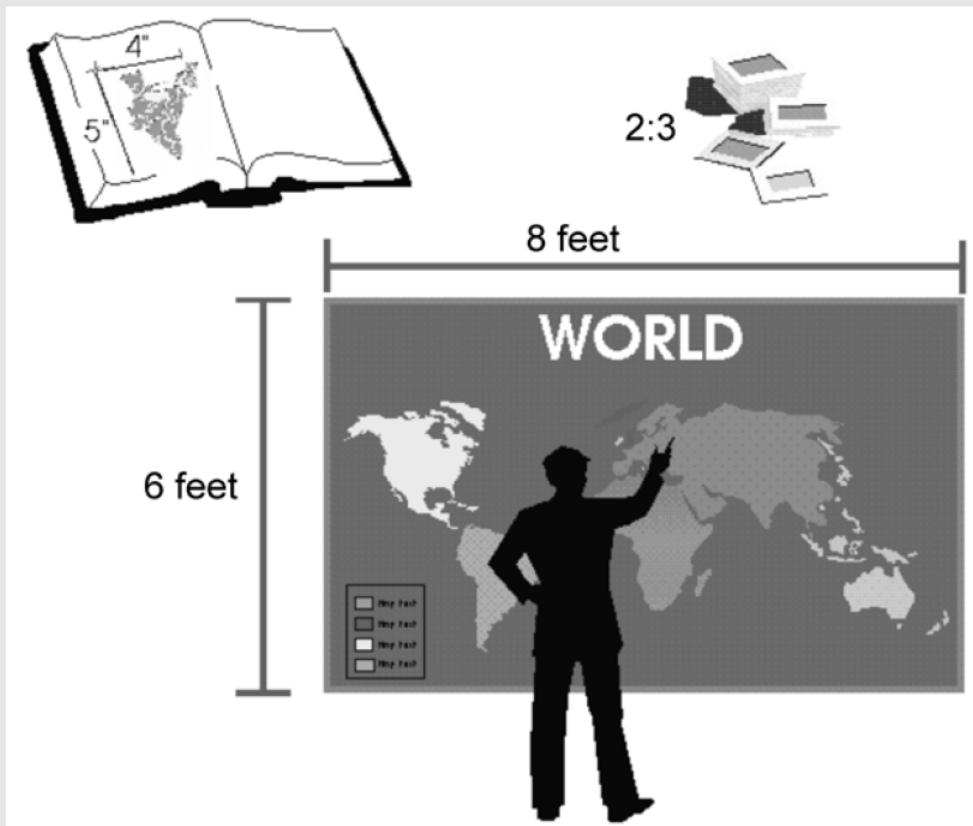
Chapter 3

GISs Roots in Cartography



HUMAN COGNITION OF THE SPATIAL WORLD

- The study of cognition is about knowledge: its acquisition, storage and retrieval, manipulation, and use by humans and other intelligent creatures.
- Cognition includes sensation and perception, thinking, imagery, reasoning and problem-solving, memory, learning, and language.
- Spatial cognition deals with the cognition of spatial properties of the world,
  - including location, size, distance, direction, shape, pattern, movement, and inter-object relations.

SENSING AND PERCEIVING THE WORLD

- Perception is the active acquisition of knowledge about the self and the world through the senses.
- Characteristics of the perceived world:
  - Locational perspective – world perceived from a point-of-view, incomplete access to world
  - Redundancy of information (*e.g.,* depth cues of interposition and linear perspective)
  - Constancy (colour, size, position, shape) – objects, events, and background maintain many characteristics even as viewing conditions change
  - Meaningfulness – tendency to perceive meaningful, familiar objects and events

COGNITIVE MAPS

- Cognitive maps are internal representations of the world and its spatial properties stored in memory (also called ‘mental maps’).
  - Like – what’s out there, what are its attributes, where it is, how to get there.
  - It consists of discrete pieces determined by physical, perceptual, or conceptual boundaries.
    - They are hierarchically organized pieces with multiple levels of pieces differing in status (*e.g.,* size).
- It is not a unitary representation with a constant scale neither completely integrated.
  - It is not like a cartographic map in the head.
  - The cognitive maps have distortions.

SPATIAL LEARNING AND DEVELOPMENT

- Spatial knowledge is learned via one or more *media of acquisition.*
  - Direct sensorimotor experience, maps, models, photos and drawings, movies and videos, verbal and written language.
- Traditional theory of developmental sequence in spatial knowledge of the world inspired by Piagetian theory; consists of 3 stages or elements, acquired over time:
  - First is ‘landmark knowledge’: unique patterns of perceptual events that identify a place.
  - Second is ‘route knowledge’: sensorimotor routines that connect ordered sequences of landmarks; little or no metric spatial knowledge.
  - Third is ‘survey knowledge’: two-dimensional layout knowledge of simultaneous interrelations of locations; allows detouring, shortcutting, and creative navigation.

- Information-processing approach inspires an alternative sequence of continuous and quantitative increase in extent, accuracy, and completeness of sometimes crude metric spatial knowledge.

Navigation

- Navigation is coordinated and goal – directed route following through space.
- It consists of 2 components: locomotion and way-finding.
  - Locomotion is guidance through space in response to local sensorimotor information in immediate surrounds.
    - It finds support surfaces, avoid obstacles and barriers, follow beacons, move through openings.
  - Way-finding is planning and decision – making in response to non – local information, undertaken to reach goal.

USING AND LEARNING MAPS

- The main purpose of cartographic maps is to communicate geographical information and support geographical problem-solving.
  - Maps use convenient scales and viewing perspectives (we can perceive all from a single viewpoint).
  - Maps highlight and clarify relevant properties; omit or downplay irrelevant properties.
  - But projections, generalizations, exaggerations, omissions may mislead or distort knowledge in a map.
  - Similarly, perspective translation from overhead to terrain-level view may be confusing or interpretation of symbols may be difficult or misleading.
- However, training and experience with maps changes the way they are perceived and interpreted.

Spatial Language

- Spatial information often expressed verbally, giving verbal directions, spatial descriptions in stories, road signs, and computer queries.
- Producing spatial language often requires translation of nonverbal spatial knowledge, which can alter the knowledge.
- Language expresses mostly non – quantitative or imprecise quantitative (fuzzy) information about space; connections and general location more important
  - for example, we say ‘turn left at the railway station’, not ‘turn 80° after you go 1.4 kilometres’.
- Quantitative precision usually unnecessary or even confusing for verbal communication but context is critical in interpreting spatial language.

Relevance to GIS

- Cognitive issues touch on all three major functions of GIS: the storage, representation, and analysis of earth-referenced data.
- Some examples of cognitive issues in GIS:
  - How experts and laypeople conceptualize and reason about geographical space, and how GIS can be designed and taught to support both classes of users.
  - How people express spatial information in natural language (such as English), and how this can be used to understand communication with a GIS in natural language (such as a navigation computer inside a car).
  - How interfaces should be designed to promote accurate and efficient communication of spatial and geographic information, such as scale, uncertainty, and network structure.

GIS AND SPATIAL COGNITION

- GIS are tools for supporting human decision-making and tools to help people acquire spatial information, learn about geography.
- The interface between the GIS and the user is a filter which determines how successfully information can be transferred.
  - The effective user interfaces depends on how people learn and reason with spatial information.
- Maps are the main source of data for GIS, the traditions of cartography is fundamentally important to GIS.
- GIS has roots in the analysis of information on maps, and overcomes many of the limitations of manual analysis.

DEFINING A MAP

- Map is a representation, normally to scale and on a flat medium, of a selection of material or abstract features on, or in relation to, the surface of the Earth – by the International Cartographic Association.
- The term ‘map’ is often used in mathematics to convey the notion of transferring information from one form to another,
- The term ‘map’ is used loosely to refer to any visual display of information, particularly if it is abstract, generalized or schematic.

- Cartography is very much a process of abstraction in which features of the real world are generalized or simplified to meet the demands of the theme and audience.
- Production of a map requires:
  - selection of the few features in the real world to include,
  - classification of selected features into groups (*i.e.,* roads, houses, railways),
  - simplification of jagged lines like river meandering,
  - exaggeration of features to be included that are to small to show at the scale of the map and
  - symbolization to represent the different classes of features chosen.
- Maps provide useful ways of displaying information in a meaningful way.

Box 4: What is a map?

Map is a representation, normally to scale and on a flat medium, of a selection of material or abstract features on, or in relation to, the surface of the Earth.

Cartographic Abstraction

Production of a map requires: ***selection*** of the few features in the real world to include, ***classification*** of selected features into groups (*i.e.,* bridges, houses, railways), ***simplification*** of jagged lines like coastlines, ***exaggeration*** of features to be included that are to small to show at the scale of the map and ***symbolization*** to represent the different classes of features chosen.

Role of Maps

Traditionally, maps have four roles today

- Data display-maps provide useful ways of displaying information in a meaningful way.
- Data storage - as a means of storing data.
- Spatial indexes - a map can show *the boundaries* of areas (*e.g.,* land use zones, soil or rock types) and identify each area with a label.
- Data analysis tool - maps are used in analysis to make or test hypotheses and examine the relationship between two distributions using simple transparent overlays.

Changeover to Computer Mapping

Impetus for change began in two communities

1. Scientists wishing to make maps quickly to see the results of modelling, or to display data from large archives already in digital form, *e.g.,* census tables.
2. Cartographers seeking to reduce the cost and time of map production and editing.

GIS and Computer Cartography

Computer cartography has a primary goal of producing maps, systems have advanced tools for map layout, placement of labels, large symbol and font libraries, interfaces for expensive, high quality output devices. However, it is not an analytical tool, therefore, unlike data for GIS, cartographic data does not need to be stored in ways which allow, for example, analysis of relationships between different themes such as population density and housing prices or the routing of flows along connecting highway or river segments.

Contd...

GIS's Advantage over Maps

Data Storage

- Spatial data stored in digital format in a GIS allows for rapid access for traditional as well as innovative purposes.
- The nature of maps creates difficulties when used as sources for digital data.
- Most GIS take no account of differences between datasets derived from maps at different scales.
- Idiosyncrasies (*e.g.*, generalization procedures) in maps become "locked in" to the data derived from them.

Data Indexes

- This function can be performed much better by GIS due to the ability to provide multiple and efficient cross-referencing and searching.

Data Analysis Tool

- GIS is a powerful tool for map analysis.
- Traditional impediments to the accurate and rapid measurement of area or to map overlay no longer exist.

Data Display Tool

Electronic display offers significant advantages over the paper map

- Ability to browse across an area without interruption by map sheet boundaries.
- Ability to zoom and change scale freely.
- Potential for the animation of time dependent data.
- Display in "3 dimensions" (perspective views), with "real-time" rotation of viewing angle.
- Potential for continuous scales of intensity and the use of colour and shading independent of the constraints of the printing process, ability to change colours as required for interpretation.

TYPES OF MAPS

Two types of maps

- *Topographic map:*
  - These maps are a reference tool, showing the outlines of selected natural and man-made features of the Earth, often acts as a frame for other information.
  - 'Topography' refers to the shape of the surface, represented by contours and/or shading, but topographic maps also show roads and other prominent features.
- *Thematic map:*
  - These maps are a tool to communicate geographical concepts such as the distribution of population densities, climate, land use etc.
  - Thematic maps are important in GIS.
  - The boundaries are different for each map as they are determined by the variation of the attribute being mapped.
    - *e.g.,* breaks of soil type occur independently of breaks of vegetation.

- Map type is not just characteristics of the map but can be determined by use.
  - *e.g., can* look at *distribution* of major roads on a general-reference topographic map or *can* find specific *location* of observation units (like district) on a thematic map.
- The classification of maps can be made on the basis of :
  - content of the map (climate, socio-economic...)
  - form of the map (dot, choropleth, animated...)
  - display technology used (electronic, paper,...)
  - production technology used (manual, automated,...)
  - scale of the map (large, medium, small)
  - resolution of the map (country, state,...) etc.

OTHER REPRESENTATIONS OF THE WORLD

- Maps are not the only representation of the world; others include: air photos, satellite imagery, drawings and artwork, verbal description, tables etc.
- The uniqueness of the map among representations of the world lies in content as well as
  - area shown is selective (unlike air photos, satellite imagery, snapshots)
  - maker has control over emphasis, (unlike air photos, satellite imagery, snapshots)
  - emphasis is on spatial relations (unlike drawings and artwork, in which spatial relations support some other message)
  - it is an analogue of what is represented (unlike words, tables, and digital data).

GIS AND COMPUTER CARTOGRAPHY

- Computer cartography
  - It has a primary goal of producing maps.
  - Systems have advanced tools for map layout, placement of labels, large symbol and font libraries, interfaces for expensive, high quality output devices.
  - It is not an analytical tool, therefore, unlike data for GIS, cartographic data does not need to be stored in ways which allow analysis of relationships between different themes.
- GIS
  - In GIS spatial data stored in digital format allows for rapid access for traditional as well as innovative purposes.
  - The nature of maps creates difficulties when used as sources for digital data but most GIS take no account of differences between datasets derived from maps at different scales.

- The prime differences between a GIS and computer cartography are in their functional components:
  - A GIS contains these four components :
    - a. Input b. Database c. Analysis d. Output
  - A mapping (cartographic) system can be described in three components:
    - i. Input ii. Map design iii. Output

Table 3.1: Comparison between computer cartography and GIS.

Computer Cartography	GIS
Feature type	Area (m^2)
Boundary colour	Perimeter (m)
Pattern	Land use/Land cover
Fill colour	Residential
Design level	Average plot area

- This difference is best shown in a software query that lists element attributes.
  - A cartographic query gives information on design features, while a GIS query yields details or parameters about the features themselves, where the data are stored in a GIS database.
  - A query to display where residential areas are or the houses with area of more than 500 square meters
    - It is possible in GIS but is not possible in computer cartography.
- Capabilities of different mapping softwares (Table 3.2)
  - Each may perform better than a GIS at their specialty, but only a GIS has all four components.

Table 3.2: Capabilities of different mapping softwares.

Mapping Software	Capabilities			
	Means of data input	*Database management system (DBMS)*	*Analysis capability*	*Graphics output*
CAD Computer Assiting Drawing *e.g.* AutoCad, Microstation	√	X	X	√
GDS Graphic Design System *e.g.* CorelDraw, Illustrator	√	X	X	√
DBMS Database Management System *e.g.* Oracle, Sybase	√	√	X	X
DIPS Digital Image Processing System *e.g.* PCI, Erdas	√	√	√	√
GIS Geographical Information System *e.g.* ArcGis, Ilwis, MapInfo	√	√	√	√

Table 3.3: Major advantages and disadvantages of computer cartography.

Advantages	Disadvantages
Lower cost for simple maps, faster production	Relatively few full - scale systems have been shown to be truly cost - effective in practice.
Greater flexibility in output easy scale or projection change maps can be tailored to user needs	Computer methods do not ensure production of maps of high quality. There is a perceived loss of regard for the 'cartographic tradition' with the consequent production of 'cartojunk'.
Other uses for digital data	High capital cost, though this is now much reduced.

MAPPING CONCEPTS, FEATURES & PROPERTIES

- A map represents geographic features or other spatial phenomena by graphically conveying information about locations and attributes.
  - Locational information describes the position of particular geographic features on the Earth's surface, as well as the spatial relationship between features.
  - Attribute information describes characteristics of the geographic features represented.
- The basic objective of mapping is to provide
  - Descriptions of geographic phenomenon
  - Spatial and non spatial information
  - Map features like Point, Line, & Polygon

Map Features:

- Locational information is usually represented by points, lines and areas.
  - Points for features such as wells and schools
  - Lines for features such as streams, roads and contour lines
  - Areas for features such as lakes, cultivated lands and census tracts.

Scale in Digital Maps:

- In digital mapping, the term scale is used to indicate the scale of the materials from which the map was made.
  - For example, if a digital map is said to have a scale of 1:100,000, it was made from a 1:100,000-scale paper map.
  - Digital maps retain the same accuracy and characteristics as their source maps.
- Data collected at a specific scale are suitable for mapping and analysis only at similar scales
  - At smaller scales, large scale data are too complex (but could be generalised).
  - At larger scales, small scale data are too generalized (detail cannot be ‘added’).

