UNIT -2

DESIGNING OF SPATIAL DATABASE

2. Designing of Spatial Database: Identification of Geographic features – attributes & data layer – Defining the storage parameters for each attribute – ensuring of co-ordinate registration – map projection – Transformation. - **12 Hrs.**

Spatial Database : Definition

- A spatial database is a database that is optimized to **store and query data related to objects in space, including points, lines and polygons.**

- While typical databases can understand various **numeric and character types of data**, additional functionality needs to be added for databases to process spatial data types. These are typically called **<u>geometry or feature.</u>**

Spatial Databases Background

- Spatial databases provide structures for **storage and analysis of spatial data**
- Spatial data is comprised of objects in multi-dimensional space
- Storing spatial data in a standard database would require excessive amounts of space

Spatial Databases Background (Cont.)

- Queries to retrieve and analyze spatial data from a standard database would be long and cumbersome leaving a lot of room for error
- Spatial databases provide much more efficient storage, retrieval, and analysis of spatial data

Types of Data Stored in Spatial Databases

- **Two-dimensional data** examples
  - Geographical
  - Cartesian coordinates (2-D)
  - Networks
  - Direction

Types of Data Stored in Spatial Databases (Cont.)

- **Three-dimensional data** examples
  - Weather
  - Cartesian coordinates (3-D)
  - Topological
  - Satellite images

Spatial Databases Uses and Users

- **Three types of uses**
  - **Manage spatial data**
  - **Analyze spatial data**
  - **High level utilization**

Spatial Databases Uses and Users (Cont.)

- A few examples of users
  - Transportation agency tracking projects
  - Insurance risk manager considering location risk profiles
  - Doctor comparing Magnetic Resonance Images (MRIs)
  - Emergency response determining quickest route to victim
  - Mobile phone companies tracking phone usage

Spatial Database Management System

- Spatial Database Management System (SDBMS) provides the capabilities of a traditional database management system (DBMS) while allowing special storage and handling of spatial data.

Spatial Database Management System (Cont.)

- SDBMS:
  - Works with an **underlying DBMS**
  - Allows spatial data models and types
  - Supports **querying language specific to spatial data types**
  - Provides handling of spatial data and operations

Spatial Query Language

- Number of specialized adaptations of SQL
  - **Spatial** query language
  - **Temporal** query language (TSQL2)
  - **Object** query language (OQL)
  - **Object oriented structured** query language (O_2SQL)

Features of Spatial Databases (Cont.) - query types

The following query types and many more are supported by the Open Geospatial Consortium:

- **Spatial Measurements**: Finds the distance between points, polygon area, etc.

Features of Spatial Databases (Cont.) - query types

- **Spatial Functions**: Modify existing features to create new ones, for example by providing a **buffer around them, intersecting features,** etc.

- **Spatial Predicates**: Allows true/false queries such as 'is there a residence located within a mile of the area we are planning to build the landfill?'

Features of Spatial Databases (Cont.) - query types

- Constructor Functions: Creates new features with an SQL query specifying the vertices (points of nodes) which can make up lines. If the first and last vertex of a line are identical the feature can also be of the type polygon (a closed line).
- Observer Functions: Queries which return specific information about a feature such as the location of the center of a circle

Types of queries - PostGIS

The function names for queries differ across geodatabases. The following list contains commonly used functions built into PostGIS, a free geodatabase which is a PostgreSQL extension (the term 'geometry' refers to a point, line, box or other two or three dimensional shape):

Types of queries - PostGIS (Cont.)

1. **Distance**(geometry, geometry) : number
2. **Equals**(geometry, geometry) : boolean
3. **Disjoint**(geometry, geometry) : boolean
4. **Intersects**(geometry, geometry) : boolean
5. **Touches**(geometry, geometry) : boolean
6. **Crosses**(geometry, geometry) : boolean

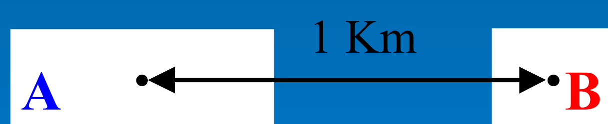
Types of queries - PostGIS (Cont.)

7. **Overlaps**(geometry, geometry) : boolean
8. **Contains**(geometry, geometry) : boolean
9. **Intersects**(geometry, geometry) : boolean
10. **Length**(geometry) : number
11. **Area**(geometry) : number
12. **Centroid**(geometry) : geometry

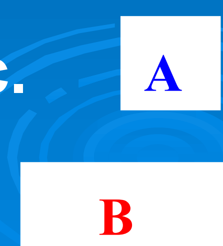
Spatial Relations

- **Topological Relations**: containment, overlapping, etc. [Egenhofer et al. 1991]

- **Metric Relations**: distance between objects, etc. [Gold and Roos 1994]

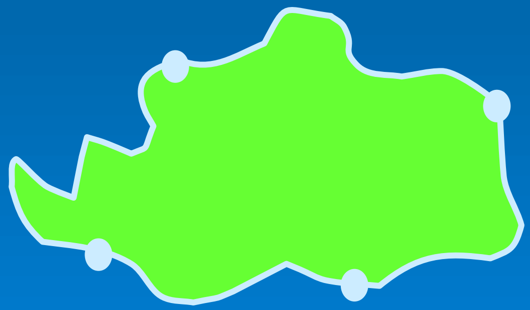


- **Direction Relations**: north of, south of, etc. [Hernandez et al. 1990; Frank et al. 1991]



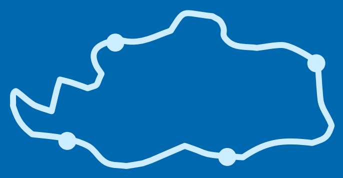
Topological Relations

- **Topological relations are defined using *point-set topology* concepts, such as *boundary* and *interior.***

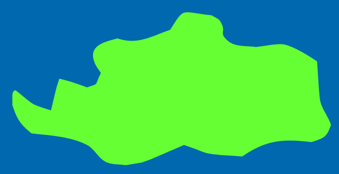


Topological Relations (Cont.)

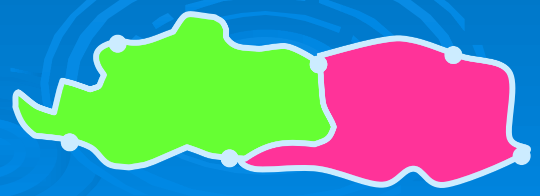
- **For example:**
  - **the *boundary* of a region consists of a set of curves that separate the region from the rest of the coordinate space**



  - **The *interior* of a region consists of all points in the region that are not on its boundary**



- **Given this, two regions are said to be *adjacent* if they share part of a boundary but do not share any points in their interior**



Spatial Relations Model

- **An abstract model** (or conceptual model): is a theoretical construct that represents something, with a set of variables and a set of logical and quantitative relationships between them.
- Models in this sense are constructed to enable reasoning within an idealized logical framework about these processes and are an important component of scientific theories.

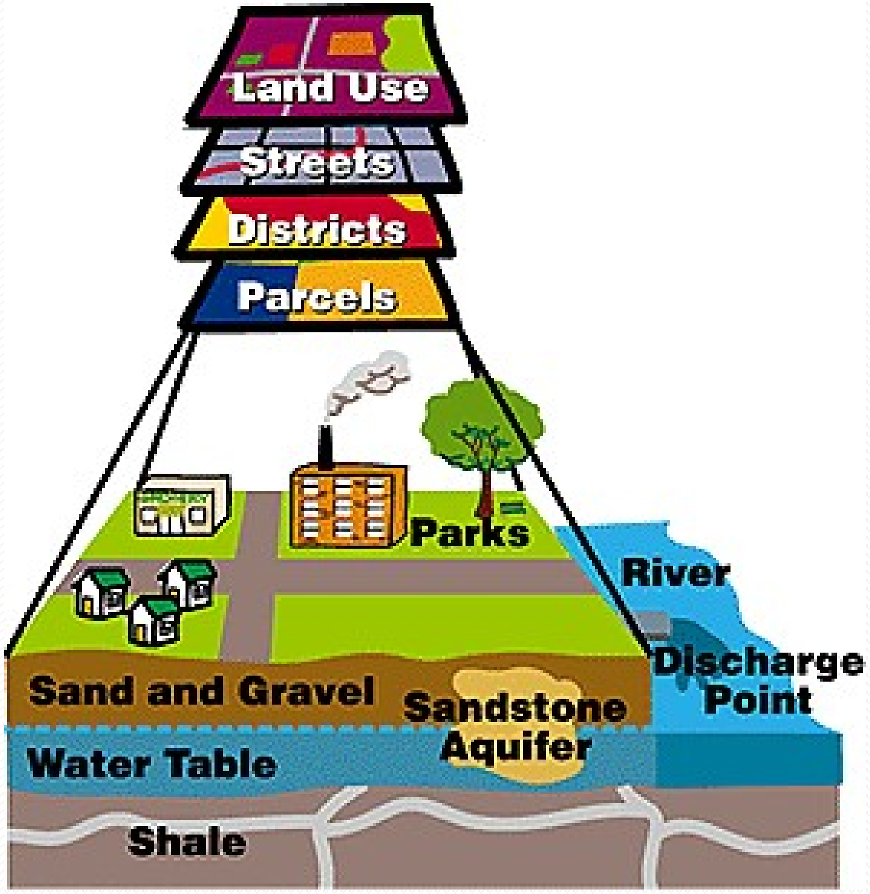
GIS DATA MODELS

GIS DATA MODELS

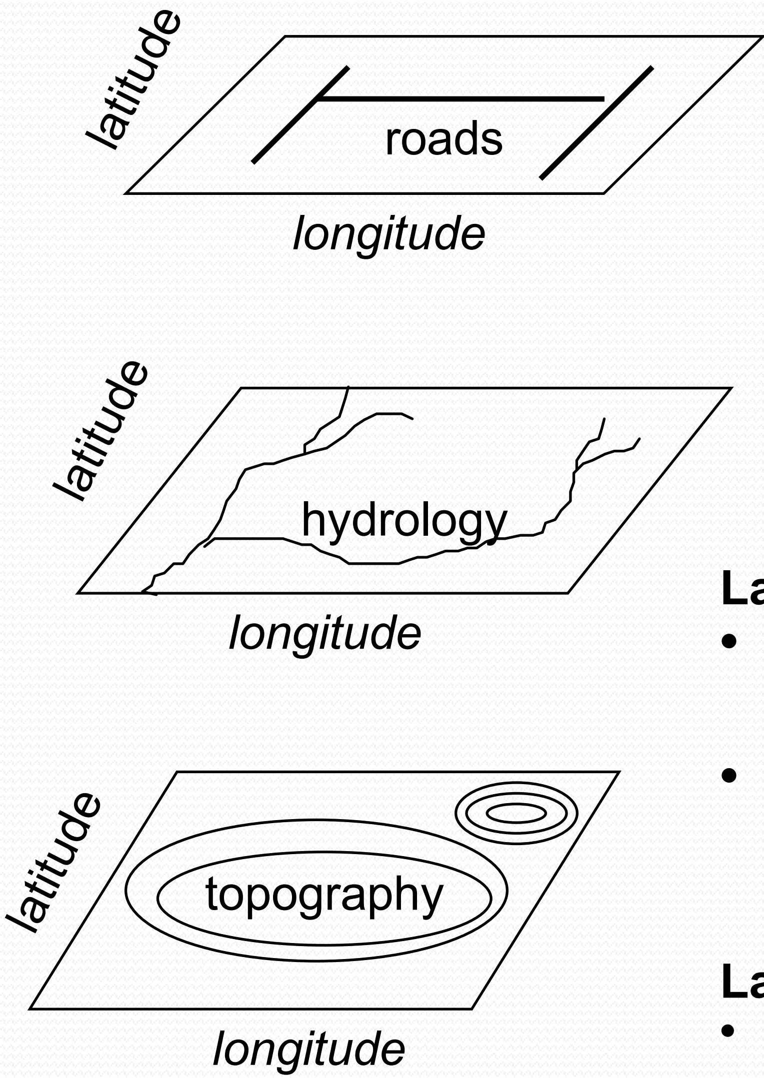
A GIS stores information about the world as a collection of thematic layers that can be linked together by geography.

This may look simple but extremely powerful in solving many real-world problems

The thematic layer approach allows us to organize the complexity of the real world into a simple representation



THE GIS MODEL: EXAMPLE



Here we have three <u>layers</u> or themes:
--roads,
--hydrology (water),
--topography (land elevation)

They can be related because precise geographic coordinates are recorded for each theme.

Layers are comprised of two data types

- ***Spatial data*** **which describes location (where)**
- ***Attribute data*** **specifing what, how much,when**

Layers may be represented in two ways:

- **in *vector* format as points and lines**
- **in *raster(or image)* format as pixels**

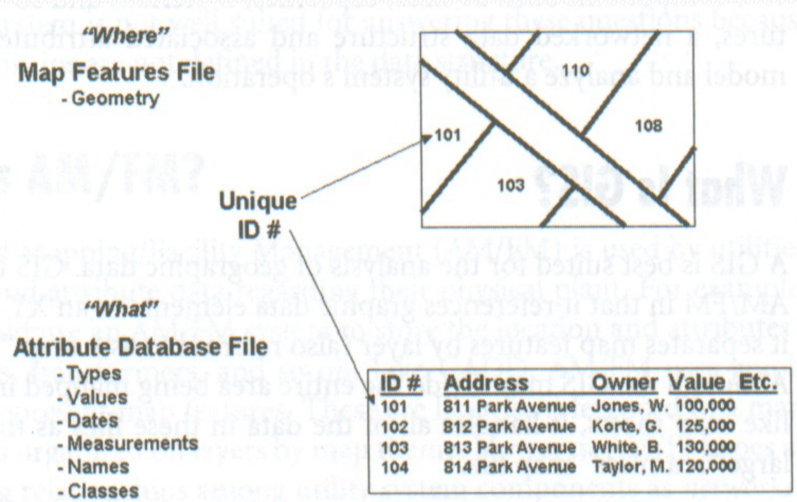
GIS DATA TYPES

GIS DATA TYPES

GIS technology utilizes two basic types of data. These are:

Spatial data describes the absolute and relative location of geographic features.

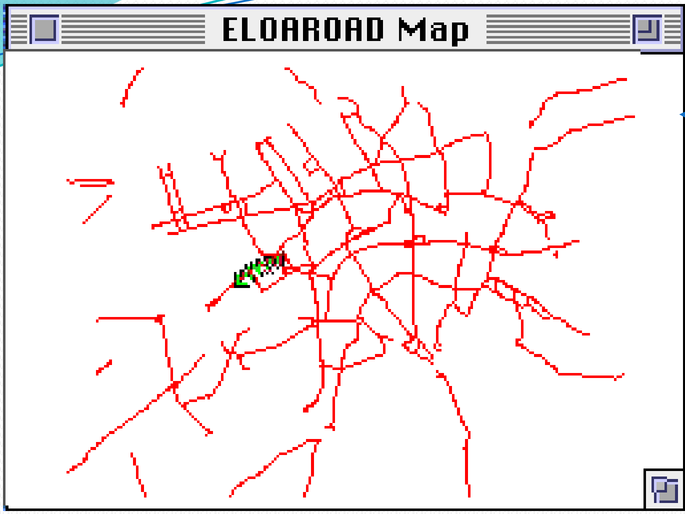
Attribute data describes characteristics of the spatial features. These characteristics can be quantitative and/or qualitative in nature. Attribute data is often referred to as tabular data.

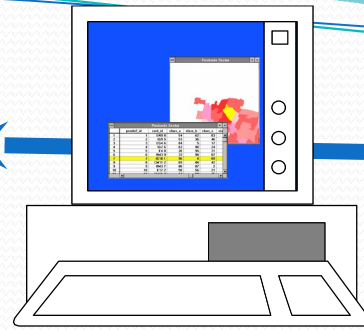


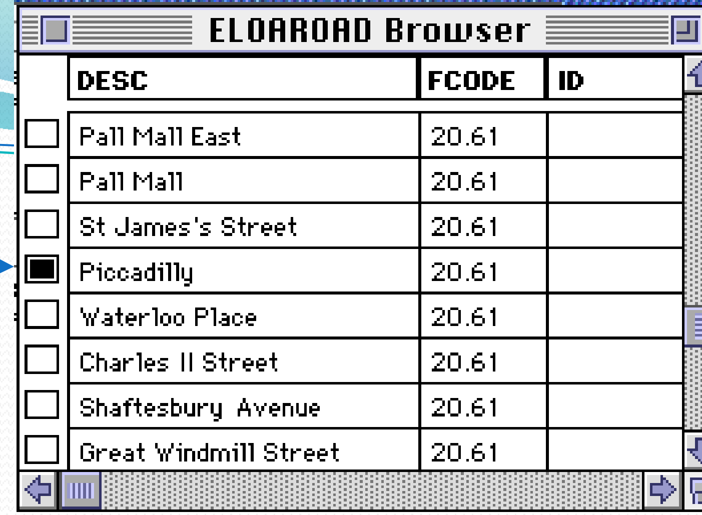
The coordinate location of a forestry stand would be spatial data, while the characteristics of that forestry stand, e.g. cover group, dominant species, crown closure, height, etc.,

- **Spatial data *(where)***
  - specifies location
  - stored in a *shape file, geodatabase* or similar geographic file
- **Attribute (descriptive) data** *(what, how much, when)*
  - specifies characteristics at that location, natural or human-created
  - stored in a data base *table*

GIS systems traditionally maintain spatial and attribute data separately, then "join" them for display or analysis

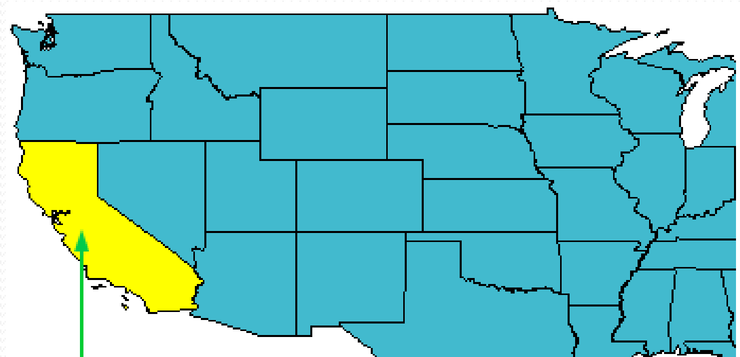


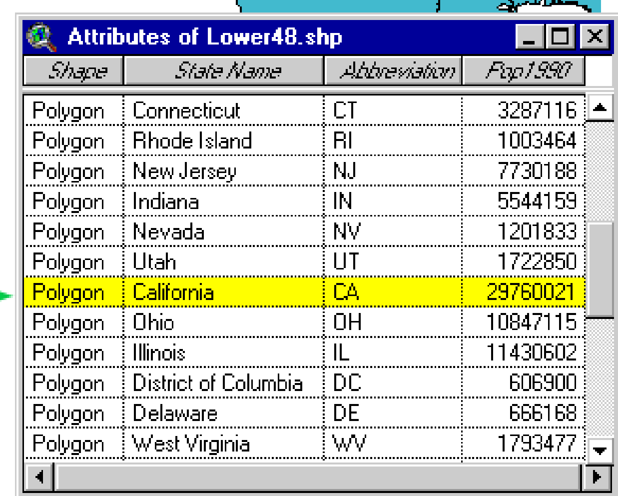




ELOAROAD Browser

	DESC	FCODE	ID
☐	Pall Mall East	20.61	
☐	Pall Mall	20.61	
☐	St James's Street	20.61	
■	Piccadilly	20.61	
☐	Waterloo Place	20.61	
☐	Charles II Street	20.61	
☐	Shaftesbury Avenue	20.61	
☐	Great Windmill Street	20.61	

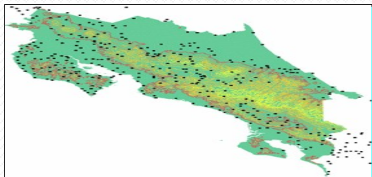
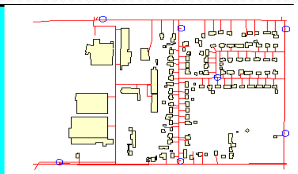




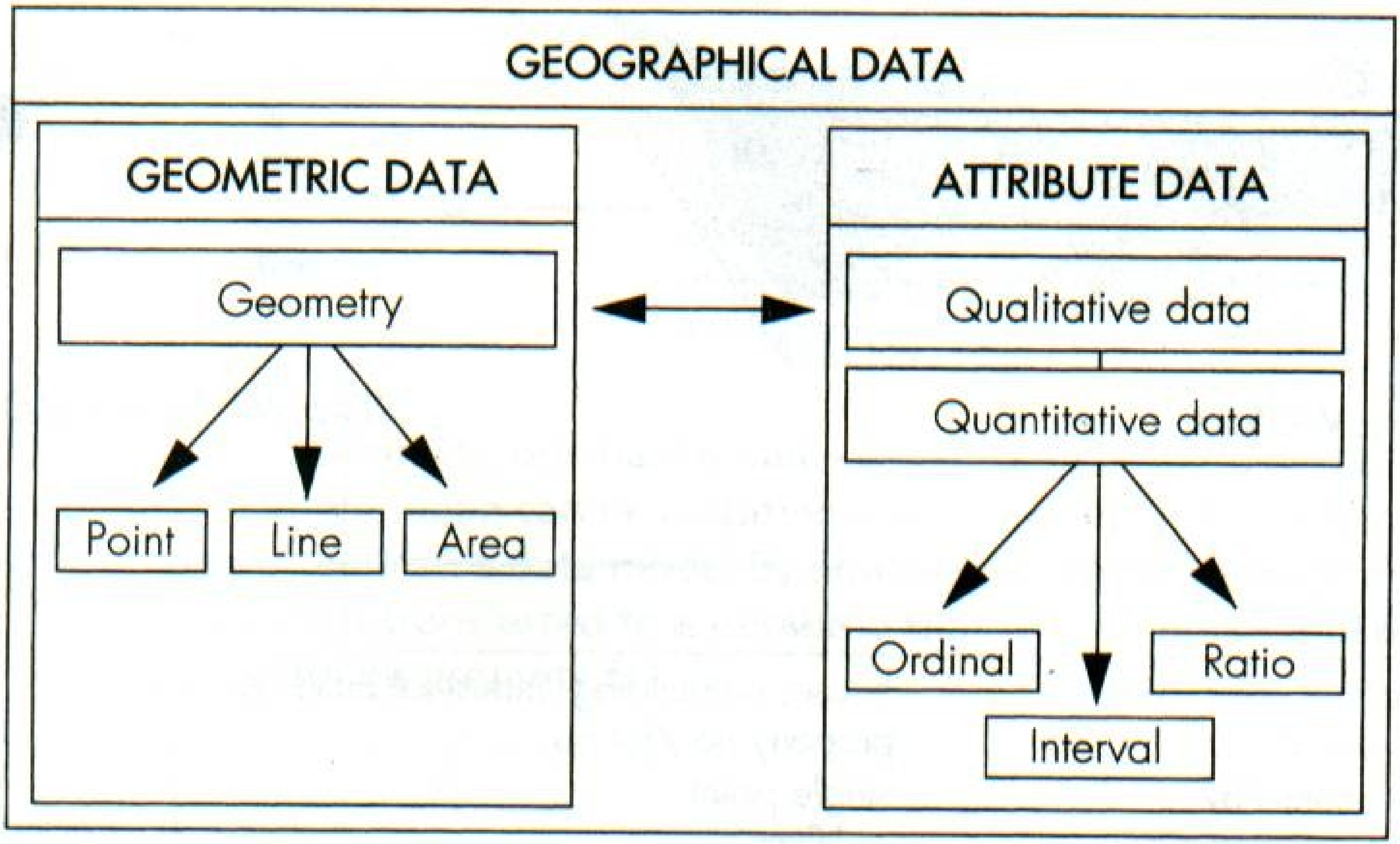
Attributes of Lower48.shp

Shape	State Name	Abbreviation	Pop1990
Polygon	Connecticut	CT	3287116
Polygon	Rhode Island	RI	1003464
Polygon	New Jersey	NJ	7730188
Polygon	Indiana	IN	5544159
Polygon	Nevada	NV	1201833
Polygon	Utah	UT	1722850
Polygon	California	CA	29760021
Polygon	Ohio	OH	10847115
Polygon	Illinois	IL	11430602
Polygon	District of Columbia	DC	606900
Polygon	Delaware	DE	666168
Polygon	West Virginia	WV	1793477

Types of data

Spatial	non-spatial/ Attribute
 Maps	Schematic diagrams 
 Images	Oblique photographs 
Videography	Films
KT1 2EE RH8 9AA SW1P 3AD Postcodes/ZIP codes	Financial statements £12,000 23.45 56789 £23,456 12.45 23456 £45,987 29.57 87634

SPATIAL & ATTRIBUTE DATA



VECTOR DATA TYPES

VECTOR DATA TYPES

- **All spatial data models are approaches for storing the spatial location of geographic features in a database.**
- **Spatial locations are explicit**
- **Relationships between entities/objects are implicit**
- **Vector storage implies the use of vectors (directional lines) to represent a geographic feature.**

All real world features are represented by three vector types

Point: features are defined by one coordinate pair, a vertex.

Each vertex consists of an X coordinate and a Y coordinate

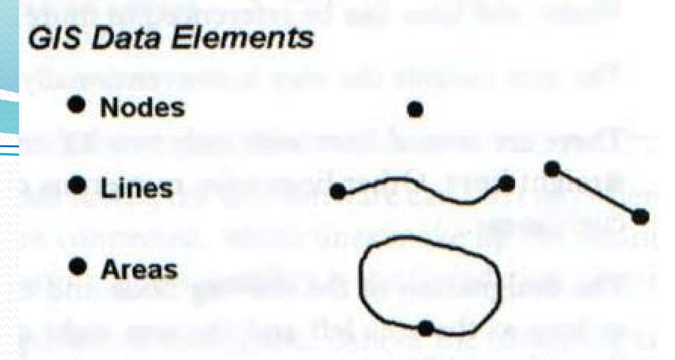
Line: Vector lines are often referred to as arcs and consist of a string of vertices terminated by a node.

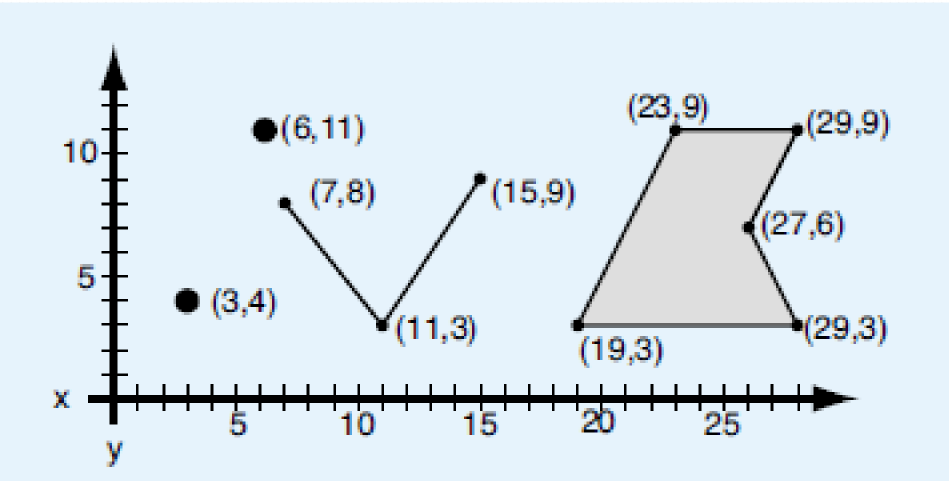
A node is defined as a vertex that starts or ends an arc segment

Polygon: Polygonal features are defined by a set of closed coordinate pairs.

In vector representation, the storage of the vertices for each feature is important, as well as the connectivity between features, e.g. the sharing of common vertices where features connect.

VECTOR DATA: TYPES



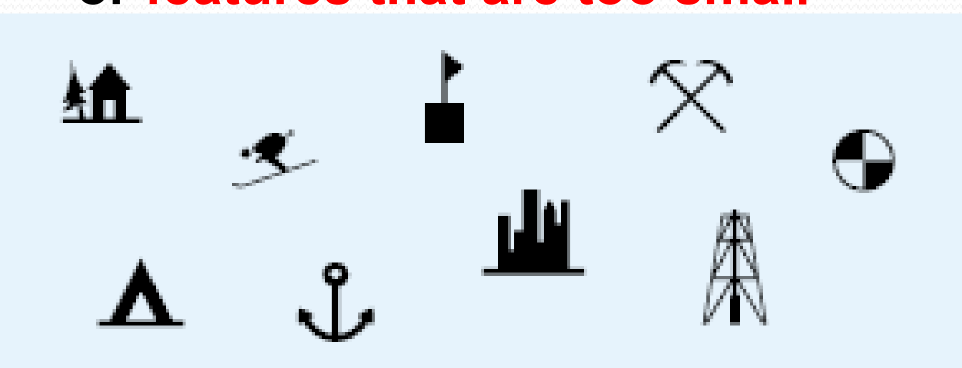


Vector data are defined spatially as either:

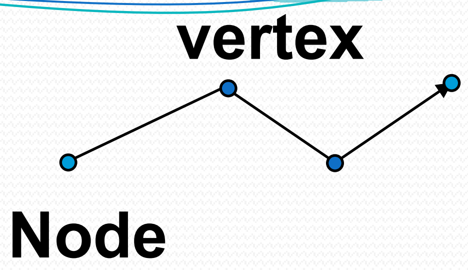
(x_1,y_1)

Point/Node **- a pair of x and y coordinates**

Points represent geographic features that have no area or length, or features that are too small

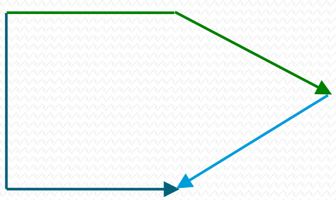


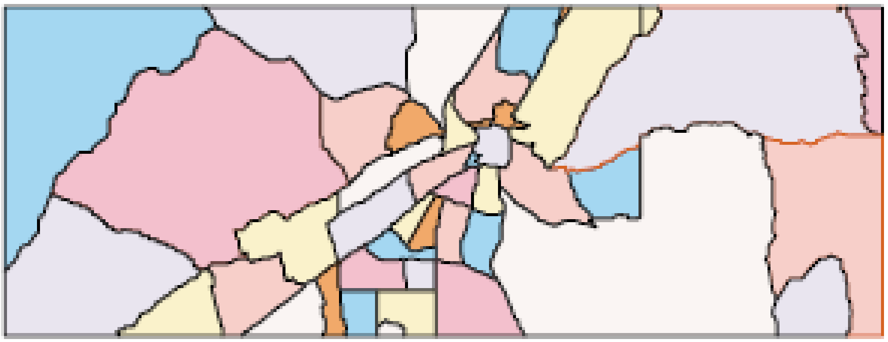
Line **- a sequence of points**



Linear features represent objects that have length but no area, or features whose shapes are very narrow at a given map scale.

Polygon/Area **- a closed set of lines**





Polygons enclose areas that meet a user-specified set of common characteristics for the phenomena being represented.

