The fractal dimension of space

Subhash Kak
April 5, 2020

Abstract. This paper provides a new theoretical basis for the existence of fractal phenomena and scale invariance that is known to occur for diverse physical systems. Rather than see these emerge from the nature of interactions in matter, we apply information theoretic considerations to show that space has a dimension of value $e$ which means that it is three-dimensional only approximately, and the information in the three-dimensional representation is off from the optimal value of $e$ by 0.006 percent. One surprising consequence of $e$-dimensions is the Cantor set along with its generalization to two and three dimensions that are often the starting points to demonstrate characteristics of self-similarity and scale invariance. Some implications of $e$-dimensionality for a model universe are investigated.

Introduction

Our understanding of physical reality is based on certain assumptions about space and time. Space is the expanse in which objects and events have relative position and direction, and it is variously seen as an entity, a relationship between entities, or part of a conceptual framework. Although our intuitions about a three-dimensional space are built up on our everyday terrestrial experience, we apply it to scales that range from the smallest structure to the largest span some seventeen orders of magnitude. As a projection of spacetime, space is the “container” in which the motions and interactions of the objects are accounted for.

Although, locally, space appears to be three-dimensional, it may not be exactly so. This has been investigated within the framework of critical phenomena, many-body systems, quantum field theory, large-scale structure of the universe, and other applications. The proposed theories dealing with deviation from three-dimensionality were partly motivated by the fact that some key quantities like the Gaussian integral and the Laplace operator retain their form when the integer value of dimensions is replaced by a non-integer value. Stillinger provided an axiomatic basis for rules for computation in spaces with non-integer dimensions and further showed that wave mechanics and statistical mechanics readily transform into non-integer dimension spaces. His axioms showed that three mutually perpendicular lines provide neither a necessary nor a sufficient condition for dimension to equal three.

Non-integer dimensional spaces $D = 4 - \varepsilon$ of space-time and $\varepsilon$-expansion of dimensional continuation are actively used in the theory of critical phenomena and phase transitions.
in statistical physics and quantum theory\textsuperscript{11}. In quantum theory, the divergences are parameterized and then removed by renormalization to obtain finite physical values. In statistical mechanics, scale invariance is a feature of phase transitions. Near a phase transition or critical point, fluctuations occur at all length scales, and thus one looks for an explicitly scale-invariant theory to describe the phenomena. Scale-invariant statistical field theories are formally very similar to scale-invariant quantum field theories.

Current theories look at the distribution of structures in space as a consequence of matter interactions. But there appear limits to how much of the nature of reality can be accessed by mathematical analysis of observational data\textsuperscript{12,13} because of paradoxes of logic\textsuperscript{14} as well as others related to interpretation\textsuperscript{15}.

This paper looks at the problem from a new perspective and propose that there is an additional component to the information associated with the distributions that reflects intrinsic structure of space. This idea, which has not been investigated before, complements observational data, and we propose this as the logical ground on which our understanding of space should rest.

We present a mathematical argument to show that there is a limit to the amount of information that physical dimensions can carry. We show that the optimal dimensionality of space is $e$, which can be approximated by the familiar three dimensions. This approach in which information is fundamental is complementary to determining the structure from empirical laws of physics. We show that the $e$-dimensional space is approximately equivalent to the three-dimensional generalization of the Cantor set, and describe some implications for a model universe.

**Geometry of information**

From a practical perspective, most physical phenomena are studied at different scales and it is assumed that these do not influence each other, so that one need not know molecular structure of a fluid when considering disturbances in it. But there do exist phenomena, such as second-order or continuous phase transitions, where events at many scales are interconnected. Such phenomena are characterized by large correlation length\textsuperscript{2,3}.

A case in point is water heated beyond a pressure of 218 atmospheres and temperature of 647 degrees Kelvin. In the vicinity of the critical point, the physical properties of the liquid and the vapor change dramatically, becoming ever more similar. Liquid water under normal conditions is almost incompressible, and has a low thermal expansion coefficient. These properties change near the critical point, and water becomes compressible and expandable. At higher pressures there is only a single, undifferentiated fluid phase, and water cannot be made to boil no matter how much the temperature is raised.
Near the critical point water develops fluctuations in density at all possible scales. The fluctuations take the form of drops of liquid thoroughly interspersed with bubbles of gas, and there are both drops and bubbles of all sizes from single molecules up to the volume of the specimen. Precisely at the critical point the scale of the largest fluctuations becomes infinite, but the smaller fluctuations are in no way diminished. Any theory that describes water near its critical point must consider the entire spectrum of length scales and, therefore, be representative of the fundamental nature of space and reflect its true dimensions\textsuperscript{10}.

