysis

Once the data entry stage has been completed, all the information is stored in a file that is used as input by the adjustment program (CADJ). This program can be run immediately, while the operator waits for the results, or it can be submitted to run later as a batch process. In either case, CADJ runs without operator intervention.

Any problems in the data found during the analysis of the adjustment are noted in an output file that the operator can examine. In addition, questionable data items are marked as such in the results file created by CADJ. This file has exactly the same format as the input file and can be read by CDEP. When a plan is plotted on the graphics screen by CDEP, any data flagged as questionable is highlighted so the operator can check it. He may then edit the data, if necessary, and resubmit the coordinate computation, or he may accept or reject the plan.

### 6.5 Results of Testing

The programs CDEP and CADJ have been used to enter and analyze two survey plans provided by the Alberta Bureau of Surveying and Mapping (ABSM). The first plan entered is that shown in Figure 1.1. During the coordination of this plan by ABSM, using their existing procedures (see Chapter 8), they determined that there was a problem getting the observations to close in lots 10 and 11. The analysis results from CADJ isolated this problem to the distance between the two lots. It is in error by almost ten feet.

The second plan used to test CDEP is a large subdivision consisting of 71 lots and 8 roads. This plan contains over 200 points and 400 observed values. This plan was entered and checked in less than three hours. No errors were found on the plan.

This limited testing has shown that the general approach proposed in this thesis, for cadastral survey plan data entry, works. In the testing process the programs met the original objective of a fast, complete, robust, and simple system. Further testing is planned in cooperation with the Alberta Bureau of Surveying and Mapping.

### 6.6 Summary

The implementation of the ideas and concepts presented in Chapters 1 to 5 has been successful and has proven their validity. The programs that have been written are powerful tools for the entry, into computer readable form, of all the information shown on a cadastral survey plan. The use of the dual menu systems has resulted in a program that is simple and natural to use, while the extensive use of automated procedures has made the program very powerful and fast.

Using the method of least squares for the computation of the coordinates of the plan points was successful. The inclusion of the topological constraints makes the results match the original design to a much higher level than would otherwise be the case. The statistical analysis of the adjustment provides a rigorous and thorough examination of the plan for errors.

# Chapter 7

## CADASTRAL PLAN FITTING

When the coordinates of the points contained in a legal survey plan are computed on a plan by plan basis, discrepancies in the coordinates of points common to the two plans will occur. These common points usually occur at the boundaries of the plans. The discrepancies may be small, resulting from the inexact nature of surveying, or they may be large, resulting from undetected blunders in the data. Whatever the size of the discrepancy, it must be resolved so that there is only one set of coordinates for the point. It is also important that, in the resolution of these discrepancies, the nature of the survey plan is not changed, that is straight lines must remain as straight lines, circular curves stay circular, etc.

### 7.1 Basic Approaches to Plan Fitting

The objective of the plan fitting process is to extend the idea of consistent homogeneous coordinates from the plan level to the data base level. In other words, all the coordinates stored in the data base should refer to a common, continuous datum. Since there will be points used in more than one plan, and because of the inexact nature of surveying they will not have identical coordinates computed in each plan, some method of fitting the two plans together is required. There are three options regarding which coordinates to adjust in the process of plan fitting that need to be examined. These are:

a) Adjust both the coordinates for the plans already in the data base and those for the plan being added to the data base;

b) Adjust only the coordinates of the plans in the data base to fit the coordinates of the new plan.

c) Adjust only the coordinates of the plan being added to fit the plans already in the data base; and

The first approach is the mathematically correct approach as it allows the simultaneous consideration of all the observations shown on the legal survey plans. It suffers from the serious drawback that an adjustment for the coordinates of all the legal survey points in a province would have to be carried out every time a plan, or set of plans, is added to the data base. In Alberta there are approximately 1,350,000 land titles [Giffen, 1986]. If, on average, each parcel is defined by four points and each point is used in the definition of two parcels, then there would be 2,700,000 points in the data base. The re-adjustment of this entire data base would involve solving for 5.4 million unknowns, an inconceivable project.

In practice, it is not actually necessary that the coordinates of all the points in the data base be solved for each time a new plan is added. The significant effect on the coordinates of the existing points is localized to some region surrounding the new plan. The problem lies in identifying, before the coordinates are recomputed, the extent of the region. If the area that is recomputed is too small, then more data will have to be extracted and the recomputation repeated, while if the area chosen is larger than required then computer resources are wasted. Compounding the problem of identifying the local area that is affected is the fact that the distance metric in this situation is not the cartesian distance from the new plan. It is, in fact, related to the relative accuracies between the points common to the new and existing plans, and the other existing points in the data base. Points having a smaller relative accuracy are "closer". The relative accuracy information is not currently available nor is it expected to be available in the future. It would require the storage of too great a volume of information.

The second approach listed above suffers from the same problem as the first. There are just too many coordinates to change. Its attractiveness comes from the assumption that, by forcing the older data to fit to the newer data, the quality of the older data will be improved. This is based on the assumption that the newer data are better than the older data. This is not necessarily true. There may be errors in the newer survey [Robinson, 1982]. In addition, the "older" survey may only be a few days older!

