DRAFT / Feb 2014 / bj

From the World of Color Toward a World of Darkness

I. Inrtroduction

It is a pleasure to participate in this meeting, devoted to reviewing the emergence of the concept of color and the resultant QCD gauge theory. It is also a pleasure to have had a small part in that history. My role has been absorbed into the folk history of the subject, especially with regard to the connection with Feynman's parton model. I do not want here to go over all that again. But I must mention that the foundation for what I did lay in Murray Gell-Mann's current algebra, and the consequences that followed from that seminal idea. Of special importance were sum rules derived by Steve Adler and Sergio Fubini, among others. Those pioneers deserve in my opinion more recognition.

While it would have been safer to reminisce for this entire talk, I have made the opposite choice. Since retirement, I have focused my physics efforts on the problem of dark energy. I have found myself drawn in a strange direction, and would like to sketch it out to you here. I think there may be some lessons from the past that can be applied to this problem, including the history relevant to this meeting. After describing my endeavors, I will return to that at the end of the talk.

II. Darkness and Gravity

There are two basic scales in the theory of gravity. The Planck scale of 10^-33 cm (or 10^28 eV) is determined by Newton's constant. The cosmological scale of 10^28 cm (or 10^-33 eV) is determined by the cosmological constant. But there is an intermediate scale associated with the density of dark energy out there in empty space. The natural scale associated with dark energy is halfway in between, logarithmically speaking (.003 eV or .003 cm).

The theme of this talk is the suggestion that there is another such intermediate scale, to be associated with the number density of conjectured topological structures present in the gravitational vacuum state. This second scale is, again logarithmically speaking, two thirds of the way from the cosmological scale to the Planck scale, namely 10^-13 cm (or 10^8 eV). Such a scale has been noticed in the past, most notably by Zeldovich in 1967. But there has not been very much hot pursuit of this observation. In my opinion there should be more.

My own pursuit begins with a version of general relativity called the MacDowell-Mansouri extension of the first-order Einstein-Cartan formulation. Its selling points include the following :

1) It is a genuine Yang-Mills gauge theory, based on the gauge group O(4,1).
2) Its Lagrangian is quadratic in the field strengths F built from the O(4,1) gauge potentials A.
3) The metric tensor is "emergent", i.e. it is built from a portion of the O(4,1) gauge potential.
4) The formalism requires that the cosmological constant must be nonvanishing and positive.

5) The MacDowell-Mansouri formalism is most naturally re-expressed in terms of the first-order Einstein-Cartan formalism. And that formalism is required if one includes Dirac particles within general relativity. This would seem to be a necessity if one wants to include quarks and leptons in a future theory that unifies gravity with the standard model.

The novelty of this description is that when the theory is expressed in Einstein-Cartan language, there appear three terms. The smallest of the three is the cosmological-constant term. The Einstein term, leading via the variational principle to the field equations, is larger. But it is the third, "largest", term which creates the novelty. It is called the Euler or Gauss-Bonnet topological invariant. Its Lagrangian is a time derivative of the quantity I call darkness, and from general principles does not affect the Einstein field equations at all. But the coefficient of this term is indeed enormous: it is the famous factor of 10^120 that pervades all discussions of the dark-energy problem that we face. While this Euler topological term appears as the leading term in the expansion of the MacDowell-Mansouri description in terms of the Einstein-Cartan description, one might take the point of view that this term is totally irrelevant, because it is only the equations of motion that matter.

I evidently take the point of view that the Euler/Gauss-Bonnet term does matter. We know, especially from experience with QCD, that at the quantum level such topological terms can make a difference. In the case of QCD it has to do with strong CP violation, instantons, and the effect of all that on the structure of the QCD vacuum.

III. The Topological Term in FRW Cosmology

In order to understand better the nature of this purported darkness degree of freedom, I have proceeded by considering how darkness behaves in simple solutions of the Einstein equations, rather than searching for a grand generalization. In particular, FRW cosmology is very easy to consider. Its geometry is controlled by a single parameter, the scale parameter a(t) that describes the time dependence of the size of the universe. It turns out that darkness scales with the comoving volume, in a way analogous to how the total energy scales. Therefore one can define a density of darkness, which turns out to be the cube of adot divided by the cube of a, all multiplied by the ubiquitous factor of 10^120.