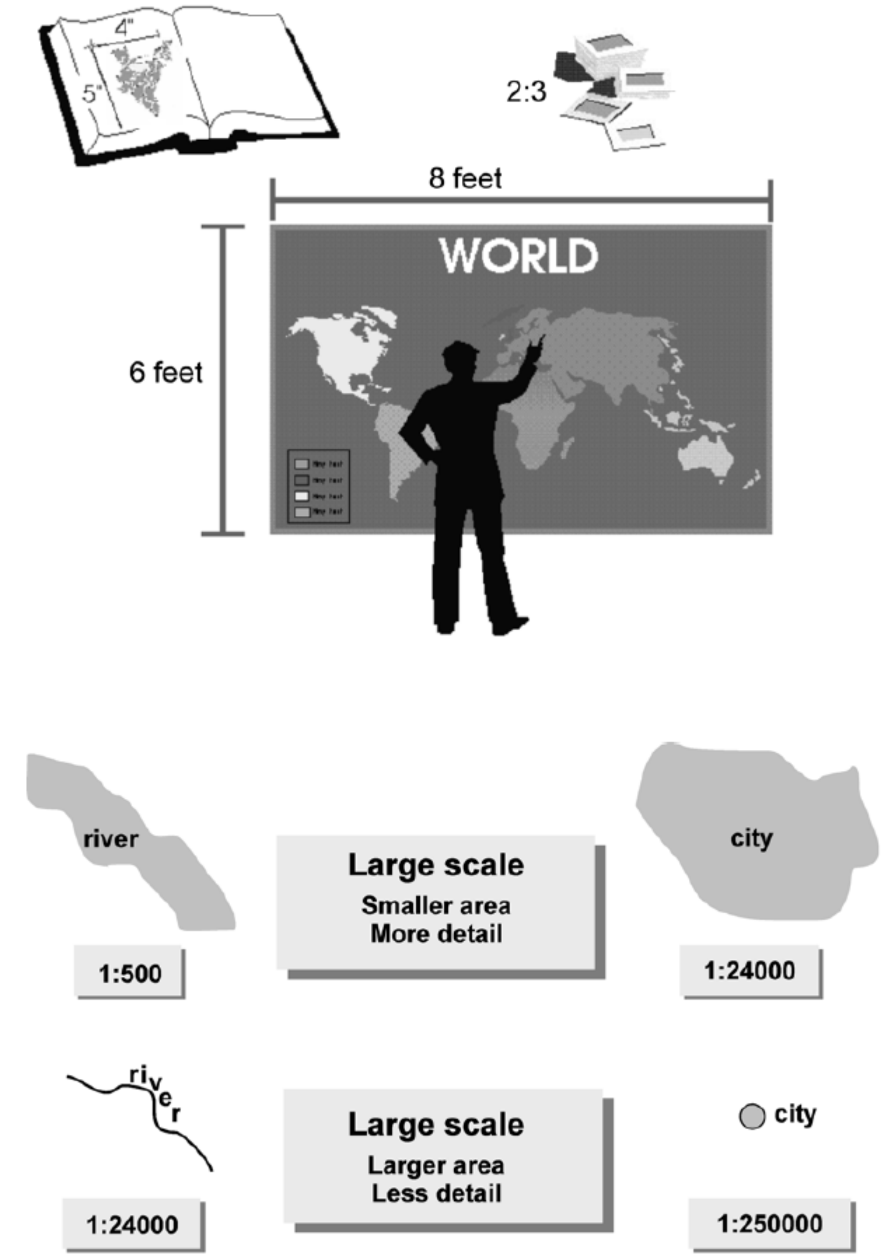


Figure 3.1: Map scale determines the size and shape of features.

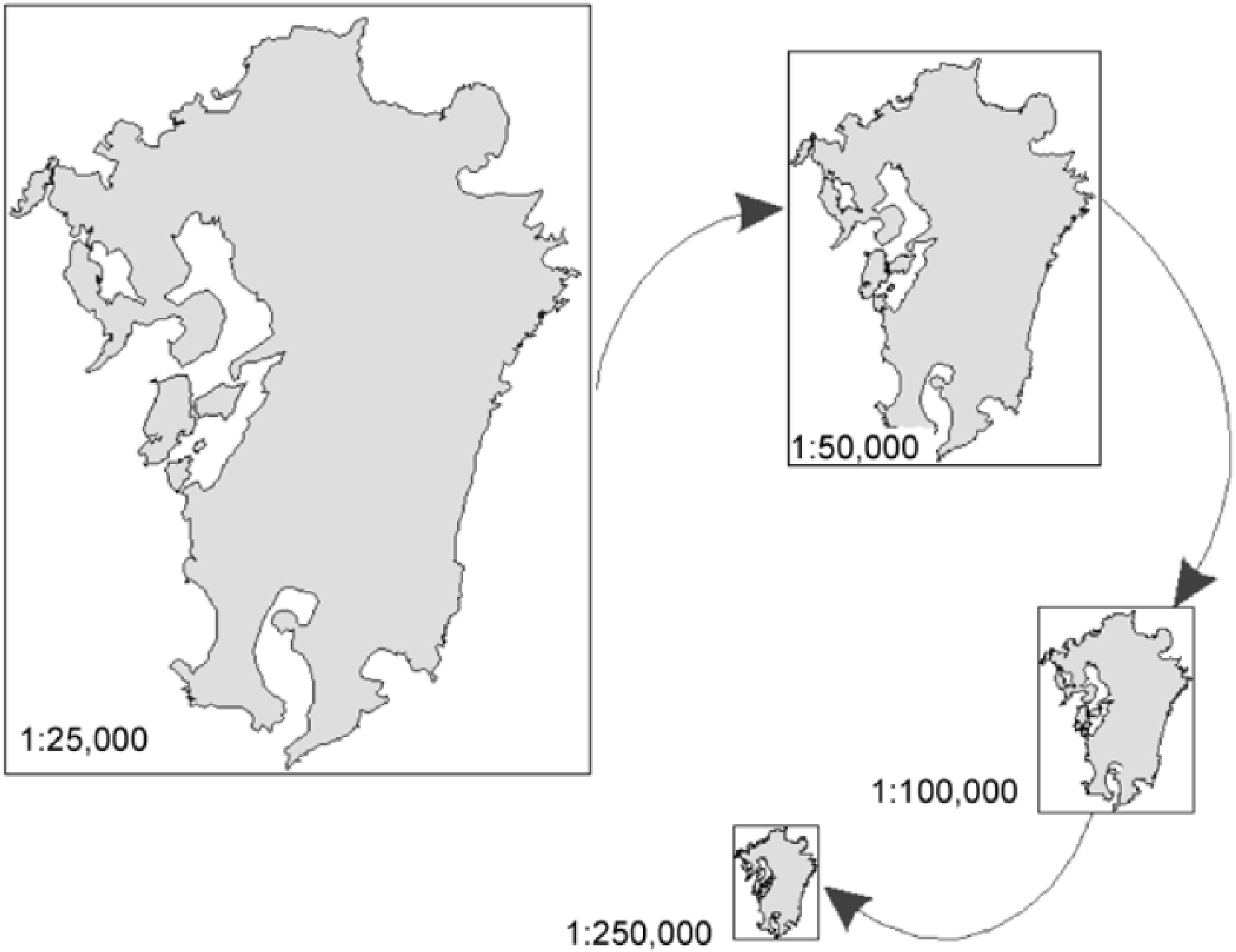


Figure 3.2: The details are blurred as the scale decreases.

Map Resolution:

- Map resolution refers to how accurately the location and shape of map features can be depicted for a given map scale.
- Scale affects resolution.
  - In a larger-scale map, the resolution of features more closely matches real-world features because the extent of reduction from ground to map is less.
  - As map scale decrease, the map resolution diminishes because features must be smoothed and simplified, or not shown at all.

Map Accuracy:

- ***Absolute accuracy***
  - Absolute accuracy of a map refers to the relationship between a geographic position on a map and its real-world position measured on the surface of the earth.
  - Absolute accuracy is primarily important for complex data requirements such as those for surveying and engineering-based applications.
- ***Relative accuracy***
  - Relative accuracy refers to the displacement between two points on a map (both distance and angle), compared to the displacement of those same points in the real world.
  - Users with simple data requirements generally need only relative accuracy.

- ***Attribute accuracy***
  - Attribute accuracy refers to the precision of the attribute database linked to the map's features.
  - Attribute accuracy is most important to users with complex data requirements.
- ***Currency***
  - A map's Currency refers to how up-to-date it is.
  - Currency is usually expressed in terms of a revision date, but this information is not always easy to find.
- ***Complete***
  - A map is Complete, if it includes all the features a user would expect it to contain.
  - Completeness and currency usually are related because a map becomes less complete as it gets older.

- The most important issue to remember about map accuracy is that the more accurate the map, the more it costs in time and money to develop.
- It is important to consider map accuracy to ensure that our data is not used inappropriately.
- Any number of factors can cause error. Note these sources can have at cumulative effect.

$$E = f(f)+f(l)+f(c)+f(d)+f(a)+f(m)+f(p)+f(rms)+f(mp)+u$$

Where,

f = flattening the round Earth onto a two-dimensional surface (transformation from spherical to planar geometry)

l = accurately measuring location on Earth (correct project and datum information)

c = cartographic interpretation (correct interpretation of features)

d = drafting error (accuracy in tracing of features and width of drafting pen)

a = analog to digital conversion (digitizing board calibration)

m = media stability (warping and stretching, folding. Wrinkling of map)

p = digitizing processor error (accuracy of cursor placement)

rms = Root Mean Square (registration accuracy of ties)

mp = machine precision (coordinate rounding by computer in storing and transforming)

u = additional unexplained source error.

Map Extent:

- The aerial extent of map is the area on the Earth’s surface represented on the map.
- It is the limit of the area covered, usually defined by rectangle just large enough to include all mapped features.
- The size of the study area depends on the map scale.

Database Extent:

- The aerial extent of a database is the limit of the area of interest for the GIS project.
- This usually includes the areas directly affected by the organization’s responsibility (such as assigned administrative units) as well as surrounding areas that either influence or are influenced by relevant activities in the administrative area.

Data Automation:

- Map data is collected, automated and updated as series of adjacent map sheets or aerial photograph.
- To combine these smaller sheets into larger units or study areas, the co-ordinates of coverage must be transformed into a single common co-ordinate system.
- Once in a common coordinate system, attributes are associated with features.
- Then as needed map sheets for layer are edge matched and joined into a single coverage for our study area.

TYPES OF INFORMATION IN A DIGITAL MAP

- Three general types of information can be included in digital maps
  - Geographic information, which provides the position and shapes of specific geographic features.
  - Attribute information, which provides additional non-graphic information about each feature.
  - Display information, which describes how the features will appear on the screen.

THE SHAPE OF THE EARTH

- Eratosthenes, a Greek geographer, gave the notion of spherical earth in second century B.C.
- But now researchers have confirmed that earth’s surface is not spherical or flat rather it is oblate ellipsoidal.
  - The radius to the poles is slightly less to equator (approximately 21 kilometres lesser).
  - The flattening of the ellipse for the earth is only 1/297, but it necessary to take care in calculations for plotting accurate maps on large scales.

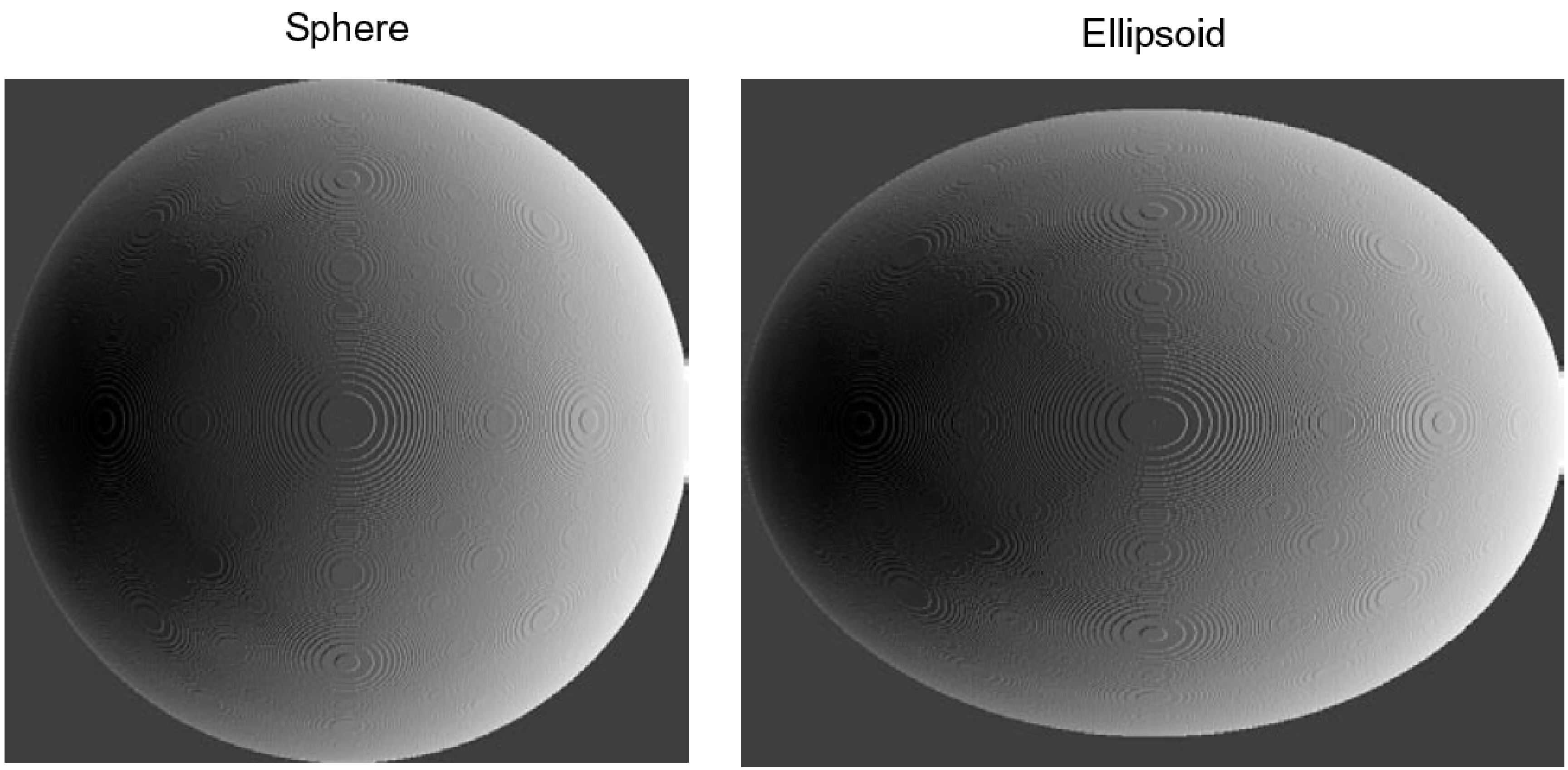


Figure 3.3: Earth shape: sphere or ellipsoid.

DATUMS

- Datums define the reference systems that describe the size and shape of the earth, and the origin and orientation of the coordinate systems used to map the earth.

Datum Types

- 1. *Horizontal:*
  - Datums that define the relationship between the physical earth and horizontal coordinates such as latitude and longitude.
  - Examples include the North American Datum of 1927 (NAD27) and the European Datum 1950 (ED50).
- 2. *Vertical:*
  - Datums that define level surfaces.
  - Examples include the National Geodetic Vertical Datum of 1929 (NGVD29) and the North American Vertical Datum of 1988 (NAVD88).

  - Some are based on sea-level measurements and levelling networks (NGVD29), others on gravity measurements (NAVD88).
- 3. *Complete:*
  - Datums that describe both vertical and horizontal systems.
  - Some, such as World Geodetic System 1984 (WGS-84), also describe other parameters such as the rotation rate of the earth and various physical constants such as the angular velocity of the earth and the earth’s gravitational constant.

