CHIASMUS

CHIASMUS

PROCEEDINGS
ANPA 37 & 38

MAY 2018

EDITOR:
JOHN CERES AMSON

Copyright © 2018, Alternative Natural Philosophy Association
and the Authors.

All rights reserved

ISBN: 978-0-9574922-3-3

Published May 2018 by ANPA

PRODUCTION NOTE: Papers submitted as LaTeX source have a common typographic style, and papers submitted as MS Word documents retain most of the author's typographic style.

To the memory of our

Pierre Noyes

1923 — 2016

Scholar
Innovator
Teacher
Rebel
Mentor

and, above all,

Friend

— in gratitude —

Alternative Natural Philosophy Association

PATHETIC FALLACY REVISITED

BASS ROCK, calm block
why do you obsess me so?
To-day, clean edges like a knife
or flying mound à la Magritte (that shock of mighty mass and air)
you float on bands of Fifeshire light
a magic pop-up Treasure Island kidnapped from a Story-Book
To-morrow, shroud yourself in mist
your steep hoar-head, Rembrandt's Old Man
accosting clouds with rival snow, your martyrs' dungeons to the South
(doubly hidden where I stand) yearning like my avid mouth
to storm your silent citadel and make you speak
the endless Protean shapes you show

Before new flows of ice entomb
your faithful vigil of these shores,
open your choked-up throat again and blaze your old volcanic fires!
Hear my vain chains of metaphor stretched to you across the Forth!
Respond to me or let me go!

'NO!', you boom in your deep bass voice
'NO!', you pound from your hollow caves
your vast indifferent catacombs beyond the reaches of the sun
letting the tempestuous winds slip by your blind, unyielding brow
and Art's beleaguering commands
stream from your steeply tilted bulk
(sea-bird yourself now still you dive
below your crumbling avian hive) like vain cries off a gannet's back

NO! NO! NO! NO!

Christine Crow

CONTENTS

EDITORIAL — JOHN C. AMSON
Chiasmus. ix
CHimera. xiii

RICHARD AMOROSO
1 Why the Current SETI Protocol Will Fail. 1

JOHN C. AMSON
2 Binary Necklace Systems. 43

GEOFFREY CONSTABLE
3 Pions in Nucleons. 87

GRENVILLE J. CROLL
4 Order and Disorder in the Torah. 93

ANTHONY M. DEAKIN & LOUIS H. KAUFFMAN
5 Cosmological Theories Of The Extra Terms. 107

HERB DOUGHTY
6 AlphaGo and Friends — In Context. 203

MICHAEL HEATHER & NICK ROSSITER
7 Formal Representation of "Process & Reality" in the Metaphysical Language of Category Theory: Whitehead's Relational Theory of Space. 217

LOUIS H. KAUFFMAN

8 Iterants, Braiding and the Dirac Equation. 255

MARTIN LOWE

9 The Introduction and Development of Symbolic Representation in Mathematics during the Late Renaissance. 315

MICHAEL MANTHEY

10 Awareness Lies Outside Turing's Box. 339

N. ROSSITER & M. HEATHER & M. BROCKWAY

11 Monadic Design for Universal Systems. 369

PETER ROWLANDS

12 3-Dimensionality. 401

13 Fundamental Errors in Physics. 447

NINA SOTINA

14 The Causal (Deterministic) Approach to Quantum Mechanics. 479

NINA SOTINA & NADEZHDA LVOV

15 Coherence Domains in Water. Change of Water Properties in Living Organisms. 491

STEPHEN WOOD

16 Encountering Birds: A Phenomenology of Jizz. 497

EDITORIAL

CHIASMUS

Let us not ask "What can ANPA do for me?"
but "What can I do for ANPA?"

Such a rhetorical chiasmus might suggest that we know what ANPA is. As a four-letter acronym it stands for many things, as any diligent search on the internet will confirm. All but one is irrelevant here.
The exception being the acronym for our

Alternative Natural Philosophy Association.

Are we any the wiser ?

Is this an Alternative Association concerned with Natural Philosophy ? Or an Association concerned with an Alternative Philosophy of Nature ? Or perhaps one concerned with a Philosophy of an Alternative Nature ?

What is the epistemological force of this adjective Alternative ? Alternative to what ? What possible situation obliged the founders of our Association to accept any need to include the loaded signifier "Alternative" when choosing a prescriptive name ?

I would like to question What ? Who ? is being alternative to Whom ? to What ? I would like to question Why ? this Alternative. What possible confusion of human thinking obliges one cohort of thinkers to describe themselves as Alternative to another cohort of thinkers ? Are we all not thinkers in the same human mould ? Are we not all seeking to try to understand the Universe we find ourselves in, and our part in that Universe ?

Should we not consider the possibility that ANPA is simply the four-letter acronym for

A Natural Philosophy Association.

What could be more meaningful ? More truthful ? More significant in this factional, fractured age of the aggressively selfish intellect ?

It has been said[1], with bleak pithiness and much perception, that:

> We urgently need to reorganise our universities so that they become devoted to seeking and promoting wisdom by rational means as opposed to just acquiring knowledge, as at present.
>
> Modern science began as natural philosophy. In the time of Newton, what we call science and philosophy today – the disparate endeavours – formed one mutually interacting, integrated endeavour of natural philosophy to improve our knowledge of the universe, and to improve our understanding of ourselves as a part of it. Profound discoveries were made, indeed one should say unprecedented discoveries. It was a time of astonishing intellectual excitement and achievement.
>
> And then natural philosophy died. It split into science on the one hand, and philosophy on the other [...] and the split is now built into our intellectual landscape. [...]
>
> Rigour, sheer intellectual good sense, and decisive argument demand that we put the two together again, and rediscover the immense merits of the integrated enterprises of natural philosophy. This requires an intellectual revolution, with dramatic implications for how we understand our world, how we understand and do science, and how we understand and do philosophy. There are dramatic implications, too, for education.

The need for this recombining of science and philosophy is ever more poignantly brought out by the statistical evidence for the growth of multifarious organisations world-wide devoting themselves to disagreement, disparagement, and – sadly – angry polemics and *ad hominem* attacks on thinkers who are judged by one sect to be ignorant fools (and worse) belonging to another sect.

It does not take much searching to find that prominent amongst those organisations who receive opprobrium or rejection or plain despisement are these:

1. *Alternative Natural Philosophy Association* (ANPA) Founded 1979
2. *R. M. Santilli Foundation* Founded 1981
   `http://santilli-foundation.org/` `http://www.galileoprincipia.org/`

1 Nicholas Maxwell, *In Praise of Natural Philosophy : A Revolution for Thought and Life.* McGill-Queens University Press (2017). See also ANPA Proc. 31, pp. 282–302 (2011)

3. *Natural Philosophy Alliance (NPA)* Founded 1994.
   `http://www.connect-360.com/npa/about/`

4. *Editions dAssailly* Founded 2004, Jean de Climont
   `https://editionsassailly.com/les_videos.htm`

5. *The John Chappell Natural Philosophy Society (CNPS)* Founded 2014
   `http://www.naturalphilosophy.org/site/about/`

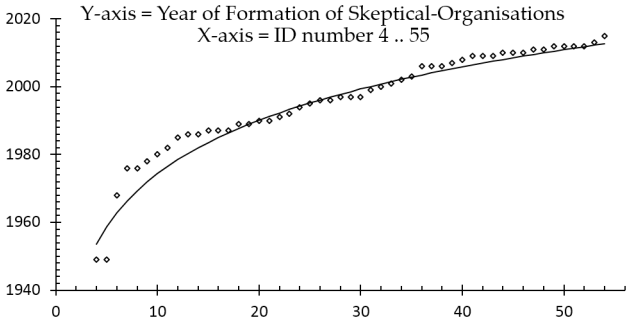
Philosophical and empirical "scepticism" has had a long and honourable rôle to play in our attempts to understand ourselves and our universe. But for some its meaning is too easily debased to imply "scorn". Perhaps a more disturbing aspect is the proliferation of self-important, condemnatory self-styled "Sceptic" organisations who seek to question the emergence of thinking, which, by their own definitions, is there only to be exposed as pernicious, ridiculous, stupid, condemnable. For example, one might take note of the

Stichting Skepsis Founded 1987
`https://en.wikipedia.org/wiki/Stichting_Skepsis`

There is list of 55 such "Sceptical Organisations" to be found at
`https://en.wikipedia.org/wiki/List_of_skeptical_organizations`

They began as early as 1856, 1881, 1903; the 4th in 1949 — the timewise distribution of their Founding-Year is shown by me in the following diagram. Curiously, their emergence seems approximately to follow a Logarithmic law :

$$\text{Founding-Year} = 1922.11 + 22.7\ \text{Log(N)}$$



Perhaps one might at least be cynically pleased that their growth is merely logarithmic, despite a sense that it could become exponential any time soon in our increasingly censorious and profit-seeking societies.

Where does our ANPA organisation fit into all this ? The clearest picture was given to us in a 2014 article[2] in which all 444 contributions that appeared in ANPA Proceedings in the thirty-three years 1979–2012 were observed to fall within seven different categories, of which the purely "Natural Philosophy" category claimed 187 (42 percent) of the total.

What is perhaps of further significance is that in this Double-Volume of ANPA Proceedings there are 16 contributions of which 15 can also be readily categorized as "Natural Philosophy" in their own right.

It pleases me that this supports my contention that our ANPA attitude is simply "Natural Philosophy" with no falsely defensive need to claim that these many contributions are in any way "Alternative".

A truly important Chiasmus with increasingly
significant consequences.

Long may we continue so.

John Amson, President, ANPA

2 Michael Horner, *A Management View of ANPA (East) 1979 to 2012*; in: *Scientific Essays in Honor of H Pierre Noyes on the Occasion of His 90th Birthday.* Eds. John C. Amson, Louis H. Kauffman, World Scientific Publishing Company (2014).

CHIMERA

The O.E.D. definition of a "Chimera" is
A thing which is hoped for but is illusory or impossible to achieve.

In this short note I draw attention to a recently discovered flaw in the construction of the classical 'Combinatorial Hierarchy' (CH). The CH had been created by Frederick Parker-Rhodes, Ted Bastin, Clive Kilmister and myself, in the mid 1960s. This flaw, bluntly put, means that the CH is a chimera.

The idea of the CH grew from the innovative discussions about an "Information Preserving Structure" consisting of a hierarchy of "Levels". This took place in 1960 and was reported by Ted Bastin and Mary Hesse in detail[1]. Formal details and the interpretation of our CH were first reported in our pre-ANPA paper[2] in 1979, and the algebraic aspects of the CH were described by me in the Appendix to that paper.

The basic idea of our CH was that 'information' at one 'Level L' could be 'preserved' at the next higher 'Level L+1' by identifying certain structural objects in L with uniquely assigned structural objects in L+1.

To give a representational model of our CH we chose to use an n-dimensional vector-space of bit-strings (with bit-components $0, 1$) of length n as one Level, and the 2^n-dimensional vector space over $\{0, 1\}$ of all linear operators acting on that lower Level. The structural objects in L were chosen to be what we called 'discriminately-closed-subsets' (DCSSs) : these were seen to be nothing more than vector

1 Bastin, Ted *et al.* *Self-Organization and the notion of Level. A Summary of Three Discussions held in The Tower Mill, Burnham Overy Staithe, Norfolk, and in Cambridge, September – October 1960*; in: *ANPA Proc. 32*, pp.3–23, (2011)

2 Bastin, Ted *et al.* *On the physical interpretation and mathematical structure of the combinatorial hierarchy*. Intern.J. of Theoretical Physics, 18, pp. 445–488, (1979)

subspaces in L with the o-vector deleted. The pool of structural objects in L+1 from which we were to draw uniquely asignable objects to match those in L consisted of all Invertible Linear Operators (Automorphisms) acting on L. As our way of assigning unique objects A in L+1 with each DCSS D in L, we chose to identify the Eigenspace of A in L with that DCSS, which together with the o-vector, made up that Eigenspace.

This worked well with the very lowest levels of dimension $2^0 = 1$ and $2^1 = 2$. It also worked, empirically, with a bit-space of dimension 3. Fatally, we assumed that it would work with the next level of dimension $2^2 = 4$, the Automorphsisms on which could be represented by 4×4 invertible bit-matrices, of which there are 20160 out of a total of 65366 bit-matrices.

Our mistake was not to recognise that any two 'similar' invertible bit-matrices have the the same eigen-spaces (to within isomorphism). By 'isomorphism' is meant that "two" vector subspaces which appear to be different under a change-of-basis (as happens with similar matrices) contain precisely the same 'vectors', and it is only their 'coordinate-representations' that change with the change of basis.

In our 4-dimensional Level L there are $2^4-1 = 15$ vector-subspaces, and hence 15 DCSSs. And each one is patiently waiting to be assigned to a unique Automorphism in the $4\times4 = 16$-dimensional Level L+1, whose Eigen-space is that DCSS-plus-o-vector.

The problem is that there are only 14 distinct Similarity Equivalent classes of Automorphims in L+1 (also known as 'Conjugacy Classes', for those who think that way). That vast 20160 strong collection of 4×4 invertible bit-matrices collapses into just 14 Similarity Classes, with varying amounts in each :

$1+112+210+1344+2520+1120+1680+2880+2880+105+3360+1344+1344+1260 = 20160$

There are 15 — just one-too-many — DCSSs in our 4-dimensional Level L to be matched as Eigen-spaces belonging to the 14 distinct Automorphims in the next-up 16-dimensional Level L+1.

Our hoped-for CH cannot continue beyond that 4-dimensional Level L. Our CH has become a chimera.

Ironically, those four 'physics numbers' $2^2-1 = 3$, $2^3-1 = 7$, $2^7-1 = 127$, $2^{127}-1 \sim 10^{38}$, (Mersenne-primes) which arose as the counts of the sizes of our now non-existent CH, do still arise from the simpler Hierarchy of subspaces which I have proposed elsewhere[3] and which are associated with that other well-known Von-Neumann Hierarchy. Moreover, unlike our defunct CH, these other Hierarchies exhibit the possibility of growth beyond our earlier four fixed Levels.

John Amson, President, ANPA

3 Amson, John *'PuCH', A Purely'Combinatorial' Combinatorial Hierarchy*, in: *ANPA Proc. 31*, pp. 40–57, (2011)

ACKNOWLEDGEMENT

Once again,
the dedication and skill,
of our colleague Anton Vrba,
in transforming a collocation of documents,
submitted in many diverse formats,
into a harmoniously unified publication,
is gratefully acknowledged.

WHY THE CURRENT SETI PROTOCOL WILL FAIL

RICHARD L. AMOROSO

Noetic Advanced Studies Institute
Escalante Desert, Beryl, UT 84714-5401 USA
amoroso@noeticadvancedstudies.us

ABSTRACT**.** In this work arguments are taken from Theology, Philosophy and Astrophysics/Physical Cosmology; license taken because of perceived conceptual overlap. In a highly speculative chapter in the monograph The Holographic Anthropic Multiverse, it was postulated that the current SETI protocol could fail. One of the main metaphors, was that aiming a telescope/microscope at someone's head would yield very little information about the content of their mind relative to the current standard model. It was postulated that some form of microscopic to cosmic conformal scale-invariance applied, which allowed a simple postulate that dark energy/matter was a gravitational effect from the rest of the causally separated multiverse beyond our observed finite Hubble sphere. Causal separation from the rest of the multiverse seems evident in 4-space; but in the echelons of 3rd regime unified field mechanics (UFM), a weak causal connection is manifest. A radical prediction extrapolated from $E = mc^2$, that the mass-equivalence of the earth is tantamount to the mass of intelligence of the Earth over its 5-billion-year history. Surprisingly, putting this supposition in a simple proportionality equation along with the mass of the sun, the result was the Hubble radius, or 'local' limit of observation. In terms of the anthropic multiverse model, this meant that one could postulate an infinite number of nested Hubble spheres each with their own fine-tuned the laws of physics beyond the event horizon. Several other details are examined. But what led to the possibility of a rigorous investigation of the proposal that the current SETI protocol will fail came via a surprising insight regarding the Poincaré dodecahedral AdS^5-dS^5 x S^5 wraparound universe cosmology. We already had in place the idea that additional dimensions (XD) were not of the microscopic Kaluza-Klein form, invisible because they are curled up at the Planck scale; but rather of infinite size (LSXD) and not seen because of a process called subtractive interferometry related to the virtual nature of 'observer

centric' perceived reality. When these additional parameters, especially subtractive interferometry, are applied to wraparound universe cosmology, it becomes straightforward to develop a rigorous model for the failure of the current SETI protocol. A brief introduction to a putative multiverse Q-Telescope able to successfully make multiversal SETI observations is given.

Biological organization [*is*] *the true foundation of reality* – Brian Josephson [1].

1. Phaedrus' Angry Charging Bull – On the Horns of a Dilemma

We enter the arena, but since we are unable to sing the bull to sleep; we wish we could throw sand in the bull's eyes to win the argument… Beware, 'all ye who enter here', by far the most speculative document the author, 'proud' of being able to 'color outside the box' (as required) has ever written; this work however, gives even him pause. Most say, quantum mechanics is stranger than we have yet imagined; I believe the role/nature of the observer is stranger yet and the salient reason its fundaments have not yet been discovered. Thus, this paper introduces an element of the observer that may have even given Descartes a moment of ratiocination.

If the physicality of the universe is not an *ex nihilo* Darwinian naturalistic Big Bang cosmology following a Friedman-Robertson-Walker (FRW) metrical solution to Einstein's gravitational field equations as the current vogue suggests; but a radically different Big Bang basis derived from anthropic principles guiding its temporal evolution from an eternal/atemporal multiverse perspective, then in this context, we explore the possibility that the Search for Extraterrestrial Intelligence (SETI) which formally began in 1959 (originally called the Order of the Dolphin) will require a radically different search protocol in order to have any hope of achieving success.

The highly-ordered symmetry conditions of anthropic multiverse cosmology suggest that conformal scale-invariance is an inherent property. This could mean that 'other worlds' exhibit the property of the 1st person - 3rd person barrier. This conjecture suggests that the current SETI strategy will fail because current search protocols are based on Big Bang naturalism in the context of the Copenhagen interpretation of quantum theory which is governed by the uncertainty

principle (The observer can only view one complement of an observation at a time because of the Pauli exclusion principle). This exclusive property according to the principle of invariance would apply to both other minds and other worlds. The current cosmological perspective does not include any putatively required addressable anthropic parameters of 'consciousness' relative to the physics of an observer. To repeat, in order to ameliorate any surprise: What is suggested is that other worlds in 4-space have properties similar to 'other minds' relative to an Earth observer which keeps conscious content blocked or hidden from detection by the limit of 4D Minkowski space based instrumentation by the quantum uncertainty principle. Why anthropic properties are required is explored. An outline for implementation of an alternative strategy based on new cosmological principles inherent in unified field parameters as represented in a Holographic Anthropic Multiverse is delineated. Design for an inter-dimensional radio Q-telescope utilizing entanglement with an HD 'ontological wave function of the multiverse' that Anthropic Multiverse cosmology might require is presented in a preliminary manner.

Sometimes scientific enquiry follows lines of inductive reasoning just to explore the spurious conclusions. This work is a logical extension of anthropic multiverse parameters in penultimate form; if it's SETI speculation is wrong it remains an interesting diversion to follow the curious thread of logical deductions. If it turns out correct then we are prophets dreaming of the day we can watch *Bronson Beta* TV or listen to *Ethnarch Gamma* radio! We cross the line of usual conservatism in order to leave no stone unturned, especially as SETI researchers have stated that under the current methodology it will probably take several generations to achieve success. With the pace of improving technology this would only take about two generations to check every star in the Milky Way Galaxy.

> The first episode of *I Love Lucy* was broadcast on October 15, 1951. About 0.0002 seconds later, the signal glided over the rooftops of the farthest city suburbs, and headed into space. It's still going. Every day, that first installment passes through an additional 4 thousand trillion trillion trillion cubic kilometers of the cosmos. Given that stars in our galactic neighborhood are separated by about 4 light-years, it's easy to figure that roughly 10 thousand star systems have been exposed to "I Love Lucy" in the past five decades [2].

When first researching this paper the ideas seemed outrageous; we suspect this or worse may initially be the case for you also dear reader. The endeavor was initially conducted more as a recreational exercise to see what tinkering provided rather than as a serious avenue of scientific inquiry. We still remain a bit dumbfounded at the final result but have learned to embrace it in view of the broad explanatory power of other aspects (medical breakthrough) of the new paradigm. We do not wish to belittle it by our sarcasm or by passing it off as errant twaddle; but we want to make it clear we realize the challenges it presents within the currently popular Big Bang / cognitive science mindset and wish to state this obvious fact. Finally, this work is a radical blend of theoretical anthropic cosmology, philosophy of science and epistemological theology. We hope you will enjoy at least the exercise of trying to keep an open mind as we have done; which is after all supposed to be the guiding light of scientific inquiry!

2. Anthropic Multiverse Cosmology

Many questions remain unanswered in cosmology. Most salient to our issues here relate to standard model Big Bang interpretations of the basis of CMB, redshift and the nature of expansion/inflation/quintessence. The Casimir (especially), Zeeman, Sagnac and Aharonov-Bohm effects provide serious evidence for the existence of a covariant polarized Dirac vacuum [3,4] which allows non-Doppler interpretations of redshift supporting anthropic multiverse cosmology. The Cosmological horizon problem, the horizon is a measure of the distance from which one could possibly retrieve information. This observable constraint is due to various properties of general relativity, the expanding universe, and the physics of Big Bang cosmology. Cosmological horizons set the size and scale of the observable universe.

2.1. *Friedmann–Lemaître–Robertson–Walker Metric (FLRW)*

The FLRW (Friedman-Lemaître-Robinson-Walker metric), an exact solution of Einstein's field equations, $G_{\mu\nu} + \Lambda g_{\mu\nu} = 8\pi G / c^4 \left(T_{\mu\nu} \right)$ derived independently during the 1920s-1930s by those named [5-10], is considered the basis of Standard Model Big Bang cosmology; and describes a geometrically homogeneous isotropic universe that can be

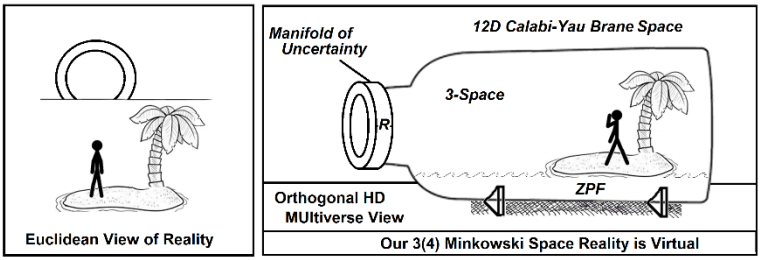


Fig. 1. Euclidean-Minkowski reality is virtual, Beyond The 4D Standard Model Lies Infinite Size Dimensionality (LSXD) 'Hidden' by Uncertainty.

either contracting or expanding [11,12]. A generalized metric with these conditions is $c^2 d\tau^2 = c^2 dt^2 - a(t)^2 d\Sigma^2$ where Σ ranges over uniformly curved 3D space that could be elliptical, Euclidean or hyperbolic. $d\Sigma$ commoving distance and is independent of *t* which is a function *a*(*t*) called the scale-factor.

For hyperspherical coordinates *r* is proportional to the radial distance (Euclidean distance of the curvature), such that $d\Sigma^2 dr^2 + S_k(r)^2 d\Omega^2$ where $d\Omega$ remains:

$$S_k(r) = \begin{cases} \sqrt{k}^{-1} \sin\left(r\sqrt{k}\right), & k > 0 \\ r, & k = 0 \\ \sqrt{|k|}^{-1} \sinh\left(r\sqrt{|k|}\right), & k < 0. \end{cases} \quad (1)$$

k represents the set of curvatures $\{-1, 0, +1\}$.

According to Big Bang dogma, comoving distance factors can rule out the expansion of the universe, giving a distance that does not change in time due to the expansion of space (may change due to other local factors, such as the motion of a galaxy within a cluster). Comoving distance and proper distance are defined to be equal at the present time; therefore, the ratio of proper distance to comoving distance now is 1. At other times, the scale factor differs from 1. The Universe's expansion results in the proper distance changing, while the comoving distance is unchanged by this expansion because it is the proper distance divided by that scale factor.

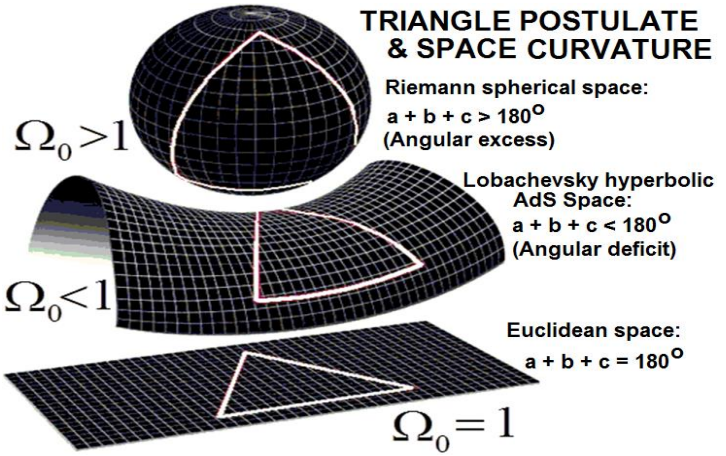


Fig. 2a. Possible shape of space for the FLRW metric. In our anthropic multiverse model the curvature of space oscillating around zero curvature, $\Omega_0 = -1 \leftrightarrow 0 \leftrightarrow 1$, providing a framework for wraparound cosmology.

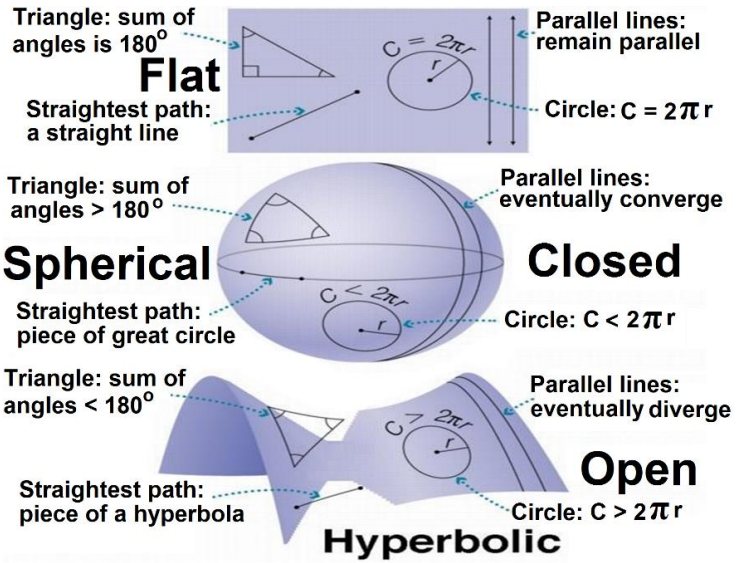


Fig. 2b. Topologies with positive, negative and zero curvature. Preliminary WMAP data suggests $\Omega_0 > 1$ for our Hubble universe allowing possibility for dodecahedral AdS^5-dS^5 x S^5 wraparound universe cosmology.

For our model, the cosmological constant, Λ can be interpreted as arising from a form of energy which has negative pressure, equal in magnitude to its (positive) energy density: $p = -\rho c^2$. Such form of energy can be generalized as dark energy and dark matter arising from the rest of the multiverse beyond out Hubble sphere. Like Ω_0, Λ cosmologically and $\hbar$ microscopically, can oscillate ± 0 in a complementary fashion. This is illustrated with our M-Theoretic derivation of a variable string tension, $T_s = \hbar + \Delta T_0$ which oscillates from asymptotic $\hbar$ to the Larmor radius of the hydrogen atom [13].

2.2 *de Sitter/Anti-de Sitter Space Multiverse Cosmology*

A de Sitter space in 4D Minkowski spacetime is analogous to a sphere in ordinary Euclidean 3-space [11,12]. The n-dimensional de Sitter space, (dSn), is the Lorentzian manifold analog of an n-sphere (with a canonical Riemannian metric); it is maximally symmetric, has constant positive curvature, and is simply connected for n of at least 3. For general relativity, de Sitter space is the maximally symmetric vacuum solution of Einstein's field equations with a positive cosmological constant, Λ (corresponding to a positive vacuum energy density and negative pressure). When n = 4 (3 space dimensions plus time), it is a cosmological model for the physical universe and a valid solution of the FLRW metric equations.

An n-dimensional anti-de Sitter space (AdSn) is a maximally symmetric Lorentzian manifold with constant negative scalar curvature. Manifolds of constant curvature are most familiar in 2D cases, where the surface of a sphere is a surface of constant positive curvature, a flat (Euclidean) plane is a surface of constant zero curvature, and a hyperbolic plane is a surface of constant negative curvature. The cases of spacetime of constant curvature are de Sitter space (positive), Minkowski space (zero), and anti-de Sitter space (negative). As such, they are exact solutions of Einstein's field equations for an empty universe with a positive, zero, or negative cosmological constant, respectively (see Fig. 2a,b).

Anti-de Sitter space generalizes to any number of space dimensions. In HD, it is best known for its role in the AdS/CFT correspondence, which suggests that it is possible to describe a force

in quantum mechanics (like electromagnetism, the weak force or the strong force) in a certain number of dimensions (for example four) as with a string theory where strings/branes exist in an anti-de Sitter space, with one XD. Unlike our universe, claimed be expanding at an accelerating rate, anti-de Sitter space is neither expanding nor contracting; instead it looks the same at all times [12-14]. In more technical language, one says that anti-de Sitter space corresponds to a universe with a negative cosmological constant, whereas the real universe has a small positive cosmological constant.

We embrace the Poincaré Dodecahedral 'wraparound' model, but as an anti-de Sitter/de Sitter ($AdS^5 \times dS^5$) duality, so that the $\Omega_0 = -1 \leftrightarrow 0 \leftrightarrow 1$ oscillation may occur. Preliminary data from WMAP supports an Anti-de Sitter ($AdS_5 \times S_5$) 'Poincaré Dodecahedral wraparound' model (suggesting Ω is > 1 for our Hubble universe) keeping the PDS model alive so far; but more definitive Planck satellite data will take several more years as the 2013 observation data was not taken at the required frequencies. For example, 3-space appears infinitely flat or Euclidean, but there are numerous observationally flat topologies that are not actually flat. The simplest shape for our reality is a 3-sphere with positive curvature. When we draw tori, they appear curved only because we embed them in 3-space. A flat torus can be made from a square with its edges wrapped around to join seamlessly. The criterion is that the angles of a triangle add up to 180°. If our observed Hubble sphere, H_R is a flat torus observations in certain directions allow one to see oneself in the distance (wraparound) [15-19].

This is compatible with the topology of a Big Bang 3-sphere postulated to be like a 3D expanding balloon; one with a 12 to 15 billion light year radius where light hasn't had sufficient time to cross yet. If the symmetry of curvature is not broken there are three possibilities: 3-sphere topologies with positive, negative and zero curvature as in Fig. 2a,b. But a 3D flat torus (cube with opposite faces joined) has zero curvature. There are also hyperbolic 3D spaces with negative curvature. In a finite 3-space like our H_R purported self-observations in various directions would allow us to work out the curvature and shape of the H_R which for us is considered closed and finite in time; but open and infinite in the atemporal LSXD anthropic multiverse cosmology it is embedded in.

The spatial region observed from any local point is a circular disk that increases in size with time. When it grows to the same size as the H_R it begins to overlap. At this moment because of relativity it would be possible to observe the same object in many directions in the overlapping portions of space. Age of the universe predictions from Big Bang cosmology suggest that if H_R is a 3-sphere overlapping should have begun and that these overlaps would form large circles in the sky, circles because the intersection of two 3-spheres is a circle [18]. Recent WMAP and Planck satellite observations have observed these putative circles in initial support of Poincaré Dodecahedral Space (PDS) [19-21].

Weeks claims, if Ω is sufficiently close to 1, the dodecahedral model is dead; but if Ω is found to be near 1.02 and bounded away from 1, then the dodecahedral model or some variation of it will likely be correct, whether or not matching circles are found [20].

> *… we see all the way back to the final stages of the big bang itself, when the whole universe was filled with a blazing hot plasma similar to the outer layers of the modern Sun. In principle, we see this plasma in all directions; it fills the entire background of the sky. So why don't we notice it when we look up at the night sky? The catch is that the light from it—originally visible or infrared—over the course of its 13.7 billion-year voyage from the plasma to us has gotten stretched out as part of the overall expansion of the universe. Specifically, the universe has expanded by a factor of about 1100 from then till now, so what was once a warm reddish glow with a wavelength around 10,000 angstroms is now a bath of microwaves with a wavelength of about a millimeter* [20].

According to Luminet, satellite measurements add support to a cosmology of finite space with pairs of so-called matched circles in the cosmic microwave background; in particular, a finite universe with a spherical Poincaré dodecahedral wraparound space topology, that is folded back on itself. The data suggests the horizon sphere wraps around the universe and intersects [15-19].

> Such circles have been searched in WMAP data … using various statistical indicators and massive computer calculations. First, Cornish et al. [21] claimed to have found no matched circles on angular sizes greater than 25, and thus rejected the PDS hypothesis. Next, Roukema et al. [22] performed the same analysis for smaller circles, and found six pairs of matched circles distributed in a dodecahedral pattern, each circle on an angular size about 11°. This implies $\Omega_0 = 1.010 \pm 0.001$ for $\Omega_m = 0.28 \pm 0.02$, values which are perfectly consistent with the PDS model. Finally, Aurich et al. [23] performed a very careful search

> for matched circles and found that the putative topological signal in the WMAP data was considerably degraded by various effects [24], so that the dodecahedral space model could be neither confirmed nor rejected ... The controversy still went up a tone when Key et al. [25] claimed that their negative analysis was not disputable, and that accordingly, not only the dodecahedral hypothesis was excluded, but also any multiply-connected topology on a scale smaller than the horizon radius. Since such an argument of authority, a fair portion of the academic community believes the WMAP data has ruled out multiply-connected universe models. However, at least the second part of the claim is wrong. The reason is that they searched only for antipodal or nearly-antipodal matched circles. But Riazuelo et al. [26] have shown that for generic multiply-connected topologies (including the well-proportioned ones, which are good candidates for explaining the WMAP power spectrum), the matched circles are generally not antipodal; moreover, the positions of the matched circles in the sky depend on the observer's position in the fundamental polyhedron. The corresponding larger number of degrees of freedom for the circles search in the WMAP data generates … computer time, up to values … out-of-reach of the present facilities. It follows that the debate about the pertinence of a multiply connected well-proportioned space model to reproduce CMB observations is fully open [19].

Vigier postulated a 4D model of TBS in hydrogen extending QED [27,28]. The theory seemed untenable until it was extended to a 12D M-Theoretic brane model with correspondence to the newly postulated LSXD regime of the Unified Field, U_F [29-31] in conjunction with anthropic multiverse cosmology [13]. It remains difficult to know the fundamental polyhedron (topology/geometry) of our cosmology made from observations within it; and astrophysicists continue to struggle with this problem.

2.3. *Standard Model of Cosmology – Alternative Interpretations*

As we have hinted, there are obviously numerous interpretations of these metrics; we reviewed them to illustrate that the alternative interpretation we have chosen is putatively valid. The de Sitter fluctuation duality (AdS_5-dS_5) which we build on is considered a derivative of the FLRW metric. We follow Luminet [15-19] in deriving our view of a dodecahedral AdS_5-dS_5 wraparound multiverse cosmology for our SETI investigation. This is not the only valid interpretation for a Big Bang alternative cosmology without an *ex nihilo* creation or expansion:

- CMB is not a Big Bang remnant, but the ‘cavity’ of the spacetime vacuum is an inherent natural QED cavity blackbody-CMB

equilibrium [32]
- Photon mass (tired light) anisotropy [33].
- Quasar luminosity as gravitational shock waves [34].
- Duality of Newton's and Einstein's gravity [35].

Except for the three maximally symmetric cases, Minkowski, de-Sitter, and anti-de Sitter spacetime, defining the space-like hypersurface can be naturally done with, e.g. a constant density-hypersurface, in the FRW world model. Spacetime everywhere is the same where $\kappa = 0$ for Minkowski spacetime, $\kappa > 0$ for de-Sitter spacetime and $\kappa < 0$ for anti-de Sitter spacetime.

The most interesting development arose from work with Kafatos [13,14] where an alternative interpretation arose for the energy of expansion/inflation/quintessence contrived in Big Bang cosmology. From this work $\dot{R} \equiv c$, usually used for the rate of universal expansion, is alternatively interpreted as relativistically internalized in the 'singularity' of spacetime as a 'continuous-state, which dramatically was the key to developing 3rd regime UFM [30,31].

An alternative to M-Theory is also used to overcome its current misses: 1) Search for a unique 11D to 4D compactification to the standard model, 2) discarding of variable string tension originally part of the early hadronic string model, 3) That XD are microscopic Planck-scale Kaluza-Klein curls and 4) The main thrust for a quantum gravity for which there is no *a priori* reason. Integration occurs in the 3rd regime.

3. An Anthropic Perspective for SETI Epistemology

Currently two standard techniques are utilized in the search for extra-solar planets. So far most planets have been discovered by observing tiny wobbles in a stars radial velocity due to the gravitational tug of an orbiting planet. This method favors the detection of very massive Jupiter class planets. The other technique, called the *transit method*, observes the slight dimming in a stars light when a planet passes over its disk. Technological improvements in the sensitivity of digital cameras and spectrometers continue to enhance both techniques significantly, which at the time of writing includes over 5,000 candidate and confirmed planets.

It seems reasonable to assume that if extraterrestrial intelligence exists at a technological level sufficient to broadcast radio signals into space, an Earth based search could be conducted by searching the sky with radio telescopes of sufficient sensitivity. This appears inherently obvious from the perspective of a naturalistic Big Bang cosmology and was the driving consideration for implementing the SETI program. But by 2050 the SETI program will be considered a failure if negative results remain; proponents have stated this search protocol would then have to be rethought.

Is the situation different in an anthropic cosmology; that is the challenging question? If it is, thinking metaphorically as queried above, as strange as it may seem; aiming a radio antenna at the sky may be no different than aiming a microscope or telescope at someone's head in order to read the content of their mind. This is another epistemological conundrum. Cognitive psychologists who currently consider the brain tantamount to mind contend that forthcoming improvements in f-MRI brain scans will be able to fully resolve consciousness. Therefore, an f-MRI enhanced radio telescope should be sufficient to resolve a signal, right? No, nonsense; but it hints at the direction of the solution. We have addressed this issue in detail elsewhere [36-40].

The suggestion explored here is that from the anthropic multiverse perspective SETI protocols must consider an additional parameter for resolving the signal because 'other world' will be caught under the 'other minds' 1st person 3rd person barrier that entails causal and therefore information separation. This is a stretch even for us to accept; future empirical tests will resolve the issue. Quantum theory discovered the nonlocal Einstein, Podolsky, Rosen (EPR) correlation. A completed form of Cramer-de Broglie-Bohm quantum theory with an HD M-Theoretic form of EPR entanglement that can be used to design a radical new type of inter-dimensional Q-telescope for manipulating and accessing causally separated anthropic domains. This relates to the nature of the observer, a challenge considered distasteful and therefore ignored so far in physical theory and epistemology in general [13].

f-MRI or MRI, are types of NMR where nuclei in a magnetic field absorb and re-emit electromagnetic radiation, usually photons in brain/body biochemistry such that the activity reveals medical conditions. This methodology of image formation is called a form of

quantum holography. There is an impending form of 'noetic holography' [41], that in terms of scale-invariance could be called the Q-Telescope for the human soul – spirit and the body or combination of Descartes *res extensa* and *res cogitans* [36].

We proceed in terms of the nature of the observer relative to the common philosophical question: *Is there a sound in the forest if no one is there to hear it*? Skipping the quantum mechanical argument as to whether there is a forest with no observer; the answer is no by definition. Sound is defined as a sensory input requiring a biological system; likewise, a tape recorder left in the forest has not recorder sound either until correlated with a consciousness. Ok, now we approach the horns of the dilemma. A usual em-field of propagating photons contains no EPR correlation that we can detect. With simultaneous emission of photon pairs by parametric down-conversion; nonlocally correlated EPR pairs are created for which we can detect the correlation. But this gets us no further than our 'eggs erroneous' nonsensical f-MRI quantum holographic enhanced radio telescope. So here is the obverse of F. Crick's 'Astonishing Hypothesis' thesis [42]. In the LSXD brane-world, matter takes on a profoundly different character. Imagine the de Broglie-Bohm pilot wave-quantum potential extended to the form of a super-quantum potential in the realm of Bohm's postulated super-implicate order, our 3rd regime of UFM [30,31]. The imminent completion of our understanding of matter is more than just the stringy topology of dual mirror symmetric Calabi-Yau 3-tori; albeit this is the vessel, but the brane structure is mediated by topological charge – or the entangled noeon flux of mind energy [36]. The 2nd element of this thesis is that a dodecahedral wraparound cosmology creates by inherent cosmological principles of scale-invariance a so-called 1st person – 3rd person barrier causing failure of 3-space SETI protocols!

Before delving into this thesis, we wish to discuss some additional epistemological considerations, as stated by theoretical physicist P. Rowlands,

> *Physics appears to be the only source of fundamental knowledge about the natural world. No other system of thought has shown any of its systematic explanatory or predictive power* [43].

The tools of physical science have been the pragmatic use of logic, reductionism and empiricism. But this can now be considered a myopic view because currently physics is devoid of underlying principles of 'consciousness' associated with not just the role, but also the fundamental nature of the observer. We have ignored the fact that the observer is imbedded in and made out of the same materials as the instrumentation. We have postulated that perceived reality itself is an intermediate collapsed or limited state governed by the quantum uncertainty principle. The time has come to intricately inspect the bias any inherent limitations the virtual nature of the observer's reality intrudes on the pragmatic foundations of empiricism. This means to complete epistemology the tools associated with transcendence must be included. Recall that thousands of years ago, Greek philosopher Plato stated that 'no matter how great one's intelligence, or how great one's wisdom, noetic insight (transcendent) from the cosmos is still greater' [44]. We are suggesting that in order to complete the tools of epistemology, transcendence or at least the role of anthropic properties should be utilized as a tool in guiding scientific theory formation relative to implementing access to 3rd regime natural science [36].

In contrast to the methods of scientific inquiry, theology also attempts to describe the nature of reality utilizing faith and revelation or transcendence. Scientific inquiry by definition should not discount one any more quickly than the other as long as pragmatic verification can occur. Myopic scientific interpretation has historically made as many horrendous mistakes as rigid bias in theology. We suggest that Theology and Science are not mutually exclusive as they became around the time of Galileo because of both the failure of sound logic as in the case of gravity for different weight objects and the narrow mindedness of the contemporary ruling theocracy, but opposite ends of a long continuum of human epistemology. We are not suggesting theology should replace science or that the empirical stance should be weakened in any way, only that theology and transcendence can be used as a viable tool to temper scientific theory formation and interpretation in some form of 'empirical metaphysics' [36,44]. Case in point, the principles introduced here could have been utilized in the year 2,000 rather than after the proposed 2,050 change in the SETI protocol, for a possible savings of many thousands of man hours and many millions of dollars. Not to mention that some future generation

instead of our generation gets to enjoy the wonders of extrasolar intelligence.

SETI researchers have suggested three possible reasons for the protocols failure:

1) Technical intelligence is very rare,
2) Technical societies are short lived,
3) Failure occurred because the correct search strategy is not utilized. (The one we address here)

We use this paper as an axiomatic test case for future hindsight into the viability of an 'empirical metaphysics [36,44,45] in evaluating the utility of the putative integration of science and theology, but not to return to the bias shown in the following quotation:

> *For the first time since the Dark Ages," physicists Paul Ginsparg and Sheldon L. Glashow wrote 12 years ago, "we can see how our noble search may end, with faith replacing science once again* [46].

4. The Curious Drake Equation Enumerating Extra-Solar Intelligence

The Drake Equation [47-49], first formulated by radio astronomer Frank Drake in 1961 purports to estimate the number of communicative civilizations in the Milky Way Galaxy. Drake's initial solution to the Drake Equation estimates 10,000 communicative civilizations in our galaxy of about 250 to 400 billion stars. But Drake's equation is based on various assumptions. If different values to the assumptions are utilized intelligence can just as readily be suggested to be the general rule not the statistical exception. At the time of writing about 5,000 extra-solar planets have been discovered; most of which are Jupiter size found by the minute stellar wobble their gravitational mass creates on the star orbited [50]. But on 7 March 2009 NASA launched the Kepler extrasolar planet finder which will search the Milky Way for Earth-size planets utilizing the transit method. Kepler's photomultiplier sensitivity allows searches for Earth-size planets around stars up to 3,000 light years distance.

The Drake Equation:

$$N = R^* f_p n_e f_l f_i f_c L, \quad (2)$$

where

- **N** = The number of communicative civilizations
- **R*** = The rate of formation of suitable stars (such as our Sun)
- $\mathbf{f_p}$ = Fraction of those stars with planets. (Current evidence indicates planetary systems may be common for stars like the Sun.)
- $\mathbf{n_e}$ = The number of Earth-like worlds per planetary system
- $\mathbf{f_l}$ = The fraction of those Earth-like planets where life actually develops
- $\mathbf{f_i}$ = The fraction of life sites where intelligence develops
- $\mathbf{f_c}$ = The fraction of communicative planets (those on which electromagnetic communications technology develops)
- **L** = The "lifetime" of communicating civilizations

Using one of the interactive sites in [46,47] one may experiment with the parameters of the Drake equation to explore the range of possibilities.

Following the principle that life is the rule in an anthropic cosmology we propose corollary §1

§1 All Stars Have an Anthropic Zone.
Every star has an anthropic zone associated with it. Since 'gravity is caused by the movement of spirit' [51] or inertia of the unified field if you prefer; the anthropic action principle optimizes the formation of planets compatible with life in this region; but of course, this doesn't necessarily mean a star has planets in this zone; but anthropically that shouldn't be ruled out either as an additional inherent evolutionary force.

5. Reviewing Holographic Anthropic Multiverse Parameters

Anthropic Multiverse cosmology is not a naturalistic Big Bang cosmology but a form of continuous-state eternal/atemporal timeless multiverse derived by extending Einstein's original static universe model [52] to include HD SUSY M-Theoretic parameters. Anthropic Multiverse cosmology has properties reminiscent of Kant's antinomy of spacetime which he proposed as a solution to the argument between

Newton and Leibniz as to whether the universe was open or closed [53]. A fundamental Anthropic Multiverse premise is that the observed Hubble radius, H_R the Einstein 3-sphere of our perceptual reality is closed and finite temporally; and open and infinite atemporally. Thus, the topology of Anthropic Multiverse cosmology is like a continuously transforming HD hyper-Klein bottle with the possibility for an infinite number of Hubble type spheres 'holographically' nested within it, each of which might have a fine-tuned variance of the laws of physics [13]. The paradigm was developed by extending the Wheeler-Feynman-Cramer radiation/transactional interpretation models [54,55] and the de Broglie-Bohm ontological models to an HD regime commensurate with our version of SUSY-M-Theory parameters [36-40], but not interpreting Everett's many worlds condition as a duplicate parallelism but as additional and unique in their own right.

For Kant there are four antinomies connected with

1. the limitation of the universe in respect of space and time,
2. the theory that the whole consists of indivisible atoms (whereas, in fact, none such exist),
3. the problem of freedom in relation to universal causality
4. the existence of a necessary being

all of which by pure reason contradict the empirical, as thesis and antithesis. This was part of Kant's critical program of determining limits to science and philosophical inquiry [53].

The inherent purpose of an Anthropic Multiverse is the evolution of life and consciousness, therefore intelligent life is the rule not the accidental exception as often considered by Big Bang cosmologists. The difference in evolution for an Anthropic Multiverse is that evolution is not random but guided by a teleological action principle. A major premise of Anthropic Multiverse cosmology in this respect is that cosmology is a self-organized complex system. Such systems are driven by an external action principle which in this case is the anthropic teleology. This action of course applies to each nested Hubble sphere. Thus, the horizon of knowledge moves outside the temporal bounds of H_R where theory is still silent about other aspects of the Multiverse; but since the model is empirically testable we hope to soon begin to unravel new mysteries. In the Bible's Book of

Genesis, it says, "*Let there be light*"; which seems to logically correlate with a Big Bang ignition. In terms of an infinite cosmology, this would suggest the Hubble radius was observer related and occurred simultaneously everywhere without expansion.

