EMBEDDING REALISTIC SURVEYS IN SIMULATIONS THROUGH VOLUME REMAPPING

JORDAN CARLSON AND MARTIN WHITE
Department of Physics, University of California, Berkeley, CA 94720, USA
Received 2010 February 19; accepted 2010 August 19; published 2010 September 20

ABSTRACT

Connecting cosmological simulations to real-world observational programs is often complicated by a mismatch in geometry: while surveys often cover highly irregular cosmological volumes, simulations are customarily performed in a periodic cube. We describe a technique to remap this cube into elongated box-like shapes that are more useful for many applications. The remappings are one-to-one, volume-preserving, keep local structures intact, and involve minimal computational overhead.

Key words: large-scale structure of universe – methods: numerical

Online-only material: color figures

1. INTRODUCTION

Numerical simulations have become an indispensable tool in modern cosmological research used for investigating the interplay of complex physical processes, studying regimes of a theory which cannot be attacked analytically, generating high-precision predictions for cosmological models, and making mock catalogs for the interpretation and analysis of observations. Such simulations traditionally evolve the matter distribution in periodic cubical volumes, which neatly allow them to approach the homogeneous Friedmann solution on large scales. The use of a periodic volume also allows the long-range force to be easily computed by fast Fourier transform methods in many popular algorithms (e.g., particle-mesh, particle-particle-particle-mesh, or tree-particle-mesh algorithms).

Surveys, on the other hand, often cover cosmological volumes that are far from cubical in shape, and making a mock catalog which includes the full geometrical constraints of the observations is difficult. One approach is to simulate a sufficiently large volume that the survey can be embedded directly within the cube, but this often means that large parts of the computational domain are unused. An alternative is to trace through the cube across periodic boundaries so as to generate the desired depth, with various rules for avoiding replication or double-counting of the volume (if desired). A third approach is to run a simulation in a non-cubical geometry. If the side lengths are highly disproportionate this can lead to its own numerical issues, and in addition it makes it difficult to reuse a given simulation for many applications.

In this paper, we present a new solution to this problem which allows one to embed a hypothetical survey volume inside a cosmological simulation while limiting wasted volume and artificial correlations. The method is based on the simple observations that (1) a cube with periodic boundary conditions is equivalent to an infinite three-dimensional (3D) space with discrete translational symmetry and (2) the primitive cell for such a space need not be a cube. We show that one may take the primitive cell to be a cuboid1 of dimensions $L_1 \times L_2 \times L_3$ for a discrete but large choice of values $L_1 \geq L_2 \geq L_3$. The possible choices, subject to the constraint $L_1 L_2 L_3 = 1$, are illustrated in Figure 1.

Our approach leads to a one-to-one remapping of the periodic cube which keeps structures intact, does not map originally distant pieces of the survey close together, and uses no piece of the volume more than once. It complements existing techniques for generating mock observations (e.g., Blaizot et al. 2005; Kitzbichler & White 2007) or for ray-tracing through simulations (e.g., White & Hu 2000; Vale & White 2003; Hilbert et al. 2009; Fullana et al. 2010). Though not ideal in all cases, our remapping procedure may be seen as a general-purpose alternative that neatly skirts many of the complications involved with previous methods.

We begin in Section 2 with a mathematical description of the remapping, and explain what choices of dimensions are possible. In Section 3, we describe how to implement this remapping numerically. We present a few useful examples in Section 4, and conclude in Section 5 with a discussion of some of the advantages and limitations of this method.

2. MATHEMATICAL DESCRIPTION

Consider a unit cube with periodic boundary conditions. By tiling copies of this cube in all directions, each point $(x, y, z) \in \mathbb{R}^3$ corresponds to a point in the canonical unit cube $[0, 1]^3$ via the map

$$(x, y, z) \mapsto (x \mod 1, y \mod 1, z \mod 1),$$

where "$x \mod 1"$ is just the fractional part of $x$ if $x \geq 0$, or one minus this fractional part for $x < 0$. Since we can identify each point in $\mathbb{R}^3$ with a point in $[0, 1]^3$, any region in space corresponds to a sampling of the unit cube. We are primarily interested in bijective samplings, i.e., regions that according to the above equivalence cover each point of the unit cube once and only once. We call such a sampling a remapping of the unit cube. Equivalently, a remapping may be thought of as a partitioning of the unit cube into disjoint regions, which are then translated by integer offsets and glued back together along periodic boundaries.

