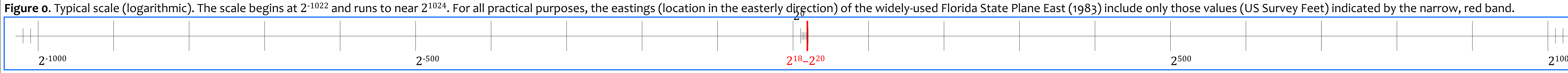
SHEET	DRAWING TITLE	DRAWN BY
1 OF A SERIES	**Sub-Nanometer Scale GIS: Number Systems for GIS**	D. Michael Parrish
PROJECT G	**The Program for Universal Computational Freedom** https://github.com/users/dmparrishphd/projects/1	PRESENTED AT The 28th South Florida GIS Expo 2021-08-26–27

Sub-Nanometer Scale GIS
Number Systems for GIS

THE PROGRAM FOR UNIVERSAL COMPUTATIONAL FREEDOM

Figure 0. Typical scale (logarithmic). The scale begins at 2^{-1022} and runs to near 2^{1024}. For all practical purposes, the eastings (location in the easterly direction) of the widely-used Florida State Plane East (1983) include only those values (US Survey Feet) indicated by the narrow, red band.



There is a disconnect, typically, between actual GIS problems and the representation of geographic **locations** in hardware addressing those problems. Locations are often stored as **floating-point** values: this **wastes** significant **storage**, **time**, and **energy**.

0000111100001111000011110000111100001111000011110000111100001111

Figure 1. Typical **64-bit–word** representation of floating-point values uses 53 bits of precision (p, black), eleven bits of exponent (e, red), and one sign bit (s, blue) to represent values $(-1)^s(1+p/2^{53})\times2^{e-1023}$.

For example, in data sets using the NAD83 HARN Florida East State Plane and units of US Survey Feet, eastings cover only two orders of magnitude, from roughly 2^{18} to less than 2^{20}. In other words, although only one bit is needed for the exponent (e.g., to select between 18 or 19), eleven bits are used. This **wastes** about **16%** of space (10 out of 64 bits).

In the eastern half of the state plane, there are 34 bits to the right of the ones' place, corresponding to a precision of 2^{-34} ft (about 5.8×10^{-11} ft $\approx 1.8\times10^{-11}$ m $= 1.8\times10^{-2}$ nm $=$ 18 pm).