Reference Ellipsoids

- Reference ellipsoids are defined by either *semi-major* (equatorial radius) and *semi-minor* (polar radius) axes, or the relationship between the semi-major axis and the flattening of the ellipsoid (expressed as its *eccentricity).*
- Many reference ellipsoids are in use by different nations and agencies.
- Reference ellipsoids are identified by a name and often by a year for example, the Clarke 1866 ellipsoid is different from the Clarke 1858 and the Clarke 1880 ellipsoids.

Table 3.4: Selected reference ellipsoids.

Ellipse	Semi-major axis	Flattening
Airy 1830	6377563.396	299.3249646
Bessel 1841	6377397.155	299.1528128
Clarke 1866	6378206.4	294.9786982
Clarke 1880	6378249.145	293.465
Everest 1830	6377276.345	300.8017
Fischer 1960 (Mercury)	6378166	298.3
Fischer 1968	6378150	298.3
G R S 1967	6378160	298.247167427
G R S 1975	6378140	298.257
G R S 1980	6378137	298.257222101
International	6378388	297.0
Krassovsky 1940	6378245	298.3
WGS 60	6378165	298.3
WGS 66	6378145	298.25
WGS 72	6378135	298.26
WGS 84	6378137	298.257223563

Geodetic Datums

- Topographic and sea-level models attempt to model the physical variations of the surface:
  - The *topographic surface* of the earth is the actual surface of the land and sea at some moment in time.
  - *Sea level* can be thought of as the average surface of the oceans, though its true definition is far more complex.
- *Gravity models* and ***geoids*** are used to represent local variations in gravity that change the local definition of a level surface.
  - *Gravity models* attempt to describe in detail the variations in the gravity field. The importance of this effort is related to the idea of *levelling.*
  - *Geoid models* attempt to represent the surface of the entire earth over both land and ocean as though the surface resulted from gravity alone.

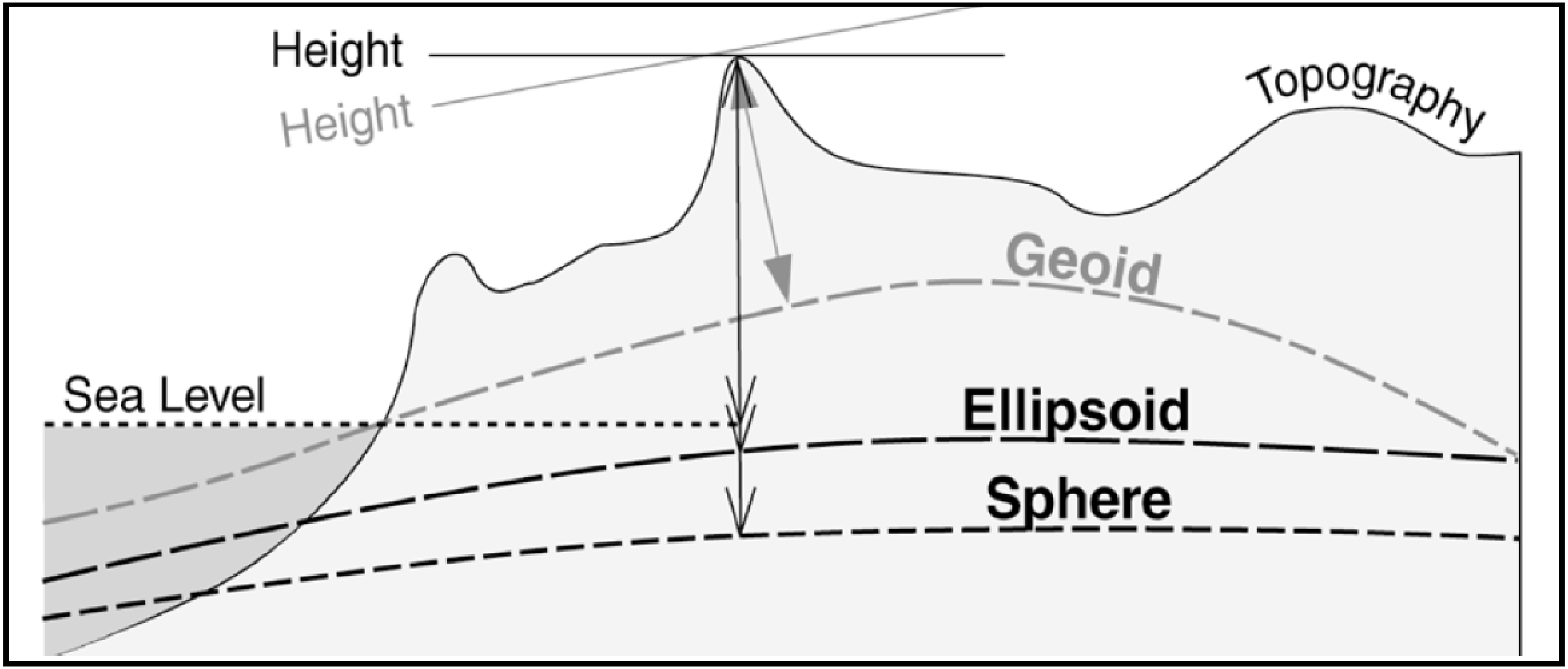


Figure 3.4: Elevations defined with reference to a sphere, ellipsoid, geoid, or local sea level will all be different. Even location as latitude and longitude will vary somewhat.

- Geodetic datums define reference systems that describe the size and shape of the earth based on these various models.
- They are the central concern of the science of ***geodesy.***
- Datums have evolved from those describing a spherical earth to ellipsoidal models derived from years of satellite measurements.
- Modern geodetic datums range from *flat-earth models,* used for plane surveying to *complex* models, used for international applications, which completely describe the size, shape, orientation, gravity field, and angular velocity of the earth.
- Different nations and international agencies use different datums as the basis for coordinate systems.
- Reference system can be divided into two groups:
  - *Global systems* can refer to positions over much of the Earth.
  - *Regional systems* have been defined for many specific areas, often covering national, state, or provincial areas.

GENERAL COORDINATE SYSTEMS

Plane coordinate system – Cartesian coordinates

- Cartesian coordinates are determined by locating an origin there after setting two axes through origin in fixed directions, at right angles to each other.
  - By convention these are usually identified as x and y, where x is horizontal and y vertical (x is east, y is north).

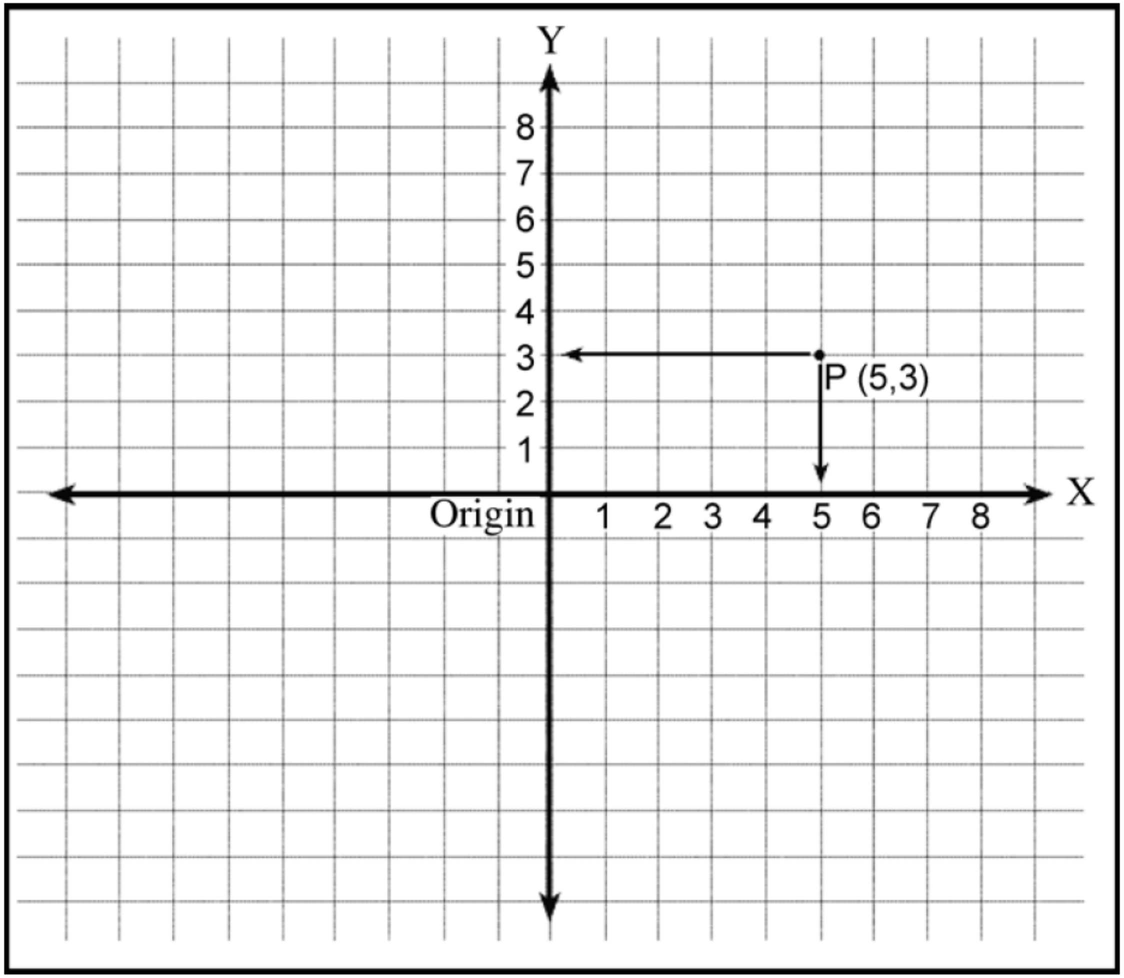


Figure 3.5: The Cartesian coordinate system.

Storing coordinates

- In a GIS, coordinates must be stored in the computer as numbers, there are two important concepts that need to be considered:
- 1. *Integer vs. real numbers*
    - Integers are whole numbers, optionally preceded by ‘ – ’ to indicate negation and are discrete.
    - Real numbers can be expressed as decimal numbers and are continuous.
- 2. *Computer precision*
    - In the computer, the number of digits which can be stored for each value is limited by the hardware.
    - Integers are normally stored using 16 bits of memory and can have a range from – 32767 to +32767.
    - Floating point numbers can use single or double precision. Single precision commonly allocates 32 bits, or 4 bytes. Double precision commonly allocates 64 bits or 8 bytes.

Precision of Cartesian Coordinates

- The number of significant digits required for a specific project when using Cartesian coordinates depends on two measures:
  - Size of the study area
  - Resolution (accuracy) of measurement

Plane coordinate system – Polar coordinates

- Polar coordinates use distance from origin (r) and angle from fixed direction (q), usually fixed direction is north and angle is measured clockwise from it.
- Polar coordinates are useful for measuring from some fixed point such as the center of the city or when using data from sources such as ground surveys and radar.
- To translate from (r, q) to (x, y)

  $x = r \sin(q)$ $y = r \cos(q)$

  $r = (x^2 + y^2)$ $q = \arctan(x/y)$

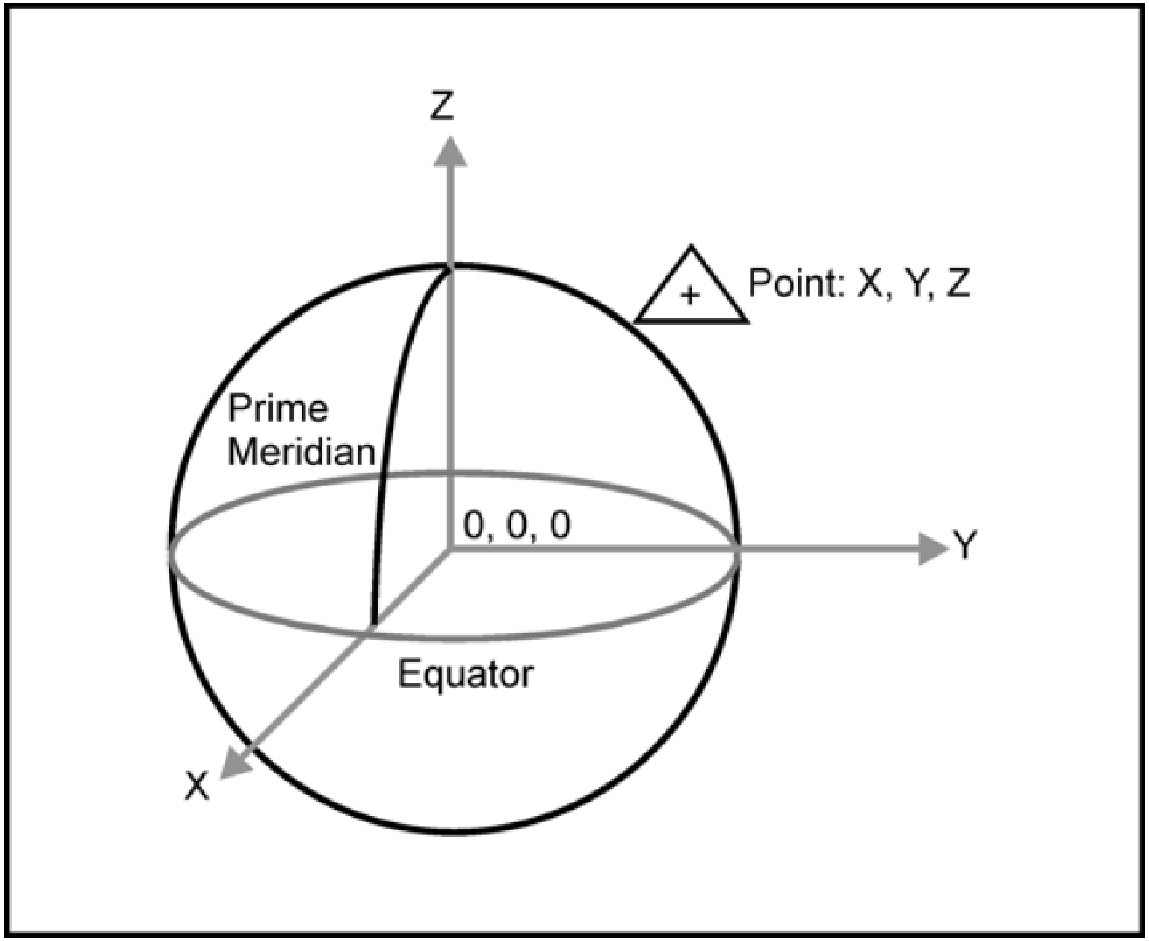


Figure 3.6: Earth Centered, Earth Fixed (ECEF) Cartesian coordinates can also be used to define three dimensional positions.

EARTH COORDINATE GEOMETRY

Rotation of the Earth:

- The spinning of the earth on its imaginary axis is called *rotation.*
- The North and South poles represent the axis of spin and are fixed reference points.
  - If the North Pole was extended, it would point to a fixed star, the North Star (Polaris).
- Any point on the earth's surface moves with the rotation and traces imaginary curved lines are *Parallel of Latitude.*

The Equator:

- If a plane bisected the earth midway between the axis of rotation and perpendicular to it, the intersection with the surface would form a circle. This unique circle is the *equator.*
- The equator is a fundamental reference line for measuring the position of points around the globe.
- The equator and the poles are the most important parts of the earth's coordinate system.