A general class of such remappings may be constructed by applying shear transformations to the unit cube. We first illustrate the idea in two dimensions (2D), where the transformations may be easily visualized, and then state the appropriate generalization to 3D. Start with a continuous field laid down within the unit square, which may be thought of as a parallelogram spanned by the vectors $u_1 = (1, 0)$ and $u_2 = (0, 1)$. We suppose

---

1 To be precise, by cuboid we mean here a rectangular parallelepiped, i.e., a parallelepiped whose six faces are all rectangles meeting at right angles. For practical purposes, a cuboid is simply a box that is not a cube.
Figure 2. Possible dimensions $L_1 \times L_2 \times L_3$ for cuboid remappings, subject to the conditions $L_1 \geq L_2 \geq L_3$, $L_1 L_2 L_3 = 1$, and $L_1 < 7$. The choice used for the Stripe 82 remapping described in Section 4 is indicated by a red circle.

(A color version of this figure is available in the online journal.)

the field to be periodic, and tile copies throughout the plane (see Figure 2(a)). Now imagine taking the unit square and shearing it along its top edge, leaving the underlying field in place (Figure 2(b)). The result is a new parallelogram defined by the vectors

$$\mathbf{u}_1 = (1, 0), \quad \mathbf{u}_2 = \mathbf{u}_2 + m \mathbf{u}_1 = (m, 1),$$

where $m$ is some real scalar that controls the extent of the shear. Since shear transformations are area-preserving, this parallelogram has unit area. In fact, since $\mathbf{u}_1$ is a lattice vector (i.e., a vector of translational invariance), this parallelogram covers each point of our field once and only once, and hence defines a valid remapping.

Now observe that if $m$ is an integer, then $\mathbf{u}_2 = (m, 1)$ will also be a lattice vector. In that case we again have a parallelogram with edge vectors $\mathbf{u}_1$ and $\mathbf{u}_2$ that are both lattice vectors. We can now shear this parallelogram along its right edge (Figure 2(c)), giving a new parallelogram with edge vectors

$$\mathbf{u}_1' = \mathbf{u}_1 + n \mathbf{u}_2 = (1 + mn, n), \quad \mathbf{u}_2' = (m, 1).$$

Once again this parallelogram covers each point of our field exactly once, and if $n$ is an integer then $\mathbf{u}_1'$ and $\mathbf{u}_2'$ are both lattice vectors. We may repeat this process, applying integer shears alternately to the top and right edges of the parallelogram. In general, any pair of integer-valued 2D vectors $\mathbf{u}_1$ and $\mathbf{u}_2$ that span a parallelogram of unit area may be obtained by such a sequence of shear transformations. Mathematically, the condition that $\mathbf{u}_1$ and $\mathbf{u}_2$ span unit area is simply

$$\det \begin{pmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{pmatrix} = 1,$$

i.e., $\mathbf{u}_1$ and $\mathbf{u}_2$ must be the rows of an integer-valued $2 \times 2$ matrix with unit determinant.

While remapping the unit square into a parallelogram may be useful in certain cases, parallelograms are generally too awkward to be useful. Instead, given lattice vectors $\mathbf{u}_1$ and $\mathbf{u}_2$ spanning unit area, we may apply one last shear to “square up” the parallelogram into a rectangle (Figure 2(d)). Explicitly, we let

$$\mathbf{e}_1 = \mathbf{u}_1, \quad \mathbf{e}_2 = \mathbf{u}_2 + \alpha \mathbf{u}_1,$$

and choose $\alpha$ so that $\mathbf{e}_1$ and $\mathbf{e}_2$ are orthogonal. The rectangle defined by $\mathbf{e}_1$ and $\mathbf{e}_2$ now covers each point of the unit square exactly once (Figure 2(e)).