The Einstein/FRW equation tells us that the square of adot(t) divided by the square of a(t), multiplied by the square of the Planck mass, determines the energy density. So we can see that the amount of energy associated with a single unit of darkness will scale as a(t) divided by adot(t), with the cosmological constant Lambda as coefficient. In the future, when dark energy dominates the expansion, it follows that the energy per unit of darkness is independent of time and extremely small, of order 10^-33 eV. It is hard to imagine a smaller value of the energy than that, because it would take longer than the age of the universe to accurately measure it, just from the uncertainty principle.

Not only is there a fundamental energy scale associated with a unit of darkness in dark-energy-dominated spacetime, but also a fundamental volume. An easy computation shows that the size scale associated with a single unit of darkness is, as already advertised, 10^-13 cm. It is large because of the large coefficient of the topological term in the MacDowell-Mansouri action.

In going back into the past, the darkness density increases rapidly. It in fact was Planckian when the Big Bang temperature was of order the darkness scale of 10 – 100 MeV. This suggests that the MacDowell-Mansouri description in some sense is incomplete. At larger energy scales or smaller distance scales, the formalism requires modification. This is only one such indication, and we will return to this issue in a later section. However, we here emphasize that this apparent deficiency does not necessarily mean that the formalism predicts a breakdown of the Einstein equations at such energy scales, but only that the formal structure is deficient there.

IV. Dark Energy and the Cosmic Web

As the universe expanded and cooled, the darkness density decreased from the Planck scale to its present value characterized by the darkness scale of 10^-13 cm. Dark energy-dominated regions of space began to appear only after large-scale structure formation was well along. Within the cosmic web--- consisting of nodes, filaments, and sheets of matter---- there appeared low density regions called voids. In the center of the large voids, isolated regions dominated by dark energy have by now appeared. In the future (after a few efoldings of accelerated expansion), almost all of space will be dominated by dark energy. Regions containing matter will be surrounded by dark-energy-dominated space instead of the other way around.

Actually, at the present time, we are experiencing a "percolation transition". The boundary surface separating matter-dominated space from dark-energy-dominated space now has long range connectivity and a complex topology. I think the region near this boundary surface deserves special attention. It can be defined as follows: Compute the spacetime curvature (two gradients of the Newtonian potential) and square it. If it is large in comparison with the square of the cosmological constant Lambda, one is in the matter dominated region; if it is small one is in the "heart of darkness", namely within a dark-energy dominated void.

Simulations can readily study the details of the evolution of these darkness-dominated voids. Different models of dark energy should affect the boundary regions with maximum sensitivity. Applying this knowledge to the real cosmos will of course be much more challenging. How best to do it would probably be learned via experience with the simulations. I recognize that such programs already exist. But to my knowledge, the identification of the boundary surface as described above has not yet been done.

V. The Topological Term Near a Matter Source

Place a proton in the middle of a dark-energy void. The boundary surface defined above occurs at a radius of about 20 cm. from the proton. I call it the sphere of influence of the proton. One nice way of defining it is in terms of the period of a hypothetical test particle orbiting the proton (under the influence of only the gravitational force), with an orbital radius of order 20 cm. It is of order the lifetime of the universe, i.e. the doubling time for the expansion of the universe. Evidently if the orbital radius were larger the test particle would be swept outward by the cosmic acceleration.

Another way of defining the sphere of influence is to compute the total dark energy that would be inside the sphere of influence in the absence of the source, and then to demand that it be of order the mass of the source. While this estimate is correct, it does not correctly determine the total amount of darkness within the sphere of influence. That quantity is much larger. In fact, with this geometry it is not hard to compute the darkness density as one goes from the sphere of influence inward toward the proton. One finds that the darkness density increases like the inverse 9/2 power of the distance from the source. It becomes Planckian at a distance of order the proton radius. If one replaces the proton by a source of nuclear matter density, such as a lead nucleus or a neutron star, the same thing happens. Within a factor two or so of the radius of the source, the darkness density becomes Planckian. This again indicates that one must regard the MacDowell-Mansouri description as incomplete beyond that point.

Evidently most of the darkness within the sphere of influence

and outside the radius of the source is near the source. It scales linearly with atomic number, and for a proton is of order 10^60. On the other hand, the amount of darkness within the sphere of influence in the absence of the source would be only of order 10^40 per nucleon; that quantity again scales linearly with atomic number.