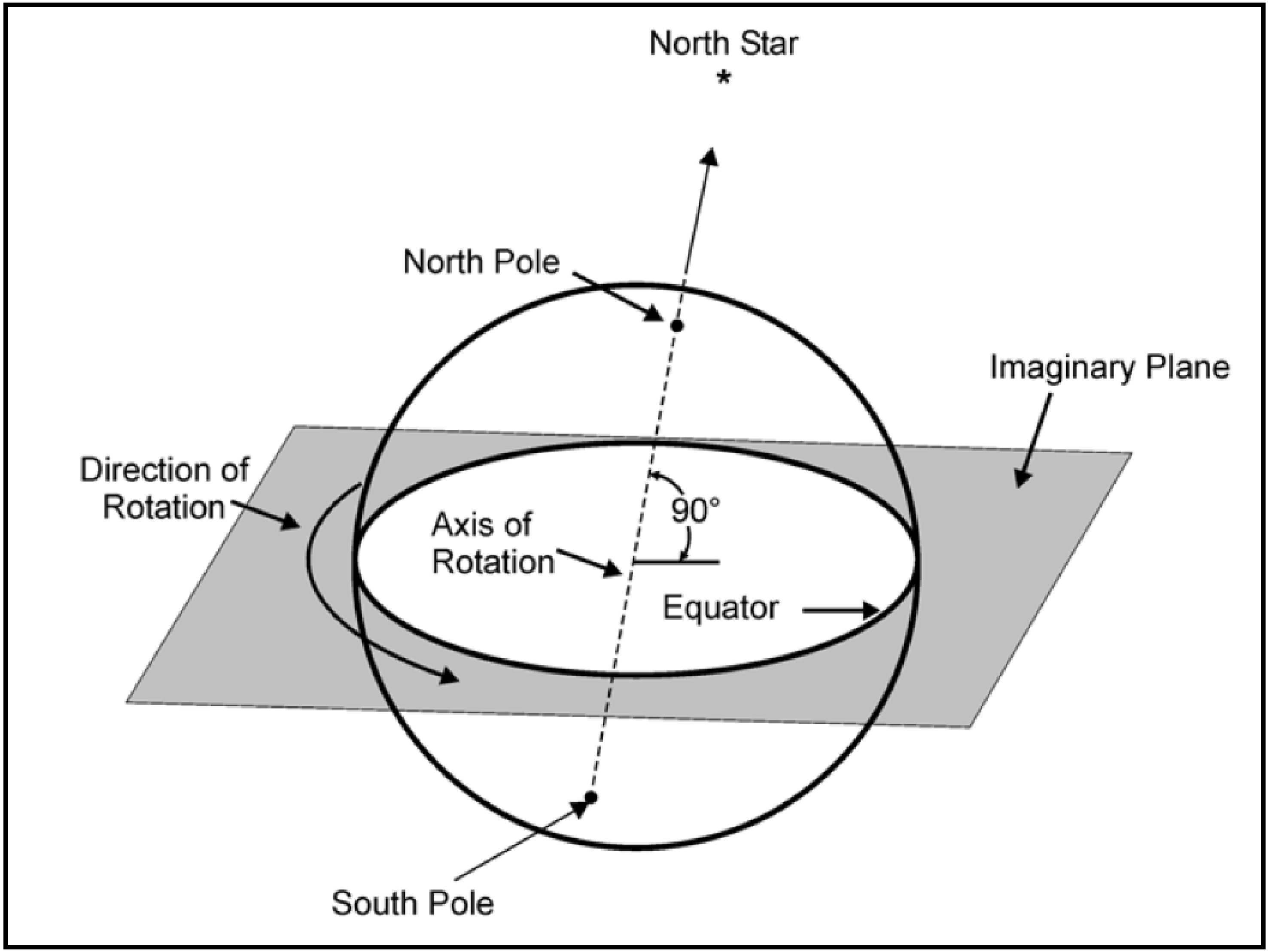


Figure 3.7: Location of the equator, north and south poles, and the imaginary axis of rotation.

The Geographic Grid:

- The spherical coordinate system with latitudes and longitudes used for determining the locations of surface features.
  - Parallels: east-west lines parallel to the equator.
  - Meridians: north-south lines connecting the poles.
  - Parallels are constantly parallel, and meridians converge at the poles.
  - Meridians and parallels always intersect at right angles.

Parallels of Latitude:

- *Parallels of latitude* are all small circles, except for the equator.
- They are true east-west lines, always parallel, any two are always equal distances apart and an infinite number can be created.
- Parallels are related to the horizontal x-axes of the Cartesian coordinate system.

Meridians of Longitude:

- *Meridians of longitude* are halves of great circles, connecting one pole to the other.
- All run in a true north-south direction, spaced farthest apart at the equator and converge to a point at the poles, an infinite number can be created on a globe.
- Meridians are similar to the vertical y-axes of the Cartesian coordinate system.
- The Point P has a latitude of 30 degrees North and a longitude of 20 degrees West. (Figure 3.8)

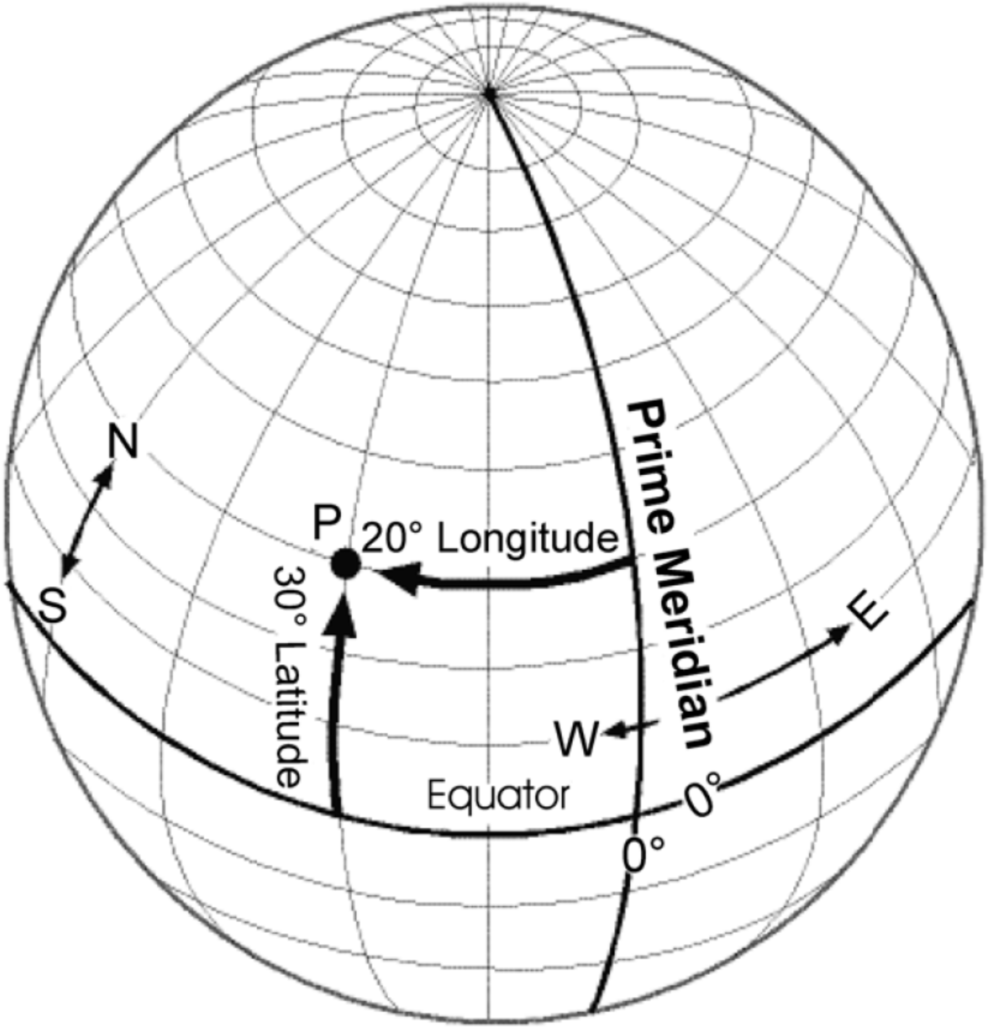


Figure 3.8: The Geographic grid.

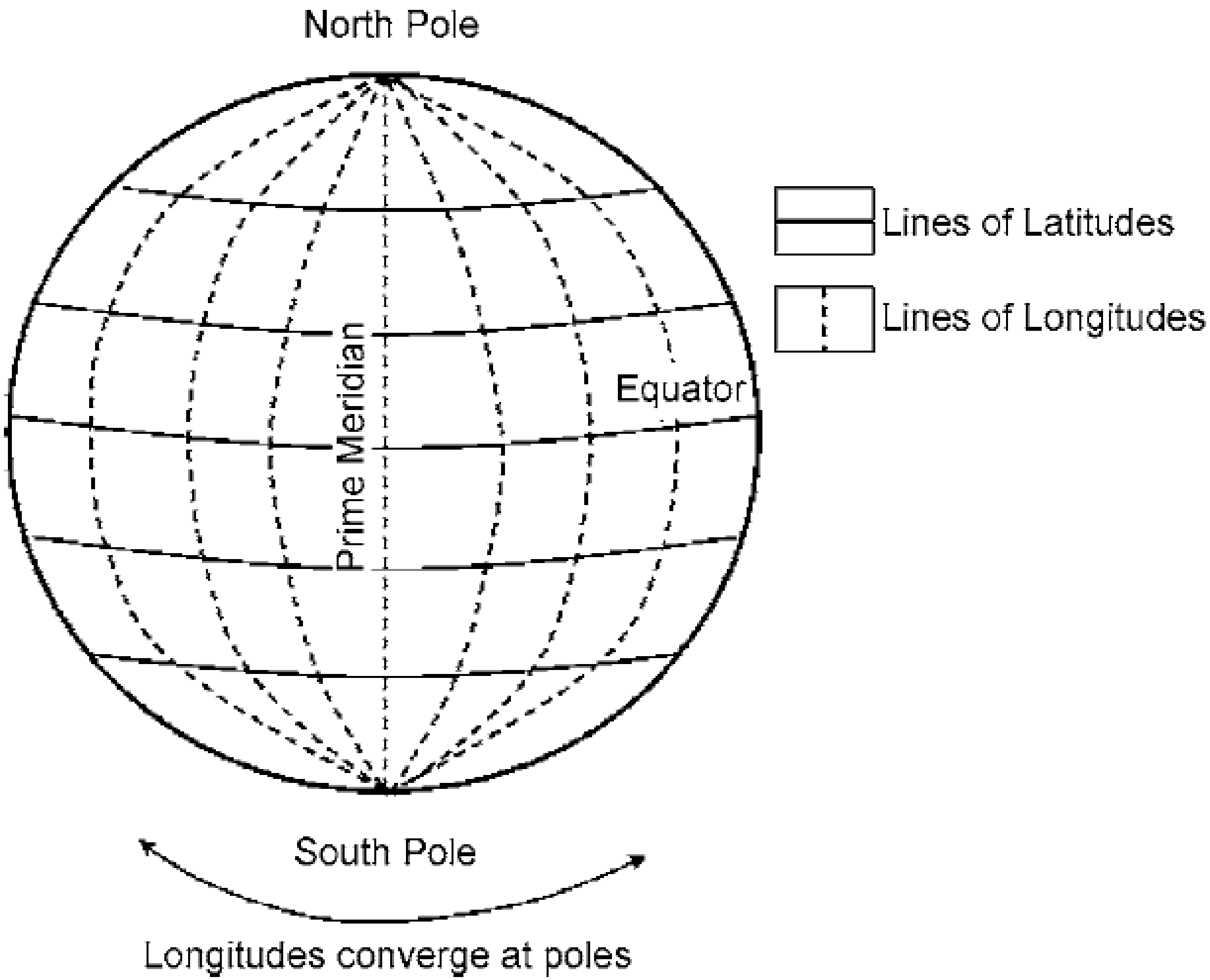


Figure 3.9: Parallels of latitude and Meridians of longitude.

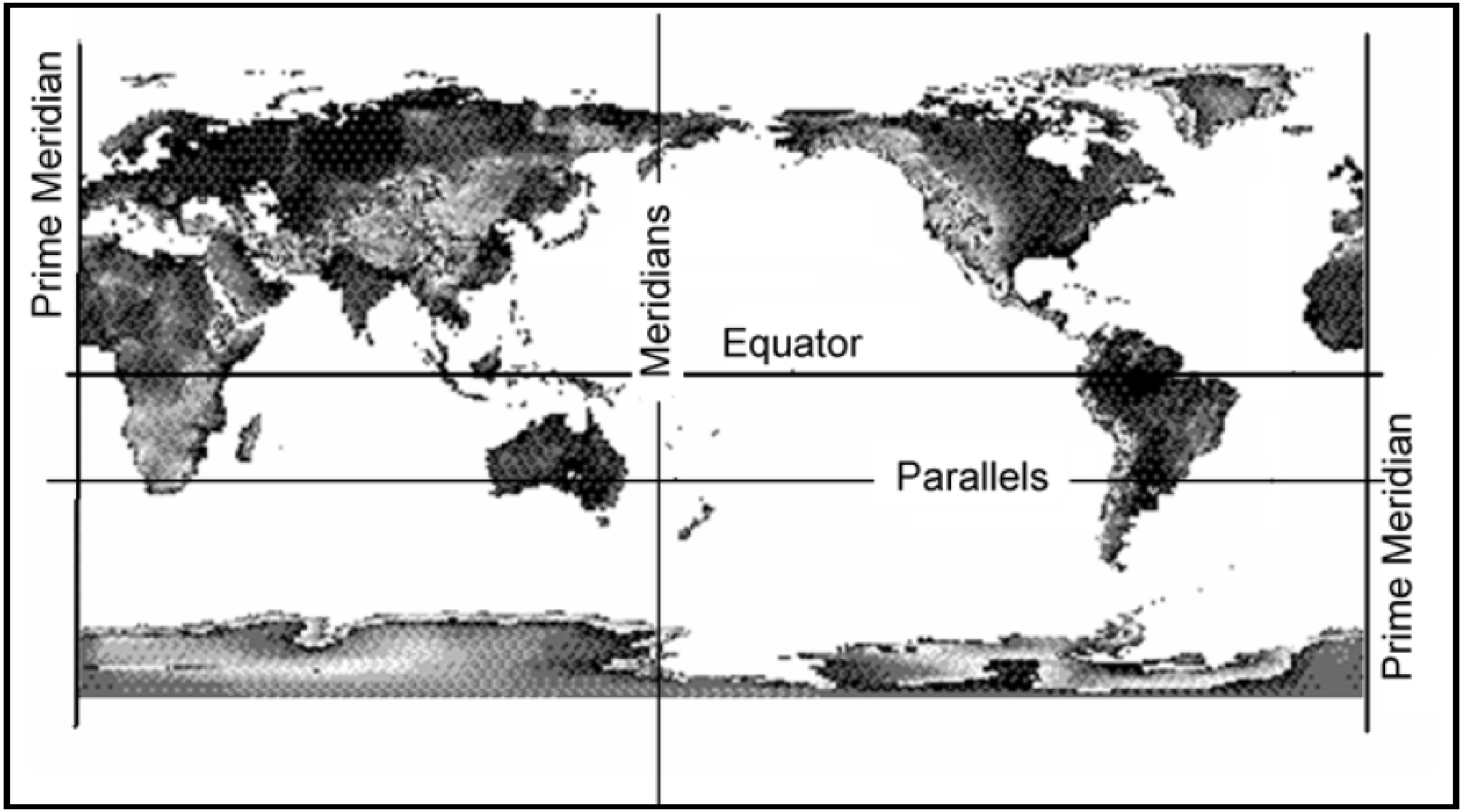


Figure 3.10: Geographic coordinates.

Degrees, Minutes, and Seconds:

- Angular measurement is used in addition to simple plane geometry to specify location on the earth's surface.
- This is based on a *sexagesimal scale*: A circle has 360 degrees, 60 minutes per degree, and 60 seconds per minute.
  - For example, 45° 33′ 22″ (45 degrees, 33 minutes, 22 seconds).
- It is often necessary to convert this conventional angular measurement into decimal degrees.
  - For example, 45° 33′ 22″ is converted into 45.55°
    (33′ 22″ is 2,002 seconds. the ratio: 2,002/3,600 = 0.55 °)

Great and Small Circles:

- A *great circle* is a circle formed by passing a plane though the exact center of a sphere.
  - It is the largest circle that can be drawn on a sphere's surface.
  - An infinite number of great circles can be drawn on a sphere.
  - Great circles are used in the calculation of distance between two points on a sphere.
- A *small circle* is produced by passing a plane through any part of the sphere other than the center.