The generalization to 3D is straightforward. By applying integer shears to the faces of the cube, we may obtain any parallelepiped with edges given by integer vectors $\mathbf{u}_1$, $\mathbf{u}_2$, $\mathbf{u}_3$ satisfying

$$\det \begin{pmatrix} u_{11} & u_{12} & u_{13} \\ u_{21} & u_{22} & u_{23} \\ u_{31} & u_{32} & u_{33} \end{pmatrix} = 1.$$

Again this space of possibilities corresponds to the space of integer-valued $3 \times 3$ matrices with unit determinant. We apply two final shears to square up this parallelepiped into a cuboid, by choosing coefficients $\alpha$, $\beta$, and $\gamma$ such that

$$\mathbf{e}_1 = \mathbf{u}_1, \quad \mathbf{e}_2 = \mathbf{u}_2 + \alpha \mathbf{u}_1, \quad \mathbf{e}_3 = \mathbf{u}_3 + \beta \mathbf{u}_1 + \gamma \mathbf{u}_2$$

are mutually orthogonal. This gives a remapping of the unit cube into a cuboid with side lengths $L_i = |\mathbf{e}_i|$, satisfying $L_1 L_2 L_3 = 1$. Moreover, since $\mathbf{e}_1$ is still a lattice vector, this cuboid is periodic across the faces perpendicular to this vector.

The dimensions of the final cuboid, as well as the orientation of the cuboid with respect to the axes of the unit cube, are completely determined by the edge vectors $\mathbf{e}_1$, $\mathbf{e}_2$, and $\mathbf{e}_3$, which are in turn uniquely determined by the choice of an integer-valued $3 \times 3$ matrix with unit determinant. Such matrices may be generated easily by brute force computation, and compiled into a list of possible cuboid remappings. The dimensions of some of the allowed remappings are illustrated in Figure 1 as a scatter plot in $L_1$ and $L_2$, with $L_3$ given implicitly by the unit volume condition $L_1 L_2 L_3 = 1$.

(A color version of this figure is available in the online journal.)
Figure 3. Regions in the unit cube and how they “unfold” to produce the wide, thin geometry shown in Figure 4.

(A color version of this figure is available in the online journal.)

3. NUMERICAL ALGORITHM

The goal of our remapping procedure is to provide an explicit bijective map between the unit cube and a cuboid of dimensions $L_1 \times L_2 \times L_3$. We will refer to points in the unit cube by their simulation coordinates $x \in [0, 1]^3$, and their remapped positions in a canonical, axis-aligned cuboid by remapped coordinates $\mathbf{r} \in [0, L_1] \times [0, L_2] \times [0, L_3]$.

The reverse map $\mathbf{r} \mapsto x$ is simple. The edge vectors $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ discussed previously describe how to embed an oriented cuboid within an infinite tiling of the unit cube. Let $\hat{n}_i = \mathbf{e}_i/L_i$ be unit vectors along these edges. Then given $(r_1, r_2, r_3) \in [0, L_1] \times [0, L_2] \times [0, L_3]$, the point

$$\mathbf{p} = r_1 \hat{n}_1 + r_2 \hat{n}_2 + r_3 \hat{n}_3$$

lies within this oriented cuboid, and this maps to a point in the unit cube according to Equation (1).

The forward map $\mathbf{x} \mapsto \mathbf{r}$ is slightly more complicated. Consider again the oriented cuboid with edge vectors $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$, embedded within an infinite tiling of the unit cube. Each tile (i.e., each replication of the unit cube) may be labeled naturally by an integer triplet $\mathbf{m} \in \mathbb{Z}^3$, where the canonical unit cube has $\mathbf{m} = (0, 0, 0)$. The intersection of the cuboid with a tile is called a cell, of which only a finite number will be non-empty. Each non-empty cell is uniquely labeled by the triplet $\mathbf{m}$. (The four cells of the 2D example from the previous section are indicated in Figure 2(e).)