VI. Visualizing the Topological Flow

I crave better intuition in dealing with this very curious situation. My experience with topological issues is limited. But in the case of QCD I find that, in dealing with instantons and the corresponding issues regarding vacuum topology, the choice of gauge makes a big difference. (In that case it is temporal gauge which gives me the best intuition.) In this case, the choice of coordinate system is analogous. I find that by far the Painleve-Gullstrand metric provides the best intuition. I am in fact surprised it is not found in most elementary texts on general relativity. The Painleve-Gulstrand description features a velocity field v(x,t), known in the trade as “shift”. In the simple geometries we consider here, this velocity field is time-independent, and can be identified with that of test particles of zero total energy falling in radially, from just inside the sphere of influence, toward the source. On the other hand, outside the sphere of influence, the velocity field naturally describes similar test particles accelerating radially outward due to the dark energy.

It turns out that these separate coordinate patches can be smoothly merged, because all components of the Riemann tensor match and are even functions of the velocity field and its gradients. However, the Gauss-Bonnet term is most naturally interpreted in terms of a stationary flow of the quantity I call darkness, with the velocity given by the Painleve-Gullstrand description above. The situation at the boundary resembles the mid-Atlantic ridge on Planet Earth. Consequently, the sphere of influence could also be renamed the “rift zone”.

Not only can one identify the outer boundary in such geophysical terms. The purported stationary flow of topology in the neighborhood of the inner boundary must somehow be controlled if one disallows (as I choose to do) superPlanckian darkness densities. I call this inner boundary the “subduction zone”. The next section is devoted to a feeble defense of such a supposition.

VII. Limitations of the Macdowell-Mansouri Formalism and a Possible Fix

The examples we have given indicate that something goes wrong with the MacDowell-Mansouri formalism when the distance scale is smaller than the darkness scale of 10^-13 cm. What might it be? In searching for an answer, one is drawn to the basic origin of this very large scale: it is the coefficient of 10^120 in front of the Gauss-Bonnet term. This coefficient also appears in front of the original O(4,1) Lagrangian as well. It is a fundamental feature of the starting point. Therefore it is very reasonable to presume that we have indeed started in the middle of the story.

A very straightforward approach for a fix is to keep the same basic geometrical structure of this action, but to extend it into small, compact extra dimensions a la string theory. If one does this, it follows that for every two extra dimensions there is one more field strength F to be multiplied into the action. The optimal choice seems, for both practical and politically-correct reasons, to be six extra dimensions. The action is generalized to an O(10,1) gauge group living in a ten-dimensional Minkowskl space. The new Lagrangian goes as the fifth power of the generalized field strength. If the three extra field strengths each have vacuum expectation values of order the Planck scale, and if the size of the extra dimensions is of order the darkness scale, 10^-13 cm, then the factor of 10^120 is generated. This gives an a posteriori reason why the formalism differs in the two regimes, as well as motivating the notion of rift and subduction zones. In some sense the darkness flows in and out of the extra dimensions in those regions.

With such a low energy scale for the extra dimensions, one probably has to exclude all Kaluza-Klein excitations (including gravity); this six-dimensional “bulk” should be dynamically passive. This is a feature possessed by topological insulators. And indeed the structure of the action is very Chern-Simons in character, supporting this premise. But the devil is in the details. And in the case of topological insulators, the details contain a great deal of subtlety. We can expect no less here.

VIII. Topological Flow in the Cosmic Web

Up to this point the idea has only been developed in the context of very simple, symmetrical model systems. It is very important to find a generalization. Furthermore, the description we have in hand so far is probably just a very rough sketch or cartoon of the real situation---assuming of course that what we are dealing with is not nonsense. That admittedly is a big assumption.

However, there is one simplifying element in going ahead. There has to be a valid description for the Newtonian limit of the gravitational theory. And this limit is a very practical one: dark energy is only relevant for low-redshift cosmology, and for the cosmology of the future, when exponential expansion of the universe leads to its domination by a space filled almost completely with darkness. While the present-day evolution of the cosmic web is complex, the underlying equations of motion are not. This is especially true for simulations, which generally use point sources for the matter component. Such point sources must be excluded from the MacDowell-Mansouri description anyway; one must work only outside of the subduction zones. Therefore the Laplacian of the gravitational potential can be taken to vanish everywhere inside the generalized rift zone boundary which separates the dark-energy region of space from the matter-dominated region.