SPATIAL DATA STORAGE MODEL

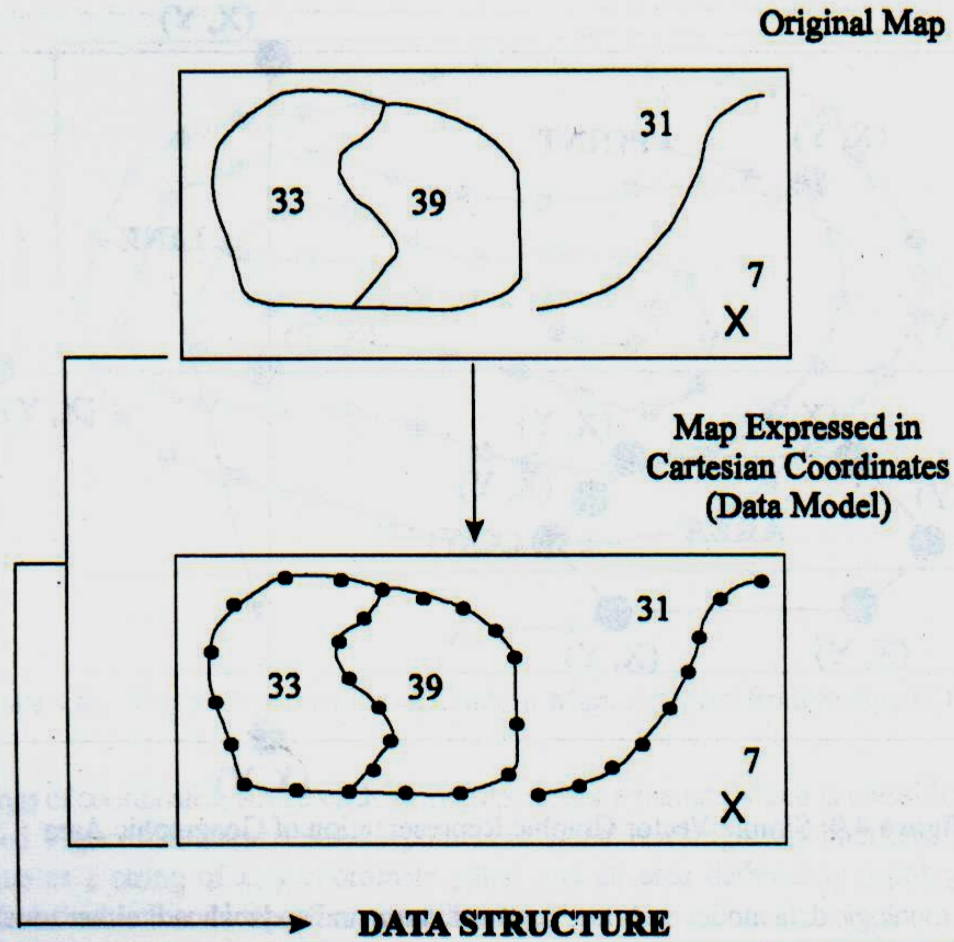
VECTOR DATA MODEL

VARIETY OF VECTOR MODELS

- Spaghetti model
- Topological model (most common)
- Triangulated irregular network (TIN)
- Dime files and TIGER files
- Network model
- Digital Line Graph (DLG)
- Shapefile (ArcView/ArcGIS; ESRI)
- Others: HPGL, PostScript/ASCII, CAD/.dxf

SPAGHETTI

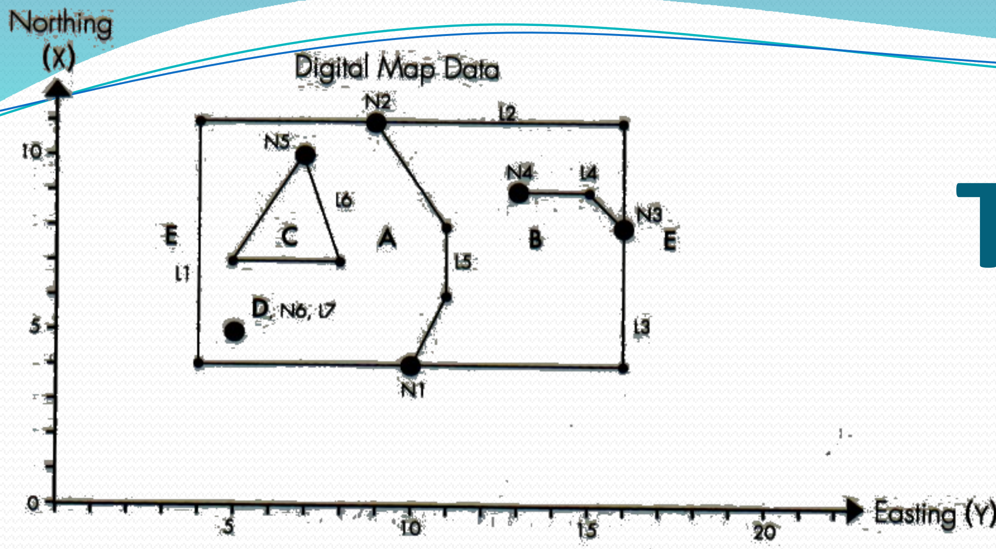
- Non-topological



FEATURE	NUMBER	LOCATION
POINT	7	X Y (SINGLE POINT)
LINE	31	X_1Y_1 X_2Y_2 ... X_NY_N (STRING)
POLYGON	33 39	X_1Y_1 X_2Y_2 ... X_NY_N (CLOSED LOOP) X_1Y_1 X_2Y_2 ... X_NY_N (CLOSED LOOP)

Figure 4.10: The "Spaghetti" Data Model. Source: Adapted from drawing presented by Dangermond (1983).

TOPOLOGICAL



(a)

(b)

POLYGON TOPOLOGY

Polygon	Links
A	L1, L5
B	L2, L3, L5
C	L6
D	L7
E	L1, L2, L3

(c)

NODE TOPOLOGY

Node	Links
N1	L1, L3, L5
N2	L1, L2, L5
N3	L2, L3, L4
N4	L4
N5	L6
N6	L7

(d)

LINK TOPOLOGY

Links	Start node	End node	Left polygon	Right polygon
L1	N1	N2	E	A
L2	N2	N3	E	B
L3	N3	N1	E	B
L4	N3	N4	B	B
L5	N2	N1	B	A
L6	N5	N5	A	C
L7	N6	N6	A	A

(e)

LINK COORDINATES

Link	Coordinates			
L1	4,10	4,4	11,4	11,9
L2	11,9	11,16	8,16	
L3	8,16	4,16	4,10	
L4	8,16	9,15	9,13	
L5	11,9	8,11	6,11	4,10
L6	10,7	7,8	7,5	10,7
L7	5,5			

The most popular method of retaining spatial relationships among features is to explicitly record adjacency information in what is known as the topologic data model.

The topologic data structure is often referred to as an *intelligent data structure* because spatial relationships between geographic features are easily derived when using them.

Topology is a mathematical concept that has its basis in the principles of feature adjacency and connectivity.

All geometric objects are represented by nodes and links

- **(a) The objects’ attributes and relationships can be described by storing nodes and links in three tables:**
- **(b) a polygon table**
- **(c) a node topology table and**
- **(d) a link topology table**
- **(e) an additional table gives the objects’ geographical coordinates and is stored separately from the attribute data files**

Why Topology Matters

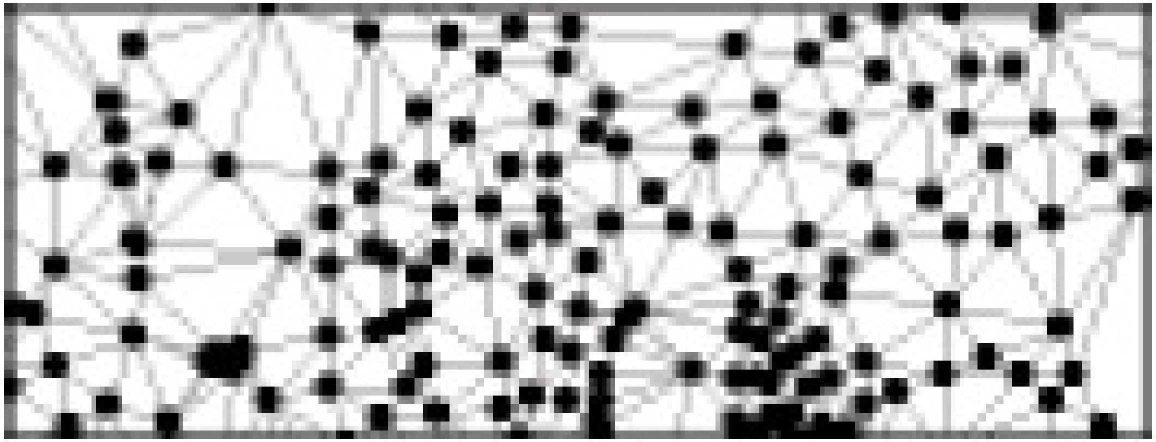
- **Connections & relationships between objects are independent of their coordinates**
- **Overcomes major weakness of spaghetti model – allowing for GIS analysis (Overlaying, Network, Contiguity, Connectivity)**
- **Requires all lines be connected, polygons closed, loose ends removed.**

TIN (TRIANGULATED IRREGULAR NETWORKS)

A Triangulated Irregular Network (TIN) is an efficient and accurate model for representing continuous surfaces

For each point, GIS software creates an optimal network of triangles, called a Delaunay triangulation. In a TIN, each triangle is created to be as close to equilateral as possible.

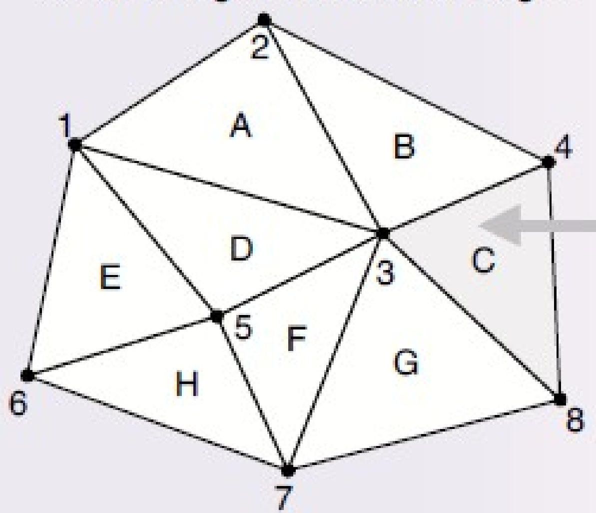
Each triangle forms a face with a gradient slope.



TIN topological data structure manages information about the nodes that comprises each triangle and the neighbours to each triangle

Topology in a TIN

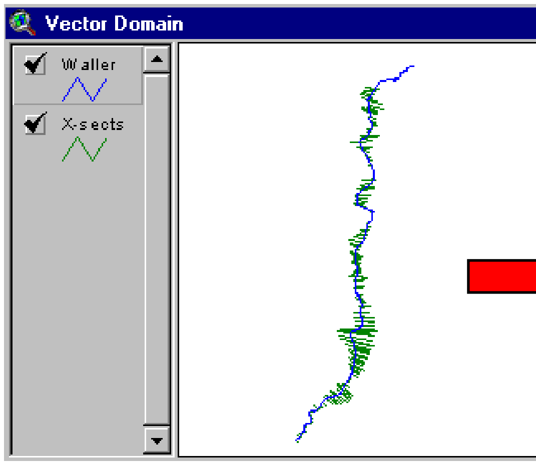
A TIN is a topological data structure that manages information about the nodes that comprise each triangle and the neighbors to each triangle.

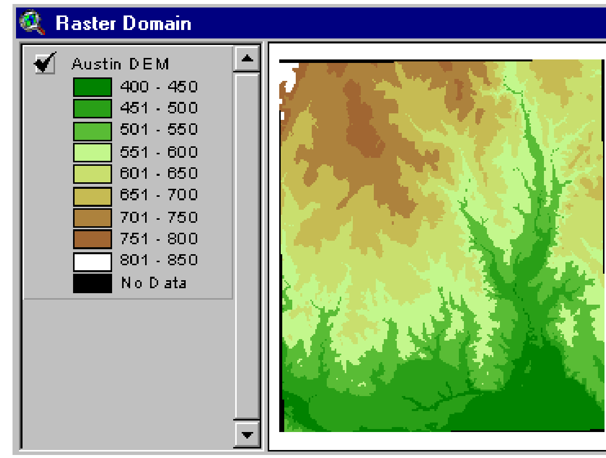


Triangle	Node list	Neighbors
A	1, 2, 3	–, B, D
B	2, 4, 3	–, C, A
C	4, 8, 3	–, G, B
D	1, 3, 5	A, F, E
E	1, 5, 6	D, H, –
F	3, 7, 5	G, H, D
G	3, 8, 7	C, –, F
H	5, 7, 6	F, –, E

Triangles always have three nodes and usually have three neighboring triangles. Triangles on the periphery of the TIN can have one or two neighbors.

Triangulated Irregular Network (TIN)

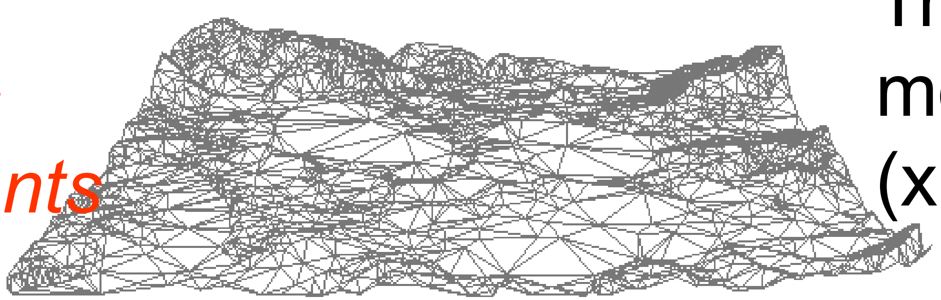




Inside Channel:
Vector cross-sections

Outside Channel:
Raster DEM

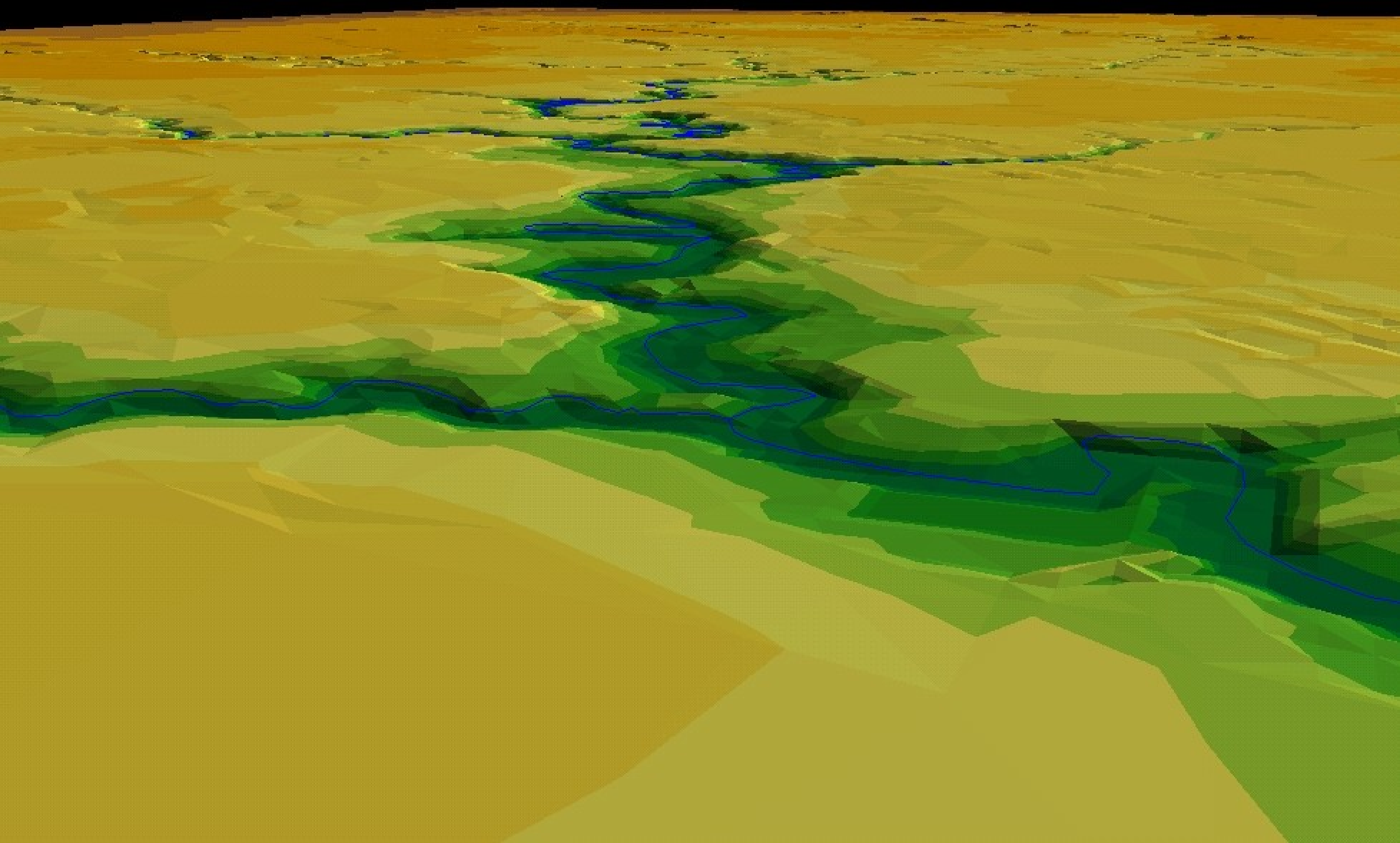
Triangles as elements

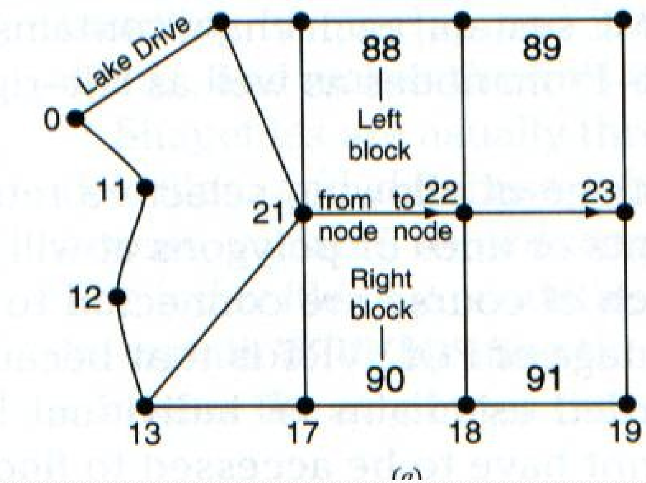


Triangular mesh of (x,y,z) points

Result: TIN of channel and landscape

Three-Dimensional View of a TIN
(Clear Creek and Cowarts Creek Confluence)

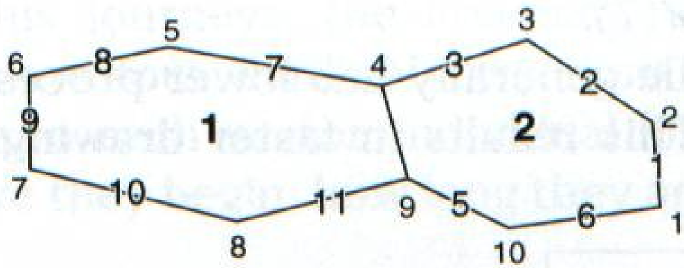




From Node		To Node		Blocks		Left Addr		Right Addr	
Map	Node	Map	Node	Left	Right	Low	High	Low	High
3	21	3	22	88	90	111	1999	102	198

- **GBF/DIME (geographic base file/dual independent map encoding):**
- **created by US Bureau of the Census to automate storage of street map data.**
- **Each segment ends when it either changes direction or intersects another line**
- **Nodes identified with codes**
- **Directional codes assigned "from node; to node" – facilitates error checking**
- **street address & UTM coordinates for each link are explicitly defined – facilitating address matching**
- **Disadvantage: no order in which line segments occur in system – searching sequential slow**

- **POLYVRT (POLYgon conVeRTer):**
  - **developed by Peucker & Chrisman & later implemented at Harvard Lab for Computer Graphics**
  - **In basic topological model each type of entity is stored separately as : points, lines and polygons.**
  - **Here all are interelated linked in a hierarchical data structure :points relating to lines, in turn related to polygons, all through the use of pointers.**
  - **Each collection of line segments, collectively called *chains* begins & ends w/specific nodes (intersection between 2 chains)**
  - **advantage: reduces storage required, faster retrieval**
  - **disadvantage: difficult to detect incorrect pointers for any given polygon until actually retrieved**



Polygons		Chain list	Chains						
Polygon name	Polygon pointer	Chain list	Chain name	Chain points (X,Y strings)	Chain length	From node	To node	Left polygon	Right polygon
1	.	4	4	x,y, x,y ...	.	4	9	2	1
2	.	5	5	.	.	.	.	.	.
.	.	6	6	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.

(c)

- **TIGER formats are from the enumeration maps of the US Census Bureau; designed for use with the 1990 US census**

- **Consist of an arc/node arrangement, w/separate files for points, lines, & areas linked by cross reference.**

- **TIGER terminology calls points *zero cells*, lines *one cells*, and areas *two cells*.**

- **By cross indexing some features can be encoded as landmarks e.g. rivers roads, buildings**

- **TIGER files exist for entire US**

TIGER

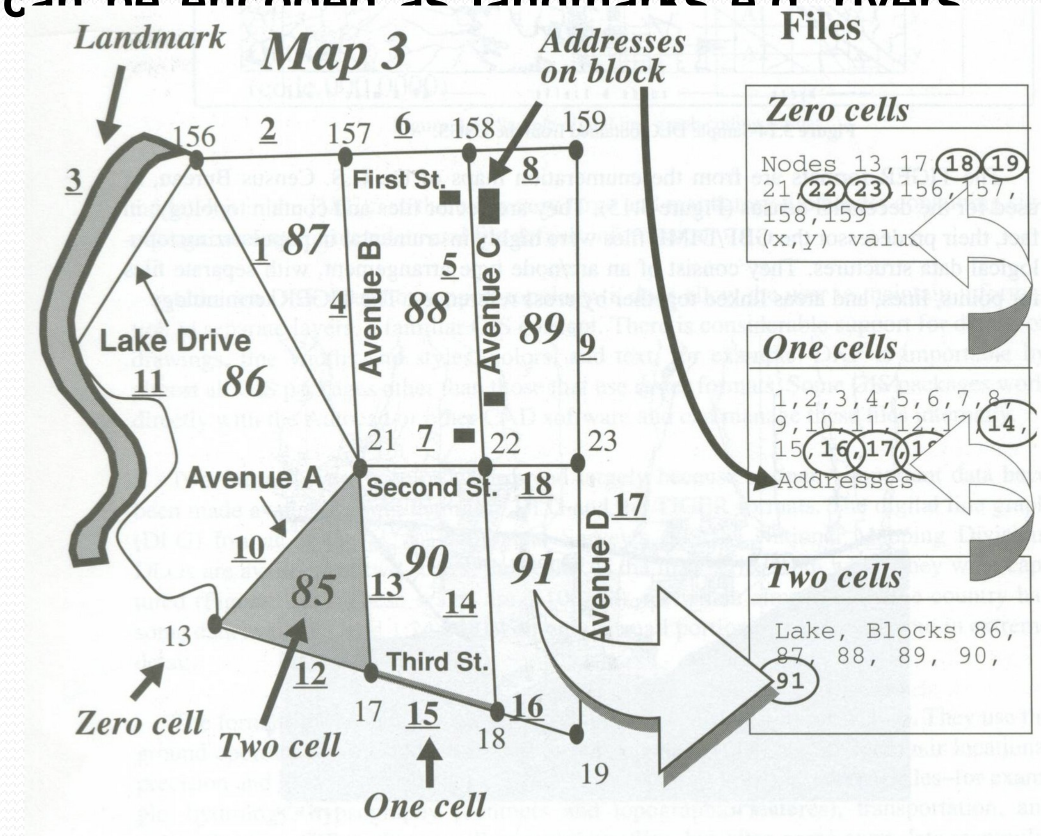


Figure 3.16 The U.S. Census Bureau's TIGER data structure.

- DLG format of the US Geological Survey's (USGS) National Mapping Division
- Features handled in separate files – hydrology, hypsography (contours & topographic features), transportation, political, etc.

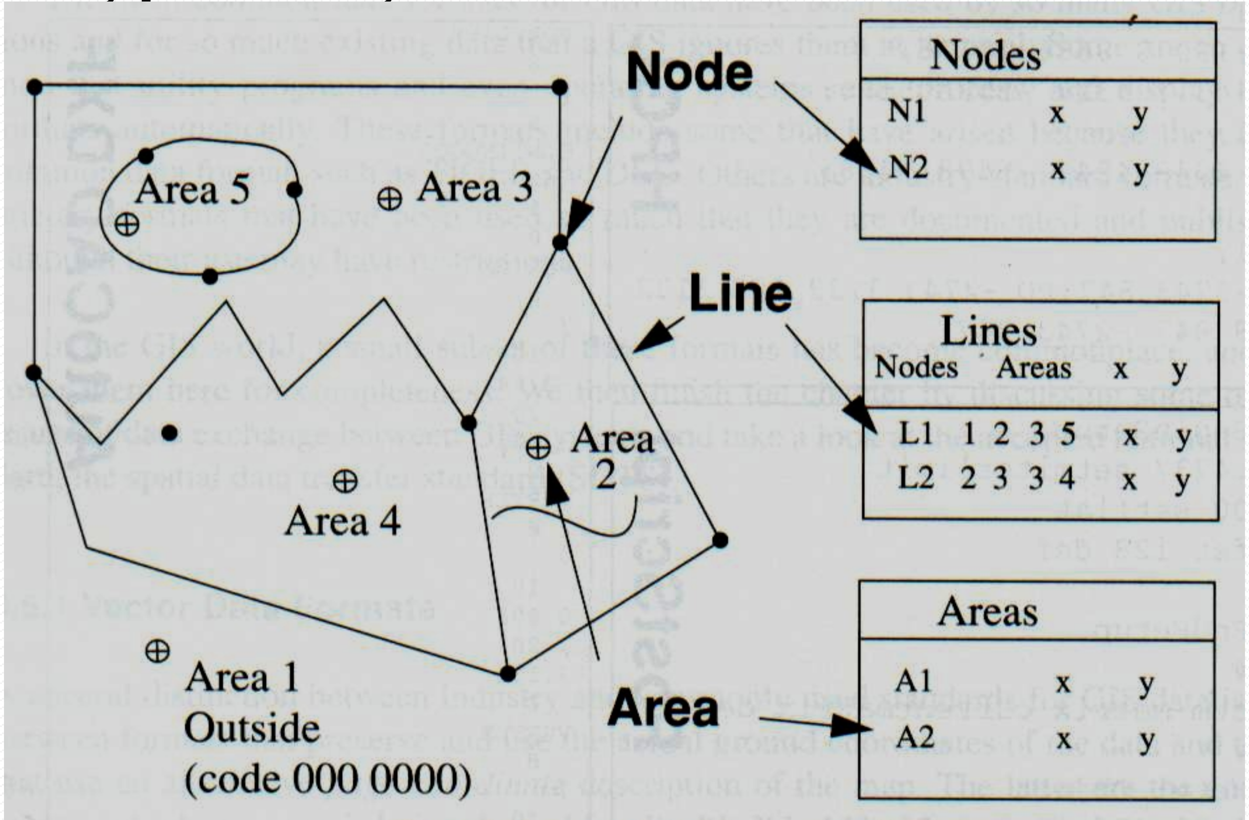
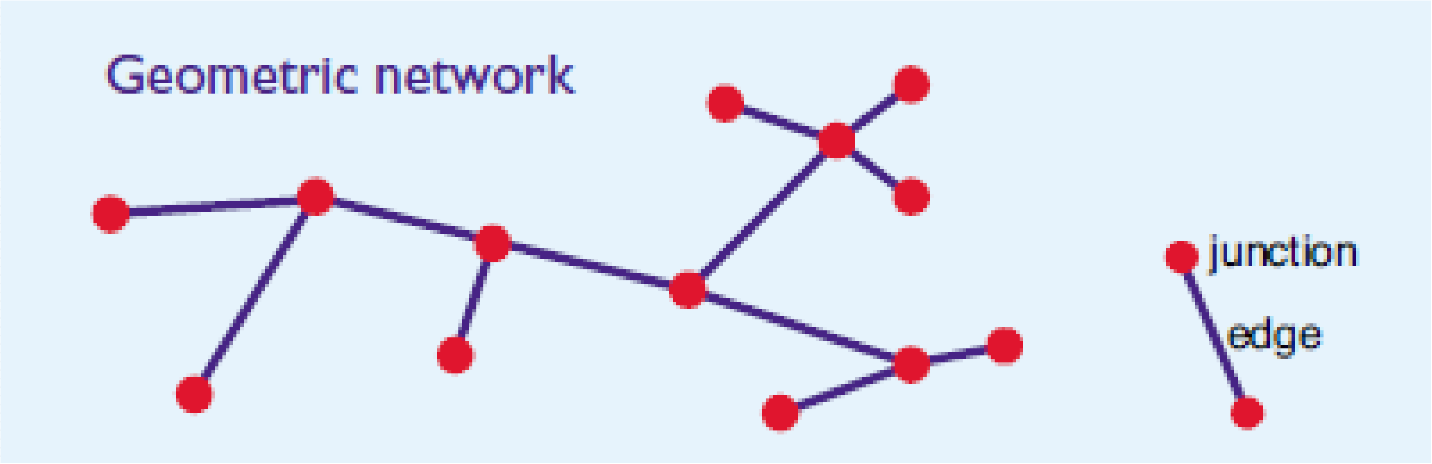


Figure 3.13 Sample digital line graph coding format.

DLGs

NETWORK

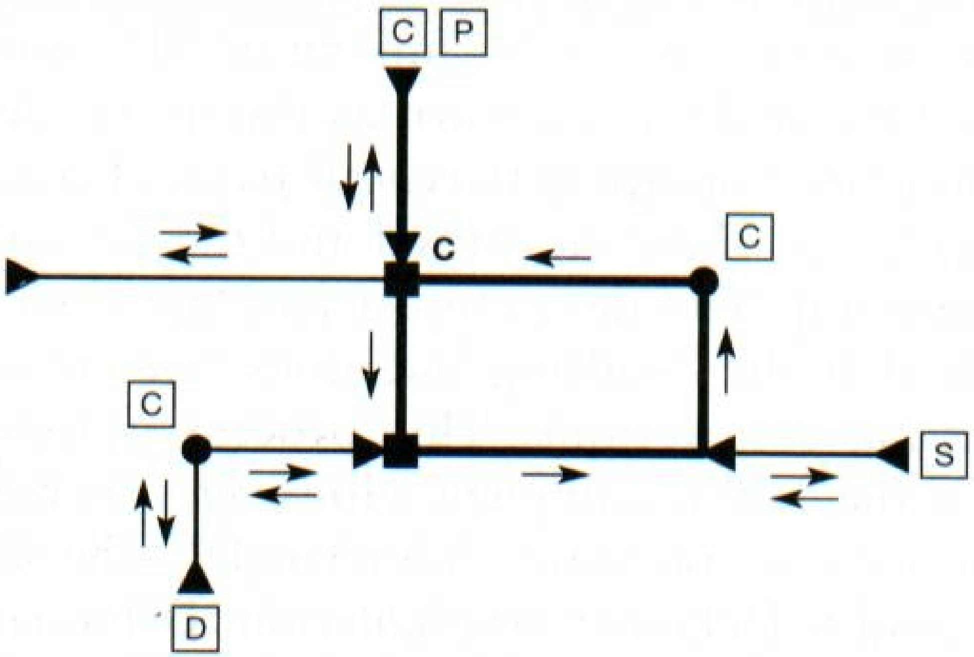
A network contains *edges that have nodes at their* endpoints. A node can be connected to one or many edges. This assemblage of edges and nodes is called a *geometric network.*

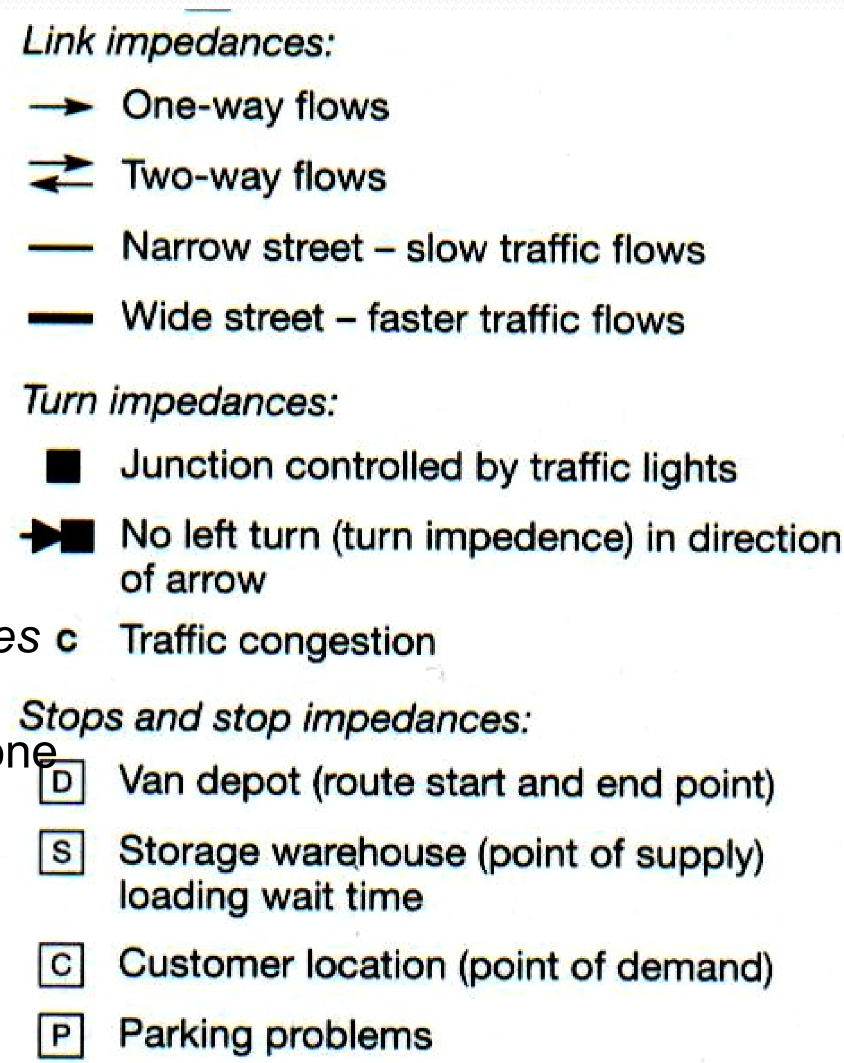


For transportation analysis, stream networks, switches/valves in utility networks, rail/air routes.