This leaves the third approach, that of fitting the new plan to the existing plans. This is not the theoretically rigorous solution; the first option is that. The approach of fitting the new plan to the existing plans is, however, the best solution in terms of its being achievable while still resulting in a satisfactory solution. The major disadvantage of this approach is that any errors in the existing data will result in a distortion of the new plan.

Whatever approach is used, it is essential that the coordinates of the common points be identical after the fitting process has been completed. If the fitting process results in only an approximate fit, discontinuities will be introduced into the data and there will still be two separate sets of coordinates with no known relationship between them.

## 7.2 Topological Requirements in Plan Fitting

One other requirement of the plan fitting process is that it should maintain the topology of the entities forming the plan. Of primary concern here is that the nature of the connections between points does not change and that the shapes of the polygons remain the same as they were. The specific criteria that have been identified and addressed in this research are:

a) Straight lines should remain straight after the plan fitting process;

b) Circular curves should remain circular;

c) Parallel lines should remain parallel;

d) The shape of the lots should not change; and

e) No new line intersections should be introduced.

The first two criteria can be met exactly by transforming only the end points of the lines and curves and then later proportioning the intermediate points. This approach also has the advantage of reducing the number of points that will have to be transformed as part of the plan fitting process.

Criteria (c) and (d) above cannot, in general, be met exactly. Using any type of coordinate transformation process, a certain amount of distortion will be introduced into the plan. The transformation method chosen should minimize such distortions. It should be noted that the distortions introduced will be small in any case. The objective of the plan fitting process is to resolve small discrepancies between the coordinates of common points while maintaining the relative positions between the points in the plan. Only if there are large discrepancies in the coordinates will significant distortions be introduced. Such large discrepancies will only result from errors or blunders in the data and, as such, must be corrected prior to the plan fitting process.

## 7.3 Three Approaches to Plan Fitting

In this section three approaches to the fitting together of several plans are presented. All three of the approaches have application to the process of plan fitting under differing circumstances but, for reasons to be shown, the third method is being recommended for use in the majority of cases.

### 7.3.1 The Photogrammetric Block Adjustment Approach

Using this approach each survey plan is treated as an irregularly shaped photogrammetric model. The common points between the plans are considered to be tie points and a photogrammetric block adjustment program is used to fit the plans together and obtain a single set of coordinates for the common points. This approach is being proposed for use in Québec [Giroux, 1987].

In his paper Giroux [1987] describes the use of the block adjustment program SPACE-M in two experiments to test this approach. SPACE-M is an independent model, photogrammetric block adjustment program [Blais, 1979a] that is widely used and available in Canada. In the first test, four subdivision plans with a total of 50 lots were adjusted using nine control points and nine common points as tie points. The second test involved the adjustment of 15 subdivision plans containing a total of 180 individual lots. The subdivision plans were from 22 to 33 years old. A total of 54 points were identified on the ground and surveyed to obtain their coordinates within the geodetic framework. A total of 49 common points were identified and used as tie points. The results of these two experiments, as reported by Giroux, are summarized in Table 7.1. (The discrepancy in the number of control points for the second experiment is not explained in his paper.)

| Experiment Number | Number of Plans / Lots | Control Points: Number | RMS X / Y (cm) | Tie Points: Number | RMS X / Y (cm) |
|---|---|---|---|---|---|
| 1 | 4 / 50 | 9 | 19 / 14 | 9 | 16 / 19 |
| 2 | 15 / 180 | 43 | 13 / 10 | 49 | 4 / 5 |

Table 7.1
Results of Québec SPACE-M Experiments

SPACE-M was developed for aerotriangulation purposes. As such it solves for the coefficients of a seven parameter (one scale factor, three translations, and three rotations) 3-D similarity transformation for each model to transform the three dimensional model coordinates into ground coordinates. In its application to the plan fitting process the third dimension is held fixed, resulting in effectively solving only for a four parameter (one scale factor, two translations, and one rotation), two dimensional similarity transformation for each survey plan. The nature of a similarity transformation is that there is no change in shape of the object being transformed. The result is that all five criteria listed in Section 7.2 are fully met by this process.

Using the similarity transformation, however, fails to meet the prime objective of the plan fitting process. This approach does not obtain unique coordinates for each and every point. There are still two sets of coordinates for the common points. The tie point RMS values shown in Table 7.1 are the RMS values of the differences in common point coordinates after they have been transformed using the similarity transformation parameters for each of the plans. The SPACE-M approach to resolving these discrepancies is simply to use the weighted average of all the transformed coordinates for a point as the final coordinates of the point [Blais and Chapman, 1983]. This results in a discontinuity in the plans. In addition, once this averaging process is done, the five criteria listed in Section 7.2 are no longer met, since the averaging process introduces discontinuous transformations of the coordinates.

The final point against the block adjustment approach is one of applicability. SPACE-M is a large program designed for aero-triangulation. It uses a seven parameter transformation model and has to be "fooled" into solving for only a four parameter model for the plan fitting problem. For each plan, all seven parameters are still solved for, resulting in the solution of a set of equations that is 1.75 times as large as it needs to be. Since the effort required to solve a set of equations is proportional to $n^2$ this corresponds to three times the computational effort actually required.

