The rHEALPix Discrete Global Grid System

R G Gibb

Informatics Team Leader, Landcare Research, Private Bag 11052,
Palmerston North 4442, New Zealand

E-mail: gibbr@landcareresearch.co.nz

Abstract. In light of the new standard for Discrete Global Grid Systems (DGGS) currently being proposed by the Open Geospatial Consortium (OGC), we revisit our earlier work on defining the rHEALPix DGGS. This paper briefly restates the definition of the rHEALPix DGGS and shows how the cell geometries, and schema for cell identifiers are both compliant with the standard. We then show that the fundamental topological operators, required by the standard can be implemented directly by manipulation of the cell identifiers without reference to or knowledge of traditional coordinates.

1. Introduction

Early work characterising discrete global grid systems (DGGS) [1; 2] focused on identifying a broad set of ideal geometric properties for DGGS and establishing measures to assess the deviation from the ideal across the globe. Using these measures different DGGS geometries could be compared and DGGS identified to meet the geometric needs of particular applications. In the decade since this early work practitioners realised that geometry alone wasn't sufficient for defining a useful DGGS, and that for instance a cell addressing scheme was also needed [3]. The OGC DGGS Standards Working Group [4] has gone further and now proposes that standard DGGS must "define a hierarchical tessellation of equal area cells that both partition the entire Earth at multiple levels of granularity and provide a global spatial reference frame." but also provide "encoding methods to: address each cell; assign quantized data to cells; and perform algebraic operations on the cells and the data assigned to them."

The new definition for DGGS contrasts with earlier usage in two important ways.

(i) Early work [1; 2; 5] focused on defining a set of 14 geometric criteria—the 'Goodchild Criteria'—for grids that should be taken into account in designing a good DGGS. Since it is mathematically impossible [1] to completely fulfill all the 'Goodchild Criteria' the OGC standard identifies an essential subset of geometric criteria that can be simultaneously satisfied by a number of different approaches to DGGS design.

(ii) While early work focused primarily on geometry, the OGC standard recognises the emerging demands of spatial big-data analysis, linked spatial data and DGGS interoperability need a new digital spatial reference system for the whole-earth. Despite pre-dating the era of the internet and bigdata, Goodchild also foresaw the need for an addressing system and efficient algorithms supporting common grid operations in his criteria [5], though he didnt stipulate what those operations should be. The OGC standard therefore identifies a specific set of fundamental operations that the spatial reference system's addressing scheme should support.

The HEALPix DGGS introduced in [6] to support cosmic microwave background radiation studies, defines a family of related DGGS each with the following features:

(i) Cells that are hierarchical, congruent, aligned for odd values of an integer parameter N_{side}, and constant aperture, which makes it easy to implement with efficient data structures, and easy to establish cell addressing schemes.

(ii) At every resolution their grid cells have equal areas, which provide equal probabilities for statistical analyses, and ensure environmental properties assigned to each cell have equivalent meaning.

(iii) At every resolution their grid cells cover the entire globe, so that area is preserved throughout the hierarchy, and every cell has a nucleus that can be used as a reference point.

(iv) At each resolution their k nuclei lie on only $\Theta(\sqrt{k})$ parallels of latitude, which makes computing spherical harmonics fast.

(v) Their planar projections have low average angular and linear distortion.

In our original paper [3], while preserving features (i)–(v), we extended the HEALPix DGGS from a purely spherical solution suitable for astronomical use to:

(vi) Ellipsoids of revolution such as the WGS84 ellipsoid, which is more suitable for terrestrial use.

Finally we elaborated on the work of [7, Section 3.1], by rearranging the underlying HEALPix map projection into what we called the rHEALPix map projection and built a DGGS for ellipsoids of revolution based upon it, which we called the rHEALPix DGGS. The rHEALPix DGGS also has features (i)–(vi) and additionally:

(vii) Its base polyhedron—the cube—is a regular polyhedron.

(viii) Its planar projection consists of horizontal-vertical aligned nested square grids, which are widely used in remote sensing and in environmental modelling.