NETWORK

For transportation analysis: stream networks, switches/valves in utility networks, rail/air routes.





A network contains *edges that have nodes at their* endpoints. A node can be connected to one or many edges. This assemblage of edges and nodes is called a *geometric network.*

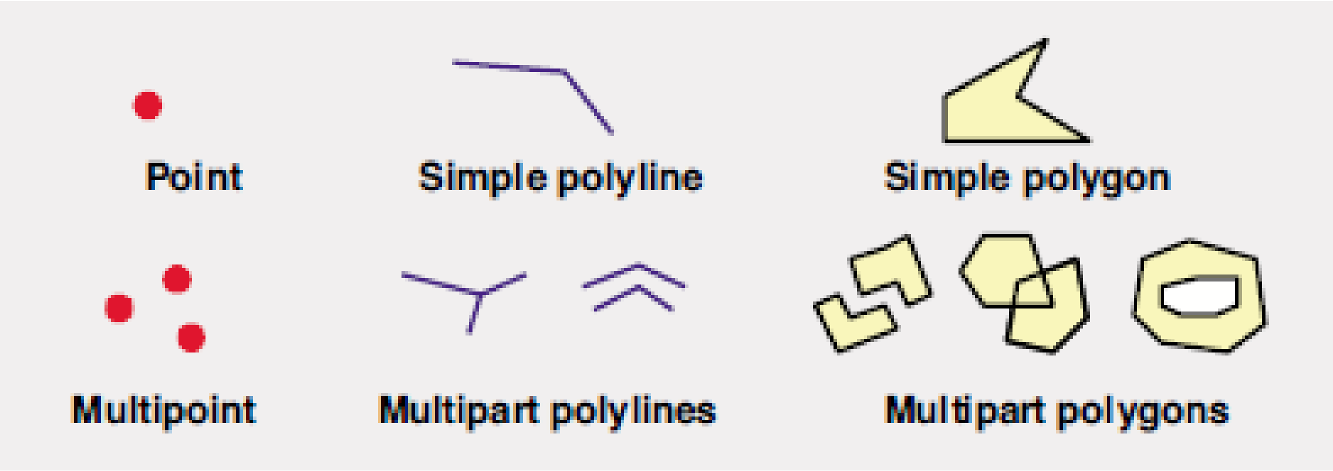
Shared edges in a topology

SHAPEFILE

Value	Shape type
0	Null shape
1	Point
3	PolyLine
5	Polygon
8	MultiPoint
11	PointZ
13	PolyLineZ
15	PolygonZ
18	MultiPointZ
21	PointM
23	PolyLineM
25	PolygonM
28	MultiPointM
31	MultiPatch

- **Non-topological vector model**
- **Stores geometry & attribute information for geographic features in a data set**
- **Geometry for a feature is stored as a *shape* comprising a set of vector coordinates linked to their attributes;**

there are 14 shape types (see slide) each describing particular entities or entity combinations



Spatial data is stored in binary files. Attribute data is stored in dBASE tables.

Shapefiles usually comprise 3 separate & distinct types of files:
main files, index files, and database tables.

- **Main file (e.g., counties.shp) is a direct access, variable record length file that contains the shape as a list of vertices.**
- **Index file (e.g., counties.shx) contains character length & offset (spaces) information for locating the values**
- **Database table/dBase (e.g., counties.dbf) that contains the attributes that describe the shapes.**
- **Shape file generally requires less processing power than topological counterpart.**

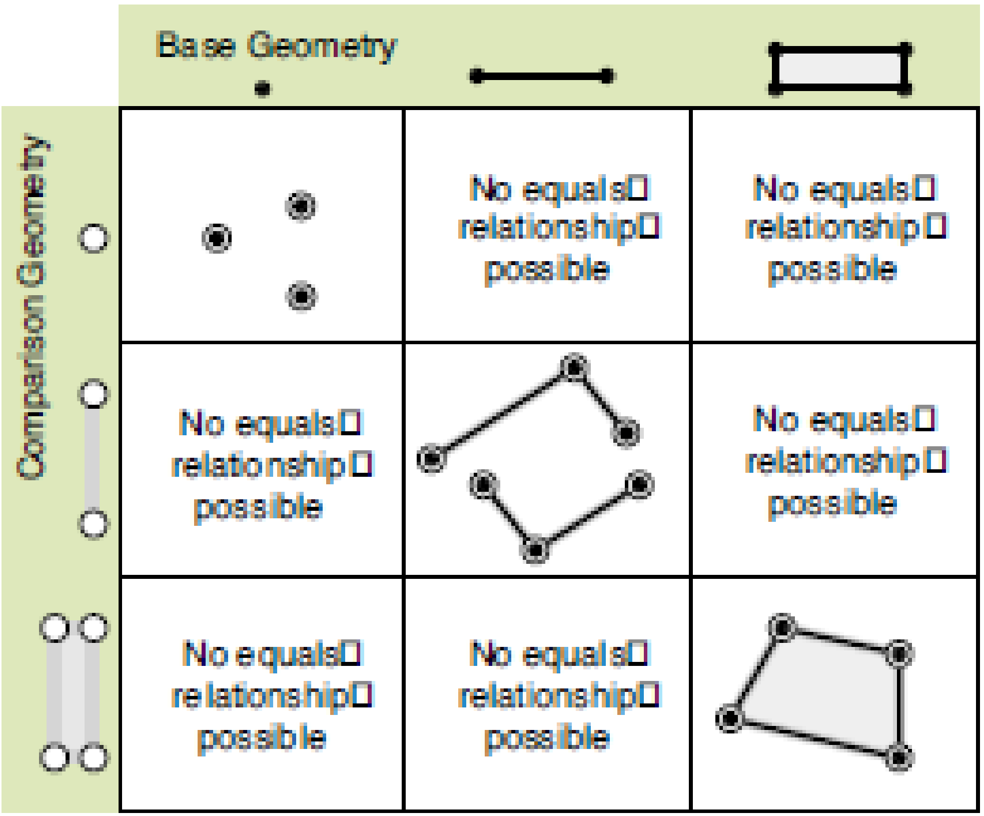
SPATIAL RELATIONSHIPS

SPATIAL RELATIONSHIPS

The GIS system includes a set of Boolean operators that test the spatial relationships between a base geometry and a comparison geometry. These operators can be applied to points, multipoints, polylines, and polygons

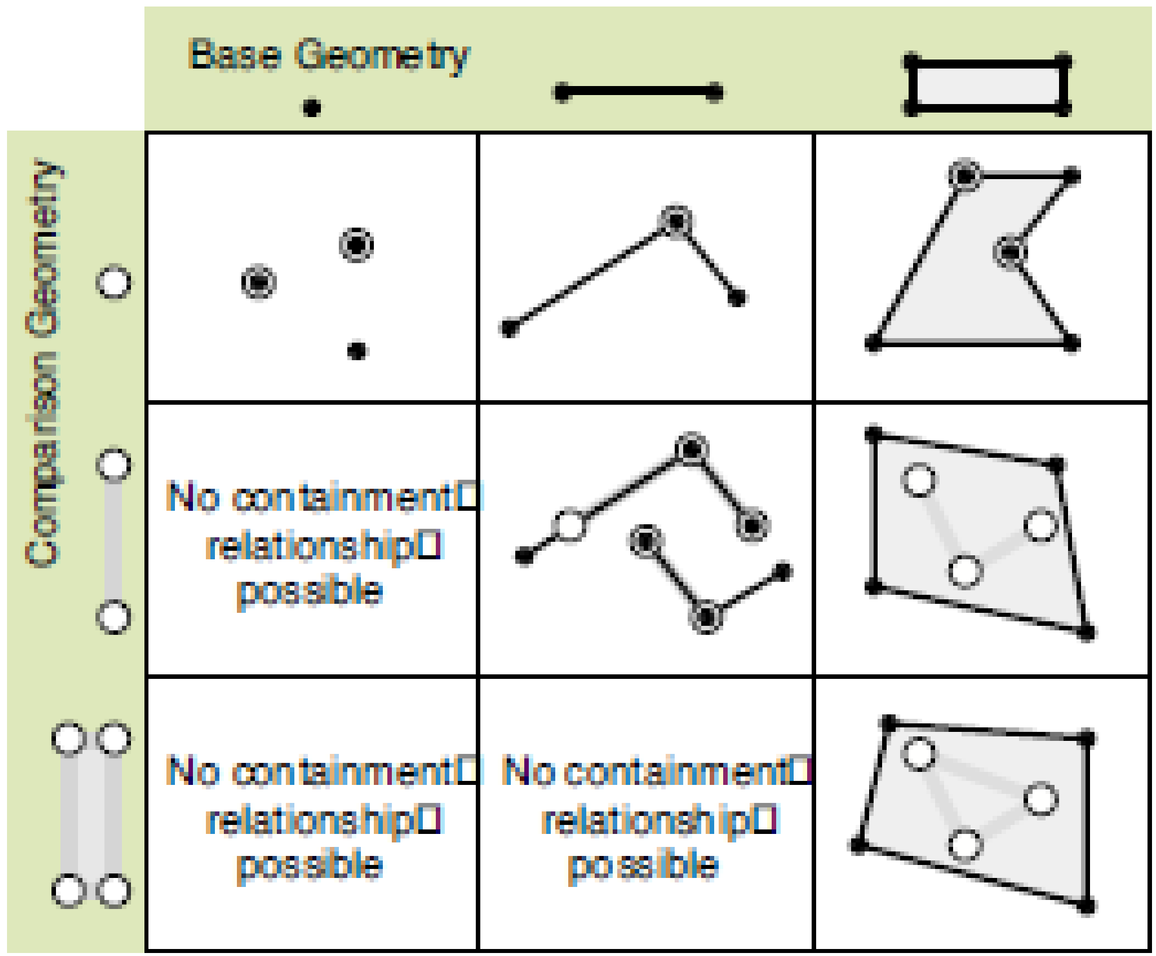
Equals

Does the base equal the comparison geometry?



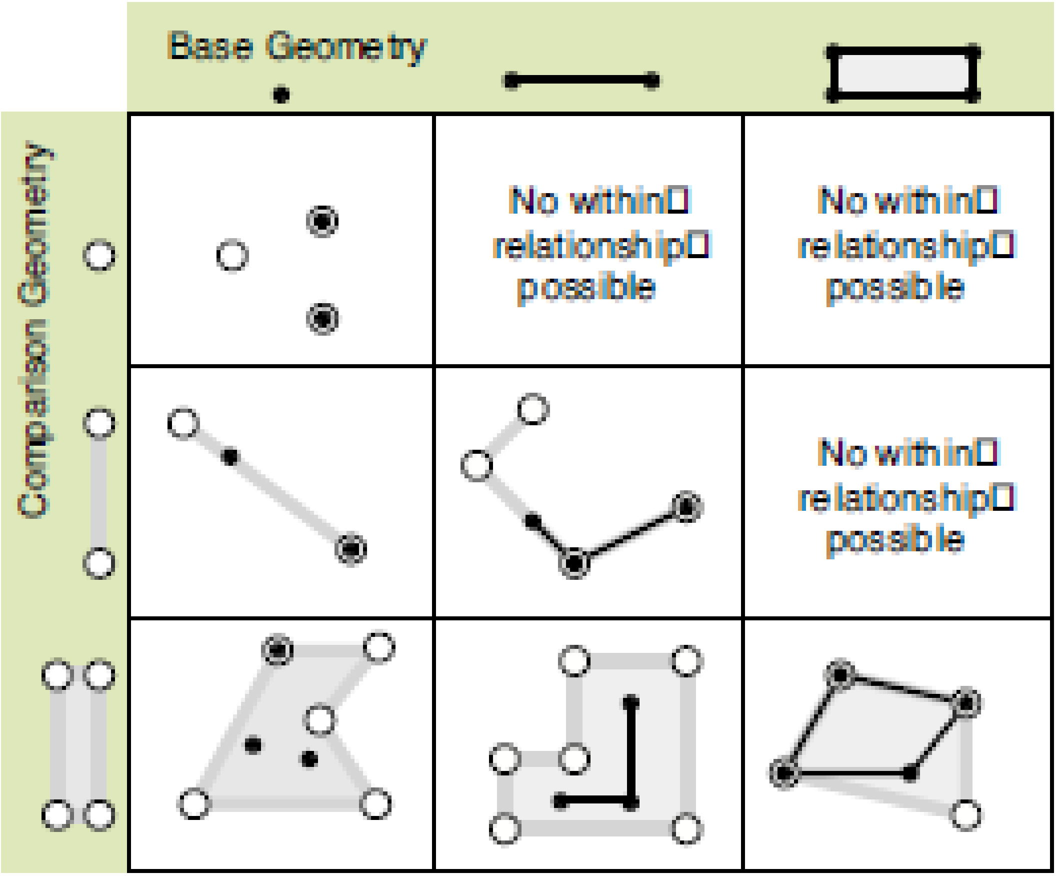
Contains

Does the base geometry contain the comparison geometry?



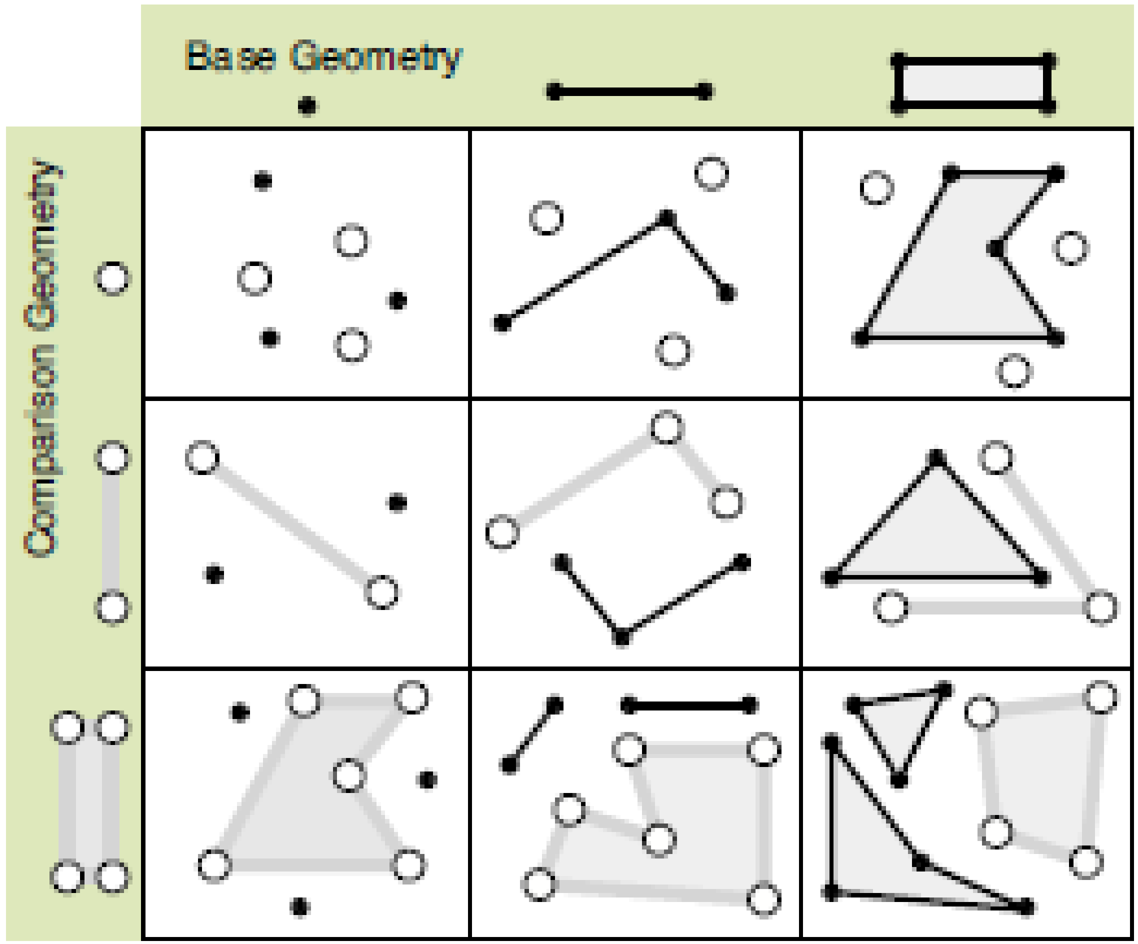
Within

Is the base within the comparison geometry?

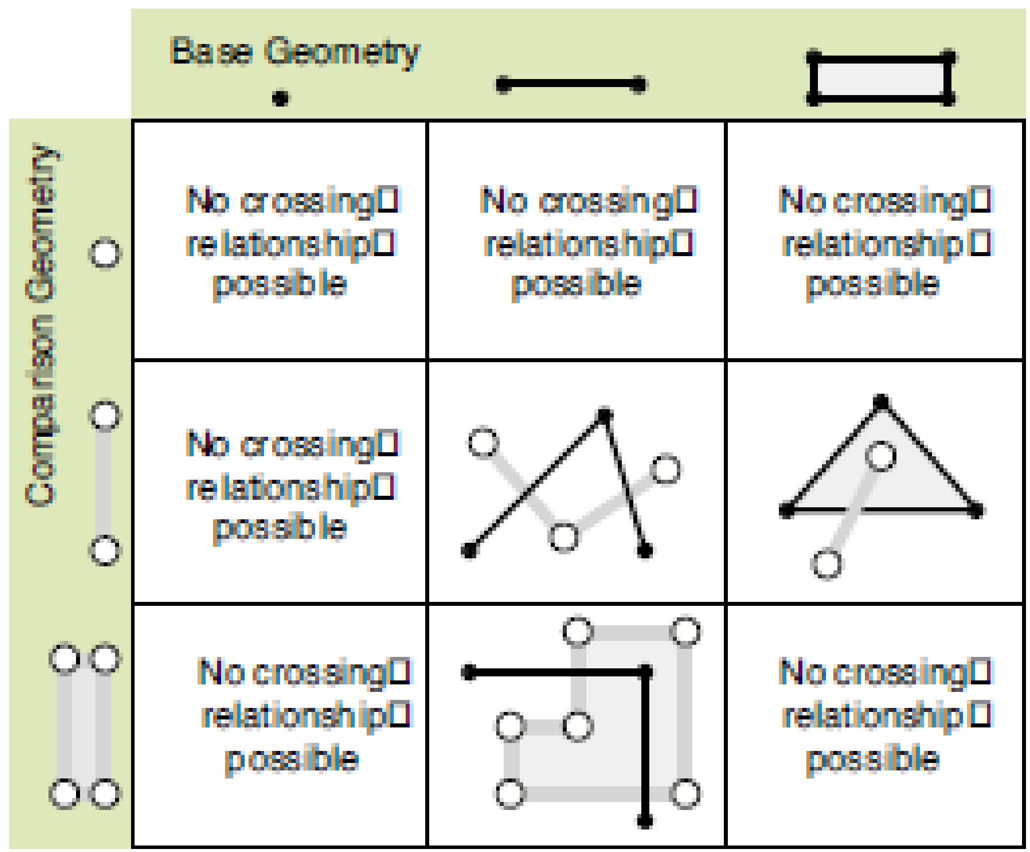


Disjoint

Is the base geometry disjoint from the comparison geometry?

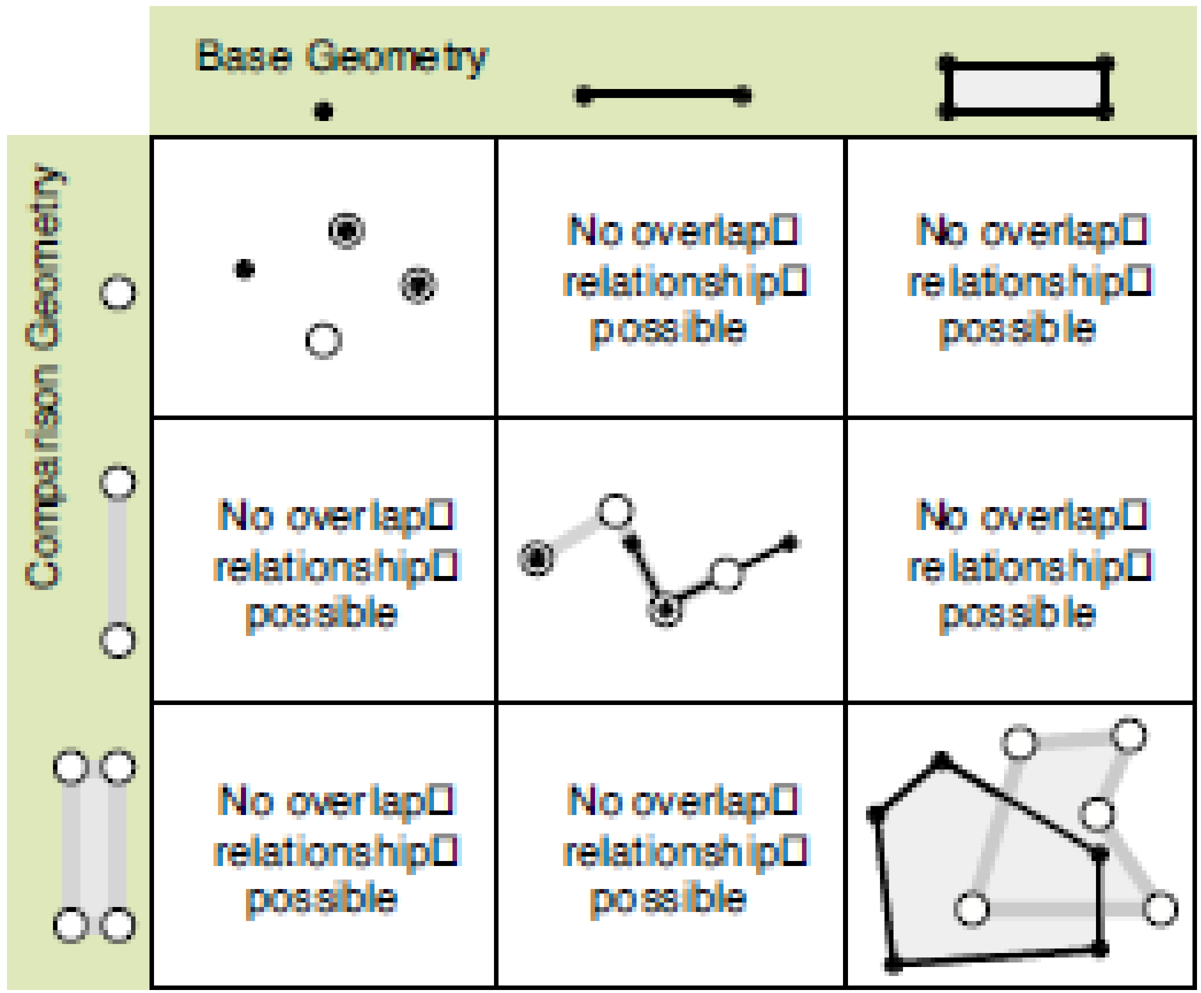


Crosses
Does the base geometry cross the comparison geometry?



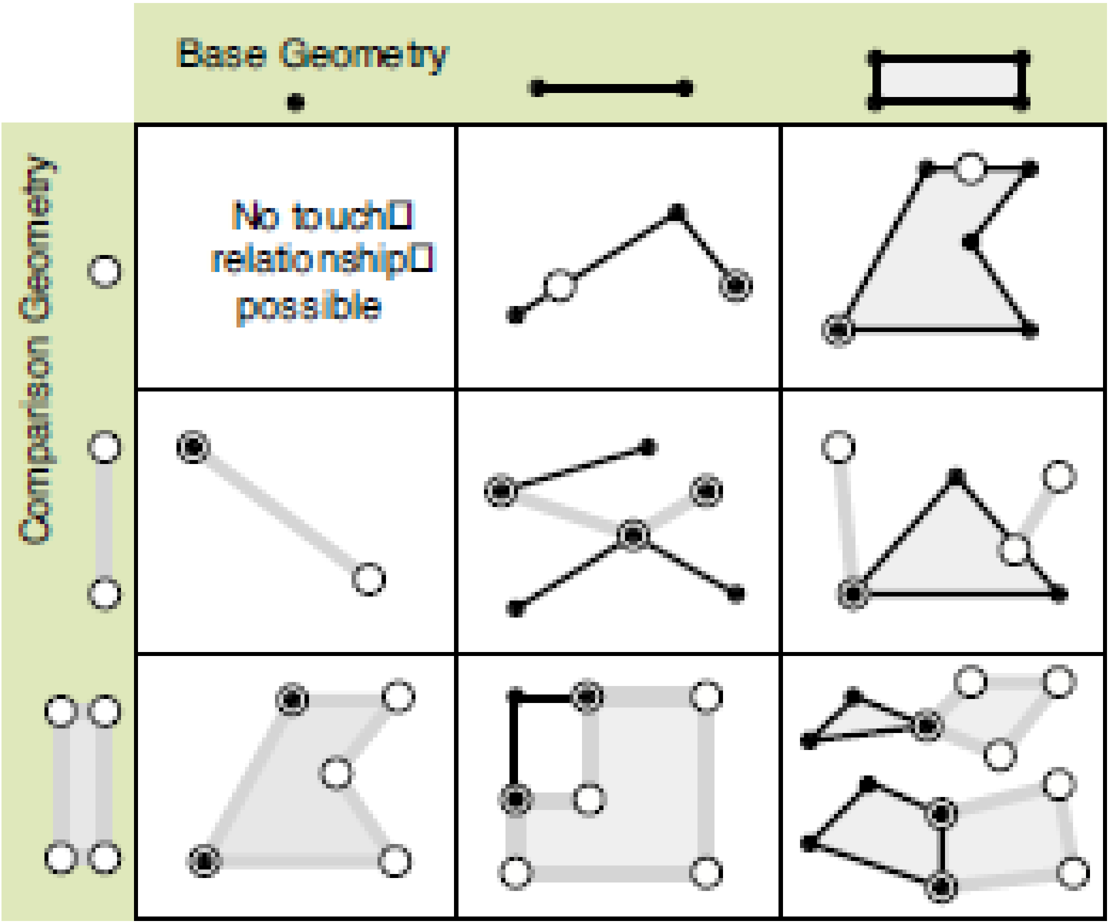
Overlaps

Does the base geometry overlap the comparison geometry?



Touches

Does the base geometry touch the comparison geometry?



Locational Data

The capture of locational data for entities (e.g. by digitising from a paper map) can result in numerous errors. ESRI
suggests a useful checklist of objectives when capturing data in vector mode:

1. All entities that should have been entered are present.
2. No extra entities have been digitised.
3. The entities are in the right place and are of the correct shape and size.
4. All entities that are supposed to be connected to each other are.
5. All polygons have only a single label point to identify them.
6. All entities are within the outside boundary identified with registration marks.

- This provides a good indication of the types of problem that might arise. Entities (i.e. points, lines, polygons) may simply be overlooked when digitising, or may be entered more than once. An arc missing between two nodes may result in two polygons being captured as a single polygon. An arc inadvertently digitised twice may result in a **sliver line**. Vertices inaccurately digitised may result in lines having the wrong shape or, if the vertex in question is a node, may result in a **dangling node**. The dangling node may either **undershoot** its correct location, resulting in a gap, or it may **overshoot** its intended location, resulting in a cul-de-sac (and an intersection not identified as a node). Vertices digitised in the wrong sequence may result in **weird polygons** or a **polygonal knot**. If digitising polygons then it is obviously important to have the correct number of labels points in the correct locations. Too few label points will result in some polygons not having associated attribute data, whilst too many label points may result in a polygon having the wrong attribute data associated with it.

- Digitising errors do not necessarily indicate a lack of accuracy when digitising points. They may also arise if the snapping tolerance is incorrectly set. For example, dangling nodes frequently arise if the snapping tolerance is set too low. However, if the snapping tolerance is set too high, nodes may be snapped to the wrong points. Apart from causing lines to have the wrong shape, this could result in topological inconsistencies.

- If the data are topologically encoded, then the digitising software can run a number of checks to identify potential problems. For example, the software can check how many line segments enter each node. If only one line segment enters a node then it can be identified as a dangling node. If two line segments enter a node then it is referred to as a **pseudo node**. Both situation can be flagged as potential problems. However, dangling nodes may reflect genuine cul-de-sacs in a road system, or the sources of tributaries in a river system; whilst a pseudo node may identify a polygon completely enclosed within another polygon (e.g. a lake) or a change in attribute along a line (e.g. single lane road to dual carriageway). The first situation is sometimes referred to as a **spatial pseudo node** and the second as an **attribute pseudo node**.

Data Processing Errors

- Further errors may be introduced during data processing. For example, if converting data from raster to vector mode, vector mode lines which should be straight may take on a stepped appearance. There are various smoothing algorithms which may be used to smooth out angular lines, but there is no way of knowing whether the smoothed lines are actually any more accurate - the net effect of smoothing the lines may be to introduce further errors by making them artificially smooth. Vector to raster conversions may result in topological errors being introduced or even in the creation or loss of small polygons. Raster coverages created from the same vector coverage will tend to vary depending upon relatively arbitrary decisions about cell size, the orientation of the raster and the location of the origin.

- Interpolation of data values in a continuous field from sample points will result in different values depending upon the choice of method of interpolation and other decisions made with regard to the parameters used. The number of sample points will also have a fairly obvious influence upon the reliability of the resulting estimates. When analysing field data it is therefore necessary to bear in mind that the estimated data values are not necessarily accurate.

- It is important to realise that computers may introduce errors when processing data due to limitations placed upon the precision of numbers arising from the way in which they are stored in a computer. When working with numbers requiring a large number of significant digits, calculations done by computers may result in a high degree of inaccuracy. This problem is becoming less serious with the availability of 32-bit and 64-bit machines, provided that the software has been programmed to take advantage of the extra precision. If you are working with high precision numbers you should confirm that both the computer and the software can support the degree of precision required because the implications may be more serious than simply rounding numbers to a smaller number of significant digits.