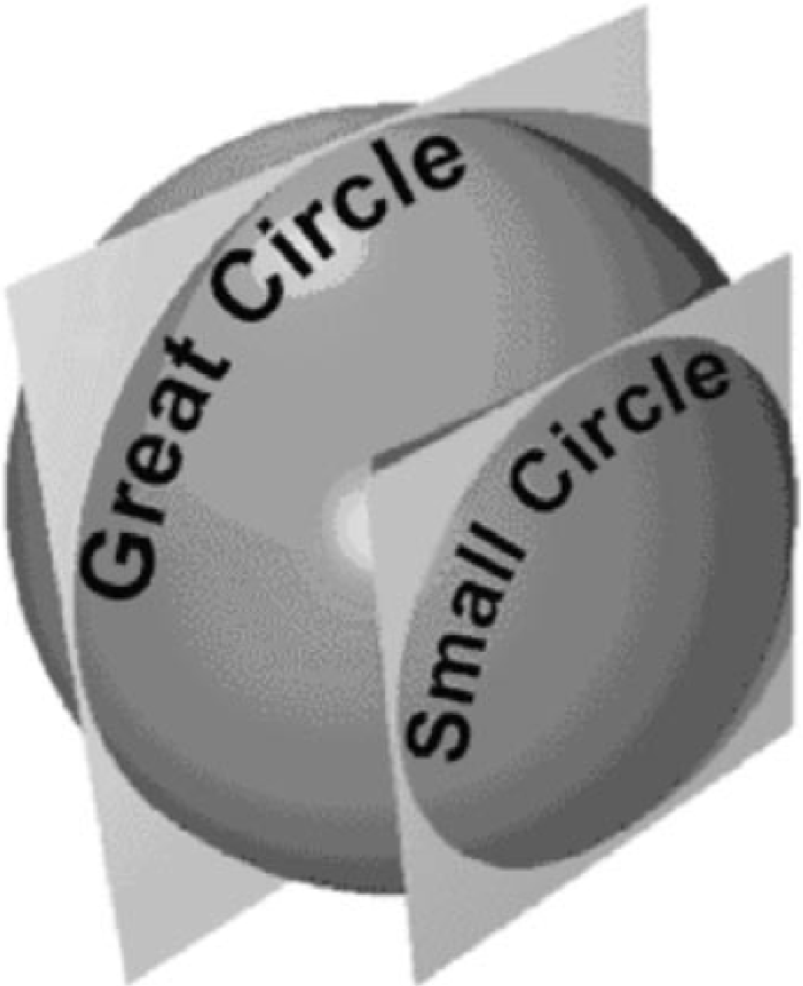


Figure 3.11: Great and small circles.

LATITUDE AND LONGITUDE AND LOCATIONS

Latitude

- Authalic Latitude
  - It is based on a spherical earth.
  - It measures the position of a point on the earth's surface in terms of the angular distance between the equator and the poles.
    - A point in the northern hemisphere 28 degrees north of the equator is labelled Lat. 28° N.
  - The north or south measurement of latitude is actually measured along the meridian which passes through that location. It is known as an *arc of the* meridian.

- Geodetic Latitude
  - It is based on an ellipsoidal earth.
  - The ellipsoid is a more accurate representation of the earth than a sphere since it accounts for polar flattening.
  - In the WGS 84 Ellipsoid, the length of 1° of latitude is not the same everywhere as it is on the sphere.
    - At the equator, 1° of latitude is 110.57 kilometers (68.7 miles).
    - At the poles, 1° of latitude is 111.69 kilometers (69.4 miles).

Latitude and Distance

- Parallels of latitude decrease in length with increasing latitude.
- Length of parallel at latitude x = (cosine of x) * (length of equator).
  - For example, the length of the parallel at latitude 60° is one half the length of the equator.
- Since, the variation in lengths of degrees of latitude varies by only 1.13 kilometers (0.7 mile), the standard figure of 111.325 kilometers (69.172 miles) can be used.
  - For example, anywhere on the earth, the length represented by 3° of latitude is (3×111.325) = 333.975 kilometers.

Longitude

- Longitude measures the position of a point on the earth’s surface east or west from a specific meridian, the *prime meridian.*
- The most widely accepted prime meridian is based on the *Bureau International de l’Heure (BIH) Zero Meridian.* It passes through the old Royal Observatory in Greenwich, England.
  - The prime meridian has the angular designation of 0° longitude.
- The length of a degree of longitude along a meridian is not constant because of polar flattening.
  - At the equator, the approximate length is determined by dividing the earth’s circumference (24,900 miles) by 360 degrees *i.e.,* 111.05 kilometers (69 miles).
  - The meridians converge at the poles, and the distance represented by one degree decreases.

Longitude and Distance

- Measurement along meridians of longitude accounts for the earth's polar flattening and degree lengths along meridians are not constant.
  - For example, 111.325 kilometers (69.172 miles) per degree at the equator, while 16.85 kilometers (10.47 miles) per degree at 80° North and 0 kilometers at the poles.
- The distance between meridians of longitude on a sphere is a function of latitude. The *Mathematical expression* is:

  Length of a degree of longitude = cos (latitude) × 111.325 kilometers.
  - For example, 1° of longitude at 40° N = cos (40°) × 111.325 = 85.28 km.

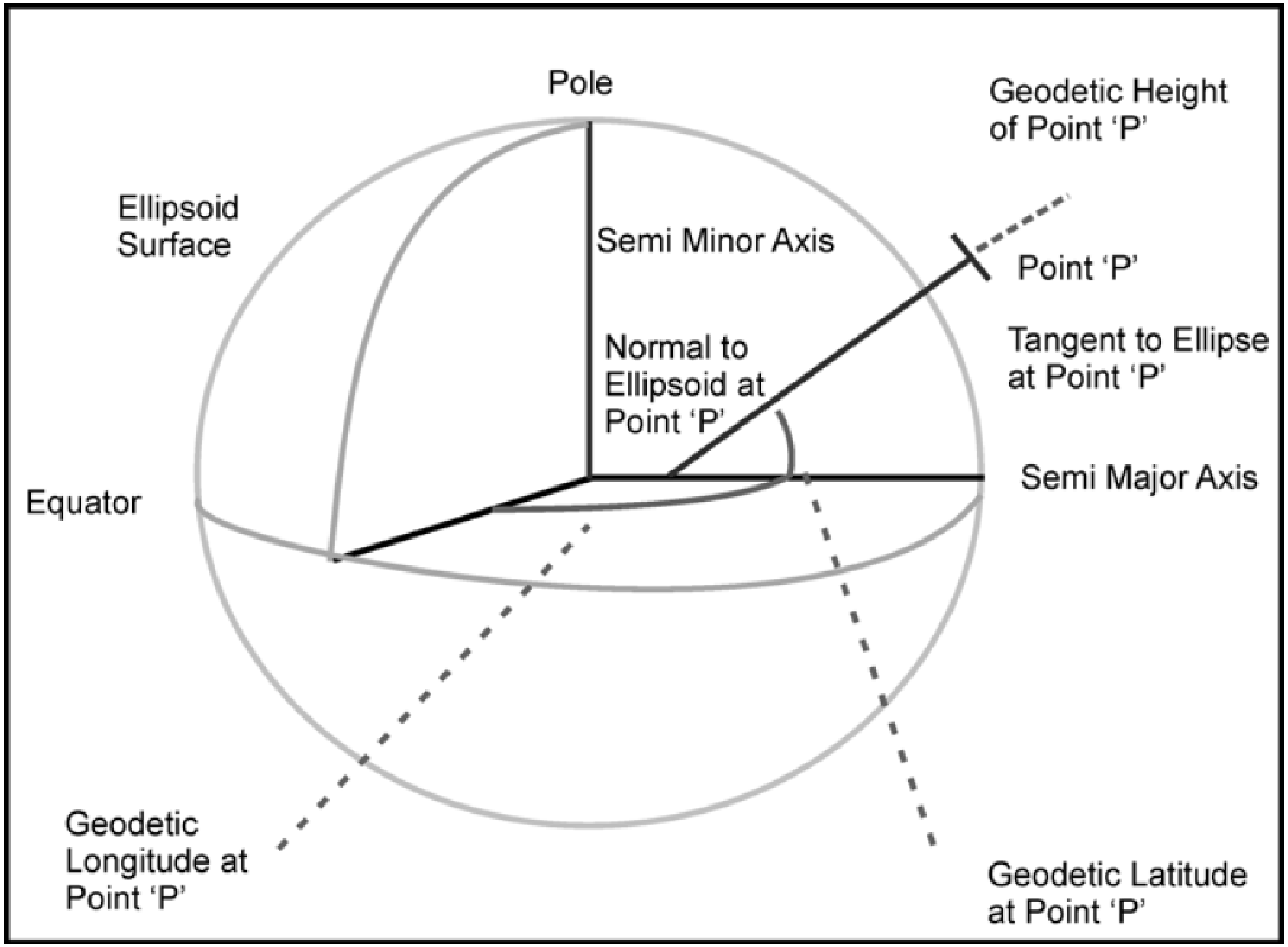


Figure 3.12: Geodetic latitude, longitude, and height.

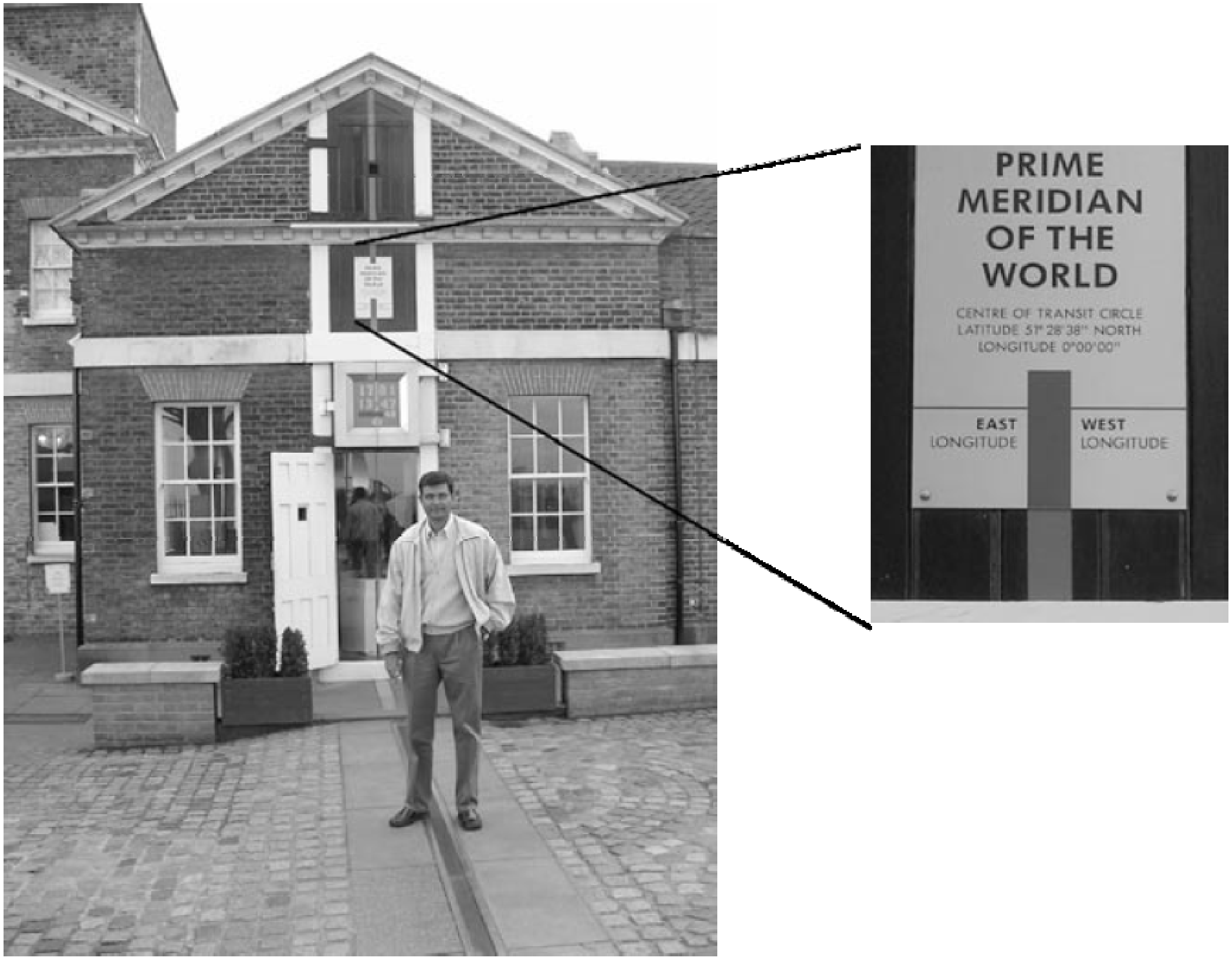


Figure 3.13: The author at Royal Observatory in Greenwich, England.

Table 3.5: Length of a degree of geodetic latitude and geodetic longitude.

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

Earth-Based Locational Reference Systems

- Reference systems and map projections extend the ideas of Cartesian and polar coordinate systems over all or part of the earth.
- Earth-based reference systems are based on various *models* for the size and shape of the earth.
  - Earth shapes are represented in many systems by a *sphere.*
- However, precise positioning reference systems are based on an *ellipsoidal earth* and *complex gravity models.*

MAP PROJECTIONS

- A map projection is a system in which locations on the curved surface of the earth are displayed on a flat sheet or surface according to some set of rules.
- Mathematically, projection is a process of transforming global location to a planar position.

Map projections and GIS

- Generally maps collected from different sources are in different projections, requiring transformation of maps to make coordinates compatible, thus, mathematical functions of projections are needed in a GIS.
- Often GIS are used for projects of global or regional scales so consideration of the effect of the earth's curvature is necessary.
- Angles, areas, directions, shapes and distances become distorted when transformed from a curved surface to a plane.
  - All these properties cannot be kept undistorted in a single projection.
  - Usually the distortion in one property will be kept to a minimum while other properties become much distorted.

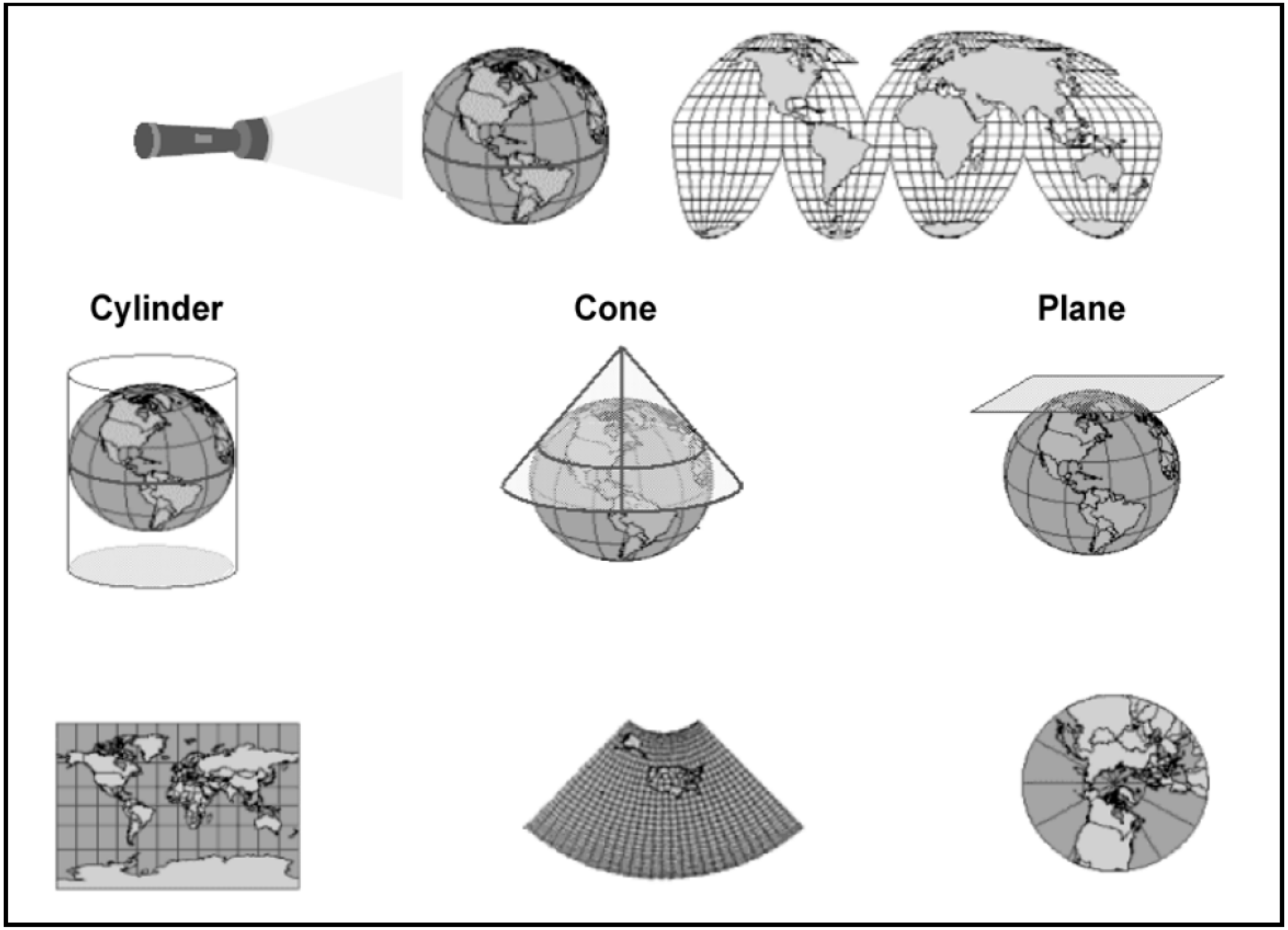


Figure 3.14: Map projections convert curved surface of the earth into a flat surface.