The difference in the two scenarios demands radically different search protocols for detecting extra-solar intelligence.

6. A Different Strategy for SETI Success

Even if an EPR[1] correlation was somehow employed; under current methodologies EPR superposition is unable to teleport detailed information from one location or domain to another. The EPR condition only arises under the simultaneous emission of photons (photon pairs) which as a result are correlated in time. It is suggested that neither people or planets naturally maintain such a correlation and the current interpretation of quantum theory is not sufficient to develop such a superposition from causally separated entities by any known method of parametric up or down-conversion. The EPR paradox therefore illustrates both why a Big Bang SETI strategy will fail and as will be illustrated below leads to the explanation of why a Anthropic Multiverse SETI strategy holds more promise.

This anthropic condition for SETI is an issue of both philosophy of science and theology. For example, although cognitive brain scientists remain hopeful, the myriad forms of brain scan technology cannot achieve the analogous result of resolving consciousness. Even utilizing the highly anticipated new MRI or fMRI advances will fail because 'mental information' is not contained in the energetic atomic or molecular states of brain neural biochemistry. Consciousness is not only a brain state, but a teleological cosmological condition [36-40]. The proper criteria for a successful SETI program must be aligned with a LSXD Cartesian dualistic model of mind-body interactionism [36].

This is the historical debate over the sufficiency of *Biological Mechanism*[2] versus the need for an additional life principle [36,45]. At

[1] EPR paradox – From a thought experiment devised by Einstein, Podolsky and Rosen which later in relation to experimental tests proved the existence of 'quantum nonlocality', a condition of long range entanglement or superposition.

[2] Biological Mechanism – The philosophical position that the principles of Chemistry and Physics are sufficient to describe all life; no additional life principle is required.

the moment, this issue remains highly controversial and many would insist that the 'telescope/microscope' metaphor applies neither to the cosmological issues pertaining to SETI research nor especially not to information content in the brain/mind. The critique revolves around the philosophical issue of Biological Mechanism. To insist at the present level of scientific progress that the mechanistic model of mind or the Big Bang cosmology is incorrect is highly unpopular. Unfortunately, at this moment nearly every scientist or SETI researcher would therefore summarily label the premises given here as nonsense, believing that solar systems or planets are nothing like brains; and to even consider such a claim would be absurd. No current standard model of science (of which there are up to a dozen depending on how one counts them) can predict this conundrum because they remain naturalistic and the problems addressed here therefore are outside this scope of influence.

In Anthropic Multiverse cosmology, the Hubble radius is an observational limit based on a 'tired light' redshift [33,56-65] rather than a Doppler redshift. Hubble discovered redshift not expansion of the universe. With the alternative interpretation, the Universe is not limited merely to the ~15 billion light year radius Hubble sphere of the Big Bang model that we observe. The Multiverse has the potential for an infinite number of holographically nested Hubble-type spheres in causal separation each with their own fine-tuned laws of physics [14].

Here is where the conceptual problem arises for SETI as it is currently employed. In a Conscious or anthropic universe, it appears that each intelligent system (by this we mean planetary system with intelligent life) has its own natural laws that are out of phase with the cosmological conditions of other causally separated multiverse intelligent systems (The same might be said for people) making oscillations of the background matter, not the mental content (like the quantum fluctuations of brain chemistry explored by fMRI) the only material viewed. So, the discussion here is centered on the problem of how to get at the 'Qualia' or endogenous 'conscious light' (qualia) of intelligence rather than just the thermodynamic energetics of the atomic structure of the rocks and gasses there.

7. Theological Arguments – Adam's Reckoning Versus Cosmological Age of the Earth

The age of the Earth based on scientific radiometric dating has an upper limit of about 4.567 billion years based on rock in the Earth's crust which is also compared to moon rock and Martian meteors. The best-known absolute dating technique is carbon-14 dating, which archaeologists prefer to use. However, the half-life of carbon-14 is only 5730 years, so the method cannot be used for materials older than about 70,000 years. Radiometric dating involves the use of isotope series, such as rubidium/strontium, thorium/lead, potassium/argon, argon/argon, or uranium/lead, all of which have very long half-lives, ranging from 0.7 to 48.6 billion years. Subtle differences in the relative proportions of the two isotopes can thus give good dates for rocks of any age. Radiometric dating is based on solidified rock. Should this be the valid origin of dating? It is purported that the Martian core took about .5 billion years to solidify. The Earth's core has not solidified because of its larger mass. Thus, extrapolating the age of the Earth as a molten ball before the crust solidified could add a billion years to the Earth's age; or do we not call the molten ball before crust solidification Earth? This analysis is made to contrast the scientific age of the Earth in opposition to the literal 'young earth' theological doctrine with a different type of calculation making correspondence to the scientific model. Although not a rock-hard theory it at least gives suggestive support to the Anthropic cosmology put forward here [66].

Theologically, after the creation as delineated in the Book of Genesis of the Judeo-Christian Bible/Torah, Adam (the 1st man) was '*given his reckoning'* in the Garden of Eden. This is interpreted to mean that instead of existing in God's timeless eternal frame; Adam was given a different Earthly temporal clock or reckoning where "*A day with the Lord is a 1000 years with man*" [67] and a lifespan of "*four score years*". Anthropic Multiverse cosmology suggests that Adam's 'clock' is relative only to Earthly intelligence and therefore the fine-tuning relative to Earthly laws of physics, so that both God and other cosmic civilizations are out of phase (therefore invisible and causally separated) as it were by both the clock and geometric structure of our reality for the Earth observer [66].

It is possible to make a simple 'age of the Earth' calculation to illustrate this in contrast to the myopic view of the young-earth

creationists who claim the Earth is literally only 7,000 years old by utilizing the Judeo-Christian premises for the calculation:

- A day with the Lord is a 1000 years with man
- God created the world in seven days
- The creation occurred first spiritually and second temporally.

Therefore, there is a modicum of credibility of doing a simple calculation using 'Gods time'. Thus, 7 days x 1,000 man years per day = 7,000 years which is the usual literal point of view taken for the period of man's existence. But Adam was not given his reckoning until after he was placed on the Earth. so instead, if we consider the creation as 1,000 God years, multiplying 365.4 Man days per year x 7000 God days x 1,000 = 2,557,800,000 billion Man years x 2 because of the dual spiritual and temporal creations = 5,115,600,000, we come up with a proper order of magnitude for the scientific age of the Earth. Oh, you noticed the number appears to be a half billion years off. You forgot that the scientific reckoning date begins from the radio-dating time when rock solidified. It seems logical that it took half a billion years of cooling to solidify. The calculation follows directly from scripture; is it just a coincidence that it arrives at the correct answer of ~ 4.6 billion years?

Now let's fudge in a different way. If we keep the spiritual creation as the 7,000 God years and reduce the temporal to 6,000 we get an age of the Earth of 4,750,200,000 billion years well within acceptable error for the putative radiometric dating methods. But how can we justify this concatenation? Easy. We go back now to make correspondence to the 'young Earthers' in terms of Adam given his reckoning and interpret the omitted 1,000 years as the age of man for the fulfillment of a society of Gods people. Adam to Abraham to Moses to Christ to present age being the 6,000 years with the 7^{th} to come as the imminent Judeo-Christian-Islam Millennium. No offense if the reader wishes to call this 'hogwash'. It has as much valid logic as any other consideration; and its purpose is only to be thought provoking and foster debate on the utility of some form of the anthropic principle in cosmology.

This has more pertinent meaning in terms of the first line of Genesis: "*In the beginning God created the Heaven and the Earth*" – Genesis 1:1. This is taken to mean God created this earth and its

heaven. (A pocket universe in the Multiverse?) We realize that moot theological dogma is not considered scientific; but it can be used as a philosophical basis for developing a science of cosmology. Thus, in a Holographic Multiverse this implies room for an infinite number of nested Hubble spheres each with their own fine-tuned laws of physics [13,14]. Notably, Genesis 1:1 supports the SETI opacity of Poincaré dodecahedral wraparound cosmology!

Four Theological Arguments can be used to support failure of the current SETI search protocol.

1. Eternal Multiverse cosmology versus Temporal Big Bang Cosmology. In an anthropic universe, a teleological action principle guides evolution. The argument is that even quantum theory is insufficient for success. A model utilizing the noetic unified field is required.
2. Genesis 1:1 states that God created[3] both '*the Earth and <u>its</u> Heaven*' suggesting the laws of physics may be different for each civilization because the nested Hubble spheres are in causal separation requiring a different Q-telescope method for SETI success.
3. Scripture also states that '*Adam is given his reckoning'* preaching that Man's time is not God's time suggesting his time rate is different than that of other Hubble spheres (In a holographic anthropic multiverse a Hubble sphere represents the observational limit of a particular civilization).
4. Anthropic Principle (AP) arguments exist against Steady State cosmology. They suggest that ETI should have already filled a timeless universe and currently be in our solar system [68]. The argument is logical for a naturalistic Big Bang cosmology; but not for Anthropic Multiverse cosmology. Earthlings are barely able to coexist with each other. By putative teleological considerations the organizer of intelligence may have a planets millennium (judgment day) begin before a civilization is able to use 'warp drive' space travel to visit other solar systems and interfere with their societies natural evolution [29,69]. This argument may seem unacceptable to some but one is forced to admit that it is as logically valid as AP arguments against Steady State cosmology.

[3] Not *ex nihilo* (out of nothing) creation, but 'organized' (Hebrew *Bereshith*) from existing raw materials, Genesis 1:1.

There is one scripture supporting this argument albeit from obscure Mormon theology:

> *Is not the reckoning of God's time, angel's time, prophet's time, and man's time, according to the planet on which they reside? … Yes. But there are no angels who minister to this earth but those who do belong or have belonged to it* [70].

This suggests that God will not allow external interference in the evolution of other planetary societies and could be a reason other stars are so many light years distant from each other.

Validity of a Big Bang cosmology is hinted at in Genesis: "*Let there be light*", but not an *ex nihilo* random Darwinian Big Bang. A theologically based Big Bang scenario requires dodecahedral wrap-around cosmology, gleaned from the Genesis verse: '*God created the heavens AND the Earth*'. It is useful to be reminded of the new U_F component (spirit of God, Super quantum potential) perfusing matter (see Sect. 3) which is invisible to the current SETI protocol by phase interference as applied to the definition of 'sound in the forest' and the horizon of observation by redshift attenuation of light.

That observed reality is virtual seems obvious. Like Plato's Analogy of the Cave, we assume that an observer sitting in a movie theater watching a continuous stream of images on the screen in front of him, produced from the motion of microscopic discrete frames of film passing through a projector behind him, represents the observers place in reality.

> *Which truth shineth … the light of the sun … the light of the moon … the light of the stars … And the light which shineth, which giveth you light, is through him who enlighteneth your eyes, which is the same light that quickeneth your understandings; Which light proceedeth forth from the presence of God to fill the immensity of space. The light which is in all things, which giveth life to all things, which is the law by which all things are governed, even the power of God who sitteth upon his throne, who is in the bosom of eternity, who is in the midst of all things* [71].

8. Absurd Pseudo-Scientific Argument

Now some very radical theoretical inference is introduced. The renowned physicist J.A. Wheeler proposed what he called the geon concept, a ball of photonic light of sufficient size to gravitationally self-cohere into a sphere (not a star with matter, but a ball of pure light). The propagation of photons in the group would be in circular rays such that there would be no loss [72]. According to the special theory of relativity [73], in terms of the laws of conservation of energy and momentum, even though photons are considered to have zero rest mass; they have an equivalent *moving-mass*, $m = E/c^2$ that is able to facilitate the cohesion of the Wheeler geon [72].

French physicist Louis de Broglie proposed that the same wave properties exhibited by em-phenomena applies also to all material particles [74-76]. Recall first that in dualistic models of consciousness the mind is like the light (rather than the usual electricity) running an optical computer. If the de Broglie wave of the Earth were considered to be like a Wheeler geon representing the light of all of Earth's intelligence over all time, then the mass of the Earth according to Einstein's formula for special relativity, equating matter and energy $E = mc^2$, when converted to energy could represent the geon energy mass for all life on Earth over the approximately 5 billion years of its existence.

Our moment to moment view of reality is in time. This is like the centroid of a wave or internal spinning mass-particle moment of a propagating wave according to the standard concept of wave-particle duality. The envelope (wave) is timeless and the internal motion (particle) is temporal. Therefore, the total mass-energy of the wave can be used to theoretically represent the total mass-energy of planetary intelligence, which can be considered reminiscent of Jung's Collective Unconscious [77] which most transpersonal psychologists consider to contain physically real archetypes. Idiot savants and past-lives regression seem to suggest this is the case

The 14 billion light-year Hubble radius, H_R is the limit of observation. In Big Bang Cosmology this radius is said to be the current limit of expansion of the universe in terms of a Doppler redshift. In Anthropic Multiverse cosmology, the interpretation is different and based on a so-called 'tired light' mechanism. During the centroid moment of photon propagation, the internal motion causes an

inertial gravitational spin coupling to the spacetime vacuum. This coupling is responsible for the observed cosmological Redshift [33]. This means that at the Hubble radius all visible light has attenuated to zero energy signifying the limit of observation. This 'mass of intelligence' would correlate with a classical 3(4)D gravitational force causing a perceptual collapse of the 11(12)D eternal space of Anthropic Multiverse cosmology reducing it to the more limited temporal reality we observe as the external world. What I'm trying to say is that the photonic mass of the Earth's geon of intelligence provides a force collapsing observational space to the Hubble radius for the temporal based Earth observer which is a limited virtual temporal 3-space reality.

Let us try to explain this more clearly. When an EPR experiment is performed, simultaneously produced parametric down-converted photon pairs are required to provide the correlated system. These photons are propagating through space at the speed of light. Imagine instead a stationery system with an eternal conformal present [78] as might be viewed by a HD observer from the point of view of Einstein's pre-relativistic thought of 'what it would be like to ride on a photon' – Time would stand still, i.e. one could circumnavigate the whole universe without the passage of time. In Anthropic Multiverse cosmology both the multiverse and human beings contain an inherent duality of temporality and eternity. In this duality the invisible core is a 12D node of the unified field from which our 3(4)D view continuously flows as a standing-wave subspace.

The Pauli exclusion principle is illustrated in the uncertainty principle by the fact that components of the quantum wave function do not commute in the lesser Copenhagen regime. Correlated photon pairs are produced by what is called parametric down-conversion [79]; although theoretically described science currently has no method for parametric up-conversion [80] to correlate unpaired photons. EPR superposition is the most primitive form of these correlations. The point being made is that a higher level unified correlation is required to obtain 'conscious' information like through telepathy or 'revelation from God'. This information is scale-invariant so that what we are describing at the photonic and human level is also postulated to occur at the cosmological scale. All the information is contained in the hologram, but by the 1st person - 3rd person barrier macroscopic uncorrelated systems are not able to share internal information. This

is a form of collapse of the wave function at the level of reality itself which our awareness is naturally coupled to.

This is a converse of the observational limit in physical cosmology where the ~14 billion light year radius of the Hubble sphere H_R is caused by light redshifted to complete attenuation. This relies on an alternative interpretation of Hubble's 1940 discovery. Big Bangers interpret his discovery of the red shift as an indicator of expansion of the universe, when actually all he discovered is a redshift distance effect that can just as reasonably be interpreted to be non-Doppler as Doppler.

In summary, the gravitational mass of Earthly intelligence over all time represents the force that gives the Hubble sphere its observed radius, i.e. in a conscious universe U_R is infinite, but the innate force of the observer creates in the continuous creation of his reality, the observational limit which for Earthlings is ~14 billion light years. This means that within another Anthropic Multiverse, Hubble type nested sphere elsewhere in the multiverse, that mass of intelligence would be different, the laws of physics different by fine tuning and therefore the Hubble radius different for the denizens there! This is the crux of our premise for the failure of the current SETI protocol. These other intelligent domains would be 'out of phase' with our reality and therefore not observable with a standard telescope or microscope. This is considered a standard feature of wave mechanics. This is similar to making a quantum measurement; only a portion of the information is observable and only ontological versions propose methods for obtaining complete information beyond uncertainty [13]. This is part of an ongoing debate as to whether the Schrödinger wave function provides a complete description of reality.

Therefore, in this model the Barnard's star geon of intelligence would have a radically different gravitational mass (statistically highly improbable that they would be the same; representing the highly improbable only case where the current SETI protocol could work[4]. It still would not work because this would be like an EPR experiment that cannot be used to teleport information, creating the Hubble type sphere created by the denizens living there. This would relate to the 'reckoning' given to the 'Adam' there, which must be taken into

[4] Because simultaneously created entangled EPR photon pairs would be required – this is why the Q-telescope postulated here is required to produce such a correlation.

account for our reception of Alpha Centauri technological broadcasts according to the theoretical model presented here. There are several kinds of coordinate transformation, the Galilean transformation for small velocities, the Lorentz-Poincairé transformations for relativistic velocities and a newly proposed noetic transformation for ontological UFM LSXD considerations [81].

9. Cosmology of the Anthropic Principle (AP)

The development of Anthropic Multiverse cosmology aligns with arguments for the Anthropic Principle which states emphatically that the observed universe is designed to accommodate intelligent beings [82-84]. Subtle fine-tuning also seems to play an important part. For example:

- A star must remain on the main sequence for about 10 billion years in order for there to be sufficient time for planet formation and a 'guided' evolution of intelligent life; which for Earth as a likely average case took about 4.5 billion years.
- If the Earth spins much faster the atmosphere will fly off.
- If the Earth spins much slower opposite sides will alternately burn and freeze.
- If the mass of the Earth is very large gravitational forces will be too strong to support life as we know it. The linear increase in body size and mass requires a quadratic increase in bone strength; suggesting height would be less than a meter and leg diameter the circumference of an average man's waist today.
- If the Earth were too small (like Mars for example) the molten core would solidify, the dynamo would stop rotating attenuating the geomagnetic field causing the atmosphere to fly off into space since there would be no charge to trap the gasses.

It's interesting to speculate on possible environments of extrasolar planets. Living conditions would depend on the distance of a planet from a star and the stars spectral qualities. For a binary star system which is very common, if a dim massive red giant had a smaller bright star orbiting it, the red giant might be far enough away from a planet orbiting the smaller star to have little effect on it. If one assumed that the second star had a significant effect on the planet, it would probably

result in the seasons being based on the proximity of the secondary star. Summer would occur while the planet was between the two stars, and daylight would be continuous. Winter would occur when the planet was on the far side of the primary star, and there would be the usual darkness at night. Habitable worlds of such a system would have a highly elliptical orbit that was much further away from the primary star during the summer because of the gravity of the secondary star. The elliptical orbit would regulate the annual temperature: the heat of two stars would be tolerable because the planet stays farther away from one of the heat sources [85].

In order for life to be the rule rather than the exception, these fine-tuned details unique to each star system would be far easier to meet with an anthropic action principle guiding planets evolution and star formation. Gee, maybe there is an Occam's razor for living systems. Soon enough we will have sufficient data for Titius-Bode profiles for the planets of numerous star systems to further insight into these interesting conundrums.

The Weak Anthropic Principle (WAP) [68] asks where and when are good conditions for intelligent life realized. The Strong Anthropic Principle (SAP) extends the AP to explain the actual nature of physical constants. Some suggest that the SAP leads to teleology or belief in a 'Divine Purpose'. The SAP can be extended to apply to an ensemble of Hubble spheres in a multiverse with varied laws of physics for each Hubble-like domain. In this respect the SAP puts constraints on the physical laws of the universe.

Darwinian style evolutionary biologists argue that man is probably unique in the universe because of the number of improbable evolutionary steps required to produce a Homo Sapiens. There is also the space travel argument against ETI in our galaxy. If ETI existed in our galaxy, by the Principle of Mediocrity (Earth evolution being typical) [68], ETI would have evolved sufficiently to be in our solar system because within 300 million years they would have explored the whole galaxy. AP arguments against Steady-State cosmology, where time is meaningless, state that ETI would also fill the whole universe [68].

However, another plausible theological argument follows from anthropic teleology. After thousands of years of the evolution of human consciousness history remains continuously littered with the

destruction of man getting along with man. One may make three reasonable assumptions by anthropic design.

1) Other stars are far away, probably by design. Proxima Centauri, the closest known star to Earth is 4.2 light years away, requiring some form of advanced Star Trek style warp drive technology which for us is still science fiction, but not necessarily for much longer [29,69].
2) In Judeo-Christian theology one might assume the prophesized millennium begins before a civilization develops warp drive capacity taking human intelligence off planet. (based on non-intervention by interplanetary 'Nazi' marauders and scripture hinting at the former situation).
3) Fully developed intelligence has no need for physical travel. A form of inter-dimensional superposition would be tantamount to providing an ontological holographic presence invisible to us but not to them.
4) We worship the 'God of Technology'; Intelligent life doesn't necessitate technology.

10. Anthropic Multiverse Calculations

We assume axiomatically from earlier work [14] that the universe is conformally scale-invariant from microcosm to macrocosm (i.e. no change in physical laws). But like most regimes there is a limit; because of fine-tuning parameters in the Multiverse, such a limit also applies to 'our' Hubble sphere, H_R. Also from prior work a formula for the coherent noetic force of consciousness, F_N (an anthropic potential) [86] was derived from Newton's law as $F_N = \aleph / \pm\rho$ in general unexpanded form, where $\aleph$ equals the 'ontological' energy mediating a conscious noetic field in noeons, the exchange unit of the unified field extrapolated from the Einstein (the physical unit defined as a mole of photons) and $\pm\rho$ represents a complex or rotational radius derived from momentum[5]. The formalism becomes complex

[5] Note that this noetic formalism takes the precise form as the fundamental equation for string tension ($T = E/L$) in M-theory and could be said to represent an alternative manner for its derivation [ham].

when expanded and applied to a variety of SUSY domains in its range of application; for example, $\aleph / \pm\rho$ becomes alternating dynamic-static Casimir boundary conditions [87-89] in spacetime brane domains of continuous-state symmetry breaking. This noetic formula is utilized to perform a proportionality order of magnitude calculation to see if the mass of the universe, M_U calculated from the philosophy of standard physics can be considered commensurate to the alternative derivation of mass from the Anthropic Multiverse model. The purpose of the calculation is to demonstrate the possibility that the collapse of the universe from the 11(12)D eternal realm to the subspace of our observed temporal reality of 3(4)D occurs by or is a property of the action of consciousness.

We make the axiomatic assumptions that:

1. All spirit is matter; and that all matter associated with a living system, under proper conditions may evolve to become spirit. (The latter being the task of theEarth's creation).
2. That the mass of the Earth, M_E is equivalent to the mass of intelligence for all life living during all time (~ 5 billion years) of the Earth's existence [66].

The mass of the universe, M_U is not nearly as precise as M_E and is derived from the Eddington and Dirac 'large number hypotheses' [14,90] for the suspected number of nucleons $N_p \sim 10^{80}$ in the universe. Since the mass of an average nucleon is $1.6\times10^{-24} g$, multiplication tell us $M_U \approx 10^{57} g$. This will be utilized as the standard M_U for our H_R to be compared below with the mass derived from a conscious universe, M_{UC}.

The first part of the alternative calculation is to derive the energy of Earth's intelligence I_E from its mass m_E by using Einstein's well known mass energy relation formula $E = mc^2$. With simple multiplication in *cgs* units using M_E as $6\times10^{27} g$ and the speed of light *c* as $3\times10^{10} cm/\sec$, we arrive at a magnitude of 5.4×10^{128} for I_E.

Next, we find the radius of the Earth E_R from its circumference $C = 2\pi R$. If the circumference of the earth is taken to be ~25,000 miles, converting to cgs we find $E_R = 1.6\times10^9 cm$.

Finally, from the Hubble radius, H_R of the universe ~14 billion light years; we convert to *cm* to again be in proper *cgs* form for our purposes. Since the speed of light, *c* is $3\times10^{10} cm/\sec$, and 60s x 60m x 24h x 365.4d seconds are in a year, we find H_R to be $9.14\times10^{17} cm$.

Using the noetic universe proportionality formula:

$$\frac{I_U}{H_R} \propto \frac{I_E}{E_R}, \qquad I_U = \frac{H_R \times I_E}{E_R} \tag{3}$$

we find for the intelligence of the universe

$$I_U = \frac{9.14\times10^{17} cm \times 5.4\times10^{128} cm}{1.6\times10^9 cm} = 3.1\times10^{137}. \tag{4}$$

Using Einstein's mass-energy equivalency formula $E = mc^2$ again we convert to the complex mass

$$m = \frac{E}{c^2} = \frac{3.1\times10^{137}}{3\times10^{10}\cdot 3\times10^{10}} = 3.33\times10^{116} \tag{5}$$

Finally, we assume this is a complex mass related to the wave function of the universe from the Wheeler-DeWitt equation (6) [91-94].

If we interpret result (5) in terms of Cramer's transactional model of quantum theory [55] where all off diagonal advanced-retarded components of the wave function are considered to be physically real standing wave future-past elements; we may use the standard quantum equation $\Psi = \psi\psi^*$ and take the square root of the complex mass, M_U of the $M_U^{'}$. This gives the result for real Euclidean/Minkowski space the mass $M_U \sim 10^{58} g$ which, within acceptable limits of error

for cosmological numbers, is the same mass of the universe calculated by conventional cosmological means!

We use this phantasmagoric albeit simple calculation to suggest that for an anthropic multiverse a quantum condition of our virtual reality acts as a force of collapse of the wave function (reality is like a continuous measurement) such that H_R, the limit of observation, is determined by the mass of intelligence, I_U or I_E. This is the basis for fine-tuning in nested holographic anthropic pockets where each pocket has its own laws of physics based on this anthropic principle. So, to be clear, this suggests that in another creation with a different mass of intelligence, the H_{ROC} would have a different radius. We did not address this condition in the dodecahedral wraparound section; but it is possible this element could produce a phase factor more significant than the subtractive interferometry of PDH circle pairs.

11. Wave Function of the Universe (WFU)

A number of authors have postulated a wave function of the universe satisfying what is known as the Wheeler-DeWitt equation (WDE) with the simplistic unexpanded form undefined in the general case

$$H\Psi = 0 \tag{6}$$

which is a supposed gravitational form of the Schrödinger equation suggesting the universe is described by a quantum wave function rather than a classical spacetime [91-95]. This equation has become the basis for theories of quantum gravity. The WDE puts a constraint on the Hamiltonian, $\hat{H}(x)\left|\psi\right\rangle = 0$ if quantizing general relativity so that $\left|\psi\right\rangle$ although the wavefunction of the universe, is not a spatial wavefunction as defined for a 3-space surface normalized to unity, but for configurations on all spacetime, and the Hamiltonian, $\hat{H}$ no longer determines system evolution. Cosmologically this universal wave function, Ψ depends on the action of gravity and matter on a 3-torus for a closed universe (in our cosmology, the universe is closed and finite in time (H_R), but open and infinite multiversally [13].

There are three main proposals – The Hartle-Hawking wave function [91,95], the Linde wave function [96] and the tunneling wave

function [97,98]. While this avenue has spurred years of creative thinking, it remains insufficient to solve the problem. Many scientists currently believe we live in a quantum universe. This is not true; we do not live in a quantum universe any more than we live in a Newtonian universe. We live in a multiverse that is a continuous anthropic state (like a complex HD standing wave) of complementary classical, quantum and unitary parameters [54,55] requiring an ontological interpretation of the WFU [99]. Feynman has said that there might not be a quantum gravity [100] for which no *a priori* reason that there should be one exists, and Anthropic Multiverse cosmology also suggests that this is true. Mach's principle is also a factor. The quantum regime has a limit in the same way the Newtonian regime does. The quantum world and gravity do not integrate with each other, but with the unified field. The purpose of the quantum regime is to separate the reduced or limited Newtonian absolute space (AS) from the 3rd regime HD AS [30,31]. This separation allows our temporal reality to 'surf' as it were on the face of eternity; and is why for Earthly observer's reality appears 3D.

12. Dimensionality by Subtractive Interferometry

The technology exists today to implement astrophysical subtractive interferometry; but it will probably wait some twenty-five years before NASA-ESA will budget such a program. For SETI research a wide baseline interferometer is set up with telescopes on opposite sides of the Sun. Such a telescope array is able to subtract out stellar coronas which normally block 'M-class' planets from view which orbit too close to the star to be observed by conventional telescopes.

Subtractive interferometry is required for two reasons. M-class planets are much too small to cause a stellar wobble which is the method used to discover the roughly five-hundred Jupiter sized extra-solar planets. When the stellar corona is subtracted out by interferometric techniques, spectroscopy may then be used to detect water and chlorophyll. It is these planets that are of paramount interest and where SETI searches should be concentrated for highest efficiency.

The technology to construct a pair of precision telescopes able to remain in formation while maintaining the required stability for observations is an engineering challenge, but researchers at NASA and

the European Space Agency (ESA) have programs to meet the challenge. The ESA is developing a formation-flying interferometer, called Darwin, planned for a joint NASA-ESA mission before 2020.

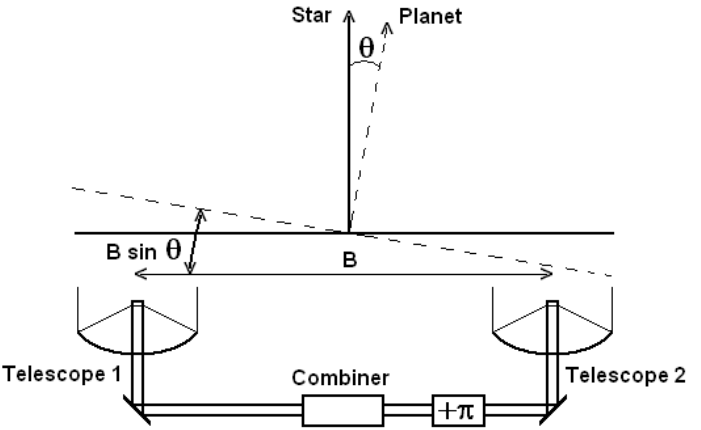


Fig. 3. Two telescope space-based nulling interferometer, simplest configuration for producing a dark fringe in the line of sight when properly phased. This destructive interference enables planet detection by subtracting stellar coronas when focusing the null region on the star's disk. Redrawn from [101].

13. Trans-Dimensional Q-Telescope – A New UFM Multiverse SETI Technology

Anthropic Multiverse cosmology proposes that a new type of interdimensional (universe to multiverse) Q-Telescope is required for an anthropic multiverse in order for SETI to be successful for the radical reasoning we have introduced above. The Q-Telescope requires the completed 12D version of Quantum Theory that allows utilization of the Noetic unified field to obtain the information of other planetary civilizations [13,81]. This information is lost to temporal 4D reality because it is only a knot crossing subspace [81] of the higher LSXD reality and some anthropic information is lost. Analogously this is the same as solving the problem of the 1st person - 3rd person barrier in philosophy of mind and is also similar to the EPR paradox where communication cannot be sent between 'Alice & Bob'.

The new technology proposed is named a Q-Telescope (Q to represent both the mental Qualia and the Q charge of radio) to be designed to take these new considerations into account and perform a new type of advanced multiversal EPR-like measurements on Proxima Centauri. Currently EPR correlated photons are only created or entangled by simultaneous emission of photon pairs from excited atoms such as mercury. We currently do not have the ability to entangle uncorrelated photons in an interferometer for example by any known method of parametric upconversion [79,80] to obtain entanglement in the nonlocal eternal present with no temporal separation. What is the relation between the two types of photons?

§2. In an anthropic multiverse because of inherent conformal invariance, the 1st person – 3rd person barrier said to exist between other minds is postulated to exist between extra-solar societies and our own because of the 1) phase variation of fine-tuning subtractive interferometry and 2) PDH wraparound (shield) cosmology making this information inaccessible in the usual 4D reality of standard measurement protocols.

How can this be logical since we can observe these planets? In the same manner, we can observe other people but not observe their minds through telescope or microscope, only the 3-space matter they are made of.

It is now technologically feasible (although NASA says it will be at least 25 years before budgeting) to deploy a dual satellite system to utilize what is called ‘subtractive interferometry’. This entails putting a satellite on each side of the sun. Currently astronomers have discovered over 500 extra-solar Jupiter class planets by observing the slight wobble their orbits cause in the stars they circle. Earth type planets, the ones deemed most likely to produce life as we know it, are too small to create stellar wobbles and close enough to their stars to be unobservable because they are hidden behind the stellar coronas. Subtractive Interferometry has the ability to cancel out these stellar coronas so that powerful telescopes can search for Earth class planets. Once Earth type planets are observed; and it is speculated that they are common, spectroscopic observations can look for water and bioluminescence. Earth class planets with water are currently anticipated to be the most likely venues to produce technological societies.

So rather than broadband radio frequency searches across the whole sky as currently being performed by SETI programs; only the Earth class planets discovered by subtractive interferometry would be investigated. These planets would be searched with the radically different approach using the new Q-Telescope protocol technology. The Q-Telescope would be designed to utilize

1. A multiversal correlated temporal adjustment accounting for the 'Adam's reckoning' and
2. Phase resonance protocols coupling to the 'Intelligent geon' state of quantum computing systems.

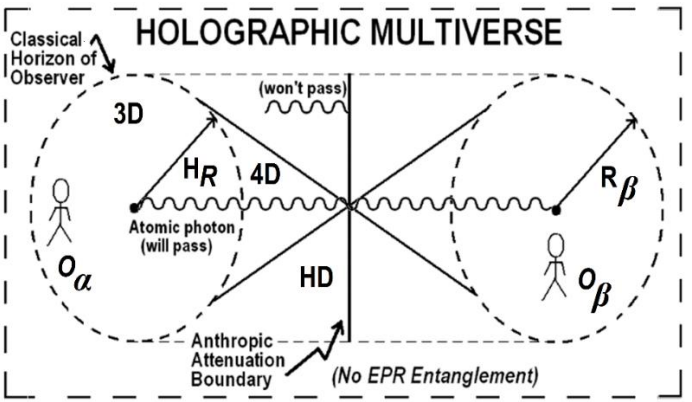


Fig. 4. In an Anthropic Multiverse causally separated nested Hubble spheres, H_R, R_β may have no EPR correlation for the 3-space H_R Earth SETI observer because of the different basis for the temporal 'reckoning' of the denizens so that atomic photons will pass but anthropic photons will not.

According to Fig. 4, while 'material' or atomic light in M_4 passes between the two causally separated (mind component) observers; because each observers awareness has a different 'reckoning' [67] consciousness related phenomena like TV or radio program waves will be out of phase and require an HD brane-world parametric up-conversion EPR-like process to synchronize the reckoning times as in the entangled state of simultaneously emitted photons in parametric down-conversion by Hg atoms [79,80]. We admit this seems ridiculously strange to current thinking but we get halfway there by

the clock paradox of special relativity where one twin space traveling relativistically away from Earth may age only 5 years while his Earthbound twin could age hundreds of years [102]. Could they observe each other? The addition of LSXD to 3rd regime M-Theory creates a paradigm shift [13,81,103-105].

> *Understanding holography in space-times other than Anti-de Sitter space, e.g. flat space and cosmological space–times. This throws up new conceptual challenges. For example, the boundary of Anti-de Sitter space is time-like, but the boundary of cosmological space–times is often space-like and for flat-space it is null. Extending holography to these cases, will surely be worthwhile* [106].

Because of the manner in which the skeptical scientific mind works, I suppose SETI-I is doomed to take some decades to fail before many will pay even a scintilla of serious attention to the views presented here [107]. But instead of beginning to design a Q-Telescope 25 years from now when NASA and SETI programs might be ready to listen; perhaps if scientists start thinking about it now Q-Telescope designs could be ready at the same time. Because I for one don't want to have to wait for Proxima Centauri radio or Barnard's star TV any longer than necessary [108]!!

Still if the SETI-I protocol fails as anticipated, empirical tests are required. The Lord God certainly does seem to work in mysterious ways; because if we had not delved into this whimsical work (speaking now in terms of conformal scale-invariance for certain physical principles) we would not have realized until much much later that these strange cosmological principles also apply directly to the microscopic scale in terms of the new noetic class medical paradigm we have been developing…[109].

References

[1] Josephson, B. (2017) Biological organization as the true foundation of reality, in R.L. Amoroso, G. Albertini, L.H. Kauffman & P. Rowlands (eds.) Unified Field Mechanics II, London: World Scientific.
[2] Shostak, S. (2006) Listening for ET's Television, www.space.com/searchforlife/seti_shostak_tv_061109.html.
[3] Dirac, P.A.M. (1952) Nature (London) 169, 702.
[4] Hainzl, C., Lewin, M., & Séré, É. (2005) Existence of a stable polarized vacuum in the Bogoliubov-Dirac-Fock approximation., Communications in mathematical physics, 257:3, 515-562.

[5] Friedmann, A. (1922) *Über die Krümmung des Raumes, Zeitschrift für Physik A*, 10 (1): 377–386.
[6] Friedmann, A. (1924) *Über die Möglichkeit einer Welt mit konstanter negativer Krümmung des Raumes, Zeitschrift für Physik A*, 21 (1): 326–332; English trans. Gen. Rel. Grav. 1999 vol. 31, 31–40.
[7] Lemaître, G. (1931) Expansion of the universe, A homogeneous universe of constant mass and increasing radius accounting for the radial velocity of extra-galactic nebulæ, Monthly Notices of the Royal Astronomical Society, 91: 483–490; Trans. Lemaître, G. (1927) *Un univers homogène de masse constante et de rayon croissant rendant compte de la vitesse radiale des nébuleuses extra-galactiques, Annales de la Société Scientifique de Bruxelles*, A47: 49–56.
[8] Lemaître, G. (1933) *l'Univers en expansion, Annales de la Société Scientifique de Bruxelles*, A53: 51–85.
[9] Robertson, H.P. (1935-37) Kinematics and world structure I,II,III, Astrophysical Journal, 82: 284–301; 83: 187–201, 83: 257–271, respectively.
[10] Walker, A.G. (1937) On Milne's theory of world-structure, Proceedings of the London Mathematical Society 2, 42 (1): 90–127.
[11] de Sitter, W. (1917) On the relativity of inertia: Remarks concerning Einstein's latest hypothesis, Proc. Kon. Ned. Acad. Wet., 19: 1217–1225.
[12] de Sitter, W. (1917) On the curvature of space, Proc. Kon. Ned. Acad. Wet., 20: 229–243.
[13] Amoroso, R.L. & Rauscher, E.A. (2009) The Holographic Anthropic Multiverse: Formalizing the Ultimate Geometry of Reality, Singapore: World Scientific.
[14] Kafatos, M, Roy, S & Amoroso, R. (2000) Scaling in cosmology and the arrow of time, in R. Buccheri, V. di Gesu & M. Saniga (eds) Studies on the Structure oF Time, Dordrecht: Kluwer Academic.
[15] Luminet, J-P, Weeks, J.R., Riazuelo, A., Lehoucq, R., & Uzan, J.P. (2003) Dodecahedral space topology as an explanation for weak wide-angle temperature correlations in the cosmic microwave background, Nature, 425(6958) 593-595.
[16] Luminet, J-P (2008) The wraparound universe, Wellesley: AK Peters.
[17] Luminet, J-P (2005) A cosmic hall of mirrors, https://arxiv.org/ftp/physics/papers/0509/0509171.pdf;
[18] Luminet, J-P (2007) Geometry and Topology in Relativistic Cosmology, https://arxiv.org/pdf/0704.3374.
[19] Luminet, J-P (2008) The Shape and Topology of the Universe, arXiv:0802.2236v1 [astro-ph].
[20] Weeks, J. (2004) The Poincaré Dodecahedral Space and the Mystery of the Missing Fluctuations, Notices of the AMS Vol. 51, No. 6., 610-619.
[21] Cornish, N.J., Spergel, D.N., Starkman, G.D. & Komatsu, E. (2004) PRL 92, 201302.
[22] Roukema, B.F., Lew, B, Cechowska, M., Marecki, A. & Bajtlik, S. (2004) Astron. Astrophys. 423, 821.
]23] Aurich, R., Lustig, S.& Steiner, F. (2006) MNRAS 369, 240.
[24] Then, H. (2005) M.N.R.A.S. 373, 139.
[25] Key, S.J., Cornish, N.J., Spergel, D.N. & Starkman, G.D. (2006) astro-ph/0604616.

[26] Riazuelo, A., Uzan, J-P, Lehoucq, R. & Weeks, J. (2004) Phys Rev. D69,103514.
[27] Dragić, A., Marić, Z. & Vigier, J-P (2000) New quantum mechanical tight bound states and 'cold fusion' experiments, Physics Letters A: 265;163-167.
[28] Vigier, J-P (1993) New Hydrogen (Deuterium) Bohr Orbits, Proceedings ICCF4, Hawaii, Vol 4, p. 7.
[29] Rauscher, E.A. & Amoroso, R.L. (2011) Orbiting the Moons of Pluto: Complex Solutions to the Einstein, Maxwell, Schrödinger, and Dirac Equations, Singapore: World Scientific.
[30] Amoroso, R.L. Rowlands, P. & Kauffman, L.H. (eds.) Physics of Reality: Space, Time, Matter, Cosmos, London: World Sci.
[31] Amoroso, R.L., Kauffman, L.H. & Rowlands, P. (2016) Unified Field Mechanics: Natural Science Beyond the Veil of Spacetime, London: World Scientific.
[32] Amoroso, R.L., & Vigier, J-P (2003) The origin of CMBR as intrinsic blackbody cavity-QED resonance inherent in the dynamics of the continuous state topology of the Dirac vacuum, in Gravitation and Cosmology: From the Hubble Radius to the Planck Scale, Dordrecht: Kluwer, 27-38.
[33] Amoroso, R.L., Kafatos, M. & Ecimovic, P. (1998) The origin of cosmological redshift in spin exchange vacuum compactification and nonzero rest mass photon anisotropy, in G. Hunter, S. Jeffers & J-P Vigier (eds.) Causality & Locality in Mod. Physics, Dordrecht: Kluwer.
[34] Amoroso, R.L. (2009) On the Possibility of Relativistic Shock-Wave Effects in Observations of Quasar Luminosity, in R.L. Amoroso (eds.) The Holographic Anthropic Multiverse: 326-350, Singapore: World Scientific; http://vixra.org/pdf/1506.0211v1.pdf.
[35] Amoroso, R.L. (2012) Unified Geometrodynamics: A Complementarity of Newton's and Einstein's Gravity, in R.L. Amoroso (eds.) The Physics of Reality: Space, Time, Matter, Cosmos, 152-163; London: World Scientific; http://vixra.org/pdf/1403.0919v1.pdf.
[36] Amoroso, R.L. & Pribram, K.H. (2009) The Complementarity of Mind and Body: Realizing the Dream of Descartes, Einstein and Eccles, Cambridge: MIT Press.
[37] Amoroso, R.L. (2006) The Holographic Conscious Multiverse, in R.L. Amoroso, B. Lehnert & J-P Vigier (eds.) Beyond the Standard Model: The Search for Unity in Physics, Orinda: Noetic Press.
[38] Amoroso, R.L. (2003) The Fundamental Limit and Origin of Biological Systems, Noetic Journal 4:1; 24-32.
[39] Amoroso, R.L. & Amoroso, P.J. (2004) The Fundamental Limit and Origin of Complexity in Biological Systems: A New Model for the Origin of Life, in D.M. Dubois (ed.) Computing Anticipatory Systems, AIP Conf. Proceedings V. 718, pp. 144-159, Melville: Am Inst. Phys.
[40] Amoroso, R.L. (2002) Developing the cosmology of a continuous state universe, in R.L. Amoroso, G. Hunter, M. Kafatos & J-P Vigier (eds.) Gravitation & Cosmology: From the Hubble Radius to the Planck Scale, Dordrecht: Kluwer.
[41] Amoroso, R.L. (2018) Millennial Science: The New Age of Discovery's 'Conscious' Technologies, in I. Fredriksson (ed.) Applications of Consciousness, Essays on Technologies, Therapies and Arts, McFarland.

[42] Crick, F. (1994) The Astonishing Hypothesis: The Scientific Search for the Soul, New York: Touchstone.
[43] Rowlands, P. (2007) Zero to Infinity: The Foundations of Physics, Singapore: World Scientific.
[44] Plato (1989) The Collected Dialogues of Plato: Including the Letters, E. Hamilton & H. Cairns (eds.) L. Cooper (trans.) Princeton: Bollingen Series LXXI.
[45] Meehl, P.E. (1966) The compleat autocerebroscopist: A thought experiment on professor Feigel's mind-body identity thesis, in P.K. Feyerabend, & G. Maxwell, Mind, Matter and Method; Essays in Philosophy and Science in Honor of Herbert Feigl, pp. 103-180.
[46] Musser, G. (1998) String Instruments, Sci. Am., V.10, pp. 17-19.
[47] Drake, F. & Sobel, D. (1992) Is Anyone Out There? New York: Delacorte Press.
[48] Beyond the Drake Equation (does calculations) www.station1.net/DouglasJones/drake.htm#dolphin
[49] www.as.utexas.edu/astronomy/education/drake/drake.html; or www.classbrain.com/artmovies/publish/article_50.shtml; or www.pbs.org/wgbh/nova/origins/drake.html; http://frombob.to/drake.html.
[50] Exoplanet search count, http://exoplanet.eu
[51] Smith, J. (1974) The Doctrine and Covenants, D&C 88: 11-13, Salt Lake City: The Church of Jesus Christ of Latter-Day Saints.
[52] Einstein, A. (1917) *Kosmologische Betrachtungen zur allgemeinen Relativitätstheorie, Sitzungsberichte der Preussischen Akademie der Wissenschaften,*142–152; Also Cosmological considerations on the general relativity theory, pp. 177–188; (1923) in The Principle of Relativity, a collection of original memoirs on the special and general theory of relativity, H.A. Lorentz, A. Einstein, H. Minkowski, & H. Weyl, with notes by A. Sommerfeld; W. Perrett & G. B. Jeffery (trans.) London: Methuen & Co.; reprinted (1990); also (1952) New York: Dover.
[53] Kant, E. (1787) *Kritik der reinen Vernunft*, Critique of Pure Reason (1996) W. Pluhar (trans.) Indianapolis: Hackett.
[54] Wheeler, J.A. & Feynman, R. (1945) Rev. Mod. Physics, 17, 157.
[55] Cramer, J. (1985) The Transactional interpretation of Quantum Mechanics, Rev. Mod. Phys 58, 647-687.
[56] Bass, L. & Schrödinger, E. (1955) Proc. R. Soc. Lon A 232,1-6.
[57] Einstein, A. (1917) Ann. Phys., vol. 18, p. 121
[58] Zwicky, F. (1929) Proc. Nat. Ac. Sc., Washington, 15, 773.
[59] Finlay-Freundlich, E. (1953) *Göttinger Nachrichten*, 7, 95-102.
[60] Finlay-Freundlich, E. (1954) Philosophical Magazine, 45, 303-319.
[61] Finlay-Freundlich, E. (1954) Proceedings of the Physical Society A 67, 192-193.
[62] Born, M. (1954) *Nachr. Ak. Wiss. Göttingen*, 7, 102.
[63] Pecker, J-C., Roberts, A.P., & Vigier, J-P (1972) Non-velocity redshifts and photon-photon interactions, Nature, v. 237, p. 227-229).
[64] Pecker, J-C & Vigier, J-P (1988) A Possible Tired-Light Mechanism, Apeiron, No. 2, pp 19-23.
[65] Vigier, J-P (1990) Evidence for nonzero mass photons associated with a vacuum-induced dissipative redshift mechanism, IEEE Trans Plasma Sci, 18:1;64-

72.
[66] Amoroso, R.L. (2016) Age of the Earth: 'Young Earthers' versus evolutionists, Sci God J, Vol 7, No 7, p. 433-6, http://scigod.com/index.php/sgj/article/view/492/540; vixra.org/pdf/1410.0102v1.pdf.
[67] Bible, New Testament, King James Version, 2 Peter 3:8.
[68] Barrow, J.D. & Tipler F.J. (1996) The Anthropic Cosmological Principle, Oxford: Oxford Univ. Press.
[69] Amoroso, R.L. (2011) SSE Talks - Faster-than-light Space Travel? https://www.youtube.com/watch?v=ynxD-BaDxRk&t=1s.
[70] Smith, J. (1974) The Doctrine and Covenants, D&C 130: 4-5, Salt Lake City: Church Jesus Christ of Latter-Day Saints.
[71] Smith, J. (1974) The Doctrine and Covenants, D&C 88:7-11, Salt Lake City: Church Jesus Christ of Latter-Day Saints.
[72] Wheeler, J.A. (1955) Geons, Physical Review, 97:2, 511-536.
[73] Shadowitz, A. 1968, Special Relativity, New York: Dover.
[74] De Broglie, L. (1927) *J. Physique, 6e serie* 8, 225.
[75] Kulhánek, J. (1965) Propagation of de Broglie waves in space-time, Il Nuovo Cimento, Volume 38:3; 1955-1965.
[76] Feoli, A. & Scarpetta, G. (1998) De Broglie Matter Waves from the Linearized Einstein Field Equations, Foundations Phys L, 11:4; 395-403.
[77] Jung, C. (1970) Collected Works of C. G. Jung, Princeton: Princeton Univ. Press.
[78] Franck, G. (2000) Time and presence, in Science and The Primacy of Consciousness, R.L. Amoroso et al, (eds.) Orinda: Noetic Press.
[79] Marshall, T.W. (2002) Engineering the vacuum, pp. 459-468, in R.L. Amoroso, G. Hunter, M. Kafatos & J-P Vigier (eds.), Gravitation & Cosmology: From the Hubble Radius to the Planck Scale, Dordrecht: Kluwer Academic
[80] Marshall, T.W. (1998) Parametric up conversion of the vacuum; http://arxiv.org/PS_cache/quant-ph/pdf/9803/9803054.pdf.
[81] Amoroso, R.L (2017) Universal Quantum Computing: Supervening Decoherence, Surmounting Uncertainty, London: World Scientific.
[82] Sagan, K & Drake, F. (1975) The search for extraterrestrial intelligence, Scientific American, May
[83] Sagan, K. (1994) Search for extraterrestrial life, Sci Am, October.
[84] Tarter, J.C. & Chyba, C.F. (1999) Is there life elsewhere in the universe? Scientific American, 281:6, pp. 118-123.
[85] www.answerbag.com/q_view/506418.
[86] Amoroso, R.L (2000) Derivation of the fundamental equation of consciousness, Part I, Boundary conditions, Noetic J 3:1, pp. 91-99.
[87] Schwinger, J. (1992) Casimir energy for dielectrics, Proc. Nat. Acad. Sci. USA 89, 4091-3.
[88] Schwinger, J. (1993) Casimir light: The source, Proc. Nat. Acad. Sci USA 90, 2105-6.
[89] Schwinger, J. (1994) Casimir energy for dielectrics: spherical geometry, Proc. Nat. Acad. Math. Psych. 41:64-67, SF: W.H. Freeman.