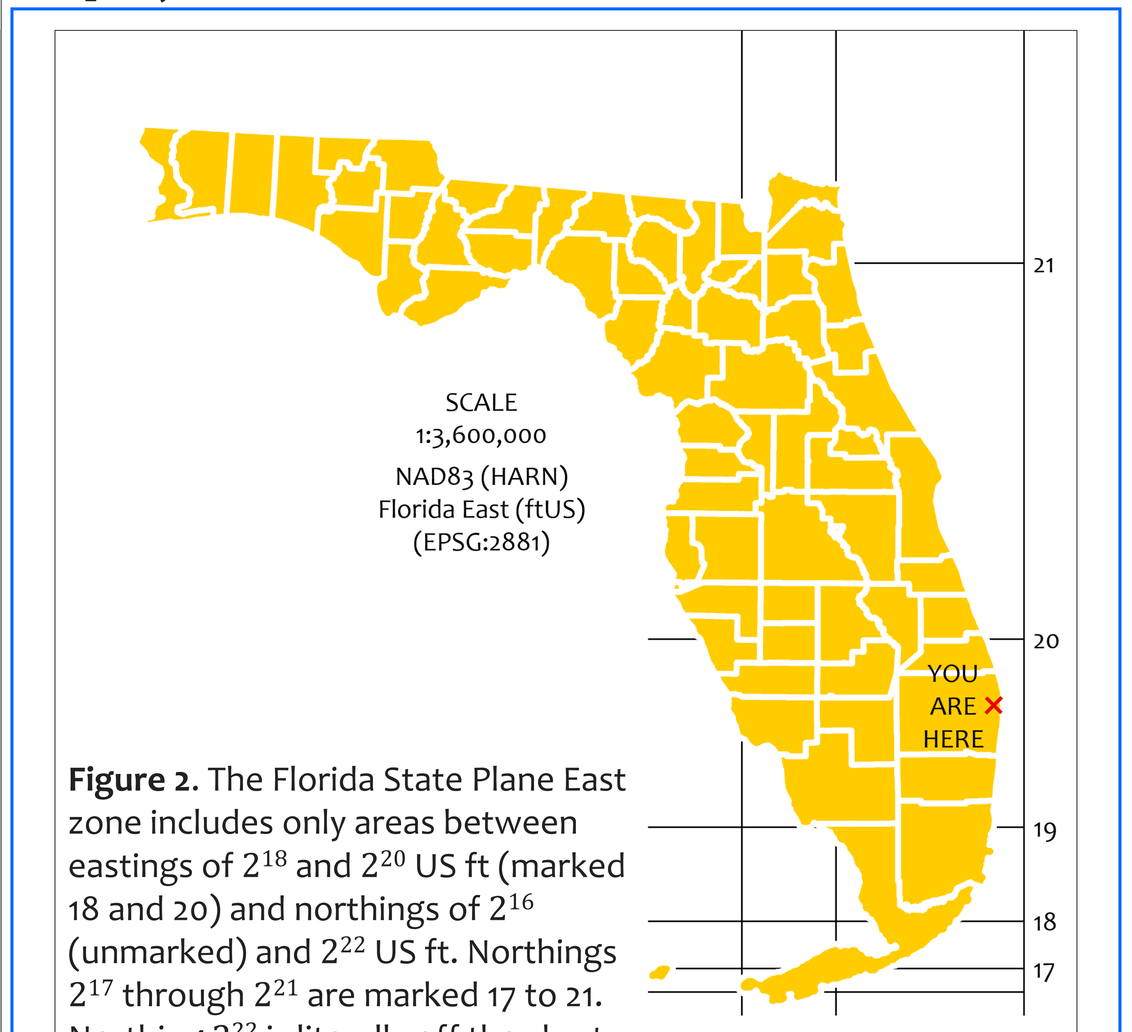


Figure 2. The Florida State Plane East zone includes only areas between eastings of 2^{18} and 2^{20} US ft (marked 18 and 20) and northings of 2^{16} (unmarked) and 2^{22} US ft. Northings 2^{17} through 2^{21} are marked 17 to 21. Northing 2^{22} is literally off the chart. $2^{20}\approx10^6 = 1{,}000{,}000$.

What's the Alternative? There are many. For vector data we suggest the following.

Store location data using 64-bit **scaled integers.** Divide the area of interest into a **regular grid** of arbitrary **grid size** $\Delta x = \Delta y$.

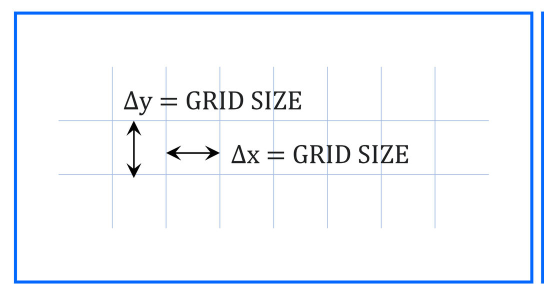


Figure 3. Space may be organized into a regular grid with **grid size** $\Delta x = \Delta y$.

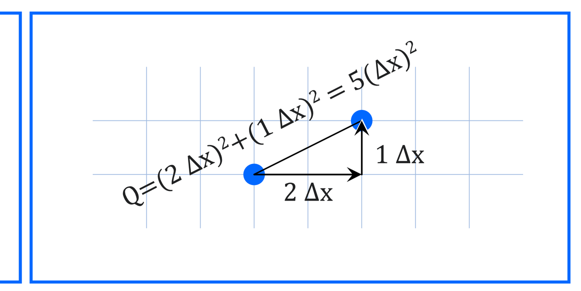


Figure 4. The separation between two **points** can be quantified using component **vectors** or by **quadrance.** Quadrance can be thought of as the *square* of distance; it can be computed **exactly** for any two points.

Points are represented by coordinates each ranging from $1-N$ to $N-1$ (one of $2N-1$ distinct values). If $N = 2^{60}\approx10^{18}$ and grid size is 1 pm $= 10^{-12}$ m, then N represents over one million meters. Such choices would be suitable for many regional areas of interest, including the Florida East State Plane.

There are at least two ways in which to represent the **separation between two points**. One is by a **vector**, composed of directed distances along gridlines (e.g., east and north components). Another is by **quadrance**, which can be thought of as the *square* of distance.

Precise, though inexact, locations are summarized using parallelograms whose vertices are **points**. Such objects are termed **p-parallelograms** ("p-" for "point").

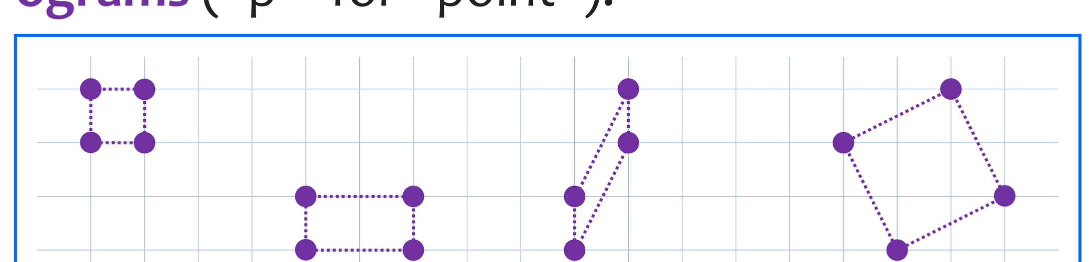


Figure 5. Certain collections of **points** form **p-parallelograms. P-parallelograms** need not be squares or rectangles (the skewed figure at center is a p-parallelogram). Neither do they need be aligned with the grid. The dotted "lines" are included for clarity; they are not part of any **p-parallelogram.**