Features (i)–(iii) and (vii) are sufficient to meet the cell geometry requirements of the proposed OGC DGGS standard, and features (i), (iv) and (viii) significantly assist in designing an indexing system with efficient conversion to and from latitude longitude coordinates, which we describe below.

2. Defining the rHEALPix DGGS

The rHEALPix DGGS can be thought of as a mapping of an ellipsoid of revolution onto a regular polyhedron, namely a cube, see figure 1, followed by a symmetric hierarchical partitioning of the polyhedral faces along with a choice of nuclei, followed by the inverse mapping of the result back onto the ellipsoid, see figure 2. It is therefore an example of a cubic geodesic DGGS [2]. The rHEALPix DGGS and its associated mathematics have been fully described by us in [3].

To support transfer of data into and out of both rHEALPix and HEALPix we have implemented the mathematics for both projections in the PROJ.4 cartographic library (from release 4.8), allowing us to use standard raster manipulation tools. Using the PROJ.4 library as a link to traditional GIS libraries provides a fundamental building block for satisfying the proposed OGC DGGS Standard's requirement for provision of operations for assigning data to and retrieving data from cells.
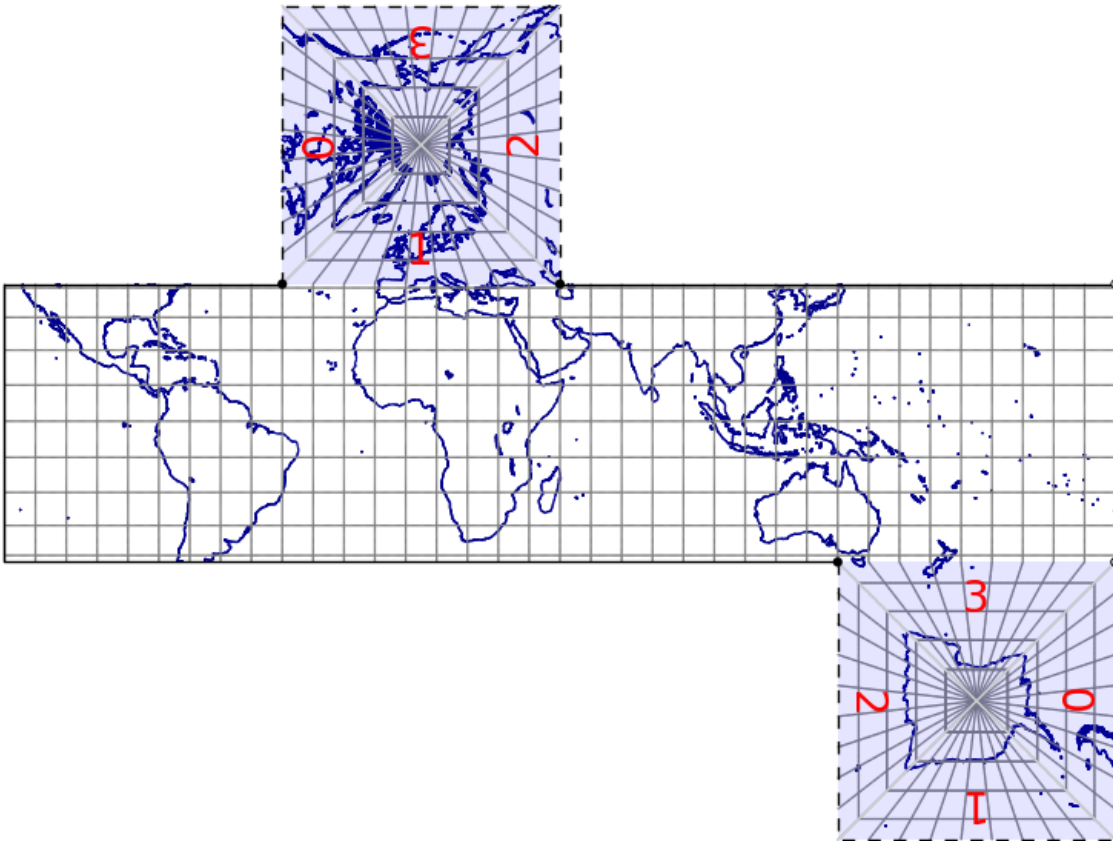
Figure 1. The (1,3)-rHEALPix projection of the WGS84 ellipsoid with a 10° graticule. To illustrate some (r)HEALPix projection options, the prime meridian has been shifted to 50° east longitude to move all cube's corners into water, Arctic polar triangle boundary cuts to less densely populated regions, and rearranged polar triangles to the (1,3) position to optimise Europe and New Zealand views.
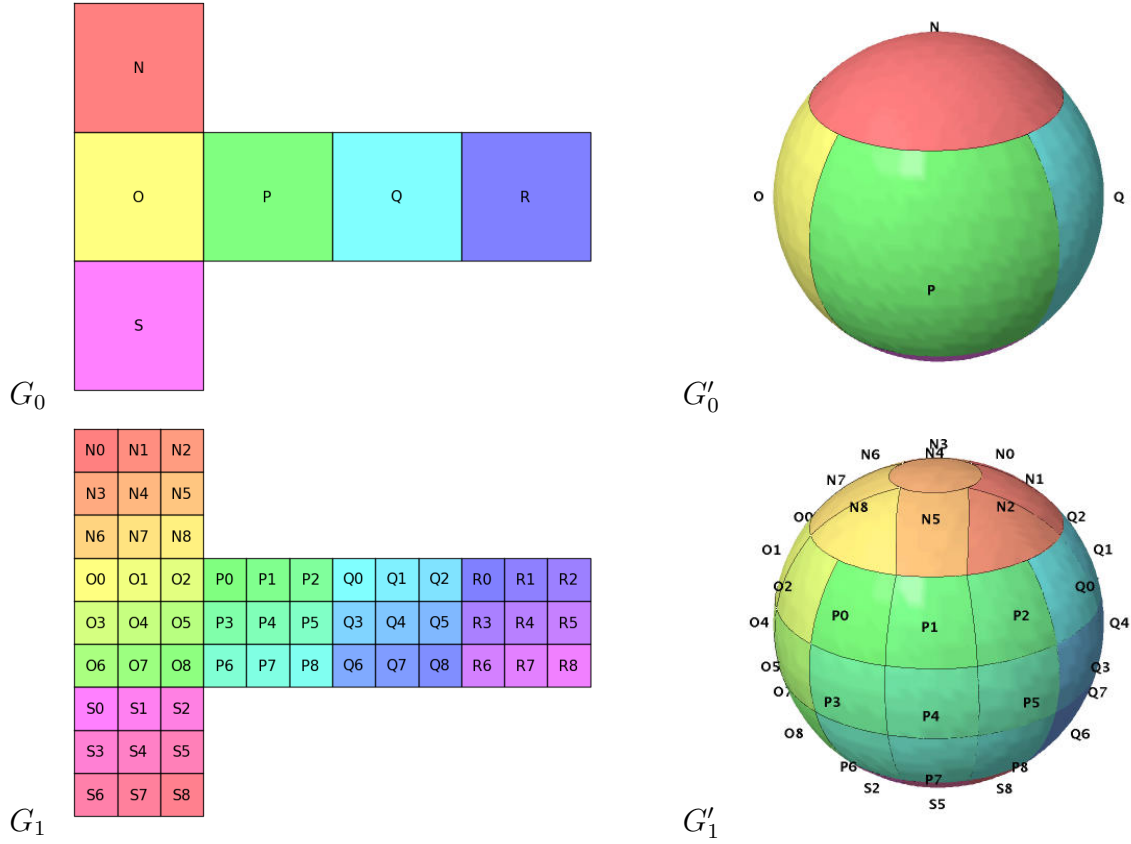
Figure 2. The first two planar and ellipsoidal grids for the (0, 0)-rHEALPix map projection with $N_{side} = 3$ and cells labelled with their IDs.

From here forward we focuses on a (0,0)-rHEALPix DGGS with $N_{side} = 3$, producing aligned hierarchies with the smallest change of area between resolutions. Table 1 provides cell dimensions for a selection of resolutions using this DGGS.