[90] Carter, B. (1974) Large number coincidences and the Anthropic Principle, in M.S. Longair (ed.) Confrontation of Cosmological Theory with Astronomical Data, Dordrecht: Reidel.
[91] Hartle, J.B. & Hawking, S.W. (1983) Phys. Rev. D28, 2960.
[92] Halliwell, J.J. (2000) Trajectories for the wave function of the universe from a simple detector model; arXiv:gr-qc/0008046.
[93] Vilenkin, A. (1998) The quantum cosmology debate, arXiv:gr-qc/9812027.
[94] Smith, Q. (1997) The Ontological Interpretation of the Wave Function of the Universe, The Monist, 80:1,160-185.
[95] Penrose, R. (2005) The Road to Reality, New York: Knopf.
[96] Linde, A.D. (1984) Lett. Nuovo Cimento, 39, 401
[97] Vilenkin, A. (1984) Phys. Rev. D30, 509.
[98] Vilenkin, A. (1986) Phys. Rev. D33, 3560.
[99] Callender, C. & Weingard, R. (1994) The Bohmian model of quantum cosmology, PSA, vol. 1, D. Hull et al. (eds.) pp. 218-27.
[100] Feynman, R.P. (1971) Lectures on Gravitation, Pasadena: California Inst of Technology.
[101] Velusamy, T. et al (2001) Single and double bracewell nulling interferometer in space, Proc. Toward Other Earths: DARWIN/TPF; and the Search for Extrasolar Terrestrial Planets (2003) Heidelberg, Germany, (ESA SP-539.
[102] Bohm, D. (1965) The Special Theory of Relativity, New York, W.A. Benjamin.
[103] Trivedi, S.P. (2001) Holography, black holes and string theory, Current science, 81:12; 1582 -1590.
[104] Randall, L. (2010) Warped Passages. Randall, L. (2006). Warped passages: unravelling the universe's hidden dimensions. Penguin UK.
[105] Randall, L. (2007) The case for extra dimensions. Physics Today, 60(7), 80-81.
[106] Randall, L. & Sundrum, R. (1999) An alternative to compactification, Physical Review Letters, 83(23), 4690; https://arxiv.org/pdf/hep-th/9906064.pdf.
[107] Basalla, G. (2006) Civilized Life in the Universe: Scientists on Intelligent Extraterrestrials, Oxford: Oxford Univ. Press.
[108] Amoroso, R.L & Vigier, J-P (2012) Evidencing ‘tight bound states’ in the hydrogen atom: empirical manipulation of large-scale XD in violation of QED, in R.L. Amoroso (eds.) The Physics of Reality: Space, Time, Matter, Cosmos, 254-272, London: World Scientific; http://rxiv.org/pdf/1305.0053v2.pdf.
[109] Amoroso, R.L. (2010) A seminal introduction to imminent parameters of noetic medicine, in W. Carvalho & Pelizoli, M. (eds.) Proc. 1st Symp. for ‘Cure and Consciousness’ Nov. 6-8, Recife, Brazil.

BINARY NECKLACE SYSTEMS

JOHN C. AMSON

School of Mathematics and Statistics
University of St Andrews, Scotland
Contact : Anstruther, KY10 3DY, Fife, Scotland
john.amson@uwclub.net

A REPORT ON WORK IN PROGRESS

ABSTRACT: A novel algebraic system of Binary Necklaces, which is non-Archimedean normed, non-associative, and non-distributive, is introduced, explored and developed. An application to Binary Cellular Automata is noted.

NECKLACES

The popular notion (in Combinatorial Theory) of a necklace is that it is a 'canonical representative' of an equivalence class of 'bitnecklaces' which can be visualised as 'strings-of-black-and-white-beads' subject to certain symmetry rules.

As a string-of-beads a bit necklace can be associated with a plane polygon whose vertices are the beads. Such a bitnecklace laid on a flat surface is not changed by rotating it; the associated polygon has cyclic rotational symmetry about its centre.

[Note that a bitnecklace may be changed by 'flipping it over' — raising it off the table, turning it over, and replacing it. It is not changed by flipping if and only if the associated polygon has dihedral (reflection) symmetry about one of its diagonals (i. e. if and only if it exhibits 'mirror-symmetry', *alias* 'handed-ness', 'chirality', or 'enantiomorphism'). Plane polygons can be deemed to be equivalent under rotations and reflection, i. e. 'dihedral symmetry equivalence'

under rotations-and-reflection. However, in this paper, we do not consider bitnecklaces with dihedral symmetry, so-called 'BitBracelets'; they are left aside for a future paper.]

In this paper, plane polygons are deemed to be equivalent under rotations : i. e. they possess 'cyclic symmetry equivalence'.

Plane polygons can then be collected together into cyclic equivalence classes, any one member of which may be regarded as the 'canonical' or 'representative' polygon of its equivalence class.

In this paper I shall use the novel notion of requiring the canonical representative of an equivalence class of bitnecklaces / polygons to be the one which has the least decimal value of the associated string-of-bits when the latter is regarded as a binary number in big-endian format. The notion has some similarities with those in 'Lyndon Word' theory, but is differrent from and not to be confused with that theory. My unusual 'least decimal value' technique allows us to introduce a wide range of structural properties which lead us to refer to the resulting quite elaborate systems as "Binary Necklace Systems".

I shall introduce and begin to explore many such properties in this paper whilst at the same time recognising that there are very many more still awaiting formulation and development.

Of special interest is that each generation of a 1-dimensional 2-colour Cellular Automaton with (cyclical) wrap-around edge conditions can be regarded as a Binary Necklace.

This suggests that a finite collection of generations of such a Cellular Automaton can be equipped with the kind of normed algebraic structure being introduced in this paper. This notion, and the extension of the theory to 'Neckloids' (Necklaces-of-Necklaces, . . .), will be considered in a later paper.

1 BITSTRINGS, BITNECKLACES, NECKLACES

1.1 BITSTRINGS

We recall that a *bitstring* of *length* n is an ordered list (written using round brackets) $(\, b_n, \ldots, b_1 \,)$ of bits b_i $(i = n \ldots 1)$, each *bit* being associated uniquely with an integer in the finite **2**-field $\{0, 1\}$. The ordering is inherited from the naturally ordered set of integers $\mathbf{Z}$.

Note particularly, that I prefer to write bitstrings from Right to Left — this makes bitstrings immediately compatible with the familiar 'big-endian' format of binary numbers — with the most significant digit at the Left. Bitstrings can also be written as $(\, b_i \,)_{i \in \{n \ldots 1\}}$ or as $(\, b_i \,)$ (when convenient).

Bitstrings can be 'compared' — specifically can be added and multiplied — 'component-wise':

$$\begin{aligned} (\, a_i \,) + (\, b_i \,) \quad &= \quad (\, c_i \,) \stackrel{\text{def}}{=} (\, a_i +_2 b_i \,) \quad \text{i.e.} \, ((a_i - b_i)^2 - a_i\, b_i \,) \\ &\qquad\qquad (\, +_2 \text{ is integer addition mod } 2 \,) \\ (\, a_i \,) \times (\, b_i \,) \quad &= \quad (\, c_i \,) \stackrel{\text{def}}{=} (\, a_i \times_2 b_i \,) \quad \text{i.e.} \, (\, a_i\, b_i \,) \\ &\qquad\qquad (\, \times_2 \text{ is integer multiplication mod } 2 \,) \end{aligned}$$

And it is known that the set of all bitstrings of a fixed length n forms a commutative ring over the **2**-field; an algebraic ring isomorphic to a Boolean ring when bit-addition and bit-multiplication are interpreted as Boolean `XOR` and `AND`.

1.2 BITNECKLACES

A *bitnecklace* of *circumference* $n \geqslant 1$ is a bitstring of length n in which the linear index ordering is replaced by a *cyclical* ordering

$$1 < 2 < \cdots < n < 1 \,.$$

A bitnecklace will now be denoted (with square brackets) by $[\,b_n \ldots b_1\,]$ (written again from Right to Left); or as $[\,b_i\,]_{i\in\{n\ldots1\}}$ or as $[\,b_i\,]$ if or when convenient.

" $\approx$" : The set of all bitnecklaces of fixed circumference n form an equivalence class under the cyclical permutation of the indices of it members. Thus for each cyclical permutation $\kappa \in \mathfrak{C}^{(n)}$ (the group of cyclic symmetries of size n) the bitnecklaces $[\,b_n \ldots b_1\,]$ and $[\,b_{\kappa(n)} \ldots b_{\kappa(1)}\,]$ are (cyclically) equivalent (and written as $[\,b_n \ldots b_1\,] \approx [\,b_{\kappa(n)} \ldots b_{\kappa(1)}\,]$). For example,

$$\{\,[\,b_3\,b_2\,b_1\,],\ [\,b_2\,b_1\,b_3\,],\ [\,b_1\,b_3\,b_2\,]\,\}$$

is the equivalence class of all bitnecklaces of circumference n=3, with 3 bits (*cf.* 'beads') b_1, b_2, b_3.

The cardinalities of these bitnecklace equivalence classes (i. e. the number of bitnecklaces in each equivalence classes) are given by a well-known combinatoric formulae which is shown, with examples, in Appendix **C**.

" $\hookrightarrow$" : It is sometimes useful to denote the cyclical-rotation of equivalent bitnecklaces with a 'hook-arrow', e. g.
$[1,1,0,0] \hookrightarrow [0,1,1,0] \hookrightarrow [0,0,1,1,] \hookrightarrow [1,1,0,0]$

Definition 1.1:
Let $\mathbf{b} = [\,b_n \ldots b_1\,]$ be a bitnecklace of circumference $n \geqslant 1$.

(a) By the **Hamming Norm** $\|\mathbf{b}\|$ of $\mathbf{b}$ we mean the integer $\|\mathbf{b}\| = \sum_{i=0}^{n} b_i$ *(where the summation is ordinary integer-summation of bits as integers).*
(b) By the **Binary Value** $\mathtt{Bin}(\mathbf{b})$ of $\mathbf{b}$ we mean the Binary Number $b_n \ldots b_2 b_1$ where each bit b_i now represents a binary digit, and b_1 is the least significant binary digit *(***'big-endian'** *format).*
(c) By the **Decimal Value** $\mathtt{Dec}(\mathbf{b})$ of $\mathbf{b}$ we mean the conversion to its decimal value of the binary value $\mathtt{Bin}(\mathbf{b})$.
(d) By the **Significance** $\mathtt{Sif}(\mathbf{b})$ of $\mathbf{b}$ we mean index k of the most significant binary digit in the Binary Value

$\mathtt{Bin}(\mathbf{b})$ of $\mathbf{b}$, thus $b_k = 1$ and for each index $j > k$ we have $b_j = 0$ and for at least one index $r \leqslant k$ we have $b_r = 1$.

REMARKS 1.2:

1. Plainly $\|\mathbf{b}\|$ is the number of 1-valued bits b_i in $\mathbf{b}$; and each cyclically equivalent bitnecklace to $\mathbf{b}$ has the same Hamming Norm, i. e. Hamming Norm is invariant under cyclic symmetry.
2. $\|\mathbf{b}\| = \sum_{i=1}^{n} b_i \quad \leqslant \quad \sum_{i=1}^{n} 2^{b_i} = \mathtt{Dec}(\mathbf{b})$.
3. $\mathtt{Sif}(\mathbf{b}) = 1 + \mathtt{trunc}(\log[2](\mathtt{Dec}(\mathbf{b})))$.
4. Cyclically equivalent bitnecklaces do not necessarily have the same Binary Value or Decimal Value or Signifcance. i. e. Binary and Decimal Values and Significance are not invariant under cyclic symmetry.

$$\begin{gathered} e.\,g. \quad [1,0,0,0] \approx [0,0,0,1], \\ \|[1,0,0,0]\| = \|[0,0,0,1]\| = 1, \\ \mathtt{Bin}([1,0,0,0]) = 1000, \qquad \mathtt{Bin}([0,0,0,1]) = 0001, \\ \mathtt{Dec}([1,0,0,0]) = 16, \qquad \mathtt{Dec}([0,0,0,1]) = 1, \\ \mathtt{Sif}([1,0,0,0]) = 4, \qquad \mathtt{Sif}([0,0,0,1]) = 1. \end{gathered}$$

1.3 NECKLACES

DEFINITION 1.3:

(a) By a **Necklace** x *of circumference* $n \geqslant 0$ we mean the member *with least Decimal Value* of the equivalence class of bitnecklaces $[x_n, \ldots, x_1]$ under cyclical symmetry, and denoted by $x = \langle x_n, \ldots, x_1 \rangle$, or $\langle x_i \rangle_{i \in \{n \ldots 1\}}$, or $\langle x_i \rangle$ (with bent brackets).
The set of all Necklaces of circumference $n \geqslant 0$ is denoted by $\mathsf{NK}^{[n]}$.
The circumference of a Necklace x is denoted by $\mathtt{Circ}(x)$.
In the exceptional case when $n = 0$, then $x = \langle\ \rangle$ is the 'Empty Necklace' and $\mathsf{NK}^{[0]}$ is the 'Empty Necklace Set'.

(b) By the **dual** x' of a necklace $x = \langle x_i \rangle_{i \in \{n \ldots 1\}}$ we mean the member with *least Decimal Value* of the equivalence

class of bitnecklaces $[x'_n, \ldots, x'_1]$ bit-by-bit-dual to those bitnecklaces $[x_n, \ldots, x_1]$ in the equivalence class for the Necklace x. (*N.b.* each dual bit $x'_i = 1 + x_i$, i. e. $0 \leftrightarrow 1$).

(c) The necklace x of circumference n which has all its terms $x_i = 0$ is called the **zero-necklace** of circumference n and denoted by $0 = \langle 0, .^n., 0 \rangle$.

(d) The necklace x of circumference n which has all its terms $x_i = 1$ is called the **unity-necklace** of circumference n and denoted by $U = \langle 1, .^n., 1 \rangle$. [U and 0 are dual.]

(e) Necklaces with circumference $n = 0$ or $n = 1$ are said to be 'trivial'; otherwise they are 'non-trivial'.

DEFINITION 1.4:
By the **Hamming Norm** $\|x\|$, **Binary Value** $\mathtt{Bin}(x)$,
Decimal Value $\mathtt{Dec}(x)$, **Significicance** $\mathtt{Sif}(x)$,
of a *necklace* $x = \langle x_i \rangle$ we mean that value, respectively, of the bitnecklace $[x_i]$ corresponding to the *necklace* x.

REMARKS 1.5:
(1) The Decimal Value $\mathtt{Dec}(x)$ of every non-zero Necklace x of circumference $n \geqslant 1$ is odd : $n = 2m + 1$, $m \geqslant 0$.
(2) For each odd integer $n = 2m + 1$, $m \geqslant 0$ there exists at least one Necklace x (of sufficiently large circumference) with $\mathtt{Dec}(x) = n$:

E.g. take $n = 73$ (in binary 1001001 with 7 digits), then we have infinitely many Necklaces $x = \langle 0, \ldots, 0, 1, 0, 0, 1, 0, 0, 1 \rangle$ with $\mathtt{Dec}(x) = 73$ and circumference $K \geqslant 9 = (7 + 2)$, and $\kappa \geqslant 1$ leading bits '0'.

(Note that the appropriate necklaces must have large enough circumference — the bitnecklace $[1, 0, 0, 1, 0, 0, 1]$ is also cyclically equivalent to a Necklace $y = \langle 0, 0, 1, 0, 0, 1, 1 \rangle$, circumference $K = 7$ and $\mathtt{Dec}(y) = 19 \neq 73$; and $z = \langle 0, 0, 1, 0, 0, 1, 0, 1 \rangle$ circumference $K = 8$ and $\mathtt{Dec}(z) = 37 \neq 73$.)
(3) There is an alternative definition of a 'Necklace' which involves 'Lyndon Words' ([LyW]). Lyndon Word necklaces arise as the the unique minimum element in the lexicographical ordering of all its cyclical rotations. In the case of big-endian binary strings this not the same as our ordering by Decimal Value : for example, consider the 3-bit-string 101 :

its lexicographical ordering is	$101 < 011 < 110$
but its decimal value ordering is	$011 < 101 < 110$
	i. e. $3 < 5 < 6$.

2 ADDING, MULTIPLYING, CONCATENATING, SNAPPING NECKLACES

Because of the cyclic nature of a bitnecklace it is not feasible to speak of the addition or multiplication 'component-wise' of two of them (as we would in the case of e. g. strings or vectors) — there is no way of uniquely associating which component of one bitnecklace is to be matched with which component of the other one.

In our 2005 paper [AB][1], Keith Bowden and I showed that one possible way of dealing with this dilemma was to consider all n possible cyclical matchings of the components of the two bitnecklaces of circumference n simultaneously and then to construct a single output 'sum' or 'product' bitnecklace *via* suitably modified versions of inner-products. This method though ingenious was unsatisfactory, e. g. addition was not surjective : many bitnecklaces in the set of all n-circumference bitnecklaces, were not the sum of any two bitnecklaces in that set.

Instead of attempting to add or multiply bitnecklaces as such, we shall now propose a way of adding and multiplying **necklaces** — cyclically-equivalence-class **representatives** — a procedure which is much simpler, and which exhibits novel algebraic properties. Here is the construction for doing this:

CONSTRUCTION 2.1:
Let $x = \langle x_i \rangle_{i \in \{n \ldots 1\}}$ and $y = \langle y_i \rangle_{i \in \{n \ldots 1\}}$ be two Necklaces of the same circumference n.

1 *N.b.* Bitnecklaces *etc.* were there called "bit-hoops".

" + " *Construct*

(i) the class of all bitnecklaces cyclically equivalent to the bitnecklace $z = [z_i]$ with each $z_i = x_i +_2 y_i$, (where $+_2$ is 'integer addition mod 2'), and

(ii) the Necklace $s = \langle s_i \rangle_{i \in \{n \ldots 1\}}$ representative of that equivalence class.

Plainly, s is that bitnecklace whose decimal value is least[2] of all bitnecklaces equivalent to z.

" x " Repeat a similar *construction* but with z replaced by w where $w = [w_i]$ with each $w_i = x_i \times_2 y_i$, (where $\times_2$ is 'integer multiplication of bits 0, 1'), and s is replaced by $p = \langle p_i \rangle_{i \in \{n \ldots 1\}}$, that bitnecklace whose decimal value is least of all bitnecklaces equivalent to w.

CONSTRUCTION 2.2:

" cat " Let $x = \langle x_i \rangle_{i \in \{n \ldots 1\}}$ and $y = \langle y_j \rangle_{j \in \{m \ldots 1\}}$ be two Necklaces of circumferences n, m respectively

Construct

(i) the class of all bitnecklaces cyclically equivalent to the bitnecklace $z = [z_k]_{k \in \{K \ldots 1\}}$, where $K = n + m$,
and $z_k = y_j$ for each $k = i$, $1 \leqslant j \leqslant m$
and $z_k = x_i$ for each $k = j + n$, $1 \leqslant i \leqslant n$.

(ii) the Necklace $c = \langle c_k \rangle_{k \in \{K \ldots 1\}}$ representative of that equivalence class. (Note that 'cat' is formally cat-LEFT.)

" pad " In the above, let us take the first necklace x to be 0_n the zero-necklace of circumference $n > 0$, and carry through the same Construction for the Necklace $c = \langle c_k \rangle_{k \in \{K \ldots 1\}}$ with $K = n + m$ again. (Note that 'pad' is formally pad-LEFT.)

" snap " Let $x = \langle x_i \rangle_{i \in \{n \ldots 1\}}$ be a Necklace of circumference $n \geqslant 2$, and choose integer K such that $1 \leqslant K \leqslant n$.

Construct

(i) the class of all bitnecklaces cyclically equivalent to the

2 *Cf.* the related notion of 'least lexicographic order' in 'Lyndon Word' theory ([LyW]) as noted above (Def. 1.5).

bitnecklace
$u = [u_k]_{k\in\{n\cdots K+1\}}$, $u_k = x_k$ $(n \geqslant k \geqslant K{+}1)$, respectively
$v = [v_h]_{h\in\{K\cdots 1\}}$, $v_h = x_h$ $(K \geqslant h \geqslant 1)$.
(ii) the Necklaces $L = \langle u_k \rangle_{k\in\{n\cdots K+1\}}$, respectively
$R = \langle v_h \rangle_{h\in\{K\cdots 1\}}$, representative of those respective equivalence classes — with the proviso that in the exceptional cases :
where $K = 1$, then $L = x$ again,
and R is the Empty Necklace, or
where $K = n$, then L is the Empty Necklace,
and $R = x$ again.

Definition 2.3:
1. Let $x = \langle x_i \rangle_{i\in\{n\ldots 1\}}$ and $y = \langle y_i \rangle_{i\in\{n\ldots 1\}}$ be two Necklaces of the same circumference n.
(a) By the **sum** $x \boxplus y$ we mean the necklace $s = \langle s_i \rangle_{i\in\{n\ldots 1\}}$ constructed as above at (2.1).
(b) By the **product** $x \boxtimes y$ we mean the necklace $p = \langle p_i \rangle_{i\in\{n\ldots 1\}}$ constructed as above at (2.1).

2. Let $x = \langle x_i \rangle_{i\in\{n\ldots 1\}}$ and $y = \langle y_j \rangle_{j\in\{m\ldots 1\}}$ be two Necklaces of circumferences n, m respectively, and let $K = n + m$.
(c) By the **concatenation** $x\,\mathtt{cat}\,y$ *(or* $\mathtt{cat}(x, y)$ *)* we mean the necklace $c = \langle c_k \rangle_{k\in\{K\ldots 1\}}$ of circumference $K{=}n{+}m$ constructed as above at (2.2) "cat", ('cat-LEFT')

3. Let 0_n be the zero-necklace of circumference $n > 0$, and let $y = \langle y_j \rangle_{j\in\{m\ldots 1\}}$ be a Necklace of circumference m, and let $K = n + m$.
(d) By the **padout** $0_n\,\mathtt{pad}\,y$ *(or* $\mathtt{pad}(0_n, y)$*)* we mean the necklace $c = \langle c_k \rangle_{k\in\{K\ldots 1\}}$ of circumference $K{=}n{+}m$ constructed as above at (2.2) "pad", ('pad-LEFT')

4. Let $x = \langle x_i \rangle_{i\in\{n\ldots 1\}}$ be a Necklace of circumference $n \geqslant 2$, and choose integer K such that $1 \leqslant K \leqslant n$.
(e) By a **snap** $\mathtt{SNAP}(x, K)$ of x at K, we mean the pair of necklaces $\{L, R\}$ of respective circumferences $(n{-}K)$ and K, constructed as above at (2.2) "snap".

REMARK 2.4:
Note that a 'padout' of a necklace is still a 'necklace' since the (least) Decimal Value of any such padout remains invariant. Note also that the LeftSnap and the RightSnap of a necklace are also still 'necklaces', by construction.

EXAMPLES 2.5:

1. Let $\mathtt{x} = \langle 0,0,1,1 \rangle$, $\mathtt{Dec} = 3$; $\mathtt{y} = \langle 0,1,0,1 \rangle$, $\mathtt{Dec} = 5$;
   then $\mathtt{x} \boxplus \mathtt{y} = \langle 0,0,1,1 \rangle$, $\mathtt{Dec} = 3$; and
   $\mathtt{x} \boxtimes \mathtt{y} = \langle 0,0,0,1 \rangle$, $\mathtt{Dec} = 1$.

2. Let $\mathtt{x} = \langle 1,1,0 \rangle$, $\mathtt{Circ} = 3$, $\mathtt{Dec} = 6$;
   $\mathtt{y} = \langle 0,1,0,1 \rangle$, $\mathtt{Circ} = 4$, $\mathtt{Dec} = 5$;
   then (bitnecklaces) $[1,1,0]\ \mathtt{cat}\ [0,1,0,1] = [1,1,0,0,1,0,1]$,
   $\mathtt{Dec} = 101$;
   whence (Necklaces) $\mathtt{x}\ \mathtt{cat}\ \mathtt{y} = \langle 1,1,0 \rangle\ \mathtt{cat}\ \langle 0,1,0,1 \rangle =$
   $\langle 0,0,1,0,1,1,1 \rangle = \mathtt{c}$,
   $\mathtt{Circ} = 4{+}3 = 7$, $\mathtt{Dec} = 23$, *(least decimal value)*.

3. Let $0_4 = \langle 0,0,0,0 \rangle$, $\mathtt{Circ} = 4$, $\mathtt{Dec} = 0$,
   and $\mathtt{y} = \langle 0,1,1 \rangle$, $\mathtt{Circ} = 3$, $\mathtt{Dec} = 6$;
   then (bitnecklaces)
   $[0,0,0,0]\ \mathtt{pad}\ [0,1,1] = [0,0,0,0,0,1,1]$, $\mathtt{Dec} = 3$;
   whence (Necklaces) $0_4\ \mathtt{pad}\ \mathtt{y} = \langle 0,0,0,0 \rangle\ \mathtt{pad}\ \langle 0,1,1 \rangle =$
   $\langle 0,0,0,0,0,1,1 \rangle = \mathtt{c}$
   $\mathtt{Circ} = 4{+}3 = 7$, $\mathtt{Dec} = 3 = \mathtt{Dec}(\mathtt{y})$, *(unaltered)*.

4. Let $\mathtt{x} = \langle 0,1,0,0,1,0,1 \rangle$, $\mathtt{Circ} = 7$, $\mathtt{Dec} = 37$, $\mathtt{Sif} = 6$;
   take $\mathtt{K} = 3$, $(7 > 3 > 1)$,
   then (bitnecklaces)
   $\mathtt{SNAP}\big([0,1,0,0,1,0,1],\ 3\big) = \big\{\ [0,1,0,0],\ [1,0,1]\ \big\}$
   whence (Necklaces) $\mathtt{SNAP}\big(\mathtt{x},\ 3\big) = \{\mathtt{L}, \mathtt{R}\}$ where
   $\mathtt{L} = \langle 0,0,0,1 \rangle$, $\mathtt{Circ} = 4 = (\mathtt{K}{-}3)$, $\mathtt{Dec} = 1$, $\mathtt{Sif} = 1$,
   $\mathtt{R} = \langle 0,1,1 \rangle$, $\mathtt{Circ} = 3 = \mathtt{K}$, $\mathtt{Dec} = 3$, $\mathtt{Dec} = 2$.

FACTS 2.6:
Let $\mathtt{x}$, $\mathtt{y}$ be necklaces, $\mathtt{0}$ and $\mathtt{U}$ the Zero and One necklaces, all of same circumference n (and let $'$ denote a dual necklace). Then :

(a) $x \boxplus y = y \boxplus x$; (b) $x \boxtimes y = y \boxtimes x$;
(c) $x \boxplus x = 0$; (d) $x \boxtimes x = x$;
(e) $x \boxplus 0 = x = 0 \boxplus x$; (f) $x \boxtimes 0 = 0 = 0 \boxtimes x$;
(g) $x \boxplus U = x' = U \boxplus x$; (h) $x \boxtimes U = x = U \boxtimes x$;

(i) $x \boxplus x' \neq U$ (in general); (j) $x \boxtimes x' \neq 0$ (in general);

(u) $0' = U = U''$; $U' = 0 = 0''$;
(v) $U \boxplus 0 = U = 0 \boxplus U$; (w) $U \boxtimes 0 = 0 = 0 \boxtimes U$;

Proof. Under the method employed to construct the sum and product of two Necklaces from their associated equivalence classes of bitnecklaces, each of the (a)...(h) identities hold for the corresponding sum and product operations on the bitnecklaces, and extracting the least decimal-valued bitnecklace as Necklace representative does not disturb those identities.
The non-identities (i),(j) are confirmed by counter-examples (see the following Remarks).
The (u),(v),(w) facts are special cases involving the zero- and the unity-Necklaces. □

Remarks 2.7:
(1) Necklace addition $\boxplus$ and multiplication $\boxtimes$ are both commutative (2.6.a,b).
(2) The 0 Necklace is the $\boxplus$–Identity and the $\boxtimes$–Zero (2.6.e,f); and U is the $\boxtimes$–Unit (2.6.h).
(3) Necklace addition $\boxplus$ is of order 2 (2.6.c), whence each x is its own $\boxtimes$–Inverse (its own 'Negative'); and under multiplication $\boxtimes$ each Necklace x is idempotent (2.6.d).
(4) Multiplication by U is the involutory **dualization** operation $\mathtt{Dual}(x) = x'$, (2.6.g); and $\mathtt{Dual}(\mathtt{Dual}(x)) = x$, so that $\mathtt{Dual} \circ \mathtt{Dual} = \mathtt{Id}$ the **Identity** operation $\mathtt{Id}(x) = x$.

The following remarks apply specifically to non-trivial necklaces.

REMARKS 2.8:

(1) Necklace non-identities (i),(j) are confirmed by these examples using information in Appendix **N4** :

(i) $x \boxplus x' \neq \mathtt{U}$; take $n = 4$, and $x \boxplus x' \equiv 1 \boxplus 7 = 3 \neq 15 \equiv \mathtt{U}$;

(j) $x \boxtimes x' \neq \mathtt{0}$; take $n = 4$, and $x \boxplus x' \equiv 1 \boxtimes 7 = 1 \neq \mathtt{0} \equiv \mathtt{0}$.

(2) Necklace multiplication $\boxtimes$ is not distributive over addition $\boxplus$, as is shown by the following simple counter-example (using information in Appendix **N3**).

$\mathtt{Let\ necklaces}\ \ x = 1 : \langle 0,0,1\rangle = z,\ \mathtt{and}\ y = 0 : \langle 0,0,0\rangle,$
$\quad \mathtt{then}\ \ x \boxtimes (y \boxplus z) \equiv 1 \boxtimes (1 \boxplus 3) = 1 \boxtimes 1 = 1,$
$\quad \mathtt{whilst}\ (x \boxtimes y) \boxplus (x \boxtimes z) \equiv (1 \boxtimes 1) \boxplus (1 \boxtimes 3)) = 1 \boxplus 1 = 0.$

(3) Necklace addition $\boxplus$ is not distributive over multip. $\boxtimes$,

$\mathtt{Let\ necklaces}\ \ x = 1 : \langle 0,0,1\rangle = y,\ \mathtt{and}\ z = 3 : \langle 0,1,1\rangle,$
$\quad \mathtt{then}\ \ x \boxplus (y \boxtimes z) \equiv 1 \boxplus (0 \boxtimes 1) = 1 \boxplus 0 = 1,$
$\quad \mathtt{whilst}\ (x \boxtimes y) \boxplus (x \boxtimes z) \equiv (1 \boxplus 1) \boxtimes (1 \boxplus 0)) = 0 \boxtimes 1 = 0.$

(4) Necklace addition $\boxplus$ is <u>not</u> associative ,

$\mathtt{Let\ necklaces}\ \ x = 1 : \langle 0,0,1\rangle = y,\ \mathtt{and}\ z = 3 : \langle 0,1,1\rangle,$
$\quad \mathtt{then}\ \ x \boxplus (y \boxplus z) \equiv 1 \boxplus (1 \boxplus 3) = 1 \boxplus 1 = 0,$
$\quad \mathtt{whilst}\ (x \boxplus y) \boxplus z \equiv (1 \boxplus 1) \boxplus 3 = 0 \boxplus 3 = 3.$

(5) Necklace multiplication $\boxplus$ is <u>not</u> associative ,

$\mathtt{Let\ necklaces}\ \ x = 3 : \langle 0,0,0,0,1,1\rangle : y = 23 : \langle 0,1,0,1,1,1\rangle,\ \mathtt{and}$
$z = 27 : \langle 0,1,1,0,1,1\rangle,$
$\quad \mathtt{then}\ \ x \boxtimes (y \boxtimes z) \equiv 3 \boxtimes (23 \boxplus 27) = 3 \boxtimes 13 = 1,$
$\quad \mathtt{whilst}\ (x \boxtimes y) \boxtimes z \equiv (3 \boxtimes 23) \boxtimes 27 = 3 \boxtimes 27 = 3.$

[Nevertheless, necklace multiplication $\boxtimes$ is associative for necklaces of circumference $\leqslant 5$; the necklaces in the above c/example have circumference 6.]

FACTS 2.9:

Let x, y be necklaces, $\mathtt{0}$ and $\mathtt{U}$ the Zero and One necklaces, all of same circumference n (and let $'$ denote a dual necklace). Then :

(a1) $\quad \|x\| = 0 \Leftrightarrow x = \mathtt{0}$;

(a2) $\quad \|x \boxplus y\| \leqslant \|x\| + \|y\|$ ('Hamming Norm' is a norm);

(a3) $\|\mathtt{x} \boxplus \mathtt{y}\| \leqslant \max(\|\mathtt{x}\|, \|\mathtt{y}\|)$ (non-Archimedean norm);
(a4) $\|\mathtt{x} \boxtimes \mathtt{y}\| \leqslant \min(\|\mathtt{x}\|, \|\mathtt{y}\|)$;
(a5) $\|\mathtt{x} \boxtimes \mathtt{y}\| \leqslant \|\mathtt{x}\|\,\|\mathtt{y}\|$.

(b1) $\mathtt{Dec}(\mathtt{x}) = \mathtt{0} \Leftrightarrow \mathtt{x} = \mathtt{0}$;
(b2) $\mathtt{Dec}(\mathtt{x} \boxplus \mathtt{y}) \leqslant \mathtt{Dec}(\mathtt{x}) + \mathtt{Dec}(\mathtt{y})$ ('Dec' is a norm);
(b3) $\mathtt{Dec}(\mathtt{x} \boxplus \mathtt{y}) \leqslant \max(\mathtt{Dec}(\mathtt{x}), \mathtt{Dec}(\mathtt{y}))$ (non-Archimed.);
(b4) $\mathtt{Dec}(\mathtt{x} \boxtimes \mathtt{y}) \leqslant \min(\mathtt{Dec}(\mathtt{x}), \mathtt{Dec}(\mathtt{y}))$;
(b5) $\mathtt{Dec}(\mathtt{x} \boxtimes \mathtt{y}) \leqslant \mathtt{Dec}(\mathtt{x})\,\mathtt{Dec}(\mathtt{y})$.

(c) $\|\mathtt{x}\| \;?\; \|\mathtt{x}'\|$; $\mathtt{Dec}(\mathtt{x}) \;?\; \mathtt{Dec}(\mathtt{x}')$.

Proof.
(a1) and (b1) are immediate from the definitions; (a2) follows immediately from (a3); and since the addition of two necklaces involves the addition ('XOR-ing') of two associated bitnecklaces, the number of 1-bits in the sum can at best remain equal to the maximum number of 1-bits in each of the summands or else be reduced below that maximum — which is what is asserted in (a3).
(a4) Since the multiplication of two necklaces involves the multiplication ('AND-ing') of two associated bitnecklaces, the number of 1-bits in the product can at best remain equal to the minimum number of 1-bits in each of the factors or else be raised above that minimum — which is what is asserted in (a4).
(a5) follows immediately from (a4).
(b2) follows immediately from (b3); and (b3) follows from a similar ('XOR-ing') argument used for (a3).
(b4) follows from a similar ('AND-ing') argument used for (a4).
(b5) follows immediately from (b4).
(c) Simple counter-examples (*cf.* Appendix **N**) show that there is no usefully consistent relationship between $\|\mathtt{x}\|, \|\mathtt{x}'\|$; and similarly for $\mathtt{Dec}(\mathtt{x}), \mathtt{Dec}(\mathtt{x}')$. □

Remarks 2.10:
(1) From the additive relations (a3) and (b3) we see that each set $\mathtt{NK}^{[n]}$ of necklaces of circumference n equipped with resp. the '$\|.\|$ Norm', or the 'Dec Norm' possesses the (discrete) non-Archimedean

Normed Space condition.
(2) From the multiplicative relations (a5) and (b5) we see that in each set $\mathtt{NK}^{[n]}$ of necklaces of circumference n equipped with resp. the '$\|.\|$ Norm', possesses the (discrete) 'Normed Algebra' multiplicative condition.

3 ORDERING NECKLACES

DEFINITION 3.1:
Let $x = \langle x_i \rangle_{i\in\{n\ldots1\}}$ and $y = \langle y_i \rangle_{i\in\{n\ldots1\}}$ be two Necklaces of the same circumference n; and let $\mathtt{Dec}(x)$, $\mathtt{Dec}(y)$ be their decimal values.
(a) We say that x *precedes* y, and write
$$x \prec y \quad \Leftrightarrow \quad \mathtt{Dec}(x) < \mathtt{Dec}(y)$$
(b) Should equality be also possible we write
$$x \preceq y \quad \Leftrightarrow \quad \mathtt{Dec}(x) \leqslant \mathtt{Dec}(y)$$
(c) By the *minimum* of x and y, denoted by $\min(x, y)$, we mean
x if $x \preceq y$, or y if $y \preceq x$.
(d) By the *maximum* of x and y, denoted by $\max(x, y)$, we mean
y if $x \preceq y$, or x if $y \preceq x$.

FACTS 3.2:
Let x, y, z be necklaces, $\mathtt{0}$ and $\mathtt{U}$ the Zero and One necklaces, all of same
circumference n; and let $\mathtt{NK}^{[n]}$ be the set of all necklaces of circumference n. Then :

(a1) $x \preceq x$;
(a2) if $x \preceq x$ and $y \preceq x$, then $x = y$;
(a3) if $x \preceq y \preceq z$, then $x \preceq z$;
(a4) $x \prec y \quad \Leftrightarrow \quad x \preceq y$ and $x \neq y$;
(a5) $\preceq$ is a total order on $\mathtt{NK}^{[n]}$,
and hence $(\mathtt{NK}^{[n]}, \preceq)$ is a chain ;

(b) $\mathtt{0} \preceq x \preceq \mathtt{U}$;
(c) $x \boxplus y \preceq \max(x, y)$;
(d) $x \boxtimes y \preceq \max(x, y)$;

Proof.
(a1)–(a5) immediate, from the definitions and the total order on the Integers.
(b) is immediate : $\mathtt{Dec}(\mathtt{0}) = 0 \leqslant \mathtt{Dec}(\mathtt{x}) \leqslant n = \mathtt{Dec}(\mathtt{U})$.
(c),(d) follow immediatley from Facts (2.9.e,f) . □

Because necklace addition $\boxplus$ and multiplication $\boxtimes$ are not generally associative, nor distributive, any notion of using these as internal laws of composition on sets of necklaces of a given circumference to form analogues of additive groups, multiplicative groups, or rings *etc.* is not possible. Despite this we can still form structures which have some resemblance to groups, rings, *etc.*. To avoid any possibility of misunderstanding, I shall use alternative names for these novel kinds of structures. Thus we have "NPacks, NBoxes, NIdeals" rather than Groups, Rings, Ideals, and so on.

4 NPACKS AND NBOXES OF NECKLACES

DEFINITION 4.1:
(a1) By an **NPack** $\mathtt{NP}(n)$ *of circumference* n we means the set of all Necklaces of fixed circumference n $(n = 0, 1, 2, \ldots)$ equipped with an internal law of composition which is either

• *commutative Necklace addition* $\boxplus$ — in which case we refer to it as an 'additive NPack', and may also denote it by $\mathtt{NP}(\boxplus, n)$, or by $\mathtt{a_NP}(n)$; or
• *commutative Necklace multiplication* $\boxtimes$ — in which case we refer to it as a 'multiplicative NPack', and may also denote it by $\mathtt{NP}(\boxtimes, n)$, or by $\mathtt{m_NP}(n)$.
An NPack $\mathtt{NP}(n)$ is said to be 'trivial' in case $n = 0$ or $n = 1$, and 'non-trivial' otherwise.

(a2) By a **sub-NPack** we mean a subset of an NPack which is closed w.r.t. its internal law of composition ($\boxplus$ or $\boxtimes$).
A sub-NPack is 'proper' `iff` it is a proper subset of the NPack itself.

(b1) Given an additive sub-NPack H of an NPack $NP(n)$,
by an **H-Co-NPack** we mean a subset gH of necklaces in $NP(n)$ where $gH = \{ g \boxplus h \mid h \in H \}$ and Necklace $g \in NP(n)$.

(b2) By the **Quotient-NPack** $NP(n)/H$ of an NPack $NP(n)$ w.r.t. an additive sub-NPack $H \subseteq NP(n)$ we mean the set of H-Co-NPacks in $NP(n)$.

(c1) By an **NBox** $NB(\boxplus, \boxtimes, n)$ *of circumference* n — or $NB(n)$ for brevity — we means the set of all Necklaces of fixed circumference n $(n = 0, 1, 2, \ldots)$ equipped with two internal laws of composition, *viz.* commutative Necklace addition $\boxplus$ and commutative Necklace multiplication $\boxtimes$.
A Necklace Box $NB(n)$ is said to be 'trivial' in case $n = 0$ or $n = 1$, and 'non-trivial' otherwise.

(c2) By a **sub-NBox** we mean a subset of an NBox which is closed with respect to both its internal laws of composition ($\boxplus$ and $\boxtimes$).
A sub-NBox is 'proper' `iff` it is a proper subset of the NBox itself.

(d) By an **NIdeal** $NI(n)$ we mean a proper subset of an NBox $NB(\boxplus, \boxtimes, n)$ which is itself an additive NPack $NP(\boxplus, n)$ (an additive sub-NPack) and such that for each necklace $x \in NI(n)$ and each necklace $y \in NB(\boxplus, \boxtimes, n)$ we have their product $x \boxtimes y \in NI(n)$.
Being proper, no NIdeal can contain the $\boxtimes$–Identitiy Necklace U.

REMARKS 4.2:
(1) Each trivial NPack $NP(\boxplus, 0)$, $NP(\boxtimes, 0)$, $NP(\boxtimes, 1)$, $NP(\boxtimes, 1)$ is (trivially) an algebraic Group.
No non-trivial NPack $NP(\boxplus, n)$, $NP(\boxtimes, n)$ is an algebraic Group.
(2) Each trivial NBox $NB(\boxplus, \boxtimes, 0)$, $NB(\boxplus, \boxtimes, 1)$ is (trivially) an algebraic Ring.
No non-trivial NBox $NB(\boxplus, \boxtimes, n)$ is an algebraic Ring.

These facts are immediately obvious from the absence of the $\boxplus, \boxtimes$ distributive laws, and the $\boxplus$ associative law, in non-trivial necklace systems, as remarked in (1.5) above.
(3) Each NPack $\mathsf{NP}(\boxplus, n)$ when equipped with either of the non-Archimedean norms $\|.\|$, $\mathtt{Dec}(.)$, becomes the 'Necklace Analogue' of a non-Archimedean Normed Space.
(4) Each NBox $\mathsf{NB}(\boxplus, \boxtimes, n)$ when equipped with either of the non-Archimedean norms $\|.\|$, $\mathtt{Dec}(.)$ becomes the 'Necklace Analogue' of a non-Archimedean Normed Algebra.
Both (3) and (4) follow from the relevant facts in (2.9).
(5) Suppose $\mathsf{H} = \{h_1, \ldots, h_m\}$, $(0 < m < n)$ be an additive sub-NPack of an NPack $\mathsf{NP}(n) = \{g_1, \ldots, g_n\}$. It need not happen that m divides n : e. g. (see Appendix **N6**), the NPack $\mathsf{S} = \{0, 1, 3, 5, 7, 15\}$ has 6 necklaces and the additive sub-NPack $(17) = \{0, 1, 3, 5\}$ has 4 – which does not divide 6; whilst the additive sub-NPack $\mathsf{H} = (7) = \{0, 1, 3\}$ has 3 and $6 \div 3 = 2$.
Next form the 6 subsets $\mathsf{gH} = \{\, \mathsf{g} \boxplus \mathsf{h} \mid \mathsf{h} \in \mathrm{H} \,\}$, where the Necklaces $\mathsf{g} \in \mathsf{NP}(n)$:

g	gH	?
0	{0, 1, 3}	not disjoint from H
1	{1, 0, 1}	*do.*
3	{1, 1, 0}	*do.*
5	{5, 1, 3}	*do.*
7	{7, 3, 1}	*do.*
15	{15, 7, 3}	is disjoint from H, and is a $\boxplus$Sub-NPack.

Hence, 15H is the (only) H-Co-NPack for the additive sub-NPack H={0, 1, 3}. Moreover, the set $\{$ H={0, 1, 3} , K={15, 7, 3} $\}$ is the the Quotient-NPack S / H of the $\boxplus$NPack S = {0, 1, 3, 5, 7, 15} w.r.t. H.

And, from the following addition table, where $\boxplus'$ denotes componentwise $\boxplus$-addition,

H={0,1,3}	$\boxplus'$	H={0,1,3}	=	H={0,1,3}
H={0,1,3}	$\boxplus'$	K={3,7,15}	=	{3,3,3}
H={3,7,15}	$\boxplus'$	K={0,1,3 }	=	{5,3,3}
K={5,7,15}	$\boxplus'$	K={5,7,15}	=	H={0,1,3}

we see that the Quotient-NPack S/H, being not closed under $\boxplus'$-addition, cannot be an additive NPack.
Which shows that necklaces systems do not have the notion, analogous to the group theory notion of a *coset*.
(6) We also see from the (general) possibilty (illustrated in the previous remark (4)) of the size of an $\boxplus$Sub-NPack failing to divide the size of the parent $\boxplus$NPack, that this means there is nothing comparable to "Lagrange's Theorem" (group theory) for additive sub-NPacks in Necklace Systems.

* * * * * *

Despite the facts that Necklace addition and multiplication are neither associative nor distributive, the removal of brackets does not affect compound orderings, as we now demonstrate:

FACTS 4.3:
Let x, y, z be necklaces in an NBox $\mathsf{NB}(\boxplus, \boxtimes, n)$. Then :

(a1) $x \boxplus (y \boxplus z) \preceq \max(x, y, z)$.
(a2) $(x \boxplus y) \boxplus z \preceq \max(x, y, z)$.
(b1) $x \boxtimes (y \boxtimes z) \preceq \max(x, y, z)$.
(b2) $(x \boxtimes y) \boxtimes z \preceq \max(x, y, z)$.

(c11) $x \boxtimes (y \boxplus z) \preceq \max(x, y, z)$.
(c12) $(x \boxtimes y) \boxplus (x \boxtimes z) \preceq \max(x, y, z)$.
(c21) $x \boxtimes (y \boxplus z) \preceq \max(x, y, z)$.
(c22) $(x \boxplus y) \boxtimes (x \boxplus z) \preceq \max(x, y, z)$.