The renormalization group is used for investigating the changes of a physical system under scale transformation. It is intimately related to scale invariance and conformal invariance, symmetries in which a system is characterized by self-similarity. In renormalizable theories, the system at one scale will generally be seen to consist of self-similar copies of itself when viewed at a smaller scale, with different parameters describing the components of the system. These parameters may be variable couplings which measure the strength of various forces, or mass parameters themselves. The components themselves may appear to be composed of more of the self-same components as one goes to shorter distances. Renormalization is an \textit{ad hoc} method of associating structure with scale\textsuperscript{4,5}.

One normally uses general relativity to investigate geometric properties of space and time, or spacetime. It defines the curvature of spacetime as directly related to the energy and momentum of matter and radiation that are present. But that doesn’t address the question of dimensionality of space and, therefore, it shall not be considered any further here. In general relativity, gravitational fields are geometric perturbations in spacetime, and one might ask if dimension itself could serve as a field variable\textsuperscript{10}. The gas clouds of the interstellar medium have a fractal structure, the origin of which has generally been thought to lie in turbulence or in self-gravity\textsuperscript{16,17}. But there is no theory of the origin of such structures and long-range correlations. Fractal structure arising out of space provides a new way of looking at the problem.

\textbf{The main result: the optimal number of dimensions}  
One can think of information dimension of an object M as the amount of information necessary to specify the position of a point in M. It is also related to the representation of information to an appropriate base where the base-e representation was investigated recently\textsuperscript{18}.

\textit{Proposition 1}. The amount of information required to specify a point in a space represents the \textit{information dimension} of that space.
Information dimensions are relevant to both physical and abstract spaces. A solid is three-dimensional because one needs three coordinates to specify any point inside.

Not all physical shapes require integer dimensions. To see this, consider measuring a shape by a cube and then use smaller cubes with the scaling factor of $\varepsilon$, so that if $N$ such smaller cubes are to be used, then we can write:

$$N = \varepsilon^{-D}$$

The dimensionality associated with the shape will then be:

$$D = \lim_{\varepsilon \to \infty} \left( -\frac{\log N}{\log \varepsilon} \right)$$

Now we ask the question of the dimensionality of a general space. If space were $d$-dimensional, we could label the dimensions as 1, 2, 3, … $d$. The probability of the use of each of the $d$ dimensions may be taken to be the same and equal to $1/d$, and the information associated with each dimension is $\log d$.

Clearly, the location information will be greater if the dimensionality is higher. But the increase in information must be squared off against the extra burden entailed by the use of the larger set of dimensions. For two-dimensional space, the information value of each dimension is $\ln 2 = 0.693$ nats (=1 bit); for three-dimensional space, it is 1.099 nats (=1.585 bits); and for ten-dimensional space, it is 2.303 nats (=3.322 bits).

The efficiency of the representation of information per dimension is:

$$E(d) = \frac{\ln d}{d}$$

Its maximum value is obtained by taking the derivative of $E(d)$ and equating that to zero, which yields $d_{optimum} = e = 2.71828\ldots$. In other words:

**Theorem 1.** The optimal number of information dimensions associated with space is $e$.

Table 1 gives the value of $E(d)$ in bits for $d$ ranging from 2 to 10, together with the additional value for the optimum $d=e$:

<table>
<thead>
<tr>
<th>$d$</th>
<th>2</th>
<th>$e$</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>8</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E(d)$ bits</td>
<td>0.500</td>
<td>0.531</td>
<td>0.528</td>
<td>0.500</td>
<td>0.465</td>
<td>0.375</td>
<td>0.331</td>
</tr>
</tbody>
</table>
The efficiency for \( e \) dimensions is 0.531 bits whereas for \( d=3 \) it is 0.528 bits. The next best value coming at the bases 2 and 4 (where it is 0.500 bits). The three-dimensional space is off from the \( e \)-dimensional optimal space by about 0.003, or about 0.6 percent.

One may propose that since our cognitions are based on counting, we associate the nearest integer space of 3 dimensions to space.