Lines—not to be confused with line segments—are each represented by three **integers.** A point (x, y) is on the line (c, a, b) if and only if

$$c + a\,x + b\,y = 0$$

Furthermore—by design:

$$1-8N \le a \le 8N-1,\ 1-8N \le b \le 8N-1, \text{ and}$$
$$1-8N^2 \le c \le 8N^2-1$$

These constraints tolerate lines which pass through 1) any two **points**, 2) the centroids of any two **p-parallelograms**, or 3) any one **point** and the centroid of any one **p-parallelogram.**

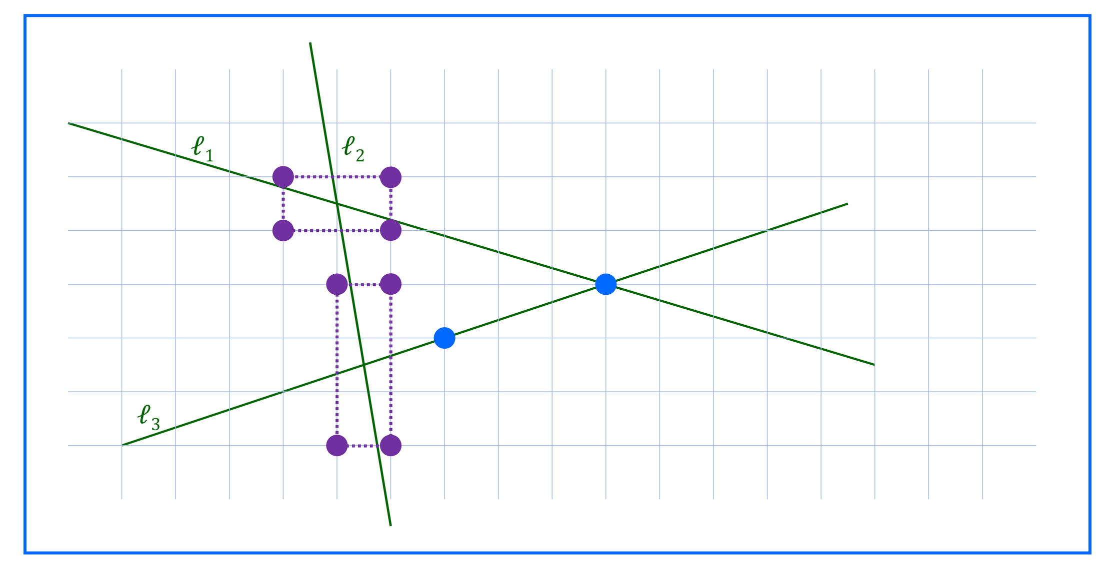


Figure 6. Lines extend indefinitely. **Line** ℓ_1 passes through a **point** and the centroid of a **p-parallelogram. Line** ℓ_2 passes through two **p-parallelograms. Line** ℓ_3 passes through two **points.**

Exclusive and **inclusive half-planes**, are similar to **lines** in their representation. The *corresponding* relations (inequalities) are

$$c + a\,x + b\,y < 0 \quad\text{and}\quad c + a\,x + b\,y \le 0$$

A point (x, y) is in a **half-plane** if and only if the corresponding relation is satisfied.

Precise, inexact boundaries are represented with pairs of parallel **half-planes** called **swaths.**