Finding an appropriate intuition for dealing with the general field configurations present in the cosmic web is nontrivial. Since the quasistatic limit appropriate for the astrophysical application looks so much like electrostatics, it would seem that simple electrodynamic theory might suffice as a source of intuition. But it falls short. Instead, the fact that the gravitational description is based on a nonabelian gauge theory appears to be a crucial element, even in this simplified limit. I do expect that the O(3) subgroup present in not only O(4,1) but also the Einstein-Cartan O(3,1) will suffice as a provider of intuition. But even this is quite far away from what one finds in, say, Jackson's book.

At present I have a candidate for the darkness-current 4-vector appropriate to the Newtonian limit. The time component is constructed by taking the square root of the determinant of the"tidal" tensor formed by taking two gradients of the Newtonian potential (whose trace vanishes, thanks to the Poisson equation). The space components of the current are formed by taking the gradient of the square of the gravitational-force vector (itself the gradient of the Newtonian potential), i.e. the "gravistatic energy density". The flow velocity is the ratio of the space components to the time component, and comes out in accord with what was used in the explicitly computed, highly symmetrical examples. I do not yet have the mathematics under full control. But the intuition is very much in line with intuition I gained some time ago by looking at the loop quantum gravity formalism in the Newtonian limit. I now intend to revisit that subject in more detail.

IX. Lessons From the Past

The emergence of non-Abelian gauge theories is perhaps the most central development in particle theory in the last half century. The case of QCD is of course what we here celebrate. But it is striking that each relevant gauge theory is distinctly different from previous cases: QCD confines, QED does not. Electroweak theory has the Higgs mechanism, the others do not. And in the MacDowell-Mansouri version of gravity, topological issues are much more central than in previous cases. We must be prepared for surprising new features---even though the unifying theme in all these examples is local gauge invariance.

In dealing with the potential surprises, I believe that choice of descriptive language can be crucial. In the case of the history of quantum theory, the rather arcane Hamilton-Jacobi version of nonrelativistic mechanics provided a more direct portal into the quantum world than F = ma, or (initially, at least) a Lagrangian approach. I am of course assuming that the relatively arcane MacDowell-Mansouri description of gravity is the analog. While this is dangerous, I feel that, at the very least, it is an avenue that deserves thorough exploration.

The history of quantum theory also showed that a stepwise approach to the problem was very beneficial. Special and general relativity appeared in the midst of the old quantum theory revolution. But it turned out to be prudent to set all that aside, and to get the nonrelativistic limit under firm control before attacking the more general problem. I am hoping that the analogous non-relativistic limit here can serve a similar purpose, and help to keep the formalism grounded close to the real-life situation.

But even with the best descriptive language, the remaining obstacles in creating the quantum theory were huge. The appropriate mathematical tools (in particular Hamiltonian mechanics) were in the hands of the great pioneers like Bohr, Einstein, Sommerfeld, Ehrenfest, and Born. But it was deBroglie, a young latecomer, who set things in motion with the notion of matter particles as waves, generalizing the pre-existing notion of electromagnetic waves as particles. This idea, so simple and intuitively obvious in hindsight, was apparently very difficult to generate. The mathematics was there, but for its time the physics was too novel. The message is that thinking outside the box regarding the physics of dark energy is a central challenge, almost certainly more central than dealing with the mathematics of dark energy.

As things stand, a major challenge in understanding darkness better will be to cross the distance scale from the infrared side to the ultraviolet side. Here the experience from QCD may be of help. The language used on the infrared side (pions, nucleons, slightly broken chiral symmetry, etc.) is in general very distinct from that used on the other side (quarks, gluons, color). But there do exist descriptive elements which are robust on each side of the divide, most notably the local current operators built from the quark fields. Thanks mainly to Murray Gell-Mann, they provided a reliable way of exploring the avenue connecting infrared to ultraviolet. I anticipate such a situation will also prevail in the case of the darkness problem. And the situation I outlined in the previous section suggests to me that, even in the context of Newtonian gravity, there is an O(3) nonAbelian gauge symmetry controlling the properties of darkness. This may create a new set of color-like descriptive elements, which may help to illuminate this mysterious realm of dark energy and darkness.