Tissot's Indicatrix

- This is a convenient way of showing distortion.
- If a tiny circle drawn on the surface of the globe, on the distorted map the circle will become an ellipse, squashed or stretched by the projection.
- The size and shape of the Indicatrix will vary from one part of the map to another, the Indicatrix is used to display the distorting effects of projections.

Figure of the Earth

- The figure of the earth is a geometrical model used to generate projections; a compromise between the desire for mathematical simplicity and the need for accurate approximation of the earth's shape. The common types are:
- a. *Plane*
  - It assume the earth is flat (use no projection) and used for maps only intended to depict general relationships or for maps of small areas.
  - At scales larger than 1:10,000 planar representations has little effect on accuracy.
  - Planar projections are usually assumed when working with air photos.
- b. *Sphere*
  - It assumes the earth is perfectly spherical thus does not truly represent the earth's shape.

Ellipsoid

- This is the figure created by rotating an ellipse about its minor axis.
- The ellipsoid models the fact that the earth’s diameter at the equator is greater than the distance between poles, by about 0.3%.
- At global scales, the difference between the sphere and ellipsoid are small.
  - The difference is unlikely to affect mapping of the globe at scales smaller than 1:10,000,000.

Planar or Azimuthal Projections

- A flat sheet is placed in contact with a globe, and points are projected from the globe to the sheet.
- Mathematically, the projection is easily expressed as mappings from latitude and longitude to polar coordinates with the origin located at the point of contact with the paper.
- The examples are:
  - Stereographic projection
  - Gnomic projection
  - Lambert’s azimuthal equal-area projection
  - Orthographic projection

Conic Projections

- The transformation is made to the surface of a cone tangent at a small circle (tangent case) or intersecting at two small circles (secant case) on a globe.
- Mathematically, this projection is also expressed as mappings from latitude and longitude to polar coordinates, but with the origin located at the apex of the cone.
- The examples are:
  - Alber’s conical equal area projection with two standard parallels
  - Lambert conformal conic projection with two standard parallels
  - Equidistant conic projection with one standard parallel

Cylindrical Projections

- These projections are developed by transforming the spherical surface to a tangent or secant cylinder.
- Mathematically, a cylinder wrapped around the equator is expressed with x equal to longitude, and the y coordinates some function of latitude.
- The Example is Mercator projection.

Non-Geometric Projections

- Some projections cannot be expressed geometrically, they have only mathematical descriptions.
- The examples are Molleweide and Eckert etc.

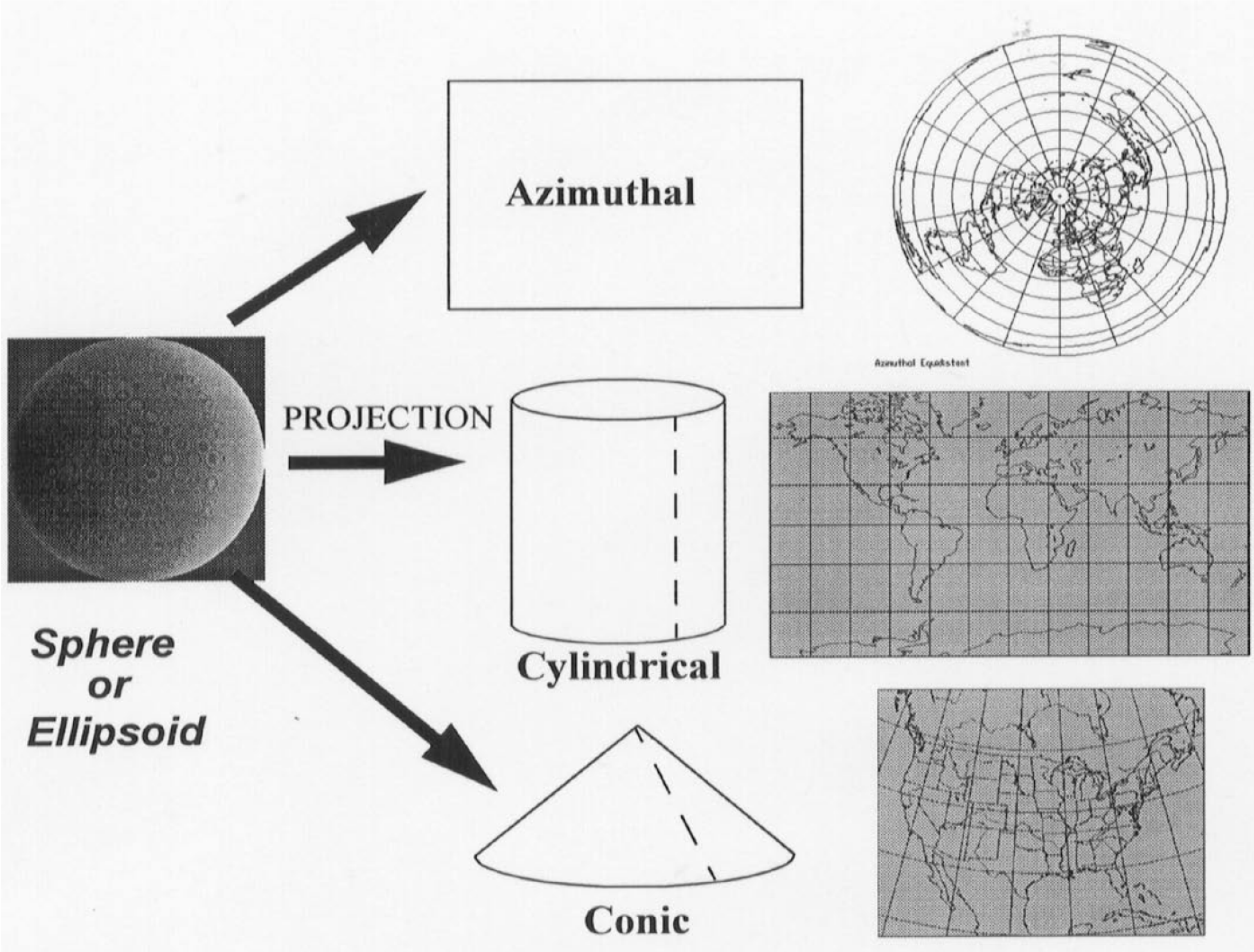


Figure 3.15: The earth can be projected in many ways, but basically into three shapes that can be unrolled into a flat map. A flat plane, a cylinder and a cone.

Geometric Analogy

- The points or lines where a developable surface touches the globe in projecting from the globe are called standard points and lines, or points and lines of zero distortion.
- If the developable surface touches the globe, the projection is called **tangent** and if the surface cuts into the globe, it is called **secant**.
- Where the surface and the globe intersect, there is no distortion.
- Where the surface is outside the globe, objects appear bigger than in reality and scales are greater than 1.
- Where the surface is inside the globe, objects appear smaller than in reality and scales are less than 1.

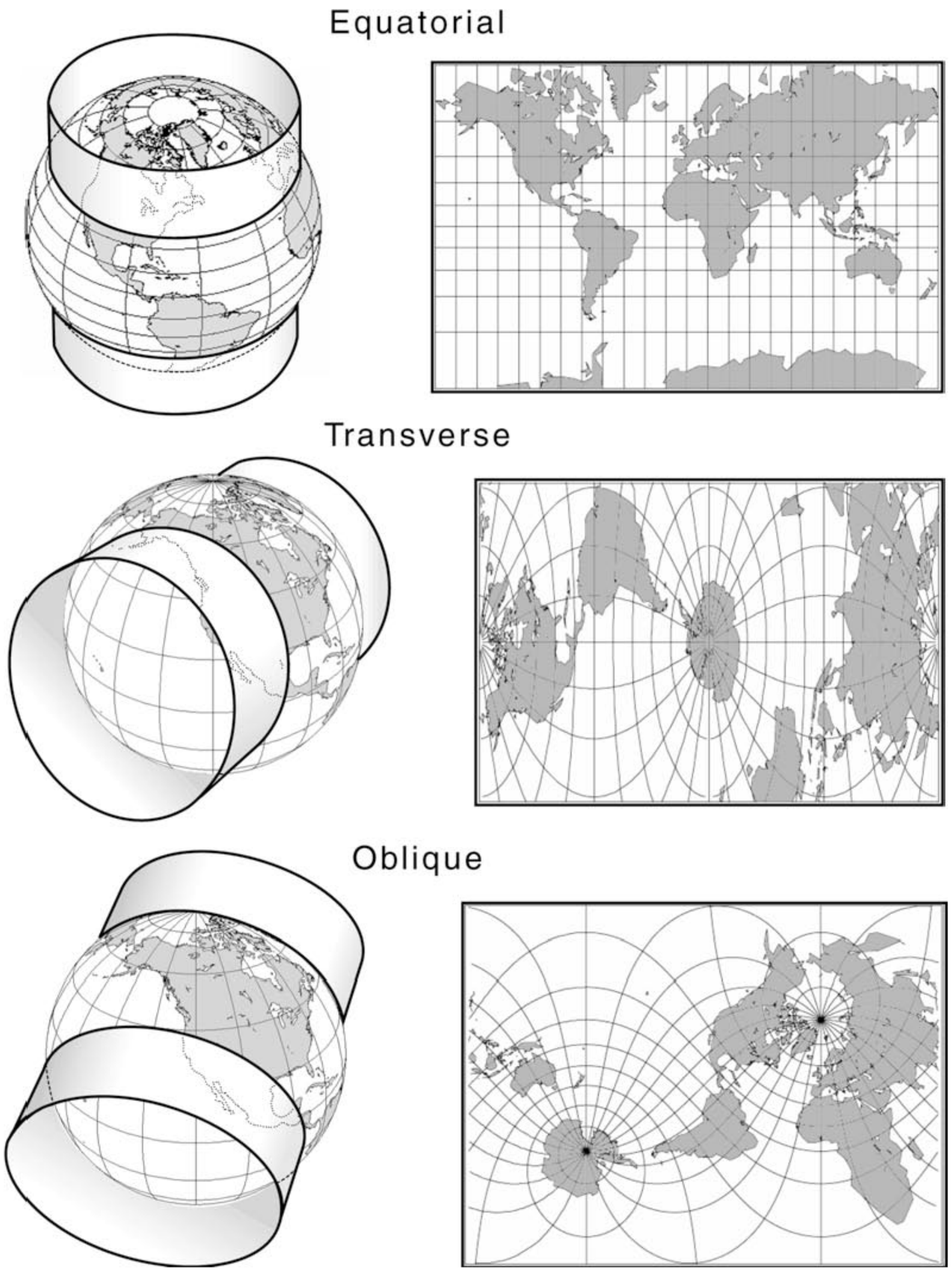


Figure 3.16: Variations on the Mercator projection shown as **Secant**.

Conformal (Orthomorphic) Projections

- A projection is conformal if the angles in the original features are preserved, over small areas the shapes of objects will be preserved.
    - Preservation of shape does not hold with large regions (*i.e.,* Greenland in Mercator projection).
- A line drawn with constant orientation (*e.g*., with respect to north) will be straight on a conformal projection, is termed a rhumb line or loxodrome.
- Parallels and meridians cross each other at right angles.
- The Tissot Indicatrix is a circle everywhere, but its size varies.
- Conformal projections cannot have equal area properties.
    - So some areas are enlarged, generally, areas near margins have a larger scale than areas near the center.

Equal Area (Equivalent) Projections

- The representation of areas is preserved so that all regions on the projection will be represented in correct relative size.
- Equal area maps cannot be conformal, so most earth angles are deformed and shapes are strongly distorted.
- The Indicatrix has the same area everywhere, but is always elliptical, never a circle (except at the standard parallel).

Equidistant Projections

- We cannot make a single projection over which all distances are maintained.
- Thus, equidistant projections maintain relative distances from one or two points only,
  - *i.e.,* in a conic projection all distances from the center are represented at the same scale.

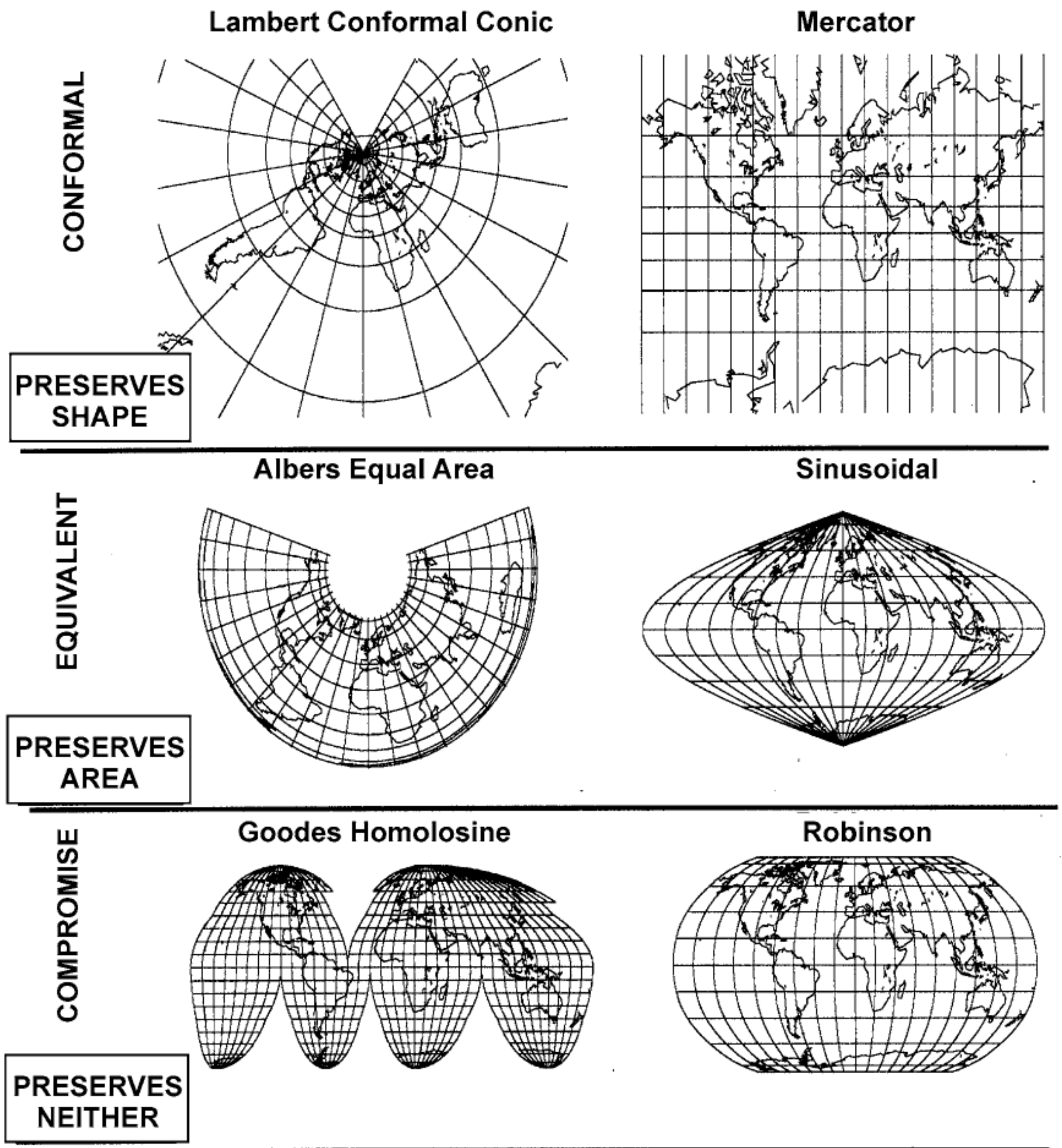


Figure 3.17: Examples of projections classified by their distortions.