(d11) $(x \boxtimes y) \boxplus z \preceq \max(x, y, z)$.
(d12) $(x \boxtimes z) \boxplus (y \boxtimes z) \preceq \max(x, y, z)$.

(d21) $(x \boxplus y) \boxtimes z \preceq \max(x, y, z)$.
(d22) $(x \boxplus z) \boxtimes (y \boxplus z) \preceq \max(x, y, z)$.

Proof. Immediate, from repeated use of the ordering facts (3.2.c,d), e. g. $x \boxtimes (y \boxplus z) \preceq \max(x, (y \boxplus z)) \preceq \max(x, \max(y, z)) \preceq \max(x, y, z)$;

and $(x \boxplus y) \boxtimes z \preceq \max((x \boxplus y), z) \preceq \max(\max(x, y), z) \preceq \max(x, y, z)$. □

Definition 4.4:
Let $\mathbf{S} = \{x_1, \ldots, x_H\}$ be a set of $H > 0$ necklaces x_h, $h = 1 \ldots H$ all of the same circumference n.
(a) By the **a-Hull** of the set $\mathbf{S}$, denoted by $\mathsf{NH}(\boxplus, \mathbf{S}, n)$ we mean the smallest additive NPack $\mathsf{NP}(\boxplus, n)$ which contains each necklace $x_h \in \mathbf{S}$.
(b) By the **m-Hull** of the set $\mathbf{S}$, denoted by $\mathsf{NH}(\boxtimes, \mathbf{S}, n)$ we mean the smallest multiplicative NPack $\mathsf{NP}(\boxtimes, n)$ which contains each necklace $x_h \in \mathbf{S}$.
(c) By the **Hull** of the set $\mathbf{S}$, denoted by $\mathsf{NH}(\boxplus, \boxtimes, \mathbf{S}, n)$ we mean the smallest NBox $\mathsf{NB}(\boxplus, \boxtimes, n)$ which contains each necklace $x_h \in \mathbf{S}$.

Examples 4.5: Using sample necklaces shown in Appendix **N5** (in Decimal Value notation), let $x_1 \equiv 3$, $x_2 \equiv 7$ and $\mathbf{S} = \{3, 7\}$; then, by inspection, we see that :
1. the a-Hull of $\mathbf{S}$ is $\mathsf{NH}(\boxplus, \mathbf{S}, n) = \{0, 1, 3, 5, 7\}$;
2. the m-Hull of $\mathbf{S}$ is $\mathsf{NH}(\boxtimes, \mathbf{S}, n) = \{0, 1, 3, 5, 7\}$;
3. the Hull of $\mathbf{S}$ is $\mathsf{NH}(\boxplus, \boxtimes, \mathbf{S}, n) = \{0, 1, 3, 5, 7\}$.

5 NECKLACE SYSTEMS

Definition 5.1:
By a **Necklace System** we mean a set $\mathbf{S} = \{x_1, \ldots, x_H\}$ of H necklaces x_h, $h = 1 \ldots H$ $(H > 0)$ all of the same circumference n, equipped

with at least one internal law of composition.
By an **Ordered Necklace System** we mean a Necklace System equipped with an ordering compatible with the system's internal law(s) of composition.

EXAMPLES 5.2:
Each of the following is a Necklace System; and when equipped with the ordering inherited from the ordering $\preceq$ (Definition 3.1) is an Ordered Necklace System :

1. Each additive NPack $\mathsf{NP}(\boxplus, n)$,
2. Each multiplicative NPack $\mathsf{NP}(\boxtimes, n)$,
3. Each NBox $\mathsf{NB}(\boxplus, \boxtimes, n)$,
4. Each NIdeal $\mathsf{NI}(n)$;
5. All sub-NPacks and sub-NBoxes.

5.1 MORPHISMS OF NECKLACE SYSTEMS

DEFINITION 5.3:
Let $\mathbf{S} = \{s_1, \ldots, s_H\}$ $(H > 0)$ and $\mathbf{T} = \{t_1, \ldots, t_K\}$ $(K > 0)$ be two Necklace Systems, each equipped with one or other or both of its laws $\boxplus$, $\boxtimes$, and ordering $\preceq$.
By a **morphism** between $\mathbf{S}$ and $\mathbf{T}$ we mean a mapping $f : \mathbf{S} \to \mathbf{T}$ which preserves the law(s) and ordering in these necklace systems.

$$f(s_i \boxplus_{\mathbf{S}} s_j) = f(s_i) \boxplus_{\mathbf{T}} f(s_j);$$
$$f(s_i \boxtimes_{\mathbf{S}} s_j) = f(s_i) \boxtimes_{\mathbf{T}} f(s_j);$$
$$f(s_i \preceq_{\mathbf{S}} s_j) = f(s_i) \preceq_{\mathbf{T}} f(s_j).$$

Such a morphism is an **endomorphism** if the mapping f is not surjective $f(\mathbf{S}) \subsetneq \mathbf{T}$, $(H < K)$, and is an **isomorphism** if the mapping f has an inverse $f^{-1} : \mathbf{T} \to \mathbf{S}$ which is also a morphism (so that f is one-to-one and onto, an $H = K$).

6 INJECTIVE AND PROJECTIVE HULLS OF NECKLACES

6.1 INJECTIVE HULLS

DEFINITION 6.1:
Let $\mathbf{S} = \{x_1, \ldots, x_H\}$ be a set of H $(H > 0)$ necklaces x_h, $h = 1 \ldots H$ all of the same circumference n.
Let $\mathtt{PAD}(\mathbf{S}, H, K)$ be the set $\{y_1, \ldots, y_H\}$ constructed from the necklace set $\mathbf{S}$ by replacing each necklace $x_h \in \mathbf{S}$ by its padout necklace $y_h = 0_K \ \mathtt{pad}\ x_h$ each of circumference $L = K{+}H$, $K > 0$.
(a) By the **a-InjectHull** of a subset $\mathbf{S}$ of an Npack $\mathtt{NP}(\boxplus, H)$ into an NPack $\mathtt{NP}(\boxplus, L)$, where $L = K{+}H > H$, we mean the a-Hull $\mathtt{NH}\big(\boxplus, \mathtt{PAD}(\mathbf{S}, H, K),\ L\big)$ in the NPack $\mathtt{NP}(\boxplus, L)$.
(b) By the **m-InjectHull** of a subset $\mathbf{S}$ of an Npack $\mathtt{NP}(\boxtimes, H)$ into an NPack $\mathtt{NP}(\boxtimes, L)$, where $L = K{+}H > H$, we mean the m-Hull $\mathtt{NH}\big(\boxtimes, \mathtt{PAD}(\mathbf{S}, H, K),\ L\big)$ in the NPack $\mathtt{NP}(\boxtimes, L)$.
(c) By the **InjectHull** of a subset $\mathbf{S}$ of an NBox $\mathtt{NB}(\boxplus, \boxtimes, H)$ into an NBox $\mathtt{NB}(\boxplus, \boxtimes, L)$, where $L = K{+}H > H$, we mean the Hull $\mathtt{NH}\big(\boxplus, \boxtimes, \mathtt{PAD}(\mathbf{S}, H, K),\ L\big)$ in the NBox $\mathtt{NB}(\boxplus, \boxtimes, L)$.

FACTS 6.2:
For each pair of circumferences n, m; $0 \leqslant n < m$, we have :

(a1) each NPack $\mathtt{NP}(\boxplus, n)$ is endomorphically identifiable with a sub-NPack of the NPack $\mathtt{NP}(\boxplus, m)$ (proper, if $2 < n$).
(a2) each NPack $\mathtt{NP}(\boxtimes, n)$ is endomorphically identifiable with a sub-NPack of the NPack $\mathtt{NP}(\boxtimes, m)$ (proper, if $2 < n$).
(a3) each NBox $\mathtt{NB}(\boxplus, \boxtimes, n)$ is endomorphically identifiable with a sub-NBox of the NBox $\mathtt{NB}(\boxplus, \boxtimes, m)$ (proper, if $2 < n$).

(b1) The set $\{\,\mathsf{NP}(\boxplus, n)\,|\,n = 0, 1, 2, 3, \ldots\}$ of all additive NPacks of necklaces of every circumference is a *'Chain'* with respect to the a-InjectHull identifications $\mathsf{NP}(\boxplus, n) \rightarrowtail \mathsf{NP}(\boxplus, n{+}1)$ with $\mathsf{NP}(\boxplus, n) \rightarrowtail \mathsf{NH}\Big(\boxplus, \mathsf{PAD}(\mathsf{NP}(\boxplus, n), n, 1),\ n{+}1\Big)$.

(b2) The set $\{\,\mathsf{NP}(\boxtimes, n)\,|\,n = 0, 1, 2, 3, \ldots\}$ of all multiplicative NPacks of necklaces of every circumference is a *'Chain'* with respect to the m-InjectHull identifications $\mathsf{NP}(\boxtimes, n) \rightarrowtail \mathsf{NP}(\boxtimes, n{+}1)$ with $\mathsf{NP}(\boxtimes, n) \rightarrowtail \mathsf{NH}\Big(\boxtimes, \mathsf{PAD}(\mathsf{NP}(\boxtimes, n), n, 1),\ n{+}1\Big)$.

(b3) The set $\{\,\mathsf{NB}(\boxplus, \boxtimes, n)\,|\,n = 0, 1, 2, 3, \ldots\}$ of all NBoxes of necklaces of every circumference is a *'Chain'* w.r.t. the InjectHull identifications $\mathsf{NB}(\boxplus, \boxtimes, n) \rightarrowtail \mathsf{NB}(\boxplus, \boxtimes, n{+}1)$, with $\mathsf{NB}(\boxplus, \boxtimes, n) \rightarrowtail \mathsf{NH}\Big(\boxplus, \boxtimes, \mathsf{PAD}(\mathsf{NB}(\boxplus, \boxtimes, n), n, 1),\ n{+}1\Big)$.

Proof.
The details of proof are straight-forward but tedious, and can perhaps be best illustrated by reference to the following Examples (6.3). □

Examples 6.3: Using sample necklaces shown in Appendix **N5** (in Decimal Value notation), let $x_1 \equiv 3$, $x_2 \equiv 7$, and let

$\mathbf{S} = \{x_1, x_2\} \equiv \{3, 7\}$ be the set of these $H = 2$ necklaces. Next, padout-by-1-bit each of them so that

$y_1 = 0_1\,\mathsf{pad}\,x_1$, $\mathtt{Circ} = 6, \mathtt{Dec} = 3$, and

$y_2 = 0_1\,\mathsf{pad}\,x_2$, $\mathtt{Circ} = 6, \mathtt{Dec} = 7$,

and write $\mathbf{T} = \{y_1, y_2\} \equiv \{3, 7\}$. Thus $\mathbf{S} \subset \mathsf{NP}(\boxplus, n)$ and $\mathbf{T} \subset \mathsf{NP}(\boxplus, n{+}1)$.

Then, by inspection of Tables in Appendices **N 5**, **N 6**, we see that

1. the a-Hull of **S** in $\mathsf{NP}(\boxplus, n)$ is $\mathsf{NH}(\boxplus, \mathbf{S},\ n) = \{0, 1, 3, 5, 7\}$, a set of necklaces with circumferences n in $\mathsf{NP}(\boxplus, n)$,

2. the a-InjectHull of $\mathbf{S}$ in $\mathsf{NP}(\boxplus, n{+}1)$, is the a-Hull
$\mathsf{NH}(\boxplus, \mathbf{T}, n{+}1)$ of $\mathbf{T}$, $\{0, 1, 3, 5, 7\}$,
a set of necklaces with circumferences $n{+}1$ in $\mathsf{NP}(\boxplus, n{+}1)$.

Whence we can make the endomorphic identification
$\mathbf{S} \rightarrowtail \mathbf{T}$ of $\mathbf{S} \subset \mathsf{NP}(\boxplus, n)$ with $\mathbf{T} \subset \mathsf{NP}(\boxplus, n{+}1)$.

This example can be similarly extended, to many other illustrative examples involving additive and multiplicative NPacks, and also NBoxes as well.

6.2 PROJECTIVE HULLS

Recall (Def.(2.3) that we can 'snap' a single necklace $x = \langle x_i \rangle_{i \in \{n \ldots 1\}}$, of circumference $n \geqslant 2$, into a pair of necklaces L and R of respective circumferences $\lambda = (n{-}K)$ and $\rho = K$, where the 'snap-point' is K, where $0 \leqslant K \leqslant n$.
We shall refer to these two snapped necklaces as the *LeftSnap* and *RightSnap*, respectively, of x.

We now consider the set of all the LeftSnaps, respectively, RightSnaps, of the necklaces in a set of necklaces; and how we may form their respective Hulls.

DEFINITION 6.4:
Let $\mathbf{S} = \{x_1, \ldots, x_H\}$ be a set of H, $(H > 0)$ necklaces x_h,
$h = 1 \ldots H$, all of the same circumference n.
Let $\mathsf{SNAPS}(\mathbf{S}, \lambda, \rho)$ be the partition $\mathbf{S} = \mathbf{S}_\lambda \sqcup \mathbf{S}_\rho$ of the set $\mathbf{S}$ into its two disjoint subsets $\mathbf{S}_\lambda$ and $\mathbf{S}_\rho$, which consist respectively of all the LeftSnaps and the RightSnaps of the necklaces in $\mathbf{S}$.
(a-L) By the **a-LeftProjectiveHull** of a subset $\mathbf{S}$ of an Npack $\mathsf{NP}(\boxplus, H)$ we mean the a-Hull $\mathsf{NH}\big(\boxplus, \mathbf{S}_\lambda, \lambda\big)$ in the NPack $\mathsf{NP}(\boxplus, \lambda)$.
(a-R) By the **a-RightProjectiveHull** of a subset $\mathbf{S}$ of an Npack $\mathsf{NP}(\boxplus, H)$ we mean the a-Hull $\mathsf{NH}\big(\boxplus, \mathbf{S}_\rho, \rho\big)$ in the NPack $\mathsf{NP}(\boxplus, \rho)$.
(b-L) By the **m-LeftProjectiveHull** of a subset $\mathbf{S}$ of an Npack $\mathsf{NP}(\boxtimes, H)$

we mean the a-Hull $\mathsf{NH}\left(\boxtimes, \mathbf{S}_\lambda, \lambda\right)$ in the NPack $\mathsf{NP}(\boxtimes, \lambda)$.
(b-R) By the **m-RightProjectiveHull** of a subset **S** of an Npack $\mathsf{NP}(\boxtimes, H)$ we mean the a-Hull $\mathsf{NH}\left(\boxtimes, \mathbf{S}_\rho, \rho\right)$ in the NPack $\mathsf{NP}(\boxtimes, \rho)$.

(c-L) By the **LeftProjectiveHull** of a subset **S** of an NBox $\mathsf{NB}(\boxplus, \boxtimes, H)$ we mean the Hull $\mathsf{NH}\left(\boxplus, \boxtimes, \mathbf{S}_\lambda, \lambda\right)$ in the NBox $\mathsf{NB}(\boxplus, \boxtimes, \lambda)$.
(c-R) By the **RightProjectiveHull** of a subset **S** of an NBox $\mathsf{NB}(\boxplus, \boxtimes, H)$ we mean the Hull $\mathsf{NH}\left(\boxplus, \boxtimes, \mathbf{S}_\rho, \rho\right)$ in the NBox $\mathsf{NB}(\boxplus, \boxtimes, \rho)$.

Examples 6.5: Using sample necklaces shown in Appx.**N6**
let $x_1 = \langle 0,0,0,1,1,1\rangle \equiv 7$, $x_2 = \langle 0,0,1,1,0,1\rangle \equiv 13$,
$x_3 = \langle 0,1,0,1,1,1\rangle \equiv 23$, and let $\mathbf{S} = \{x_1, x_2, x_3\} \equiv \{7, 13, 23\}$ be the set of these $H = 3$ necklaces.

(Ex.1) Choose an arbitrary snap-point K=ρ=4 and form the Left-Snaps and RightSnaps of each of these necklaces in **S**.
Recall that this implies that each snapped part of a necklace becomes converted into a 'necklace' (with least Decimal Value), by construction (see 2.2). Thus we have "snapped" **S** into its disjoint union $\mathbf{S}_{(\lambda=2)} \sqcup \mathbf{S}_{(\rho=4)}$, with

$$\mathbf{S}_{(\lambda=2)} = \{v_1, v_2, v_3\} = \{\langle 0,0\rangle, \langle 0,0\rangle, \langle 0,1\rangle\} \equiv \{0,0,1\}$$
$$\mathbf{S}_{(\rho=4)} = \{w_1, w_2, w_3\} = \{\langle 0,1,1,1\rangle, \langle 0,1,0,1\rangle, \langle 0,1,1,1\rangle\} \equiv \{7,5,7\}$$

Using the addition table in **N6** we see that the additive left projective Hull $\mathsf{NH}\left(\boxplus, \mathbf{S}_\lambda, \lambda\right)$ of $\mathbf{S}_{(\lambda=2)}$ is reduced to just $\{\langle 0,0\rangle, \langle 0,1\rangle\} \equiv \{0,1\}$ in the NPack $\mathsf{NP}(\boxplus, \lambda)$, whilst the additive right projective Hull $\mathsf{NH}\left(\boxplus, \mathbf{S}_\rho, \rho\right)$ of $\mathbf{S}_{(\rho=4)}$ is reduced to $\{\langle 0,0,0,0\rangle, \langle 0,0,0,1\rangle, \langle 0,1,1,1\rangle\} \equiv \{0,1,7\}$ in the NPack $\mathsf{NP}(\boxplus, \rho)$.

Using the multiplication table in **N6** we see that the multiplicative left projective Hull $\mathsf{NH}\left(\boxtimes, \mathbf{S}_\lambda, \lambda\right)$ of $\mathbf{S}_{(\lambda=2)}$ is reduced to $\{\langle 0,0\rangle, \langle 0,1\rangle\} \equiv \{0,1\}$ in the NPack $\mathsf{NP}(\boxtimes, \lambda)$, again, whilst the multiplicative right projective Hull $\mathsf{NH}\left(\boxtimes, \mathbf{S}_\rho, \rho\right)$ of $\mathbf{S}_{(\rho=4)}$ is reduced to $\{\langle 0,1,0,1\rangle, \langle 0,1,1,1\rangle\} \equiv \{5,7\}$ in the NPack $\mathsf{NP}(\boxtimes, \rho)$.

Thus, in this instance, the left projective hull $\mathsf{NH}\left(\boxplus, \boxtimes, \mathbf{S}_\lambda, \lambda\right)$ remains the same : $\{\langle 0,0\rangle, \langle 0,1,1\rangle\} \equiv \{0,1\}$, in the NBox $\mathsf{NB}(\boxplus, \boxtimes, \lambda)$. But the right projective hull $\mathsf{NH}\left(\boxplus, \boxtimes, \mathbf{S}_\rho, \rho\right)$ becomes :
$\{\langle 0,0,0,0\rangle, \langle 0,0,0,1\rangle, \langle 0,1,0,1\rangle, \langle 0,1,1,1\rangle\} \equiv \{0,1,5,7\}$,
in the NBox $\mathsf{NB}(\boxplus, \boxtimes, \rho)$

(Ex.2) Now choose a different snap-point, say K=ρ=2, and form the LeftSnaps and RightSnaps of each of those necklaces in **S**.

Thus we have "snapped" **S** into its disjoint union $\mathbf{S}_{(\lambda=4)} \sqcup \mathbf{S}_{(\rho=2)}$, with

$\mathbf{S}_{(\lambda=4)} = \{v_1, v_2, v_3\} = \{\langle 0,0,0,1\rangle, \langle 0,0,0,1\rangle, \langle 0,1,0,1\rangle\}$
$\equiv \{1,1,5\}$
$\mathbf{S}_{(\rho=2)} = \{w_1, w_2, w_3\} = \{\langle 1,1\rangle, \langle 0,1\rangle, \langle 1,1\rangle\} \equiv \{3,1,3\}$

It is left as an Easy Exercise for the Reader, using the addition and multiplication tables in **N6**, to form the all six of their various projective hulls.

oooooooooooooooooooooooooooooooooooooooooooooooooooooooo

...continued next page...

APPENDIX N

Addition and Multiplication tables for necklaces of circumferences 3,4,5,6.

N3 : NECKLACES OF CIRCUMFERENCE 3

Here are the 4 Necklaces of circumference 3, with their Decimal Values (and Hamming Norms, as subscripts) :

$$0 : \langle 0,0,0\rangle_{(0)} \quad 1 : \langle 0,0,1\rangle_{(1)} \quad 3 : \langle 0,1,1\rangle_{(2)} \quad 7 : \langle 1,1,1\rangle_{(3)}$$

and their ($\boxplus$) and ($\boxtimes$) Tables, and the their Dual Necklaces (indexed by those Decimal Values) :

$\boxplus$	0	1	3	7
0	0	1	3	7
1	1	0	1	3
3	3	1	0	1
7	7	3	1	0

$\boxtimes$	0	1	3	7
0	0	0	0	0
1	0	1	1	1
3	0	1	3	3
7	0	1	3	7

x	0	1	3	7
x'	7	3	1	0
$x \boxplus x'$	7	1	1	7
$x \boxtimes x'$	0	1	1	0

STRUCTURE IN THE SET $\mathbf{S} = \{0,1,3,7\}$ OF ALL 4 NECKLACES OF CIRCUMFERENCE 3.

In **S** there are $2^4-1 = 15$ non-empty subsets, $2^{4-1} = 8$ subsets containing the ⊞-Zero 0 , and $2^{4-2} = 4$ containing the ⊠-Unit U but not the ⊞-Zero 0. There are

6 additive ⊞-SubNPacks, 4 multiplicative ⊠-SubNPacks,
2 ⊞—⊠-SubNBoxes, and 2 NIdeals.

N3 STRUCTURE TABLE :: SUBSET 'IDX' FROM 1 TO 15

Idx	SubSet	⊞SubNPk	⊠SubNPk	⊞⊠SubNBx	NIdeal
1	{0}	Y	-	-	-
2	{0,1}	Y	Y	Y	Y
3	{0,3}	Y	-	-	-
4	{0,7}	Y	-	-	-
5	{0,1,3}	Y	Y	Y	Y
6	{0,1,7}	-	-	-	-
7	{0,3,7}	-	-	-	-
8	{0,1,3,7}	Y	-	-	-
9	{1}	-	Y	-	-
10	{3}	-	-	-	-
11	{7}	-	-	-	-
12	{1,3}	-	Y	-	-
13	{1,7}	-	-	-	-
14	{3,7}	-	-	-	-
15	{1,3,7}	-	-	-	-
	Totals	6	4	2	2
(100%)	PerCent	40%	27%	13%	13%

N4 : NECKLACES OF CIRCUMFERENCE 4

Here are the 6 Necklaces of circumference 4, with their Decimal Values (and Hamming Norms, as subscripts) :

$$0 : \langle 0,0,0,0\rangle_{(0)} \qquad 1 : \langle 0,0,0,1\rangle_{(1)} \qquad 3 : \langle 0,0,1,1\rangle_{(2)}$$
$$5 : \langle 0,1,0,1\rangle_{(2)} \qquad 7 : \langle 0,1,1,1\rangle_{(3)} \qquad 15 : \langle 1,1,1,1\rangle_{(4)}$$

and their ($\boxplus$) and ($\boxtimes$) Tables (indexed by those Decimal Values) :

$x \boxplus y$	0	1	3	5	7	15
0	0	1	3	5	7	15
1	1	0	1	1	3	7
3	3	1	0	3	1	3
5	5	1	3	0	1	5
7	7	3	1	1	0	1
15	15	7	3	5	1	0

$\boxtimes$	0	1	3	5	7	15
0	0	0	0	0	0	0
1	0	1	1	1	1	1
3	0	1	3	1	3	3
5	0	1	1	5	5	5
7	0	1	3	5	7	7
15	0	1	3	5	7	15

And the Dual Necklaces[3], *etc.* :

x	0	1	3	5	7	15
x'	15	7	3	5	1	0
$x \boxplus x'$	15	3	0	0	3	15
$x \boxtimes x'$	0	1	3	5	1	0
$x \boxplus u$	15	7	3	5	1	0
$x \boxtimes u$	0	1	3	5	7	15

3 Note that $3 \approx [0011]$ and $5 \approx [0101] \hookrightarrow [1010] = [0101]'$ are each self-dual necklaces in this instance.

STRUCTURE IN THE SET $\mathbf{S} = \{0, 1, 3, 5, 7, 15\}$ OF ALL 6 NECKLACES OF CIRCUMFERENCE 4.

In $\mathbf{S}$ there are $2^6 - 1 = 63$ non-empty subsets, $2^{6-1} = 32$ subsets containing the $\boxplus$-Zero $\mathtt{0}$, and $2^{6-2} = 16$ containing the $\boxtimes$-Unit $\mathtt{U}$ but not the $\boxplus$-Zero $\mathtt{0}$.

The following Table lists the additive and multiplicative structures possessed by these subsets of $\mathbf{S} = \{0, 1, 3, 5, 7, 15\}$.

...continued next page...

N4 STRUCTURE TABLE PART 1 :: SUBSET 'IDX' : 1 TO 32

Idx	SubSet	⊞SubNPk	⊠SubNPk	⊞⊠SubNBx	NIdeal
1	{0}	Y	-	-	-
2	{0,1}	Y	Y	Y	Y
3	{0,3}	Y	-	-	-
4	{0,5}	Y	-	-	-
5	{0,7}	Y	-	-	-
6	{0,15}	Y	-	-	-
7	{0,1,3}	Y	Y	Y	Y
8	{0,1,5}	Y	Y	Y	Y
9	{0,1,7}	-	Y	-	-
10	{0,1,15}	-	Y	-	-
11	{0,3,5}	Y	-	-	-
12	{0,3,7}	-	-	-	-
13	{0,3,15}	Y	-	-	-
14	{0,5,7}	-	-	-	-
15	{0,5,15}	Y	-	-	-
16	{0,7,15}	-	-	-	-
17	{0,1,3,5}	Y	Y	Y	Y
18	{0,1,3,7}	Y	Y	Y	Y
19	{0,1,3,15}	-	Y	-	-
20	{0,1,5,7}	-	Y	-	-
21	{0,1,5,15}	-	Y	-	-
22	{0,1,7,15}	-	Y	-	-
23	{0,3,5,7}	-	-	-	-
24	{0,3,5,15}	Y	-	-	-
25	{0,3,7,15}	-	-	-	-
26	{0,5,7,15}	-	-	-	-
27	{0,1,3,5,7}	Y	Y	Y	Y
28	{0,1,3,5,15}	-	Y	-	-
29	{0,1,3,7,15}	Y	Y	Y	-
30	{0,1,5,7,15}	-	Y	-	-
31	{0,3,5,7,15}	-	-	-	-
32	{0,1,3,5,7,15}	Y	Y	Y	-
	subTotals	17	16	8	6

N4 STRUCTURE TABLE PART 2 :: SUBSET 'IDX' : 33 TO 63

Idx	SubSet	⊞SubNPk	⊠SubNPk	⊞⊠SubNBx	NIdeal
33	{1}	-	Y	-	-
34	{3}	-	-	-	-
35	{5}	-	-	-	-
36	{7}	-	-	-	-
37	{15}	-	-	-	-
38	{1,3}	-	Y	-	-
39	{1,5}	-	Y	-	-
40	{1,7}	-	Y	-	-
41	{1,15}	-	Y	-	-
42	{3,5}	-	-	-	-
43	{3,7}	-	-	-	-
44	{3,15}	-	-	-	-
45	{5,7}	-	-	-	-
46	{5,15}	-	-	-	-
47	{7,15}	-	-	-	-
48	{1,3,5}	-	Y	-	-
49	{1,3,7}	-	Y	-	-
50	{1,3,15}	-	Y	-	-
51	{1,5,7}	-	Y	-	-
52	{1,5,15}	-	Y	-	-
53	{1,7,15}	-	Y	-	-
54	{3,5,7}	-	-	-	-
55	{3,5,15}	-	-	-	-
56	{3,7,15}	-	-	-	-
57	{5,7,15}	-	-	-	-
58	{1,3,5,7}	-	Y	-	-
59	{1,3,5,15}	-	Y	-	-
60	{1,3,7,15}	-	Y	-	-
61	{1,5,7,15}	-	Y	-	-
62	{3,5,7,15}	-	-	-	-
63	{1,3,5,7,15}	-	Y	-	-
	subTotals	0	16	0	0
(63)	TOTALS	17	32	8	6
(100%)	PerCent	27%	51%	13%	10%

SUMMARY :
In the set $\mathbf{S} = \{0,1,3,5,7,15\}$ of all 6 Necklaces of circumference 4 there are 17 additive ⊞-SubNPacks, 32 multiplicative ⊠-SubNPacks, 8 ⊞–⊠-SubNBoxes, and 6 NIdeals.
[The ⊠ unit necklace $\mathsf{U} = 15 = \langle 1,1,1,1\rangle$ cannot belong to a (proper) NIdeal.]

N5 : NECKLACES OF CIRCUMFERENCE 5

Here are the 8 Necklaces of circumference 5, with their Decimal Values (and Hamming Norms, as subscripts) :

$$0 : \langle 0,0,0,0,0\rangle_{(0)} \qquad 1 : \langle 0,0,0,0,1\rangle_{(1)} \qquad 3 : \langle 0,0,0,1,1\rangle_{(2)}$$
$$5 : \langle 0,0,1,0,1\rangle_{(2)} \qquad 7 : \langle 0,0,1,1,1\rangle_{(3)} \qquad 11 : \langle 0,1,0,1,1\rangle_{(3)}$$
$$15 : \langle 0,1,1,1,1\rangle_{(4)} \qquad 31 : \langle 1,1,1,1,1\rangle_{(5)}$$

and their (⊞) and (⊠) Tables (indexed by those Decimal Values) :

⊞	0	1	3	5	7	11	15	31
0	0	1	3	5	7	11	15	31
1	1	0	1	1	3	5	7	15
3	3	1	0	3	1	1	3	7
5	5	1	3	0	1	7	5	11
7	7	3	1	1	0	3	1	3
11	11	5	1	7	3	0	1	5
15	15	7	3	5	1	1	0	1
31	31	15	7	11	3	5	1	0

⊠	0	1	3	5	7	11	15	31
0	0	0	0	0	0	0	0	0
1	0	1	1	1	1	1	1	1
3	0	1	3	1	3	3	3	3
5	0	1	1	5	5	1	5	5
7	0	1	3	5	7	3	7	7
11	0	1	3	1	3	11	11	11
15	0	1	3	5	7	11	15	15
31	0	1	3	5	7	11	15	31

And the Dual Necklaces :

x	0	1	3	5	7	11	15	31
x′	31	15	7	11	3	5	1	0
x⊞x′	31	7	1	7	1	7	7	31
x⊠x′	0	1	3	5	1	5	1	0

STRUCTURE IN THE SET $\mathbf{S} = \{0, 1, 3, 5, 7, 11, 15, 31\}$ OF ALL 8 NECKLACES OF CIRCUMFERENCE 5.

In $\mathbf{S}$ there are $2^8-1 = 255$ non-empty subsets, $2^{8-1} = 128$ subsets containing the ⊞-Zero 0 , and $2^{8-2} = 64$ containing the ⊠-Unit U but not the ⊞-Zero 0. There are :

23 additive ⊞-SubNPacks, 112 multiplicative ⊠-SubNPacks,
12 ⊞⊠-SubNBoxes, and 10 NIdeals.

[The ⊠ unit necklace $\mathsf{U} = 31 = \langle 1, 1, 1, 1, 1 \rangle$ cannot belong to a (proper) NIdeal.]

The 10 NIdeals are :

$$\{0,1\}$$
$$\{0,1,3\} \qquad \{0,1,5\}$$
$$\{0,1,3,5\} \qquad \{0,1,3,7\}$$
$$\{0,1,3,5,7\} \qquad \{0,1,3,7,15\}$$
$$\{0,1,3,5,7,11\} \qquad \{0,1,3,5,7,15\}$$
$$\{0,1,3,5,7,11,15\}$$

N6 : NECKLACES OF CIRCUMFERENCE 6

Here are the 14 Necklaces of circumference 6,
with their Decimal Values (and Hamming Norms, as subscripts) :

$$0 : \langle 0,0,0,0,0,0\rangle_{(0)} \qquad 1 : \langle 0,0,0,0,0,1\rangle_{(1)} \qquad 3 : \langle 0,0,0,0,1,1\rangle_{(2)}$$
$$5 : \langle 0,0,0,1,0,1\rangle_{(2)} \qquad 7 : \langle 0,0,0,1,1,1\rangle_{(3)} \qquad 9 : \langle 0,0,1,0,0,1\rangle_{(2)}$$
$$11 : \langle 0,0,1,0,1,1\rangle_{(3)} \qquad 13 : \langle 0,0,1,1,0,1\rangle_{(3)} \qquad 15 : \langle 0,0,1,1,1,1\rangle_{(4)}$$
$$21 : \langle 0,1,0,1,0,1\rangle_{(3)} \qquad 23 : \langle 0,1,0,1,1,1\rangle_{(4)} \qquad 27 : \langle 0,1,1,0,1,1\rangle_{(4)}$$
$$31 : \langle 0,1,1,1,1,1\rangle_{(5)} \qquad 63 : \langle 1,1,1,1,1,1\rangle_{(6)}$$

and their ($\boxplus$) and ($\boxtimes$) Tables and Dual Necklace Tables (indexed by those Decimal Values) :

...continued next page...

(⊞) and Dual Necklace Tables

⊞	0	1	3	5	7	9	11	13	15	21	23	27	31	63
0	0	1	3	5	7	9	11	13	15	21	23	27	31	63
1	1	0	1	1	3	1	5	3	7	5	11	13	15	31
3	3	1	0	3	1	5	1	7	3	11	5	3	7	15
5	5	1	3	0	1	3	7	1	5	1	9	15	13	23
7	7	3	1	1	0	7	3	5	1	9	1	7	3	7
9	9	1	5	3	7	0	1	1	3	7	15	9	11	27
11	11	5	1	7	3	1	0	3	1	15	7	1	5	13
13	13	3	7	1	5	1	3	0	1	3	13	11	9	11
15	15	7	3	5	1	3	1	1	0	13	3	5	1	3
21	21	5	11	1	9	7	15	3	13	0	1	7	5	21
23	23	11	5	9	1	15	7	13	3	1	0	3	1	5
27	27	13	3	15	7	9	1	11	5	7	3	0	1	9
31	31	15	7	13	3	11	5	9	1	5	1	1	0	1
63	63	31	15	23	7	27	13	11	3	21	5	9	1	0

x	0	1	3	5	7	9	11	13	15	21	23	27	31	63
x′	63	31	15	23	7	27	13	11	3	21	5	9	1	0
x⊞x′	63	15	3	9	0	9	3	3	3	0	9	9	15	63

(Note that $7 \approx [000111]$ and $21 \approx [010101]$
are self-dual necklaces in this instance.)

...continued next page...

(⊠) and Dual Necklace Tables

⊠	0	1	3	5	7	9	11	13	15	21	23	27	31	63
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	0	1	1	1	1	1	1	1	1	1	1	1	1	1
3	0	1	3	1	3	1	3	1	3	1	3	3	3	3
5	0	1	1	5	5	1	1	5	5	5	5	1	5	5
7	0	1	3	5	7	1	3	5	7	5	7	3	7	7
9	0	1	1	1	1	9	9	9	9	1	1	9	9	9
11	0	1	3	1	3	9	11	9	11	1	3	11	11	11
13	0	1	1	5	5	9	9	13	13	5	5	9	13	13
15	0	1	3	5	7	9	11	13	15	5	7	11	15	15
21	0	1	1	5	5	1	1	5	5	21	21	5	21	21
23	0	1	3	5	7	1	3	5	7	21	23	13	23	23
27	0	1	3	1	3	9	11	9	11	5	13	27	27	27
31	0	1	3	5	7	9	11	13	15	21	23	27	31	31
63	0	1	3	5	7	9	11	13	15	21	23	27	31	63

x	0	1	3	5	7	9	11	13	15	21	23	27	31	63
x′	63	31	15	23	7	27	13	11	3	21	5	9	1	0
x⊠x′	0	1	3	5	7	9	9	9	3	21	5	3	1	0

...continued next page...

In the set $\mathbf{S} = \{0, 1, 3, 5, 7, 9, 11, 13, 15, 21, 23, 27, 31, 63\}$ of all 14 Necklaces of circumference 6 there are $2^{14}-1 = 16383$ non-empty subsets, $2^{14-1} = 8192$ necklace subsets containing the $\boxplus$-Zero `0` , and $2^{14-2} = 4096$ necklace subsets containing the $\boxtimes$-Unit `U` but not the $\boxplus$-Zero `0`.

PROBLEM : Determine
how many additive $\boxplus$-SubNPacks,
how many multiplicative $\boxtimes$-SubNPacks,
how many $\boxplus$-$\boxtimes$-SubNBoxes, and
how many NIdeals,
there are in $\mathbf{S}$.

APPENDIX C

C1 : SOME FORMULAE FOR COUNTING NECKLACES

ZEDNECKLACES ($n \geqslant 1$, $k \in \mathbf{Z}$, $k \geqslant 2$)

The number of necklaces with n beads of $k \geqslant 2$ colours equivalent under the cyclic group of order p (i.e. '**zednecklaces**') is given [see: [GR], [LW]] by

$$\begin{aligned} C(n,k) &= \frac{1}{n} \sum_{d|n} \phi\left(\frac{n}{d}\right) k^d \\ &= \frac{1}{n} \sum_{d|n} \phi(d)\, k^{n/d} \end{aligned} \tag{1}$$

where the summation control $d|n$ means that the summation is over just those 'integers d that divide integer n'.

Here, ϕ is the EULER 'TOTIENT' FUNCTION $\phi(m)$, the number of integers k with $1 \leqslant k \leqslant m$ having only 1 as a divisor in common with m, which is given in closed form [see : [AS]] by:

$$\phi(m) \;=\; m \prod_{i=1}^{r} \left(1 - \frac{1}{p_i}\right) \tag{2}$$

for any positive integer m expressed as a product of prime numbers $m = p_1^{a_1}\, p_2^{a_2} \ldots p_r^{a_r}$. In particular, if m is a power of 2 ($m = 2^{\alpha}$), then $r = 1 = i,\ p_i = 2$ and $\phi(m) = \phi(2^{\alpha}) = \frac{m}{2} = 2^{\alpha-1}$. And, if m is prime, then $r = 1 = i,\ p_i = 2$ and $\phi(m) = m\ \left(1 - \frac{1}{m}\right) = m - 1$.

Whence, if $n = p$ is prime, then the only divisors d of p are 1 and p, and in this case we have the closed formula:

$$\begin{aligned} C(p,k) \;&=\; \frac{1}{p}\ \left(\ \phi(p)\, k^1 + \phi(1) \times k^p\ \right) \\ &=\; \frac{1}{p}\ \left(\ (p-1)\, k + 1 \times k^p\ \right) \\ &=\; \frac{k}{p}\ \left(\ (p-1) + q^{(p-1)}\ \right) \end{aligned} \tag{3}$$

The number of necklaces with n beads of $k \geqslant 2$ colours equivalent under the dihedral group of order n (i.e. 'flip-over-able zednecklaces', also known as '**zedbracelets**') can be given [see : [LW], page 464] in terms of the number $C(p,q)$ by

$$(\text{even } n) \qquad D(n,k) \;=\; C(n,k)/2 \;+\; k^{n/2}(k+1)/4 \tag{4}$$

$$(\text{odd } n) \qquad D(n,k) \;=\; C(n,k)/2 \;+\; k^{((n+1)/2)}/2 \tag{5}$$

BITNECKLACES ($n \geqslant 1,\ k = 2$)

The number of necklaces with n beads of $k = 2$ colours equivalent under the cyclic group of order n (i.e. '**bitnecklaces**')

$$C(n,2) \equiv C(n) = \frac{1}{n} \sum_{d|n} \phi\left(\frac{n}{d}\right) 2^d$$
$$= \frac{1}{n} \sum_{d|n} \phi(d)\, 2^{n/d} \tag{6}$$

And if n is prime then (see (3))

$$C(n) = \frac{2}{n} \left((n-1) + 2^{(n-1)} \right) \tag{7}$$

For example, with $n = 7$,
$C(7) = \frac{2}{7} \left(6 + 2^6 \right) = \frac{2}{7} \times 70 = 20$;
Again, with $n = 137$,
$C(137) = \frac{2}{137} \left(136 + 2^{136} \right) = \frac{2}{137} \left(136 + 8.711..\times 10^{40} \right)$
$= 1.27171220\ldots\times 10^{39}$;
whilst $C(136) = 0.64053\ldots\times 10^{39}$;
and $C(138) = 2.52499\ldots\times 10^{39}$.

...continued next page...

BitBracelets ($n \geqslant 1,\ k = 2$)

The number of necklaces with n beads of $k \geqslant 2$ colours equivalent under the dihedral group of order n (i.e. '**zedbracelets**') is given in terms of the number of bitnecklaces in a bitnecklace-cyclic-class, $C(n)$, by

$$(\text{even } n) \qquad D(n) = C(n)/2 + 3 \cdot 2^{((n-4)/2)} \tag{8}$$

$$(\text{odd } n) \qquad D(n) = C(n)/2 + 2^{((n-1)/2)} \tag{9}$$

(That $D(n) = C(n)$ for $n = 1, 2$ is plain by inspection.)

Some more illustrations are given in the following table.

Partial Census of BitNecklaces and BitBracelets

in Cyclic $C(n)$ & Dihedral $D(n)$ equivalence classes of circumference n and the Euler totient function $\phi(n)$

Note that :
prime integers n are marked n^*, and $\phi(n \text{ prime}) = n - 1$.
Note also that for $n = 2^\alpha$ we have:
$C(n) \sim 2^{2^\alpha - \alpha} = 2^\beta \sim 10^{0.3 \times \beta}$,
$D(n) \sim 10^{0.3 \times \beta}/2$,
$\phi(n) = 2^{(\alpha-1)}$.

Part 1 of 2

$[a^2](2^b)$	n	bitNecklaces C(n)	bitBracelets D(n)	$\phi(n)$
(2^1)	2*	3	3	1
	3*	4	4	2
$[2^2](2^2)$	4	6	6	2
	5*	8	8	4
	6	14	13	2
	7*	20	18	6
(2^3)	8	36	20	4
$[3^2]$	9	60	46	6
	10	108	78	4
	11*	188	126	10
	12	352	224	4
	13*	632	380	12
	14	1,182	687	6
	15	2,192	1,224	8
$[4^2](2^4)$	16	4,116	2,200	8
	17*	7,712	4,112	16
	18	14,602	7,685	6
	19*	27,596	14,310	18
	20	52,488	27,012	8
	21	99,880	50,964	12
	22	190,746	96,909	10
	23*	364,724	184,410	22
	24	699,252	352,698	8
$[5^2]$	25	1,342,184	675,188	20

Part 2 of 2

$[a^2](2^b)$	n	bitNecklaces $C(n)$	bitBracelets $D(n)$	$\phi(n)$
(2^5)	32	134,219,796	67,159,050	16
$[6^2]$	36	1,908,881,900	954,637,558	12
$[7^2]$	49	11,488,774,559,636	5,744,404,057,034	12
$[8^2](2^6)$	64	$2.88230...\times10^{17}$	$1.44115...\times10^{17}$	32
▶	127*	▶ $1.33969...\times10^{36}$	$0.66984...\times10^{36}$	126
(2^7)	128	$2.65846...\times10^{36}$	$1.32923...\times10^{36}$	64
▶	137*	▶ $1.27171...\times10^{39}$	$0.63586...\times10^{39}$	136
$[16^2](2^8)$	256	$4.52313...\times10^{74}$	$2.26156...\times10^{74}$	128
(2^9)	512	$2.61871...\times10^{151}$	$1.30936...\times10^{151}$	256
$[32^2](2^{10})$	1024	$1.75556...\times10^{305}$	$0.87778...\times10^{305}$	512

7 BIBLIOGRAPHY

[AS] M. Abramovitch and I. A. Segun, *Handbook of Mathematical Functions*, Dover, New York, 1968 (1968, 5th printing).
See page 826, and especially the Tables on pages 840–843.
NB: The *NIST Handbook of Mathematical Function*, Ed. Frank W.J. Olver *et al.*, CUP (2010), the 'successor' to the above Handbook, has less information on the Euler Totient Function, *etc.* (compare above-mentioned pages and tables with §27.2 and Table 27.2.2, p.639, *ibid.*)

[AB] John Amson and Keith Bowden. *A Combinatoric Bit-Hoop System*, ANPA Proceedings No 26, 407–440. (2005).

[GR] E. N. Gilbert and John Riordan. *Symmetry types of periodic sequences*, Illinois J. Math. Volume 5, Issue 4, 657–665, (1961).

[LW] Lint, J.H. and Wilson,R.M., *A Course in Combinatorics*, CUP, (1992, reprinted 1996) See especially Chapter 35 : *Pólya theory of counting*;

[LyW] `Wikipedia https://en.wikipedia.org/wiki/Lyndon_word [2018]`

PIONS IN NUCLEONS

GEOFFREY CONSTABLE

106 Kings Road
Berkhamsted
Herts HP4 3BP

The magnetic moments of the proton and neutron have long been considered anomalous. It is proposed that such anomalies might be explained by the presence of pions in nucleon structures.

1 THE HISTORY.

The history of experiments and theoretical studies into nuclear magnetic moments is long. The anomalously large magnetic moment of the proton was first observed by Stern in 1933 in Hamburg, using a much modified and more sensitive apparatus than that used with Gerlach to resolve beam splitting due to electron spin.

The nuclear magneton, the magnetic moment one might expect from a nucleon, is given by

$$\mu_N = \frac{e\hbar}{2M_p},$$

where μ_N is the nuclear magnetron, and M_p is the mass of the proton. The Stern experiment showed that the measured magnetic moment of the proton was larger than this value by a factor of almost three. (Note: the magnetic moment of a particle generally varies reciprocally with particle mass. Hence, the magnetic moment of the electron is some three orders of magnitude greater than that of the proton).

In 1934, Stern, (Pittsburg), and Rabi, (New York), measured the magnetic moments of both the proton and the deuteron, and deduced

thereby that the magnetic moment of the neutron is significant but negative. This finding was verified, using studies of the hyperfine structure of atomic spectra, by Bacher at Ann Arbor (1933) and by Tamm and Altshuler 1934 in the Soviet Union. The Soviet result that $mu_n = -0.5\mu_N$ was viewed with surprise and was not widely accepted.

In 1936 Rabi and collaborators, with a much improved technique involving molecular beams, and also using the principle of magnetic resonance, produced accurate measurements of the proton magnetic moment. Comparing these with that of the deuteron confirmed the negative value for the neutron.

The first accurate determination of the magnetic moment of the neutron was made in 1940 at Berkeley, California, by Alvarez and Bloch. Their technique was to employ the powerful neutron emission from a cyclotron. Due to their large power of penetration, neutrons cannot readily be collimated by slits and the researchers had to rely on a cadmium tube as collimator, protected by large blocks of paraffin wax to eliminate stray neutrons.

Subsequent experiments have determined the magnetic moments of both the proton and neutron to very high accuracy – see later. However, despite an extended programme of theoretical studies, these values remain unexplained and, therefore, anomalous.

For example, it was thought at first that Quantum Electrodynamics (QED), which had so successfully explained the magnetic moment of the electron, might do the same for the proton and the neutron. Despite many papers that described, amongst other things, the magnetic moments generated by the cloud of virtual particles (pions and photons) that might arise from local nuclear and electromagnetic forces, predictions of nucleon magnetic moments (particularly that of the neutron) failed to match reality.

Hopes of solving this mystery revived in the 1960s with the arrival of the quark model. It was quickly decided that the neutron had one up quark (charge $2/3e$) and two down quarks (charge $-2/3e$). Quarks were believed to possess magnetic moments and these were apportioned on the assumption that each quark had a mass one third that

of a nucleon. At first, this seemed to provide results that were good – see following table.

Baryon	Magnetic moment of quark model	Computed μ_n	Observed μ_n
P	$4/3\mu$ — $1/3\mu_d$	2.79	2.793
n	$4/3\mu_d$ — $1/3\mu_u$	−1.86	−1.913

Alas, it was later realised that the quark masses had been significantly over-estimated, and should be only about 1% of a nucleon mass - most of the mass of a nucleon, according to the standard model, being accumulated in gluon fields.