A similar approach using similarity transforms to fit two cadastral plans together is proposed by Gagnon [1988]. He has written a program specifically for cadastral plan fitting using a four parameter transform. This avoids the excessive CPU usage problem but still suffers from the other problems with the use of similarity transforms.

### 7.3.2 The Conformal Polynomial Approach

A second approach to the fitting of a plan to those already loaded into a data base is to use rubber sheeting [Burrough, 1986]. There are many approaches to rubber sheeting, some of which are continuous in nature, e.g. Blais [1979b] and Lodwick [1981], and some of which are discontinuous [Gillman, 1985]. The rubber sheeting described here is continuous and uses a complex polynomial of full degree to map the coordinates of the points on the new plan to those on the existing plans [Mepham, 1987b]. The results of this transformation will be an exact mapping of the common points, with a corresponding mapping of the points that appear in the new plan only. The steps involved in this approach are:

a) Identify the points that are common to the new and existing plans;

b) Compute the coefficients of a (n-1)th order complex polynomial using the coordinates of the n common points;

c) Transform the coordinates of the points on the new plan using the coefficients of the polynomial.

The coordinate transformation equation is simply:

$$P^{*} = C_0 + C_1 \cdot P + C_2 \cdot P^2 + C_3 \cdot P^3 + \dots + C_{(n-1)} \cdot P^{(n-1)} \qquad 7.1$$

where: $P^{*}$ represents the coordinates of the point transformed into the same system as the existing plans (expressed as a complex number),

$P$ represents the coordinates of the point from the new plan,

$C_0, \dots, C_{(n-1)}$ are the coefficients of the polynomial, and

$n$ is the number of points common to the new and existing plans.

Once the coefficients of the polynomial have been determined it is a simple process to apply Equation 7.1 to all of the points in the new plan. The coefficients ( $C_0, \dots, C_{(n-1)}$ ) may be computed by using Equation 7.1 as the mathematical model and treating the coefficients as the unknowns and the two sets of coordinates as knowns. One equation is written for each common point leading to n equations in n unknowns [Chapman and Blais, 1987]. This system of equations may be solved for using the method of least squares as presented in Chapter 7 but is most efficiently solved for directly. As a practical consideration in the solving for, and evaluation of, the polynomial the first step should be the subtraction of the average new coordinate from all of the coordinates used. This value is then added back to the transformed coordinates as a final step. This will avoid problems with exponent overflow resulting from raising large numbers to high powers in the evaluation of Equation 7.1.

There are several advantages to this approach. The foremost of these is that the transformed coordinates of the new points exactly equal the existing coordinates for those points. This is a result of the use of a full degree polynomial. The use of complex algebra in the transformation process simplifies matters greatly. The determination of what coefficients to solve for to obtain an exact fit is simply a matter of counting the number of common points and solving for the coefficients of a polynomial of that degree. This is not true for real polynomials. If, for example, the number of real polynomial coefficients required is greater than that required for a full fourth order polynomial but fewer than required for a full fifth order polynomial, the question arises as to what fifth order terms to leave out. There is no simple solution to this problem.

A second advantage of the use of the complex polynomial is that it is conformal [Kreyszig, 1972]. The angle preserving nature of conformal transformations helps to minimize the shape distortions introduced into the plan by the transformation process. The final advantage of this approach is its mathematical simplicity. The computations can be performed on mini and micro computers.

Unfortunately there are also several serious drawbacks to this approach. Evaluation of this approach using the five requirements listed in Section 7.2 reveals that the use of conformal polynomials fails to satisfy any of them. The first requirement, keeping straight lines straight, can be met by transforming only the end points of the straight line and then proportioning the intermediate points according to their relative distance along the line. A similar approach can be used for the second requirement, keeping circles circular. The additional problem that is introduced here is that in transforming the end points of a circular arc and the centre point of that arc, there is no guarantee that the distances between the centre point and the end points will remain equal. In general the two distances will be unequal and this obviously conflicts with the definition of a circle. The simplest remedy for this discrepancy is, after the transformation, to change the coordinates of the centre point to those of the nearest point to the transformed centre point, that is also equidistant to the transformed end points. This can easily be done if the centre point is used only in a single arc, but if there are several parallel arcs it may be impossible.

The third and fourth requirements, maintaining parallelism and shape, will not be met using a polynomial transformation. The conformal nature of the complex polynomial transformation will result in a minimization of the distortions but will not eliminate them. Avoiding the introduction of new line intersections, requirement five, is also impossible when using the polynomial approach. This will occur when common points are misidentified, resulting in a mapping of point A to point B and vice versa. This is a specific example of a more general problem with the use of full degree polynomial transformations.

Earlier it was stated that one of the advantages of using a full degree polynomial is that the fit between the transformed coordinates of the common points on the new plan and the existing coordinates of the common points would be exact. Unfortunately this is true whether or not the plans fit together. If, to use an extreme example, one of the coordinates were out by 100 metres, this would be modelled in the transformation, and all of its neighbouring points would be shifted appropriately. Since the computation of the coefficients involves $n$ equations in $n$ unknowns there are no residuals that can be examined in an attempt to detect the blunder. The only way to detect such large changes in position is to examine the transformed coordinates of the points with respect to their original values.