Table 3.6: Common map projections: Their properties and their application areas.

Projection	Properties	Application
Albers equal area	Equal area: conformal along standard parallel	Small regional and national maps
Azimuth equidistant	Equidistant: true direction from centre	Air and sea navigation, large scale maps in the equatorial areas.
Lambert conformal conical	Conformal: locally true direction	Navigation, US – state plane system
Mercator	Conformal: true dirction	Navigation, world maps
Equidistant conical	Equidistant along standard parallel and central meridian	Mid latitude areas with east – west extent, atlas mapping for smaller countries
Polyconic-conical	Equidistant along each parallel and central meridian	Topographic maps, Survey of India maps, USGS
Sinusoidal-cylindrical	Equal area, true direction along central meridian and equator	World maps
Stereographic-planar	Conformal: true direction	Navigational maps
Transverse Mercator–cylindrical	Conformal: locally true direction	Topographic mapping for areas with north south extents

Universal Transverse Mercator (UTM)

- UTM provides georeferencing at high levels of precision for the entire globe.
- Established in 1936 by the International Union of Geodesy and Geophysics, it is adopted by many national and international mapping agencies.
- It is commonly used in topographic and thematic mapping, for referencing satellite imagery and as a basis for widely distributed spatial databases.
- Each UTM zone is identified by a number. UTM zone *numbers* designate individual 6° wide longitudinal strips extending from 80° South latitude to 84° North latitude as distortions at the poles is too large.
  - For example, Zone 14 has a central meridian of 99° west longitude. The zone extends from 96° to 102° west longitude.

- Coordinates
  - They are expressed in meters, eastings (x) are displacements eastward while northings (y) express displacement northward.
  - The central meridian is given an easting of 500,000 m.
  - The northing for the equator varies depending on hemisphere,
    - In the northern hemisphere, the equator has a northing of 0 m
    - while in the southern hemisphere, the equator has a northing of 10,000,000 m.
- Distortions
  - To reduce the distortion across the area covered by each zone, scale along the central meridian is reduced to 0.9996.
  - This produces two parallel lines of zero distortion approximately 180 km away from the central meridian.

Table 3.7: UTM zones and their extents

Zone no.	Central meridian	Bounding meridians	Zone no.	Central meridian	Bounding meridians	Zone no.	Central meridian	Bounding meridians
1	177°W	180° - 174°W	21	57°W	60° - 54°W	41	63°E	60° - 66°E
2	171°W	174° - 16°W	22	51°W	54° - 48°W	42	69°E	66° - 72°E
3	165°W	168° - 162°W	23	45°W	48° - 42°W	43	75°E	72° - 78°E
4	159°W	162° - 156°W	24	39°W	42° - 36°W	44	81°E	78° - 84°E
5	153°W	156° - 150°W	25	33°W	36° - 30°W	45	87°E	84° - 90°E
6	147°W	150° - 144°W	26	27°W	30° - 24°W	46	93°E	90° - 96°E
7	141°W	144° - 138°W	27	21°W	24° - 18°W	47	99°E	96° - 102°E
8	135°W	138° - 132°W	28	15°W	18° - 12°W	48	105°E	102° - 108°E
9	129°W	132° - 126°W	29	09°W	12° - 06°W	49	111°E	108° - 114°E
10	123°W	126° - 120°W	30	03°W	06° - 00°W	50	117°E	114° - 120°E
11	117°W	120° - 114°W	31	03°E	00° - 06°E	51	123°E	120° - 126°E
12	111°W	114° - 108°W	32	09°E	06° - 12°E	52	129°E	126° - 132°E
13	105°W	108° - 102°W	33	15°E	12° - 18°E	53	135°E	132° - 138°E
14	99°W	102° - 96°W	34	21°E	18° - 24°E	54	141°E	138° - 144°E
15	93°W	96° - 90°W	35	27°E	24° - 30°E	55	147°E	144° - 150°E
16	87°W	90° - 84°W	36	33°E	30° - 36°E	56	153°E	150° - 156°E
17	81°W	84° - 78°W	37	39°E	36° - 42°E	57	159°E	156° - 162°E
18	75°W	78° - 72°W	38	45°E	42° - 48°E	58	165°E	162° - 166°E
19	69°W	72° - 66°W	39	51°E	48° - 54°E	59	171°E	166° - 172°E
20	63°W	66° - 60°W	40	57°E	54° - 60°E	60	177°E	172° - 180°E

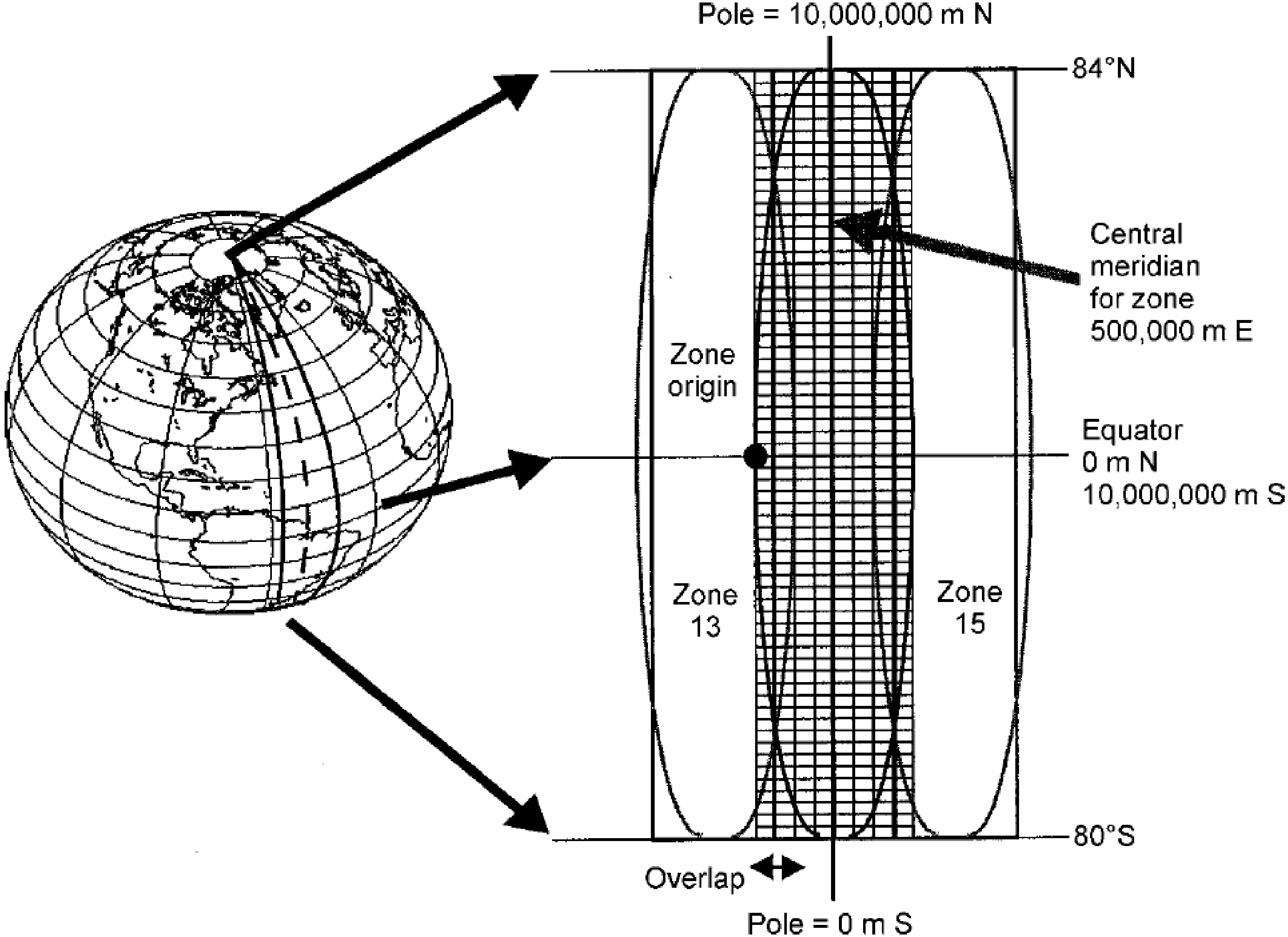


Figure 3.18: Overlap in UTM projection.

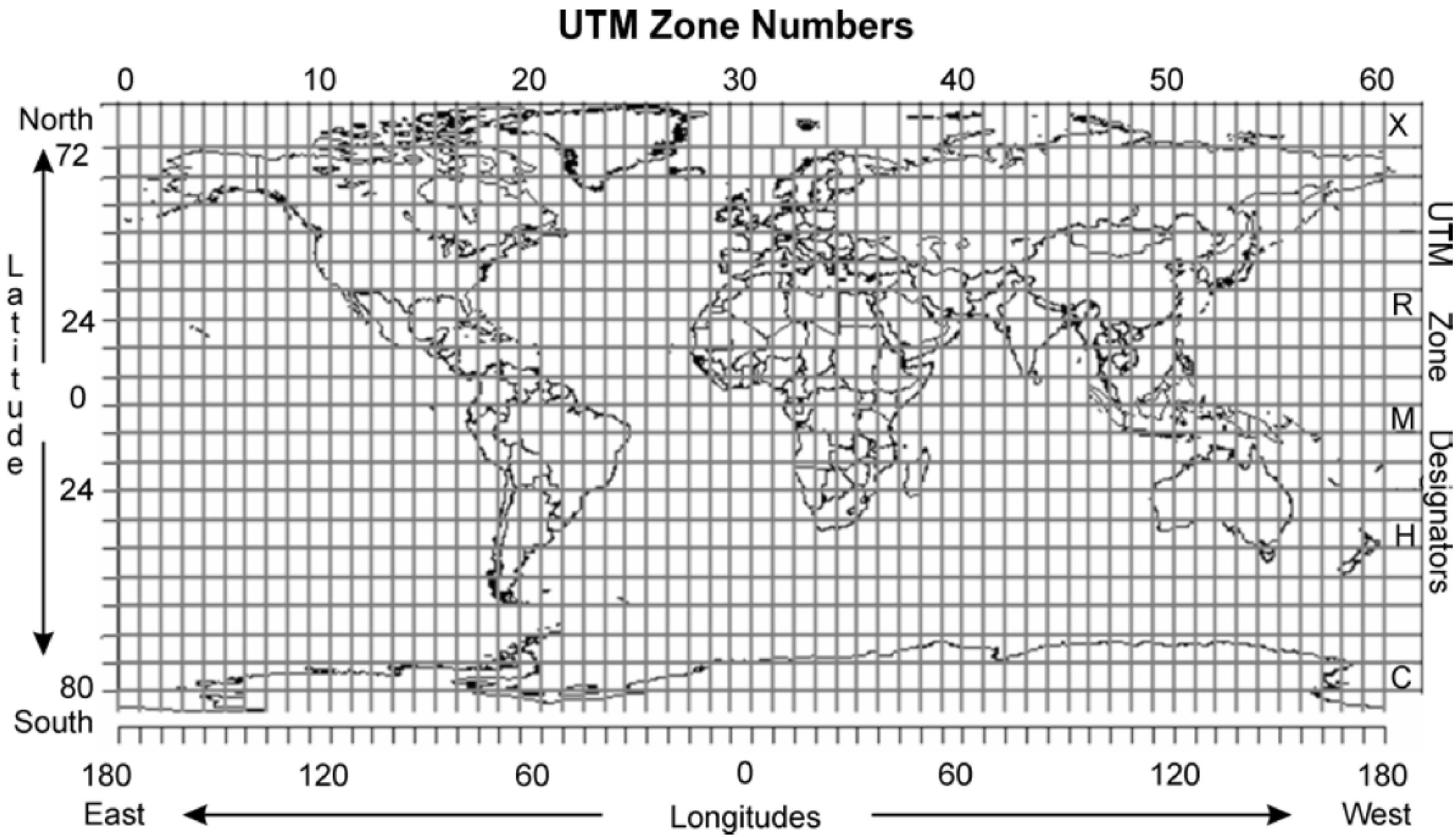


Figure 3.19: Universal transverse Mercator system.

World Geographic Reference System (GEOREF)

- The World Geographic Reference System is used for aircraft navigation.
- GEOREF is based on latitude and longitude.
- The globe is divided into twelve bands of latitude and twenty-four zones of longitude, each 15° in extent.
  - These 15° areas are further divided into one degree units identified by 15 characters.

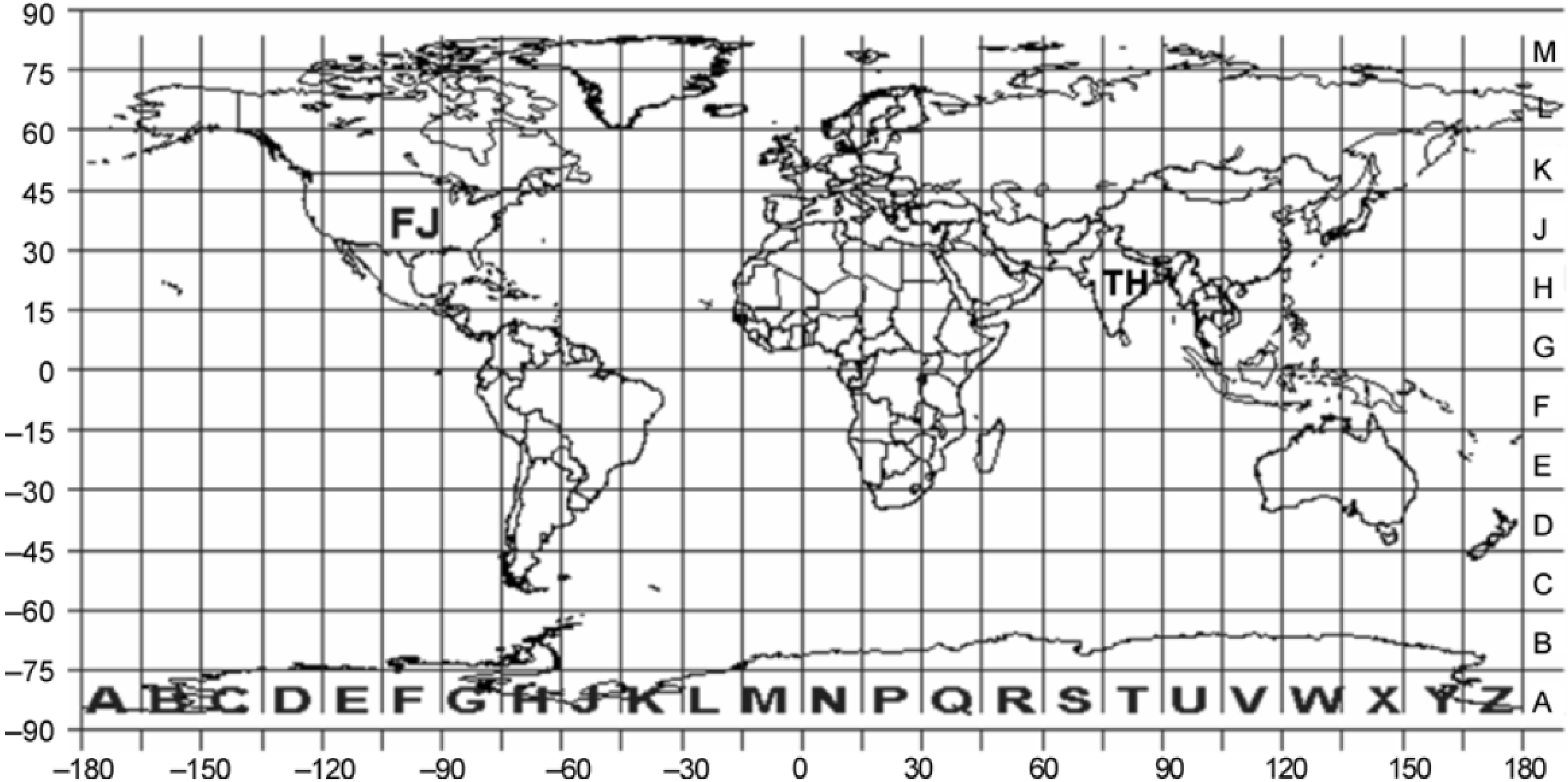


Figure 3.20: World geographic reference system (GEOREF).