Attempts to find a theoretical understanding of proton and neutron magnetic moments continue to the present. For example, a recent attempt using the nuclear physics associated with lattice quantum chromodynamics (QCD) demonstrated for the first time the ability to calculate the magnetic structure of some of the lightest nuclei. Large amounts of computing power were needed and the outcome is published in Physical Review Letters under the title, *Pinpointing the magnetic moments of nuclear matter* 12 Feb 2015 (Kathy Kincaid).

That said, the team admit that, *The theoretical magnetic moments of these nuclei at 'unphysically' large masses are remarkably close to those found in physical experiments*. If realistic quark masses had been chosen, the results of the study might have been different. Although such work is promising and deserves support, as of now the mystery of the magnetic moments of the proton and neutron remains unsolved.

...continued next page...

2 THE FACTS

	Mass (MeV)	Spin	Charge	Magnetic moment (μ_N)
Proton	938.2720813	1/2	$+1e$	2.7928473508
Neutron	939.5654133	1/2	$0e$	-1.91304273
Pion π^+	139.57019	0	$+1e$	
Pion π^-	139.57019	0	$-1e$	
Pion π^0	134.9766	0	$0e$	

The neutron decays to a proton, plus electron, plus (it is believed but not confirmed so far) a neutrino.

3 THE ARGUMENT

Assume, first, that both the proton and the neutron have some form of structure and are not simply homogeneous bodies of matter. (This is widely believed to be so and is not controversial). Assume, too, that the change from neutron to proton involves a change in just one element of the neutron structure, and that the change from neutron magnetic moment to proton magnetic moment can be attributed to this change alone.

The net change is given by the addition

$$1.91304272 + 2.7928473508 = 4.70589007\mu_N$$

a seemingly unremarkable result. The sum is now divided by the fraction $\pi/4$. [1] The division yields 5.99172534 surprising close to the integer six.

1 Although there is an element of empiricism is this action, is not entirely arbitrary or casual. When using digital methods to calculate the fine structure constant and the magnetic moment of the electron, series such as $1 + 1/2 + 1/3 + 1/4 \ldots\ldots$ and $1 + 1/2^2 + 1/3^2 + 1/4^2 = \pi^2/6$ had to be used in this way. To use the series $1 - 1/3 + 1/5 - 1/7 \ldots. = \pi/4$ in a similar way is not totally strange.

Next, suppose the nuclear magneton were calculated in terms of the neutron mass, not that of the proton. This thought involves dividing the number above by the ratio between these two masses, namely 0.99862347, yielding 5.99998450 also remarkably close to the integer six — a fact that must have meaning.

On the assumption that this number really is six, the next step is to ask what element in the neutron structure could deliver a change of magnetic moment of six nuclear magnetons. As magnetic moment varies reciprocally with mass, the answer to this question is an element with a mass one sixth that of the neutron, i.e. a particle with a mass of 156.59 MeV.

At first sight this result is puzzling. Such a mass does not identify any known particle. However, experience gained while studying the electron teaches that it is normal for sub-particles to be in motion, and such motion is likely to be rapid and reciprocally quantised (i.e. $c/2, c/3, c/4 \ldots$). As such, relativistic effects come into play.

The first guess is that the particle travels at half the speed of light. To calculate the rest mass 156.59 MeV needs to be multiplied by $\sqrt{1 - v^2/c^2}$ or $\sqrt{3/4}$ or 0.866025. This yields a rest mass of 135.61 MeV, a good match with that of the pion.

So the argument now leads to the concept of a neutron that consists of six orbiting and/or rotating pions, some charged and others not, the orbits arranged to provide the measured magnetic moment and half integral spin, but no charge. When a charged pion changes to an uncharged pion (with the emission of an electron), the particle becomes a proton with proton magnetic moment, charge, and spin.

Such a paragraph is easy to write but the process can be hard to visualise. One problem is that a pion has no spin or magnetic moment, so how can six pions produce a proton that has half integral spin and a magnetic moment? A helpful thought is that there may be subtleties with regard to the structure of the pion. Perhaps it has a spinning nucleus and an orbiting electron – no nett spin but one charge and, given a suitable binding of pions, an magnetic moment as well. If the electron is emitted, half integral spin, a $1e$ charge deficit and a different magnetic moment remain.

4 SUMMARY

Experimental evidence for the thoughts expressed above is in short supply. But it is widely believed that virtual pions exist in the nucleus as a means of transmitting force and, indeed, Yukawa predicted the existence and mass of the pion on just such a basis. That said, it is acknowledged that the thoughts expressed in this brief paper are tentative and exploratory.

ORDER AND DISORDER IN THE TORAH.

GRENVILLE J. CROLL
grenvillecroll@gmail.com

We propose to transform the characters of a small part of the Torah into the binary domain. We intend to sequence the resulting bit arrays and place them in a logical order, then reflect that ordering within the original Hebrew domain. We will examine the order and disorder of these bit arrays and sequences using the BiEntropy function in order to test for useful properties and to develop heuristics to possibly assist this cryptographic project.

1 INTRODUCTION

We have long held a working hypothesis [Croll, 2007b] that the Torah – the first five books of the Hebrew Bible – is a cryptogram. More specifically, we assume that the Torah is a transposition cipher. Our use of the Koren digital edition of the Torah [Computronic, 1997] is axiomatic. We have previously demonstrated, at least in principle [Croll 2007a], that it is possible to algorithmically permute a small part of the Torah into multiple alternate readable sequences. Some of these derivative Biblical Hebrew sequences were commercially translated into English, however they lacked lucidity and end-to-end readability. Nevertheless, in our view, the entire Torah is a long anagram that is intended to be decoded. We have identified some circumstantial evidence to support our point of view [Croll, 2011]. Furthermore, there is a long tradition that the letters of the Torah should one day be placed into an alternative disposition, thereby creating a New Torah.

Our previous work came to an end some years ago due to insurmountable combinatorial complexity. We had intended to use the small part of the Torah located between the inverted letter nuns (Numbers 10, 35&36), which we have named T2, as a key library. Our plan was to permute T2 using our Algorithm Three and then apply the parameters used to generate any lucid permutations of T2 – the keys - to the entire Torah (denoted T1T2T3) in the hope of achieving a plain text decryption. We were able to plan to do this because the dimensionality of T2

(the number of prime factors of its length) is the same as the dimensionality of T1T2T3, with in addition one dimension (prime factor of the length) in common. Unfortunately this plan could not succeed as there is no quantitative definition of lucidity and therefore no limit to the number of potential keys. Nor is there any means of determining the order of application of a sequence of keys, making the sequence of application of a set of non-commutative keys (see Appendix Three) subject to a factorial and thereby intractable number of permutations.

Many years ago we identified another course of action that constrained the number of keys to a small number. There remained the intractable problem that there was still a factorial number of permutations of the keys. However, within the small number of keys that we identified, we noticed that there was a simple way of sequencing some of the keys.

Recent work on the nature of order and disorder led us to the invention of our BiEntropy function [Croll, 2013] which operates on binary strings. Its success in a variety of disparate fields (in particular estimating the degree of order and disorder in the prime constant) led us to wonder if transforming the Torah to binary could be a useful strategy. We immediately noticed that transforming permutations of T2 to binary significantly increased the possibility of generating a long set of sequenced keys with a well identified start and end point to the sequence.

In this paper we describe the particular set of keys we have identified, the binary transformations we plan to subject them to together with some simple illustrative examples. We hope to describe in a later paper the results of our work.

2 HEBREW ELOHIM KEYS

2.1 Basic permutations of T2

There are 85 Hebrew characters between the inverted letter nuns. Our *a priori* Torah decryption methodology suggests that they should be laid out in 5 rows of 17 character as per Figure 1. Note that the central "pivot" column, marked with an asterisk, comprises the characters HH!'A (See Appendix One for a Hebrew transliteration table).

Switching rows 01234 to 43021 – per our algorithm, changes the order of the characters in the central column to A'H!H or "One God" in English. 'H!H is the Tetragrammaton or singular Divine name. Notice also that the word APR'L appears bottom to top in the first column. Figure 2 was first drawn in April 1999.

Figure 1 – Original Text of T2

```
                *
ר מ א י נ ו ר א ה ע ס נ ב י ה י ו   0
א ו צ פ י ו ה ו ה י ה מ ו ק ה ש מ   1
פ מ כ י א נ ש מ ו ס נ י ו כ י ב י   2
י ה ב ו ש ר מ א י ה ח נ ב ו כ י נ   3
ל א ר ש י י פ ל א ת ו ב ב ר ה ו ה   4
```

Figure 2 – Row order switched to 43021

```
                *
ל א ר ש י י פ ל א ת ו ב ב ר ה ו ה   4
י ה ב ו ש ר מ א י ה ח נ ב ו כ י נ   3
ר מ א י ו נ ר א ה ע ס נ ב י ה י ו   0
פ מ כ י א נ ש מ ו ס נ י ו כ י ב י   2
א ו צ פ י ו ה ו ה י ה מ ו ק ה ש מ   1

^APRעL          ^AעH!H
```

In our previous work we generated about 15 billion permutations of T2, of which 850 passed a set of four deterministic and statistical readability tests. Although it was possible to increase the number of permutations generated to achieve higher readability scores, it proved impossible to find a way of easily translating or otherwise evaluating the resulting sequences for lucidity and so the work came to a natural end.

2.2 The plural Divine Name

Having noticed that the central or "pivot" column of T2 contained the characters of the singular Divine name, we were curious to see if any permutations of T2 per our algorithm, contained the plural Divine name "ALH'M" – Elohim, which is five characters long. There are other divine names and a wide choice of five character words or short

expressions which could have been (but were not) subject to our enquiry as candidates for the central pivot column of the permutations of T2.

Since the number of permutations of T2 available to us was very large, we used a monte carlo method to pseudo-randomly generate permutations of T2. We tested each permutation to see if the central characters of T2 contained the sequence ALH'M. Using a series of computer runs of differing durations, we were able to determine that it was very likely that there are exactly 64 permutations of T2 containing Elohim in the central pivot column. We provide an example in Figure 3.

Figure 3 – An Elohim Key

									*						
ה	ר	מ	א	י	ו	נ	ר	א	ה	ע	ס	נ	ב	י	ה	י
י	ל	א	ר	ש	י	י	פ	ל	א	ת	ו	ב	ב	ר	ה	ו
א	נ	י	כ	ו	ב	נ	ח	ה	י	א	מ	ר	ש	ו	ב	ה
י	מ	ש	ה	ק	ו	מ	ה	י	ה	ו	ה	ו	י	פ	צ	ו
ו	פ	מ	כ	י	א	נ	ש	מ	ו	ס	נ	י	ו	כ	י	ב

There is a small set of integer parameters associated with each T2 key that specifies exactly how to generate it from the original version of T2. This small set of parameters can also be applied to the whole Torah to transform it in its entirety into an alternative disposition of indeterminate utility.

2.3 Key Sequencing

By definition each Elohim key has the Hebrew ALH'M in the central column. However we noticed that there were many occasions where the first and last columns of one Elohim key matched either the last or first column of another Elohim key. It was thus possible to sequence up to six Elohim keys together in this way just as in the ancient table game of Dominoes. We illustrate this in Table 1 below. We give each left or right edge sequence which has a match (e.g. APR'L) a two character code. (e.g. AB). The fifteen central columns are denoted by the dot. Unique keys are denoted XX.

Thus there are 18 key pairs, 15 unmatched keys, two sets of 3 and 4 repeated keys and one sequence of 6 keys making 64 total. All the keys are unique, even though their left and right hand edges may match. A brief calculation of the *a posteori* probability of being able to find a sequence of six linked keys in a random trial is very low. See Appendix One for Hebrew character probabilities.

Table 1 – Sequencing the Elohim Keys

KeyPair	AB.AA	AA.AB				
KeyPairs	AC.AD	AC.AD	AC.AD	AC.AD		
KeyPair	AE.AF	AF.AE				
KeyPair	AG.AH	AH.AG				
KeyPair	AI.AJ	AJ.AI				
KeyPair	AK.AL	AL.AK				
KeyPairs	AM.AN	AM.AN	AM.AN			
KeyPair	AO.AP	AP.AO				
KeyPair	AQ.AZ	AZ.AQ				
KeyPair	AU.AT	AT.AU				
KeySeq	XX.AT	AT.AS	AS.AU	AU.AR	AR.AW	AW.XX
KeyPair	BA.BB	BB.BA				
KeyPair	BC.BD	BD.BC				
KeyPair	BF.BE	BE.BF				
KeyPair	BG.BH	BH.BG				
KeyPair	BI.BJ	BJ.BI				
KeyPair	BK.BL	BL.BK				
KeyPair	XX.BM	BM.XX				
KeyPair	BN.BO	BO.BN				
KeyPair	BP.BQ	BQ.BP				
KeyPair	BR.BS	BS.BR				
Unique	XX.XX	15 keys unique and unsequenced				

2.4 Proof of Perfection

We have not yet investigated the application of the Elohim keys to the whole Torah. Prior to doing so, we would wish to investigate an additional long term *a priori* hypothesis of ours that application of shorter key sequences may transform the Torah into a highly ordered state suggestive of perfection. For example, use of a key or a key pair as above might transform the Torah into a simple series of repeating letters. Alternatively, one or more of the key pairs above might transform the Torah into a different document and then back again into the same document, the original Torah. The two sets of keys AM.AN (of which

there are three) and AC.AD (of which there are four) might each generate the same derivative documents. With perhaps the subtle difference that whereas all the resulting letters were the same and in the correct order, the origin of each letter might differ. This ordering might appear across the dimensions in a manner similar to the "Topological Interlock" mechanism outlined in our previous work [Croll, 2007b]. We would expect these types of ordered states to be specific to the Torah.

2.5 Plain text decode of entire Torah

Use of six Elohim keys such as the six sequenced keys as above could potentially generate approximately 10^{30} derivative documents. This could conceivably imply sufficiently many degrees of freedom to transform the Torah into a New Torah. We have not yet attempted to apply these six keys to the whole Torah, as this would be only a singular action within a vast combinatorial space. Conversely, use of all 64 Elohim keys implies exploring a space of 64! combinations. Use of the 15 unique keys alone implies 15! or 1.3E12 documents to investigate. The readability tests of our previous work (QPT, Word & Path tests) enable us to easily and quickly test for systematic readability of a derivative document without having to examine all of it.

Our preference therefore is to attempt to increase the number of sets of sequenced keys in order that there are sufficiently many degrees of freedom between the Torah and any New Torah. A small library of long sequences of keys provides an alternative to a computationally intensive search within an intractable number of unsequenced keys.

It is for the above reasons that we now propose to investigate the binary transformations of the Elohim keys.

3 BINARY TRANSFORMATION

In order to preserve the integrity of our decryption principles and strategy, we use a simple substitution cipher [Schneier, 1996] to transform the Torah into the binary domain. Since there are 22 characters of the Hebrew Alefbet, there are 2^{22} ways of allocating 1's and 0's to the

Hebrew characters. These include the trivial possibilities of substituting every Hebrew characters by a 0 or every Hebrew character by a 1.

An alternative would be to choose a substitution method that resulted in the total number of 1's and 0's in the binary Torah being exactly equal. Or, since the Torah (T1T2T3) is of length 304,805, which is odd, the total number of 1's and 0's should differ by one. This is a simple computation and we show in Table 2 (Appendix Two) half of the various possibilities, of which there are 84 in total. The frequencies of the characters of the Koren edition of the Torah are given in Appendix One.

4 BINARY ELOHIM KEYS

4.1 Transformation to Binary

Use of the binary character allocations in Table 2 (Appendix Two) enables us to transform the Hebrew Elohim keys to binary. We show in Figure 4 two such Binary Elohim keys transformed using the fifth row (labeled 4) of Table 2 (Appendix Two). For brevity the left and right edges of each key are labelled in decimal according to the numeric value of the binary in the edge column. Thus in the first key the left edge 01001 = 9 and the right edge 10011 = 19. Note that the ninth column of each key is 01101, corresponding to ALH'M, obtainable by reference to Appendix One.

Figure 4 – Two Binary Elohim keys

```
     00001000011000011
     11000110110100110
  9  00101010110101010 19
     01001000000000001
     10110011111000011

     11010001011111110
     01000011100000001
 19  01000110101010110  9
     11100101000010000
     10000001101000001
```

The above two Elohim keys 09.19 and 19.09 are linked by their matching edges. We believe that a circularity in their use may be implied. There are 84 possible binary transformations through Table 2 and its 1's complement. Thus in our view there are 84 sets of 64 Elohim keys. Note that, coincidentally, the number of unique permutations of T2 via our Algorithm One is 84.

4.2 Linked Binary Elohim Keys

Figure 5 – Linked Binary Elohim keys

Seq Start	03.22	22.00	00.22	22.03	03.01	01.02	02.02	02.30
	30.20	20.02	02.17	17.13	13.00	00.11	11.06	06.16
	16.18	18.23	23.18	18.16	16.08	08.01	01.09	09.19
	19.09	09.01	01.16	16.04	04.04	04.10	10.14	14.10
	10.29	29.10	10.04	04.24	24.26	26.00	00.21	21.12
	12.08	08.05	05.28	28.05	05.27		Seq End	

Unique	00.26		Unique	18.25		
Unique	02.30		Unique	21.00		
Unique	02.30		Pair	23.08	08.01	
Pair	07.20	20.30	Pair	23.08	08.12	
Unique	15.09		Unique	23.08		
Unique	16.06	06.11	Unique	23.08		
Pair	17.02	02.30	Unique	26.24		

Figure 4 makes clear how binary keys and by implication the corresponding Hebrew keys may be linked. We show in Figure 5 above the linkages possible using the full set of 64 binary Elohim keys including the two of Figure 4 above.

The linkages are determined from the binary transformation. The production parameters used to generate the sequence of keys can then be applied in the same order directly to T1T2T3 or to T2 in Hebrew or binary.

4 ENTROPY TESTING AND SEARCH HEURISTICS

4.1 Entropy Testing

Figure 6 - Ordered Elohim Keys

	00000000000000000	
	11111111111111111	
10	00000000000000000	10
	11111111111111111	
	00000000000000000	
	01010101010101010	
	01010101010101010	
0	01010101010101010	0
	01010101010101010	
	01010101010101010	
	11111111011111111	
	11111111011111111	
31	11111111011111111	31
	11111111011111111	
	11111111011111111	
	00000000000000000	
	00000000000000000	
4	11111111111111111	4
	00000000000000000	
	00000000000000000	

One of the obvious problems of the Torah is that it is a large document. Following a binary transformation however, it is relatively easy to compute its BiEntropy. Our BiEntropy function returns a value between zero and one depending upon the degree of order and disorder in the bits of a binary string of arbitrary length. A BiEntropy of zero implies a perfectly ordered string such as 00000000 and one implies a disordered string such as 11010001. In the search for proof of perfection it would be relatively straightforward to detect patterned behavior in the digits of a binary Torah, whether in the linear flow of the digits or across the dimensions as per the topological interlock described in

our previous work. Patterned behavior would result in a lower BiEntropy. Elohim keys exhibiting ordered behaviors such as those illustrated in Figure 6 above would be of particular interest. They have a BiEntropy close to zero.
The *a priori* probability of the existence of binary keys such as these is extremely low (<<***O*** E-80). The patterning of the keys might be suggestive of the patterning to be expected in the whole Torah following application of the same key generation parameters.

4.2 Search Heuristics

The BiEntropy of the binary Elohim keys can be easily computed. It would be interesting to note if there was any change in the BiEntropy of the entire Torah following application of one or more Elohim keys. If so, there is a possibility that a heuristic could be used to search for consistent increases in the BiEntropy of a partly decoded document following the application of each key. We expect that a more readable document would have a higher BiEntropy than an random document. This possibility can be easily tested against our existing readability metrics.

4.3 Self Application

Note that the Elohim keys can be applied to T2 both in the original Hebrew and in the binary transformations. Interesting Elohim key transformations obtained or tested on T2 can then be applied at higher computational cost to the entire Torah.

5 SUMMARY

We have introduced our previous work on the Torah decryption problem and outlined a new strategy using a finite and potentially tractable set of decryption keys. We have shown how the keys can be easily sequenced both in the original Hebrew and even more readily in binary – and then applied to the entire Torah. We introduce the idea of entropy testing and heuristics using our BiEntropy function as a possible key search mechanism. We hope to report further progress in this work on a future occasion.

REFERENCES

Computronic Inc., (1997) Bible Codes, P.O.Box 102, Savyon, Israel

Croll, G.J., (2000a), Some A Priori Torah Decryption Principles, Proc. ANPA 28, Cambridge, http://arxiv.org/abs/0712.2605

Croll, G.J., (2000b), Algorithmic Permutation of the Torah, Proc. ANPA 27, Cambridge, http://arxiv.org/abs/0712.2606

Croll, G.J., (2011), A Chronology of Torah Cryptography, Proc. ANPA 31, Rowlands Castle, https://arxiv.org/abs/1101.5313

Croll, G.J., (2013), BiEntropy: The Approximate Entropy of a Finite Binary String, Proc. ANPA 34, Rowlands Castle, https://arxiv.org/abs/1305.0954

Schneier, B. (1996) Applied Cryptography, 2nd Edition, John Wiley & Sons

APPENDIX ONE

Frequency and probability of occurrence of the characters of the Koren Digital Edition of Torah with associated transliterations.

Hebrew Letter	English Letter	ASCII Symbol	Hebrew Symbol	Freq. of Occurrence	Probability of Occurrence
Alef	A	A	א	27059	0.089
Beit	B	B	ב	16345	0.054
Gimmel	G	C	ג	2109	0.007
Dalet	D	D	ד	7032	0.023
Hey	H	E	ה	28056	0.092

Vav	!	F	ו	30513	0.100
Zain	Z	G	ז	2198	0.007
Chet	X	H	ח	7189	0.024
Tet	T	I	ט	1804	0.006
Yud	'	J	י	31531	0.103
Kaf	K	L	כ	11968	0.039
Lamed	L	M	ל	21570	0.071
Inverted Nun	N	N	[	2	0.000
Mem	M	O	מ	25090	0.082
Nun	N	Q	נ	14126	0.046
Samech	S	R	ס	1833	0.006
Ain	Y	S	ע	11250	0.037
Pey	P	U	פ	4805	0.016
Tzadi	C	W	צ	3962	0.013
Kuf	Q	X	ק	4695	0.015
Reish	R	Y	ר	18125	0.059
Shin	$	Z	ש	15595	0.051
Tav	T	[	ת	17950	0.059

				304807	

APPENDIX TWO

Converting the Torah (T1T2T3) to binary ensuring that the number of 1's and 0's differ by only 1, since the length of the Torah is odd. For brevity, half the letter allocations are shown as the other half are symmetric whereby 1 and 0 are reversed. There are exactly 84 ways in total in which a binary character allocation can be done. The sequence/id is the decimal equivalent of the adjacent 22 bit binary sequence. The final column shows the number of 1's in the binary allocation.

Table 2 – Binary Character Allocations

	CHAR	Al	Be	Gi	Da	He	Va	Za	Ch	Te	Yo	Ka	La	Me	Nu	Sa	Ay	Pe	Tz	Qo	Re	Sh	Ta	#
	ASCII	A	B	C	D	E	F	G	H	I	J	L	M	O	Q	R	S	U	W	X	Y	Z	[	1's
	SEQ/ID	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	
0	135784	0	0	0	0	1	0	0	0	0	1	0	0	1	0	0	1	1	0	0	1	1	1	8
1	202010	0	0	0	0	1	1	0	0	0	1	0	1	0	1	0	0	0	1	1	0	0	1	8
2	244901	0	0	0	0	1	1	1	0	1	1	1	1	0	0	1	0	1	0	0	1	0	0	10
3	383778	0	0	0	1	0	1	1	1	0	1	1	0	1	1	0	0	1	0	0	0	0	1	10
4	413391	0	0	0	1	1	0	0	1	0	0	1	1	1	0	1	1	0	0	1	1	1	0	11
5	446904	0	0	0	1	1	0	1	1	0	1	0	0	0	1	1	0	1	1	0	1	1	1	12
6	458155	0	0	0	1	1	0	1	1	1	1	1	1	0	1	1	0	1	0	1	0	1	0	13
7	463713	0	0	0	1	1	1	0	0	0	1	0	0	1	1	0	1	1	0	0	0	0	0	8
8	471444	0	0	0	1	1	1	0	0	1	1	0	0	0	1	1	0	0	1	0	0	1	1	10
9	494525	0	0	0	1	1	1	1	0	0	0	1	0	1	1	1	0	1	1	1	1	0	0	12
10	610022	0	0	1	0	0	1	0	1	0	0	1	1	1	0	1	1	1	0	0	1	0	1	11
11	738748	0	0	1	0	1	1	0	1	0	0	0	1	0	1	1	0	1	1	1	0	1	1	12
12	776569	0	0	1	0	1	1	1	1	0	1	1	0	0	1	0	1	1	1	1	0	0	0	12
13	856812	0	0	1	1	0	1	0	0	0	1	0	0	1	0	1	1	1	0	1	0	1	1	11
14	879990	0	0	1	1	0	1	0	1	1	0	1	1	0	1	0	1	1	1	0	1	0	1	13
15	984669	0	0	1	1	1	1	0	0	0	0	0	1	1	0	0	1	0	1	1	1	0	0	10
16	1005666	0	0	1	1	1	1	0	1	0	1	1	0	0	0	0	1	1	0	0	0	0	1	10
17	1042862	0	0	1	1	1	1	1	1	1	0	1	0	0	1	1	0	1	0	1	1	0	1	14
18	1045539	0	0	1	1	1	1	1	1	1	1	0	1	0	0	0	0	1	0	0	0	1	0	11
19	1111734	0	1	0	0	0	0	1	1	1	1	0	1	1	0	1	0	1	1	0	1	0	1	12
20	1166590	0	1	0	0	0	1	1	1	0	0	1	1	0	0	1	1	1	1	1	1	0	1	13
21	1172974	0	1	0	0	0	1	1	1	1	0	0	1	0	1	1	1	1	0	1	1	0	1	13
22	1235972	0	1	0	0	1	0	1	1	0	1	1	1	0	0	0	0	0	0	0	0	1	1	9
23	1258729	0	1	0	0	1	1	0	0	1	1	0	1	0	0	1	1	1	0	1	0	0	0	10
24	1278726	0	1	0	0	1	1	1	0	0	0	0	0	1	1	0	0	0	0	0	1	0	1	8
25	1317432	0	1	0	1	0	0	0	0	0	1	1	0	1	0	0	0	1	1	0	1	1	1	10
26	1323816	0	1	0	1	0	0	0	0	1	1	0	0	1	1	0	0	1	0	0	1	1	1	10
27	1403998	0	1	0	1	0	1	0	1	1	0	1	1	0	0	0	1	0	1	1	1	0	1	12
28	1443748	0	1	0	1	1	0	0	0	0	0	0	1	1	1	1	0	1	0	0	0	1	1	10
29	1471597	0	1	0	1	1	0	0	1	1	1	0	1	0	0	0	1	1	0	1	1	0	0	11
30	1524459	0	1	0	1	1	1	0	1	0	0	0	0	1	0	1	1	1	0	1	0	1	0	11
31	1636672	0	1	1	0	0	0	1	1	1	1	1	0	0	1	0	0	1	1	1	1	1	1	14
32	1659348	0	1	1	0	0	1	0	1	0	1	0	0	0	1	1	1	0	1	0	0	1	1	11
33	1694633	0	1	1	0	0	1	1	1	0	1	1	0	1	1	1	0	1	0	1	0	0	0	12
34	1716607	0	1	1	0	1	0	0	0	1	1	0	0	0	1	0	1	1	1	1	1	1	0	12
35	1780758	0	1	1	0	1	1	0	0	1	0	1	1	0	0	0	0	0	1	0	1	0	1	10
36	1859050	0	1	1	1	0	0	0	1	0	1	1	1	0	1	1	1	1	0	1	0	0	1	13
37	1932041	0	1	1	1	0	1	0	1	1	1	1	0	1	1	0	0	0	0	1	0	0	0	11
38	1935293	0	1	1	1	0	1	1	0	0	0	0	1	1	1	1	0	1	1	1	1	0	0	13
39	1995167	0	1	1	1	1	0	0	1	1	1	0	0	0	1	1	0	0	1	1	1	1	0	13
40	2003008	0	1	1	1	1	0	1	0	0	1	0	0	0	0	0	0	1	1	1	1	1	1	12
41	2075731	0	1	1	1	1	1	1	0	1	0	1	1	0	0	0	1	0	1	0	0	1	0	12

APPENDIX THREE

We prove by counter example that successive application of permutations of T2 to an original text are non-commutative. I.e. for Keys A and B, BA ≠ AB. We use a later variation of Algorithm Three (Algorithm Seven) which is implemented using a 27 digit decimal key.

Original Text

Maryhadalittlelambitsfleeceaswhiteassnowand
everywherethatmarywentthelambwassuretogo!!

Key A
314159265358979323846264338
Key B
327950288419716939937510582

Step 1 Key A
ewwtsr!wnreesenwMtasnheouhbhetetedybolsyaal
eaea!tftstmviildeahmilgarohrylcttsawareama

Step 2 Key B – Result 1
toislrasaegeaahalrywmnlyYdheemtetwtltwtatiehn
mesobtaeahwofeeawcvbMaiZutrleerdhsarynss

Step 1 Key B
Sethnmir!ttertaydeaatswwdacetewgwastyebniaeal!ll
seevrafrlhhtraahsmonyeeomMsaewohiblut

Step 2 Key A – Result 2
esmettunhtcesteasaytafnienewalrmoirol!ylherws!Mtt
swthelatyearohgwbrlbewhaaasaeeivamde

The two results – Result 1 and Result 2 - of steps 2 are not equal, therefore successive permutations of T2 are non-commutative.

COSMOLOGICAL THEORIES OF THE EXTRA TERMS

ANTHONY M. DEAKIN
e-mail: amdeak.deakin@btinternet.com

LOUIS H. KAUFFMAN
Dept. of Mathematics, Statistics and Computer Science, University of Illinois at Chicago, 851 South Morgan Street, Chicago, Illinois 60680 USA
Tel: 773 363 5115; Fax:312 996 1491; e-mail: Kauffman@uic.edu

Introduction

This paper purports to have: Introduced a new formulation of Quantum Mechanics, explained the apparent disconnect between Quantum Mechanics and General Relativity, explained the observed far field expansion of the Universe (Dark Energy), supplied an argument which goes towards explaining away Dark Matter (there are modelling difficulties) and not explained, on the basis of gravitational theory, the Voyager Anomaly.

0. Concepts

Constraints Theory (CT) [22, 23] is a branch of theoretical physics. It begins with Quantum Mechanics (QM) but has connections with Classical Mechanics (CM) and Cosmology. Its original purpose was to explain why certain structures appear in CM on the basis that QM is fundamental; and why these structures are often successful as a basis for predictive/ descriptive quantum calculations about the real world. But the applications of CT are even wider.

CT is based on a formulation of QM that replaces scalar observables by Hermitian operators and differentials of scalar observables by commutators. It thereby uses the Schrodinger [1] method rather than the path integral method developed by Feynman [2]. But CT does not use the structures found in a Lagrangian or a Hamiltonian

formulation of CM to construct QM; a method used originally, in different ways, by both of those authors. One has sympathy with them: For how are they to inform the problem unless they impose structure? Where is the structure to come from apart from classical Lagrangians or Hamiltonians? After all Lagrangian mechanics and Hamiltonian mechanics have been very successful in predicting/ describing how the medium to large scale Universe works.

In CT the structure comes from something *inevitable*; the quantisation of an hierarchy of *differential identities*. We quantise these by methods which are roughly what Schrodinger did with his famous hydrogen model. We thereby bring in all the baggage (of coordinates and time etc.) associated with that model. In doing so we bring in half the assumptions of CM. We can be criticised for this; but we must start somewhere! In CT we then look for recognisable structure in the relations between various operators.

CT assumes Cartesian coordinates and conjugate momenta of particles in a flat, continuous space. The space, here denote, *P* may be the ordinary 3-space of Euclid or it may be the 4-space-time of Minkowski; but it is flat. *P* may contain more than one particle; and, indeed, it may contain many particles represented as a continuous fluid. The particles in *P* are structureless points with little more than coordinates, momenta and mass assigned to them. The coordinates and momenta are assumed all to be *continuous*. The coordinates of the particles are in turn assumed to be differentiable functions of a single, continuous scalar time. This time is the proper time of a single observer and an adjacent clock both at rest at the origin.

The differential identities concern the time derivatives of a continuous, differentiable function theta. Theta is assumed to be a function of the scalar coordinates which are, in turn, assumed to be functions of the continuous time experienced by the observer. There may be more than one function theta associated with a given system of particles; there is an hierarchy of identities associated with each. The candidates for theta (operator or scalar depending on the context) are taken to be the scalar functions of the coordinates that appear in the Hamiltonian (operator or scalar depending on the context); the Hamiltonian is taken to be the *complete description* of the system. The

first four of the differential identities are as follows (the q^j are the coordinates of the particles) [20]:

$$\dot{\theta} = \dot{q}^j \theta_{,j}; \quad \theta_{,j} \equiv \frac{\partial \theta}{\partial q^j}; \quad \dot{\theta} \equiv \frac{d\theta}{dt}$$

$$\ddot{\theta} = \ddot{q}^j \theta_{,j} + \dot{q}^j \dot{q}^k \theta_{,jk}; \quad \theta_{,jk} \equiv \frac{\partial^2 \theta}{\partial q^j \partial q^k}$$

$$\dddot{\theta} = \dddot{q}^j \theta_{,j} + 3\ddot{q}^j \dot{q}^k \theta_{,jk} + \dot{q}^j \dot{q}^k \dot{q}^l \theta_{,jkl};$$

$$\ddddot{\theta} = \ddddot{q}^i \theta_{,i} + (4\dddot{q}^i \dot{q}^j + 3\ddot{q}^i \ddot{q}^j)\theta_{,ij} + 6\ddot{q}^i \dot{q}^j \dot{q}^k \theta_{,ijk} + \dot{q}^i \dot{q}^j \dot{q}^k \dot{q}^l \theta_{,ijkl};$$
$$i, j, k, l = 1, 2, ...n_c \equiv n_d n_p$$

Etc.

The Einstein summation convention is in force; and, unless otherwise stated, all indices lie in the range $[1, n_c = n_p n_d]$ where n_p is the number of particles and n_d is the dimension of the flat space *P*.

Quantising $\dot{\theta} = \dot{q}^j \theta_{,j}; \quad \theta_{,j} \equiv \frac{\partial \theta}{\partial q^j}; \quad \dot{\theta} \equiv \frac{d\theta}{dt}$ means replacing differentials by commutators and using the product rule to obtain

$$\frac{1}{i\hbar}(\Theta H - H\Theta) = \frac{1}{2}(\Theta_{,j} \frac{1}{i\hbar}(Q^j H - HQ^j) + \frac{1}{i\hbar}(Q^j H - HQ^j)\Theta_{,j})$$

where H is the Hamiltonian (Hermitian) operator and

$$h \to H; \quad \theta \to \Theta; \quad \theta_{,j} \to \Theta_{,j} \equiv \frac{\partial \Theta}{\partial Q^j}; \quad q^j \to Q^j$$

where $\to$ means 'real observable corresponding to Hermitian operator'.
Thus we systematically replace all derivatives by commutators and then the costraints are demands that the commutator representations

behave according to (some of) the rules of continuum calculus. See Appendix A for an account of the first constraint and its consequences.

The constraints are operator equations in *all* the coordinate operators, *all* the momentum operators, the Hamiltonian operator and the theta operator(s). The momentum operators are, by definition, conjugate to the coordinates; the Hamiltonian operator is, similarly, conjugate to the time. A theta operator is defined as a pure function of the coordinate operators; in the position representation it reduces to a scalar function of the scalar coordinates. The constraints (quantizations) corresponding to the scalar equations above are [20]

$$\left\{H^{:j},\Theta_{,j}\right\}=\left\lfloor H,\Theta\right\rfloor;$$

$$\left\lfloor A,B\right\rfloor\equiv\frac{1}{i\hbar}(AB-BA);\quad \{A,B\}\equiv\tfrac{1}{2}(AB+BA);\quad A^{:j}\equiv\left\lfloor Q^{j},A\right\rfloor;\quad A_{,j}\equiv\left\lfloor A,P_{j}\right\rfloor;$$

$$i\hbar\delta_{k}^{j}=Q^{j}P_{k}-P_{k}Q^{j}\quad Q\text{'s commute; }P\text{'s commute}$$

$$\left\{\left\lfloor H,H^{:j}\right\rfloor,\Theta_{,j}\right\}+\left\{H^{:j},H^{:k},\Theta_{,j,k}\right\}=\left\lfloor H,\left\lfloor H,\Theta\right\rfloor\right\rfloor$$

$$\left\lfloor A,\left\lfloor B,C\right\rfloor\right\rfloor\equiv\left\lfloor A,B,C\right\rfloor;\quad \{A,B,C\}\equiv\tfrac{1}{3!}(ABC+CBA+CAB+ACB+BCA+BAC)$$

$$\left\{\left\lfloor H,H,H^{:j}\right\rfloor,\Theta_{,j}\right\}+3\left\{\left\lfloor H,H^{:j}\right\rfloor,H^{:k},\Theta_{,j,k}\right\}+\left\{H^{:j},H^{:k},H^{:l},\Theta_{,j,k,l}\right\}$$
$$=\left\lfloor H,H,H,\Theta\right\rfloor$$
$$\left\{A,B,...\right\}\equiv\left\{AB...\right\}\equiv\tfrac{1}{n!}\sum_{perm}AB...\text{ there being }n\text{ operators}$$

$$\left\{\left\lfloor H,H,H,H^{:i}\right\rfloor,\Theta_{,i}\right\}+4\left\{\left\lfloor H,H,H^{:i}\right\rfloor,H^{j},\Theta_{,i,j}\right\}+3\left\{\left\lfloor H,H^{:i}\right\rfloor,\left\lfloor H,H^{:j}\right\rfloor,\Theta_{,i,j}\right\}$$
$$+6\left\{\left\lfloor H,H^{:i}\right\rfloor,H^{:j},H^{:k},\Theta_{,i,j,k}\right\}+\left\{H^{:i},H^{:j},H^{:k},H^{:l},\Theta_{,i,j,k,l}\right\}$$
$$=\left\lfloor H,H,H,H,\Theta\right\rfloor$$

where, in particular, the operators (expressed in position representation)

$$\Theta(\underline{Q}) \equiv \theta(\underline{q})I; \quad \Theta_{,j} = \theta_{,j}I; \quad \Theta_{,j,k} = \theta_{,jk}I; \quad \Theta_{,j,k,l} = \theta_{,jkl}I \quad \text{etc.}$$

are pure in the $\underline{Q}$; the order of the suffices is immaterial.

The following notation is used above

$$a \to A; \quad \dot{a} \equiv \frac{da}{dt} \to \dot{A} \equiv \overset{1}{A}; \quad \frac{d^n a}{dt^n} \to \overset{n}{A}; \quad n = 1, 2, ...$$

where '$\to$' means 'a real variable is represented by the Hermitian operator'.

$$[A, B] \equiv AB - BA; \quad \lfloor A, B \rfloor \equiv \frac{1}{i\hbar}[A, B]$$

$$[A, B, C] \equiv [A, [B, C]]; \quad [A, B, C, D] \equiv [A, [B, C, D]];$$
$$\lfloor A, B, C \rfloor \equiv \lfloor A, \lfloor B, C \rfloor \rfloor; \quad \lfloor A, B, C, D \rfloor \equiv \lfloor A, \lfloor B, C, D \rfloor \rfloor$$

$$p_j \to P_j; \quad q^k \to Q^k; \quad A_{,j} \equiv \lfloor A, P_j \rfloor; \quad A^{:k} \equiv \lfloor Q^k, A \rfloor$$

$$\{A_1, A_2, ..., A_n\} \equiv \tfrac{1}{n!} \sum_{perm} A_1 A_2 .... A_n$$

where the commas on the LHS are inserted, if need be, only for clarity. The order of the arguments in $\{.\}$ is immaterial. Notice that if an element inside any of the brackets $[.], \lfloor . \rfloor, \{.\}$ is null then the bracket is null.

If, in the hierarchy of constraints, the first holds then the Hamiltonian operator can be proved to be *quadratic* in the momentum operators with pre and post coefficients that are pure, free functions of the coordinate operators; see Appendix A. The proof involves the assumption that the coordinate and the momentum operators are continuous.

The equations of motion are, in general, complicated operator equations. But, in the *classical approximation* (all operators commute), Hamilton's classical equations (with all the momenta eliminated) have the appearance of geodesic equations in a Riemannian manifold. The dimension, n_c, of this space is the product of dimension, n_d of *P* and the number of particles n_p. The coordinates of a point in this space comprise the aggregate of *all* the coordinates of the particles in *P*; so there is but one point in this space that represents the particles in *P*; it is here denoted *X*. A generalisation of this space is useful. We denote by *C* the space of all the coordinate operators of the particles in *P*, similarly aggregated, in the *position representation*. Thus *C* is an ordinary continuous space of dimension n_c with one point *X* representing the particles in *P*. The connectivity of the space *C* can be guaranteed Riemannian only if the scale is large enough (sufficiently large for CM to work). If a fluid is represented in *P* then the dimension of *C* is, strictly, infinite.

If, in the hierarchy of constraints, the second *also* holds then the so called Theta Equation (TE) can be derived; see Appendix B. The TE is an operator equation. In the position representation this reduces to a fourth order PDE with a theta (an ordinary scalar function of the coordinates) as the dependent variable and the coordinates as independent variables. As asserted above a theta operator is taken to be *any* one of the *pure coordinate operator functions* that appear in the quadratic QM Hamiltonian. The reason for this assumption is that the Hamiltonian characterises the system; and, if the first constraint holds then, these functions fully characterise the Hamiltonian and hence the system. The TE is, therefore, an archetypal *field equation* in this theory; as derived it is valid both in QM and CM.

The coordinates used in the TE, however, may not form a Riemannian space. If , nevertheless, we use these coordinates to identify points in *C*, we *define* the TE on *C*. But, because *C* may be a classical artefact (see above), the TE, in that case, may be only valid in CM.

The coefficients of the quadratic terms in the classical Hamiltonian are taken to describe classical gravitational forces; the linear and zero order terms are taken to describe classical EM forces. There

is one exception to this: when the zero order term is used as a classical Newtonian gravitational potential. When the linear and zero order terms are omitted from the Hamiltonian the TE is called the Gravitational Theta Equation (GTE). If the GTE is thought of as defined on C then, for the reasons set out above, it may only be valid in macrophysical situations.

We are, in what follows, concerned only with gravitation. But talk of gravitation implies, in General Relativity (GR) at least, curvature of space-time and, necessarily, the use of curvilinear coordinates. In general *C is* curved (the fundamental tensor of C is comprised of *free* functions of all the coordinates defined in *P*); so C, providing it is Riemannian should be able to accommodate Einsteinian gravitational theory. But there is here an apparent contradiction: *P*, by definition, is *flat*; and the coordinates of a point in C cannot be the aggregate $\underline{q}$ of the coordinates of all the particles in *P* unless C is also *flat*. Let us suppose, for the moment, that this is so.

Now let us introduce a *curved* Riemannian space C' which has the same dimensionality as C and like C is continuous. Suppose that, unlike *C*, the fundamental tensor of C' is comprised of free functions of *curvilinear coordinates* $\underline{x}$. This tensor can be equal to the fundamental tensor of C only at a point P' in C' and at the corresponding point P in C. Likewise we can satisfy $\underline{x} \equiv \underline{q}$ only at those points. If we demand, in addition, that $\underline{x} \equiv \underline{q}$ and $g'^{\mu\nu}(\underline{x}) = g^{\mu\nu}(\underline{q})$ are satisfied in the *neighbourhoods* of P and P' then we have to choose the $\underline{x}$ as *Cartesian geodesics with pole* P'. Given both these circumstances the flat space C can be described as *tangential* to the curved space C' at the points P in C and P' in C'.

The TE, and hence the GTE, are valid anywhere in a flat space *C*; but, in flat *C*, the GTE has no content. If, however, the GTE is expressed in terms of the $\underline{x}$ and the fundamental tensor of C' then it will be valid in the neighbourhood of P' in C' (providing that the $\underline{x}$ are chosen as Cartesian geodesics with pole P'). This is one of the methods of bringing the flat space of conventional QM to be consistent with the curved space of GR. The two spaces are consistent only in the

neighbourhoods of the points P in C and P' in C' ; but QM applies to the *physically small*.

It can be proved that a Riemannian space cannot have curvature unless its dimension is greater than three. So we can ask the question: What tensor equation, defined in a space Riemannian C' of dimension greater than three, reduces to the GTE when the coordinates in C' are Cartesian geodesics pole P'? The answer is the Kilmister equation [7], [8]. Because tensor equations are true in any coordinate system we may use the Kilmister equation, expressed in any convenient coordinates, to examine the local consequences of the GTE (expressed in Cartesian geodesics) holding in the neighbourhood of every point in a curved Riemannian C'.

It should be noted that neither the classical TE nor the classical GTE are tensor equations. So, when these equations are stated as being true in the neighbourhood of a point, particular attention should be paid to the coordinates and the metric that have been assumed.

We now, for the most part, drop the primes and recognise that Riemannian C can be curved providing that we use *curvilinear* coordinates $\underline{x}$ instead of the *flat* coordinates $\underline{q}$. As stated above there is a theorem which states that if C is to be curved, being Riemannian, that it must have a dimension in excess of three. Note that it is hypothesised, but not proven, that, in order to derive the form of the CM Hamiltonian, we do not need to consider any of the constraints above level two.

This paper is concerned, primarily, with the Kilmister equation. This is a classical equation and therefore applies, if it applies at all, to aspects of the cosmos which can be explained by non-quantum methods. It is an ODE of fourth order; and it is satisfied by solutions of the customary classical equations of gravity which are of second order. Therefore it has extra terms in its solutions. These extra terms must be appreciable only at cosmological distances; otherwise they would not have been missed. They are thought, at first sight, to be relevant to [3], [4] and to [5]; at any rate they must produce extra physics.

The modern picture is that the Universe is appreciably *flat*; this result is based on the statistics of the deviations from uniformity of the microwave background (roughly one in 100000). The Universe, theoretically, became transparent to radiation only about 400000 years after the Big Bang. So the truth (if it is true) of the transparency dates from that epoch.

Further at most 1/6 of the matter, *sensed by gravity*, is accounted for by that which we observe with telescopes; this is also the matter which is, roughly, accounted for by *particle physics*. A total of at least 5/6 of the matter, sensed by gravity, is *dark matter* which is *not* accounted for by the present Standard Model; this is hypothesised to be a mixture of unknown particles (which are not part of the Standard Model), invisible planets and gas and dark stars (if any). The matter sensed by gravity is only 30% of the total required to produce closure; the Universe is expanding under the influence of *dark energy*. This is variously explained by Einstein's *cosmological constant* producing *vacuum energy*, by *wimps* or by *quintessence*. We stick with vacuum energy. This paper is concerned with the Kilmister equation and with explanations of [4] and [5] although not [3].

1. The Classical Space C Is Riemannian Provided That The Space P Is Riemannian

If the first constraint, in the hierarchy of constraints, holds then can be proved to require that the operator Hamiltonian is quadratic in the P_u (the Einstein summation convention is in force); see Appendix A . The proof requires the assumption that the spectra of all the coordinate and momentum operators are continuous.

(1.1)

$$H \equiv \mathsf{K}\left\{G^{uv}, P_u P_v\right\} + \left\{F^j, P_j\right\} + V; \quad Q^j \equiv q^j I; \quad G^{uv}(\underline{Q}) \equiv g^{uv}(\underline{q})I = G^{vu}(\underline{Q});$$

$$i\hbar I\delta_{jk} \equiv Q^j P_k - P_k Q^j = -i\hbar\left(q^j \frac{\partial}{\partial q^k} - \frac{\partial}{\partial q^k} q^j\right); \quad P_j P_k \equiv P_k P_j; \quad Q_j Q_k \equiv Q_k Q_j;$$

$$F^j(\underline{Q}) \equiv f^j(\underline{q})I; \quad V(\underline{Q}) \equiv v(\underline{q})I; \quad u, v, j = 1, 2, \ldots n_c \equiv n_p n_d;$$

$$\{A, B\} \equiv \frac{1}{2}(AB + BA); \quad \left\{A^{uv}, B_{uv}\right\} \equiv \frac{1}{2}\left(A^{uv}B_{uv} + B_{uv}A^{uv}\right)$$

Here the position representation is used; $\underline{q}$ denotes the aggregate of Cartesian coordinates of the particles in flat P; capital letters are used for operators (thus A is the Hermitian operator corresponding to the real observable a); the functions $g^{uv}(\underline{q}), f^j(\underline{q}), v(\underline{q})$ are free; K is a constant scalar with physical dimensions (mass)$^{-1}$ in order that H has the physical dimensions of energy and the G^{uv} have none.