The objective of the plan fitting process is to resolve the small discrepancies that are unavoidable when dealing with coordinates determined from different sets of observations. If the polynomial results in a significant change in position for a point, this is an indication that there is a problem that needs to be resolved before the coordinates are accepted and stored in the data base. The decision as to whether or not a change is significant can be based on either a global, absolute limit (e.g. 1 centimetre) or by examining the size of the shift in relation to the accuracy of the original coordinates, as determined from the covariance matrix from the adjustment. This examination must be performed on all the transformed points since, using a polynomial adjustment, the largest shift will not necessarily occur at one of the common points.

### 7.3.3 Combined Coordinate Computation and Plan Fitting

Up to this point two approaches to plan fitting have been examined, neither of which fully meet the criteria established earlier. Both these approaches are part of a stepwise approach to the coordination of a survey plan. First the coordinates of the plan are computed without regard to the existing plans surrounding the new plan and then the new plan is distorted to fit the existing data. It is this two stage approach, with its initial out of context computation stage, that leads to the problems noted in the previous sections.

A third alternative is to perform the plan coordination and fitting simultaneously. With this approach the coordinates of the common points are held fixed at their known values during the adjustment computations for the coordinates of the points on the new plan. The result is that the criteria for plan fitting will be met exactly. The coordinates of the common points will keep their current values, by nature of being held fixed at the computation stage, while the straight line, circular curve, and parallel line conditions will be met exactly as a result of the topological constraints imposed on the adjustment.

The use of the fixed points in the adjustment will complicate the analysis of the adjustment somewhat. If there are no errors or blunders identified in the statistical analysis of the adjustment then there will be no problem. If, however, the analysis detects problems there can be trouble identifying whether the problem is due to errors in the data on the new plan or a misfit between the new plan and the existing data. In the analysis of the adjustment, fixed points are treated as if they were known exactly and could not be in error. As a result, any errors in the known points will be detected as errors in the observation data. This can be confusing when trying to interpret the failures in order to determine their cause.

As an aid in the interpretation of the failures, it is recommended that whenever a plan contains rejected data after the initial adjustment, the plan should be readjusted without fixing the common points. By examining the two adjustments together it will be easier to decide whether the initial failure is a result of a misfit between the two plans or erroneous data in the new plan. As a general rule, if the second adjustment results in no rejected data, the rejections in the first adjustment are most likely due to a misfit.

Misfits between plans may be caused by either undetected errors in the existing plan data or uncontrolled extension of the positions through several plans. In either case the simplest method for resolving the misfit, and either finding the erroneous data or simply adjusting the fit, is to combine the input data for the existing plan and the new plan and simultaneously adjust the data as if they were all from one plan. If there are undetected errors in the existing data, they should be identified by the analysis of the combined data. If, however, the original misfit is due simply to uncontrolled extension of the positions, then there should be no errors detected during the analysis process.

### 7.4 Summary

The plan fitting problem must be solved if all the data stored in a cadastral data base are to maintain the correct relative positions to each other. The small discrepancies that exist between the coordinates of a point computed from two different plans will introduce discontinuities between adjacent plans making it impossible to properly determine the relative positions of points established in different surveys.

Of the three plan fitting options examined in this chapter it is clear that the combined coordinate computation and plan fitting approach is the best. It results in all coordinates in the cadastral data base being referred to a common continuous datum and meets all the topological requirements. It has the further advantage that the plan fitting occurs simultaneously with the data entry process. This elimination of a stage in the process of loading cadastral plans into a data base will result in greater accuracy and further cost savings.

## Chapter 8

# A COMPARISON OF CDEP AND THE CURRENT ALBERTA APPROACH

In Alberta a system currently exists for the production of cadastral maps in digital form. The program currently operates within the Land Information Services Division (LISD) of the Department of Forestry, Lands and Wildlife. In this chapter the procedures used are described. This is followed by an analysis of the current system and a description of how the implementation of the CDEP program presented in Chapter 6 would result in an improvement over the current situation.

## 8.1 The Current System in Alberta

At the start of this project the author visited the Alberta Bureau of Surveying and Mapping (ABSM) which was responsible, at that time, for the production of cadastral maps in digital form. There he examined the processes and procedures used for the production of the maps. (This responsibility has since been taken over by LISD.) In addition, the author also visited the cities of Edmonton and Calgary, two of the largest users of the data, to see what use they made of the information and to find out what problems, if any, they had experienced with the maps. The following presentation of the procedures is based on these visits and discussions with Mintz [1988].

When a surveyor registers a plan at the Land Titles Office (LTO) of the Alberta Attorney General, the LTO sends a copy to the mapping section of LISD. The mapping section assigns a map identifier to the plan and passes it on to the municipal integration section. Here the plan is checked for internal consistency, identifiers are assigned to all of the points and a list of coordinates is produced for all the points shown on the plan. The plan and the coordinates are then returned to the mapping section where the coordinates are displayed on an Intergraph workstation and all the line work is added. The resulting map is then forwarded to the municipality in which the subdivision is located.

