# Global addresses

August 25, 2018

Abstract

What follows was an attempt to derive a mapping from a latitude / longitude pair to a single location identifier. The solution investigated below used regular subdivision and can not guarantee all the indexed Voronoi cells to have equal area. A better solution would be to create an equally spaced spiral on a sphere and divide it in equal segments (see [4, 5, 6]). This is likely the solution W3W uses, coupled with a 45 bits identifier (3 words from a 2^{15} database, see [7, 8, 9] for wordlists). Best option available today is plus codes [10, 11, 12]: they rely on an existing point of reference and allow to specify a variable length code for fine positioning within the area of interest. Realistically, an address should be capable of disambiguate an area at most the extent of Sahara desert, or metropolitan Tokyo.

1 How many identifiers are needed?

Coordinates on planet Earth can be specified with a latitude/longitude pair of numbers, which express location as a polar mapping on a reference sphere. Earth is more similar to an oblate spheroid, however GPS coordinates provide a convenient layer of abstraction over its actual geometry.
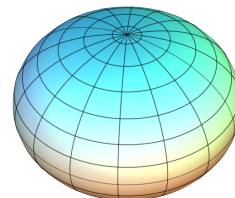
The surface area of Earth A_E is roughly 510.1 million Km^2. With each additional bit of information, we can halve the area we can index. When using B bits, the following relations hold:

$$A_E = 510.072 \cdot 10^{12} \text{ m}^2 \qquad A = \frac{A_E}{2^B} \qquad B = \log_2\left(\frac{A_E}{A}\right) \qquad (1)$$

where A is the area of each individual indexed patch. Consumer GPS module have a worst case precision of 8-10 meters, however in normal usage 5 meters is the expected deviation from the true position. With temporal integration, is not uncommon for commercial units to deviate less than 3 meters from the true position in ideal conditions. Plugging into eq. 1 the area of a circle of radius 3

meters, we obtain:

$$B = \log_2 \left(\frac{A_E}{\pi R^2} \right) = \log_2 \left(\frac{510.072 \cdot 10^{12}}{\pi \, 3^2} \right) \gtrless 44.036 \text{ bits}$$

which is a lower bound estimate of the bit budget required, since it is not possible to cover (tessellate) a surface exactly with circles.

2 Regular subdivision

One option to partition the surface of Earth into a number congruent shapes is to recursively subdivide a convex regular polyhedron inscribed in the reference sphere. This class of shapes is also known as platonic solids. All shapes but
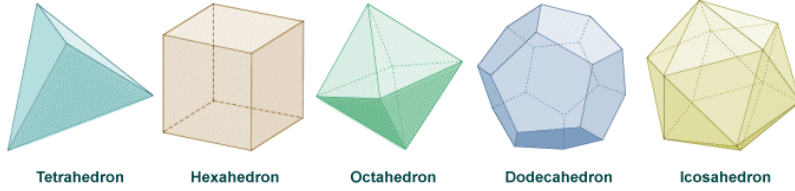
the dodecahedron have triangular or square faces, which can be conveniently subdivided into four equally sized and congruent pieces. We will ignore the dodecahedron for the rest of this document. A location can then be approximated providing an index to the closest face midpoint or vertex after subdivision.

It should be noted that regular subdivision does not guarantee equal areas; the discrepancy between the largest and smallest of them depend on the initial maximum distance between a point on the solid and the circumscribed sphere (e.g. largest on a tetrahedron, smallest for an icosahedron).

2.1 Number of faces and vertexes after subdivision

After i rounds of subdivision, the following equations specify the total number of faces F_i, edges E_i and vertexes V_i.

$$F_i = F_0 \cdot 4^i \tag{2}$$

$$E_i = \frac{\mathcal{L}}{2} \cdot F_i = \frac{\mathcal{L}}{2} \cdot F_0 \cdot 4^i = E_0 \cdot 4^i \tag{3}$$

$$V_i = 2 + E_i - F_i =$$

$$= 2 + \frac{\mathcal{L}}{2} \cdot F_i - F_i = 2 + \left(\frac{\mathcal{L}}{2} - 1 \right) \cdot F_i =$$

$$= 2 + \frac{1}{2} \cdot (\mathcal{L} - 2) \cdot F_0 \cdot 4^i = 2 + \mathcal{V} \cdot 4^i \tag{4}$$

where $\mathcal{L}$ is the number of face edges and $\mathcal{V}$ is defined as $1/2(\mathcal{L} - 2)F_0$. Table 1 summarizes the initial values for $i = 0$.

	V_0	E_0	F_0	$\mathcal{L}$	$\mathcal{V}$
Tetrahedron	4	6	4	3	2
Cube, hexahedron	8	12	6	4	6
Octahedron	6	12	8	3	4
Icosahedron	12	30	20	3	10

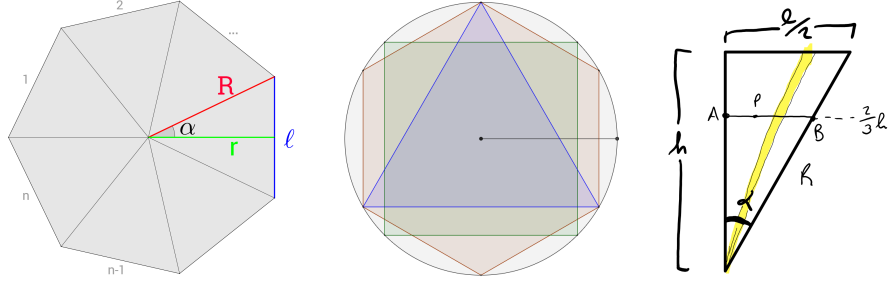Table 1: platonic solids statistics.