The classical approximation (all operators commute) to (1.1) is

(1.2) $\quad h \equiv \mathsf{K} g^{uv} p_u p_v + f^j p_j + v$ Scalar See Appendix A

This we simplify because here we are only interested in Einstein gravitation

(1.3) $\quad h \equiv \mathsf{K} g^{uv} p_u p_v; \quad f^j = 0; \quad v = 0$

So the *classical* equations of motion, defined in *flat C* and *P*, are (since Hamilton's equations are valid in CM)

(1.4a) $\quad \dot{q}^j = \frac{\partial h}{\partial p_j}; \quad \dot{p}_k = -\frac{\partial h}{\partial q^k}; \quad \dot{a} \equiv \frac{da}{ds}$

where s is the time variable. Eliminating the p_k between (1.3) and (1.4a)

(1.4b) $$\frac{d^2q^j}{ds^2}+\Gamma^j_{kl}\frac{dq^k}{ds}\frac{dq^l}{ds}=0;\quad \Gamma^l_{ij}\equiv\frac{1}{2}g^{lk}\left(g_{ik,j}+g_{jk,i}-g_{ij,k}\right)$$

the equation of a geodesic in a Riemannian space whose interval ds is defined by

(1.5) $$ds^2\equiv g_{uv}dq^u dq^v$$

So classical C is Riemannian if P is Riemannian; but both C and P are flat.

The flat space C is tangential to a curved space C' at P in C and at P' in C' if

(1.6) $$\underline{x}\equiv\underline{q};\quad g'_{uv}(\underline{x})=g_{uv}(\underline{q})\ \text{ at pole }\ \underline{x}$$

where the $\underline{x}$ are Cartesian geodesics pole P' and $g_{uv}(\underline{q})$, $g'_{uv}(\underline{x})$ are the fundamental tensors of C and C', respectively. Results (1.4b/1.5) are classical tensor equations and therefore true in any coordinate system and any pole P'. Therefore classical C' is also Riemannian.

2. The Classical Metrics Of P And C For Weak Gravity

We assume that the spaces P and C are continuous and Riemannian; indeed we have shown above that C is necessarily Riemannian if P and C are continuous and P is Riemannian. In general P has the dimension $n_d>2$ and is flat. The space C has the dimension $n_c\equiv n_p n_d$, where n_p is the number of particles. But, for the present, we discuss simpler scenarios: In these scenarios there is but one particle $n_p=1$; P has either the dimension $n_d=3$ and a Euclidean metric with Cartesian coordinates

(2.1) $$ds^2=ds_0^2;\quad ds_0^2\equiv dx^2+dy^2+dz^2$$

or the dimension $n_d=4$ and a Minkowski metric

(2.2) $ds^2 = c^2 d\tau^2 - ds_0^2; \quad |ds_0^2| << |cd\tau|$

Here τ is coordinate time and c is the speed of light. The inequality is required for Newtonian methods to be valid.

We take curvature of C' to be a symptom of gravity. As stated above it can be proved that to have curvature and be Riemannian C' must have dimensionality of at least $n_c = 4$. Since it has been assumed that $n_p = 1$ then $n_d > 3$. In consequence we assume that $n_c = 4$. If, in addition, there is but one time-like coordinate (per particle) then the other three must be space-like.

We suppose that, if P has the Minkowski metric then, given a single particle, the metric of C' is the weak gravity perturbation of Minkowski ($n_p = 1, n_c = 4$)

(2.3) $ds^2 = (1+2U)c^2 d\tau^2 + (-1+2U)(ds_0)^2; \quad |U| << 1; \quad |ds_0| << |cd\tau|$

where x, y, z, τ are *quasi-Cartesian* coordinates and time, and U is a dimensionless function of the spatial coordinates x, y, z only

(2.4) $U \equiv U(x, y, z)$

It is taken to be an invariant [9].

The metric (2.3) has a small curvature (determined by the second derivatives of U); so (2.3) is sufficient to describe a weak gravitational field from the point of view of GR [9]. We assume that U increases without limit as a particle in P is approached; that is (2.3) is valid except in a closed neighborhood that surrounds the particle. We restrict U so that it does not depend upon time because this ensures that the force, defined by

(2.5a) $\mathbf{F} \propto -\nabla U; \quad \nabla \equiv \mathbf{i}\frac{\partial}{\partial x} + \mathbf{j}\frac{\partial}{\partial y} + \mathbf{k}\frac{\partial}{\partial z}; \quad \nabla^2 \equiv \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$

(in Cartesian vector notation) is *conservative*

(2.5b) $\nabla \times \mathbf{F} = \mathbf{0}$ follows from (2.5a)

It then turns out that U is approximately proportional to the Newtonian potential [9,10].

In conformity with notation used elsewhere

(2.6) $x^1 \equiv x; \quad x^2 \equiv y; \quad x^3 \equiv z; \quad x^4 \equiv c\tau$; in C

So the metric (2.3) becomes

$$\begin{aligned} & ds^2 = (1+2U)(dx^4)^2 + (-1+2U)ds_0^2; \\ & (ds_0)^2 \equiv \sum_{J=1}^{3}\left(dx^J\right)^2; \quad |U| << 1; \quad |ds_0| << |dx^4| \end{aligned} \quad ; \text{ in } C \tag{2.7}$$

This metric is the link between Newtonian mechanics and GR. As remarked above this link applies providing that the gravity is weak ($|U| << 1$) and the speed of matter is small compared with c ($|ds_0| << |dx^4|$). We emphasise that, in this discussion, the only forces on test particles are gravitational. Result (2.7) must be regarded, from its derivation, as *classical*. NB In the metric (2.7) U is approximately proportional to the Newtonian potential and has the reverse sign to [9]. In the notation of [9]

(2.8) $\Omega \cong -c^2 U$ See [9] p. 101 et seq.

3. Motion Of A Test Particle In Weak Gravity- The Relation Between U And The Newtonian Potential

Suppose that an infinitesimal test particle is acted on by a scalar gravitational field potential $v(\underline{q}')$. Then the Newtonian Hamiltonian operator for the particle is, in the position representation,

(3.1)

$$H=\sum_{J=1}^{3}\frac{P_J'^2}{2\delta m}+V(\underline{Q}');\quad \underline{Q}'=\underline{q}'I;\quad V(\underline{Q}')=v(\underline{q}')I;$$
$$i\hbar I\delta_{JK}=Q'^J P_K'-P_K'Q'^J;\quad P_J'P_K'=P_K'P_J';\quad Q'^J Q'^K=Q'^K Q'^J;\quad J,K=1,2,3$$

where δm is the *inertial mass* of the test particle and $\underline{(.)}$ denotes aggregate (of coordinates etc.). But the coordinates and momenta (Cartesian and flat) are not necessarily the same as the coordinates and momenta $\underline{Q},\underline{P}$ referred to above; hence the primes.

Comparing (3.1) with the general case (1.1) (allowed by satisfaction of the first constraint), we see that

(3.2) $G_K^J(\underline{q}')=\dfrac{\delta_K^J I}{2\delta m\mathsf{K}};\quad f^J(\underline{q}')=0;\quad J,K=1,2,3;$ position representation

Thus the space in which H (see (3.1)) is defined has three dimensions and is truly Euclidean. Further, the spaces P and C are identical if C is flat and $n_p=1$. The classical Newtonian expression for the acceleration vector is

(3.3) $$\frac{\dot{p}_J}{\delta m}=-\frac{1}{\delta m}\frac{\partial h}{\partial q'^J}=-\frac{1}{\delta m}\frac{\partial v(\underline{q})}{\partial q'^J}$$

So infinitesimal test particles will be subject to this acceleration.

By contrast, in GR, Einstein asserts that an infinitesimal test particle moves on a geodesic in a Riemannian space [6]

(3.4)
$$\frac{d^2x^j}{ds^2}+\Gamma_{kl}^j\frac{dx^k}{ds}\frac{dx^l}{ds}=0;\quad \Gamma_{ij}^l\equiv\frac{1}{2}g^{lk}\left(g_{ik,j}+g_{jk,i}-g_{ij,k}\right);\quad (.)_{,k}\equiv\frac{\partial(.)}{\partial x^k}$$

Now, referring to (2.7), a link with Newtonian theory is the condition

$$(3.5) \qquad ds \approx \left|dx^4\right|; \quad \left|dx^4\right| >> \left|ds_0\right|; \quad cdt \approx ds_0$$

where t is the Newtonian time variable. With this (3.4) becomes

$$(3.6a) \qquad \Gamma^J_{44} = U_{,x^J} \qquad \text{See (2.7)}$$

$$(3.6b) \qquad \frac{d^2x^J}{dt^2} \approx -c^2\Gamma^J_{44} \approx -c^2U_{,x^J}; \quad J = 1,2,3; \quad \left|U\right| << 1$$

giving, approximately, Cartesian components of classical acceleration in any weak gravitational field. But this is an expression for the components of acceleration referred to coordinates that differ from those used at (3.3). This is made obvious by comparison of the metrics (2.7) and (3.2); the latter is constant and *exactly* 3-Euclidean; and the former is variable, slightly curved and *approximately* 4-Minkowskian. The latter may be made more like the former by assuming that *P* is flat but four dimensional

$$(3.7) \qquad \begin{aligned} ds^2 &= \varepsilon\left[-(dq'^4)^2 + \frac{1}{2\delta m\mathbf{K}} ds_0'^2\right]; \\ (ds_0')^2 &\equiv \sum_{J=1}^{3}\left(dq'^J\right)^2; \quad \left|U\right| << 1; \quad \left|ds_0'^2\right| << \left|dq'^4\right| \end{aligned}$$

with $\varepsilon \equiv \pm 1$ as the indicator. We can then transform (2.7) into (3.7) by only changing coordinates and, if necessary, ε.

How does v, the classical Newtonian potential, compare with U ?; see (2.7). In comparing these two variables we are contrasting a truly Newtonian case with an approximating weak field case in GR. We can make the comparison by comparing the two acceleration vectors (3.3) and (3.6b) referred to the same coordinate system.

(3.8) $$-\frac{1}{\delta m}\frac{\partial v}{\partial q'^J} \approx -c^2 \frac{\partial U}{\partial q'^J} \Rightarrow v \approx \delta m c^2 U; \quad |U| << 1$$ See (3.3/3.6b)

The last step relies on the potential being unspecified within a constant.

4. Einstein's Equations- $n_c = 4; \quad n_p = 1$

GR is essentially a geometrical theory based on Riemannian geometry. It brings in CM by noting that a particular geometrical tensor in GR has the same zero tensor divergence, as the energy-momentum tensor does in CM, in the expressing the laws of mechanics. The two tensors must be proportional in order to agree with the Newton/ Poisson theory for week gravity. GR treats measurements of time on the same footing as measurements of space. This is clearly wrong, in some sense, because we can place ourselves anywhere, in space, by an act of will; but we cannot do the same for time. Measurements of time necessarily increase; and it sweeps us along with it. Macroscopically it is something to do with the relentless increase in entropy treated by classical thermodynamics; microscopically it is something to do with QM. Yet the Einstein theory of GR has withstood all the experimental and observational tests for more than 100 years.

In GR the Einstein law of gravity for empty space (i.e., between particles) is the tensor equation

(4.1a) $R_{ab} = 0; \quad a,b = 1,2,3,4;$ R_{ab} is the Ricci tensor [6]

Alternatively the law can be expressed as

(4.1b) $G^u_v = 0; \quad G^u_v \equiv R^u_v - \frac{1}{2} R \delta^u_v;$ G^u_v is the Einstein tensor

Only when the coordinates are those of a particle does (4.1a) break down; then the RHS is a species of delta function. That is the *matter* is concentrated in the *particle*; and the curvature is *infinite* at the particle. More generally, when some of the particles are distributed

evenly and are so numerous that they can be represented by a fluid, the Einstein law is given by the equation [6]

(4.2a) $$G^a_b + \chi T^a_b = 0; \quad G^u_v \equiv R^u_v - \tfrac{1}{2} R\delta^u_v \Rightarrow G = -R \Rightarrow R^u_v = G^u_v - \tfrac{1}{2} G\delta^u_v$$

where

(4.2b)
$$\chi \equiv 8\pi \mathsf{G}/c^4 = 2.0761\times 10^{-43} m^{-1} kg^{-1} s^2;$$
$$c = 2.99792458\times 10^8 ms^{-1};$$
$$\mathsf{G} = 6.672(59)\times 10^{-11} Nm^2 kg^{-2} \equiv 6.672(59)\times 10^{-11}\ m^3 kg^{-1} s^{-2}$$

and G is Newton's constant with T^a_b as the matter-energy-momentum-stress tensor of the fluid. Another law of gravity, that Einstein suggested but later rejected (for his purposes), is

(4.2c) $$G^a_b + \chi T^a_b + \Lambda \delta^a_b = 0 \quad \text{Tensor equation}$$

where, to make (4.2c) a tensor equation, Λ is a universal constant. Note that, given (4.2c), and given a model universe empty of ordinary matter and energy

(4.3a) $$T^a_b = 0 \Rightarrow G^a_b = -\Lambda \delta^a_b \Rightarrow R^a_b = \Lambda \delta^a_b \quad \text{Einstein space}$$

More generally, when the Riemannian space is four dimensional,

(4.3b)
$$R^u_v = G^u_v - \tfrac{1}{2} G\delta^u_v = -\chi T^u_v - \Lambda \delta^u_v + \tfrac{1}{2}\left(\chi T + 4\Lambda\right)\delta^u_v = -\chi\left(T^u_v - \tfrac{1}{2} T\right) + \Lambda \delta^u_v$$

The equation (4.2c) is subject to the identity

(4.4) $$G^a_{b;a} = 0 \Rightarrow T^a_{b;a} = 0 \quad \text{Requires } \Lambda \text{ to be constant}$$

The last equation at (4.4) is the tensor expression for the classical mass-energy-momentum conservation laws in CM [6]. They require, in order that (4.2c) should be a tensor equation, that Λ should be constant.

This section raises the question of sign conventions. In Section 2 we have supposed that the signature of the Minkowski metric is $-1,-1,-1,+1$ making the interval ds real for speeds less than c. Both Eddington [9] and Spain [6] observe this convention; so, in their work, Einstein's equation (4.2c) is written as above. But in more modern work the signature of the Minkowski metric is assumed to be $1,1,1,-1$ so that, for speeds less than c, the interval ds is imaginary. The tensor T_b^a, although still real, then changes sign and Einstein's equation is written

$$G_b^a + \Lambda\delta_b^a = \chi T_b^a \tag{4.5}$$

We assume the Eddington/ Spain convention.

5. The Theta Equation, The Gravitational Theta Equation And Kilmister's Equation

If the first *two* constraints hold, in the hierarchy, then we can deduce the Theta Equation valid in QM [19]

$$\begin{aligned} &G^{vj}(\underline{Q})\left(G^{uk}(\underline{Q})\Theta_{,jku}(\underline{Q})\right)_{,v} = 0; \quad \text{irrespective of } F^j \text{ and } V; \\ &\Theta_{,j} \equiv \frac{i}{\hbar}\left(P_j\Theta - \Theta P_j\right) \quad \text{See Appendix B} \end{aligned} \tag{5.1a}$$

In the position representation this reduces to the PDE (irrespective of f^j and v)

$$g^{vj}\left(g^{uk}\theta_{,jku}\right)_{,v} = 0; \quad j,k,u,v = 1,2,...n_c; \quad \theta_{,jku} \equiv \theta_{,j,k,u}; \tag{5.1b}$$

',' denotes partial differentiation; we *choose* Cartesian coordinates. In the same representation and with the same coordinates, if we choose,

$$\theta \equiv g^{lm} \ \forall\, l,m \tag{5.2}$$

substituted we get the Gravitational Theta Equation (GTE)

(5.3) $g^{vj}\left(g^{uk}g^{lm}{}_{,jku}\right)_{,v}=0;$ GTE; See (1.1/1.3)

It is supposed that the GTE is valid in *quasi-Cartesian geodesic coordinates* $\underline{x}$ in the neighbourhood of a pole P' in C' if, at the corresponding point P in C coordinates $\underline{q}$, the space *C* is tangential to the space C'. That is

(5.4) $\underline{x}=\underline{q};\quad g'_{uv}(\underline{x})=g_{uv}(\underline{q});$ at P' in C' and P in *C*

where $g'_{uv}(\underline{x})$ is the fundamental tensor of C' and $g_{uv}(\underline{q})$ that of *C*.

In theory the GTE is valid in QM. But what kind of space is *C*? The argument leading to (1.4b) shows that it is Riemannian; but that argument relies on (1.3) which is a *classical* (that is macroscopic) equation. So we can prove that the GTE applies only to a classical Riemannian space and hence to CM.

The Kilmister Equation [7], [8], derived by the late Clive Kilmister (2006) from the GTE, is a *classical* tensor equation defined on a Riemannian manifold C'

(5.5) $K_{ab}\equiv g^{ef}(R_{ab;ef}+\frac{2}{3}R_{ae}R_{fb})=0;\qquad a,b,e,f=1,2,...n_c=4;$ see (4.2a)

where ';' denotes covariant differentiation. It is otherwise known as the K equation. The K equation reduces to the GTE at the pole of Cartesian geodesics; and, because of the choice of those coordinates, approximates the GTE in the neighbourhood of the pole. Ostensibly it applies to but a single particle. When $n_c=4, n_p=1$ it is called the relativistic K equation (RKE) and the GTE should be called the relativistic gravitational equation (RGTE); we shall not follow this usage, however, because the meaning should be clear by the context. Note that, given (4.3a) and (5.5),

(5.6a) $R^a_b = \Lambda g^a_b \Rightarrow R_{ab;ef} = 0; \quad \frac{2}{3} g^{ef} R_{ae} R_{fb} = \frac{2}{3} \Lambda^2 g_{ab} \Rightarrow \Lambda = 0$ see (5.1)

That is, when the model universe is truly empty (of all ordinary matter and of vacuum energy), the RKE requires that $\Lambda = 0$. More generally the RKE determines the elements of the Ricci tensor R^u_v and hence the fundamental tensor g_{uv} and the gravitational field. It also determines, via (4.3b), the matter tensor T^u_v subject to symmetries and boundary conditions. The K equation is an alternative law of gravity. The ordinary relativistic law of gravity (see (4.1a)) satisfies it; but is of fourth order, as opposed to second order, and therefore has more solutions.

6. Newtonian Approximations To Tensors Given Cartesian Coordinates

There follow a number of approximations of tensor quantities using Cartesian coordinates and the Newtonian scheme with the particle at the origin. Physically it is unclear whether these *can* apply only in the solar system, in our galaxy (and by extension to models of galaxies in general) or to the Universe at large.

Given the metric (2.7) it can be shown that [9]

(6.1) $R_{ab} \approx -\Delta_{ab} \nabla^2 U; \quad |U| << 1; \quad a,b = 1,2,3,4; \quad n_p = 1; \quad n_c = 4$

where the Kronecker delta Δ_{ab} takes its usual meaning. So that the law of (weak) gravity for empty space is, according to the Newtonian approximation,

(6.2a) $\nabla^2 U = 0; \quad |U| << 1$ See (4.1a)

There is another law of gravity, namely, the tensor equation

(6.2b) $R^a_b = \Lambda \delta^a_b; \quad \Lambda \neq 0;$

$\delta^a_b = 1$ if $a = b, \delta^a_b = 0$ otherwise see (4.3a)

where Λ is a (small) universal constant. Since, given the metric (2.7), and the fact that U does not depend on x^4

(6.2c)

$$R^a_b = R_{rb}g^{ra}, g^{ra} \approx -1, g^{44} \approx 1 \text{ according as } r = a < 4,\ g^{ra} = 0 \text{ otherwise}$$
$$\Rightarrow \nabla^2 U \approx \Lambda; \quad |U| << 1; \quad a,b,r = 1,2,3,4; \quad \text{See (6.1)}$$

Further

(6.3a)
$$G^i_j \equiv R^i_j - \frac{1}{2}R\delta^i_j; \quad \delta^i_j \equiv g^i_j; \quad R \equiv R^a_a; \quad \text{Definitions}$$
$$G^a_b \approx -\Lambda\delta^a_b; \quad |U| << 1; \quad \text{Newtonian approximation}$$

The identity (4.4) is therefore satisfied approximately because Λ is constant or zero. Given the metric (2.7),

(6.3b)
$$\Gamma^l_{ij,s} \equiv \left(\frac{1}{2}g^{lk}\left(g_{ik,j} + g_{jk,i} - g_{ij,k}\right)\right)_{,s}; \quad \text{Definitition}$$
$$\approx -\Delta_{li}\Delta_{lj}U_{,l,s}; \quad |U| << 1 \quad \text{Newtonian approximation}$$

Similarly, because R^a_b is small or zero,

(6.4)
$$K^a_b \approx g^{ef}R^a_{b;ef} \quad \text{Small element Ricci tensor approximation}$$
$$= g^{ef}\left(R^a_{b,e,f} + \Gamma^a_{re,f}R^r_b - \Gamma^r_{be,f}R^a_r\right) \quad \text{Assuming geodesic coordinates}$$
$$\approx g^{ef}R^a_{b,e,f} \approx -\nabla^2(\nabla^2 U) \quad \text{Newtonian approximation} \quad \text{See (6.2c)}$$

The second line of (6.4) is an expansion assuming the coordinates are geodesic; in fact the third line (see (2.5a)) requires them to be Cartesian Goedesics. So, the Newtonian approximation to the relativistic K equation, is

(6.5) $$-K^a_b \approx \nabla^2(\nabla^2 U) = 0; \quad |U| << 1 \quad \text{See (6.1/6.2c)}$$

There is another way of deducing (6.5): Simply substitute from (2.7), for the g_{uv}, into the GTE and approximate.

In all of the approximations above we have neglected second and higher order products and powers of U and its derivatives.

When the particles are numerous and continuously distributed the physical conditions that attach to Newton's theory require that

(6.6) $T_4^4 \approx c^2\rho; \quad T_s^r \to 0 \quad \text{either } r,s \neq 4,4; \quad |U| << 1$

where ρ is the matter/ energy density and the other elements of T_b^a effectively vanish. It follows that the Einstein law of gravity (4.2a), with these conditions, reduces to

(6.7a) $\nabla^2 U \approx -c^2 \frac{\chi}{2}\rho; \quad |U| << 1; \quad \chi \equiv 8\pi G / c^4$

According to (6.2c) space is suffused with a matter/ energy ρ_{00} density given by

(6.7b) $-c^2 \frac{\chi}{2}\rho_{00} = \Lambda$ There is evidence that Λ is negative

With the definitions

(6.7c) $\Omega \equiv -c^2 U$ See (2.8)

we get

(6.8) $\nabla^2 \Omega \approx 4\pi G\rho; \quad |\Omega| << c^2$ Poisson's Equation as an approximation

This is the Poisson's Equation. Here Ω is the Newtonian potential (energy per unit mass) of an infinitesimal test particle. The requirement that the Einstein law should reduce to the Newton-Poisson law

determines the value of the constant χ ; see (4.2b/6.7b). We have used (4.2a) as the Poisson expression of Newton's law, rather than (4.2c), because in Newton's theory $\Lambda = 0$.

The proofs of Poisson's equation (6.8) given in [10] and above (depending on GR) allow us to generalise the law of gravity. So, under the Newtonian scheme, (6.7a) can be regarded as fundamental. Whatever U and whatever the law of gravity (providing that gravity is a central conservative force)

$$\rho \equiv -\frac{2\nabla^2 U}{c^2 \chi}; \quad |U| << 1 \tag{6.9}$$

might be regarded as definition of mass density. This is not a satisfactory definition however. If W is a solution of

$$\nabla^2 W = 0 \tag{6.10a}$$

then it follows from (6.7a) that

$$c^2 \frac{\chi}{2} \rho = -\nabla^2 (U + W) \tag{6.10b}$$

In other words, according to (6.7a), the Newtonian potential does not uniquely determine the density; *but with any other law for which* $\nabla^2 W \neq 0$ it does. Moreover note that (6.5/7a) requires

$$\nabla^2(\nabla^2 U) = 0 \Rightarrow \nabla^2 \rho = 0; \quad |U| << 1 \Rightarrow |\Omega| << c^2 \tag{6.11}$$

By this result (6.5) requires to be a solution of Laplace's equation; this is debatable.

As we have seen in C, whether C is flat or slightly curved, Newtonian theory is a valid approximation *almost everywhere*. The exception is as follows: In P the neighbourhood of a particle corresponding to a small region in C ; we denote the aggregate of such re-

gions by *N*; so *N* is a neighbourhood of *X*. Near *X*, in *N*, the gravitational field is assumed to rise, without limit and therefore the curvature of C rises without limit. Thus, in general, Newtonian theory is not valid in *N*. The result (6.5) is, in the Newtonian approximation, an alternative to Newton's law.

When the particles are numerous and the distribution of ordinary matter-energy is sufficiently uniform the particles can be replaced by a continuous fluid in *P*. The dimension of C is then, strictly, *infinite*. A metric for *C* is valid in the neighbourhood of X but outside *N*, as heretofore, and must be almost flat in order that we may apply the Newtonian method.

According to GR, as we have seen, weak gravity can be characterised by a single scalar potential U ; see (2.7). Further, U can be due to many particles. If *C* is truly flat ($U = 0$) and the metric of P is

$$\begin{aligned} &ds^2 = (dq^4)^2 - ds_0^2; \\ &(ds_0)^2 \equiv \sum_{J=1}^{3} \left(dq^J\right)^2; \quad \left|ds_0^2\right| << \left|dq^4\right| \end{aligned} \tag{6.12}$$

(however many particles P contains) then the coordinates of *X*, used in *C*, can be chosen as an infinite repetition of those used in *P*

(6.13a)

$$\begin{aligned} &\underline{q} = \{....q^1_{\alpha-1}, q^2_{\alpha-1}, q^3_{\alpha-1}, q^4_{\alpha-1}, ...q^1_{\alpha}, q^2_{\alpha}, q^3_{\alpha}, q^4_{\alpha}, q^1_{\alpha+1}, q^2_{\alpha+1}, q^3_{\alpha+1}, q^4_{\alpha+1}....\}; \\ &\alpha = 1, 2, ...n_p \end{aligned}$$

The metric of flat *C* , using these coordinates, is therefore

$$\begin{aligned} &ds^2 = \sum_{\alpha=1}^{n_p} \left[(dq^4_{\alpha})^2 - (ds^{\alpha}_0)^2\right]; \\ &(ds^{\alpha}_0)^2 \equiv \sum_{j=1}^{3} \left(dq^J_{\alpha}\right)^2; \quad \left|ds^{\alpha}_0\right| << \left|dq^4_{\alpha}\right|; \quad n_p \to \infty; \quad 1 \le \alpha \le n_p \end{aligned} \tag{6.13b}$$

Dividing the metric (6.13b) by n_p

$$
\begin{aligned}
&d\bar{s}^2 = (d\bar{q}^4)^2 - (d\bar{s}_0)^2; \\
&d\bar{s}^2 \equiv \frac{ds^2}{n_p}; \quad (d\bar{s}_0)^2 \equiv \frac{1}{n_p}\sum_{\alpha=1}^{n_p}(ds_0^\alpha)^2; \\
&(d\bar{q}^4)^2 \equiv \frac{1}{n_p}\sum_{\alpha=1}^{n_p}(dq_\alpha^4)^2; \quad (d\bar{q}^J)^2 \equiv \frac{1}{n_p}\sum_{\alpha=1}^{n_p}(dq_\alpha^J)^2; \quad J=1,2,3
\end{aligned}
\tag{6.14}
$$

The first line of (6.14) is the metric of a Minkowskian space with coordinates $\bar{q}^1, \bar{q}^2, \bar{q}^3, \bar{q}^4$; and quantities denoted with a bar over the top are RMS values of the corresponding quantities for each of the particles in *P*.

In slightly curved C the metric appropriate to the α^{th} particle, which in this case can be treated as an infinitesimal test particle, is

$$
\begin{aligned}
&(ds^\alpha)^2 = (1+2U)(dx_\alpha^4)^2 + (-1+2U)(ds_0^\alpha)^2; \\
&(ds_0^\alpha)^2 \equiv \sum_{j=1}^{3}\left(dx_\alpha^J\right)^2; \quad |U| << 1; \quad \left|ds_0^\alpha\right| << \left|dx_\alpha^4\right|
\end{aligned}
\tag{6.15}
$$

where the dimensionless potential U is a function of *all* the particle coordinates. There are so many particles that, other than in the neighborhood of any particle (where the Newtonian formulae are invalid), U is approximately independent of the existence of any one particle. So, summing (6.16) and dividing by n_p,

$$
\begin{aligned}
&d\bar{s}^2 = (1+2U)(d\bar{x}^4)^2 + (-1+2U)d\bar{s}_0^2; \\
&(d\bar{s}_0)^2 \equiv \sum_{J=1}^{3}\left(d\bar{x}^J\right)^2; \quad |U| << 1; \quad \left|d\bar{s}_0^2\right| << \left|d\bar{x}^4\right|
\end{aligned}
\tag{6.16}
$$

This is (approximately) the metric of a slightly curved Minkowskian space; and quantities denoted with a bar over the top are RMS values of the corresponding quantities for each of the particles in P with

$$
\bar{x}^J \approx \bar{q}^J; \quad \bar{x}^4 \approx \bar{q}^4 \tag{6.17}
$$

It follows that when the system is composed of a fluid only we may use $n_c = 4$ although the dimension of C is fact infinite.

7. SS Solution Of The Newtonian Approximation To The K Equation

Given $n_p = 1, n_c = 4$ the Newtonian approximation to the relativistic K equation is (6.5). Suppose there is a particle at the origin. Outside *N*, the geometry in C, although subject to slight curvature, must *approximate* the geometry in P . We suppose, for the present discussion, that the origins of C and P coincide in the particle; and we presume that its gravitational field, as sensed by test particles, is spherically symmetric (SS) and so, defined in *C*,

$$U \equiv U(r); \quad r^2 \equiv \sum_{j=1}^{3} (x^j)^2; \quad n_c = 4, n_p = 1; \quad |U| << 1 \tag{7.1}$$

The Laplacian then reduces to

$$\nabla^2 \equiv \frac{d^2}{dr^2} + \frac{2}{r}\frac{d}{dr} \tag{7.2}$$

and (6.5) becomes

$$\left(\frac{d^2}{dr^2} + \frac{2}{r}\frac{d}{dr}\right)^2 U = 0; \quad |U| << 1 \tag{7.3}$$

A general SS solution of this ODE is

$$U = \frac{k_1}{r} + k_2 r^2 + k_3 r + k_4; \quad |U| << 1 \tag{7.4}$$

Maple 16

where k_1, k_2, k_3, k_4 are constants and r is the distance of a test particle from the origin of *C*. In order that the condition $|U| << 1$, attached to (6.5), can be satisfied we set

(7.5) $$k_4 = 0 \Rightarrow U = \frac{k_1}{r} + k_2 r^2 + k_3 r; \quad |U| << 1$$

We take it that (7.5) refers to the dimensionless potential of a test particle distant r from the origin. Because (6.5) is a Newtonian approximation (to the K equation) it is subject to the same strictures as appear at the beginning of Section 6.

The solution (7.5) is seen to include the usual Newtonian 'inverse square' contribution to the potential; but it includes also *extra terms*. To have escaped experiment these extra terms must be either very small or zero (in the solar system). We assume, in what follows, that the extra terms are *non-zero*; but that they are *small* except at huge (cosmological) distances. The extra terms may provide an approximate explanation of three recently observed and mysterious phenomena [3],[4] and [5].

One of these [4] cites very distant objects (type 1a supernova) that should, according to conventional ideas be slowing down, *as speeding up*; the radial acceleration is proportional to the distance. When Hubble found that all the galaxies where moving away from each other Einstein set his $\Lambda = 0$ because it was not needed; see (4.2c). But [4] makes use of a non-zero Λ to describe the extra acceleration. Now the Newtonian equation correspond to the Einsteinian equation (4.2a) for weak gravity; so the extra terms $k_2 r^2 + k_3 r$ must correspond to a non-zero Λ ; see the strictures, however, at the beginning of Section 6 .

If we define Λ' by the tensor gravity equation

(7.6a) $$R^a_b = \Lambda' \delta^a_b$$

for a model universe empty of ordinary matter where Λ' is a universal constant (see (4.3a)). This is a different law of gravity either from (4.1a) or the K equation (5.5). We get as the 'Newtonian' approximation, in Cartesians,

(7.6b) $\nabla^2 U \approx \Lambda'; \quad \left|U\right| << 1$ See (2.7/6.1/6.2c/6.3)

with an SS solution

(7.6c) $U = \dfrac{k_1'}{r} + k_2' + \dfrac{\Lambda' r^2}{4}; \quad \left|U\right| << 1$

where k_1', k_2' are constants of integration. In order to satisfy the condition attached to (7.6c) we put

(7.7a) $k_2' \to 0$

The terms in k_1' and $r^2\Lambda'$ must be small enough to satisfy the condition on U. Result (7.6c) is the 'Newtonian' dimensionless potential of a SS particle at the origin. If the system is of ordinary mass m'

(7.7b) $k_1' \equiv -\dfrac{\mathsf{G}m'}{c^2}$

So, comparing (7.6c) with (7.5) for *small* r,

(7.8a) $k_1 \approx k_1'$ extra terms neglected

Comparing (7.6c) with (7.5) for *large* r

(7.8b) $k_2 \approx \dfrac{\Lambda'}{4}; \quad \text{terms in } k_3, k_2', k_1' \text{ neglected}$

We have put the word 'Newtonian' in inverted commas '' because Newton's theory does not include the term Λ' in its gravitational law. According to (7.5) the radial acceleration of a test particle, for large r, is

(7.9)

$$-c^2U_{,r} = -c^2\left(-\frac{k_1}{r^2} + 2k_2r + k_3\right)\Bigg|_{r\to\infty} \approx -2c^2k_2r; \quad \left|U\right| << 1 \qquad \text{See}$$

(3.6b)

Since very distant objects have a positive acceleration, which is proportional to distance, (7.9) is positive [4]. This implies that k_2 is physically small and negative and hence Λ' is the same; see (7.8b). Note that the extra terms, in (7.5), have nothing to do with the slow drift of the perihelion of the planet Mercury; their dependence on range is wrong!

Result (7.5) applies to a single particle. If we accept (6.9), and regard Poisson's equation as fundamental, the mass/ energy density of the single particle may be defined as

(7.10a)
$$\begin{aligned}\rho(r) &= m\delta(r) + \frac{2}{c^2\chi}\nabla^2U \\ &= m\delta(r) + \frac{2}{c^2\chi}\left(\frac{d^2}{dr^2} + \frac{2}{r}\frac{d}{dr}\right)U = m\delta(r) + \frac{2}{c^2\chi}\left(6k_2 + \frac{2k_3}{r}\right)\end{aligned}$$

where $\delta(r)$ is the SS Dirac delta which has the properties

(7.10b) $\int_0^\infty 4\pi z^2\delta(z)dz = 1; \quad \delta(z) = 0$ except at $z = 0$ when it is infinite

Here $m\delta(r)$ is the particle density and equation (7.10a) is true only at

(7.11) $r = 0$

elsewhere it is

(7.12) $$\frac{2}{c^2\chi}\left(6k_2 + \frac{2k_3}{r}\right)$$

This term may well be negative.

If we accept (6.9) for the particle we must accept (see (6.11))

$$(7.13) \qquad \nabla^2\rho = \left(\frac{d^2}{dr^2} + \frac{2}{r}\frac{d}{dr}\right)\rho(r) = 0 \Rightarrow \rho(r) = \rho_1 + \frac{\rho_2}{r}$$

where ρ_1 and ρ_2 are constants of integration having the physical dimensions of density and density $\times$ length respectively. So, in order that this shall be consistent with (7.5),

$$(7.14) \qquad \rho_1 = \frac{12k_2}{c^2\chi}; \quad \rho_2 = \frac{4k_3}{c^2\chi}; \quad r \neq 0$$

Because $|U| << 1$, the solution (7.5) is not accurate, at a great distance from the origin, unless the space is almost flat. It follows that U and $\nabla^2 U$ must be, in some sense, 'small' in the region of validity.

8. Hypothesis- k_2 and k_3 Are Proportional To The Mass Of A Small Compact Portion Of Matter

Result (7.5) applies to a particle; but, because matter is made up of particles, we may sum over the particles in a *small compact* piece of matter, in the manner of Newton, providing that the fields linearly superpose and the radii are appreciably the same

(8.1)

$$U = \frac{k_1}{r} + k_2 r^2 + k_3 r \rightarrow U = \Sigma\left(\frac{\delta k_1}{r} + \delta k_2 r^2 + \delta k_3 r\right); \quad |U| << 1; \quad r \neq 0$$

where the Σ is over the particles of the small piece of matter. This implies that

(8.2a) $\quad k_1 \propto m, k_2 \propto m, k_3 \propto m$

where m is the mass of the small piece of matter. If we identify the first term in (8.1) as the ordinary Newtonian potential,

(8.2b) $k_1 \propto m$ in fact $k_1 = -\dfrac{m\mathsf{G}}{c^2}$

where m is the mass of the small, compact piece of matter. So our hypothesis means

(8.2c) $k_1 = \kappa_1 m; \quad k_2 = \kappa_2 m; \quad k_3 = \kappa_3 m; \quad \kappa_1 = -\dfrac{\mathsf{G}}{c^2};$ See (4.2b)

We conclude that for any small, compact piece of matter of mass dm the dimensionless potential dU is SS and, at radius r, is

(8.3) $$dU = dm\left(\frac{\kappa_1}{r} + \kappa_2 r^2 + \kappa_3 r\right)$$

We do not know the values of κ_2 and κ_3 but, evidently, they are independent of dm; we will assume that

(8.4) $\pm(\kappa_1 < 0, \kappa_3 > 0); \quad \left|\kappa_2 r^2\right| << \left|\kappa_3 r\right|; \quad \kappa_2, \kappa_3$ probably universal.

9. The Theory Behind (8.3) Applied To A Flat Galaxy [5]

It is customary to apply Newtonian theory to the detailed structure of galaxies save for the absolute centre where, for some or all galaxies, there is a massive black hole and where Einsteinian theory is appropriate. This is despite the facts that galaxies are of the order of $100000\,ly$ across and, in Newton's theory as opposed to Einstein's, time is universal. Probably Newton's theory works, when applied to galaxies, because the light transit time is so small compared to the age of the Universe.

By a 'flat galaxy' we mean a galaxy like our own. Judging by the observed luminosity the form includes spirals. It is a more or less

thin disc with a bright ball in the centre tailing off, in brightness, thickness and density, towards the periphery. It is postulated that it has an invisible halo, centred on the galaxy, of much greater radius than the luminous part; this halo taken to approximates an oblate spheroid and can be assumed to be rotating. According to modern theories 83% to 99% of the matter in the galaxy is in the halo. This is the so called Dark Matter [5]. Dark matter may consist of dust and gas, low luminosity stars, planets (free or otherwise), neutrinos and/ or other subatomic particles, in particular, those that are not part of the Standard Model; or we may have simply got the gravity law wrong or both. We shall examine the hypothesis about the gravity law, in detail, for the Newtonian scheme.

The model galaxy is a simplification of our galaxy. The model galaxy rotates (not necessarily as a solid body) about an axis through the centre of the galaxy and perpendicular to its plane. It could be approximated as gas of varying (radial) density. The individual stars could be 'atoms' of that gas; and there is also 'real' gas. In practice we concentrate on the *halo*. As been said at least 83% of the matter in the galaxy is theoretically in the halo; this is a cogent simplification.

Newton proved two theorems which mean that a spherical distribution of matter, for which the density only varies with radius, may treated as a point at the centre of the sphere [10]. The law of gravity can be various as long as it is central. When the test particle is *inside* the sphere the attraction is due to the matter at a *radius less than or equal to the test particle*; when the particle is *outside* the sphere the attraction is due to a point concentrated at the centre of the same mass as the sphere of matter. We approximate the halo as a perfect sphere for simplicity. This means that, although we make an error thereby, we can ignore the luminous part of the galaxy as long as we assign the mass of the whole galaxy to the halo; see Fig. 2.

Talking of the luminous part of the galaxy: Newton's gravitational law requires that two particles, the one much heavier than the other, move in an ellipse whose focus is the position of the heavier particle [10]. The simplest form of this orbit is a circle; and yet the simulations of Fig. 3 (that use the canonical Newton's law) have at least some of the stars travelling in spirals. This fact means that the

attraction is not sufficient to sustain closed orbits. The culprit is the density; it falls off with distance from the centre. This means that the Newtonian attraction falls off faster than that which is required for a closed orbit; *Bertrand's Theorem* [10] requires that the only attractions that produce closed orbits are those that vary as r or as r^{-2}. Given Newton's law the force is proportional to r when the test particle is *inside* the spherical distribution of matter and proportional to r^{-2} when the particle is *outside* the distribution.

We can, in theory, apply the argument leading to (8.3) to deduce the archetype velocity curve (speed of a test particle v as a function of its radius r). The density of the halo as a function of radius is, theoretically, required for this purpose. The article [10] calculates the connection between the parameters of a *rotating* ellipsoid of incompressible fluid. The article [19] calculates the connection between the parameters of a *stationary* spherical cloud of compressible gas. Both these calculations are complex. The halo is both rotating and a compressible gas; therefore to calculate its density is even more complex.

Our final argument, which depends on the two theorems proved by Newton, is much simpler; *but is only suggestive*. Keep in mind that, according to Newton, matter is transparent to gravitation. The result (8.3) refers to compact piece of matter of mass dm. If we make assumptions (8.4/8.2c), about κ_1, κ_2 and κ_3, we may integrate, approximately, over all the matter of the halo to produce (see (8.3/4)) the dimensionless potential at a point in the mid-plane of the galaxy distant r from the centre. There are two regimes for the dimensionless potential: one $U1(r)$ for which $r \geq a_h$ and one $U2(r)$ for which $r \leq a_h$. Here a_h is the radius of the periphery of the halo (at which the ordinary density becomes zero). We assume that the gravity is given by the Newtonian approximation to the K equation and that, in consequence, result (8.3) is satisfied:

$$U1(r) = \int_{\varepsilon}^{a_h}\int_{0}^{\pi}\int_{0}^{2\pi} g(z)\rho_h(x)x^2\sin(\theta)d\varphi d\theta dx + \text{a constant(1)};$$

(9.1a)
$$= 2\pi\int_{\varepsilon}^{a_h}\int_{0}^{\pi} g(z)\rho_h(x)x^2\sin(\theta)d\theta dx + \text{a constant(1)};$$
$$0 < \varepsilon << a_h \le r; \quad |U1| << 1$$

$$U2(r) = \int_{\varepsilon}^{r}\int_{0}^{\pi}\int_{0}^{2\pi} g(z)\rho_h(x)x^2\sin(\theta)d\varphi d\theta dx + \text{a constant(2)}$$

(9.1b)
$$= 2\pi\int_{\varepsilon}^{r}\int_{0}^{\pi} g(z)\rho_h(x)x^2\sin(\theta)d\theta dx + \text{a constant(2)}$$
$$0 < \varepsilon << r \le a_h; \quad |U2| << 1$$

(9.1c)

$$g(z) = \frac{\kappa_1}{z} + \kappa_2 z^2 + \kappa_3 z; \quad z^2 = r^2 + x^2 - 2rx\cos(\theta); \quad z > 0; \quad \kappa_1 = -\frac{G}{c^2},$$

where $\rho_h(x)$ is the ordinary density of the halo at radius x. The ordinary mass of the halo is

(9.2)
$$M_h = 2\pi\int_{\varepsilon}^{a_h}\int_{0}^{\pi}\rho_h(x)x^2\sin(\theta)d\theta dx = 4\pi\int_{\varepsilon}^{a_h}\rho_h(x)x^2dx$$

Now we approximate

(9.3a)
$$\rho_h(x) = \rho_h \quad \text{a constant}$$

(9.3b)
$$M_h = 4\pi\rho_h\left[\frac{x^3}{3}\right]_{\varepsilon}^{a_h} \approx \frac{4\pi a_h^3}{3}\rho_h$$

As a consequence we get (Maple 16)

(9.4a)

$$U1(r) \approx 2\pi\rho_h \int_{\varepsilon}^{a_h}\int_{0}^{\pi} g(z)x^2 \sin(\theta)d\theta dx + \text{a constant(1)}; \quad 0 < \varepsilon << a_h \le r$$

$$\approx \frac{1}{5}\left(5\frac{k_1}{r} + k_2\left(3a_h^2 + 5r^2\right) + k_3\left(\frac{a_h^2}{r} + 5r\right)\right) + \text{a constant(1)}; \quad \varepsilon = 0$$

$$k_n = \kappa_n M_h; \quad n = 1,2,3; \quad |U1| << 1$$

(9.4b)

$$U2(r) \approx 2\pi\rho_h \int_{\varepsilon}^{r}\int_{0}^{\pi} g(z)x^2 \sin(\theta)d\theta dx; \quad 0 < \varepsilon << r \le a_h; \quad |U2| << 1$$

$$\approx \frac{1}{5}\left(\frac{5k_1r^2 + 8k_2r^5 + 6k_3r^4}{a_h^3}\right) + \text{a constant(2)}; \quad \varepsilon = 0$$

In view of the conditions (concerning the magnitude of the dimensionless potential) attached to (9.4a/b)

(9.4c) constant(1)=constant(2)=0

The principal approximations we have made at (9.4a/b) is that we have neglected the effects of density and the gradual reduction of thickness of the visible disc with increasing distance from the centre; both are assumed constant and the thickness and density fall to zero abruptly when the edge is reached.

The dimensionless potentials $U1(r)$ and $U2(r)$ given by (9.4a/b) look very different but when

(9.5a) $r = a_h$

they are, in fact, the same (as they should be for continuity)

(9.5b) $U1(a_h) = U2(a_h) = \frac{1}{5}\left(\frac{5k_1}{a_h} + 8k_2a_h^2 + 6k_3a_h\right)$ See (9.4a/b)

When

(9.6a) $r >> a_h$

(9.6b) $U1(r) \approx \dfrac{k_1}{r} + k_2 r^2 + k_3 r$ See (8.4/7.5)

the dimensionless potential is the same as a point, at the origin, of mass M_h according to the Newtonian approximation to the K equation. As the radius decreases (subject to the condition $r \geq a_h$) the value of $U1(r)$ approaches (9.5b). When

(9.7a) $r \to 0$

the dimensionless potential is

(9.7b) $U2(r) \approx k_1 \dfrac{r}{a_h^3}; \quad \left| \dfrac{8k_2 r^5 + 6k_3 r^4}{5a_h^3} \right| << \left| k_1 \dfrac{r^2}{a_h^3} \right|$

At first $U(r)$ is negative (we adopt positive sign outside the brackets in (8.4)), with increasing r, and then climbs until, for the first time,

(9.7c) $U2(r)\big|_{r=a_h} = 0 \Rightarrow \dfrac{1}{5} \dfrac{5k_1 r^2 + 8k_2 r^5 + 6k_3 r^4}{a_h^3}\Bigg|_{r=a_h} = 0$

; see (9.5b/8.4). On account of (9.5b) this requires that (9.5a) should be satisfied and

(9.7d) $U1(a_h) = 0 \Rightarrow \dfrac{1}{5}\left(\dfrac{5k_1}{a_h} + 8k_2 a_h^2 + 6k_3 a_h \right) = 0$

Equation (9.7d) may be used to express a_h in terms of the constants $k_n;\ n = 1,2,3$. The expressions are complex and, for the purpose of illustration, we assume

(9.8a) $k_2 = 0$

then

.

(9.8b) $a_h = \frac{1}{6}\frac{^+\sqrt{-30k_1k_3}}{k_3} = \frac{1}{6}\frac{^+\sqrt{-30\kappa_1\kappa_3}}{\kappa_3} \Rightarrow \kappa_3 = -\frac{5}{6}\frac{\kappa_1}{a_h^2}$ See (8.2c/8.4)

which, if we accept (9.8a) and we accept the universality the κ_n (see (8.4)), the radii of the halos of all galaxies must be the same. So (9.8a), we must presume, does not obtain in general.