- Finally, **use errors** may arise from simply using inappropriate tools for a particular type of analysis.

Data Display Errors

- The display of data may also introduce errors. For example, the display of raster data on a vector mode device (e.g. a plotter) or the display of vector data on a raster device (e.g. a monitor or a printer) will generally introduce other small inaccuracies due to the need to round off during the conversion from one mode to the other. These errors can probably be ignored for practical purposes, but they serve as a reminder that errors can creep in at all stages in a GIS analysis.

MODELLING DATA ERRORS

- Apart from recognising that errors are likely and then taking whatever steps one can to minimise them, The treatment of errors in GIS has received relatively little attention, especially in commercial applications software, but there have been some tentative steps towards using quantitative measures of errors to provide some indication of the reliability of data in a GIS.

Positional Error Models

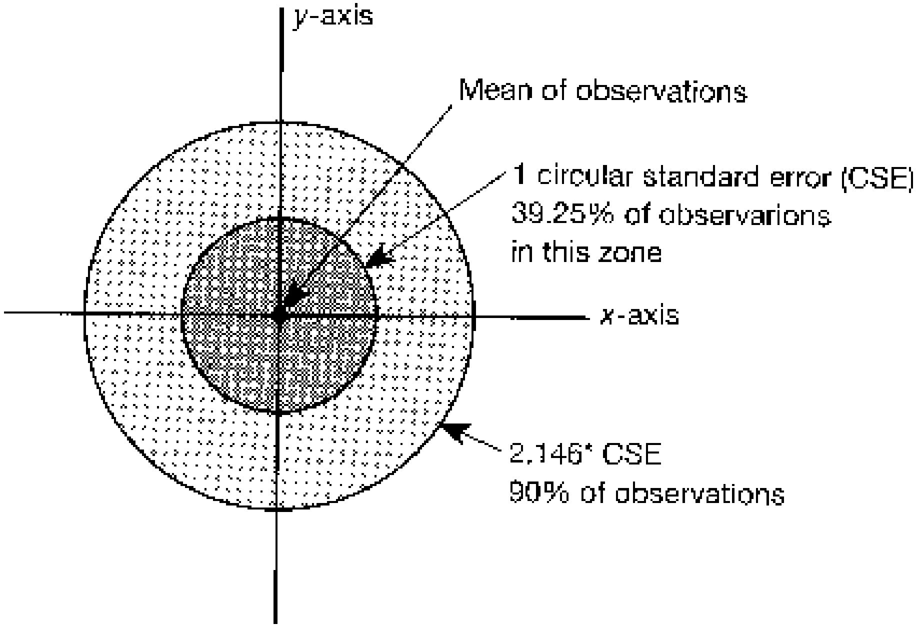
- Positional error models represent an attempt to place confidence bands around locational features. It is assumed that if the x co-ordinate of a point (or vertex) was measured repeatedly then the observed x co-ordinates would have a Normal (i.e. Guassian) distribution with an expected value (or mean) corresponding to the true value. 68 per cent of the observed co-ordinates would be within one standard error of mean, and 90 per cent would be within 1.65 standard errors. Having established the standard error for one point by experiment, the expected error associated with other points could be expressed using probabilistic confidence bands.

- There are a number of assumptions implicit in the choice of a Normal distribution to model measurement errors (e.g. it is assumed that the errors are unbiased; it is assumed that the probabilities associated with errors of differing magnitude form a continuum, etc.). If these assumptions are unrealistic, then other statistical distributions may be preferable. However, the basic approach is much the same.

Point Data

- When recording the spatial location of a point we record two co-ordinates (i.e. x and y). If the errors associated with each are assumed to be normally distributed, then the probability distribution will be a bell-shaped surface, declining at the same rate in all directions.[2] The standard error of this surface is called the **circular standard error** (CSE).

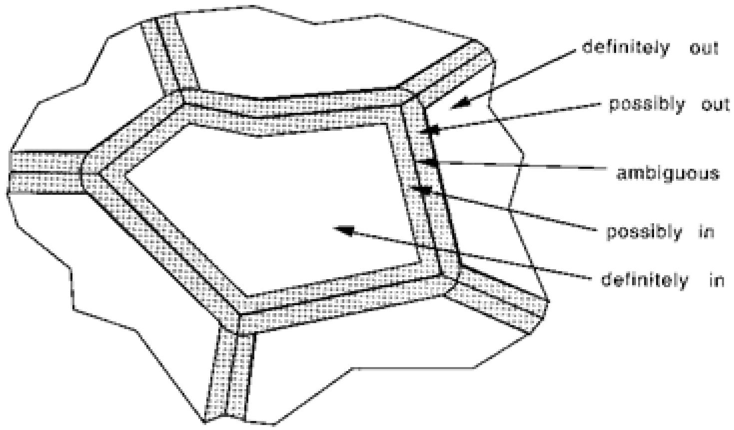
- 39.35 per cent of all points can be expected to lie within a circle with a radius of 1 CSE centred on the mean. 90 per cent of the points should be within 2.146 CSE. One way to define the accuracy of a map is to specify a circular map accuracy standard (e.g. 2.146 CSE, meaning that 90 per cent of all the observed data points will be within this distance of their locations).



Line Data

- The true location of each point on a line can be thought of as lying within a band on either side of a digitised line, where the width of the band reflects the standard error. These bands are sometimes referred to as **epsilon bands**. The original epsilon band was hypothesised as rectangular in cross-section, but a Normal distribution is more frequently assumed. However, given that each point in a line is not independent of the points which precede and follow it, it seems plausible that the digitised points forming a sequence will tend to be displaced in the same direction (i.e. there will be a bias, and the true distribution of the errors will be skewed). Some investigators have suggested the cross- sectional probability distribution should be bimodal.

__Polygon Data:__



Similar principles apply to polygons. Information on the width of the epsilon bands can be used to place confidence limits on point in polygon tests. As can be seen from the diagram, instead of points being classified as either inside

- A **Monte Carlo** approach can be used to model errors. This basically involves adding a random noise factor (which can be positive or negative) to each co-ordinate for each point before performing whatever GIS operation you need to do. The results are saved, and the whole process is repeated a large number of times (e.g. 100 times). The accumulated results can be used to calculate confidence limits for numerical answers or to draw confidence bands around features on output maps. The main problem with a Monte Carlo approach is that it requires a lot of computer resources.

- Burrough and McDonnell (Chapter 10) discuss a similar approach for evaluating the effects of measurement errors in numerical models. They also discuss the statistical theory of **error propagation**.[3] Whilst mathematically more challenging, this provides a computationally more efficient means of achieving the same objectives. The main conclusion from their review of several case studies is that even relatively small measurement errors can have a much greater impact than one might imagine. There is therefore an obvious need to develop better methods for assessing data quality and its implications.

METADATA

- The reliability of a particular set of data is dependent upon the uses to which it is put. Data which are completely inappropriate in one context may be totally adequate in a different context (or vice versa).

- Data quality is therefore to some extent a relative concept dependent upon the context. The emphasis has therefore tended to switch away from simply trying to make the data as error free as possible to providing potential users with the information which they require to make an informed decision about the adequacy of the data for a particular purpose. This information is referred to as **metadata**.

Other DATA MODELS using ATTRIBUTES

A separate data model is used to store and maintain attribute data for GIS software.

These data models may exist internally within the GIS software, or may be reflected in external commercial Database Management Software (DBMS).

A variety of different data models exist for the storage and management of attribute data. The most common are:

- **Tabular**
- **Hierarchial**
- **Network**
- **Relational**
- **Object Oriented**

TABULAR MODEL

The simple tabular model stores attribute data as sequential data files with fixed formats (or comma delimited for ASCII data)

This type of data model is outdated in the GIS arena.

It lacks any method of checking data integrity, as well as being inefficient with respect to data storage, e.g. limited indexing capability for attributes or records, etc.

HIERARCHICAL MODEL

The hierarchical database organizes data in a *tree* structure.

Data is structured downward in a *hierarchy* of tables.

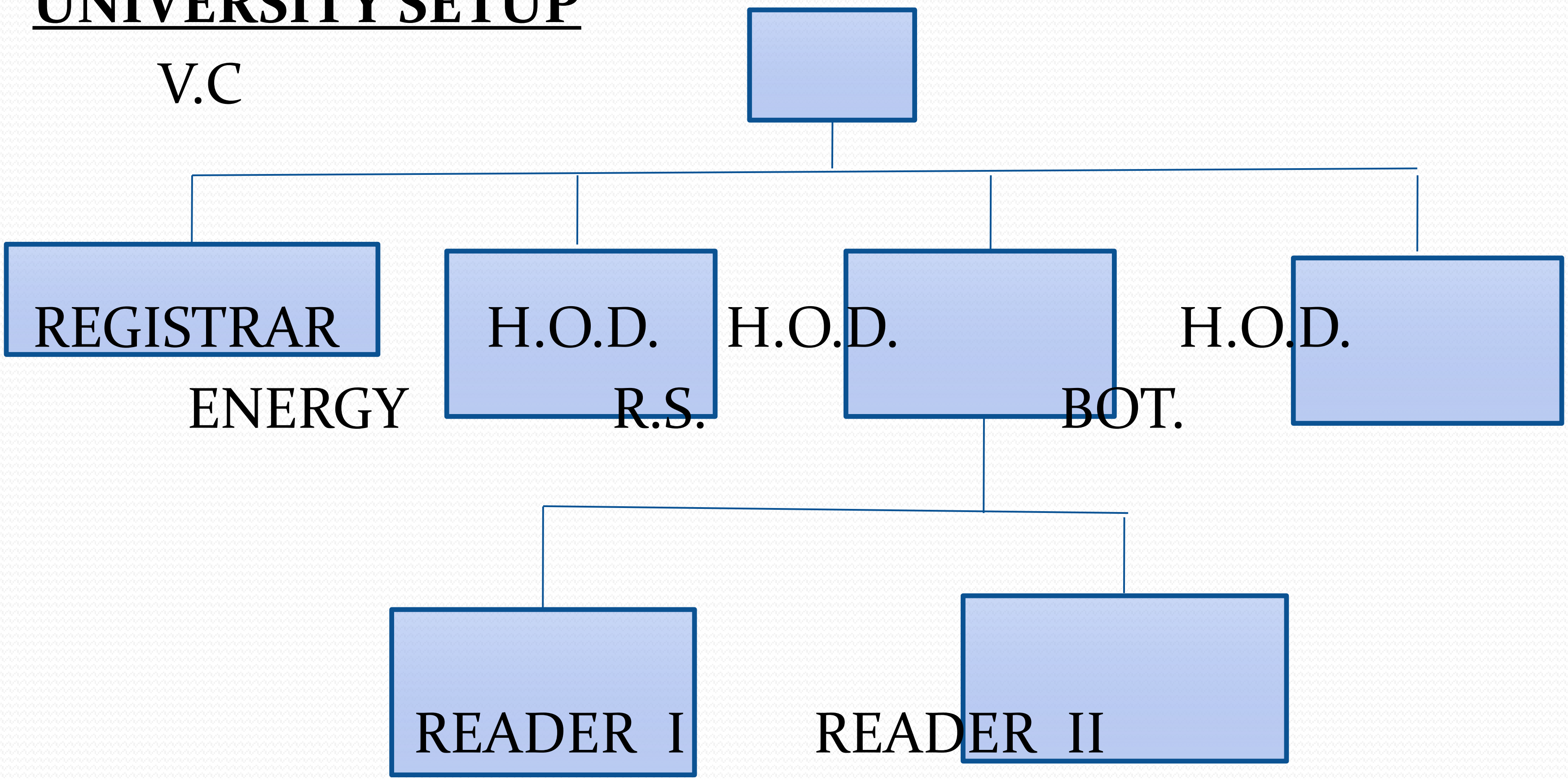
Any level in the hierarchy can have unlimited *children*, but any *child* can have only one *parent*.

Hierarchial DBMS are good for very stable datasets.

But GIS datasets are always dynamic

HIERARCHICAL DATASTRUCTURE

UNIVERSITY SETUP



Go Back

NETWORK MODEL

The network database organizes data in a network or *plex* structure.

Any column in a plex structure can be linked to any other.

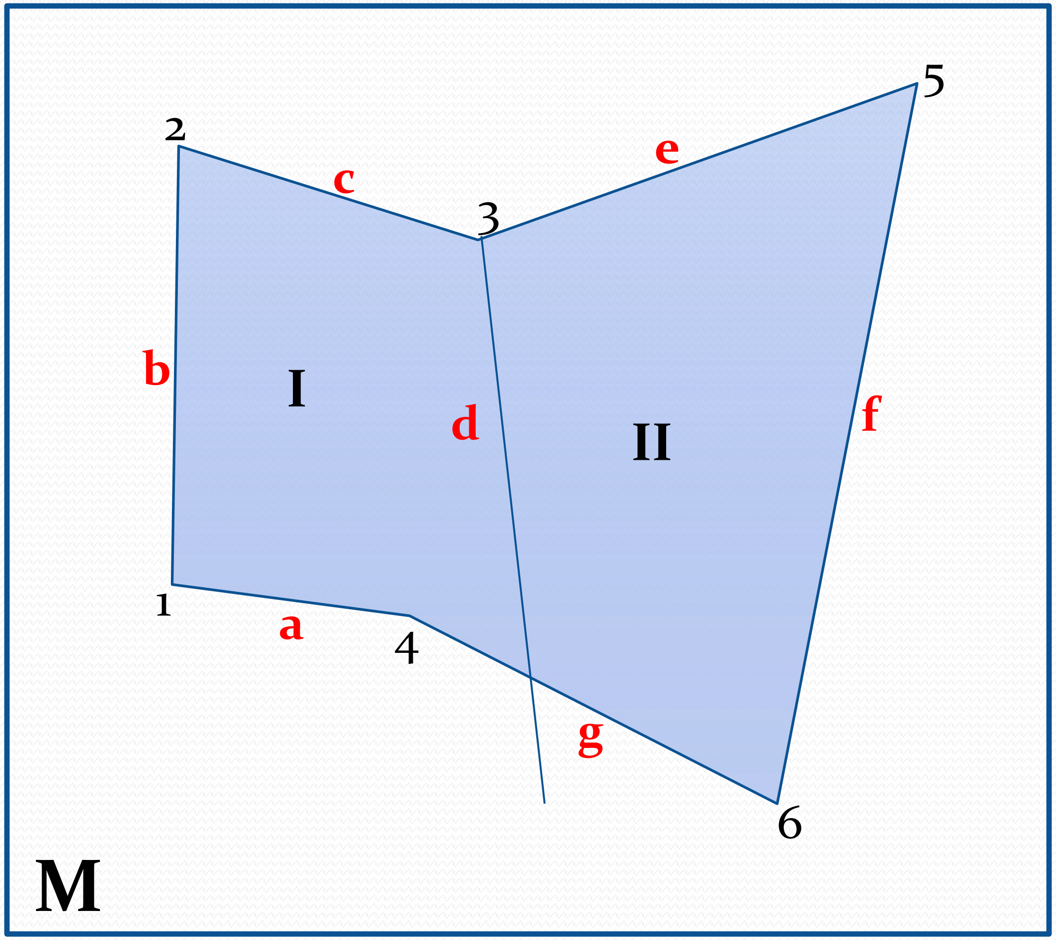
Like a tree structure, a plex structure can be described in terms of *parents* and *children*.

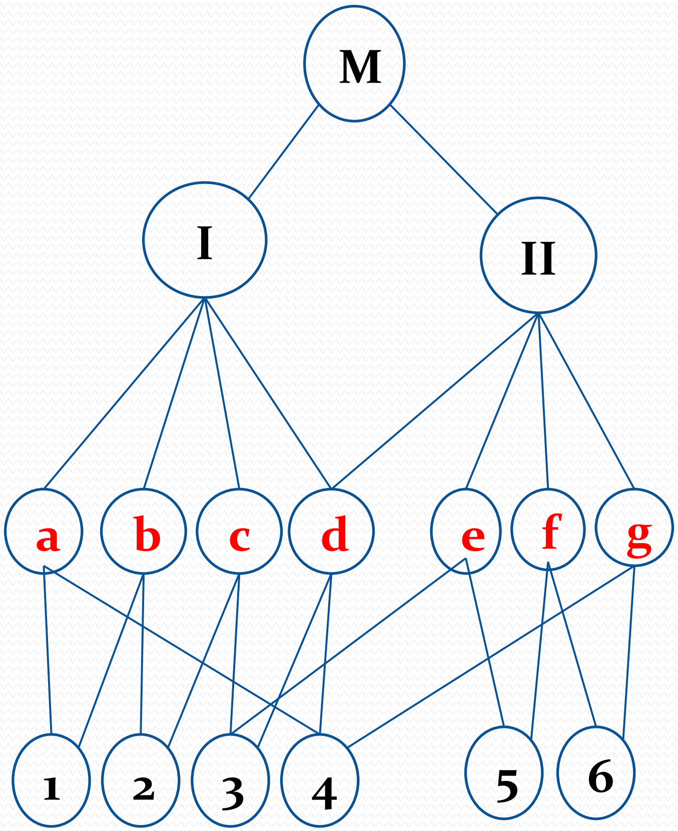
This model allows for children to have more than one parent.

Network DBMS have not found much more acceptance in GIS than the hierarchical DBMS.

They have the same flexibility limitations as hierarchical databases; however, the more powerful structure for representing data relationships allows a more realistic modelling of geographic phenomenon.

NETWORK DATA STRUCTURE





Go Back

Map M consists of 2 polygons, 7 lines (6+1 shared) and 6 nodes,

RELATIONAL MODEL

The relational database organizes data in *tables*.

Each table, is identified by a unique table name, and is organized by *rows* and *columns*.

Each column within a table also has a unique name.

Columns store the values for a specific attribute, e.g. cover group, tree height. Rows represent one record in the table. In a GIS each row is usually linked to a separate spatial feature, e.g. a forestry stand.

Accordingly, each row would be comprised of several columns, each column containing a specific value for that geographic feature.

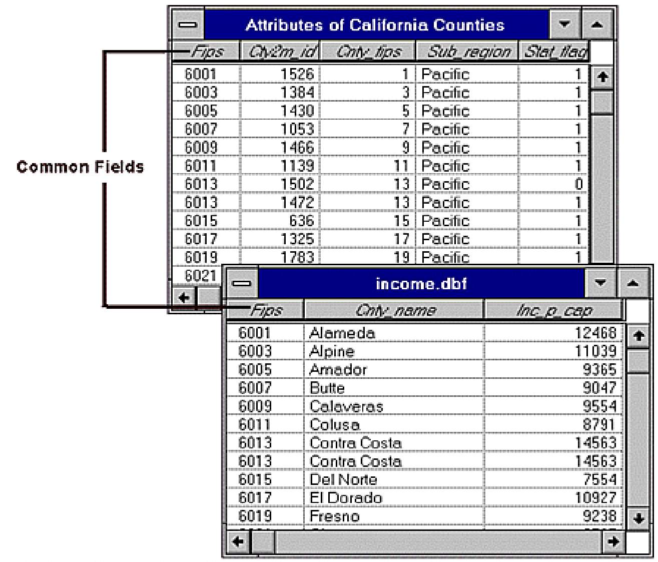
The following figure presents a sample table for forest inventory features.

This table has 4 rows and 5 columns.

The forest stand number would be the *label* for the spatial feature as well as the *primary key* for the database table.

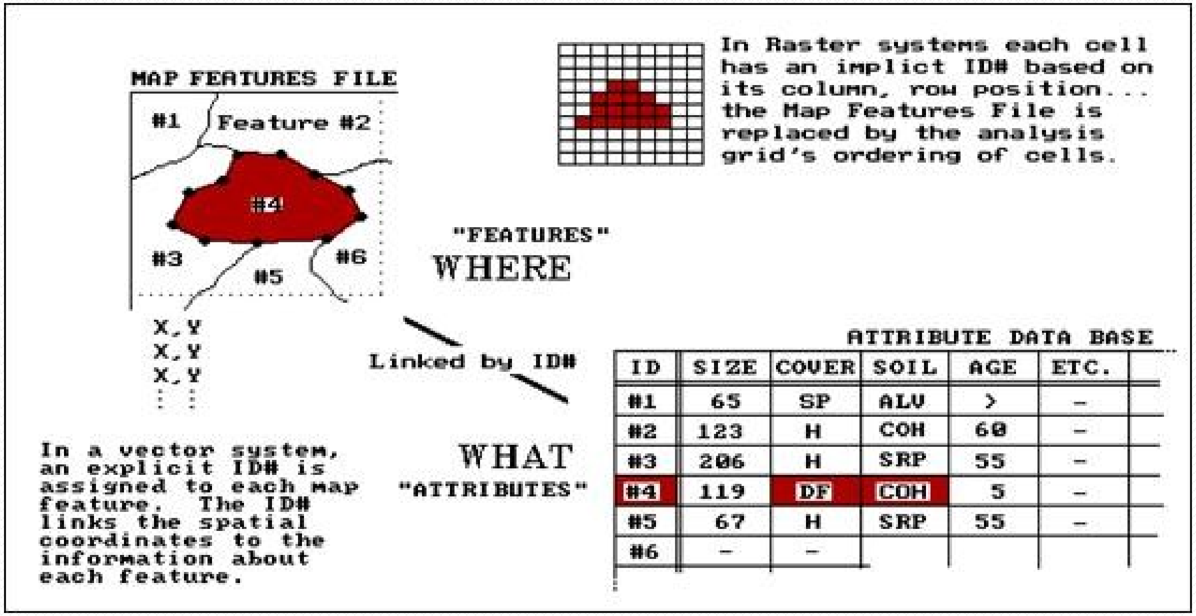
This serves as the linkage between the spatial definition of the feature and the attribute data for the feature.

UNIQUE STAND NUMBER	DOMINANT COVER GROUP	AVG. TREE HEIGHT	STAND SITE INDEX	STAND AGE
001	DEC	3	G	100
002	DEC-CON	4	M	80
003	DEC-CON	4	M	60
004	CON	4	G	120



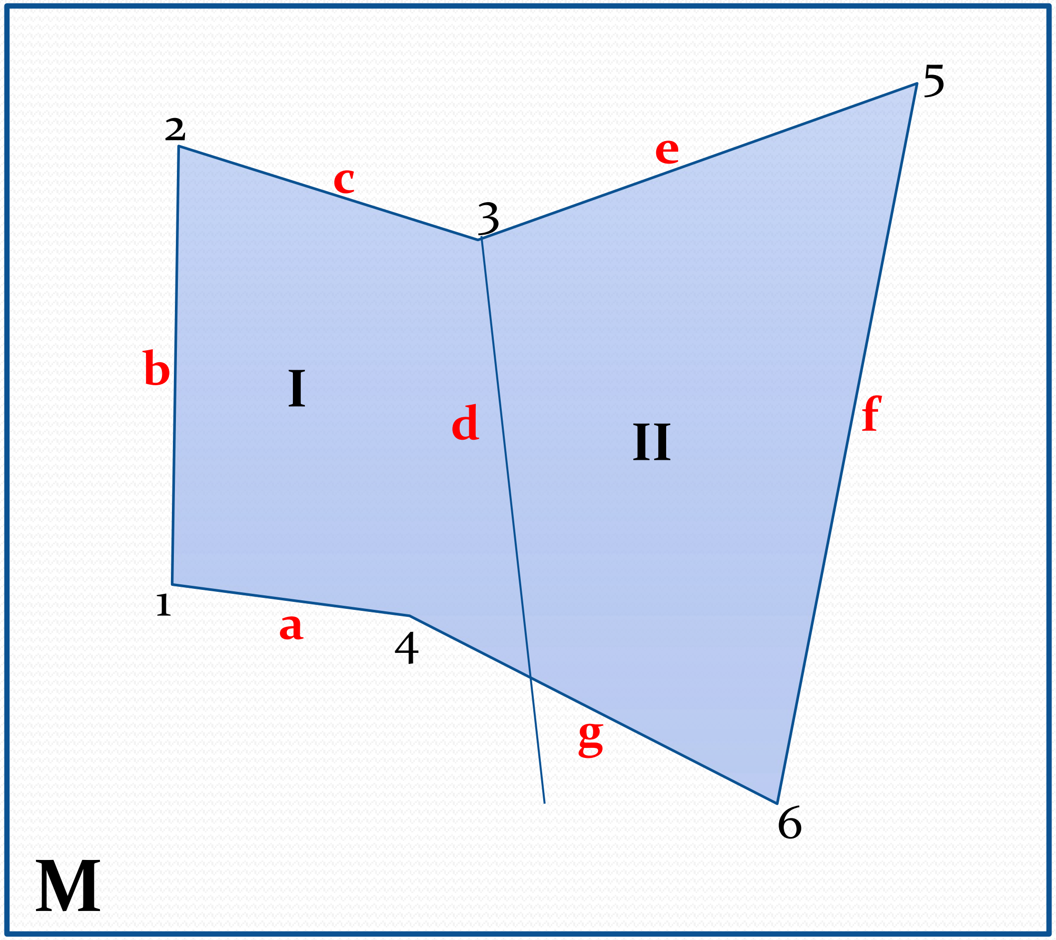
Fips	Cty2m_id	Cnty_fips	Sub_region	Stat_flag
6001	1526	1	Pacific	1
6003	1384	3	Pacific	1
6005	1430	5	Pacific	1
6007	1053	7	Pacific	1
6009	1466	9	Pacific	1
6011	1139	11	Pacific	1
6013	1502	13	Pacific	0
6013	1472	13	Pacific	1
6015	636	15	Pacific	1
6017	1325	17	Pacific	1
6019	1783	19	Pacific	1
6021				

Fips	Cnty_name	Inc_p_cap
6001	Alameda	12468
6003	Alpine	11039
6005	Amador	9365
6007	Butte	9047
6009	Calaveras	9554
6011	Colusa	8791
6013	Contra Costa	14563
6013	Contra Costa	14563
6015	Del Norte	7554
6017	El Dorado	10927
6019	Fresno	9238



ID	SIZE	COVER	SOIL	AGE	ETC.
#1	65	SP	ALV	>	–
#2	123	H	COH	60	–
#3	206	H	SRP	55	–
#4	119	DF	COH	5	–
#5	67	H	SRP	55	–
#6	–	–			

RELATIONAL DATA STRUCTURE



Map:

M	I	II

Polygons:

I	a	b	c	d
II	d	e	f	g

Lines:

I	a	4	1
I	b	1	2
I	c	2	3
I	d	3	4
II	e	3	5
II	f	5	6
II	g	6	4
II	d	3	4

Co-ordinates:

Go Back

OBJECT-ORIENTED MODEL

The object oriented data model brings a physical data model closer to its logical data model.

The data objects in a object oriented are mostly the same objects you would define in a logical data model, such as owners, buildings, parcels, and roads.

Data model lets you implement the majority of custom behaviors without writing any code.

Most behaviors are implemented through domains, validation rules, and other functions

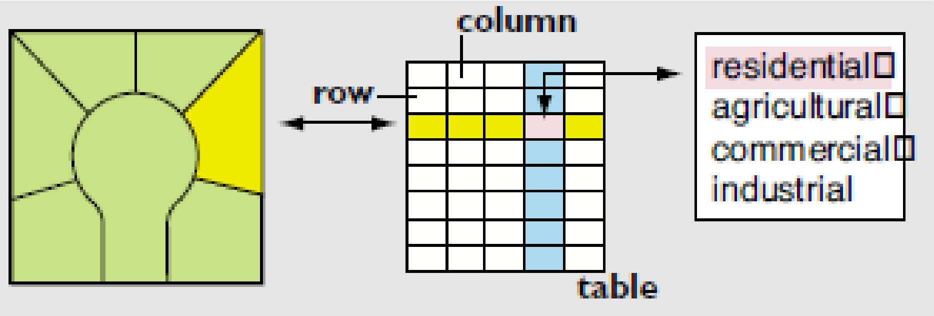
Adding and editing features

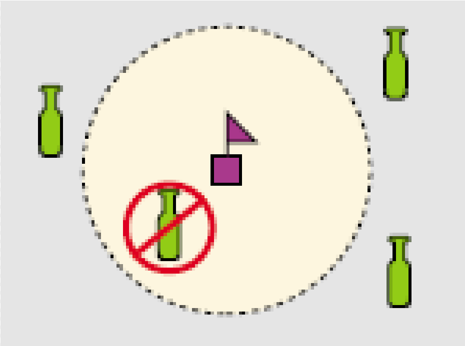
That the values you assign to an attribute fall within a prescribed set of permissible values.

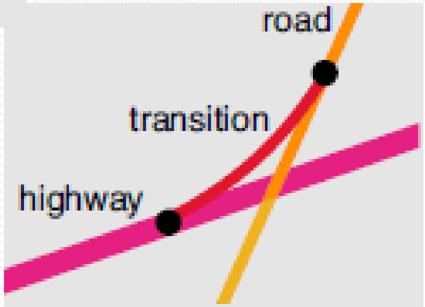
A parcel of land may only have certain land uses such as residential, agricultural, or industrial.

Placing a liquor store near a school is not permitted by law.

A city road cannot be connected to a highway without a transition segment such as an on-ramp.



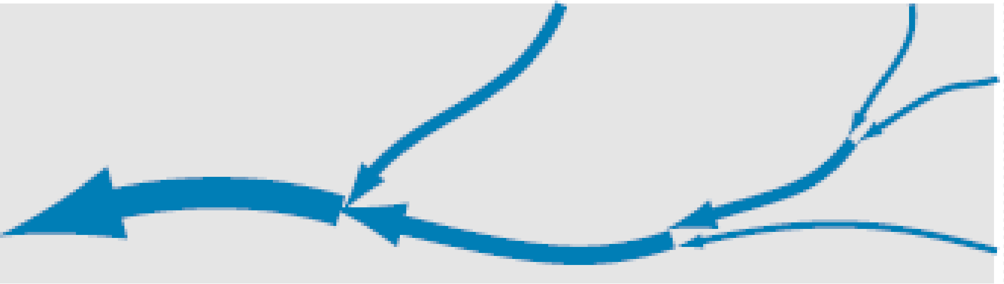


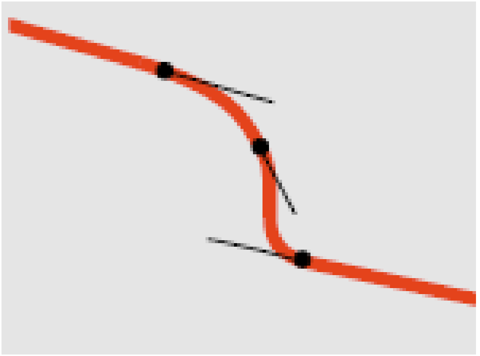


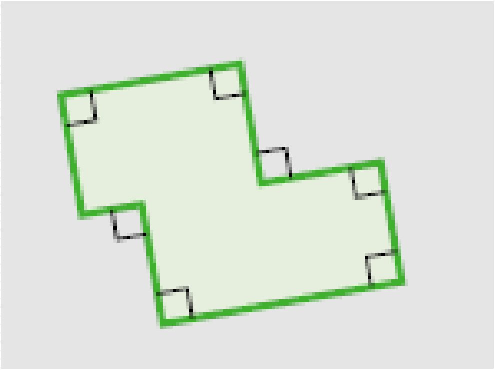
A stream system should always flow downhill. Flow down from a junction is the sum of flows upstream.

The lines and curves that make up a road should be tangent.