Figure 1: (left) ; (center) ; (right) .

2.2   Tessellation types

If using face midpoints to approximate locations, faces are subdivided into triangles or squares. When using vertexes, they are partitioned with the respective dual tessellations, hexagons and squares.

Given a regular polygon with n sides (fig. 1 left) the relationship between its area and the radius of the circumcircle can be determined from the following identities:

$$A = \frac{\ell \cdot r}{2} \cdot n \qquad r = R \cos \alpha \qquad \ell = 2 \cdot R \sin \alpha \qquad \alpha = \frac{2 \cdot \pi}{2 \cdot n} = \frac{\pi}{n}$$

$$A_n = \frac{2 \cdot R \sin \alpha \cdot R \cos \alpha}{2} \cdot n = \frac{n}{2} \sin \left(\frac{2\pi}{n} \right) R^2 \tag{5}$$

$$R_n = \sqrt{\frac{2A}{n \sin \left(\frac{2\pi}{n} \right)}} \tag{6}$$

The larger the number of edges the higher is the coverage of the reference

circumscribed circle.

$$A_3 = \frac{3\sqrt{3}}{4}R^2 \qquad \approx 1.299\,R^2 \qquad\qquad R_3 = \sqrt{\frac{4A}{3\sqrt{3}}} \qquad \approx 0.877\,\sqrt{A}$$

$$A_4 = 2R^2 \qquad = 2.000\,R^2 \qquad\qquad R_4 = \sqrt{\frac{A}{2}} \qquad \approx 0.707\,\sqrt{A}$$

$$A_6 = \frac{3\sqrt{3}}{2}R^2 \qquad \approx 2.598\,R^2 \qquad\qquad R_6 = \sqrt{\frac{2A}{3\sqrt{3}}} \qquad \approx 0.620\,\sqrt{A}$$

$$A_\infty = \pi R^2 \qquad \approx 3.142\,R^2 \qquad\qquad R_\infty = \sqrt{\frac{A}{\pi}} \qquad \approx 0.564\,\sqrt{A}$$

2.3 Mean distance from center

Possibly counter-intuitively, increasing the number of edges increases the covered area, but also the expected mean distance from the centre. For a certain bit budget, we can now compute maximum and expected error for a given tesselation.

$$A = \frac{2}{3}(0, h) \qquad B = \frac{2}{3}\left(\frac{\ell}{2}, h\right) \qquad \frac{\ell}{2} = R\sin\alpha \qquad h = R\cos\alpha$$

A more compact shape reduces the maximum absolute error, but makes the average mean distance from the centre.

$$P(t) = (1 - t) \cdot A + t \cdot B = \frac{2}{3}\left(t \cdot \frac{\ell}{2}, h\right) = \frac{2}{3}R\,(t \cdot \sin\alpha, \cos\alpha)$$

$$D(t) = \frac{2}{3}R\sqrt{t^2 \sin^2\alpha + \cos^2\alpha}$$

$$\overline{D} = \int_0^1 D(t)\,dt = \frac{2}{3}R\int_0^1 \sqrt{t^2 \sin^2\alpha + \cos^2\alpha}\,dt$$

$$= \frac{1}{3}(R + R\cos\alpha\cot\alpha\log(\tan\alpha + \sec\alpha))$$

This happens because more mass/area is distributed towards the perifery of the circumcircle (for details on the derivation see [3]).

Reference shape	Triangle	Square	Hexagon	Circle
Half-angle α (degrees)	60°	45°	30°	0°
Mean distance, unit radius	0.46	0.54	0.61	$0.\overline{6}$

2.4 Bit-budget vs. quality of approximation

Base mesh	T	O	C	T	O	C	T	O
Subdivisions	20	20	20	21	21	21	22	22
Bit budget	41	42	42.6	43	44	44.6	45	46
Worst case approx.	9.45	6.68	6.22	4.72	3.34	3.11	2.36	1.67
Mean approximation	5.76	4.08	3.36	2.88	2.04	1.68	1.44	1.02

References

[1] Spheroid (Wikipedia).
Retrieved on Jan 18th, 2016.

[2] Ten things you don't know about the Earth (BadAstronomy).
Retrieved on Jan 18th, 2016.

[3] Mean Distance from Vertex to Interior of Plane Figures.
Retrieved on Jan 18th, 2016.

[4] How to find the length of a spherical spiral.
Retrieved on Aug 24th, 2018.

[5] Making a spiral sphere (MathExchange).
Retrieved on Aug 24th, 2018.

[6] Spreading points on a disc and on a sphere.
Retrieved on Aug 24th, 2018.

[7] Deep Dive: EFF's New Wordlists for Random Passphrases.
Retrieved on Aug 24th, 2018.

[8] The Diceware Passphrase Home Page.
Retrieved on Aug 24th, 2018.

[9] PGP word list (Wikipedia).
Retrieved on Aug 24th, 2018.

[10] Plus codes for developers.
Retrieved on Aug 24th, 2018.

[11] Evaluation of Location Encoding Systems.
Retrieved on Aug 24th, 2018.

[12] Open location codes (Github).
Retrieved on Aug 24th, 2018.