The archetype v,r curve rises steeply from near zero, at a small radius, then abruptly levels out to a constant velocity (which extends to an unknown radius) [11],[12]. Fig. 1 (showing uncertainty bounds of measurements) approximates this behaviour. According to Newton's law of gravity this is impossible. According to that law beyond a certain point the velocity $v(r)$ should fall off gradually with increasing r; see the left hand of Fig 3. But the Newtonian approximation to the K equation (see (6.5/8.4)) is capable of producing, for a small distance beyond a certain radius, no force and hence a locally constant velocity; see (7.5).

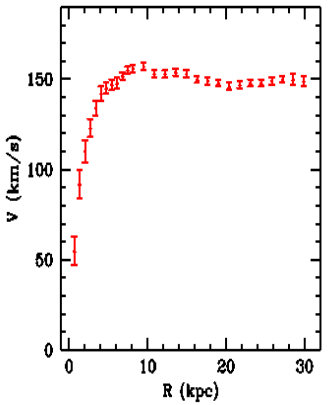


Fig.1- NGC 3198 Bergman 1989

The test particle is bound to the galaxy as long as the attractive (inwards) force is not zero. The force per unit mass of the particle $-c^2U(r),_r$ is zero at the periphery of the halo ($r = a_h$). Unfortunately neither

(9.9a) $$-c^2U1,_r\Big|_{r=a_h} = 0$$

nor

(9.9b) $$-c^2U2,_r\Big|_{r=a_h} = 0$$

is consistent with (9.7c). The equations (9.7c/d) give

(9.10a) $$k_2 = -\frac{1}{8}\frac{6a_h^2k_3 + 5k_1}{a_h^3}; \quad k_3 = k_3$$

whereas equations (9.9a/b) give

(9.10b) $$k_2 = \frac{2k_1}{a_h^3}; \quad k_3 = -\frac{15}{4}\frac{k_1}{a_h^2}$$

Result (9.10a) should give no surprise; there is effectively only one equation in two variables. Even

(9.11a) $$-c^2\,U1(r),_r\,,_r\Big|_{r=a_h} = -c^2\,U2(r),_r\,,_r\Big|_{r=a_h} = 0$$

is inconsistent with (9.10b). Equations (9.11a) give

(9.11b) $$k_2 = \frac{1}{2}\frac{k_1}{a_h^3}; \quad k_3 = -\frac{5}{4}\frac{k_1}{a_h^2}$$

The equations (9.10b) are the most important for the Newtonian theory. The others just express continuity at the edge $r = a_h$ on the plateaux of the archetype. To establish continuity, more generally, result (9.5b) should read

(9.12a) $$U1(a_h) = U2(a_h) = \frac{1}{5}\left(\frac{5k_1}{a_h} + 8k_2 a_h^2 + 6k_3 a_h \right) \neq 0$$

and result (9.11a) should read

(9.12b) $$U1(r),_r,_r\big|_{r=a_h} = U2(r),_r,_r\big|_{r=a_h} \neq 0$$

The article [12] shows that, in practice, the observed v, r curve seldom, if ever, follows the archetype; not even Fig. 1 follows the archetype strictly. For example the galaxy M33, see Fig 2, has a simple observed curve; but that curve levels off to a slight rise (out to an unknown radius), when, according to the archetype, it should level off to a constant velocity at a_h. Others show more radical departures from the archetype; these are probably due to the effect on the dimensionless potential of the spiral arms.

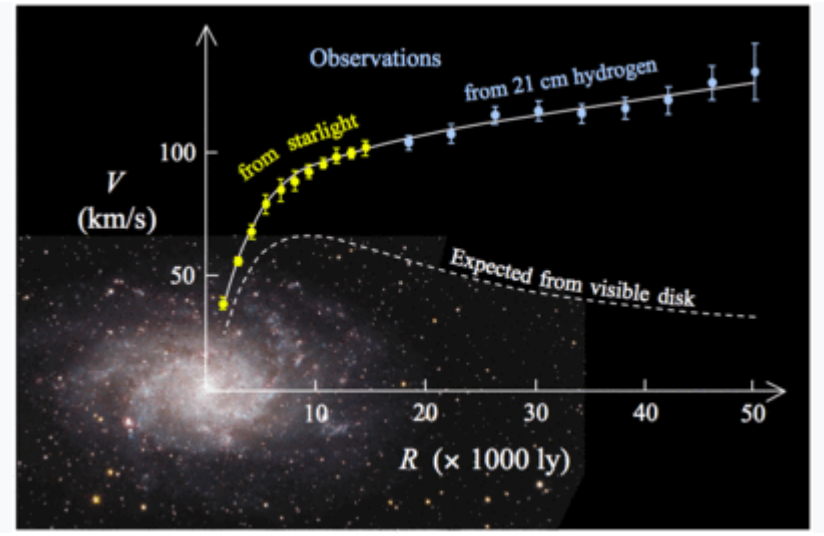


Fig 2- The v,r Curve For M33

It must be concluded that the archetype v,r curve is an over simplification.

It can be shown [10] that an orbit of a test particle, moving with speed $v(r)$ in a weak gravitational field characterised by an SS dimensionless potential with respect to the origin $U(r)$, satisfies

(9.13a)

$$\frac{d^2r}{dt^2} - r\left(\frac{d\vartheta}{dt}\right)^2 - c^2U_{,r}(r) = 0; \quad c^2 >> v^2(r) \quad \text{radial acceleration}$$

(9.13b) $$r\frac{d^2\vartheta}{dt^2} + 2\frac{dr}{dt}\frac{d\vartheta}{dt} = 0 \quad \text{transverse acceleration}$$

(9.13c) $$v(r)^2 = \left(\frac{dr}{dt}\right)^2 + \left(r\frac{d\vartheta}{dt}\right)^2; \quad v(r) > 0$$

(9.13d) $$U(r) = U1(r) \text{ or } U2(r) \text{ as appropriate}$$

where t is time and ϑ is the azimuth. This gives the v,r curve for our model and belongs to the Newtonian scheme although the dimensionless potential includes extra terms.

As is well known [10] elimination of the time between (9.3a/b) gives

(9.14a) $$\frac{d^2(1/r)}{d\vartheta^2} + \frac{1}{r} = -c^2\frac{U_{,r}(r)r^2}{h_1^2}$$

where h_1 is a constant that satisfies

(9.14b) $$\frac{d\vartheta}{dt} = \frac{h_1}{r^2}$$

For example, when the motion is approximately circular (the presence of the k_2, k_3 terms in U means, by *Bertrand's theorem*, that the orbit cannot be closed),

(9.14c) $$-c^2U_{,r} + \frac{v^2}{r} \to 0 \Rightarrow v^2 \approx rc^2U_{,r}$$

Some numerical results:

(9.15a) $$k_2 = \frac{2k_1}{a_h^3}; \quad k_3 = -\frac{15}{4}\frac{k_1}{a_h^2}; \quad a_h = 10^5 \ ly \equiv 9.4601\times10^{20} \ m$$

(assumed)

where

(9.15b)

$$\kappa_1 = -\frac{G}{c^2} = -7.4237\times10^{-28} \ mkg^{-1}; \quad k_1 = M_h\kappa_1; \quad M_h = 10^{42} \ kg$$

(9.15c)

$$k_1 = -7.4237\times10^{14} \ m; \quad k_2 = -1.7537\times10^{-48} \ m^{-2}; \quad k_3 = 3.1107\times10^{-27} \ m^{-1}$$

(9.15d) $$\kappa_2 = -1.7357\times10^{-90} \ m^{-2}kg^{-1}; \quad \kappa_3 = 3.1107\times10^{-69} \ m^{-1}kg^{-1}$$

We have used the halo, with the Newtonian approximation to the K equation as the law of gravity, to account for Dark Matter. If the halo does not exist the postulate of Dark Matter is still necessary in certain circumstances. To get the maximum, with Newton's law of gravity, up to the measured v,r curve one has to add to the mass to the mass of the visible galaxy; even then the curve does not arrive at a true plateaux and the modelling is much more difficult. Even so, we might expect to get similar values for κ_2 and κ_3.

Another circumstance where it is necessary to take account of Dark Matter is in the space *between* the galaxies. Multiple images of quasars (predicted by Einstein's theory) show much more bending of space-time than be accounted for by the apparent (luminocity of galaxies) mean density. The conclusion must be drawn that Dark Matter permeates space.

To get an idea of how difficult is to infer the existence of Dark Matter, merely from the appearance of the galaxies, is shown by the simulations of Fig. 3.

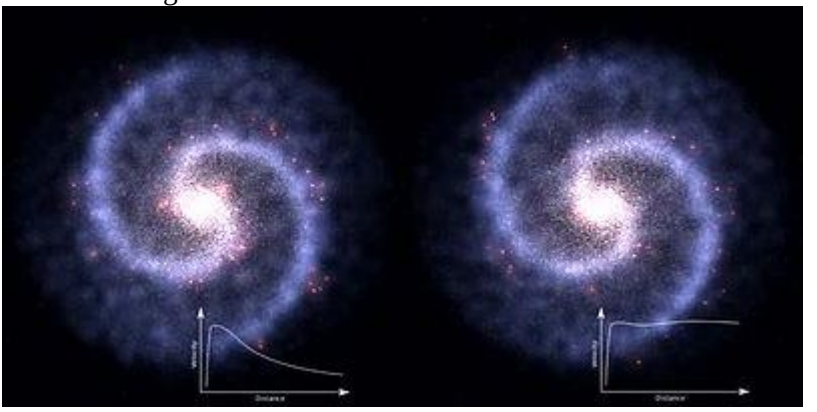


Fig. 3- Left: A simulated galaxy, with its v,r curve, without Dark Matter. Right: Galaxy with an approximately flat rotation curve that would be expected under the presence of Dark Matter.

It is stated in [10] that according to the Newtonian law G drifts and that it is getting smaller

(9.16) $$-\frac{1}{G}\frac{dG}{dt} \approx 3.6 \pm 1.8 \text{ parts in } 10^{11} \text{ per year}$$

wheras according to the Einstienian theory G is constant. Might it be that the extra terms, in the Newtonian approximation to the K equation, roughly account for the discrepancy? To test this hypothesis we need to understand how the result is arrived at. "Observations of the occultations of fixed stars by the Moon, when corrected for all known 'ordinary' causes, leads to….." result (9.16). The dimensionless potential of the Earth-Moon system is approximately

(9.17a) $$U(r) \approx \frac{k_1}{r} + k_2 r^2 + k_3 r$$ See (7.5)

The force on the Moon due to the Earth is

(9.17b) $$-c^2 U_{,r} \approx -c^2 \left(-\frac{k_1}{r^2} + 2k_2 r + k_3 \right)$$

where r is the distance from the centre of the Earth to the centre of the Moon. Substituting (9.10b) into (9.17b) we get

(9.17c) $$-c^2 U_{,r} \approx -\frac{G m_E}{r^2} \left(1 - 4\left(\frac{r}{a_h} \right)^3 + \frac{15}{4}\left(\frac{r}{a_h} \right)^2 \right)$$

Thus the force is proportional to G. The Moon's radius r is a function of time t. The dimensionless quantity r / a_h is small; far too small to have any influence on (9.16). So we have the result that Newtonian law is sufficient for (9.16)

(9.18) $$-c^2 U_{,r} \approx -\frac{G m_E}{r^2}$$

Therefore the hypothesis is negated; the argument proceeds as in [10].

10. The Sun And The Pioneer Anomaly [3]

Given r greater than the sun's radius we assume the form (see (7.5)) which is appropriate to a single particle at the origin

(10.1) $U_o = \frac{k_{1o}}{r} + k_{2o} r^2 + k_{3o} r; \quad |U_o| << 1; \quad k_{1o} = -\frac{G m_o}{c^2}; \quad k_{2o}, k_{3o}$ constants for the gravitational potential of the sun. Here m_o is the *ordinary* mass of the sun. We do not know the values to be assigned to k_{2o} and k_{3o} but it is reasonable to assume that for

(10.2) $r < 100\ AU$

the term associated with k_{2o} can be neglected. So

(10.3) $U_o \approx \frac{k_{1o}}{r} + k_{3o} r; \quad |U_o| << 1; \quad k_{1o} = -\frac{G m_o}{c^2}$

The corresponding radial acceleration of test particles can be got from

(10.4) $-c^2 \frac{\partial U_o}{\partial r} \approx -c^2 \left(-\frac{k_{1o}}{r^2} + k_{3o} \right) = -\frac{G m_o}{r^2} - c^2 k_{3o}$

Thus, in addition to the usual inverse square term, there is a constant acceleration/ deceleration depending on k_{3o}. Now it is stated in [3] that at

(10.5) $r = 86\ AU$

there appeared to be a constant attraction towards the sun of

(10.6)
$$\begin{aligned} & c^2 k_{3o} = 10^{-10} \times (\text{acceleration due to gravity at the earth's surface}) \\ & \approx 10^{-9}\ ms^{-2} \Rightarrow k_{3o} \approx 10^{-9} / 9 \times 10^{16}\ m^{-1} \approx 1.1 \times 10^{-26}\ m^{-1} \\ & \Rightarrow \kappa_{3o} \approx 1.1 \times 10^{-26}\ m^{-1} / 1.8 \times 10^{30}\ \text{kg} \approx 6.1 \times 10^{-57} m^{-1} kg^{-1} \end{aligned}$$

This is known as the Pioneer Anomaly. By this calculation the value of κ_{3o} is in perfectly reasonable; but as we do not know what it *is* we are none the wiser! But, if it is to be universal, according to the calculations in Section 9 it is much too big.

There are numerous stars, in the vicinity of the sun, that, presumably, have constant components of radial acceleration/ deceleration of similar magnitude to the sun. These motions must be presumed to have all possible directions and a range of magnitudes. Will they cancel each other out? Well, if one includes the whole Universe, presumably so. But, if one includes the whole Universe, the Newtonian assumptions, on which these calculations are based, may be invalid. All we can say is that, if the Pioneer Anomaly is gravitational then, it is *not* simply related to the constant solar term.

There are many current explanations for the Pioneer Anomaly. These divide into two classes. Firstly, there are theoretical explanations which suppose that either we have the law of gravity wrong or we wrongly applying it. For example [13] considers that the time measures at the observer and at Pioneer differ, because of quantum effects, and, in consequence, the apparent motion requires correction. Secondly, there are practical explanations that suppose that we have neglected some small physical effect in evaluating the motion. For example [14] draws our attention to the non-uniform way that the structure of Pioneer radiates heat. In consequence Pioneer receives a small impulse which, it is supposed, accounts for the anomaly. The balance of opinion seems to be converging on the 'small physical effect' as the culprit.

11. Kilmister's Equation:-Some Analysis Concerning Isotropic-Homogeneous Space; Calculations To Do With The Cosmological Metric

We have already met Kilmister's Equation derived by the late Clive Kilmister from the GTE (see (5.3)). This is a classical tensor equation defined on a Riemannian manifold *C*

(11.1) $K_{ab} \equiv g^{ef}(R_{ab;ef} + \frac{2}{3}R_{ae}R_{fb}) = 0; \qquad a,b,e,f = 1,2,...n_c;$ see (4.2a)

where ';' denotes covariant differentiation. It is otherwise known as the K equation. The K equation reduces to the GTE at the pole of Cartesian geodesics; and, because of the choice of those coordinates, approximates the GTE in the neighbourhood of the pole. When $n_c = 4, n_p = 1$ it is called the relativistic K equation (RKE) and the GTE should then be called the relativistic gravitational equation (RGTE); we shall not follow this usage, however, because the meaning should be clear by the context.

The RKE must be considered along with the Einstein equation (because the Einstein equation defines the mechanical tensor T_b^a and so brings mass into an otherwise geometric theory); see section 4. As we have already seen (see (4.3b)), for $n_c = 4$,

(11.2)
$$R_v^u = G_v^u - \tfrac{1}{2}G\delta_v^u = -\chi T_v^u - \Lambda\delta_v^u + \tfrac{1}{2}(\chi T + 4\Lambda)\delta_v^u = -\chi\left(T_v^u - \tfrac{1}{2}T\right) + \Lambda\delta_v^u$$

When space-time is *truly empty* (of 'ordinary matter' and of 'dark' vacuum energy)

(11.3a) $T_b^a = 0; \quad \Lambda = 0 \Rightarrow R_v^u = \Lambda\delta_v^u = 0 \Rightarrow R_{ab} = 0$

where (11.3a) is the Einstein law of gravity in the space between particles. The RKE gives a consistent result in that it requires that the universal constant Λ vanishes

(11.3b) $0 + \frac{2}{3}(\Lambda)^2 g_{ab} = 0 \Rightarrow \Lambda = 0$

If, however, there is no ordinary matter but there *is* vacuum energy then the law of gravity, for empty space, can still operate with $\Lambda \neq 0$

(11.4) $T_b^a = -\dfrac{\Lambda}{\chi}\delta_b^a \Rightarrow R_v^u = 0 \Rightarrow G_b^a = 0, K_b^a = 0$ See (4.1c/11.2)

This result seems to indicate that G_b^a applies to all forms of matter whereas Λ applies to vacuum (dark) energy only.

The K equation is a collection of PDEs in the g_{uv} as dependant variables and the coordinates $\underline{x}$ as independent variables. Given a fundamental tensor g_{uv} which satisfies the K equation we have a Riemannian space of the points $\underline{x}$; a K space. Study of K spaces is study of the K equation. For example we may ask: Is a K space a space of constant, non-zero Riemannian curvature κ? Such a space satisfies [6]

(11.5a) $$R_{rsmn} = \kappa\left(g_{rm}g_{sn} - g_{rn}g_{sm}\right); \quad \kappa \neq 0$$

This equation implies [15] (inner product by g^{rs})

(11.5b) $$R_{mn} = -(n_c - 1)\kappa g_{mn}; \quad \text{Einstein space}$$

That is a space of constant Riemannian curvature is an *Einstein space* [6] with a constant invariant. Substituting this into the K equation

(11.5c) $$0 + \tfrac{2}{3}\left((n_c - 1)\kappa\right)^2 g_{ab} = 0 \Rightarrow \kappa = 0 \Rightarrow R_{ab} = 0; \quad n_c > 1$$

So the answer to the above question is in the negative. A K space can be a space of constant Riemannian curvature but only if the curvature is zero; that is the space is flat [6],[15]. A constant curvature Riemannian space can be shown to be isotropic and homogeneous (Schur's Theorem) [15]. So the RKE does not permit the 4-space of space-time to be isotropic and homogeneous unless it is flat.

We need to consider, however, a related space. In GR the *cosmological metric* pertains to a 4-space which is *not*, in general, of constant curvature

(11.6a) $$ds^2 = d\tau^2 - \frac{A^2(\tau)}{c^2}\frac{(dx^1)^2 + (dx^2)^2 + (dx^3)^2}{\left(1 + kr^2/4\right)^2}; \quad k = -1, 0, 1$$

(11.6b) $$g_{uv}=0,\ u\neq v;\quad g_{JJ}=-\frac{A^2(\tau)}{c^2\left(1+kr^2/4\right)^2};\quad g_{44}=1;\quad J=1,2,3$$

where units have been chosen so that the function $A(\tau)$ has the physical dimensions of length, ds and τ have the physical dimensions of time and the coordinates x^J have no physical dimensions [16]. The 3-subspace, formed by the x^J for any given time coordinate $\tau \equiv x^4$, *is* of constant Riemannian curvature.

Going now to a physical scale which is so large, that the galaxies form individual particles of a fluid, the metric (11.6) can represent a model universe the 3-subspace of which is full of isotropic and homogeneous matter. This is a simplification of the real Universe but it suffices for the present argument. The function $A(\tau)$ is sometimes called the *radius of the universe*. The metric (11.6) is otherwise known as the Friedman–Lemaître–Robertson–Walker (FLRW) metric [17]; it forms the basis of the Big Bang model.

We may make another interpretation of (11.6) where, for illustration, we have transformed to polars

(11.7a)

$$ds^2=du^2-\frac{A^2(\tau)}{\left(1+kr^2/4\right)^2}\left(dr^2+r^2\left(d\theta^2+\sin^2\theta d\phi^2\right)\right);\quad du^2=c^2d\tau^2$$

(11.7b)

$$g_{11}=-\frac{A^2(\tau)}{\left(1+kr^2/4\right)^2};\quad g_{22}=-\frac{r^2A^2(\tau)}{\left(1+kr^2/4\right)^2};\quad g_{33}=-\frac{r^2\sin^2\theta A^2(\tau)}{\left(1+kr^2/4\right)^2};$$

$$g_{44}=1$$

Here τ has the physical dimensions of time, the function $A(\tau)$ is dimensionless, r has the physical dimensions of length as does ds and u; and k has the physical dimensions of (length)$^{-2}$. The constant k, in this interpretation, is continuous; it is the negative of the *Gaussian*

curvature of the 3-subspace [15]; and, providing it is non-zero, scales the distance r. A zero value corresponds to a zero value at (11.6a); the sign for $k \neq 0$ also corresponds to the sign at (11.6a). The quantity $A(\tau)$ is often called the *expansion/ contraction factor* of the metric (11.7); that is the 3-subspace of the r, θ, φ expands/ contracts with co-ordinate time τ unless $A(\tau)$ is constant. We assume

(11.7c) $$A(\tau) > 0$$

Given (11.7a) the tensors R^u_v, G^u_v and K^u_v turn out to be diagonal. Two of the four diagonal elements, in each case, are unique. So, for example,

(11.8a) $$G^J_J = -\frac{1}{A^2}\left(k + A'^2 + 2AA''\right) = -\chi T^J_J - \Lambda; \quad J = 1,2,3; \quad A \neq 0$$

(11.8b) $$G^4_4 = -\frac{3}{A^2}\left(k + A'^2\right) = -\chi T^4_4 - \Lambda; \quad G^u_v = 0, u \neq v; \quad A \neq 0$$

(11.9a)

$$K^J_J = \frac{1}{3}\frac{1}{A^4}\begin{pmatrix} 4A'^2k - 4A'^4 - 7AA'^2A'' + 15A^2A'A''' \\ +11A^2A''^2 + 3A^3A'''' - 4AA''k + 8k^2 \end{pmatrix} = 0; \quad A \neq 0$$

Maple 12

(11.9b)

$$K^4_4 = \frac{3}{A^4}\left(A^2A'A''' + 4A'^2k + A^2A''^2 - 5AA'^2A'' + A^3A'''' + 4A'^4\right) = 0;$$

$$K^u_v = 0, u \neq v; \quad A \neq 0$$

where

(11.9c) $$(.)' \equiv \frac{d(.)}{du}; \quad u \equiv c\tau$$

The only check on (11.8) I have been able to make is that the expression for the Einstein tensor, given by the machine, agrees with that in [16] which was written before electronic computers existed!

The two ODEs (11.9a/b) must have consistent solutions. If they have such then it is *possible* for the K equation to be satisfied; but there is no guarantee that the metric (11.7) always pertains to a K space. That G_v^u and K_v^u at (11.8/9) is independent of the coordinates r,θ,ϕ is a symptom of the fact that the 3-subspace is isotropic and homogeneous. In fact the transformation to polars at (11.7a), in the 3-subspace, is nugatory.

The elements of the mechanical tensor T_v^u and the scalar Λ determine the elements of the tensor G_v^u; see (4.2c). As has already been shown (see (11.4)), in a model universe empty of all ordinary mass/ energy but with $\Lambda \neq 0$,

$$
\begin{aligned}
& T_b^a = -\frac{\Lambda}{\chi}\delta_b^a \Rightarrow G_b^a = 0 \\
(11.10) \qquad & \Rightarrow G_4^4 = \frac{3}{A^2}\left(k + A'^2\right) = 0; \\
& G_J^J = \frac{1}{A^2}\left(k + A'^2 + 2AA''\right) = 0
\end{aligned}
$$

On the assumption that k is constant the first of these ODEs, consistent with the second, gives

(11.11) $$A(u) = \pm\sqrt{-k}u + a_0 = \frac{u}{c}a_1 + a_0; \quad a_1 = \pm\sqrt{-c^2k}$$ Maple 12

and manual

where a_1 and a_0 are constants of integration. NB It is more convenient here and in the sequel to express A as a function of the length u rather than the time τ. If a_1 is to be real then

(11.12a) $k \leq 0$ characteristic of an empty model universe

The Gaussian curvature of a closed 3-subspace is positive. This means that the 3-space of the empty model universe is open (hyperbolic) unless it is flat. The constant a_0 can be determined by the initial condition

(11.12b) $A=1$ when $\tau=0\Rightarrow u=0\Rightarrow a_0=1$

This condition assumes that at the instant 'now' is the origin of time $\tau=0\Rightarrow u=0$ and that the model universe is not expanded at that instant. So, finally, either

(11.13a) $A(u)=\pm\sqrt{-k}u+1$ or $A(u)=1\,\forall\, u;\quad k\le 0$

and *both* the ODEs (11.8) are exactly satisfied. Further Maple 16 shows that *both* the ODEs (11.9) are exactly satisfied by this solution. With the metric (11.7) the model universe, empty of ordinary matter, either expands or contracts, in proportion to the coordinate time at every point, or it is stationary (in, as it turns out, unstable equilibrium). Alternatively, we can shift the origin of time $\tau=0$ to *the beginning*, providing that

(11.13b) $A(0)=a_0;\quad a_0$ may be zero but if $a_0\neq 0$ then $1>>|a_0|$

We note a remark in [18] that 'we can have curvature without matter but not matter without curvature'.

But an empty model universe, even one with the cosmological constant non-zero, is of limited interest! To go further we begin by introducing the mean pressure $p(u)$ and the mean density of ordinary mass $\rho(u)$ (averages taken over space x^J, $J=1,2,3$ but not too far!) as functions of time/ distance u; These quantities can be defined [16], in the case (11.7), by

(11.14) $\qquad p_J\equiv -T_J^J;\quad c^2\rho\equiv T_4^4;\quad J=1,2,3;\quad T_b^a=0,\, a\neq b$

where both p_J and $c^2\rho$ have the physical dimensions of energy per unit volume. In general, by virtue of the Einstein equations (4.2c),

(11.15a)

$$G_J^J = -\frac{1}{A^2}\left(k + A'^2 + 2AA''\right) = -\chi T_J^j - \Lambda = \chi p_J - \Lambda; \quad J = 1,2,3; \quad A \neq 0$$

(11.15b) $G_4^4 = -\frac{3}{A^2}\left(k + A'^2\right) = -\chi T_4^4 - \Lambda = -\chi c^2 \rho - \Lambda$ See (11.14)

If we consider the special case (a model universe empty of ordinary matter) we get

(11.15c)
$$G_b^a = 0 \Rightarrow \text{pressure} \equiv -T_J^J = \frac{\Lambda}{\chi} \equiv p_{00J} \text{ (say)} = p_{00};$$
$$c^2 \times \text{density} \equiv T_4^4 = -\frac{\Lambda}{\chi} \equiv c^2 \rho_{00} \text{ (say)}$$
See (11.10)

Note that in Section 7 we have already shown that, in the 'Newtonian' case in order to agree with observation, Λ is negative; so ρ_{00} is positive and $p_{00J} \equiv p_{00}$ is negative in a model universe empty of ordinary matter.

We now introduce Hubble's constant [16/18]

(11.16) $$\mathsf{H}_0 = \underset{u \to 0}{Lt} \frac{\dot{A}(u)}{A(u)} = c \underset{u \to 0}{Lt} \frac{A'(u)}{A(u)} \approx 2.055 \times 10^{-18} \ s^{-1}$$

That is, wherever we put the origin ($u = 0$), the model universe and the actual Universe seems to expand (see (11.13a)). From (11.15b) we deduce Friedmann's equation

(11.17a) $k + \frac{\mathsf{H}_0^2}{c^2} = \frac{\chi c^2 \rho_0 + \Lambda}{3}$ See (7.9a)

because, by definition,

(11.17b) $A(0) = 1$ See (11.12b)

Transposing (11.17a)

$$(11.17c)\quad H_0^2 = c^2\left(\frac{\chi c^2\rho_0 + \Lambda}{3} - k\right)$$

In other words Hubble's constant increases with the mean ordinary density ρ_0 and

$$(11.18)\qquad H_0(\min) = +c\sqrt{\frac{\Lambda}{3} - k};\quad \rho_0 = 0$$

Since, in an empty model universe, $k < 0$ and $H_0(\min)$ is, by definition, real (see (11.11))

$$(11.19)\qquad \frac{\Lambda}{3} - k > 0 \Rightarrow -3k < -\Lambda \Rightarrow 3k > \Lambda$$ in an empty model universe

12. Here And Now Various Quantities Are Small So We Might Seek A Perturbation

We continue the argument for $\rho_0 > 0$ by perturbing the solution (11.11/13a) on the grounds (perhaps spurious) that, here and now, $c^2\chi\rho_0 / k,\ \chi p / k$ are small compared to unity. Many astronomers believe, however, that $k = 0$; in which case the results are spurious. If we do not make this assumption we have some hope!

We have

$$(12.2a)\qquad A(u) \equiv 1 + \frac{u}{l} + \varepsilon f(u);\quad 0 \neq |\varepsilon| << 1;\quad k \neq 0 \equiv -\frac{1}{l^2{}_1}$$

Take the positive sign in $\frac{1}{l} = \pm\sqrt{-k}$ to give

(12.2b)
$$A'(u) = \frac{1}{l} + \varepsilon f'(u); \quad A''(u) = \varepsilon f''(u);$$
$$A'''(u) = \varepsilon f'''(u); \quad A''''(u) = \varepsilon f''''(u)$$

where ε and function $f(u)$ are defined as having no physical dimensions. It is to be understood that the small density and pressure, here and now, is the perturbing agency and that is small, compared with unity, but not zero. Substitute (12.2b), for , into (11.15b)

(12.3)
$$\frac{3}{A^2}\left(k + A'^2\right) = \frac{3}{\left(1+\varepsilon f(0)\right)^2}\left(-\frac{1}{l^2} + \left(\frac{1}{l} + \varepsilon f'(0)\right)^2\right)$$
$$\approx 6\left(1 - 2\varepsilon f(0)\right)\frac{2}{l}\varepsilon f'(0) \approx 12\frac{\varepsilon f'(0)}{l} = \left(\chi c^2 \rho_0 + \Lambda\right)$$
$$\Rightarrow \varepsilon f'(0) \approx \frac{l}{12}\left(\chi c^2 \rho_0 + \Lambda\right); \quad f'(0) \neq 0$$

neglecting the terms in ε^2 and higher powers. Is $\varepsilon f'(0)$ small enough to be regarded as a perturbation of ? We opened this section with an assertion that was small (compared to unity); small enough to amount to perturbation of the uniform motion (11.11) characteristic of an empty universe. Well, utilising (11.17a),

(12.4)

is to be compared with Some astronomers believe

On the other hand, is free, so we can choose

(12.5a)

which, according to (12.4), makes as small as we like consistent with the approximation made at (12.4). If we make this choice, however, (11.17a) requires

(12.5b)

That is, the ordinary mean mass/ energy density is approximately constant and universal. This result is reminiscent of (11.15c) and corresponds to an empty model universe with

(12.5c)

So, since , the values given at (12.6a) for and hence are *spurious*. If we proceed with the formal perturbation, however, we get from (11.15a/b)

(12.6)

In other words an expanding model universe empty of ordinary matter! See (11.10/11/15).

13. Finite Expansion Of $A(u)$ As An Approximate Solution

There might be no perturbation of the solution (11.13a); if there is, all we have proved above is that, the perturbation is stable. Maple 16 can solve equations (11.8) but the solutions are immensely complex and implicit; and I cannot identify the constants of integration.

We try a Taylor's expansion about the origin 'here and now' $u = 0$

(13.1a)

$$A(u) \equiv a_0 + a_1 u + a_2 u^2 + a_3 u^3 + a_4 u^4 + a_5 u^5 + O(u^6); \quad u \geq 0$$
$$A(0) \equiv 1 \Rightarrow a_0 = 1; \quad \text{Initial Condition}; \quad 1 >> |u|; \quad \text{Approximation}$$

where the coefficients $a_0, a_1, ..., a_5$ are free. The second line of (13.1a) derives from the initial condition 'here and now'. Substituted into all four of the equations (11.8a/b) and (11.9a/b) this gives four algebraic equations for the coefficients $a_1, a_2, ...a_4$; (the coefficient a_5 happens to vanish). This gambit solves the problem that both the equations (11.9a/b) having the same solution $A(u)$. In fact the distance/ time u which we actually measure is negative. So the $A'(0)$ is negative. We might as well admit this by defining

(13.1b)

$$A(u) \cong a_0 - a_1 u + a_2 u^2 - a_3 u^3 + a_4 u^4 - a_5 u^5 + O(u^6); \quad u \le 0$$

$$A(0) \equiv 1 \Rightarrow a_0 = 1; \quad \text{Initial Condition}; \quad 1 >> |u|; \quad \text{Approximation}$$

In practice it makes no difference, to the results, which we choose of the definitions (13.1).

Because the density and pressure vary away from the origin we should write, at the least,

(13.2) $$p \cong p_0 + p_1 u + O(u^2); \quad \rho \cong \rho_0 + \rho_1 u + O(u^2); \quad |u| \to 0$$

as an approximation. This assumption will probably not be valid very far into the past or the future but it is simple. In consequence substitution of definitions (13.1/2) into equations (11.8/9) produces four polynomial equations in u; the coefficients of the powers of u in these four equations are functions of the eleven constants $a_0, a_1, ...a_4, k, p_0, p_1, \rho_0, \rho_1, \Lambda$. If we wish to solve for these constants we must produce more equations. In explanation, we get from the initial condition,

(13.3a) $$a_0 = 1 \quad \text{See (13.1)}$$

and from the definition of Hubble's constant

(13.3b) $$a_1 = \frac{H_0}{c} \quad \text{See (11.16)}$$

so that leaves nine constants $a_2, a_3, a_4, p_0, p_1, \rho_0, \rho_1, k, \Lambda$. The object is to produce nine equations in the unknowns $a_2, a_3, a_4, p_0, p_1, \rho_0, \rho_1, k, \Lambda$. The first seven are variable, depending on where we put the origin $u = 0$; the last two, k and Λ, are universal constants by definition. We can get more putative equations by any one of three alternative procedures:

Equating coefficients, of powers u^2 and higher, on both sides of some of the original polynomial equations.

Covariantly differentiating the original tensor equations (4.2c) and (5.5) and setting up more polynomial equations via the metric (11.7a/b).

Going back to the QM hierarchy of constraints and postulating that the third or higher constraint holds, deriving the equivalent classical formula, eliminating the p_k, and converting this to a tensor equation, on the assumption that the equivalent classical formula uses Cartesian geodesics, and setting up more polynomial equations via the metric (11.7a/b).

Clearly, these procedures are in ascending amounts of work. The work required by the last is enormous; unfortunately, from the point of view of the present theory, it is probably the only entirely valid one.

The first procedure has been tried; it is easily accomplished by Maple 16. It forces particular constraints on the differentials at $u = 0$; and these may not apply in practice. A similar criticism may be made of the second procedure; but at least it manipulates tensor equations. The first produces very high densities and pressures; this is entirely wrong for most of the history of the Universe.

Using the original five equations, including (13.3b), Maple 16 has given formulae for a_1, a_2, a_3, a_4, k in terms of the rest of the quantities $p_0, p_1, \rho_0, \rho_1, \Lambda$. As is to be expected these formulae do not include the quantities p_1 and ρ_1; the five equations do not include a_5 either. The formulae show that there is a solution for a model universe empty of ordinary matter

$$
\begin{aligned}
& p_0 = \frac{\Lambda}{\chi}, \rho_0 = -\frac{\Lambda}{\chi c^2} \\
& a_0 = 1 \\
& a_1 = \frac{H_0}{c} = 6.8548 \times 10^{-27}\ m^{-1} \\
& a_2 = 0, a_3 = 0, a_4 = 0 \\
& k = -\left(\frac{H_0}{c}\right)^2 = -4.6988 \times 10^{-53}\ m^{-2}
\end{aligned}
\tag{13.4}
$$

As previous work has shown this solution is exact (see (11.13a)). The last line of (13.4) corresponds to (11.15c).

Many astronomers believe that $k = 0$. That being so the formulae show that we have only to supply values for p_0 and ρ_0 to solve for Λ. The assumptions are shown first; then the results

(13.5a)

$k = 0; \quad p_0 = 0; \quad \rho_0 = 8 \times 10^{-27}\ kgm^{-3}$ closure value of density

$$
\begin{aligned}
& a_0 = 1 \\
& a_1 = 6.8548 \times 10^{-27}\ m^{-1} \\
& a_2 = -1.3823 \times 10^{-53}\ m^{-2} \\
& a_3 = 8.0244 \times 10^{-80}\ m^{-3} \\
& a_4 = -8.0798 \times 10^{-106}\ m^{-4} \\
& \Lambda = -8.3052 \times 10^{-54}\ m^{-2}
\end{aligned}
\tag{13.5b}
$$

follow. Note that Λ is negative as required, by the Newtonian case, to produce the observed far field positive acceleration; see (7.11) to (7.14).

The universal constant k is, in the above calculation, on the cusp of becoming positive. Another calculation shows that it can become positive according to the formulae. We simply give it a positive value and repeat the calculation that leads to (13.5).

(13.6a) $$k = 10^{-54}\ m^{-2}; \quad p_0 = 0; \quad \rho_0 = 8\times 10^{-27}\ kgm^{-3}$$

(13.6b)
$$a_0 = 1$$
$$a_1 = 6.8548\times 10^{-27}\ m^{-1}$$
$$a_2 = -1.3323\times 10^{-53}\ m^{-2}$$
$$a_3 = 8.0892\times 10^{-80}\ m^{-3}$$
$$a_4 = -8.0486\times 10^{-106}\ m^{-4}$$
$$\Lambda = -5.3052\times 10^{-54}\ m^{-2}$$

Such a positive value of k may be attributed to matter and hence gravitation. The coefficients in the series for $A(u)$ are hardly changed; but Λ, while remaining negative and therefore legitimate according to the observations [4], is appreciably changed. See Friedmann's equation (11.17a).

It is to be expected that the series (13.1), for $A(u)$, will converge nicely for small values of u. The series are hardly convergent for

(13.7a) $$u = |u|_{\max} = \tau_0 c$$

where τ_0 is the age of the Universe which is inferred from measurements to be

(13.7b) $$\tau_0 = 13.8\times 10^9 \text{ years}$$

so

(13.7c) $$|u|_{\max} = \tau_0 c = 1.3054\times 10^{26}\ m$$

We can conclude that the values given for $A(|u|_{\max})$ by (13.1/6b) are very approximate. This is underlined by the fact that, in this theory, the age of the model universe is given by (keeping the same origin)

(13.8) $$A(\tau_0 c) = 0$$

where, in this case, (13.8) is a quartic with real coefficients; we take only the solution which is real and positive (if any). Corresponding to (13.6b) we get

(13.9) $\tau_0 = 2.6429 \times 10^{10}$ years

which is much too large. So we cannot calculate the age of the model universe by this method.

We can translate the origin to the beginning, or close to the beginning (when, according to current theories, quantum effects take over about 400000 years from the Big Bang). But we do not know the correct form to give $A(u)$ and we do not know which values to give p_0 and ρ_0. All we know is that $\min(A(u))$ is probably not zero. A more accurate specification of (13.8), referred to the 'here and now' origin, is probably

(13.10) $A((\tau_0 + 4 \times 10^5)c) = 0; \quad \tau_0$ in years

which would make hardly any difference to the final figure. But, if we are willing to accept the current figure for the age of the Universe, (13.8) gives another equation for the a_2, a_3, a_4. With this we can solve for another variable. We choose ρ_0. The calculation becomes

(13.11a) $k = 10^{-54}\ m^{-2}; \quad p_0 = 0; \quad \tau_0 = 13.7 \times 10^9$ years

(13.11b)
$$a_0 = 1$$
$$a_1 = 6.8548 \times 10^{-27}\ m^{-1}$$
$$a_2 = -6.3932 \times 10^{-53}\ m^{-2}$$
$$a_3 = -9.6335 \times 10^{-78}\ m^{-3}$$
$$a_4 = -2.1436 \times 10^{-105}\ m^{-4}$$
$$\Lambda = -2.0774 \times 10^{-52}\ m^{-2}$$
$$\rho_0 = 1.8849 \times 10^{-26}\ kgm^{=3}$$

The full results, corresponding to (13.4) to (13.11), is given in Appendix C. Some small explanation is in order. The reader will see that from the results, from time to time, we have set up the polynomials for $A(u)$, p and ρ; the polynomials for $o1$ and $o4$ correspond to (11.9a) and (11.9b) and those for the ordinary density *den* and pressure *pres* correspond to (11.15a) and (11.15b). We have only printed these once; although we have had to regenerate them three times 'behind the scenes'. We give the versions for $u = 0$ because they are the equations that are actually solved for various purposes.

Summarising: The first solution gives the full formulae for a_1, a_2, a_3, a_4, k in terms of $\Lambda, p_0, \rho_0, H_0, c$ and χ; see Appendix C. The second solution gives the formulae for the empty model universe. The third solution gives numerical values. The fourth solution brings in the extra equation (13.8/1a) and gives numerical values. For the calculation (13.11) we have chosen the second solution (enclosed in brackets {}) because that has Λ negative and ρ_0 positive. The universal constant Λ has increased in magnitude over the value given in (13.6b) and the ordinary density ρ_0 likewise. The ordinary density value assumed at (13.6a) is approximately the closure value.

Appendix A- If Level One Of The Hierarchy Is Satisfied

A1. Quantisation

The first level in the scalar hierarchy is

(A1.1) $\dot{\theta} = \theta_{,j}\,\dot{q}^j; \quad j = 1, 2, ...n_c \equiv n_d n_p;$ Einstein convention in force

where n_d is the dimension of the flat space where quantum phenomena take place, n_p is the number of particles in that space, θ is an arbitrary pure, real function of *all* the real coordinates $q^1, q^2, ..., q^{n_c}$ and

(A1.2) $$\theta_{,j} \equiv \frac{\partial \theta}{\partial q^j}$$

The dot denotes differentiation with respect to time. All variables are continuous.

Quantising (A1.1), replacing differentials by commutators and using the product rule (see below (A1.6b))

$$\text{(A1.3a)} \quad \frac{1}{i\hbar}(\Theta H - H\Theta) = \frac{1}{2}(\Theta_{,j}\frac{1}{i\hbar}(Q^j H - HQ^j) + \frac{1}{i\hbar}(Q^j H - HQ^j)\Theta_{,j})$$

where H is the Hamiltonian (Hermitian) operator and

$$\text{(A1.3b)} \quad h \to H; \quad \theta \to \Theta; \quad \theta_{,j} \to \Theta_{,j} \equiv \frac{\partial \Theta}{\partial Q^j}; \quad q^j \to Q^j$$

where $\to$ means 'real observable corresponding to Hermitian operator'. The operator Θ is pure in all the coordinate operators Q^j. The scalar $\hbar$ has the approximate value

$$\text{(A1.3c)} \quad \hbar = \frac{6.626075\times10^{-34}}{2\pi}\ Js = 1.054573\times10^{-34}\ Js$$

The differentiation at (A1.3b) is purely formal and algebraic. Note that also that, with this proviso, and restricting the operators Θ and H to a polynomials

$$\text{(A1.4a)} \quad \Theta_{,j} \equiv \frac{\partial \Theta}{\partial Q^j} = \frac{1}{i\hbar}(\Theta P_j - P_j \Theta); \quad H^{,j} \equiv \frac{\partial H}{\partial P_j} = \frac{1}{i\hbar}(Q^j H - HQ^j)$$

where

$$\text{(A1.4b)} \quad Q^j P_k - P_k Q^j = i\hbar\delta^j_k I; \quad P_j P_k = P_k P_j; \quad Q_j Q_k = Q_k Q_j; \quad p_j \to P_j$$

So, in a more compact notation (A1.3a) can be written,

$$\text{(A1.5)} \qquad \frac{1}{i\hbar}(\Theta H - H\Theta) = \frac{1}{2}(\Theta_{,j}H^{,j} + H^{,j}\Theta_{,j})$$

where remember that θ is arbitrary and $\theta \to \Theta$.

We need to introduce more notation:

$$AB - BA \equiv \left[A, B\right]$$

$$\text{(A1.6a)} \quad \frac{1}{i\hbar}\left[A, B\right] \equiv \lfloor A, B \rfloor; \quad AB + BA \equiv (A, B);$$

$$\frac{1}{2}(AB + BA) \equiv \{A, B\} = \{AB\}$$

where A and B are any linear operators whatsoever. Generalising (A1.6a)

$$\text{(A1.6b)} \quad \frac{1}{(i\hbar)^2}(A(BC - CB) - (BC - CB)A)) \equiv \frac{1}{(i\hbar)^2}[A, B, C] \equiv \lfloor A, B, C \rfloor;$$

$$\{A, B, C, D....\} \equiv \{ABCD\} \equiv \frac{1}{n!}\sum_{perm} ABCD$$

where the product rule (second line of (A1.6b)) applies to quantum mechanics

$$\text{(A1.6c)} \quad a \to A; \quad b \to B; \quad c \to C; \quad d \to D; \quad abcd \to \{A, B, C, D\}$$

For example (A1.5) can be written

$$\text{(A1.7)} \qquad -\lfloor H, \Theta \rfloor = \{\Theta_{,j}, H^{,j}\}$$

The commas are redundant in brackets of type $\{\ \}$. If any of the operators are zero in brackets of the types $\lfloor\ \rfloor, \{\ \}, [\ ]$ then the bracket is zero.

The above notation as it appears in (A1.4) needs investigation: Let X be an Hermitian polynomial mixture of all the coordinates Q^j and all the momenta P_k that is

$$\text{(A1.8)} \qquad X \equiv X_1 + Y_1 + \{X_2 Y_2\} + \{X_3 Y_3 Z_3\} + ...$$

where X_n and Z_n are pure in all the coordinates and Y_n is pure in the momenta; therefore X_n, Y_n and Z_n are Hermitian. Then it is stated in [1] that if

(A1.9a) $$Q^j P_k - P_k Q^j = i\hbar \delta_k^j I; \quad P_j P_k = P_k P_j; \quad Q_j Q_k = Q_k Q_j;$$
$$q^j \to Q^j; \quad p_k \to P_k$$

then, where X is any operator function of all the coordinates and all the momenta,

(A1.9b) $$X_{,j} \equiv \frac{\partial X}{\partial Q^j} = \lfloor X, P_j \rfloor; \quad X^{,j} \equiv \frac{\partial X}{\partial P_j} = \lfloor Q^j, X \rfloor$$

where

(A1.9c) $$x_n(\underline{q}) \to X_n(\underline{Q}); \quad y_n(\underline{p}) \to Y_n(\underline{P}); \quad x \to X; \quad n = 1, 2, ...$$

But it is suggested here that X must have a certain form in order that it can satisfy (A1.9) for certain. A *partial proof* goes as follows:

Put

(A1.10a) $$X = X_n = (Q^j)^n \Rightarrow (X)_{,j} = \lfloor X, P_j \rfloor = n(Q^j)^{n-1}; \quad X^{,j} = 0$$

Prove by induction or otherwise that

(A1.10b) $$\frac{1}{i\hbar}(X_n P_j - P_j X_n) = n(Q^j)^{n-1}$$

This proves that $X = X_n$ can be pure polynomial in all the coordinates and that the first part of (A1.9b) is satisfied.

Put

(A1.11a) $X = Y_m = (P_j)^n \Rightarrow X_{,j} = 0; \quad X^{,j} = \frac{1}{i\hbar}(Q^j X - XQ^j) = n(P_j)^{n-1}$

Prove by induction or otherwise that

(A1.11b) $\frac{1}{i\hbar}(Q^j Y_m - Y_m Q^j) = n(P_j)^{n-1}$

This proves that $X = Y_m$ can be pure polynomial in all the momenta and that the second part of (A1.9b) is satisfied.