Resolution	Cell area (m^2)	Nominal cell edge	Number of cells
0	85.11E+12	9,225 km	6
1	9.456E+12	3,075 km	54
2	1.051E+12	1,025 km	486
..			
11	2,712	54 m	188.3E+09
12	301	18 m	1,694.6E+09
13	33.5	6 m	15,251E+09
14	3.72	2 m	137.26E+12

Table 1. Approximate cell area, nominal cell edge length, and number of cells for a selection of grid resolutions of the $N_{side} = 3$ WGS84 ellipsoidal rHEALPix grid hierarchy.

3. Constructing cell identifiers

For reference we assign unique names to the cells of the planar hierarchy G that are shared with the corresponding cells of the ellipsoidal hierarchy G'. These names echo the tree structure of the hierarchies and are defined as follows.

Let Γ be the set of all strings beginning with one of the letters $N, O, \ldots, S$ and followed by zero or more of the integers $0, 1, \ldots, 3^2 - 1$, so that Γ forms a regular language $(N|O|\cdots|S)(0|1|\cdots|8)^*$. The identifier (ID) of a cell c from G, denoted c, is the string in Γ defined recursively as follows. Assign the cells of G_0 the IDs N, O, P, Q, R, S from top to bottom and left to right. In particular, cell N contains the projection of the north pole and cell S contains the projection of the south pole. For each resolution i cell with ID s assign its 9 resolution $i + 1$ child cells the IDs $s0, s1, \ldots, s8$ following the path of a Z space filling curve from top to bottom and left to right; see figure 3(a).

(a)   (b)
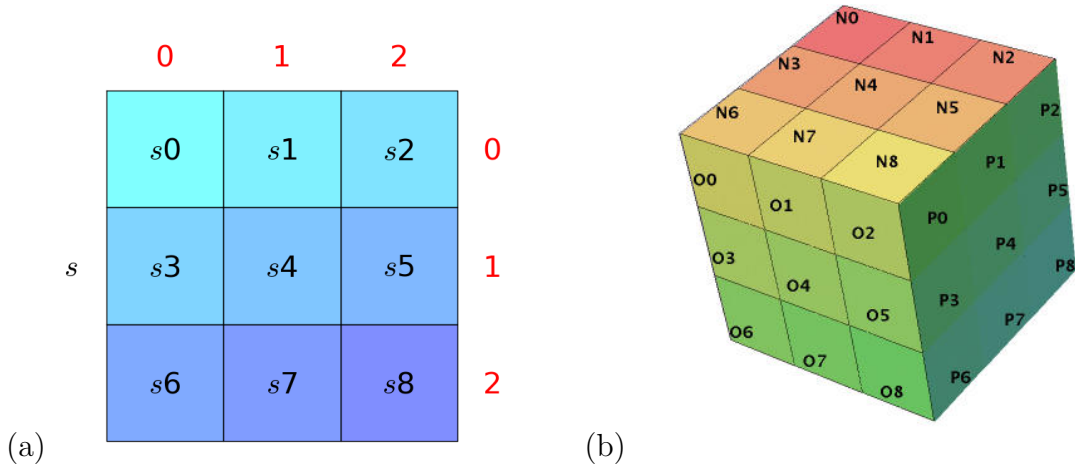
Figure 3. (a) An $N_{\mathrm{side}} = 3$, resolution i, rHEALPix planar cell with ID s showing its $N_{\mathrm{side}}^2 = 9$ resolution $i + 1$ subcells labelled with their IDs. Their row and column numbers are highlighted in red. (b) The cell adjacency structure of G_1 for the $(0, 0)$-rHEALPix map projection with $N_{\mathrm{side}} = 3$.

The proposed OGC DGGS Standard specifies that cell identifiers must uniquely identify cells across all resolutions, and that cell IDs are defined with an addressing method that uses at least one of hierarchy, space filling curve, coordinate or encoded address. The cell identifiers described here are unique with both hierarchical and space filling properties.