Building corners most often form right angles







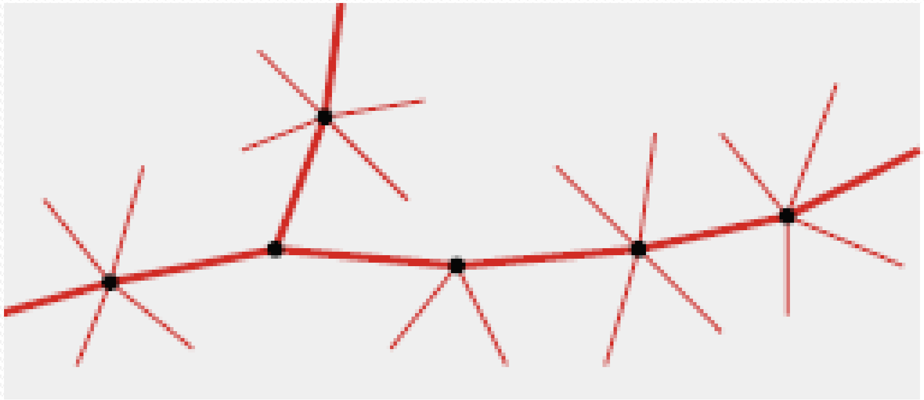
Relationships among features

All objects in the world are entangled in relationships with other objects.

The relationships fall within three general categories: topological, spatial, and general.

Topological relationships is defined when you load or edit features within a connected system

In an electric utility system, you want to be sure that the ends of primary and secondary lines connect exactly and that you are able to perform tracing analysis on that electric network

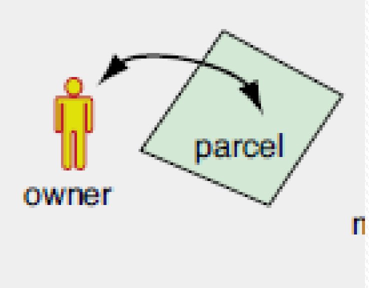


Spatial relationships are inferred from the geometry of features such as whether a feature is inside, touching, outside, or overlapping another feature.

want to determine which block contains a particular building

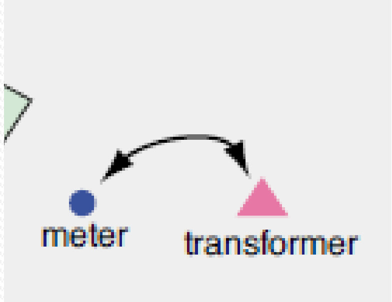
General relationship : Some objects have relationships that are not present on a map.

A parcel has a relationship to an owner, but the owner is not a feature on a map.



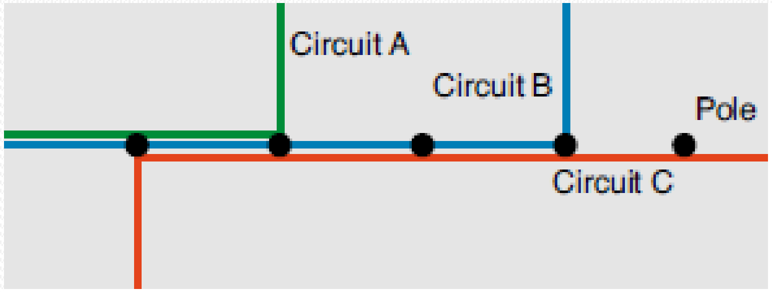
Some features on a map have relationships, but their spatial relationship is ambiguous.

A utility meter is in the general vicinity of an electric transformer, but it is not touching the transformer. The meter and the



Cartographic display

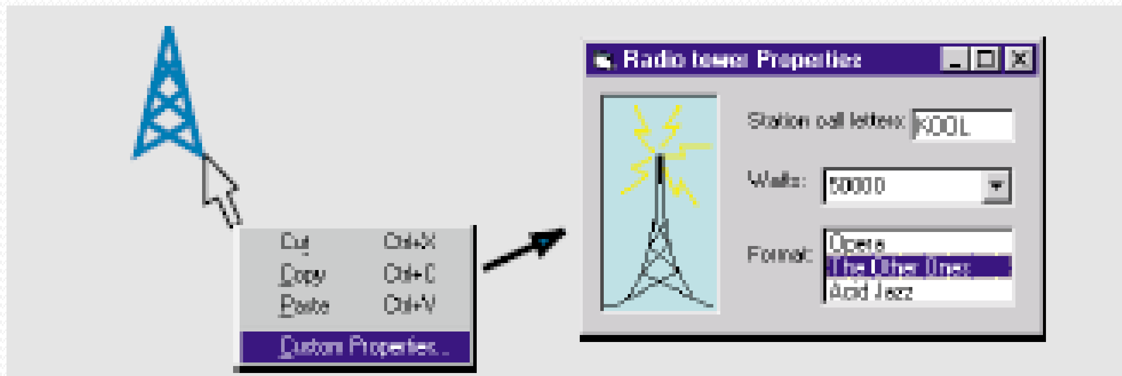
When multiple electrical wires are physically mounted on the same set of utility poles, you would like to depict them as spread in a set of parallel lines with a standard offset in map units



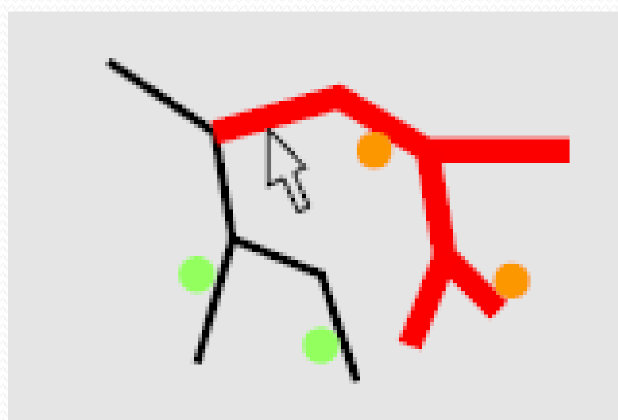
Interactive analysis

Touch a feature on a map display and invoke a form to query and update its properties.

`



Select a part of an electric network where line maintenance is planned, find all affected downstream customers, and make a mailing list to notify them.



Defining of COORDINATE SYSTEMS

To locate **any point, feature** or **area** on the **Earth's surface** or a **map** it is necessary to have concepts and definitions of direction and distance

Three main categories of referencing system:

Geographic/ Spherical (applies globally)

e.g. lines of latitude and longitude

Rectangular (applies locally)

UTM (Universal Transverse Mercator)

UPS

State Plane

e.g. Ordnance Survey National Grid system

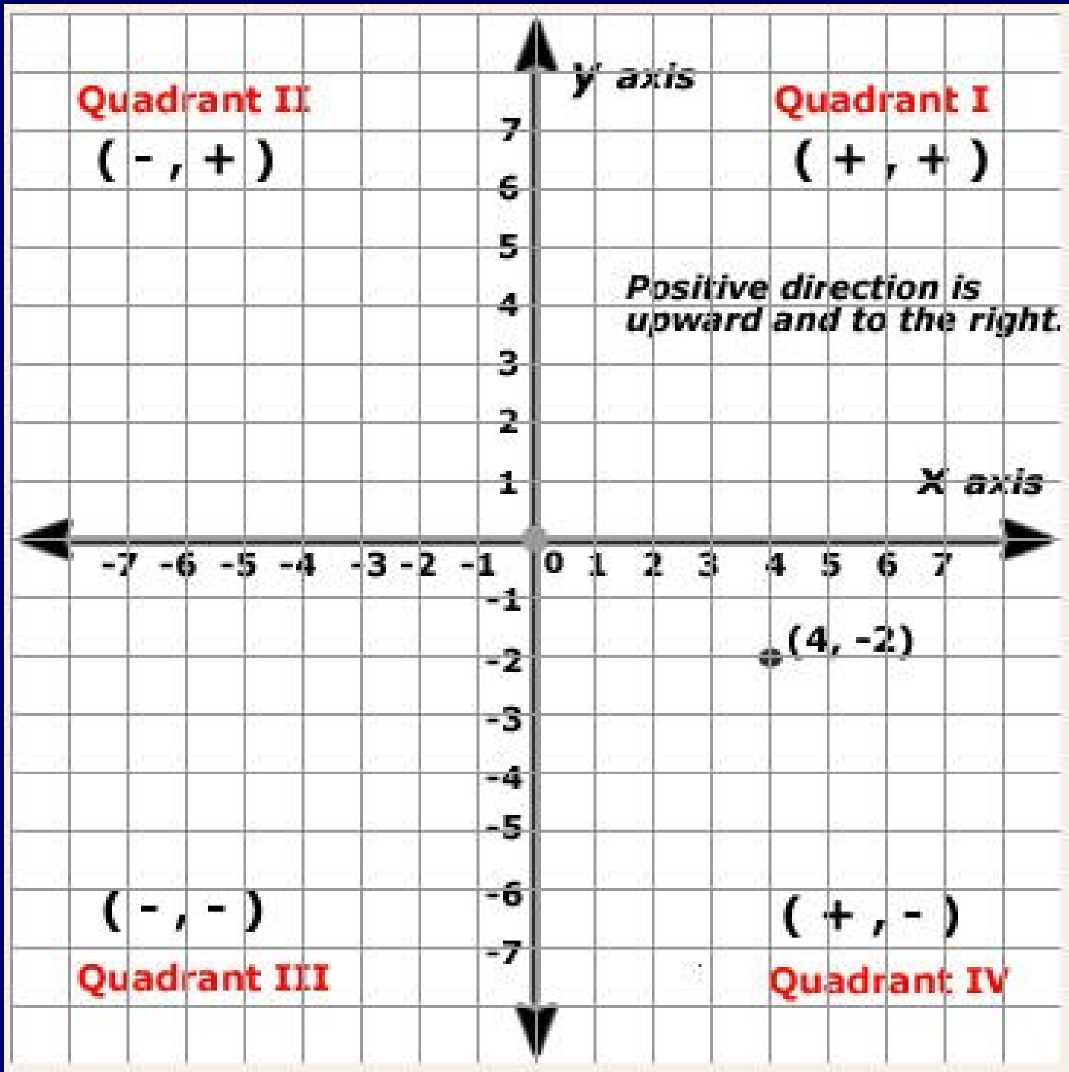
non co-ordinate (applies locally)

e.g. Post Codes, Zip Codes

COORDINATES

- In all groups we must have a point of origin and two reference line

    1. Horizontal line
    2. Vertical line



GEOGRAPHIC COORDINATES

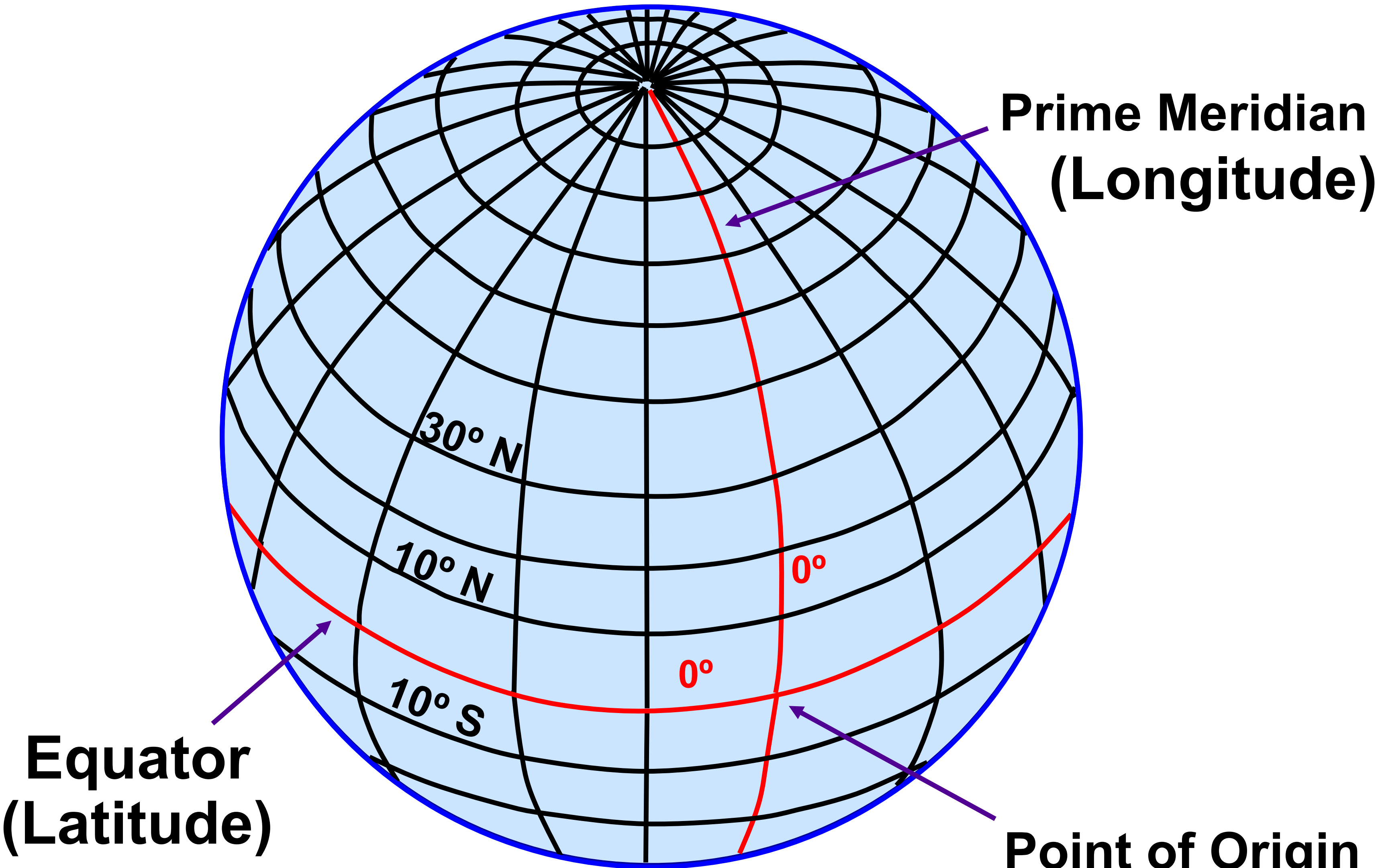
(spherical-3 dimensional)

- **The system was first devised by Hipparchus (190-120 BC)**
- **The geographical coordinates system, employing latitude and longitude**
- **Geographical coordinates system is primary location reference system for the earth**
- **It has always been used in cartography and for all basic locational calculation, such as navigation and surveying.**

Geographical Coordinate System (spherical-3 dimensional)

- Common globe referencing system
  - based on perfect sphere – not true earth shaped ellipse
- Lines of Latitude (parallels)
  - directions E/W
  - distance N/S
  - equator to poles
  - 0^0-90^0
- Lines of Longitude (meridians)
  - directions N/S
  - distance E/W
  - prime meridian (Greenwich, England)
  - 0^0-180^0

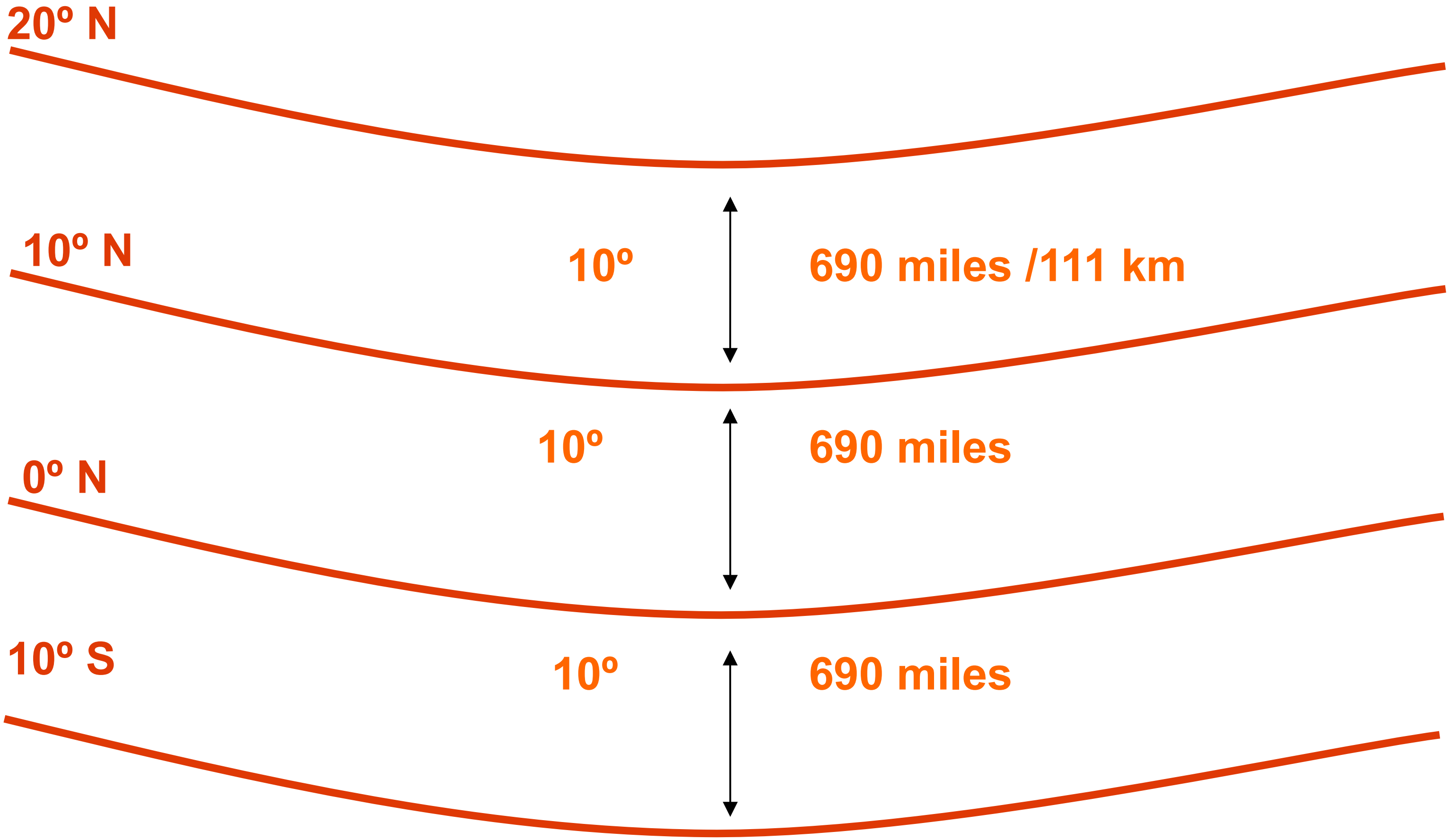
LATITUDE & LONGITUDE



LATITUDE

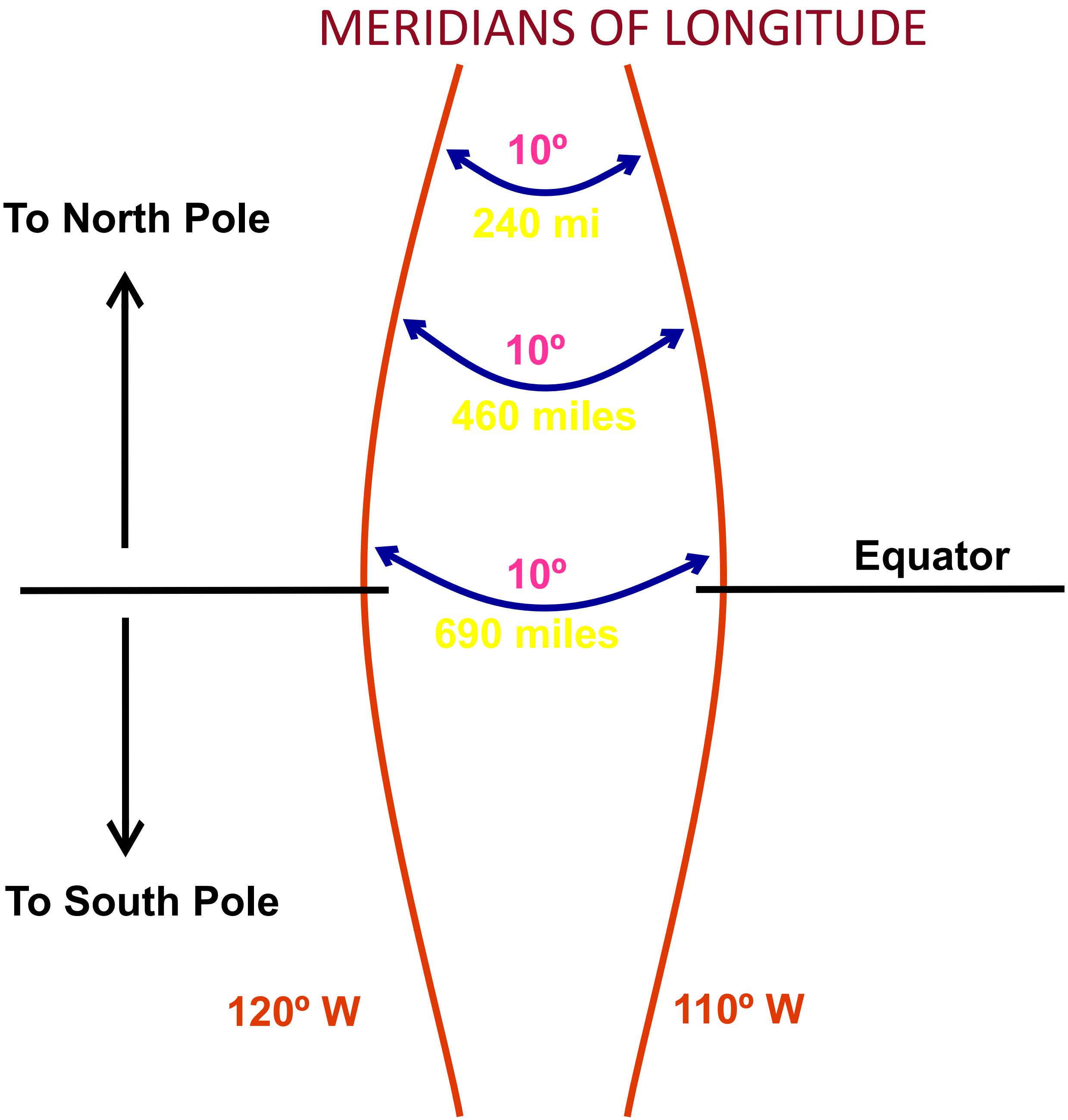
- Latitude is comprised of parallels, which are equally spaced circles around the Earth paralleling the equator.
- Parallels are designated by their angle north or south of the equator (10º, 20º, etc) .
- The equator is 0º latitude, and the North and South Poles are at 90º angles from the equator.
- The linear distance between parallel (latitude) lines never changes, regardless of their position on the Earth.
- Latitude was measured thousands of years ago using the height of the sun at noon, or by the north star position.
- 1 degree = 60 ‘ (minutes) 1’ = 60 “ (seconds)

PARALLELS OF LATITUDE



LONGITUDE

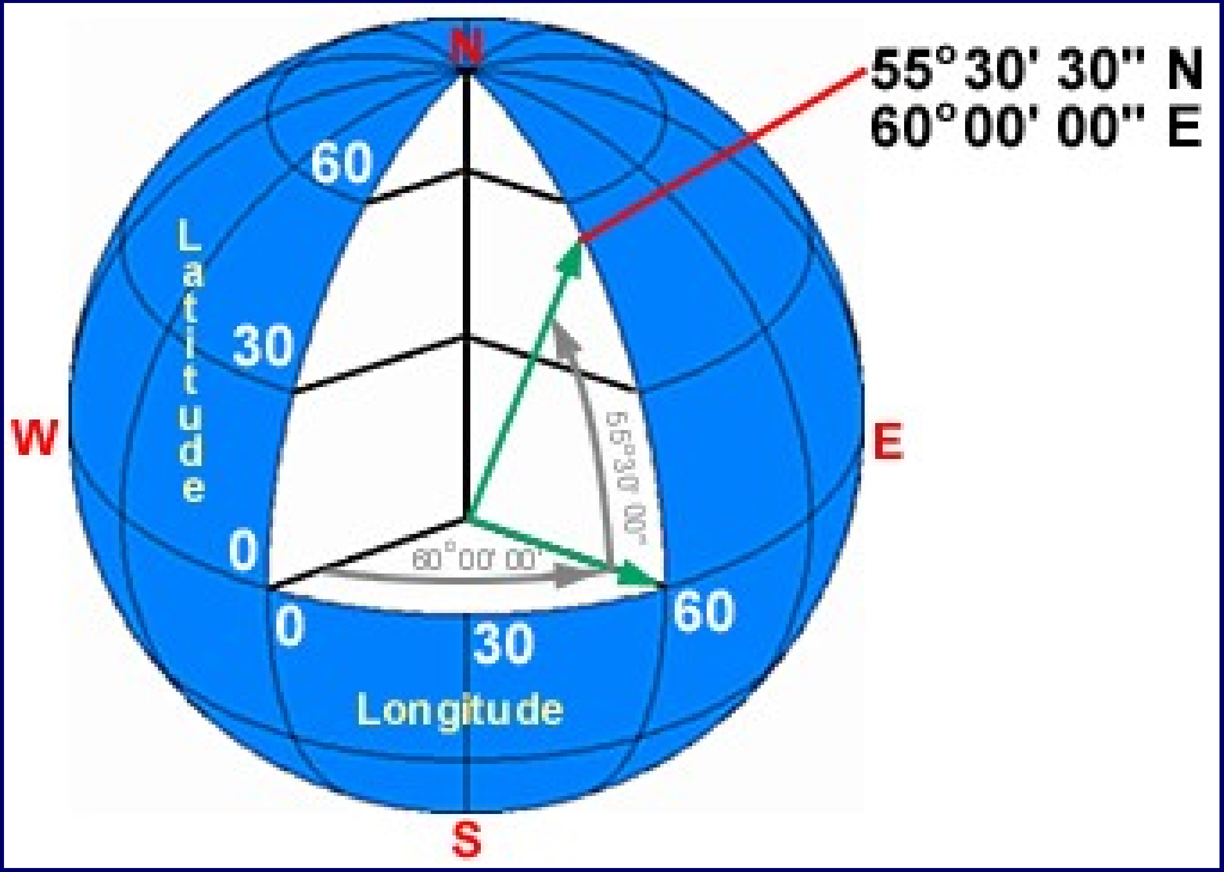
- **Longitude is comprised of meridians that form one-half of a circle, or plane.**
- **Meridians are designated by their angle west or east of the prime meridian.**
- **The Prime Meridian is designated 0º and extends from the North Pole to the South Pole through Greenwich, England.**
- **Meridians are angled, and so are not equidistant from each other at different points.**
- **1 degree longitude varies from 111 km at equator to 0 at poles (approx. half = 55.8 km, at 60 degrees latitude).**


MERIDIANS OF LONGITUDE
10º
240 mi
To North Pole
10º
460 miles
Equator
10º
690 miles
To South Pole
120º W
110º W

READING LATITUDE & LONGITUDE

- reference system uses degrees & minutes
  - 55^0 30’30” N 60^0 0’ 0” E -- reads as:
    - 55 degrees 30 minutes 30 seconds North (latitude)
    - 60^0 degrees 0 minutes 0 seconds East (longitude)
- 1st set of numbers (55^0 30’30” N) = north of the equator
- second set (60^0 0’0” E) = east of the prime meridian
- N and E tells which quarter of the globe

Reading Latitude & Longitude



latitude
longitude

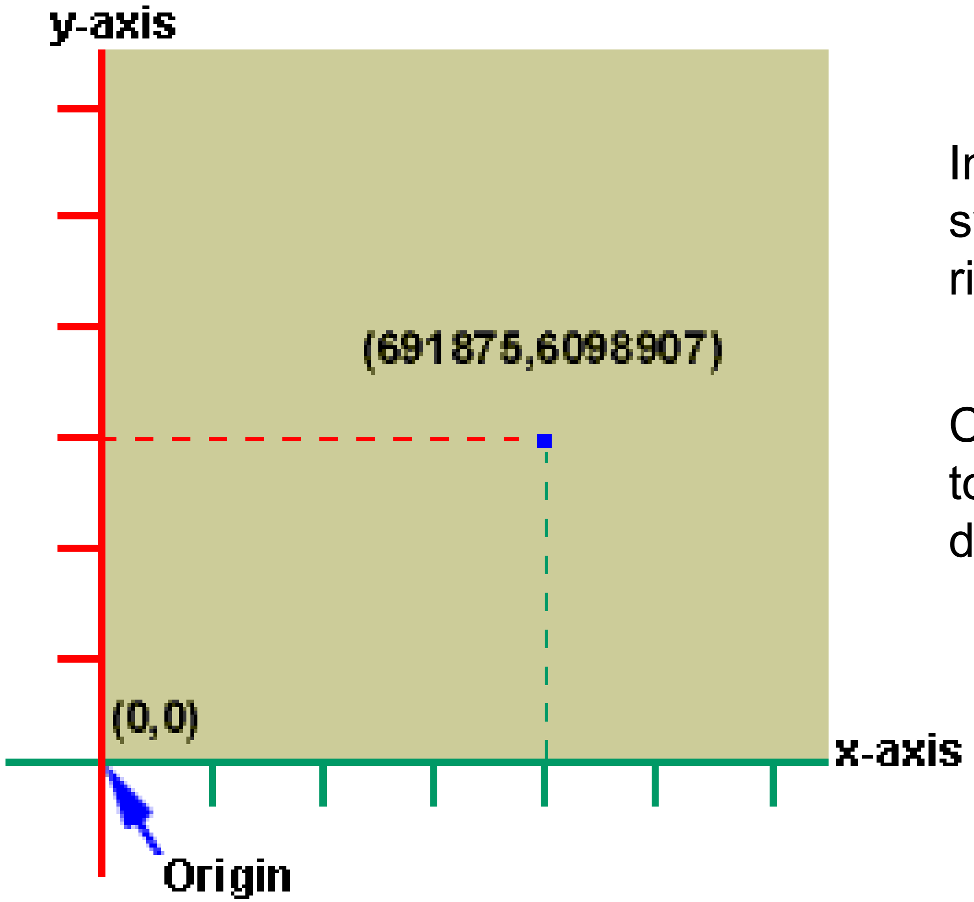
LATITUDE AND LONGITUDE

Length of One Degree of Longitude			Length of a Degree of Latitude		
Latitude	Kilometres	Miles	Latitude	Kilometres	Miles
0º	111.32	69.17	0º	110.57	68.71
10º	109.64	68.13	10º	110.61	68.73
20º	104.65	65.03	20º	110.70	68.79
30º	96.49	59.95	30º	110.85	68.88
40º	85.39	53.06	40º	111.04	68.99
50º	71.70	44.55	50º	111.23	69.12
60º	55.80	34.67	60º	111.41	69.23
70º	38.19	23.73	70º	111.56	69.32
80º	19.39	12.05	80º	111.66	69.38
90º	0.00	0.00	90º	111.69	69.40

Rectangular Coordinate Systems (2 dimensional)

- Also referred to as Planar, Cartesian, or Grid Coordinate Systems (X,Y)
- used for locating positions on a flat map of the earth’s curved surfaces
- principal axes:
  - X = horizontal (easting)
  - Y = vertical (northing)

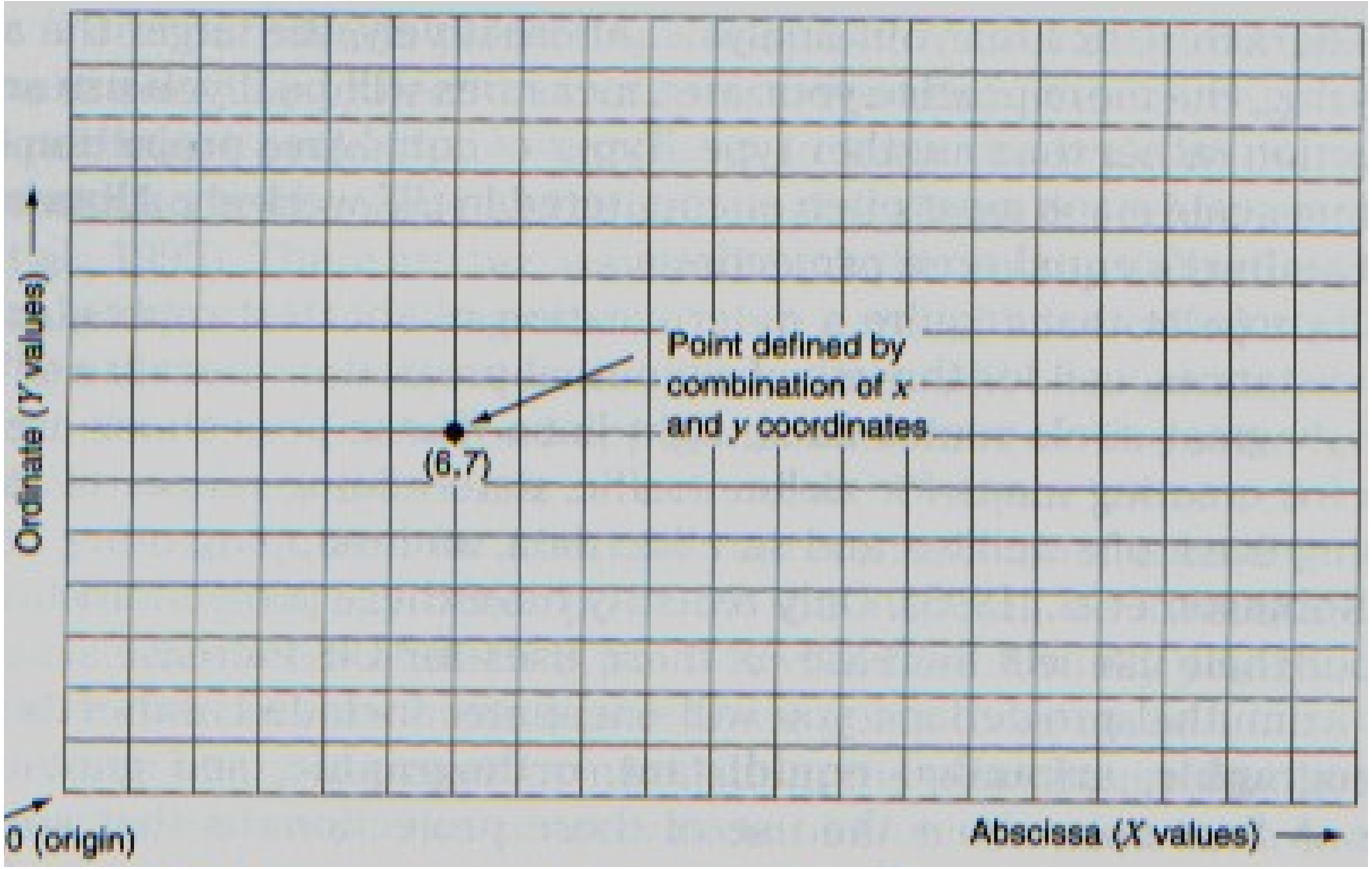
Projected Coordinate Systems (2 dimensional- Rectangular)



In a projected coordinate system, all lines intersect at right angles.