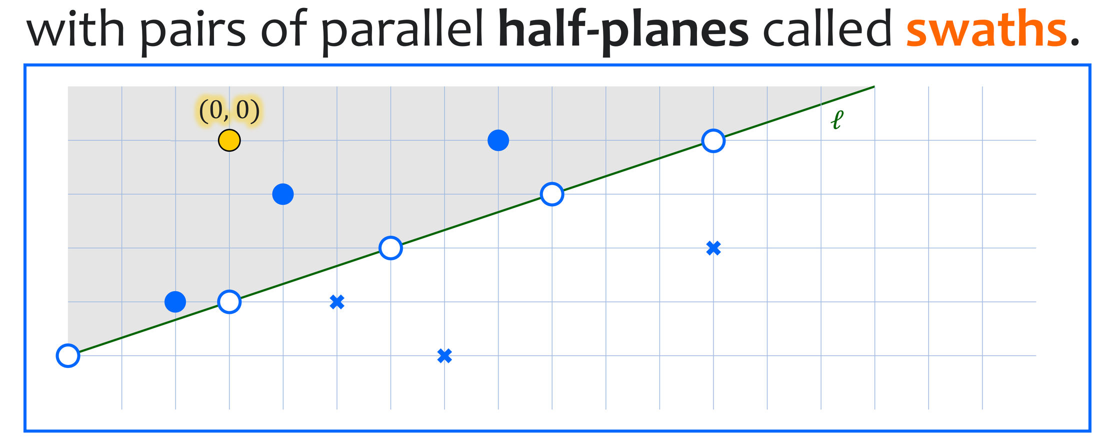
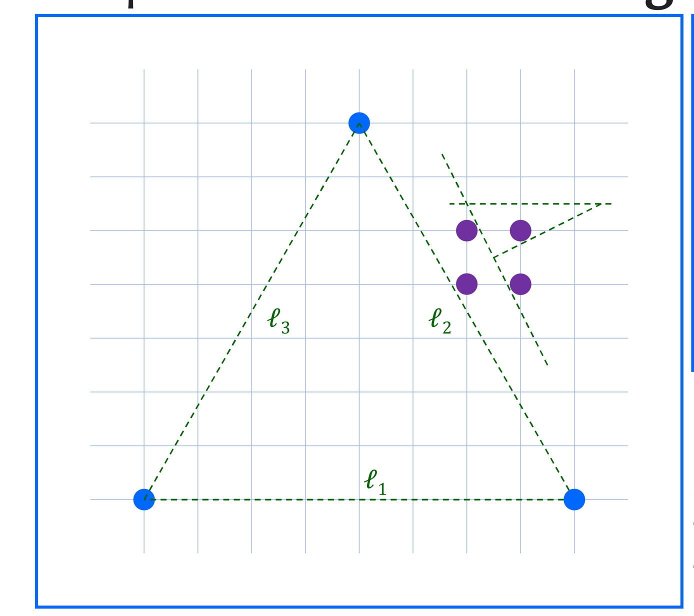


Figure 7. The **points** indicated by the solid dots (•) are among those in the exclusive half-plane with boundary **Line** ℓ. The corresponding inequality is $-9 + 1\,x + (-3)\,y < 0$. The points indicated by the hollow dots (O) are not in the exclusive half-plane; neither are the points indicated by the exes (x).

Triangles may be specified in at least two ways: as a triple of **points**—a **p-triangle**—or as a triple of **lines**—an **ℓ-triangle.**



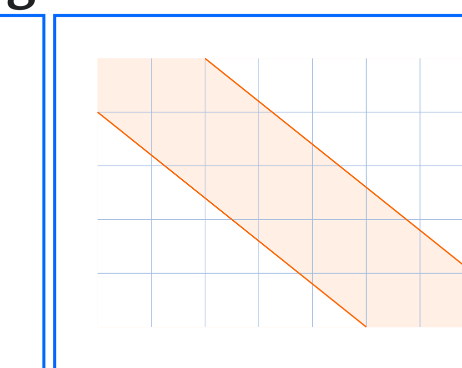


Figure 9 (above). A pair of parallel **half-planes**, forming a **swath**, represents a precise, though inexact linear feature.

Figure 8. A **triangle** can be represented by three **points** or by three **lines.** The vertices of an **ℓ-triangle** need not be (lattice) **points.**

The vertices of an **ℓ-triangle** might not coincide with (lattice) **points.** Such vertices may still be located precisely, though not exactly. A square **p-parallelogram** spanning a single grid cell provides precision at the absolute limit.

A **triangular region** may be specified using three **half-planes**, where the interior of the triangular region is in all three **half-planes.**

DISCUSSION. By using—rather than wasting—memory, precision may be extended to the picometer scale. Normalizing for precision, integer arithmetic offers potential time and energy savings vs. floating-point arithmetic.

Feature	Requirement
Integer	1
Point	2
Vector	2
Quadrance	2
P-Paralleogram	8
Line	4
Half-Plane	4
Swath	6
P-Triangle	6
L-Triangle	12

Table 1. Storage requirements for select sub-nanometer scale GIS objects. Units are **64-bit** words, as in **Figure 1.**

SEE ALSO The accompanying white paper.

ACKNOWLEDGEMENTS. Some material follows directly from Wildberger. His 2005 book *Divine Proportions: Rational Trigonometry to Universal Geometry*, Wild Egg Pty Ltd; Sydney, Australia; ISBN 097574920X; as well as his numerous recorded lectures on rational geometry (search: "wildtrig") are particularly relevant. Some inspiration and concepts come from a lecture delivered by John Gustafson entitled "Beyond Floating Point: Next Generation Computer Arithmetic" (2017-02-01; Stanford Center for Professional Development, Computer Systems Colloquium). This poster was developed using equipment at the South Florida Water Management District.