The internal consistency check and coordinate computation process is very similar to the coordinate geometry approach described in Section 1.2.2 of this thesis. The first step in the process is to collect all the existing information for the area covered by the plan and the existing plans adjacent to it. This includes hard copies of the surrounding plans and a listing of the Alberta Survey Control Markers (ASCMs) shown on the plan.

The next step in the computation process is to use the ties to the ASCMs shown on the plan, or in the accompanying field notes, to compute the positions of the plan points that are tied directly to the control. The positions of the remaining points are then computed by forming traverses between these tied points. The traverses are formed so that they pass through all the points on the plan. No attempt is made to use all the observations. The traverses are computed and their closures adjusted using proportioning, resulting in coordinates for all the points shown on the plan. The distances that are not used in the traversing are checked, by comparing the distance shown on the plan and the distance computed using the calculated coordinates. If the traverses all close properly, and the distances agree to within 15 centimetres, the plan and its coordinates are accepted.

If there are problems with the traverse closures or the distance comparisons, more traverses are formed and computed in an effort to isolate the source of the problem. The error isolation process may also involve recomputation of adjacent plans, checks for drafting errors, and recalculation of the ties. Once the observation that is in error has been isolated, it may be marked as unreliable and the processing continued, or the plan may be returned to the surveyor for correction.

When the municipal integration section finishes coordinating the plan it is returned to the mapping section. Here the points are plotted on an Intergraph workstation using the calculated coordinates, and an operator uses the workstation to add the line-work and the text to the map. This is effectively a "join the dots" operation although some automated procedures are used in an attempt to speed up operations. The end result of this process is a spaghetti map file in Intergraph's IGDS format. These data are then stored in the municipal mapping data base and copies are sent to the affected municipalities for their use.

When the municipalities receive the data, they update their municipal maps and add their data to them. The data added by the municipalities include land use, zoning, assessment, and services information. In the course of the interviews with the municipal officials responsible for using and maintaining the data, the author found that the cadastral maps are used extensively by both cities. The maps serve as the base for almost all the municipal mapping operations. Overall, the officials were very happy with the quality of the maps and they could not envision working without them. The three most common concerns expressed were:

a) The presence of errors in the data. Some users realized that errors are inevitable and had learned to work around them while others thought that the data should be perfect (an unrealistic expectation);

b) The lack of topology in the data. This was expressed as a concern about how to update the positions as a result of datum changes. With the current file formats, the point positions are hard coded into the data, with duplication, and are very hard to manipulate; and

c) The time delay in getting the data from the province after the plan has been filed. This is of particular concern in the engineering departments, who pointed out that the maps are required for the design and construction of services in the areas being developed. (It should be noted that since these concerns were raised, the province has shortened the time between plan registration and distribution of the maps to as little as 45 days).

## 8.2 The Impact of CDEP

A new approach to the solution of a problem should not be implemented just for the sake of employing new technology or ideas. The new approach must provide a significant improvement over the current solutions without causing new problems. In this section the technical impact of replacing the current cadastral mapping procedures in LISD with CDEP (including CADJ) is presented.

The most obvious change resulting from the installation of CDEP would be the merging of the functions of the municipal mapping and municipal integration sections of LISD. This is a result of the fact that, using CDEP, the coordinate computation and the line and text definition processes are simultaneous, in contrast to the current practice at LISD of separating them. The impact of this will be to reduce the time required for the plan coordination process, by having a single operator perform all the required work, without having to pass it on to a second operator for completion.

The second, and probably most important, change would be in the coordinate computation process. The operations of forming, computing, and adjusting traverses would be eliminated. Included in this is the elimination of the current, ad hoc, procedures for the isolation of errors in the data. Using the constrained least squares adjustment approach, as advocated in this thesis, in combination with statistical testing, results in a fully automated coordinate computation and error detection capability. The use of arbitrary acceptance limits, such as the 15 centimetres between the observed and computed distances, would be replaced by acceptance decisions based on sound statistical theory. The coordinates that are obtained for the points will also be more accurate, since all the information shown on the plan will be used in their calculation, not just the observations chosen during the traverse formation.

A second result of the automation of the coordinate computation process will be a significant improvement in service and turn around time without the hiring of additional personnel. This may require some realignment of tasks. Currently there are eleven people working in the municipal integration section of LISD and twelve in the municipal mapping section [Mintz, 1988]. The data entry operation using CDEP requires little additional effort over that currently required for the line and text entry process. The additional effort currently employed to perform the calculations could be transferred to data entry, resulting in a near doubling of processing capacity.

This doubling of capacity can be achieved with a minimal capital cost investment. LISD currently operates a VAX 11/785 computer system. Data entry stations, similar to the configuration used for the development of CDEP, could be connected to this computer at a cost of approximately $10,000 each. The total additional capital cost for an additional eleven data entry stations should not exceed $125,000, a small investment for the improved productivity that would be achieved.