As a geometrical object, each cell is just a convex polyhedron bounded by 12 planes: the 6 faces of the tile and the 6 faces of the cuboid. A plane may be parameterized by real numbers $(a, b, c, d)$, so that a point $(x, y, z)$ lies inside, outside, or on the plane depending on whether the quantity $ax + by + cz + d$ is less than, greater than, or equal to zero. Thus, to test whether a point belongs to a cell, we need only check if it lies inside all the planes that bound it. When translated spatially by a displacement $-\mathbf{m}$, each cell defines a region within the canonical unit cube. The collection of all such cells defines a partitioning of the unit cube, with each point in $[0, 1]^3$ being covered by exactly one cell. Our algorithm for the forward map $\mathbf{x} \mapsto \mathbf{r}$ then may be summarized as follows.

1. Determine which cell contains the point $\mathbf{x}$ (using point-plane tests).
2. Let $\mathbf{p} = \mathbf{x} + \mathbf{m}$ be the corresponding point in the oriented cuboid.
3. Define $\mathbf{r} \in [0, L_1] \times [0, L_2] \times [0, L_3]$ by $r_i = \mathbf{p} \cdot \hat{n}_i$.

4. EXAMPLES AND APPLICATIONS

The mappings described above can be used in myriad ways. The algorithm is fast enough that it can be used for on-the-fly analysis while a simulation is running, it introduces little overhead in walking a merger tree (allowing complex light-cone outputs to be built during, e.g., the running of a semi-analytic galaxy formation code), or it can be applied in post-processing to a static time output of a simulation to alter the geometry. In this section, we give a (very) few illustrative examples.

As it is so widely known we shall use as our fiducial volume the Millennium simulation (Springel et al. 2005), which was performed in a cubical volume of side length $500 \, h^{-1}$ Mpc. We begin by asking how we could embed a very wide-angle survey, such as the equatorial stripe (“Stripe 82”) of the Sloan Digital Sky Survey, in such a volume. Stripe 82 is $100^\circ$ wide and $2.5^\circ$ in height. Using the Millennium simulation cosmology, the volume within Stripe 82 out to $z \simeq 0.45$ equals the total volume within the Millennium simulation. At this depth the stripe is $1.2 \, h^{-1}$ Gpc in the line-of-sight direction, $1.9 \, h^{-1}$ Gpc transverse, but only $50 \, h^{-1}$ Mpc thin.

We can map the cubical volume of the simulation into this geometry using the transformation circled in red in Figure 1, which has $L_1 = 3.7417$, $L_2 = 2.4349$, and $L_3 = 0.1098$. A view of this transformation is given in Figure 3, where the regions within the cube are shown “unfolded” to produce a long and wide, but thin, domain into which we can embed the Stripe 82 geometry. Figure 4 shows a light cone produced from the

Figure 4. Mock catalog from the Millennium simulation remapped into the equatorial stripe of the Sloan Digital Sky Survey. The “stripe” is a thin rectangle on the sky of angular dimensions $100^\circ \times 2.5^\circ$, shown here “from above” extending out to redshift $z \simeq 0.45$, or $R \simeq 1.2 \, h^{-1}$ Gpc comoving. We show only galaxies brighter than $R = -20.5$ to avoid saturating the figure.

http://www.sdss.org
outputs of the Millennium simulation, in the Stripe 82 geometry. Specifically we show all galaxies brighter than $R = -20.5$ (to avoid saturating the figure) and with $0 < z < 0.45$ from the catalog of De Lucia & Blaizot (2007).