The efficiency at \( d=3 \) (three-dimensional space) is superior to that at \( d=2 \) (two-dimensional space) by 5.6%. After this the values decline monotonically as shown in Figure 1.

![Figure 1. Efficiency of dimensions, 2 through 10](image)

In short, the three-dimensional space is nearest to the optimal space of \( e \)-dimensions. The three dimensions represent the closest integer which is also an excellent approximation to the optimal value. Nevertheless, we acknowledge that the meaning of Theorem 1 depends on the philosophical framework that is adopted.

*Proposition 2*. The information dimension of a physical space is its physical dimensionality.

To visualize the \( e \)-dimensional space, one may consider it as shape or an abstract conception that is structured into different projections for small scale and large-scale phenomena. Its relation to the three-dimensional space is most clear at intermediate scales like the ones we encounter in everyday life.

The terrestrial observer will see the large-scale as well as small-scale structures as a continuation of the nature of space at the terrestrial level.

Axiomatic foundations of non-integer spaces have been given before by Stillinger\(^{10}\) and Wilson [3]. In addition to the usual axioms that apply to Euclidean spaces, Stillinger
needed to add two more axioms: one related to topology and another to integration measure. He further proposed that the realness of a fractional space less than 3 could be checked by experiments of sphere-packing but he acknowledged that to carry out such an experiment will not be an easy matter due to extreme constraints on accuracy.

**Theorem 2.** A unit cube in an e-dimensional space has $e^e \approx 15.154 \ldots$ sub-cubes each of side $1/e$.

*Proof.* A unit 2-cube in a 2-dimensional space is a square and the total number of sub-cubes of side $1/2$ is $2^2 = 4$, and a unit 3-cube in a 3-dimensional space has sub-cubes of side $1/3$ that equal $3^3 = 27$; generalizing, we get the result.

The e-dimensional space is smaller than the 3-dimensional space. How it maps into the larger 3-dimensional space, consider how many sub-cubes of e-dimensions $e^e$ can be fitted in a 3-dimensional unit cube.

**Theorem 3.** The number of sub-cubes of side $1/e$ that go into a 3-dimensional unit cube is $20.085 \ldots$

*Proof.* The volume of sub-cubes of side $1/e$ in a 3-dimensional space is $e^{-3}$. Therefore, the number of such sub-cubes that will go into a volume of 1 is $e^3 = 20.085 \ldots$.

Seen from the perspective of ordinary 3-dimensional space, the number of sub-cubes after $n$ iterative operations is $e^{3n}$. Therefore, applying formula (2), we get the value of the dimension to be

$$D = \lim_{\varepsilon \to \infty} (-\frac{\log e^{3n}}{\log 3^n}) = e.$$  

(4)

**An approximation to the e-dimensional space**

Now consider a deterministic model to help with the visualization of an e-dimensional space. We need a scaling transformation by which $e^3 \approx 20$ e-dimensional sub-cubes are seen as a subset of the 27 three-dimensional sub-cubes. This is done iteratively.

Taking a cue from quantum mechanics, we may speak of a *creation operator* that maps space into structure. Specifically, the $e^3 \approx 20$ sub-cubes of the e-space may be mapped by an appropriate iterative transformation in the 3-space. In other words, we need a mapping that takes us from the smaller sub-set of 20 sub-cubes to the larger 3-space of 27 sub-cubes.
Since a cube has six sides, this may be done by any mapping where one dark sub-cube is removed randomly from each of the six sides together with the one at the center. Although it is done uniformly in Figure 2, there is no reason why it cannot be done randomly. The seven extra sub-cubes represent the effect of the creation operator.

![Figure 2. A mapping of 20 e-space sub-cubes into the unit cube](image)

Let $N_n$ be the number of dark boxes, $L_n$ the length of a side of a light sub-cube, and $V_n$ the fractional volume of the dark cubes after the $n$th iteration, then

\[ N_n = 20^n \]  \hspace{1cm} (5)

\[ L_n = \left(\frac{1}{3}\right)^n = 3^{-n} \]  \hspace{1cm} (6)

\[ V_n = \left(\frac{20}{27}\right)^n \]  \hspace{1cm} (7)