One result of this improved productivity would be a significant reduction in the time delay between registration of the plan and the distribution of the updated map to the municipalities. Under the current system, cities receive the updates on a 45 or 60 day cycle, while smaller municipalities receive annual updates [Mintz, 1988]. With the implementation of the system developed in this thesis, it should be possible to reduce this turn around time to between seven and fourteen days.

The one stumbling block to the replacement of the current system with CDEP is the difference in data formats between what is currently produced and used, and that produced by CDEP. The current cadastral maps are stored on a mapsheet basis using the IGDS format, while CDEP produces files ready to be loaded into a topologically complete Land Information System, without regard to map sheet boundaries or other artificial constraints. In this sense CDEP is ahead of its time. Commercial land information systems capable of using this information to its full potential are just now starting to enter the market (e.g. Intergraph's TIGRESS system and Wild's System 9). The data from CDEP can, however, still be used in the current systems. All that is required is a translator program that will take the CDEP files and convert them to the IGDS file structure. This conversion from a topological data structure to a spaghetti data structure is relatively simple; the reverse, unfortunately, is very difficult. The one difficulty foreseen in the translation to IGDS is in the placement of the text. Some manual intervention may be required to obtain an aesthetically pleasing cadastral map.

## 8.3 Summary

The existing cadastral mapping system employed by LISD has been performing satisfactorily for several years now. It is, however, a manpower intensive process that uses ad hoc rules to determine the acceptability of the data. There is also some pressure to increase productivity while at the same time reduce costs.

In this chapter the impact of the implementation of CDEP at LISD, as a new tool for entering cadastral survey plans into a cadastral data base, and for updating cadastral maps, has been examined. The conclusion is that it would result in a near doubling of capacity while simultaneously reducing the time required to process the plans. In addition, the results are more accurate and the detection of errors more complete and thorough. One result of the increased accuracy will be that the data can be used in more applications, resulting in a more widespread use of the data, giving a better return on the investment put into the collection of the data. Finally, the implementation of CDEP is a first step towards the development of a topologically complete cadastral data base and, by collecting the data now, LISD will be in an excellent position to use and exploit the advanced LIS tools currently in commercial development.

# Chapter 9

# CONCLUSIONS AND RECOMMENDATIONS

## 9.1 Conclusions

The prime objective of this research was to design and implement a data entry system for cadastral survey plans that met the following four criteria:

a) Speed,

b) Robustness,

c) Completeness, and

d) Simplicity.

This objective has been met.

Existing approaches to the solution of the data entry problem using digitization and traverse computations were analyzed and found not to meet these criteria. Both the straight digitizing approach and the traverse computation approach were found to be incomplete and lacking in robustness. The traverse method is also very complicated, requiring much effort in the traverse formation stage.

The first step in the design of the new data entry system was to define the entities that appear on a cadastral plan and to determine the relationships that exist between them, i.e. the topology of the plan. The key entity that appears on plans is the polygon, since the purpose of a subdivision survey is to divide the land into areas that may be treated as single units. The other plan entities, points and line segments, are required to define the size, shape, and location of the polygons. An expansion of the point, line segment, and polygon entity relationship model to include super polygons, lines, and relative points was specifically developed, in order to simplify the representation of a cadastral plan in digital form.

The identification of additional topological relationships is also a critical component of the data entry system. These are the point on line, point on curve, and parallel line and curve relations. Rules were developed for the automatic detection of these relations. The inclusion of these relationships results in a more accurate and reliable determination of the coordinates of the points, as well as a better representation of the plan within the data base. They also serve as a sensitive check for missing data.

The algorithm developed for forming the simple polygons is fast and easy to use. The polygon formation process is highly automated, with the assignment of attributes being the only operator action required. The creation of the super polygons, those for the blocks and plan, requires no operator assistance other than to indicate acceptance or rejection of the polygons formed. A rejection by the operator of any of the polygons formed by the program is an indication of missing data.

The calculation of the coordinates of the points is done using the method of constrained least squares. This powerful tool was chosen, as it uses all the observations shown on the plan in the determination of the best estimate, in the statistical sense, for the point positions.

This approach also enables the inclusion of topological conditions as constraints, increasing the accuracy and reliability of the solution. The observations shown on the plan can be thoroughly tested for errors using the results from the least squares adjustment and statistical analysis techniques. This combination of the least squares method and statistical analysis contributes greatly to the robustness of the system.

All these ideas have been implemented in a prototype program. The program is menu driven, using dual menu systems, and performs extensive validity tests on the input. In testing, the program has proven to be both simple to use and very tolerant of errors, both those in the data and those made by the operator. This implementation of the ideas and concepts developed in this thesis shows that they do result in a system for the entry of cadastral survey plans into a Land Information System, that meets the four objectives identified at the outset.

Once an individual cadastral plan has been processed it must be related to all the others through the use of the same continuous homogeneous datum. To achieve this, three approaches to plan fitting were examined. The first two, similarity transformations and rubber sheeting, were shown to be unacceptable. The first introduces discontinuities into the data while the second causes unacceptable distortions. The third approach, developed for this thesis, uses the information from existing plans to constrain the new plan. It was found to be the best solution to the problem of plan fitting. It results in an exact fit without introducing any distortion, and has the added advantage that plan fitting and coordinate computation can proceed simultaneously.