Converting a spherical system to a grid system causes a distortion of the earth surface.

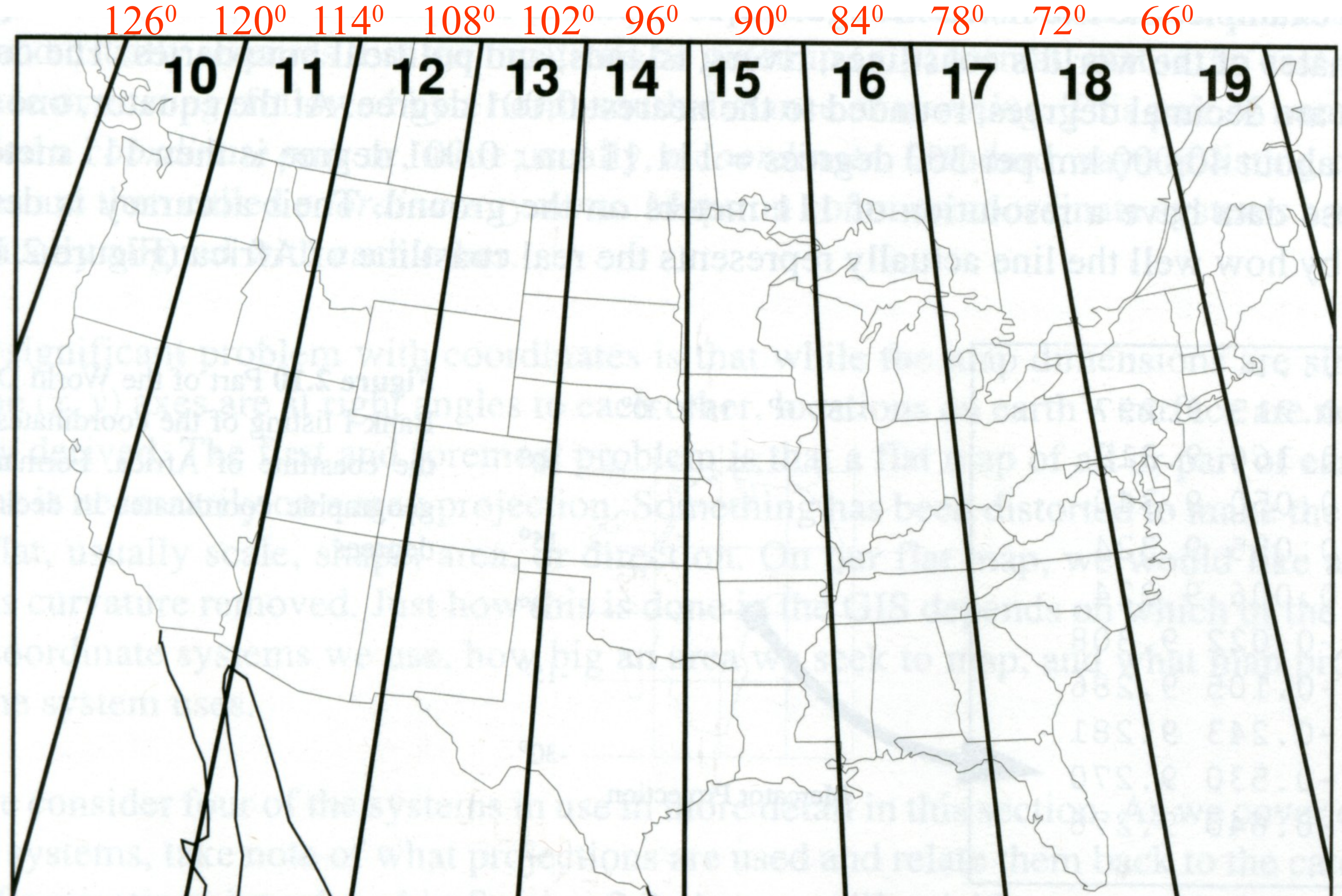
Classic Cartesian (Grid) Coordinate System



Universal Transverse Mercator System (UTM)

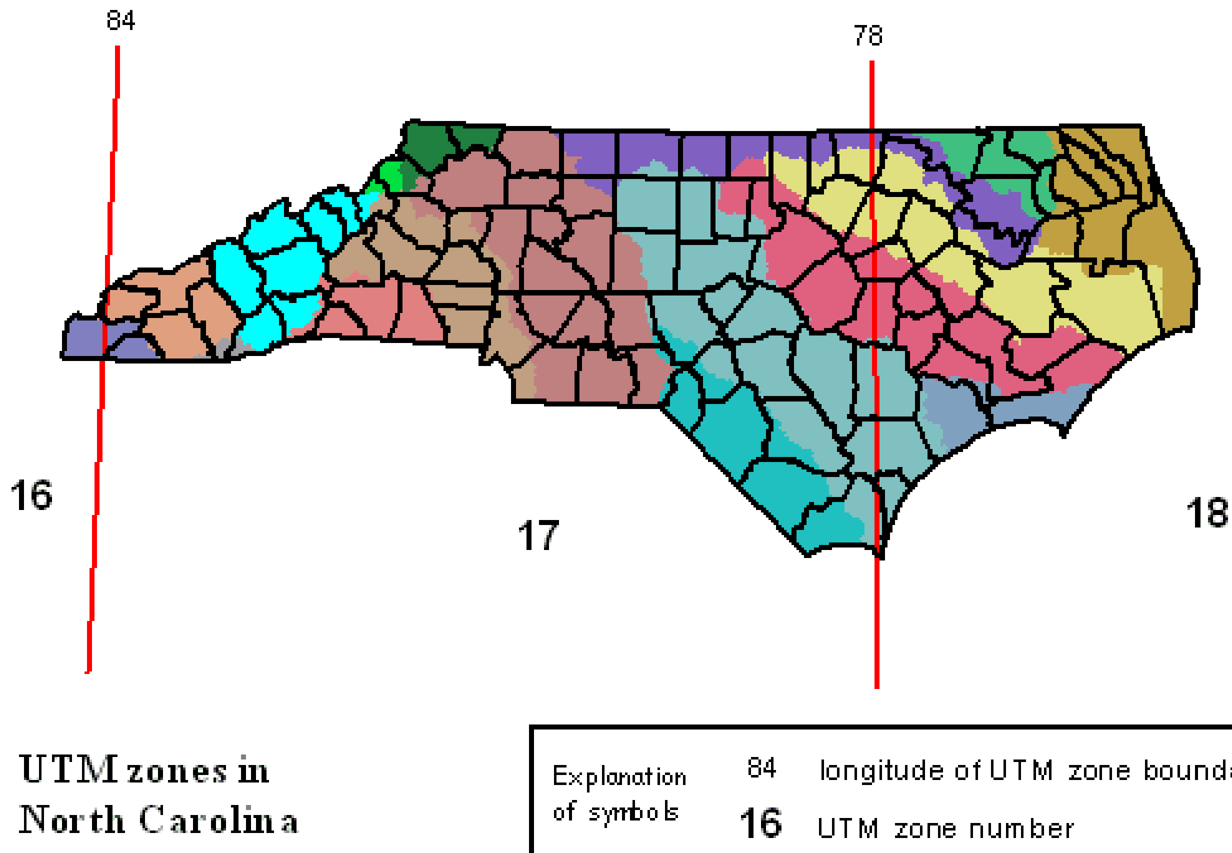
- **Most widely used grid system in GIS**
- **Derived from Transverse Mercator projection**
- **Divides earth from latitudes 84^0 N & 80^0 S into 60 numbered vertical zones 6^0 of longitude wide**
- **Each UTM zone has its own central meridian which covers 3^0 west & 3^0 east of CM**
- **Most USGS topographic maps use**
- **Areas outside of UTM grid (Polar areas) use the UPS grid system**
- **Measurement: metric system**

U.S. UTM Zones



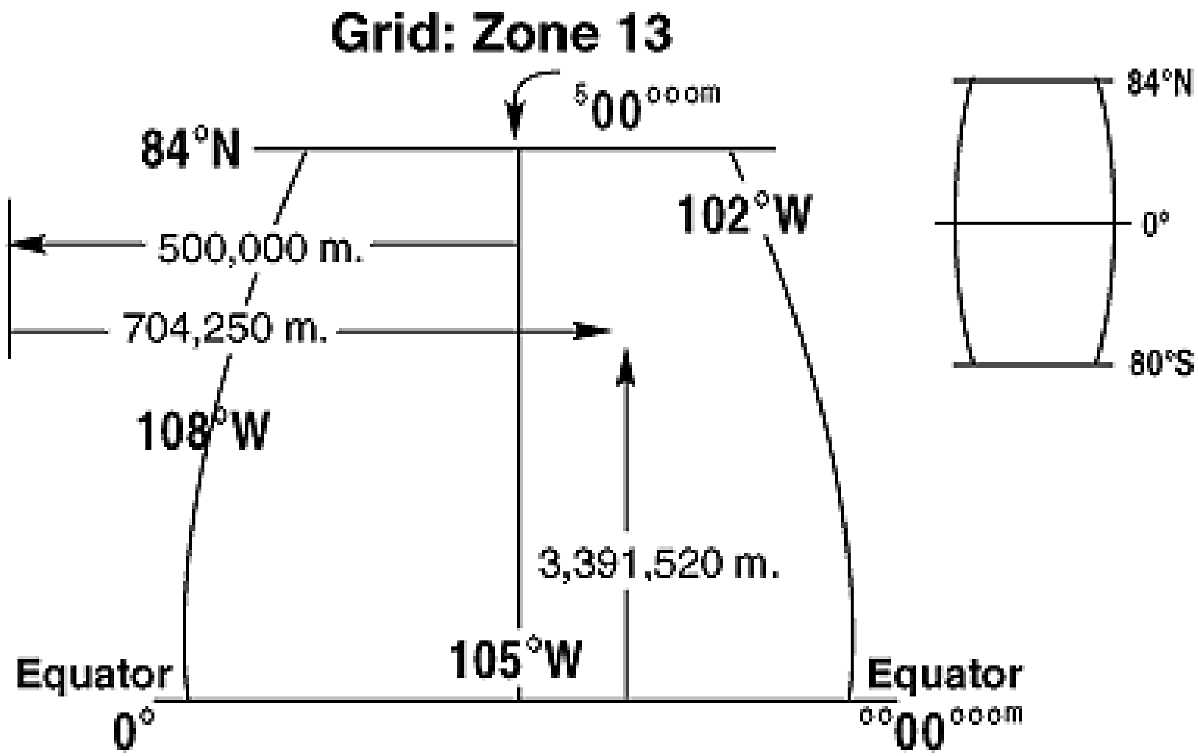
The defense Mapping Agency adopted a special grid for military use throughout the world called the Universal Transvers Mercator (UTM) grid. In this grid the world is divided into 60 north-south zones, each covering a strip 6 degrees wide in longitude. These zones are numbered consecutively beginning with zone 1, between 180 degrees and 174 degrees West longitude, progressing eastward to zone 60, between 174 degrees and 180 degrees East longitude.

UTM Zones in NC



Explanation of symbols	84	longitude of UTM zone bounday
	16	UTM zone number

Example: UTM Grid Zone 13



In each zone coordinates are measured north and east in meters. (one meter = 39.37 inches, slightly more than 1 yard). The northing values are measured continuously from zero at the Equator, in the northerly direction; the northing value of the Equator=0. Similarily done in the Southern direction, from the Equator, but with a northing value of the Equator=10,000,000. A central meridian through the middle of each 6 degree zone is assigned an easting value of 500,000 meters. Grid values to the west of this central meridian are less than 500,000 meters; to the east, more than 500,000 meters.

State Plane Coordinate System

- SPCS is a **grid system developed** by USGS in 1938
- Widely used in **public works & land surveys**
- Divides US, Puerto Rico & US Virgin Islands into over 120 zones
- Each zone has an assigned number (FIPS Code) which defines its projection parameters

State Plane Coordinate System

- **Three projections used for SPCS** determined by geometric direction of state

  - **Lambert Conformal Conic**: for states longer in E/W direction
  - **Transverse Mercator:** for states longer in the N/S direction
  - **Oblique Mercator:** for panhandle of Alaska

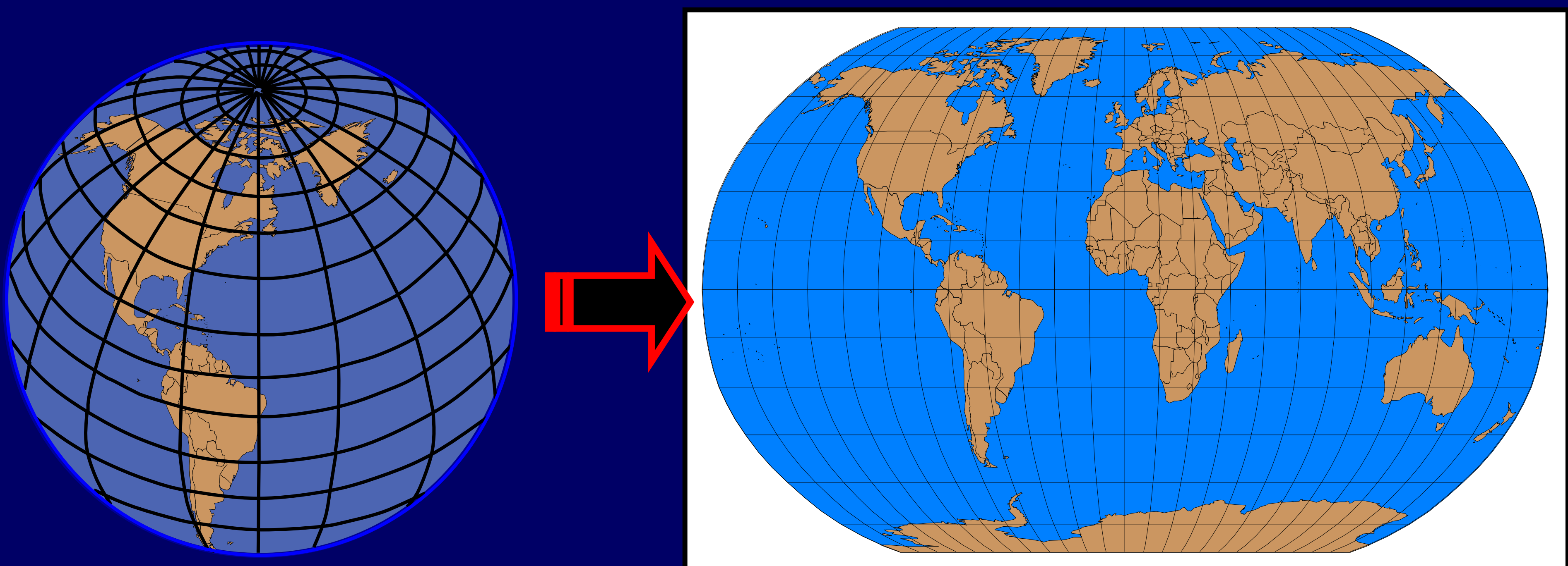
U.S. State Plane Coordinate System Zones & FIPS codes

Lambert Conformal Conic (E/W)

Transverse Mercator (N/S)

MAP PROJECTION

- A map projection is defined as the transformation of the spherical network of latitudes and longitudes on a plane surface irrespective of the method of transformation
- Projection is a systematic transfer of points on a curved surface to a flat projection surface



Projecting a Sphere Onto a Plane Three-dimensional sphere to two-dimensional flat map

KEY PROPERTIES OF MAPS

- **Shape:** shape is preserved when the scale of any point on the map is the same in any direction
- **Area:** proportional relationship between study region and area of the Earth that it represents
- **Distance:** length from center of the projection to any other place on the map
- **Direction:** angles from a point on a line to another point are portrayed correctly in all directions

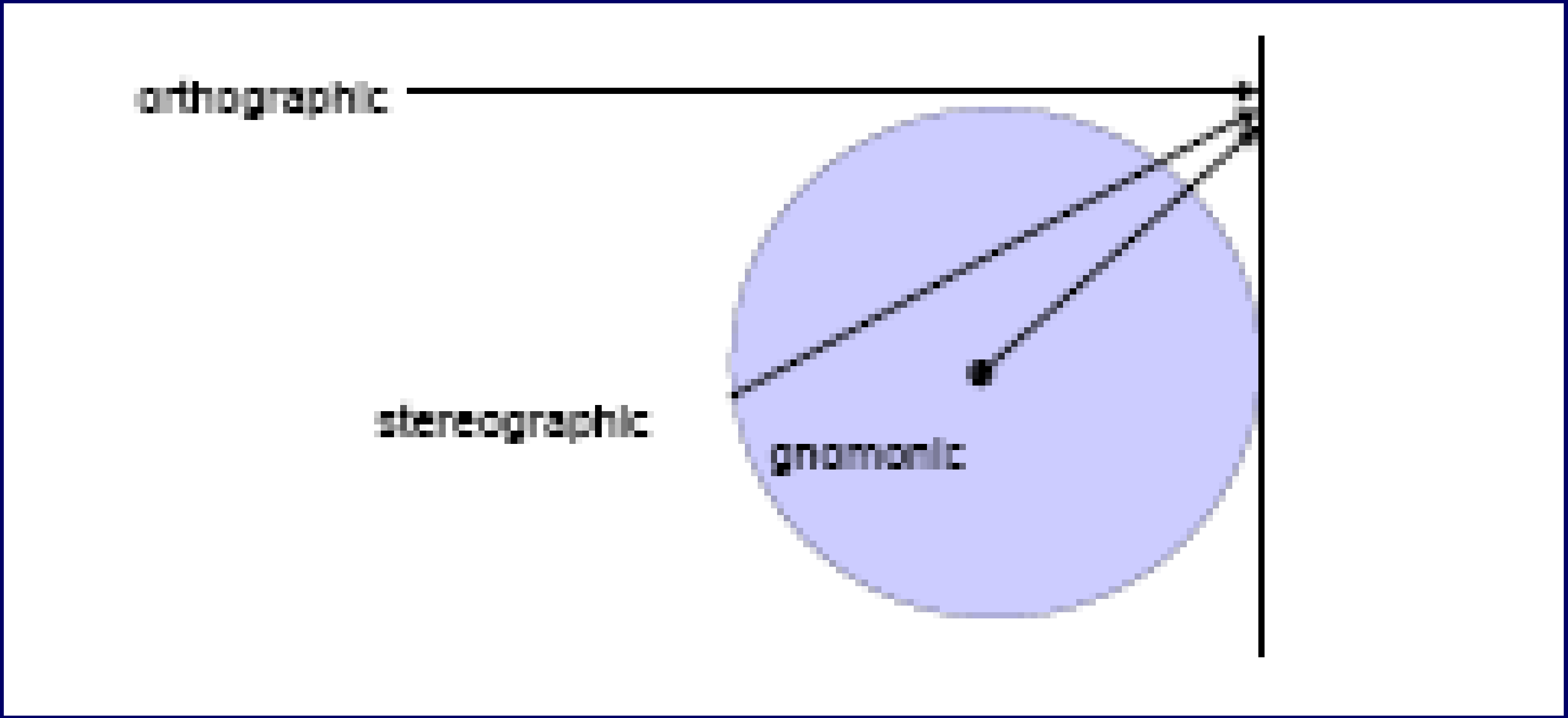
PROBLEMS WITH MAP PROJECTIONS

- Maps can preserve some of the properties but not all
- No projection can retain more than one of these properties over large portion of globe
- Every map projection distorts the earth in several ways
- Decisions of which projection to use depends on reducing distortion

PROJECTION PARAMETERS

- **Projection Center:** point of projection
- **Projection Families:** Developable Surfaces
- **Projection Aspect:** orientation of developable surfaces
- Location of developable surface

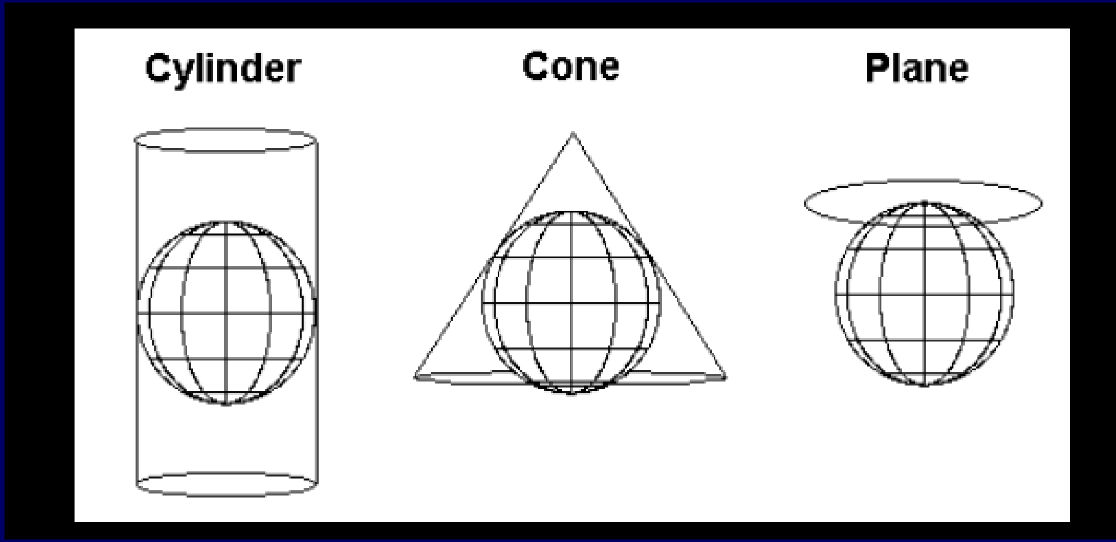
PROJECTION CENTERS



- **Gnomonic:** projection point from Earth's center
- **Stereographic:** at antipodal surface
- **Orthographic:** at infinity

3 FAMILIES OF PROJECTIONS

- Cylindrical
- Conical
- Planar



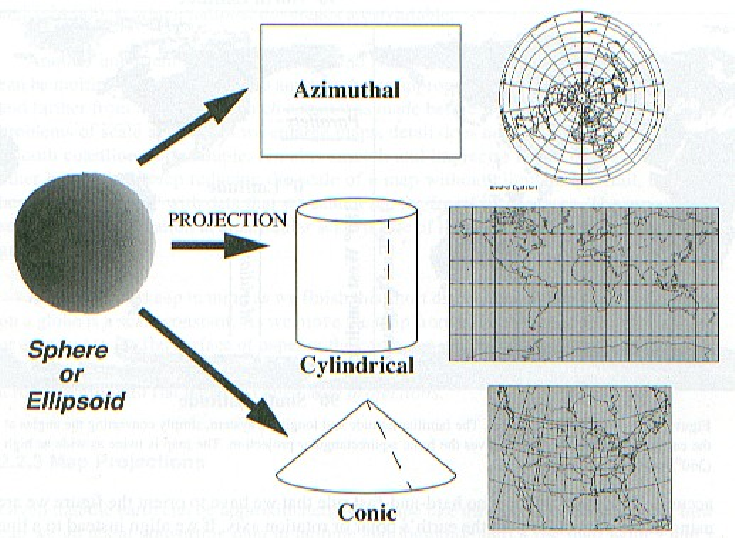
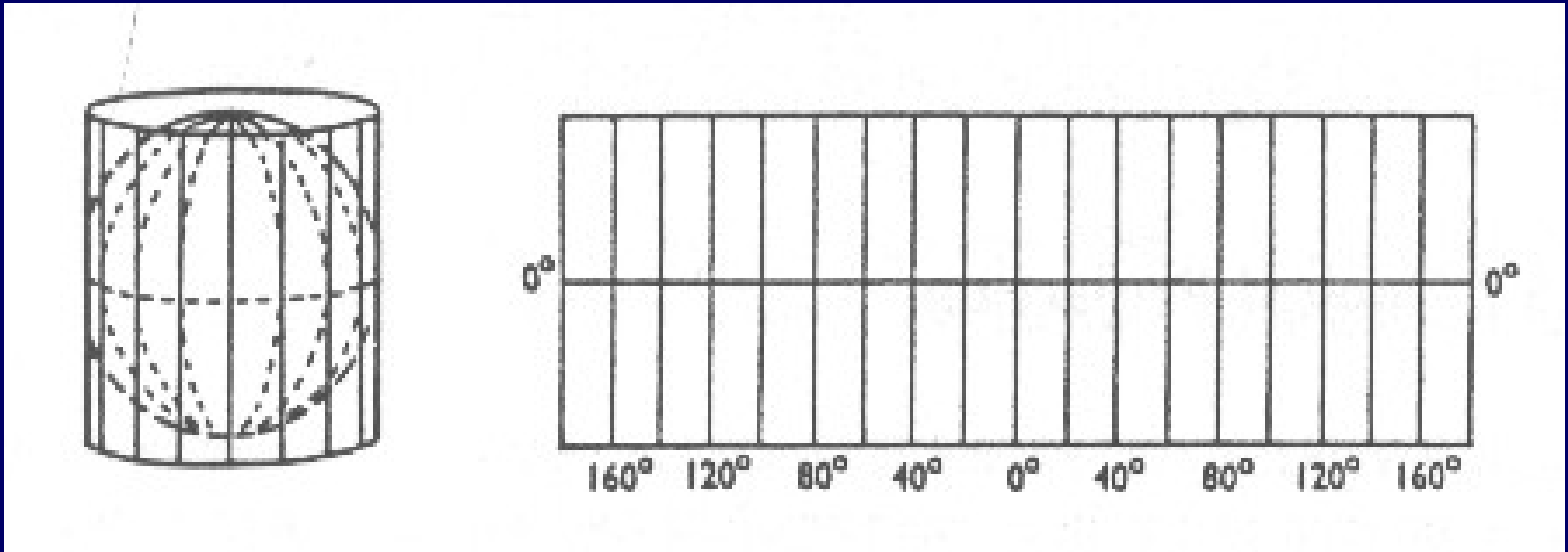


Figure 2.6 The earth can be projected in many ways, but basically onto three shapes that can be unrolled into a flat map: a flat plane, a cylinder, and a cone.