4. Cell identifiers and their topological relationships

The proposed OGC DGGS Standard proposes that DGGS should support the DE-9IM topological relationships. The cell IDs described above have a number of properties that allow us to compute both cell adjacency and DE-9IM topological relationships very efficiently directly from the cell ID without recourse to latitude longitude coordinates.

Adjacency is computed by considering each numeric digit as a $base_3$ tuple whose two digits are the row and column numbers within the parent - highlighted in red in figure 3(a). The cell IDs of adjacent cells at any chosen resolution can be found using $base_3$ fraction maths on rows and columns. Just as in $base_{10}$ maths to add 0.010 to a $base_{10}$ fraction, we only need to add 1 to the second digit past the decimal point, so we can use an equivalent construct in $base_3$ for row and column maths at any resolution by simply operating on the digit representing the desired resolution.

The cell adjacency structure of the ellipsoidal hierarchy G' induces a cell adjacency structure on the planar hierarchy G that can be visualized by folding G_0 into a cube along its cell edges, and this is how the rHEALPix DGGS can be construed as a geodesic DGGS; see figure 3(b). The row and column neighbours of $O2$ are $N8$, $P0$, $O5$, and $O1$. We need a mechanism to traverse cells from face to face as well as within faces. To accomplish this we treat faces as integers and assigned them the following (row_3, $column_4$) integer pairs N$(0_3, 0_4)$, O$(1_3, 0_4)$ P$(1_3, 1_4)$, Q$(1_3, 2_4)$, R$(1_3, 3_4)$, S$(2_3, 0_4)$. With this arrangement $base_3$ maths will result in the correct north-south face transitions between the N–O, and O–S face pairs. Similarly $base_4$ maths on the column integers will ensure correct east-west transitions between the equatorial face pairs of O–P, P–Q, Q–R and R–O. While it may seem strange to mix bases within the same number, these are strings not numbers so we can treat each digit the way we need to.

Returning to our initial example $O2$ in row,column notation is $(1_3.0_3, 0_4.2_3)$. Traversing from $O2$ to $O1$ is achieved by subtracting $.1_3$ from $O2$'s column and from $O2$ to $O5$ by adding $.1_3$ to $O2$'s row. i.e. $(1_3.0_3, 0_4.2_3 - .1_3)$ evaluates to $(1_3.0_3, 0_4.1_3) = O1$ and $(1_3.0_3 + .1_3, 0_4.2_3)$ evaluates to $(1_3.1_3, 0_4.2_3) = O5$. Similarly adding $.1_3$ to $O2$'s column we get $(1_3.0_3, 0_4.2_3 + .1_3)$ which evaluates to $(1_3.0_3, 1_4.0_3) = P0$ and subtracting $.1_3$ from $O2$'s row we get $(1_3.0_3 - .1_3, 0_4.2_3)$ which evaluates to $(0_3.2_3, 0_4.2_3) = N8$.

Traversing between faces N or S and P, Q or R requires an additional step to account for the different geometries of the planar vs ellipsoidal grids. This is most simply achieved by constructing a translation logic table from invalid ID's that are generated by the above maths to valid ones. Without giving the complete solution, note in figure 3(b) that $P0$, $P1$ and $P2$ have as their neighbours $N8$, $N5$ and $N2$ respectively. It turns out that at the junction of the N and P faces the pattern $0, 1, 2$ recursively translates to the pattern $8, 5, 2$ at every resolution so this logic can be readily coded. Similarly recurring patterns occur for the other N or S and P, Q or R transitions.

To compute DE-9IM topological relationships between cells we will often compare two IDs to determine if their most significant digits starting with $(N|O|\cdots|S)$ match, and then examine the properties of the rightmost digits that differ. We will refer to the sequence of matching digits as $ID_{\text{common parent}}$ and the two strings of remaining digits as each cell's $ID_{\text{relative child}}$. For cells $O234546$ and $O234787$ their $ID_{\text{common parent}}$ is $O234$ and their $ID_{\text{relative child}}$ are 546 and 787 respectively.