The dimension of such an iterative system will be:

\[ D = \lim_{e \to \infty} \left(-\frac{\log N_n}{\log L_n}\right) \]  \hspace{1cm} (8)

\[ = \frac{\ln 20}{\ln 3} = 2.7268 \ldots \]  \hspace{1cm} (9)

which is quite close to $e = 2.71828\ldots$

The difference between the values of $D$ and $e$ is only 0.3% and, therefore, it is a good deterministic model to visualize the $e$-dimensional space. The difference of 0.3% was due to the fact that we used the integer value of 20 rather than the exact value of $e^3 = 20.085 \ldots$

**Iterative construction**

Since the three-dimensional system is almost as efficient as the $e$-dimensional one, one would like to begin with an appropriate one-dimensional set and then generalize that to three dimensions.
It is surprising that the random mapping described above may be derived by the use of the one-dimensional Cantor set, of two kinds of elements that we label *dark* and *light*. One starts with a line segment of unit length that is *dark*, converts the middle third to *light*, then converts the middle thirds from the remaining two dark segments to *light*, and so on.

Formally, the Cantor set at the \( n \)th iteration, \( \mathcal{C}_n \), is:

\[
\mathcal{C}_n = \frac{\mathcal{C}_{n-1}}{3} \cup \left( \frac{2}{3} + \frac{\mathcal{C}_{n-1}}{3} \right), \text{ for } n \geq 1, \text{ and } \mathcal{C}_0 = [0,1] 
\]

\[
(10)
\]

![Figure 3. The Cantor set through several iterations](image)

Generalizing this to two dimensions, each cell may be transformed according to the rule of the Sierpiński carpet in which the light areas remain unchanged and each dark area is mapped into itself excepting that the middle sub-square (out of nine) is changed from dark to light:

\[
\begin{array}{c|c|c}
0 & 0 & 0 \\
0 & 0 & 0 \\
0 & 0 & 0 \\
\end{array}
\Rightarrow
\begin{array}{c|c|c}
1 & 1 & 1 \\
1 & 0 & 1 \\
1 & 1 & 1 \\
\end{array}
\]

\[
(11)
\]

Equivalently, one might use a random mapping where the 0 of the mapping for 1 is placed randomly in the matrix in the right-hand side. An example of this is given in Figure 4 where it was placed in the middle left corner in the first iteration and variously in the second iteration, and so on.

![Figure 4. The two-dimensional iterative random mapping](image)
The three-dimensional generalization of the Cantor set is the Menger sponge whose first iteration was shown in Figure 2 with the second and third iterations shown in Figure 5:

![Menger Sponge Iterations](image)

Figure 5. The evolution of the Menger sponge: (a) second iteration; (b) third iteration

**A model universe**

Consider a model universe in which the *dark* cubes at each iteration creates 7 sub-cubes of *light* cubes.

One may replace the mapping used in Figure 2 by modified forms. For example: (a) the *light* sub-cubes are not equal in size to the ones that remain *dark* (in other words, have a slower rate of conversion of dark to light), and (b) the transformation is done sequentially rather than at the same time for all regions. Doing so sequentially associates the process with a time variable. A steady-state model in which light and dark regions transform into each other at different rates may also be considered.

At the first iteration, the light density is \( \frac{7}{27} \), or approximately 0.26. In the second iteration, it will be \( \frac{7}{27} + \frac{20}{27} \times \frac{7}{27} \) which is approximately 0.45 and so on as shown in Figure 6.

![Light Accumulation Graph](image)

Figure 6. The increase in the density of light in the model universe
There need to be further investigations of stochastic versions of the model universe so that one can consider additional empirical aspects. The intrinsic properties of space may be behind phenomena such as the flyby anomaly.

**Conclusions**

Fractional spaces have been investigated for several research areas in a rather *ad hoc* manner. In this paper, we presented a theory that shows why nature must have such a characteristic.

We showed that an *e*-dimensional model for space is optimal on information theoretic grounds and since Nature is optimal, this dimension must be relevant for our investigation of physical reality. We considered some implications of this result. Since the nearest integer valued space to that of *e*-dimensions is three dimensional, we proposed that this is the reason why we perceive reality in three integer dimensions. The true dimensions of physical space can be determined by means of a careful sphere packing experiment.

We obtained the surprising result that the Cantor set approximates the one-dimensional construction of *e*-space and we showed how light might accumulate starting from a dark origin in a manner that preserves self-similarity and scale invariance.

**REFERENCES**


Subhash Kak is in the School of Electrical and Computer Engineering, Oklahoma State University, Stillwater, OK 74078, USA.

e-mail: subhash.kak@okstate.edu