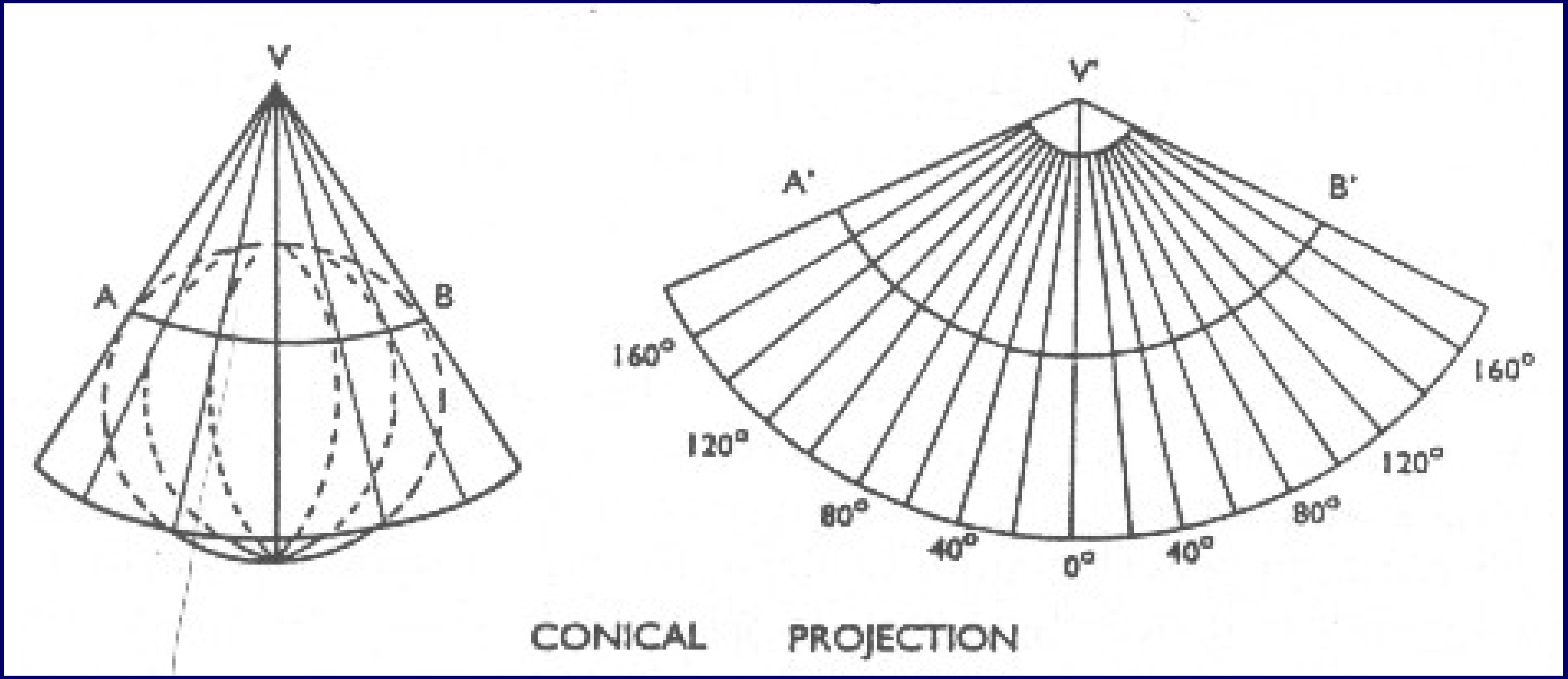
Developable surface: 2 D surface upon which map information is projected. Can be unrolled without distortion although projection will contain distortions

CYLINDRICAL PROJECTIONS



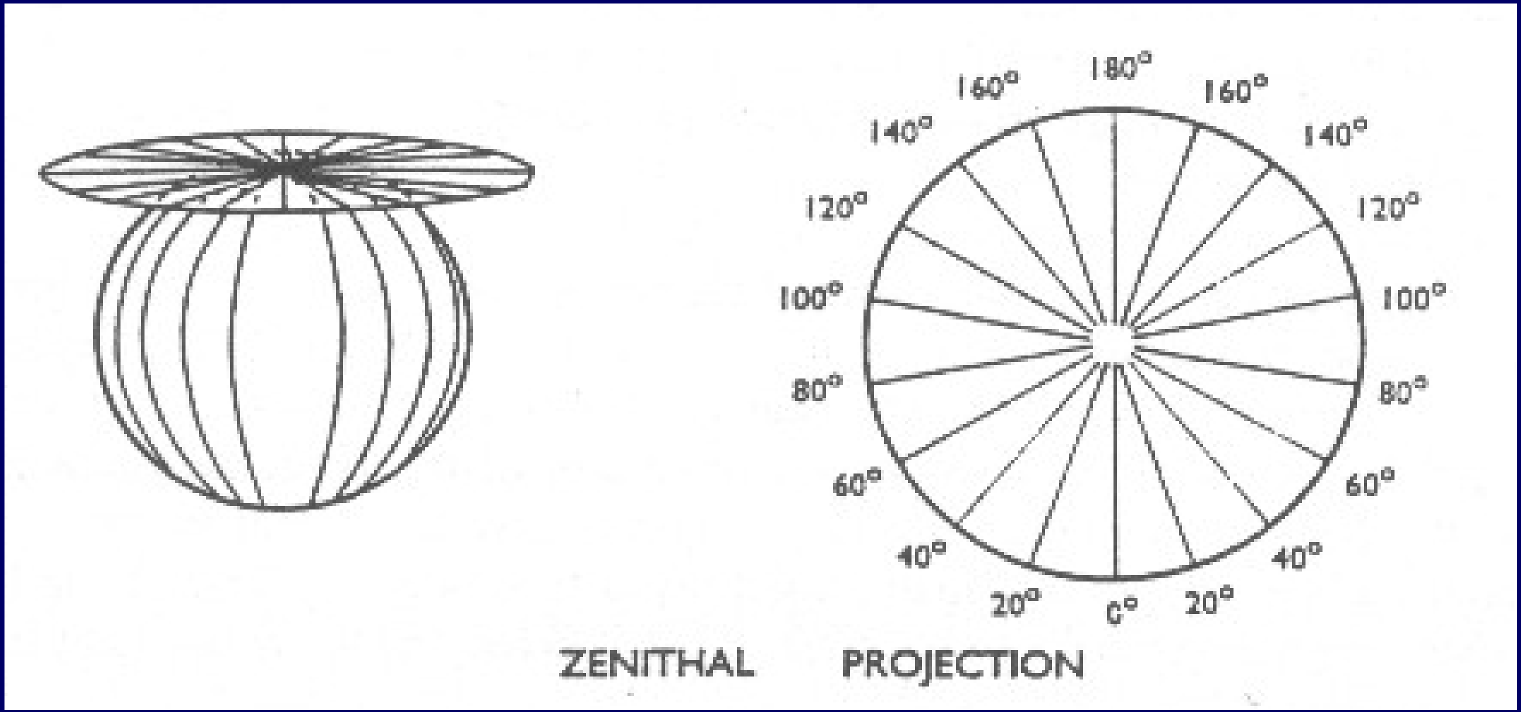
- Used by mariners for navigation
- Meridians run north - south
- Parallels run east- west
- True at equator and distortion increases towards the poles

CONICAL PROJECTIONS



- Used for mid latitude maps
- True at areas between standard parallels

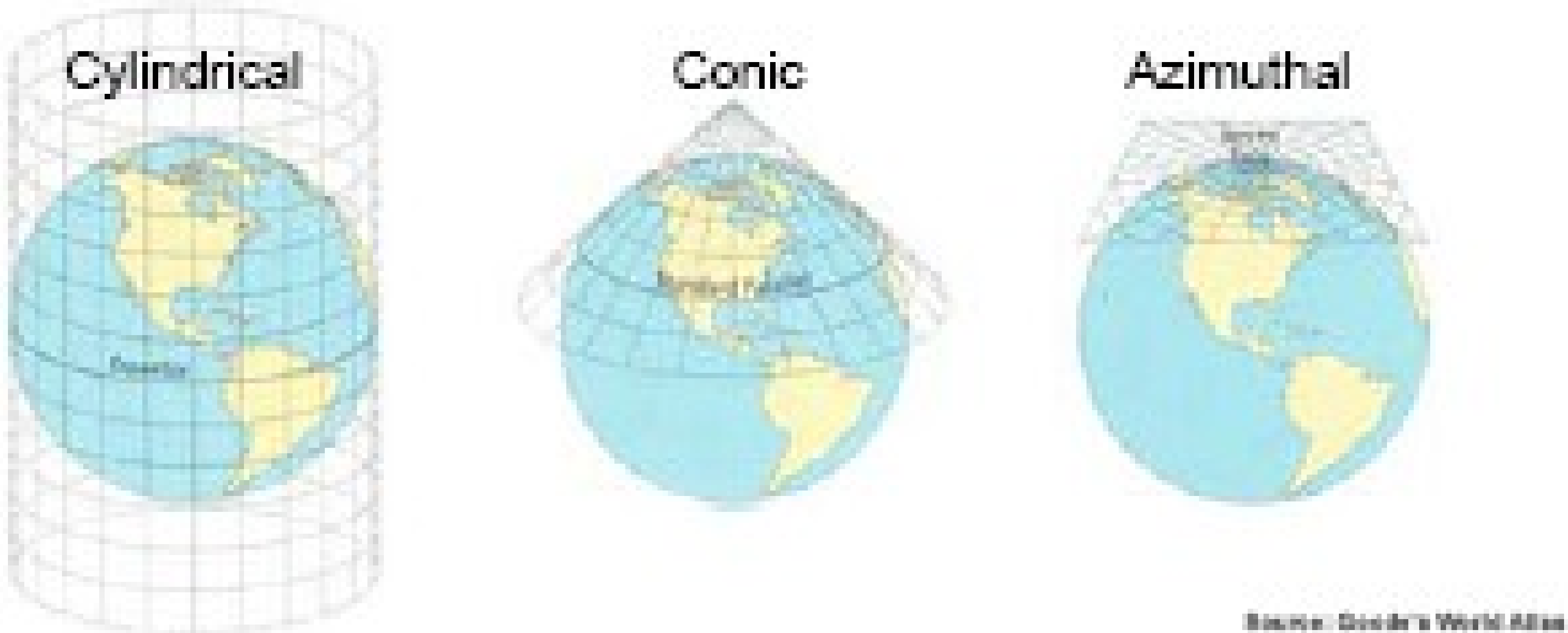
PLANAR PROJECTIONS



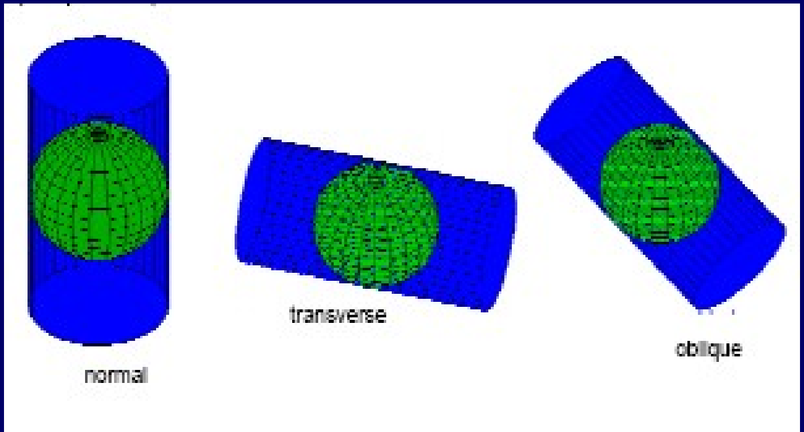
- Used to navigate flight routes
- Preserves Great Circle Lines
- In polar aspect, these maps project meridians as straight lines radiating from the poles and parallels as complete circles centered at the pole
- True at poles and distortions increases outwards

3 BASIC RULES FOR CHOOSING PROJECTION FAMILY

- Three basic rules (after Maling)
  - A country in the tropics asks for a cylindrical projection.
  - A country in the temperate zone asks for a conical projection.
  - A polar area asks for an azimuthal projection.



TYPES OF PROJECTION ASPECT



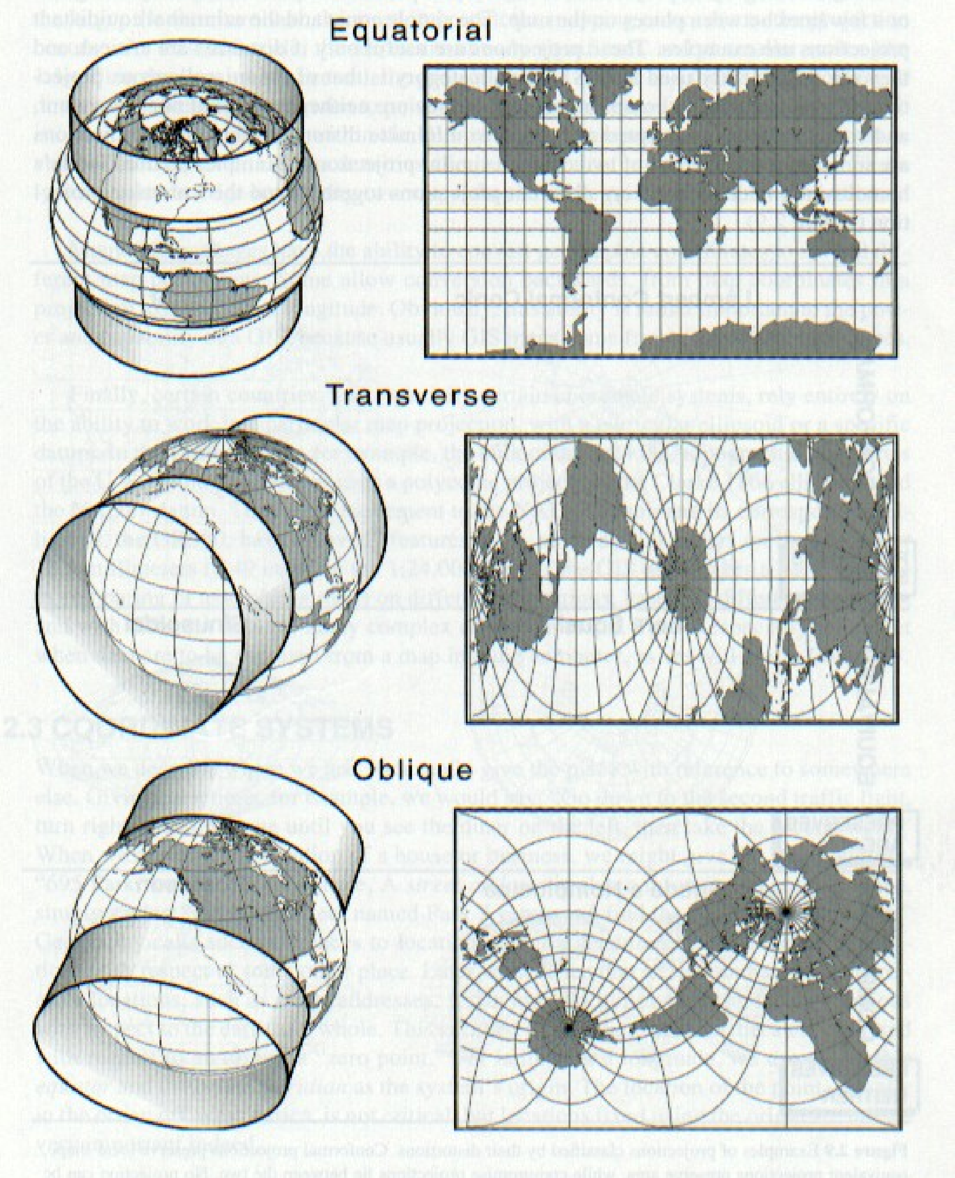


Figure 2.8 Variations on the Mercator (pseudocylindrical) projection shown as secant.

- **Normal:** oriented with polar axis
- **Transverse:** perpendicular to polar axis
- **Oblique:** all others

LOCATION OF DEVELOPABLE SURFACE

- **Tangent** : **DS touches the globe**
- **Secant** : **DS cuts into the globe**

MAP PROPERTY PRESERVATION

- If a projection preserves…
  - ➢ **SHAPE** it is called **CONFORMAL .**
  - ➢ **AREA** is called **EQUAL-AREA or Equivalent.**
  - ➢ **DISTANCE** it is called **EQUIDISTANT**
  - ➢ **DIRECTION** it is called **AZIMUTHAL**

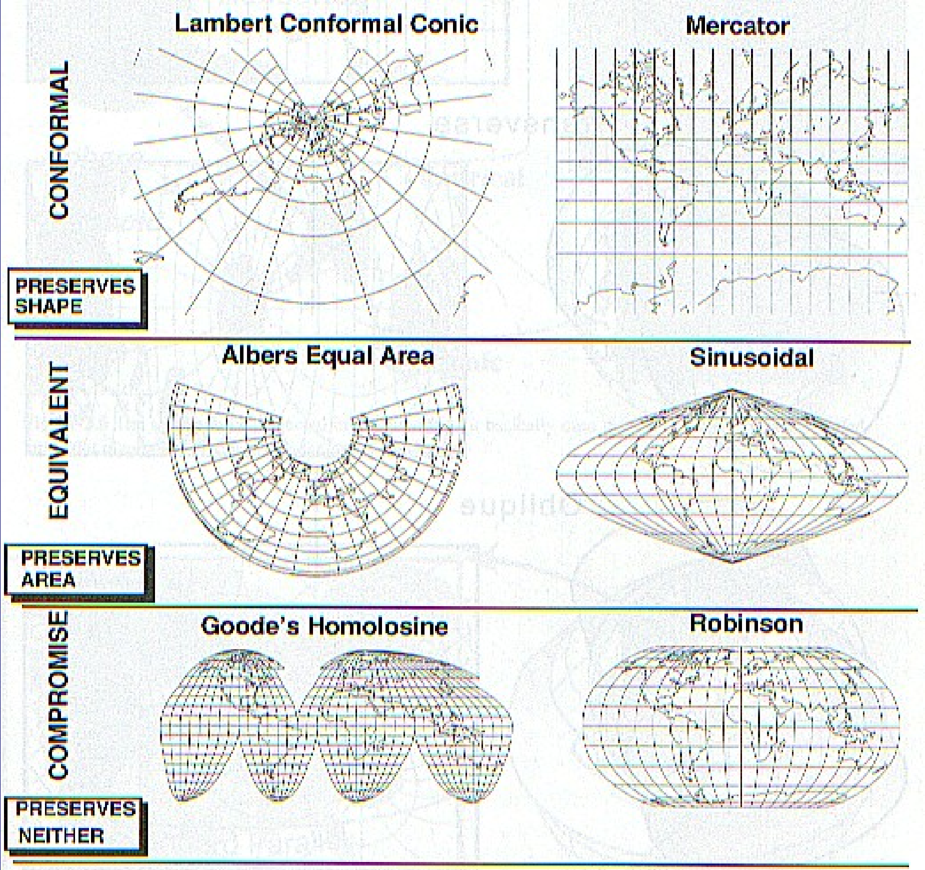


Figure 2.9 Examples of projections classified by their distortions. Conformal projections preserve local shape, equivalent projections preserve area, while compromise projections lie between the two. No projection can be both equivalent and conformal.

Coordinate Transformations

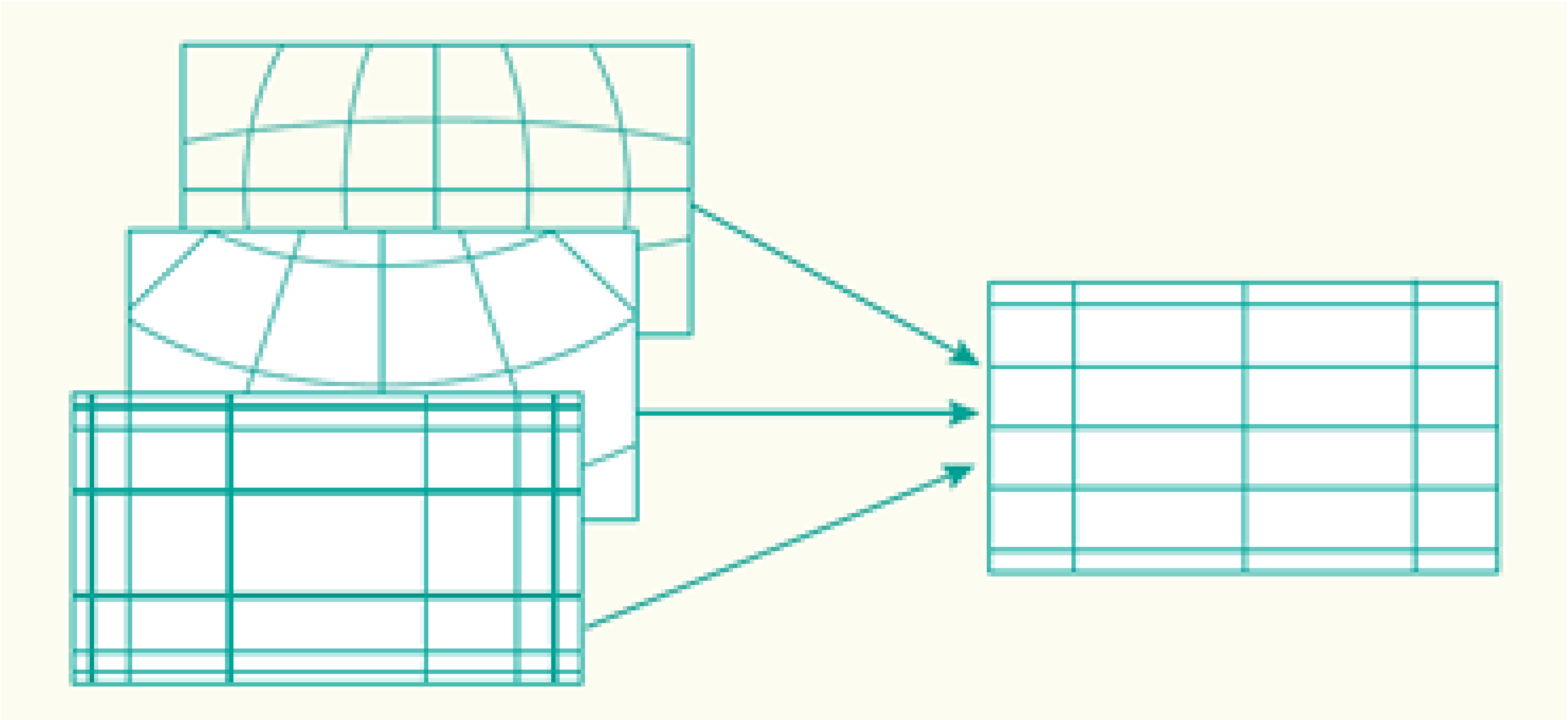
Introduction

- Map and GIS users are mostly confronted in their work with transformations from one two-dimensional coordinate system to another.

- This includes the transformation of polar coordinates delivered by the surveyor into Cartesian map coordinates or the transformation from one 2D Cartesian (x,y) system of a specific map projection into another 2D Cartesian (x,y) system of a defined map projection

- Datum transformations are also important, usually for mapping purposes at large and medium scale.

- An example, map and GIS users are often collecting spatial data in the field using satellite navigation technology and need to represent this data on published maps on a local horizontal datum



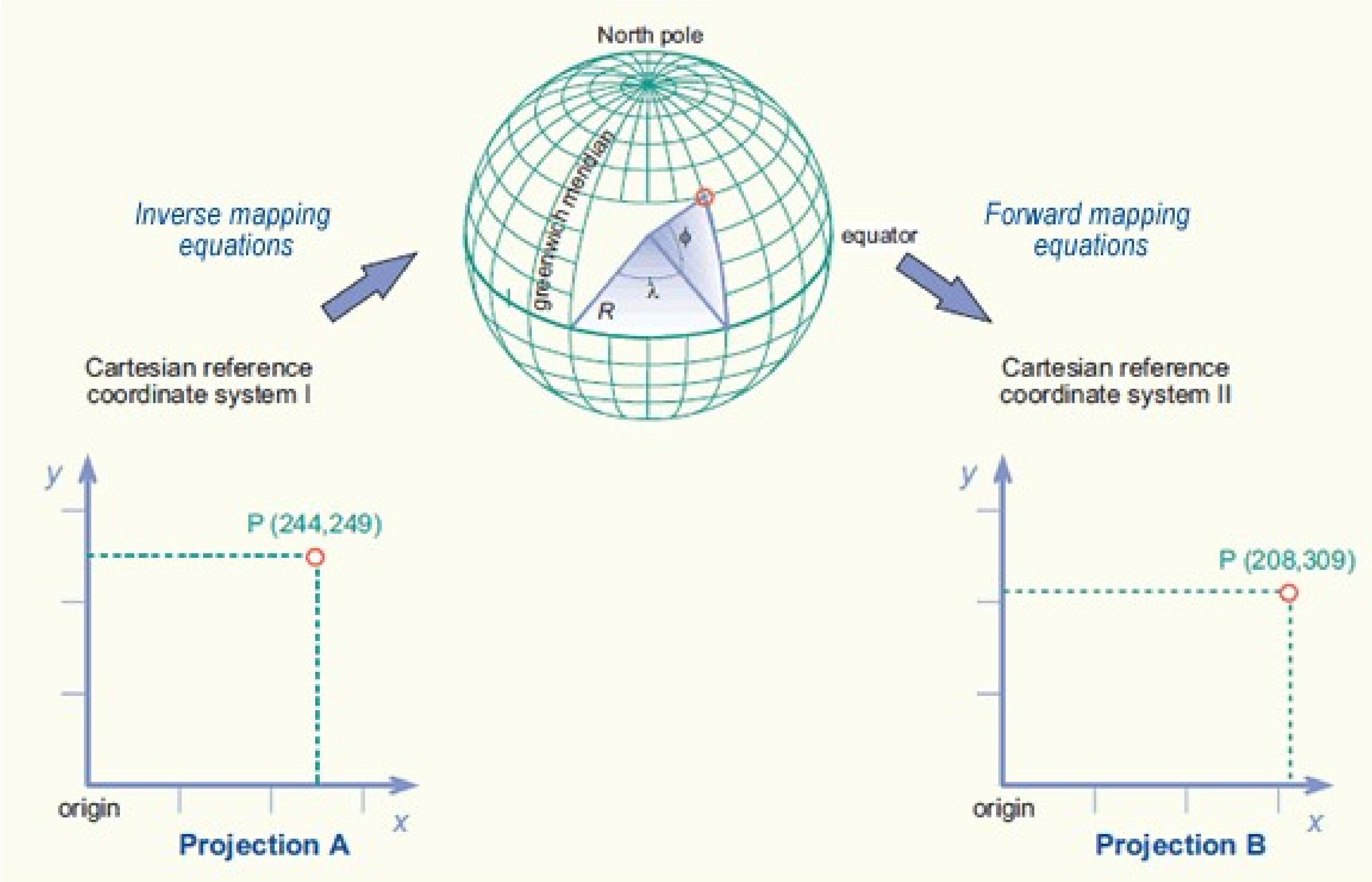
Integration of spatial data into one common coordinate system.

Changing map projection

Forward and inverse mapping equations are generally used to transform data from one map projection to another.

- The ***inverse equation*** of the source projection is used first to transform source projection coordinates (x,y) to geographic coordinates (ϕ,λ).
- Next, the ***forward equation*** of the target projection is used to transform the geographic coordinates (ϕ,λ) to target projection coordinates (x',y').

- The first equation takes us from a projection *A* into geographic coordinates.
- The second takes us from geographic coordinates (ϕ,λ) to another map projection B.



The principle of changing from one into another projection using the mapping equations.

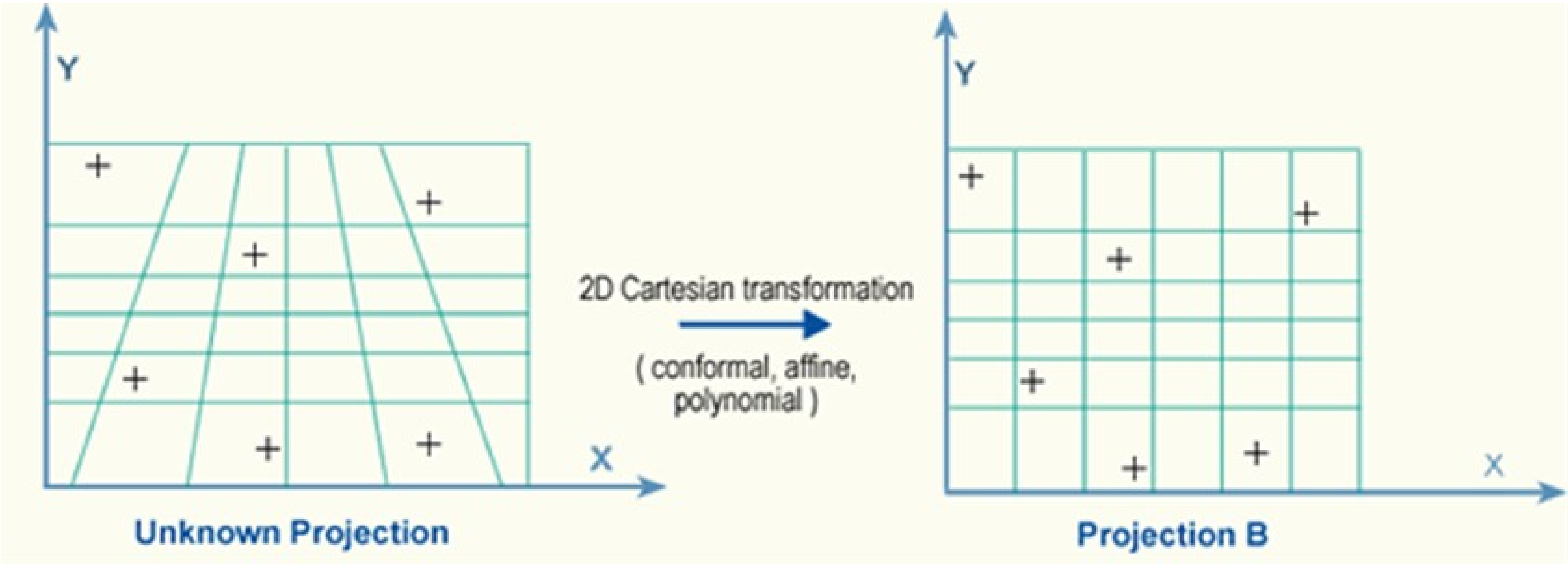
- The coordinate system (map projection) of the input data must be known to use the mapping equations for a projection change. If the coordinate system of the input data is not known it may be possible to use a 2D Cartesian transformation.

- 2D ground control points (GCPs) or common points are then required to determine the relationship between the unknown and the known coordinate system.

- The transformation may be conformal, affine, polynomial, or of another type, depending on the geometric differences between the two map projections.

- Two-dimensional Cartesian transformations have a different accuracy compared to the transformations based on projection equations. The latter take into account the Earth curvature. This is especially important in the case of large areas and small scale.

- However, if the control points are coplanar and the extent of the area is not too large, the 2D Cartesian transformation could yield a better model of coordinate relations than the presumed set of projection equations would do



The principle of changing from one unknown projection into a known projection using a 2D Cartesian transformation. A number of 2D control points are required to determine the relation between both systems.

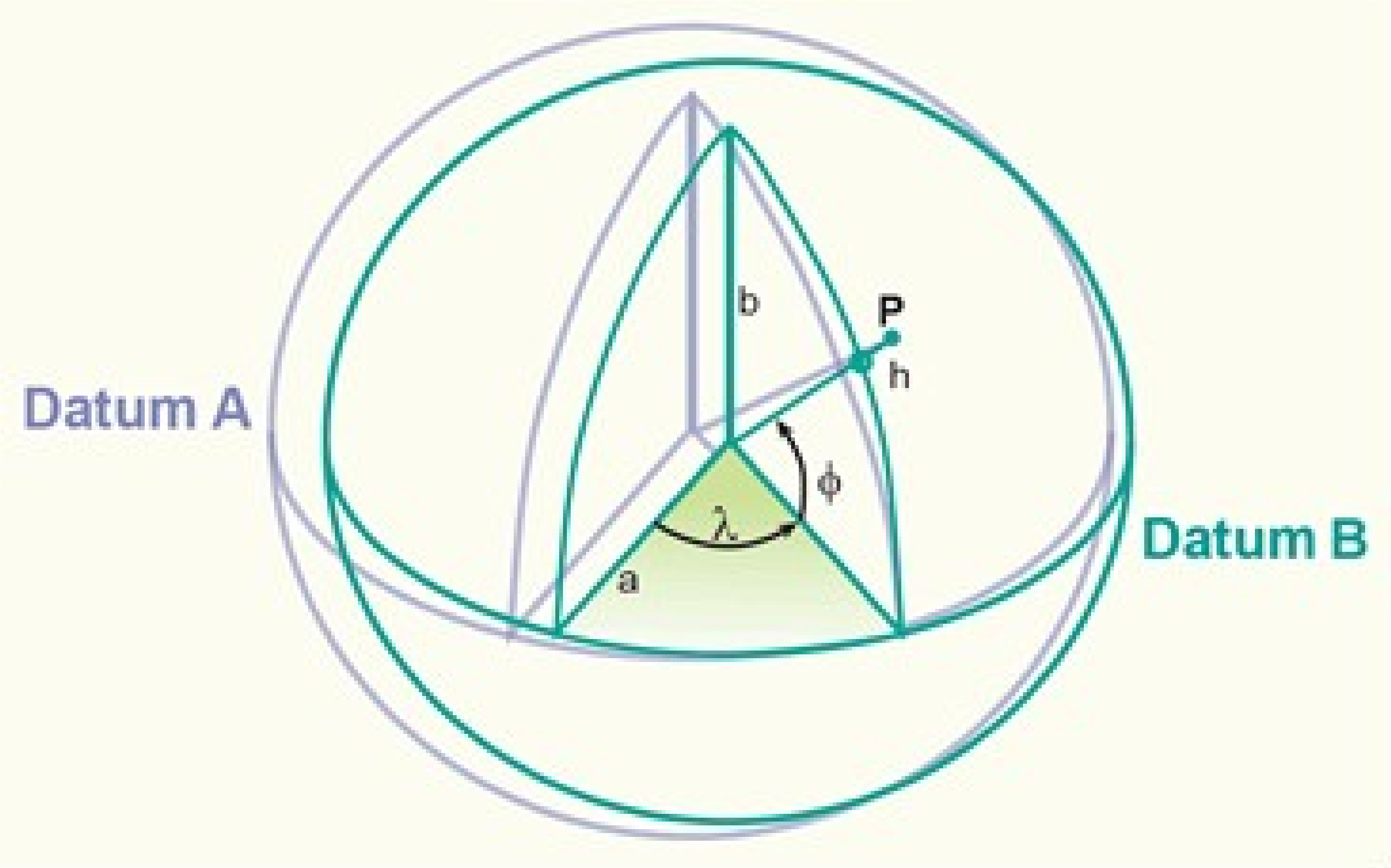
Datum transformations

- A change of map projection may also include a change of the horizontal datum (also called geodetic datum). This is the case when the source projection is based upon a different horizontal datum than the target projection.