We will also need to define three types of cell neighbour:

(i) A cell's row-neighbours are the set of cells at the same resolution, who's rows are the same and columns differ by 1. To account for transitions between the N or S

and P, Q or R faces 'differ by 1' is used in this and the next two as shorthand to include the additional translational step required to traverse between the faces,

(ii) A cell's column-neighbours are the set of cells at the same resolution, who's columns are the same and rows differ by 1,

(iii) A cell's diagonal-neighbours are the set of cells at the same resolution, who's rows and columns differ by 1,

The DE-9IM topological relationships are usually referred to by the operators: equals, within, contains, touches, disjoint, covers, coveredBy, intersects, crosses and overlaps.

The following list demonstrates how each of these relationships can be determined by manipulating cell IDs using the constructs we have described above.

(i) Cell A and cell B are equal if they have an $ID_{\text{common parent}}$ and both A and B's $ID_{\text{relative child}}$ strings are null,

(ii) If cell A and cell B share an $ID_{\text{common parent}}$ and B's $ID_{\text{relative child}}$ string is null then A is within B and B contains A,

(iii) If cell A is within cell B then A touches B on a edge if A's $ID_{\text{relative child}}$ string satisfies $(0|1|2)^*|(2|5|8)^*|(0|3|6)^*|(6|7|8)^*$. A touches B on a corner (ie two edges) if A's $ID_{\text{relative child}}$ string satisfies $(0)^*|(2)^*|(6)^*|(8)^*$. For example $O234787$ touches $O234$ on an edge and $O234666$ touches $O234$ on a corner.

(iv) If length A is greater than length B and A NOT within cell B then A touches B on an edge if at least one of A's row- or column- neighbours is within B. A touches B on a point if one A's diagonal-neighbours is within B,

(v) Cell A and cell B are disjoint if (A NOT equal B) AND (A NOT within B) AND (A NOT contain B) AND (A NOT touches B),

(vi) A covers B only if A contains B, and A is coveredBy B only if A is within B,

(vii) We note that because of the way cells are defined, no single cell can intersect, cross, or overlap another cell.

Irregular shapes may be represented geometrically by lists of cell IDs, and any level of precision can be represented by including cells of appropriately small size. It follows that the above tests may be extended to two irregular shapes by extending the above comparisons to comparisons of the members in two ID's lists, each representing one of the two irregular shapes. While this can be inefficient in a traditional serial processing environment, it is ideally suited to modern parallel processing environments.

5. Conclusion

We have revisited the original definition of the rHEALPix DGGS and identified those geometric and cell identifier properties that make it conformant with the new Discrete Global Grid System Core Standard being proposed by the Open Geospatial Consortium.

We have also demonstrated that rHEALPix DGGS cell IDs can be used directly to efficiently compute each of the DE-9IM topological relationships, which have emerged as a part of the proposed standard, but weren't explicitly part of earlier work on DGGS. This comparison of the proposed standard with an existing DGGS is a useful first step in showing how the standard can be applied.

References

[1] Kimerling J A, Sahr K, White D and Song L 1999 *Cartography and Geographic Information Science* **26** 271–287

[2] Sahr K, White D and Kimerling A J 2003 *Cartography and Geographic Information Science* **30**(2) 121–134

[3] Gibb R G, Raichev A and Speth M 2013 The rHEALPix Discrete Global Grid System URL http://dx.doi.org/10.7931/J2D21VHM

[4] Purss M, Gibb R and Peterson P 2014 *Discrete Global Grid Systems SWG* URL http://www.opengeospatial.org/projects/groups/dggsswg

[5] Goodchild M F 1994 Geographical grid models for environmental monitoring and analysis across the globe (panel sesssion). *Proceedings of GIS/LIS '94 Conference* (Phoenix, Arizona)

[6] Gorski K M, Hivon E, Banday A J, Wandelt B D, Hansen F K, Reinecke M and Bartelman M 2005 *The Astrophysical Journal* **622** 759

[7] Calabretta M R and Roukema B F 2007 *Monthly Notices of the Royal Astronomical Society* **381** 865–872 ISSN 1365-2966 URL http://dx.doi.org/10.1111/j.1365-2966.2007.12297.x

Acknowledgments