Finally the impact of implementing the cadastral data entry and cadastral adjustment programs (CDEP and CADJ) in an existing cadastral mapping agency was examined. It was found that this implementation would result in both a significant increase in productivity by the agency and an improvement in the quality of the product with no increase in costs.

## 9.2 Recommendations

The following recommendations result from the work in this thesis:

a) The CDEP and CADJ programs should be installed at the Land Information Services Division in Edmonton in order to fully evaluate them in a production environment using comprehensive data sets;

b) Should this comprehensive evaluation be successful, an analysis of the potential for the use of the programs by other provincial and municipal mapping agencies should be investigated with a view to their possible implementation. This may require some modification to the program to allow for differences in cadastral surveying practices in the various provinces;

c) A microcomputer implementation of CDEP and CADJ should be made available as soon as possible. This will make recommendation (b) above easier to realize and also make the use of the programs by small survey firms feasible; and

d) Large volumes of cadastral data currently exist in spaghetti data files. A version of CDEP should be developed that will use these files as a data source and produce the same topologically complete results as using CDEP with hardcopy plans. This will greatly facilitate the conversion of the existing files for inclusion into modern Land Information Systems currently under development.

## REFERENCES


Baker, R.F. (1987) Alberta Bureau of Surveying and Mapping, Edmonton, Canada, Personal Communication.

Bathke, William L. (1979) Land Records and Resource Information Systems: An Assessment From an Administrative Point of View, Proceedings of the 39th Annual Meeting, American Congress of Surveying and Mapping, Washington, D.C., March 18-24, pp. 380-387.

Baxter, Richard S. (1976) Computer and Statistical Techniques for Planners, Methuen and Co., London, U.K., 336 pp.

Beraha, W. (1986) LIS Programs and Initiatives at ABSM, Proceedings of the 1986 URISA Alberta Chapter Workshop, Calgary, Canada, January 30-31.

Blachut, T.J., A. Chrzanowski and J.H. Saastamoinen (1979) Urban Surveying and Mapping, Springer-Verlag, New York, N.Y., 372 pp.

Blais, J.A.R. (1979a) Least Squares Block Adjustment of Stereoscopic Models and Error Analysis, Ph.D. Thesis (UNB), The University of Calgary Surveying Engineering Publication #30001, Calgary, Canada, 208 pp.

Blais, J.A.R. (1979b) Least-squares Estimation in Geodetic Horizontal Control Densification, Collected Papers Geodetic Survey, Surveys and Mapping Branch, Energy Mines and Resources, Ottawa, Canada, pp. 31-53.

Blais, J.A.R. and M.A. Chapman (1983) Program SPACE-M: Users Guide, The University of Calgary Surveying Engineering Publication #50006, Calgary, Canada, 30 pp.

Burrough, P.A. (1986) Principles of Geographical Information Systems for Land Resources Assessment, Monographs on Soil and Resources Survey No. 12, Clarendon Press, Oxford, U.K., 193 pp.

Chapman, M.A. and J.A.R. Blais (1987) Integrating Digital Map Files, Proceedings of the 1987 URISA Alberta Chapter Spring Conference, Edmonton, Canada, April 27-28, pp. 133-144.

Chen, P. (1977) The Entity-Relationship Approach to Logical Data Base Design, No. 6 in QED Monograph Series on Data Base Management, QED Information Sciences Inc., Wellesley, Massachusetts, 33 pp.

CRC (1976) Standard Math Tables, 24th Edition, CRC Press, Cleveland, Ohio, 565 pp.

Cremont, Daniel (1980) Application of Computer Technology in Processing of Cadastral Surveying and Mapping Data, The Canadian Surveyor, Vol. 34, No. 1, pp. 21-40.

Enderle, G., K. Kansy and G. Pfaff (1987) Computer Graphics Programming: GKS - The Graphics Standard, Second Edition, Springer-Verlag, New York, N.Y., 651 pp.

Feuchtwanger, M. and T.K. Poiker (1987) Infological and Datalogical Spatial Models, in Lecture Notes in Digital Mapping and Land Information, Canadian Institute of Surveying and Mapping, Ottawa, Canada, pp. 146-167.

Gagnon, Roy (1988) Cadastral Plotting by Similarity Transformation, CISM Journal, Vol. 42, No. 2, pp. 121-125.

Giffen, Barry (1986) Alberta Land Registry: A Perspective, Proceedings of First Annual URISA (Alberta Chapter) Conference, Edmonton, Canada.

Gillman, Daniel W. (1985) Triangulation for Rubber Sheeting, Proceedings of Auto Carto 7, Washington, D.C., March 11-14, American Society of Photogrammetry, Falls Church, Virginia, pp. 191-199.

Giroux, Christian (1987) New Mathematical Method for the Integration of Subdivision Plans, Research Paper, Service du Cadastre, Gouvernement du Québec, Quebec City, Canada, 11 pp.

Government of Alberta (1980) Surveys Act, Chapter S-29, Revised Statutes of Alberta, Queens Printer, Edmonton, Canada.