Another frequently encountered situation is a “pencil beam” survey which is much longer in the line-of-sight direction than either of the (approximately equal) transverse directions. Let us consider, for example, a survey which aims to reach $z \sim 3$, or about $5 \, h^{-1} \text{Gpc}$ in the Millennium simulation cosmology. Among the many possible choices at our disposal we find a remapping with $(L_1, L_2, L_3) = (5025, 179, 139) \, h^{-1} \text{Mpc}$, which could encompass a survey of angular dimensions $2\times 0 \times 1^\circ$ out to $z \sim 3.5$. To probe even earlier epochs we could choose a remapping with $(L_1, L_2, L_3) = (7106, 145, 121) \, h^{-1} \text{Mpc}$, which allows a $1^\circ \times 1^\circ$ survey out to $z \sim 10$.

Figure 5 illustrates the abundance of possible remappings for these types of highly elongated geometries. There, as in Figure 1, we show a scatter plot of possible cuboid remappings, but now with the long axis $L_1$ converted to a maximum visible redshift using the Millennium cosmology, and the short axis $L_3$ converted to a minimum opening angle $\theta_1 = 2 \arctan(L_3/2L_1)$. It can be seen that even for very high redshifts there are many interesting remappings to choose from. Users focusing on a particular survey would most likely proceed by identifying the minimum opening angle they need, then selecting the remapping that allows for the greatest visible redshift.

5. DISCUSSION

As surveys become increasingly complex and powerful and the questions we ask of them become increasingly sophisticated, mock catalogs and simulations which can mimic as much as possible the observational non-idealities become increasingly important. Angulo & White (2010) have shown that simulations of one cosmology can be rescaled to approximate those of a different cosmology. We have introduced a remapping of periodic simulation cubes which allows one simulation to take on the characteristics of many different observational geometries. The use of such techniques enhances the usefulness of cosmological simulations, which often involve a large investment of community resources.

The methods we introduced in this paper lead to one-to-one remappings of the periodic cube which keep structures intact, do not map originally distant pieces of the survey close together, and use no piece of the volume more than once. The remapping can be done extremely quickly, meaning it can be included in almost any analysis tool with negligible overhead.

Remapping the simulation geometry is, however, not without its limitations. First and foremost, although the target geometry may have sides much longer than the original simulation, the structures will not contain the correct large-scale power since it was missing from the simulation to begin with. This problem becomes less acute as the original simulation volume becomes a fairer representation of the universe. Second, if the target geometry is too thin it is possible for points which are far apart in the survey to come from points close together in the simulation volume, leading to spurious correlations. These are analogous to the artificial correlations between survey “sides” that occur when it is embedded in a periodic cube. Simply excluding a boundary layer from the remapped volume can tame such correlations.

In addition to the description of the algorithm in this paper, we have made Python and C++ implementations of the remappings, along with further examples and animations, publicly available at http://mwhite.berkeley.edu/BoxRemap.

We thank the members of the DES simulation working group for valuable feedback and encouragement. This work was supported by NASA and the DoE.

REFERENCES

Angulo, R. E., & White, S. D. M. 2010, MNRAS, 405, 143
Blaizot, J., Wadadekar, Y., Guiderdoni, B., Colombi, S. T., Bertin, E., Bouchet, F. R., Devriendt, J. E. G., & Hatton, S. 2005, MNRAS, 360, 159
De Lucia, G., & Blaizot, J. 2007, MNRAS, 375, 2
Fullana, M. J., Arnau, J. V., Guiderdoni, B., Couchman, H. M. P., & Saez, D. 2010, ApJ, 712, 367
Hilbert, S., Hartlap, J., White, S. D. M., & Schneider, P. 2009, A&A, 499, 31
Kitzbichler, M. G., & White, S. D. M. 2007, MNRAS, 376, 2
Springel, V., et al. 2005, Nature, 435, 629
Vale, C., & White, M. 2003, ApJ, 592, 699
White, M., & Hu, W. 2000, ApJ, 537, 1

Figure 5. Possible remappings appropriate for “pencil beam” surveys, which have a narrow angular field of view but extend to high redshift. See the text for details. The two pencil beam examples mentioned explicitly in the text in Section 4 are circled in red. The figure excludes low-redshift and narrow-angle remappings that would otherwise saturate the lower left portions of the plot. (A color version of this figure is available in the online journal.)