- If the difference in horizontal datums is ignored, there will be no perfect match between adjacent maps of neighbouring countries or between overlaid maps originating from different projections. It may result in up to several hundred metres difference in the resulting coordinates.

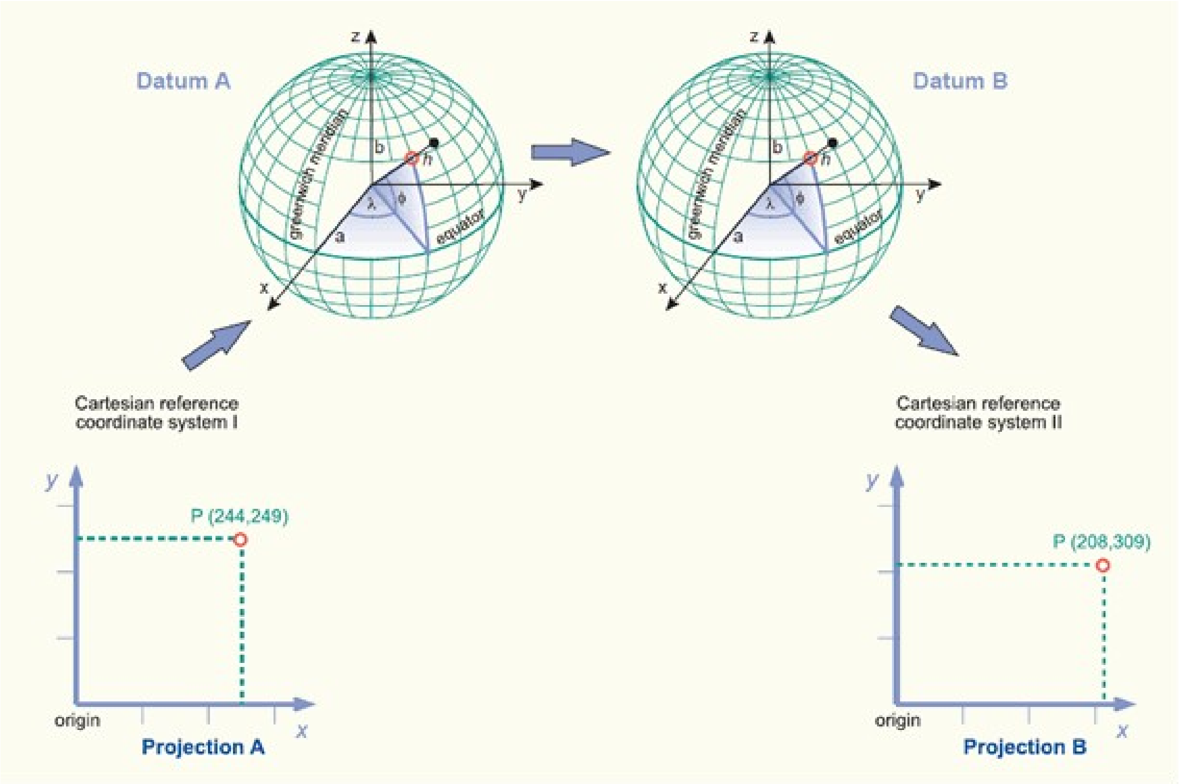
- Therefore, spatial data with different underlying horizontal datums may need a so-called datum transformation.

- ***Datum transformations*** are transformations from a 3D coordinate system (i.e. horizontal datum) into another 3D coordinate system.



Datum shift between two geodetic datums. Apart from different ellipsoids, the centres or the rotation axes of the ellipsoids do not coincide.

- Suppose we wish to transform spatial data from the UTM projection to the Dutch RD system, and that the data in the UTM system are related to the European Datum, while the Dutch RD system underlies the Amersfoort datum.
- In this example the change of map projection must be combined with a datum transformation step for a perfect match.

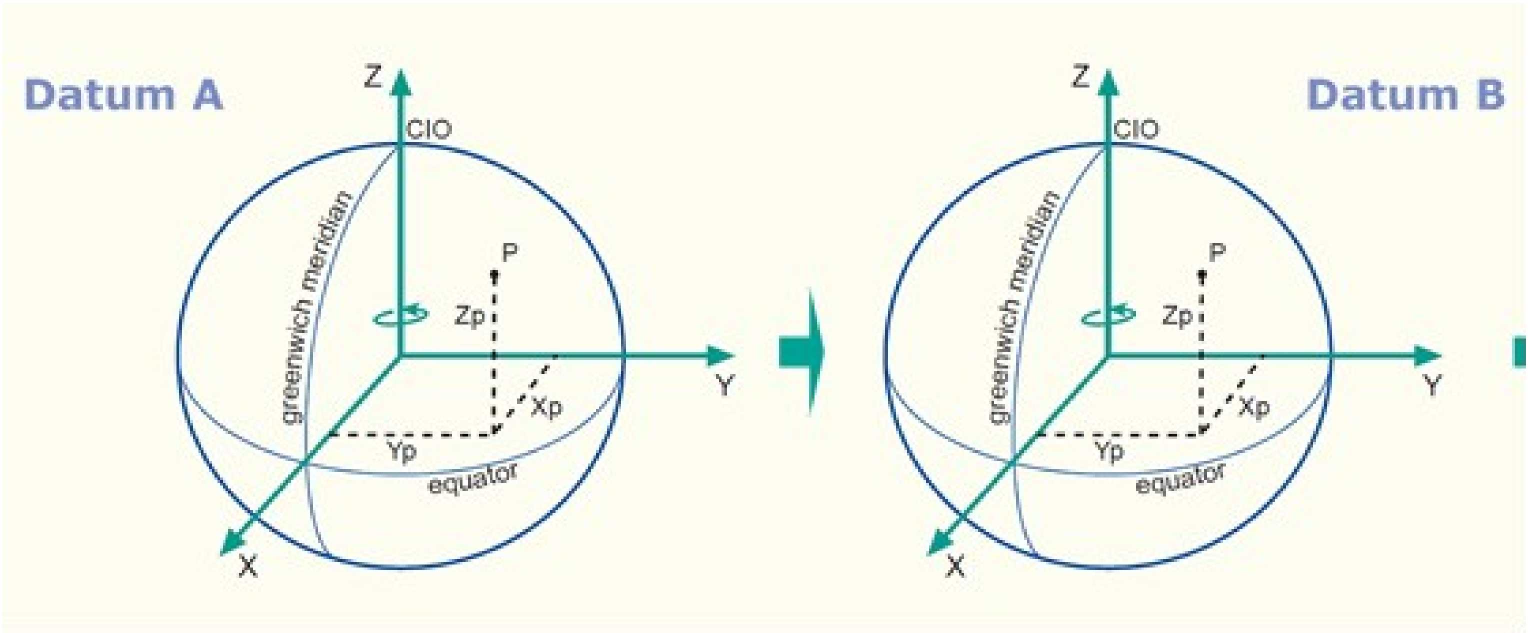


The principle of changing from one into another projection combined with a datum transformation from datum A to datum B.

- The inverse mapping equation of projection A is used first to take us from the map coordinates (x,y) of projection A to the geographic coordinates (ϕ,λ) in datum A.
- Next, the datum transformation takes us from the geographic coordinates (ϕ,λ) in datum A to the geographic coordinates (ϕ,λ) in datum B.
- Finally, the forward equation of projection B is used to take us from the geographic coordinates (ϕ,λ) in datum B to the map coordinates (x',y') of projection B.
- A height coordinate (h or H) may be added to the geographic coordinates.

Datum transformations via geocentric coordinates

- Datum transformations via the geocentric coordinates (*x,y,z*) are ***3D similarity transformations***.
- Essentially, these are transformations between two orthogonal 3D Cartesian spatial reference frames together with some elementary tools from adjustment theory.



The principle of changing from one datum into another datum via the geocentric coordinates.

The three most applied methods for a datum transformation via the 3-dimensional geocentric coordinates are:

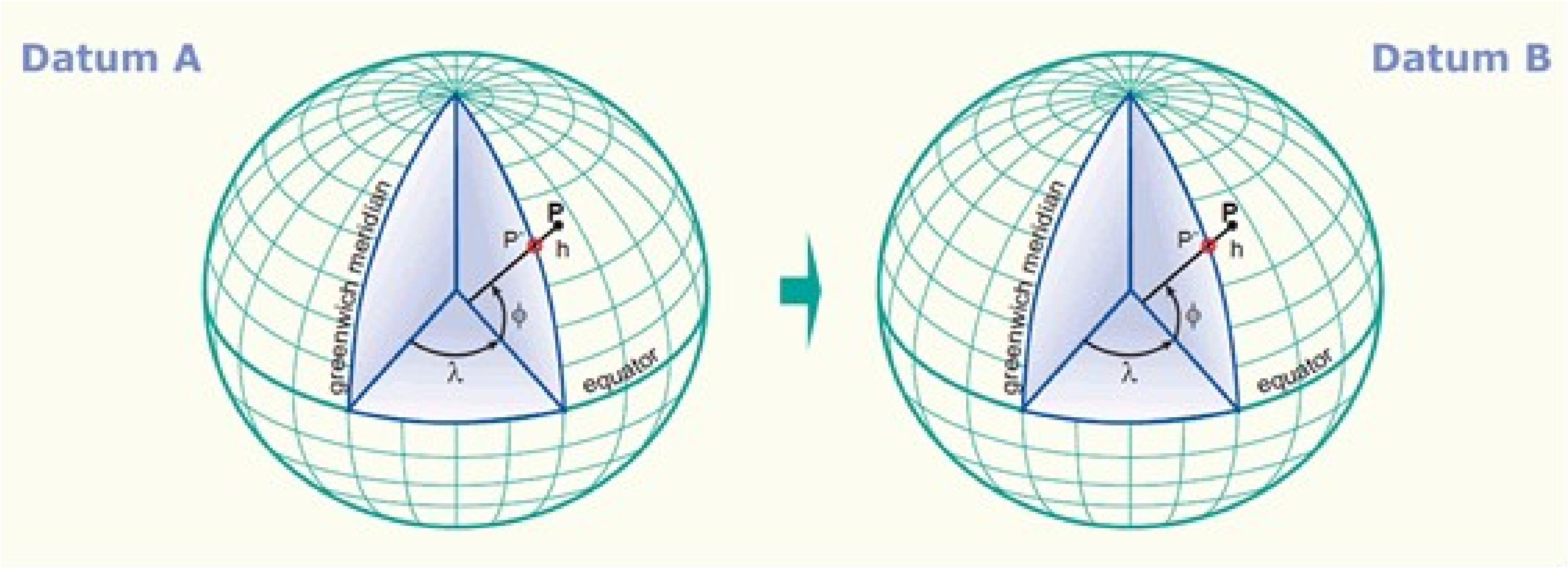
i. The ***geocentric translation***,

ii. The ***Helmert 7-parameter*** transformations (position vector or coordinate frame), and

iii. The ***Molodensky-Badekas 10-parameter*** transformation

Datum transformations via geographic coordinates

- Datum transformations via the geographic coordinates directly relate the ellipsoidal latitude (ϕ) and longitude (λ), and possibly also the ellipsoidal height (h), of both datum systems.



The principle of changing from one datum into another datum via the 3D geographic coordinates.

The applied methods for a datum transformation via the 3-dimensional geographic coordinates are:

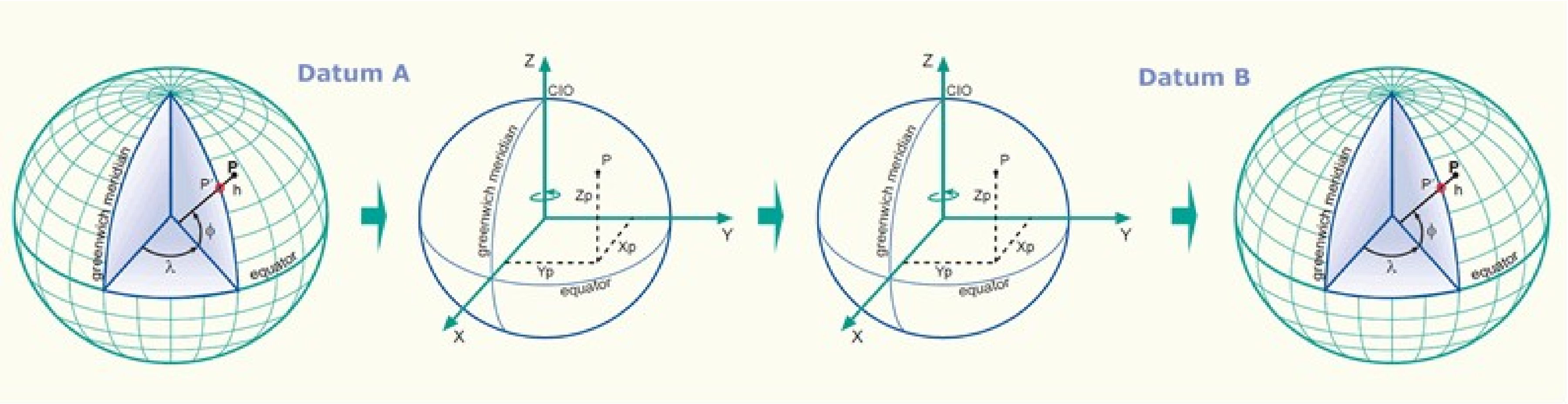
i. The ***geographic offsets***,

ii. The ***Molodensky and Abridged Molodensky*** transformation, and

iii. The ***multiple regression*** transformation

Conversions from geographic to geocentric coordinates and visa versa

➢ The three geocentric transformations decribed are usually combined with conversions between the geocentric coordinates (x,y,z) and the ellipsoidal latitude (ϕ) and longitude (λ) coordinates and height (h) in both datum systems.



The principle of a datum transformation via the geocentric coordinates. The datum transformation is combined with conversions between the 3D geographic coordinates and geocentric coordinates in both datum systems.

Geographic to geocentric conversion

- The conversion from the latitude and longitude coordinates into the geocentric coordinates is rather straightforward and turns ellipsoidal latitude (ϕ) and longitude (λ), and possibly also the ellipsoidal height (*h*), into *X*, *Y* and *Z*, using 3 direct equations.

- If the ellipsoidal semi-major axis is *a*, semi-minor axis *b*, and inverse flattening *1/f*, then

$$X = (\upsilon + h) \cos \phi \cos \lambda$$

$$Y = (\upsilon + h) \cos \phi \sin \lambda$$

$$Z = [(1 - e^2)\,\upsilon + h] \sin \phi$$

where υ is the prime vertical radius of curvature at latitude ϕ and is equal to $\upsilon = a /(1 - e^2 \sin^2 \sqrt{})^{0.5}$,
e is the eccentricity of the ellipsoid where $e^2 = (a^2 - b^2) / a^2 = 2f - f^2$.

Geocentric to geographic conversion

- The inverse equations for the reverse conversion are more complicated and require either an iterative calculation of the latitude and ellipsoidal height, or it makes use of approximating equations like those of Bowring.

- These last have millimetre precision for 'Earth-bound' points, i.e. points that are at most 10 km away from the ellipsoidal surface (any point on the Earth surface).

- An example, the Potsdam (*X,Y,Z*) coordinates of the given point in the state of Baden-Württemberg are:

X (Potsdam) = ***4,156,305.34m***

Y (Potsdam) = ***671,404.31m***

Z (Potsdam) = ***4,774,508.25m***

- These Potsdam (X,Y,Z) coordinates were computed from the ITRF (X,Y,Z) coordinates by applying a Helmert datum transformation.

- Applying the inverse equations to the given point results in the following Potsdam (ϕ,λ,h) coordinates:

ϕ *(Potsdam) =* ***48° 47' 3.2752" N***

λ *(Potsdam) =* ***9° 10' 34.3870" E***

h (Potsdam) = ***278.825m***

2D Cartesian coordinate transformations

2D Cartesian coordinate transformations can be used to transform 2D Cartesian coordinates (x,y) from one 2D Cartesian coordinate system to another 2D Cartesian coordinate system.

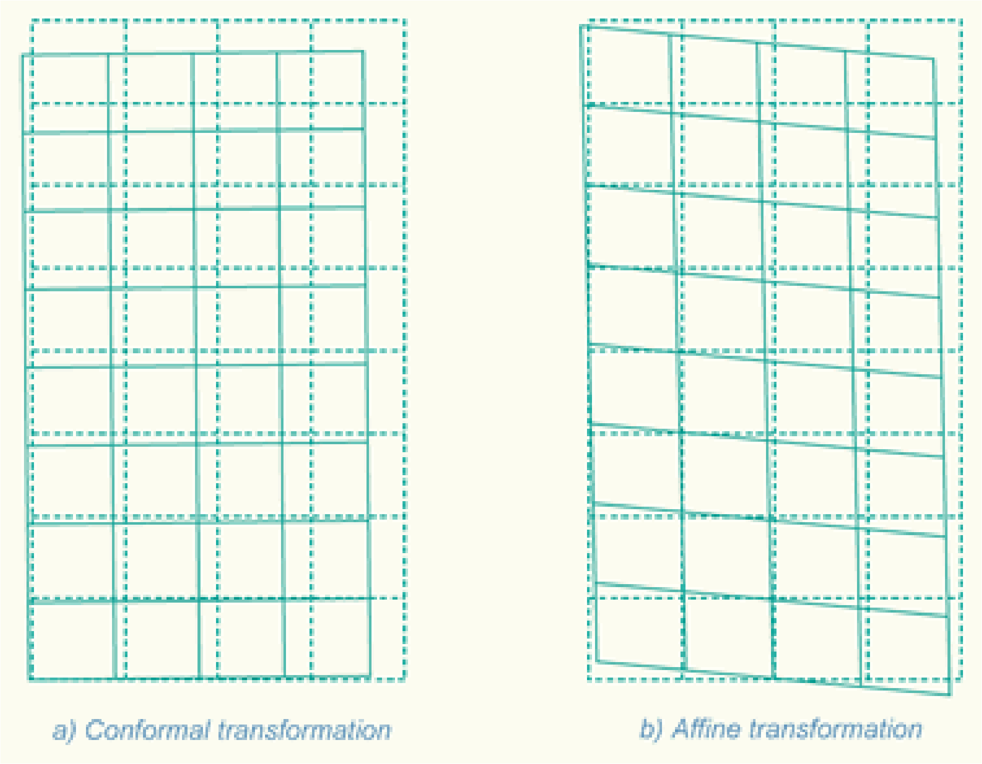
The three primary transformation methods are:

i. The ***conformal*** transformation,

ii. The ***affine*** transformation, and

iii. The ***polynomial*** transformation.

- The difference between a conformal and an affine transformation is illustrated.
- Both are linear transformations which means that the lines of the grid remain straight after the transformation.



a) *The uniform scale change of the conformal transformation retains the shape of the original rectangular grid.*

b) *The different scale in x and y-direction of the affine transformation changes the shape of the original rectangular grid, but the lines of the grid remain straight.*

iii. A ***polynomial*** transformation is a non-linear transformation and relates two 2D Cartesian coordinate systems through a *translation, a rotation* and a *variable scale change.*

The transformation function can have an infinite number of terms.

The equation is

$$X' = x_o + a_1X + a_2Y + a_3XY + a_4X^2 + a_5Y^2 + a_6X^2Y + a_7XY^2 + a_8X^3 + \ldots\ldots$$

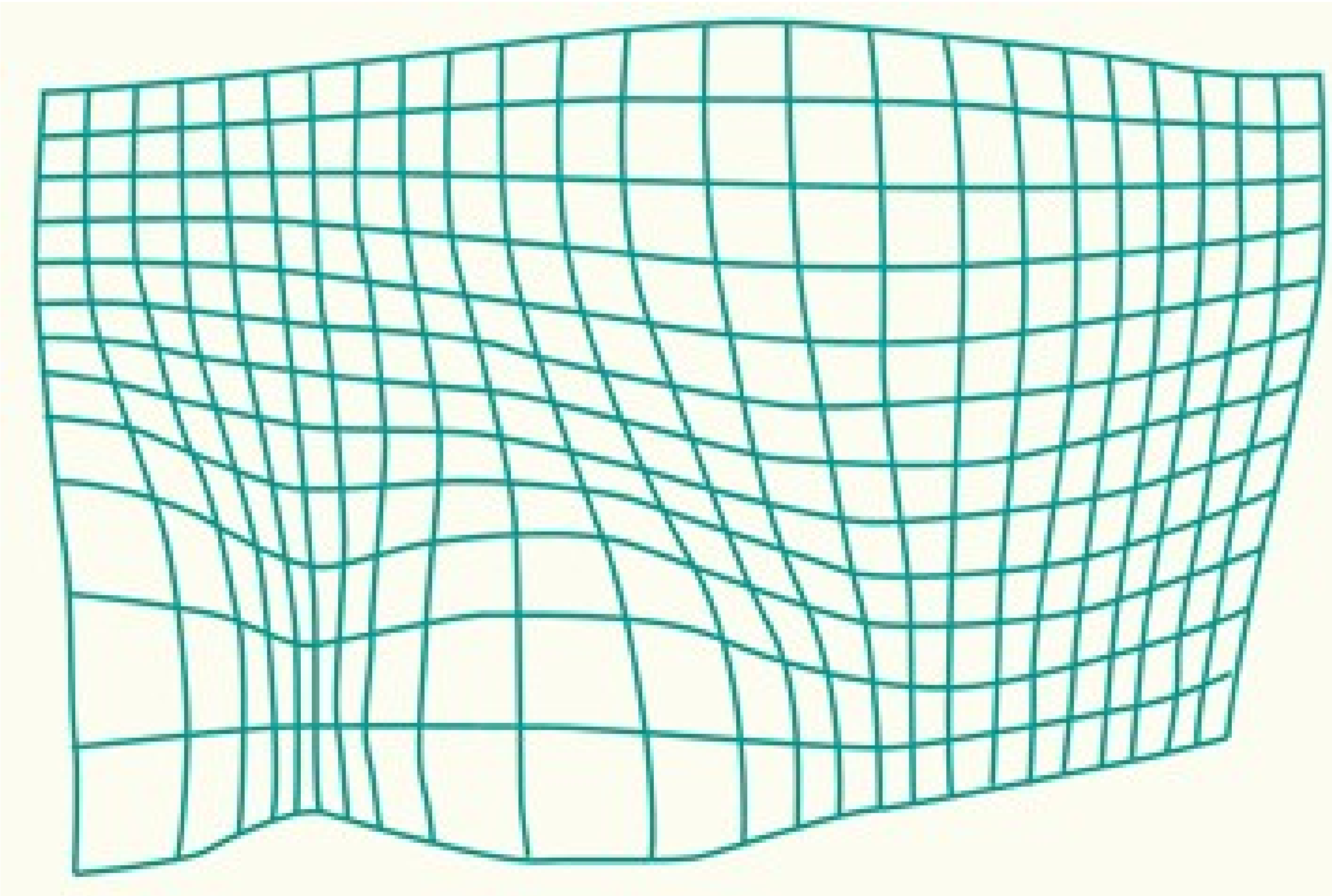
$$Y' = y_o + b_1X + b_2Y + b_3XY + b_4X^2 + b_5Y^2 + b_6X^2Y + b_7XY^2 + b_8X^3 + \ldots\ldots$$

- **Polynomial transformations** are sometimes used to geo-reference uncorrected satellite imagery or aerial photographs or to match vector data layers that don't fit exactly by stretching or rubber sheeting them over the most accurate data layer.

- The figure below shows a grid with no uniform scale distortions.

- It may occur in an aerial photograph, caused by the tilting of the camera and the terrain relief (topography). An approximate correction may be derived through a high-order polynomial transformation.

- The displacements caused by relief differences can be corrected using a Digital Elevation Model (DEM).



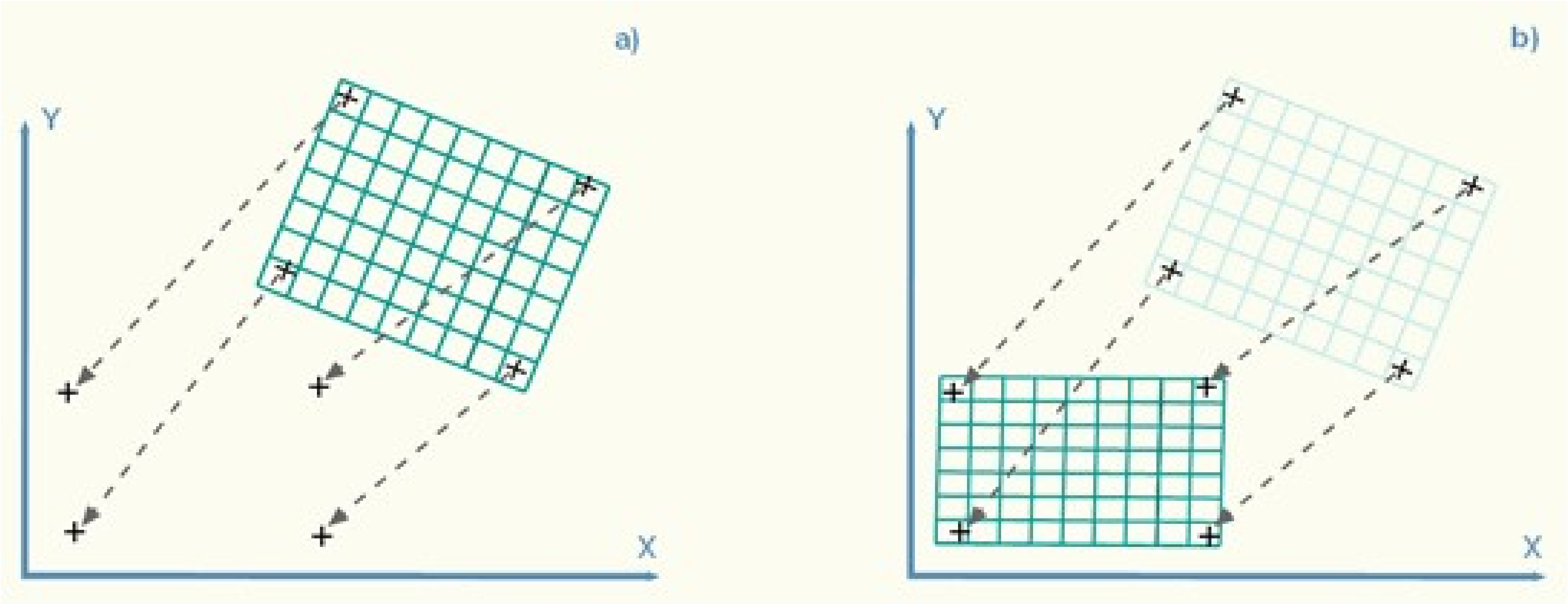
A grid with no uniform scale distortions. An approximate correction may be derived through a high order polynomial transformation

- **2D Cartesian coordinate transformations** are generally used to assign map coordinates (*x,y*) to an uncorrected image or scanned map.

- The type of transformation (usually an affine transformation) depends on the geometric errors in the data set (Fig.a).

- After georeferencing, the image can be aligned (rectified) so that the pixels are exactly positioned within the map coordinate system (figure (b).

- For each image pixel in the new coordinate system, a new pixel value has to be determined by means of an interpolation from surrounding pixels in the old image. This is called ***image resampling***.

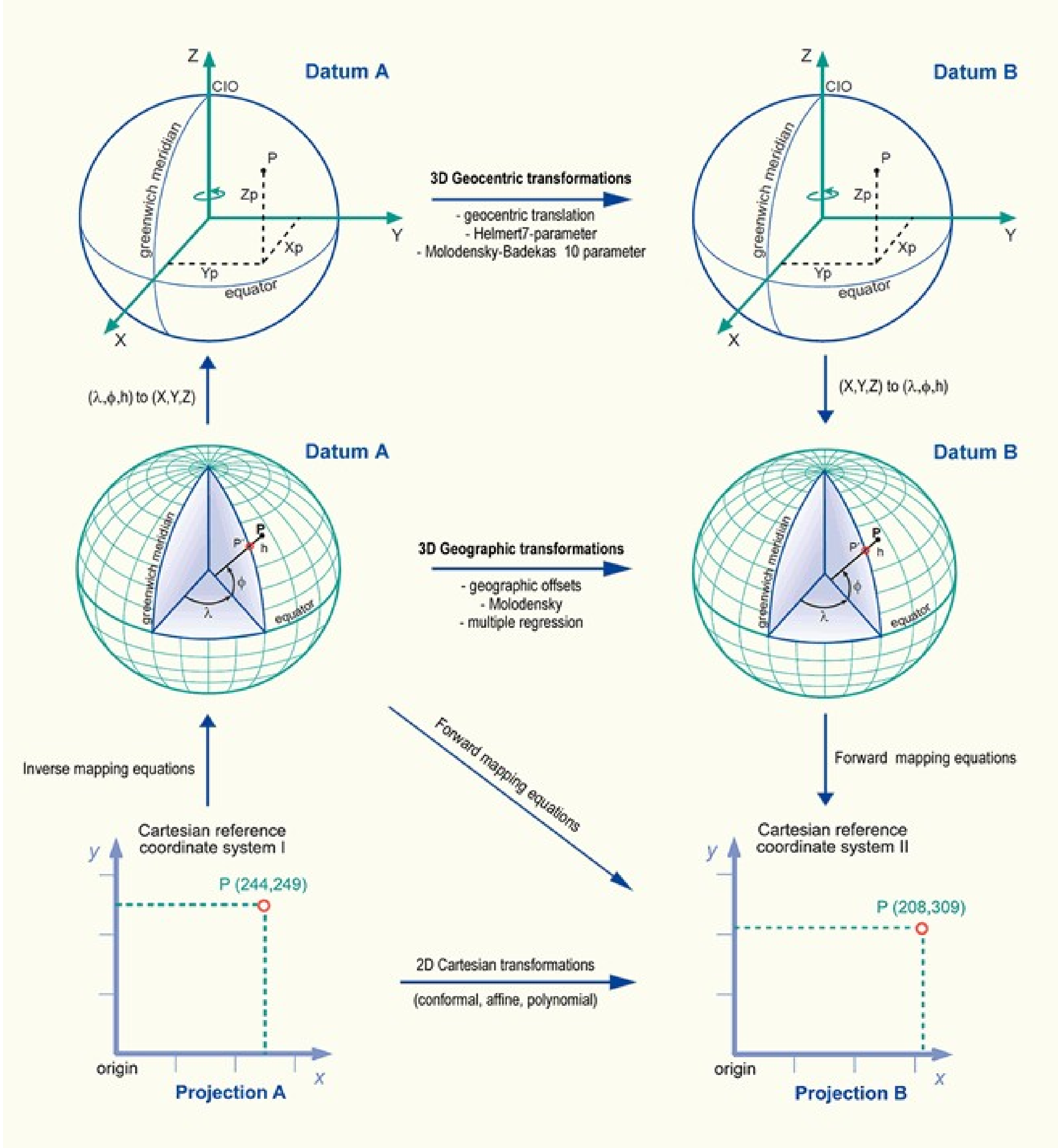


a) *Coordinates are assigned to a raster image by means of a 2D transformation using a set of control points. The image is georeferenced;*

b) *The georeferenced image is rectified to match it with the map coordinate system.*

Summary

- The several types of coordinate transformations: projection change, 3D datum transformations and 2D Cartesian transformations. An overview of these transformations is given in the figure below.
- The illustration shows how the inverse and forward mapping (or projection) equations are used to transform coordinates from one map coordinate system (projection A) to another (projection B).
- The projection change may include a datum transformation from one datum (datum A) to another datum (datum B) in case the source projection (projection A) is based upon a different horizontal datum than the target projection (projection B).
- The datum transformation may take place via a 3D geocentric transformation or directly via a 3D geographic transformation.
- Alternatively, 2D Cartesian transformations may be used to transform coordinates from one map coordinate system to another (e.g. in case the projection of the input map coordinates is unknown).



*Overview of several types of **coordinate** transformations*