Hamilton, W. C. (1964) Statistics in the Physical Sciences, The Ronald Press Company, New York, N.Y., 230 pp.

Kennedy, Ed (1986) LIS Programs and Initiatives at ABSM, Proceedings of the 1986 URISA Alberta Chapter Workshop, Calgary, Canada, January 30-31.

Kreyszig, Erwin (1972) Advanced Engineering Mathematics, Third Edition, John Wiley and Sons, New York, N.Y., 866 pp.

Leithold, Louis (1972) The Calculus with Analytic Geometry - Part 2, Second Edition, Harper and Row, New York, N.Y., 359 pp.

Lodwick, G.D. (1981) Topographic Mapping using Landsat Data, Proceedings of Thirteenth International Symposium on Remote Sensing of Environment, University of Michigan, May 11-15, ERIM, Ann Arbor, Michigan, pp. 527-533.

Lodwick, G.D. and M. Feuchtwanger (1987) Position-Based Land Information Systems, The University of Calgary Surveying Engineering Publication #10010, Calgary, Canada 223 pp.

Lodwick, G.D., S.H. Paine, M.P. Mepham and A.W. Colijn (1986) A Comprehensive LRIS of the Kananaskis Valley using Landsat Data, Proceedings of the Seventh International Symposium on Remote Sensing for Resources Development and Environmental Management, Enschede, Netherlands, Aug 25-29, pp. 927-932.

Mackenzie, A. Philip (1985) Design and Assessment of Horizontal Survey Networks, M.Sc.Eng. Thesis, The University of Calgary Surveying Engineering Publication #20011, Calgary, Canada, 201 pp.

Mepham, Michael P. (1983a) Computer Aided Survey Network Design. M.Sc.Eng. Thesis, The University of Calgary Surveying Engineering Publication #20003, Calgary, Canada, 98 pp.

Mepham, M.P. (1983b) Blunder Detection Using Graph Theory. Report from the Division of Surveying Engineering, The University of Calgary to the Alberta Bureau of Surveying and Mapping, 27 pp.

Mepham, M.P. (1987a) Automating the Inspection and Coordination of Subdivision Plans, Proceedings of the 1987 URISA Alberta Chapter Spring Conference, Edmonton, Canada, April 27-28, pp. 111-126.

Mepham, M.P. (1987b) The Resolution of Discrepancies at Common Boundaries Between Legal Survey Plans, Proceedings of the 1987 URISA Alberta Chapter Fall Conference, Calgary, Canada, November 30 - December 1, pp. 45-53.

Mepham, M.P. (1988) SCRMNU Users Guide, Unpublished Report, The Department of Surveying Engineering, The University of Calgary, Calgary, Canada, 60 pp.

Mepham, M.P., G.D. Lodwick and W.J. Mintz (in prep.) Impact of Automation of Legal Survey Plans on Surveying and Mapping Institutions.

Mikhail, Edward M. (1976) Observations and Least Squares, IEP, New York, N.Y., 497 pp.

Mintz, W.J. (1988) Alberta Bureau of Surveying and Mapping Edmonton, Alberta, Personal Communication.

Mulaku, Galcano (1987) Map Data Digitizing: Editing and Automatic Hydrological Network Reconstruction, M.Sc.Eng. Thesis, The University of Calgary Surveying Engineering Publication #20021, Calgary, Canada, 142 pp.

Newman, William M. and Robert F. Sproull (1979) Principles of Interactive Computer Graphics, McGraw Hill, New York, N.Y., 541 pp.

Petersohn, Cliff and Alan P. Vonderhoe (1982) Site-Specific Accuracy of Digitized Property Maps, Proceedings of Auto Carto 5, Crystal City, Virginia, August 22-28, American Society of Photogrammetry, Falls Church, Virginia, pp. 607-619.

Phillips, John O. (1975) Opening Address, Proceedings of the North American Conference on Modernization of Land Data Systems, Washington, D.C., April 14-17, North American Institute for Modernization of Land Data Systems, Washington, D.C., pp. 1-3.

Province of Manitoba (1986) Dept. of Natural Resources, Surveys and Mapping Branch CADCOR: A Program to Coordinate Cadastral Points, Winnipeg, Canada, 84 pp.

Rees, Paul K. and Fred W. Sparks (1969) Calculus with Analytic Geometry, McGraw-Hill, New York, N.Y., 617 pp.

Robinson, W.G. (1982) The Perils of Prorating, Proceedings of the Canadian Institute of Surveying Centennial Convention, Canadian Institute of Surveying and Mapping, Ottawa, Canada, pp. 270-290.

Vaniček, Petr and Edward J. Krakiwsky (1982) Geodesy: The Concepts, North-Holland Publishing Company, Amsterdam, Netherlands, 691 pp.

Wahl, Bernardo J. (1975) Technical Features Essential to a Modern Multi-Purpose Land Data System, <u>Proceedings of the North American Conference on Modernization of Land Data Systems</u> Washington D.C., April 14-17, North American Institute for Modernization of Land Data Systems, Washington, D.C., pp. 95-114.