Regional Systems

- Several different systems are used regionally to identify geographic location.
- Some of these are true coordinate systems, such as those based on UTM and UPS systems.
  - Others, such as the Public Land Survey systems are simply partition space.
- Many nations have defined grid systems based on Transverse Mercator coordinates that cover their territory.

- The British National Grid (BNG)
  - The British National Grid (BNG) is based on the National Grid System of England, administered by the British Ordnance Survey.
  - The BNG has been based on a Transverse Mercator projection since the 1920s. The modern BNG is based on the Ordnance Survey of Great Britain Datum 1936.
  - The true origin of the system is at 49° north latitude and 2 degrees west longitude.

- Indian Grid System
  - The Indian system follows almost the same as British system.
  - The Indian system has eight grid zones named as 00, 0I, IIA, IIB, IIIA, IIIB, IVA, IVB based on Lamberts conical orthomorphic projection with two standard parallels.
  - It covers India, Pakistan, Myanmar, Afghanistan, parts of Iran, China, Tibet and Thailand.
  - Each zone has a belt of 8° latitude.
  - The topographical maps in India are not based on lamberts projection but on polyconic projections, due to this the grid squares are not perfect squares.

- State Plane Coordinates (SPC)
  - SPCs are individual coordinate systems adopted by U.S. state agencies.
  - State plane systems were developed in order to provide local reference systems that were tied to a national datum.
  - The State Plane System 1927 was based on the North American Datum 1927 (NAD-27). The State Plane System 1983 is based on the North American Datum 1983 (NAD-83).
    - While the NAD-27 State Plane System has been superceded by the NAD-83 System, maps in NAD-27 coordinates are still in use.
    - Each state's shape determines which projection is chosen to represent that state.
  - The advantages of SPC coordinates are simpler than that of UTM and it gives a better representation than the UTM system for a state's area.
  - However, SPC are not universal from state to state and problems arises at the boundaries of projections.

GEOREFERENCING

- Methods for specifying location on the earth's surface for geographical data in a map is called as georeferencing.
- The primary requirements of a georeference are that it should be **unique**, so that there is only one location associated with a given georeference.
- It should stay ***constant*** through time, because it could create confusion if it changes.
- Data in a GIS must contain a geographic reference to a map, such as latitude and longitude.

Discrete Georeferencing

- The discrete methods – systems of georeferencing for discrete units on the earth's surface
- These methods are indirect.
  - The method provides a key or index, which can then be used with a table to determine latitude/longitude or coordinates.
  - It is important to consider the precision of these systems.
- Street Address
  - This is a common discrete method of georeferencing, here the precision of street addresses as georeferences varies greatly.
  - It is better for cities but poor for rural areas.
  - In GIS general approach is to match address to a list of streets (called address matching or 'addmatch').
  - Spelling and punctuation variations make this difficult
    - *e.g.,* Ave. or Avenue, apartment number before or after street number

- Postal Code System
  - Postal code systems have been set up in many countries, these often provide a high level of spatial precision.
  - The codes are hierarchically nested, states are uniquely identified by one or more sets of the first 2 numbers.
  - The 6 digit ZIP potentially provides a much higher level of spatial resolution,
  - But problems exist with overlapping and fragmented boundaries.
- US Public Land Survey System
  - PLSS is the basis for land surveys and legal land description over much of the US.
  - Unlike the previous systems, it is designed to reference land parcels.
  - It is commonly used by agencies such as the Bureau of Land Management and the US Forest Service, and within the oil and gas industry.

Box 5 : Commonly used systems of georeferencing

System	Domain of uniqueness	Metric / Non-metric	Example	Spatial resolution
Place name	Varies	Non metric	Hyderabad – India or Pakistan?	Varies
Postal address	Country	Non metric	11, Rose Apartments, Marris Road, Aligarh	Size of one mailbox
Postal code	Country	Non metric	202002(Aligarh, India) or WC1H OPF (London, U.K.)	Area occupied by a defined number of mailbox
Telephone code	Country	Non metric	011 (New Delhi, India)	Varies
Latitude/Longitude	Global	Metric	27°53′ North Latitude and 78°35′ East Longitude.	Infinitely fine
UTM	Zones of six degrees of longitude wide	Metric	1393267 & 3117373	Infinitely fine
State plane coordinates	USA only	Metric	55046.37 E & 75246.64 N	Infinitely fine

AFFINE AND CURVILINEAR TRANSFORMATIONS

- Coordinate transformations are required when we register different sets of coordinates for objects that may have come from maps of different (and sometimes unknown) projections.
- There are two ways to look at coordinate transformations:
  - i. move objects on a fixed coordinate system so that the coordinates change
  - ii. hold the objects fixed and move the coordinate system,

    this is the more useful way to consider transformations for GIS purposes.
- There are two major groups of transformations
  - I. **Affine transformations** are those which keep parallel lines parallel and they are a class of transformations which have 6 coefficients.
  - II. **Curvilinear transformations** are higher order transformations that do not necessarily keep lines straight and parallel and they may require more than 6 coefficients.

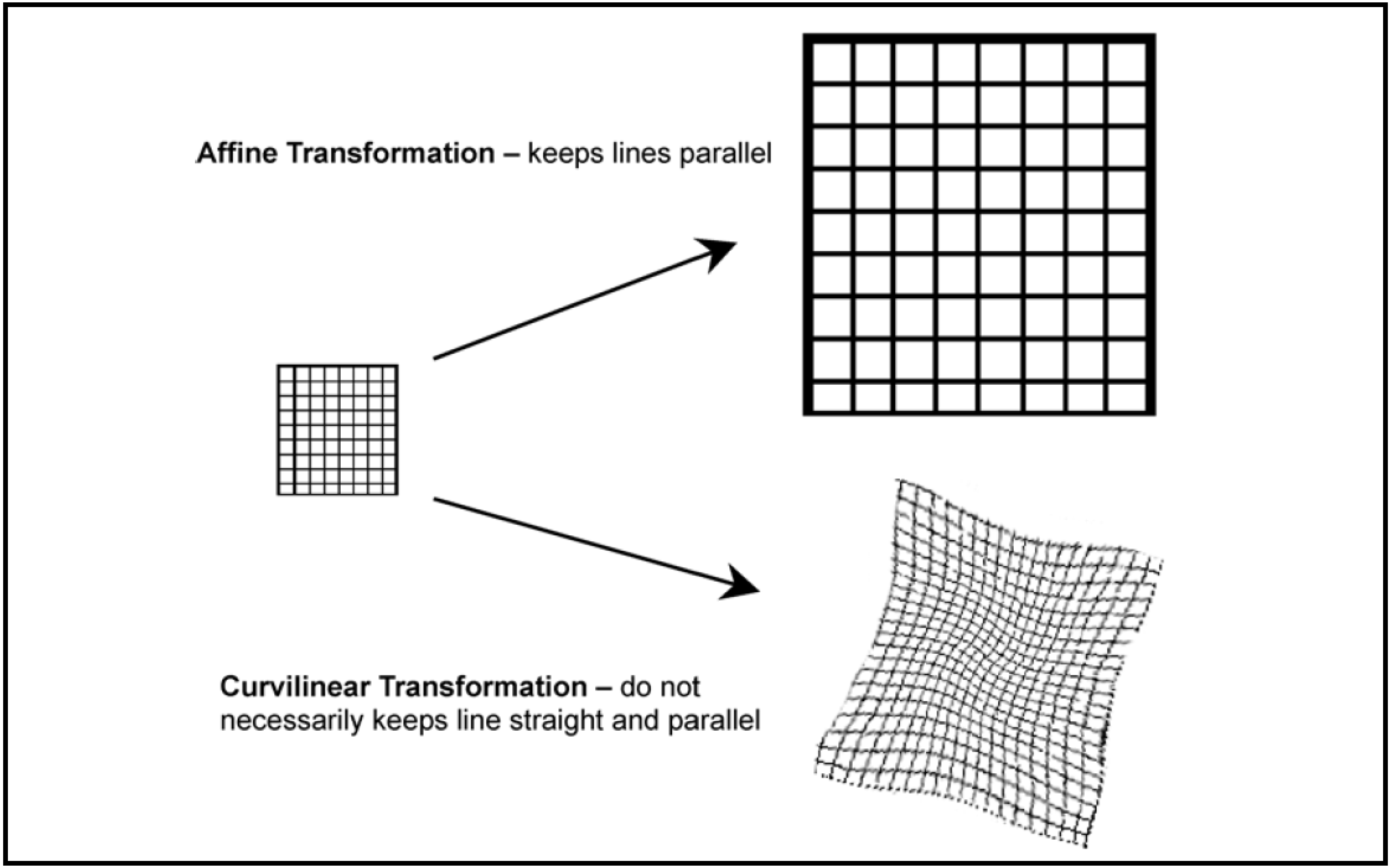
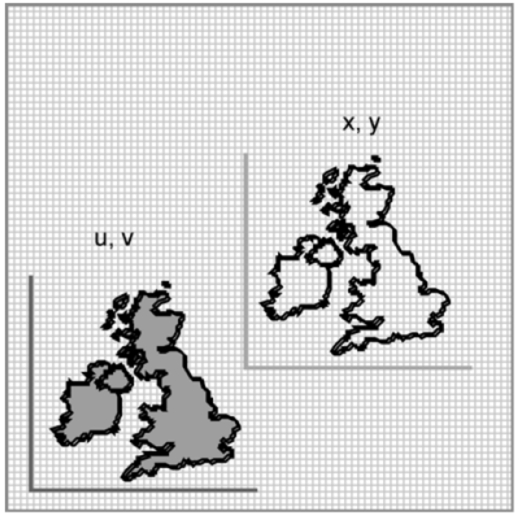


Fig. 3.21: Coordinate transformation: affine and curvilinear.

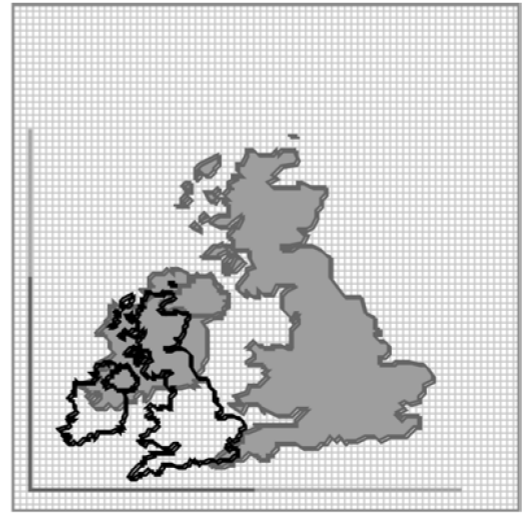
Affine Transformation Primitives

- a. Translation
  - Origin is moved, axes do not rotate
  - $u = x – a \quad v = y – b$
    - here, origin is moved a units parallel to x and b units parallel to y
- b. Scaling
  - Both origin and axes are fixed, scale changes
  - $u = cx \quad v = dy$
    - here, scaling of x and y may be different, if the scaling is different, the shape of the object will change
- c. Rotation
  - Origin fixed, axes move (rotate about origin)
  - $u = x \cos(a) + y \sin(a) \quad v = – x \sin(a) + y \cos(a)$
    - here, a is measured counterclockwise

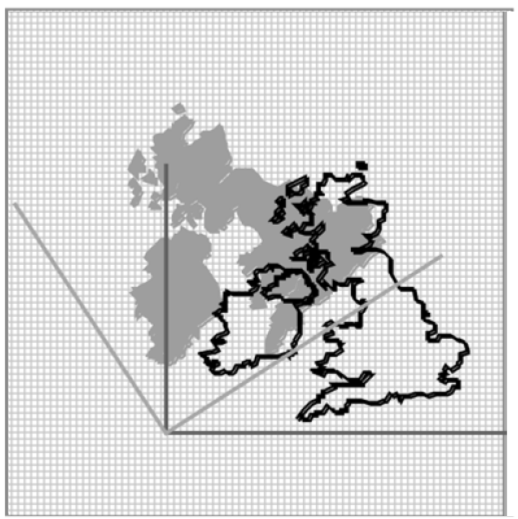
- d. Reflection
  - Coordinate system is reversed, objects appear in mirror image to reverse y, but not x
  - $u = x \quad v = c - y$
    - here, this transformation is important for displaying images on video monitors as the default coordinate system has the origin in the upper left corner and coordinates which run across and down.



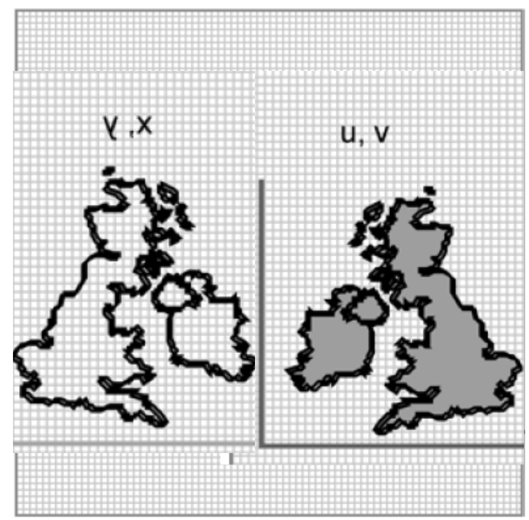
Translation (a & B)
origin is moved, axes do not rotate
$u = x - a$
$= yv - b$



Scaling (c & f)
both origin and axes are fixed,
scale changes
$u = cx$
$v = cy$



Rotation (a & d)
origin fixed, axes move
$u = x \cos(a) + y \sin(a)$
$v = -x \sin(a) + y \cos(a)$
(a = angle measured counter-clockwise)



Reflection
(b & e)
co-ordinate system is reversed, objects
appear in mirror image
$u = x$
$v = c - y$

Figure 3.22: Affine transformations.

Complex Affine Transformations

- Usually a combination of these transformations will be needed because often we cannot actually separate the needed transformations into one or more of the primitives.
- One transformation will cause changes that appear to be caused by another transformation, and the order is important.
- The combined equations are: $u = a + bx + cy \quad v = d + ex + fy$

Affine transformations in GIS

- In order to register two data sets, a set of control points or tics must be identified that can be located on both maps.
- It is necessary to have at least 3 control points since 3 points provide 6 values which can be used to solve for the 6 unknown points.
- Another precaution which is important that control points must not be on a straight line (not collinear).

Curvilinear Transformations

- Simple linear affine transformation equations can be extended to higher powers:
  - $u = a + bx + cy + gxy$ or $u = a + bx + cy + gx^2$
  - or $u = a + bx + cy + gx^2 + hy^2 + ixy$
- Curvilinear transformations usually give greater accuracy.