Put

(A1.12a) $X = \{X_n Y_m\}$

We get ostensibly

$$
\begin{aligned}
X_{,j} &\equiv \frac{\partial X}{\partial Q^j} = \frac{1}{2}\left(\frac{\partial X_n}{\partial Q^j} Y_m + Y_m \frac{\partial X_n}{\partial Q^j}\right); \quad \text{Partial Differentiation?} \\
&= \frac{1}{2}\left(\lfloor X_n, P_j \rfloor Y_m + Y_m \lfloor X_n, P_j \rfloor\right); \quad \text{Quantum Mechanics?} \\
X^{,k} &\equiv \frac{\partial X}{\partial P_k} = \frac{1}{2}\left(\frac{\partial Y_m}{\partial P_k} X_n + X_n \frac{\partial Y_m}{\partial P_k}\right); \quad \text{Partial Differentiation?} \\
&= \frac{1}{2}\left(\lfloor Q^k, Y_m \rfloor X_n + X_n \lfloor Q^k, Y_m \rfloor\right); \quad \text{Quantum Mechanics?}
\end{aligned}
\tag{A1.12b}
$$

To show that (A1.9b) is satisfied, in its entirety, we must have got the partial differentiation right; that is it is assumed here that the product rule for formal partial differentiation applies to Hermitian operators as it does to scalar functions. This is debatable. We have already shown that

(A1.13) $(X_n)_{,j} = \lfloor (X_n), P_j \rfloor; \quad (Y_m)^{,k} = \lfloor Q^k, (Y_m) \rfloor$

We cannot go any further without resolving the queries in (A1.12).

A2. H Is Quadratic In The P_j

Now (A1.3a) is linear in H ; so the various terms in H will linearly superpose providing that the coefficients are constant. Suppose that

(A2.1) $$H \equiv \frac{1}{2}(A^k(\underline{Q})P_k + P_k A^k(\underline{Q})) + B(\underline{Q}) = \{A^k, P_k\} + B$$ Hermitian

where $\underline{Q}$ denotes that the A^k and B are pure operator functions of all the coordinate operators and therefore Hermitian. Then

(A2.2) $$H^{,l} = A^l$$ See (A1.4a)

So that the LHS of (1.5) is equal to

(A2.3a)
$$\begin{aligned}
&\frac{1}{i\hbar}(\Theta H - H\Theta) \\
&= \frac{1}{i\hbar}(\Theta\{A^k, P_k\} - \{A^k, P_k\}\Theta); \quad B\Theta = \Theta B \\
&= \frac{1}{2}\lfloor \Theta, P_j \rfloor A^j + \frac{1}{2}A^j \lfloor \Theta, P_j \rfloor; \quad A^k\Theta = \Theta A^k \\
&= \frac{1}{2}(\Theta_{,j}\, A^j + A^j \Theta_{,j})
\end{aligned}$$

which is obviously equal, in this case, to the RHS of (1.5); see (2.2). So H , given by (A2.1), satisfies (A1.5).

Now suppose that we give up the tensor notation and define

(A2.4a) $$H \equiv P^n; \quad n = 1, 2, 3.... \Rightarrow H' \equiv \frac{\partial H}{\partial P} = nP^{n-1}$$ Hermitian

where

(A2.4b) $$QP - PQ = i\hbar I$$

and I is the unit operator. We also define

(A2.4c) $\Theta' \equiv \frac{\partial \Theta}{\partial Q} = \frac{1}{i\hbar}(\Theta P - P\Theta); \quad H' = \frac{\partial H}{\partial P} = \frac{1}{i\hbar}(QH - HQ);$ See (A1.4a)

Therefore

$$\frac{1}{i\hbar}(\Theta H - H\Theta) = \frac{1}{i\hbar}(\Theta P^n - P^n\Theta)$$

(A2.5a) $$= \frac{(\Theta P - P\Theta)}{i\hbar} P^{n-1} + P^{n-1}\frac{(\Theta P - P\Theta)}{i\hbar} + \frac{P\Theta P^{n-1} - P^{n-1}\Theta P}{i\hbar})$$

$$= \Theta' P^{n-1} + P^{n-1}\Theta' + \frac{P\Theta P^{n-1} - P^{n-1}\Theta P}{i\hbar}$$

(A2.5b) $$\frac{1}{2}(\Theta' H' + H'\Theta') = \frac{n}{2}(\Theta' P^{n-1} + P^{n-1}\Theta')$$

The quantities (2.5a/b) are equal only if

(A2.5c) $n = 2$

and that being so (2.4a) satisfies the appropriate version of (1.5) and

(A2.5d) $$\frac{1}{i\hbar}(\Theta H - H\Theta) = \frac{1}{2}(\Theta' H' + H'\Theta')$$

Hence H is *generally quadratic* in P ; see (A1.3a/A2.3a).

Now return to the tensor notation with the Einstein convention. Suppose that

(A2.6a) $H = \{G^{jk}(\underline{Q}), P_j P_k\} + P_j F^{jk}(Q) P_k$ Hermitian

where

(A2.6b)
$$g^{jk}(\underline{q}) \to G^{jk}(\underline{Q}); \quad g^{jk}(\underline{q}) = g^{kj}(\underline{q});$$
$$f^{jk}(\underline{q}) \to F^{jk}(\underline{Q}); \quad f^{jk}(\underline{q}) = f^{kj}(\underline{q})$$

where the g's and the f's are real free functions of all the coordinates. Now if X is a polynomial operator pure in all the coordinate operators

(A2.7)
$$X_{,jk} = \frac{1}{(i\hbar)^2}(XP_jP_k + P_jP_kX - 2P_jXP_k) \Rightarrow P_jXP_k = \{X, P_jP_k\} - \frac{(i\hbar)^2}{2}X_{,jk}$$

We have

(A2.8a)
$$\begin{aligned} &2P_jF^{jk}P_k = -P_j(P_kF^{jk} - F^{jk}P_k) + (P_kF^{jk} - F^{jk}P_k)P_j \\ &+P_jP_kF^{jk} + F^{jk}P_kP_j \\ &= i\hbar(P_kF^{jk}_{,j} - F^{jk}_{,j}P_k) + P_jP_kF^{jk} + F^{jk}P_kP_j \\ &= -(i\hbar)^2F^{jk}_{,jk} + P_jP_kF^{jk} + F^{jk}P_kP_j \end{aligned}$$

See (1.4a)

where

(A2.8b) $F^{jk}_{,j,k} \equiv F^{jk}_{,jk}$

because F^{jk} is a pure function of the $\underline{Q}$. It follows that the term $P_jF^{jk}P_k$ (see (A2.6a)) can be subsumed into the term $\{G^{jk}, P_jP_k\}$ and the term B if there is one (see (A2.6a) and (A2.1)). Therefore we may consider only

(A2.8c) $H = \{G^{jk}, P_jP_k\}$ See (A2.6a)

In practice we use, in the main text, only

(A2.9) $n_c = 4$

So, by an appropriate choice of coordinates, we may always make the matrix G^{ij}
diagonal; there are only three unique G^{ij} such that $i \neq j$ when $n_c = 4$ whereas when $n_c = 5$ there are ten. Another consequence of the linearity of (A1.5/7), with respect to H, is that the argument from (A2.4a) to (A2.6d) may be adapted to prove H is quadratic ($n = 2$) in the case

(A2.10) $H = \{G^{jj}, (P_j)^n\}$; Einstein convention in force

Therefore the most general form of the Hamiltonian operator allowed by the first level of the hierarchy is

(A2.11a) $H = s\{G^{jk}, P_j P_k\} + \{A^k, P\} + B$

where G^{jk}, A^k and B are pure operator functions of all the coordinates and therefore Hermitian and s is scalar.

(A2.11b) $g^{jk}(\underline{q}) \to G^{jk}(\underline{Q}); \quad g^{jk}(\underline{q}) = g^{kj}(\underline{q});$
$a(\underline{q}) \to A(\underline{Q}); \quad b(\underline{q}) \to B(\underline{Q})$

A3. Hamilton's Equations

If the first level in the hierarchy is satisfied the Hamiltonian operator H is quadratic in the momenta $\underline{P}$; let us assume that this is so. The classical Hamilton's equations are

(A3.1a)

$$\dot{p}_j = -\frac{\partial h}{\partial q^j}; \quad \dot{q}^k = \frac{\partial h}{\partial p_k} \Rightarrow \dot{p}_l = -p_j p_k g^{jk}_{,l} - p_r f^r_{,l} - v_{,l}; \quad \dot{q}^l = 2p_l g^{jl} + f^l$$

where

(A3.1b) $h = p_j p_k g^{jk} + p_r f^r + v$

is the scalar Hamiltonian, q^k coordinates and p_j the momenta and dot denotes differentiation with respect to time; the coefficients $g^{jk} = g^{kj}$, f^r and v are pure functions of all the scalar coordinates. The quantum mechanical version of (A3.1a) is

(A3.2a) $\dot{P}_l = -\{P_j P_k, G^{jk}_{,l}\} - \{P_r, F^r_{,l}\} - V_{,l}$; $\dot{Q}^l = 2\{P_l, G^{jl}\} + F^l$

where

(A3.2b) $H = \{P_j P_k, G^{jk}\} + \{P_r, F^r\} + V$

and $G^{jk} = G^{kj}, F^r$ and V are pure polynomials in all the coordinates. It is stated in [2] that things can be arranged so that Hamilton's equations are true in both quantum mechanics and classical mechanics; this is not generally true as the above argument shows. It is only true when the space is flat, $f^r = 0$ and the coordinates are Cartesian.

References

[1] 'Quantum Mechanics' R.A. Newing, J. Cunningham, Oliver & Boyd 1967
[2] 'Draft 12 Cosmological Theories of the Extra Terms', A.M. Deakin (available from the author)

Appendix B - If The First Two Levels Of The Hierarchy Are Satisfied - The Theta Equation And GTE

B1. Introduction

It is desirable that H and Θ should satisfy as many constraints as possible, consecutively, beginning with constraint 1. We have already proved (in Appendix A) that if H and $\theta \to \Theta$ satisfy constraint 1 then H is quadratic in the P_j, $j = 1, 2, ...n_c$

(B1.1a)

$$H = \mathsf{K}\left\{G^{uv}, P_u P_v\right\} + \left\{F^j, P_j\right\} + V; \quad \mathsf{K} \text{ scalar}; \quad \{\mathrm{A},\mathrm{B}\} = \tfrac{1}{2}(AB + BA)$$

where the operators G^{uv}, F^j and V pure in the coordinate operators Q^k. We are only interested in the gravitational case where

(B1.1b) $F^j = O$ and $V = O$

Therefore

(B1.1c) $H = \mathsf{K}\left\{G^{uv}, P_u P_v\right\}$

where in the coordinate representation

(B1.2a) $P_j = -i\hbar \dfrac{\partial}{\partial q^j}; \quad \mathrm{Q}^{\mathrm{k}} = q^k I; \quad i\hbar\delta^j_k I = -i\hbar\left(q^j \dfrac{\partial}{\partial q^k} - \dfrac{\partial}{\partial q^k} q^j \right);$

(B1.2b) $G^{uv}(\underline{Q}) \equiv g^{uv}(\underline{q})I = G^{vu}(\underline{Q}); \qquad F^j(\underline{Q}) \equiv f^j(\underline{q})I; \quad V(\underline{Q}) \equiv v(\underline{q})I$

where $\underline{q} \to \underline{Q}$ denotes the aggregate of the $q^j \to Q^j$. In view of the first sentence of this Introduction we consider, at the least, constraint 2 as well as constraint 1 should be satisfied; we expect $\theta \to \Theta$ to be restricted thereby.

B2. Constraints 1 And 2

With the notation

(B2.1a) $\lfloor A, B \rfloor \equiv \dfrac{1}{i\hbar}(AB - BA); \quad \{A, B, ...\} \equiv \dfrac{1}{n!} \sum_{perm} (A, B, ...);$

(B2.1b) $A^{:j} \equiv \left\lfloor Q^j, A \right\rfloor = \dfrac{\partial A}{\partial P_j}; \quad A_{,j} \equiv \left\lfloor A, P_j \right\rfloor = \dfrac{\partial A}{\partial Q^j}; \quad j = 1, 2, ...n_c$

(B2.1c)

$$\frac{dA}{dt} \equiv -\frac{1}{i\hbar}(HA - AH) = -\lfloor H, A \rfloor; \quad \frac{d^2A}{dt^2} \equiv \lfloor H, \lfloor H, A \rfloor \rfloor \equiv \lfloor H, H, A \rfloor$$

Constraint 1 is $(()^{:j} \equiv ()_{,j})$

(B2.2a) $Z_1 \equiv \{H^{:j}, \Theta_{,j}\} = -\lfloor H, \Theta \rfloor = \lfloor \Theta, H \rfloor$; Einstein convention in force

Constraint 2 is

(B2.2b) $Z_2 \equiv \{\lfloor H, H^{:j} \rfloor, \Theta_{,j}\} + \{H^{:j}, H^{:k}, \Theta_{,j,k}\} = \lfloor H, H, \Theta \rfloor$

By combining this with (B2.2a) we can remove all explicit reference to the $H^{:j}$ and to H. We obtain thereby an operator equation involving only the derivations of Θ. In the coordinate representation this reduces to a fourth order PDE satisfied by θ. The PDE contains no reference to either the f^u or v. It allows us, in principle, to calculate functions $\theta(\underline{q})$ that satisfy both constraints 1 and 2 given the functions g^{uv}.

Differentiate (B2.2a) with respect t to produce

(B2.3) $$-\lfloor H, Z_1 \rfloor = -\lfloor H, \{H^{:j}, \Theta_{,j}\} \rfloor = -\lfloor H, \lfloor H, \Theta \rfloor \rfloor = -\lfloor H, H, \Theta \rfloor$$

Add (B2.3) from (B2.2b) to get

(B2.4)

$$Z_2 - \lfloor H, Z_1 \rfloor \Rightarrow \{\lfloor H, H^{:j} \rfloor, \Theta_{,j}\} - \lfloor H, \{H^{:j}, \Theta_{,j}\} \rfloor + \{H^{:j}, H^{:k}, \Theta_{,j,k}\} = O$$

Now we have

(B2.5a) $$\Theta_{,j} \equiv \frac{\partial \Theta}{\partial Q^j} = \frac{1}{i\hbar}(\Theta P_j - P_j \Theta) = \lfloor \Theta, P_j \rfloor$$

(B2.5b) $$H^{:j} \equiv \frac{\partial H}{\partial P_j} = \frac{1}{i\hbar}(Q^j H - HQ^j) = \left\lfloor Q^j, H \right\rfloor = \dot{Q}^j; \quad \text{See (B2.1)}$$

(B2.6) $$\begin{aligned} &\left\{\left\lfloor H, H^{:j} \right\rfloor, \Theta_{,j}\right\} - \left\lfloor H, \left\{H^{:j}, \Theta_{,j}\right\}\right\rfloor \\ &= \left\{\ddot{Q}^j, \Theta_{,j}\right\} - \left\{\dot{H}^j, \Theta_{,j}\right\} = \left\{\ddot{Q}^j, \Theta_{,j}\right\} - \left\{\ddot{Q}^j, \Theta_{,j}\right\} = O; \quad \dot{H} = O \end{aligned}$$

The remaining term is

(B2.7a)
$$\left\{H^{:j}, H^{:k}, \Theta_{,j,k}\right\} = \left\{H^{:j}, \left\{H^{:k}, \Theta_{,j,k}\right\}\right\} = O$$
$$\Rightarrow \frac{1}{12}(H^{:j} H^{:k} \Theta_{,j,k} + \Theta_{,j,k} H^{:j} H^{:k} - 2H^{:j} \Theta_{,j,k} H^{:k}) = O; \quad \Theta_{,j,k} = \Theta_{,k,j}$$

written in full. An alternative to this is

(B2.7b) $$\frac{(i\hbar)^2}{12}\left[H^{:j}, \left[H^{:k}, \Theta_{,j,k}\right]\right] = O; \quad \Theta_{,j,k} = \Theta_{,k,j}$$

Because

(B2.8) $$\left\lfloor H^{:l}, A \right\rfloor = 2\mathsf{K} G^{ul} A_{,u}; \quad AQ^u = Q^u A; \text{ see (B2.1a)}$$

the result (B2.7b) is identical to

(B2.9a) $$(i\hbar)^2 \frac{\mathsf{K}}{6}\left[H^{:j}, G^{uk} \Theta_{,j,k,u}\right] = O$$

giving, upon further application of (B2.8),

(B2.9b) $$-\frac{\hbar^2 \mathsf{K}^2}{3} G^{vj} (G^{uk} \Theta_{,j,k,u})_{,v} = O; \quad j,k,u,v = 1,2,...n_c$$ Operator Theta Equation

The numerical factor $-\hbar^2 K^2/3$ can, of course, be cancelled; we retain this factor at (B2.9b) because the operator on the LHS is the imbalance across (B2.2b). In the coordinate representation

(B2.10a) $G^{uv} \equiv g^{uv}(\underline{q})I; \quad g^{uv} = g^{vu}; \quad \Theta = \theta(\underline{q})I$

and (B2.9b) reduces to the PDE

(B2.10b) $g^{vj}(g^{uk}\theta_{,jku})_{,v} = 0; \quad j,k,u,v = 1,2,...n_c$; Scalar Theta Equation

Notice that the theta equation does not contain the functions f^j and v. Further the theta equation is not a classical approximation; it is a purely QM result

That θ satisfies the PDE (B2.10b) raises an immediate issue: Quadratic operator H derives from constraint 1 on the assumption that θ is arbitrary; but it cannot be truly arbitrary if it satisfies (B2.10b). At most it is the general solution of given a particular n_c - space C. So is the quadratic form of the operator H valid? Yes! Solutions of (B2.10b). must be subject to complicated boundary conditions; thus θ is sufficiently arbitrary for the quadratic solution of constraint 1 to follow.

The scalar theta equation (B2.10b) can be regarded as a field equation for theta (subject to possible modification by higher constraints at levels 3 and above). Because the g^{lm} the f^l and v inform the Hamiltonian they are all candidates for θ. We thus have three versions of (B2.10b) that are putative field equations for the g^{lm} the f^l and v:

(B2.11) $g^{vj}(g^{uk}g^{lm}{}_{,jku})_{,v} = 0; \quad g^{vj}(g^{uk}f^l{}_{,jku})_{,v} = 0; \quad g^{vj}(g^{uk}v_{,jku})_{,v} = 0$

Recall that, according to the quantization axioms, these equations are true only in a flat space using flat (e.g., Cartesian) coordinates.

We support the (unproven) conjecture the if the first two levels of the hierarchy of constraints is satisfied then that is all that is required to discuss conventional CM.

Appendix C- Results For Calculations (13.4) To (13.12)- Taken From The Output Of Maple 16

The original equations are

$$A := a0 + a1\,u + a2\,u^2 + a3\,u^3 + a4\,u^4 + a5\,u^5$$

$$p := p0 + p1\,u$$

$$\rho := \rho0 + \rho1\,u$$

$$\begin{aligned} o4 := {} & \left(a0 + a1\,u + a2\,u^2 + a3\,u^3 + a4\,u^4 + a5\,u^5\right)^2 \left(2\,a2 + 6\,a3\,u\right. \\ & \left. + 12\,a4\,u^2 + 20\,a5\,u^3\right)^2 + 4\,k\left(a1 + 2\,a2\,u + 3\,a3\,u^2\right. \\ & \left. + 4\,a4\,u^3 + 5\,a5\,u^4\right)^2 + 4\left(a1 + 2\,a2\,u + 3\,a3\,u^2 + 4\,a4\,u^3\right. \\ & \left. + 5\,a5\,u^4\right)^4 - 5\left(a0 + a1\,u + a2\,u^2 + a3\,u^3 + a4\,u^4\right. \\ & \left. + a5\,u^5\right)\left(a1 + 2\,a2\,u + 3\,a3\,u^2 + 4\,a4\,u^3 + 5\,a5\,u^4\right)^2 \left(2\,a2\right. \\ & \left. + 6\,a3\,u + 12\,a4\,u^2 + 20\,a5\,u^3\right) + \left(a0 + a1\,u + a2\,u^2\right. \\ & \left. + a3\,u^3 + a4\,u^4 + a5\,u^5\right)^2 \left(a1 + 2\,a2\,u + 3\,a3\,u^2 + 4\,a4\,u^3\right. \\ & \left. + 5\,a5\,u^4\right)\left(6\,a3 + 24\,a4\,u + 60\,a5\,u^2\right) + \left(a0 + a1\,u\right. \\ & \left. + a2\,u^2 + a3\,u^3 + a4\,u^4 + a5\,u^5\right)^3 \left(24\,a4 + 120\,a5\,u\right) \end{aligned}$$

$$\begin{aligned} den := {} & k + \left(a1 + 2\,a2\,u + 3\,a3\,u^2 + 4\,a4\,u^3 + 5\,a5\,u^4\right)^2 \\ & - \frac{1}{3}\left(\chi\,c^2\left(\rho0 + \rho1\,u\right) + \Lambda\right)\left(a0 + a1\,u + a2\,u^2 + a3\,u^3\right. \\ & \left. + a4\,u^4 + a5\,u^5\right)^2 \end{aligned}$$

$$pres := k + \left(a1 + 2\,a2\,u + 3\,a3\,u^2 + 4\,a4\,u^3 + 5\,a5\,u^4\right)^2 + 2\left(a0 + a1\,u + a2\,u^2 + a3\,u^3 + a4\,u^4 + a5\,u^5\right)\left(2\,a2 + 6\,a3\,u + 12\,a4\,u^2 + 20\,a5\,u^3\right) + \left(\chi\,(p0 + p1\,u) - \Lambda\right)\left(a0 + a1\,u + a2\,u^2 + a3\,u^3 + a4\,u^4 + a5\,u^5\right)^2$$

Here and now

$$u := 0$$

$$o10 := 90\,a0^2\,a1\,a3 + 44\,a0^2\,a2^2 + 72\,a0^3\,a4 + 4\,k\,a1^2 - 4\,a1^4 + 8\,k^2 - 14\,a0\,a1^2\,a2 - 8\,a0\,a2\,k$$

$$o40 := 4\,a0^2\,a2^2 + 4\,k\,a1^2 + 4\,a1^4 - 10\,a0\,a1^2\,a2 + 6\,a0^2\,a1\,a3 + 24\,a0^3\,a4$$

$$den0 := k + a1^2 - \frac{1}{3}\left(\chi\,c^2\,\rho0 + \Lambda\right)a0^2$$

$$pres0 := k + a1^2 + 4\,a0\,a2 + \left(\chi\,p0 - \Lambda\right)a0^2$$

$$a0 := 1$$

$$a1 := \frac{H}{c}$$

The rest of the formulae

$$\left\{a2 = -\frac{1}{12}\chi c^2 \rho 0 + \frac{1}{6}\Lambda - \frac{1}{4}\chi p0,\ a3 = \frac{1}{108}\frac{1}{Hc}\left(15 c^2 H^2 \chi \rho 0 - 2\rho 0^2 \chi^2 c^6 - c^4 \Lambda \rho 0 \chi - 3 c^4 \chi^2 \rho 0\, p0 - 2 c^2 \Lambda^2 + 3 c^2 \Lambda \chi p0 - 3 c^2 \chi^2 p0^2 + 6 H^2 \Lambda + 9 H^2 \chi p0\right),\ a4 = -\frac{1}{288}\frac{1}{c^2}\left(\chi\left(36 c^2 H^2 \rho 0 - \rho 0^2 \chi c^6 - 2 c^4 \Lambda \rho 0 - 2 c^2 \Lambda p0 + c^2 \chi p0^2 + 36 H^2 p0\right)\right),\ k = -\frac{1}{3}\frac{3H^2 - \chi c^4 \rho 0 - \Lambda c^2}{c^2}\right\}$$

Results for an empty model universe

$$\rho 0 := -\frac{\Lambda}{\chi c^2}$$

$$p0 := \frac{\Lambda}{\chi}$$

$$\left\{a2 = 0,\ a3 = 0,\ a4 = 0,\ k = -\frac{H^2}{c^2}\right\}$$

Numerical results (13.5)- Assumptions

$$H := 2.055000000\,10^{-18}$$

$$c := 2.997900000\,10^{8}$$

$$aa1 := 6.854798359\,10^{-27}$$

$$A := a0 + a1\,u + a2\,u^2 + a3\,u^3 + a4\,u^4 + a5\,u^5$$

$$p := p0 + p1\,u$$

$$\rho := \rho 0 + \rho 1\,u$$

$$u := 0$$

$$o10 := 90\,a0^2\,a1\,a3 + 44\,a0^2\,a2^2 + 72\,a0^3\,a4 + 4\,k\,a1^2 - 4\,a1^4 + 8\,k^2 - 14\,a0\,a1^2\,a2 - 8\,a0\,a2\,k$$

$$o40 := 4\,a0^2\,a2^2 + 4\,k\,a1^2 + 4\,a1^4 - 10\,a0\,a1^2\,a2 + 6\,a0^2\,a1\,a3 + 24\,a0^3\,a4$$

$$den0 := k + a1^2 - \frac{1}{3}\left(\chi\,c^2\,\rho0 + \Lambda\right)a0^2$$

$$pres0 := k + a1^2 + 4\,a0\,a2 + \left(\chi\,p0 - \Lambda\right)a0^2$$

$$k := 0$$

$$\rho0 := 8.000000000\,10^{-27}$$

$$p0 := 0$$

$$a0 := 1$$

$$a1 := \frac{H}{c}$$

$$\chi := 2.076100000\,10^{-43}$$

Results for (13.5)

$$\left\{\Lambda = -8.305220790\,10^{-54},\ a2 = -1.382337033\,10^{-53},\ a3 = 8.024424778\,10^{-80},\ a4 = -8.079849533\,10^{-106}\right\}$$

Numerical results for (13.6)- Assumptions

$$A := a0 + a1\,u + a2\,u^2 + a3\,u^3 + a4\,u^4 + a5\,u^5$$

$$p := p0 + p1\,u$$

$$\rho := \rho0 + \rho1\,u$$

$$u := 0$$

$$o10 := 90\,a0^2\,a1\,a3 + 44\,a0^2\,a2^2 + 72\,a0^3\,a4 + 4\,k\,a1^2 - 4\,a1^4 + 8\,k^2 - 14\,a0\,a1^2\,a2 - 8\,a0\,a2\,k$$

$$o40 := 4\,a0^2\,a2^2 + 4\,k\,a1^2 + 4\,a1^4 - 10\,a0\,a1^2\,a2 + 6\,a0^2\,a1\,a3 + 24\,a0^3\,a4$$

$$den0 := k + a1^2 - \frac{1}{3}\left(\chi\,c^2\,\rho0 + \Lambda\right)a0^2$$

$$pres0 := k + a1^2 + 4\,a0\,a2 + \left(\chi\,p0 - \Lambda\right)a0^2$$

$$k := 1.000000000\,10^{-54}$$

$$\rho0 := 8.000000000\,10^{-27}$$

$$p0 := 0$$

$$a0 := 1$$

$$a1 := \frac{H}{c}$$

Results for (13.6)

$$\{\Lambda = -5.305220790\,10^{-54},\ a2 = -1.332337033\,10^{-53},\ a3 = 8.089213313\,10^{-80},\ a4 = -8.048751616\,10^{-106}\}$$

Numerical results for (13.9)

Solutions of quartic

$$2.500253192\,10^{26},\ -2.410983337\,10^{25} + 2.201628897\,10^{26}\,\mathrm{I},\ -1.013029447\,10^{26},\ -2.410983337\,10^{25} - 2.201628897\,10^{26}\,\mathrm{I}$$

$$age := 2.642936097\,10^{10}$$

Numerical Results for (13.11)- Assumptions

$$A := a0 + a1\,u + a2\,u^2 + a3\,u^3 + a4\,u^4 + a5\,u^5$$

$$p := p0 + p1\,u$$

$$\rho := \rho0 + \rho1\,u$$

$$u := 0$$

$$o10 := 90\, a0^2\, a1\, a3 + 44\, a0^2\, a2^2 + 72\, a0^3\, a4 + 4\, k\, a1^2 - 4\, a1^4 + 8\, k^2 - 14\, a0\, a1^2\, a2 - 8\, a0\, a2\, k$$

$$o40 := 4\, a0^2\, a2^2 + 4\, k\, a1^2 + 4\, a1^4 - 10\, a0\, a1^2\, a2 + 6\, a0^2\, a1\, a3 + 24\, a0^3\, a4$$

$$den0 := k + a1^2 - \frac{1}{3}\left(\chi\, c^2\, \rho 0 + \Lambda\right) a0^2$$

$$pres0 := k + a1^2 + 4\, a0\, a2 + \left(\chi\, p0 - \Lambda\right) a0^2$$

$$k := 1.000000000\, 10^{-54}$$

$$p0 := 0$$

$$a0 := 1$$

Value of $a1$

$$aa1 := 6.854798359\, 10^{-27}$$

Age of the Universe

$$y := 1.296038477\, 10^{26}$$

(13.10)- Extra equation

$$AA := 1.888408242 + 1.679715734\, 10^{52}\, a2 + 2.176976222\, 10^{78}\, a3 + 2.821444947\, 10^{104}\, a4$$

Results- Only take Λ negative and $\rho 0$ positive

$$\{\Lambda = 8.079505311\, 10^{-52},\ a2 = 1.899905676\, 10^{-52},\ a3 = -2.554391430\, 10^{-78},\ a4 = 1.705295387\, 10^{-105},\ \rho 0 = -3.558575675\, 10^{-26}\},\ \{\Lambda = -2.077396280\, 10^{-52},\ a2 = -6.393197215\, 10^{-53},\ a3 = -9.633534651\, 10^{-80},\ a4 = -2.143628883\, 10^{-105},\ \rho 0 = 1.884930148\, 10^{-26}\}$$

Appendix D- The K Equation Revisited

D1. Geodesic, Canonical And Cartesian Coordinates

The letter [7] raises certain questions; this appendix is an effort answer them. A derivation of the K equation is also given .

The Cristoffel symbols are defined a follows (suffices run from 1 to n_c in this case):

(D1.1a) $[ij,k] \equiv \frac{1}{2}\left(g_{ik,j} + g_{jk,i} - g_{ij,k}\right)$; first kind [1, p. 26]

(D1.1b) $\Gamma^{l}_{ij} \equiv g^{lk}[ij,k]$; second kind

(D1.1c) Suffices $= [1, n_c]$

where $()_{,j}$ denotes *partial differentiation* with respect to the j^{th} coordinate, g_{ij} is an element of the covariant fundamental tensor and g^{lk} is an element of the contravariant fundamental tensor. The fundamental tensor is transparent to covariant differentiation [1]

(D1.2) $g_{jk;l} = 0$ and $(g^{jk})_{;l} = 0$

were $()_{;l}$ denotes *covariant differentiation* with respect to l^{th} coordinate. Covariant differentiation of a product follows the usual rule for partial differentiation [1]. The connection between covariant fundamental tensor g_{ij} and the contravariant fundamental tensor g^{uv} is

(D1.3) $g^{uv}g_{vw} = \delta^{u}_{w}$ Einstein convention is in force

where $\delta^{u}_{u} = 1$ and $\delta^{\mathrm{u}}_{\mathrm{v}} = 0$ for $u \neq v$ [1].

A coordinate is *geodesic* at pole *P* if

(D1.4) $[ij,k]=0$

at pole *P*. It follows that the Christoffel symbols of the first and second kinds vanish at the pole; therefore the first covaritant derivative is equal to the first partial derivative at *P* ; see (D1.1a/b) [1].

Coordinates that satisfy, at a pole *P*,

(D1.5a) $\Gamma^{a}_{bc,d}+\Gamma^{a}_{cd,b}+\Gamma^{a}_{db,c}=0;$; constrains the $g_{ab,cd}$; $n_c=4$

are said to be *canonical*. Does the condition (D1.5a) define a valid set of coordinates? To decide that we must investigate whether or not the conditions (D1.5a) could constrain the curvature of the space (defined by the Riemann-Christoffel tensor R^{u}_{vwx}) at *P*. Because Γ^{a}_{bc} is symmetrical in the suffices b,c it follows from (D1.5a) that

(D1.5b) $\Gamma^{a}_{cb,d}+\Gamma^{a}_{dc,b}+\Gamma^{a}_{bd,c}=0$

There are thus

(D1.6) $$n_c\left[n_c+n_c(n_c-1)+\frac{n_c(n_c-1)(n_c-2)}{3!}\right]=n_c^2(n_c^2+3n_c+2)/6$$

unique conditions which constrain the

(D1.7) $n_c^2(n_c+1)^2/4$

second derivatives of the g_{uv} at *P*. It follows that there are still

(D1.8) $$n_c^2(n_c+1)^2/4-n_c^2(n_c^2+3n_c+2)/6=n_c^2(n_c^2-1)/12$$

degrees of freedom. This is also the number of unique, independent elements of the curvature tensor R^{u}_{vwx} [1]. So the conditions (D1.5a/b) do not constrain the curvature at *P*.

The counting of the unique conditions at (D1.6) goes as follows: a takes n_c values independently of b,c and d. So, in the square brackets on the LHS of (1.6), we have the contribution as b,c and d vary. There are n_c cases for which $b=c=d$. There are $n_c(n_c-1)$ cases for which two of the suffices b,c,d are equal. Because, permutation does not increase the number of conditions when the suffices b,c,d all differ, there are in total $n_c(n_c-1)(n_c-2)/3!$ cases for which none of these suffices are equal.

Conditions (D1.5a) constrain the second derivatives of the fundamental tensor at *P*. But we have yet to constrain the g_{uv} themselves. The Schrodinger quantisation rules appear to require that definitions be couched in terms of flat, Cartesian coordinates $\underline{q}$. A curved manifold C' can be made flat at *P*, and approximately flat in a neighbourhood of *P*, by requiring that *P* is the pole of local *Cartesian geodesic* (CG) coordinates. Thus the requirements of the Schrodinger definitions are met, in a neighbourhood of a point *P*, if the coordinates are chosen to be CG with pole *P*. There may be convenience of calculation if, in addition, the coordinates are chosen to be *canonical* at *P*; the coordinates are then said to be CCG (Cartesian canonical geodesic) with pole *P*.

D2. Questions About The Gravitational Theta Equation

The gravitational theta equation (GTE) can be written

$$\text{(D2.1)} \qquad \begin{aligned} &g^{vj}(g^{uk}g^{lm}_{,jku})_{,v}=0; \quad ()_{,jkuv}\equiv()_{,j,k,u,v} \text{ by convention;} \\ &j,k,l,m,u,v=1,2,...n_c; \quad n_p=1; \quad n_c=4 \end{aligned}$$

where the coordinates and the space are defined as flat. We suppose that, in a curved Riemannian manifold C', the GTE holds only at the pole P of CG coordinates. Solutions of the GTE, as it stands, then approximate the manifold in the neighbourhood of *P*. The question arises: What tensor equation reduces to (D2.1) at the pole P in the coordinates chosen ? Such a tensor equation would be valid in any

coordinate system at all points of the manifold. Clive Kilmister derived a tensor equation that purports to satisfy this condition; but his argument is based on a version of the GTE that depends upon $g_{ab,cde}$. Is this version valid? He takes the theta equation

(D2.2) $\quad g^{vj}(g^{uk}\theta_{,jku})_{,v} = 0$

and simply defines

(D2.3) $\quad \theta \equiv g_{ab}$

to get a GTE

(D2.4) $\quad g^{vj}(g^{uk}g_{ab,jku})_{,v} = 0$; Kilmister GTE

whereas it is usual to define the GTE using the coefficients of the quadratic terms of the Hamiltonian

(D2.5) $\quad \theta \equiv g^{lm}$; see (D2.1)

We need to explore the relation between the LHSs of (D2.1) and (D2.4) to answer this question. Note that, in geodesic coordinates at the pole *P*, (D2.1) simplifies to

(D2.6) $\quad g^{vj}(g^{uk}g^{lm}_{,jku})_{,v} = g^{vj}(g^{uk}g^{lm}_{,jku})_{;v}$; canonical GTE in geodesic coordinates
$\qquad = g^{vj}g^{uk}g^{lm}_{,jkuv} = 0$

Whereas, in the same coordinates, (D2.4) simplifies to

(D2.7) $\quad g^{vj}g^{uk}g_{ab,jkuv} = 0$; Kilmister's form of the GTE in geodesic coordinates

Now choose the metric (2.7) (main text). We have

(D2.8a) $\quad g_{aa} \approx -1+2U; \quad g^{ll} \approx -1-2U; \quad |U| << 1; \quad a,l = [1,3]$

(D2.8b) $g_{aa,jkuv} \approx 2U_{,jkuv}; \quad g^{ll}_{,jkuvv} \approx -2U_{,jkuv}; \quad |U| << 1; \quad a,l = [1,3]$

(D2.8c) $n_p = 1; \quad n_c = 4$

where

(D2.9) $U_{,4} = 0;$ U depends upon the first three coordinates only

The canonical GTE (D2.6) becomes

(D2.10) $g^{vj} g^{uk} g^{ll}_{,jkuv} \approx -2 g^{vj} g^{uk} U_{,jkuv} = 0$

The Kilmister form of the GTE (D2.7) becomes

(D2.11) $g^{vj} g^{uk} g_{ab,jkuv} \approx 2 g^{vj} g^{uk} U_{,jkuv} = 0$

So the two forms of the GTE are equivalent.

But what of the condition attached to (D2.8a/b)? This surely means that we have proved equivalence only for *weak gravity*. On the contrary; however strong the gravity we can always choose the pole *P* of the CG coordinates so that

(D2.12) $U = 0 \Rightarrow U_{,j} = 0; \quad U_{jk} = 0; \quad U_{jkl} = 0; \quad U_{jklm} = 0$

at *P*; hence we can write exact equalites at (D2.10/11).

D3. Lemmas

We begin to derive the K equation with certain lemmas. What follows is an identity

(D3.1a) $g_{ab,c} = g_{pb} \Gamma^p_{ac} + g_{ap} \Gamma^p_{bc}$; [1, equ. (20.4) et seq., p.27] Identity

So we have in geodesic coordinates pole *P*,

(D3.2) $g_{ab,cd} = g_{pb}\Gamma^{p}_{ac,d} + g_{ap}\Gamma^{p}_{bc,d}$; in geodesic coordinates pole *P*

Now, also at *P* in geodesic coordinates, [1, p. 49 et seq.]

(D3.3) $R^{a}_{bcd} = \Gamma^{a}_{bd,c} - \Gamma^{a}_{bc,d}$; Riemann-Christoffel or curvature tensor [1]

where, from (D1.1b/D1.5a),

(D3.4) $-\Gamma^{a}_{bc,d} = \Gamma^{a}_{cd,b} + \Gamma^{a}_{db,c}$; CCG coordinates pole *P*

so that (D3.3) becomes

(D3.5) $R^{a}_{bcd} = 2\Gamma^{a}_{bd,c} + \Gamma^{a}_{cd,b}$; CCG coordinates pole *P*

Interchange b and c

(D3.6) $R^{a}_{cbd} = \Gamma^{a}_{bd,c} + 2\Gamma^{a}_{cd,b}$; CCG coordinates pole *P*

These two (D3.5/6) give

(D3.7) $3\Gamma^{a}_{cd,b} = 2R^{a}_{cbd} - R^{a}_{bcd}$; CCG coordinates pole *P*

But

(D3.8) $-R^{a}_{bcd} = R^{a}_{cdb} + R^{a}_{dbc} = -R^{a}_{cbd} + R^{a}_{dbc}$; [1, p. 50]

Hence,

(D3.9a) $\Gamma^{a}_{cd,b} = \frac{1}{3}(R^{a}_{cbd} + R^{a}_{dbc})$; CCG coordinates pole *P*

Therefore

(D3.9b) $$\begin{aligned}&[ij,m]_{,k} = (g_{lm}\Gamma^{l}_{ij})_{,k} = g_{lm}\Gamma^{l}_{ij,k}\text{; CCG coordinates pole } P\\ &= \tfrac{1}{3}(R_{mikj} + R_{mjki})\end{aligned}$$

So, using (D3.2/9a) and the properties of R_{acdb} [1, p. 51]

(D3.10a) $$\begin{aligned}&g_{ab,cd} = g_{pb}\Gamma^{p}_{ac,d} + g_{ap}\Gamma^{p}_{bc,d}\\ &= \tfrac{1}{3}[g_{pb}(R^{p}_{adc} + R^{p}_{cda}) + g_{pa}(R^{p}_{bdc} + R^{p}_{cdb})]\text{; CCG coordinates}\\ &= \tfrac{1}{3}(R_{badc} + R_{bcda} + R_{abdc} + R_{acdb})\\ &= \tfrac{1}{3}(R_{cbad} + R_{cabd})\end{aligned}$$

pole *P*

because R_{ijkl} is skew symmetric in ij and kl. Therefore, because by definition

(D3.10b) $g^{cd}R_{cbad} = R^{d}_{bad} = R_{ba} = R_{ab}$,

(D3.10c) $g^{cd}g_{ab,cd} = \tfrac{1}{3}g^{cd}(R_{cbad} + R_{cabd}) = \tfrac{2}{3}R_{ab}$; CCG coordinates pole *P*

Further it follows from (D3.10a) that

(D3.11) $$\begin{aligned}&g^{aj}_{,de} = -g^{jk}g^{ab}g_{bk,de}\\ &= -\tfrac{1}{3}g^{jk}g^{ab}(R_{dkbe} + R_{dbke})\text{; CCG coordinates pole } P\\ &= \tfrac{1}{3}g^{ab}(R^{j}_{dbe} + R^{j}_{ebd})\end{aligned}$$

We have now the formulae necessary to express the first derivatives of the Christoffel symbols and the second derivatives of the fundamental tensor in terms of the curvature tensor. Most of them were derived by Clive Kilmister.

D.4 The K Equation

We need, in the course of the final argument, to express $g^{ef}R_{ab,ef}$ in terms of tensors evaluated in suitable coordinates. By definition

(D4.1) $R_{ab;e} \equiv R_{ab,e} - \Gamma^{l}_{ae}R_{lb} - \Gamma^{l}_{be}R_{al}$; [1, p. 34] Identity/ definition

where ';' denotes covariant differentiation. Differentiating covariantly again, but expressing the result using geodesic coordinates with pole *P*,

(D4.2a) $R_{ab;ef} = R_{ab,ef} - \mu_{abef}$; in geodesic coordinates pole *P*

where, in the same coordinates,

(D4.2b) $\mu_{abef} \equiv \Gamma^{l}_{ae,f}R_{lb} + \Gamma^{l}_{be,f}R_{al}$; in geodesic coordinates pole *P*

Therefore

(D4.3a) $g^{ef}R_{ab,ef} = g^{ef}R_{ab;ef} + M_{ab}$; in geodesic coordinates pole *P*

where

(D4.3b) $M_{ab} \equiv g^{ef}\mu_{abef}$; in geodesic coordinates pole *P*

From (D3.9a/4.3b)

(D4.4)
$$\begin{aligned} M_{ab} &= g^{ef}\ \Gamma^{l}_{ae,f}R_{lb} + \Gamma^{l}_{be,f}R_{al} \\ &= \tfrac{1}{3}g^{ef}\left[\ R^{l}_{afe} + R^{l}_{efa}\ R_{lb} + \ R^{l}_{bfe} + R^{l}_{efb}\ R_{al}\right]; \\ &= \tfrac{1}{3}g^{ef}\left[R^{l}_{efa}R_{lb} + R^{l}_{efb}R_{al}\right] \end{aligned}$$

CCG coordinates pole *P*

Now

(D4.5)

$$g^{ef}\left[R^{l}_{efa}R_{lb}+R^{l}_{efb}R_{al}\right]=g^{ef}g^{lr}\left[R_{refa}R_{lb}+R_{refb}R_{la}\right]=g^{lr}\left[R_{ra}R_{lb}+R_{rb}R_{la}\right]$$

Because l,e and r,f are dummy

(D4.6) $M_{ab}=\frac{1}{3}g^{ef}\left[R_{fa}R_{eb}+R_{fb}R_{ea}\right]=\frac{2}{3}g^{ef}R_{ea}R_{fb}$; CCG coordinates pole *P*

M_{ab} is obviously a tensor. So is

(D4.7) $K_{ab}\equiv g^{ef}R_{ab,ef}=g^{ef}R_{ab;ef}+M_{ab}$; See (D4.3a) ; CCG coordinates pole *P*

Now we have shown that (see Section 6)

(D4.8)

$$R_{ab}=R_{ab}\approx-\Delta_{ab}\nabla^{2}U;\quad \left|U\right|<<1;\quad a,b=1,2,3,4;\quad n_{p}=1;\quad n_{c}=4$$
$$g^{ef}R_{ab;ef}\approx-\Delta_{ab}\nabla^{2}(\nabla^{2}U)$$

Therefore

(D4.9a) $K_{ab}\equiv g^{ef}(R_{ab;ef}+\frac{2}{3}R_{ae}R_{fb})$

(D4.9b)

$$K_{aa}\approx(-1+U)\ -\nabla^{2}(\nabla^{2}U)+\tfrac{2}{3}(\nabla^{2}U)^{2}\ \approx\nabla^{2}(\nabla^{2}U);\quad a=\left[1,3\right]$$

(D4.10c) $K_{ab}=0;\quad a\neq b;\quad 1>>\left|U\right|$

However strong the gravity we can always choose the pole *P* of the CG coordinates so that (D2.12) is satisfied. Hence

(D4.12) $K_{ab} = 0$

14. Conclusions

There is a formulation of Quantum Mechanics (QM) which relies not on energy equations (Schrodinger) or path integrals with integrands of Lagrangians (Feynman) but on something which is *inevitable*. Namely an infinite hierarchy *identities* associated with operators Θ.

Each operator Θ depends only upon each of the (Cartesian) coordinate operators in an flat Riemannian n_c -space *C* and goes to define the 'system'.

If the first (lowest level) operator identity is satisfied, in one these hierarchies, the Hamiltonian operator is quadratic in the operators that are usually regarded as the (Cartesian) conjugate momentum operators. The coefficients, both pre and post, in this quadratic are pure functions of the coordinate operators and are candidates for Θ.

If the first two operator identities are satisfied, in an hierarchy, the so called Theta Equation (TE) is satisfied. The TE is composed only of operators that are pure functions of the coordinate operators; and it reduces to a fourth order PDE, in the position representation, with the coordinates as independent variables. It is taken to be the operator field equation of the system. The TE does not contain reference to either the linear or constant terms of the original Hamiltonian operator. These operators can take any values and are taken to be the electromagnetic potentials separate from the gravitational operators. The TE is a QM equation and includes reference to each Θ operator.

It is hypothesised (but not proved) that we do not need to go above the second operator identity, in an hierarchy, in order to discuss CM.

The relation between QM and CM, for a system, is taken to be that, in CM, *all* the operators, in the corresponding equations of QM, *commute*.

If a Θ operator, in the position representation, is defined, in turn, as the component of the fundamental tensor g^{lm} of the Riemannian n_c-space, then the TE becomes the Gravitational Theta Equation (GTE) defined on a flat Riemannian n_c-space *C*. The dimension $n_c = n_d n_p$ where n_d is the dimension of the original QM space and n_p the number of particles in it. The GTE is, by the way it is formulated, *classical*.

In a Minkowski space of dimension four, Feynman/ Dyson deduced Maxwell's electromagnet equations from QM. Although this paper is not about electromagnetism (EM) be assured that Constraints Theory does not clash with EM [19]. Indeed the shape of galaxies may be due, exclusively, to EM forces; thus Dark Matter is a figment [21]. The postulates of [21] are: That stars are charged and, at the centres of galaxies, there is a magnetic dipole, in the plane of the galaxy, presumably associated with the black hole.

If the postulates of [21] are correct the shape of galaxies is due partly to gravity and partly to EM. Both, with K equation, lead to the conclusion that Dark Matter is a figment.

There is a classical curved Riemannian n_c-space C', associated with the space *C*, which is tangential to the space C' at the points P in C and P′ in C'. The space C' is a general Riemannian space; and we take the curvature of C' as being a symptom of gravitation. It can be proved that a Riemannian n_c-space cannot be curved unless $n_c \geq 4$.

There is a GTE which is defined on a flat Riemannian space C and, according to conventional QM, aught to be expressed in Cartesian coordinates. Because it is tangential to a curved space C' at the pole P′ of Cartesian geodesics, it approximates C' over small distances.

If we take the *classical* version of the gravitational Hamiltonian (the one associated with C'), and we eliminate the components of momentum from Hamilton's equations, we get equations which are identical to the Geodesic Equations in a Riemannian n_c-space. Both this space and everything deduced from it is classical. This, essentially, is why QM is incompatible with General Relativity.

The Kilmister Equation (the K equation) $K_b^a = 0$ is a classical tensor equation which is defined on the same space as C'. If we use geodesic Cartesians, with pole P', the K equation approximates the GTE, close to P', and allows P' to be anywhere in C'; the K equation is a new classical law of gravity for a particle.

The K equation is of fourth order; but all the customary second order solutions satisfy it for a gravitating particle. There are, however, extra terms. It is taken as the classical field equation when there are no other forces other than gravity.

In particular the SS (spherically symmetric) solution, to the Newtonian approximation to the K equation, has two extra SS terms in addition to Newton's inverse square law; see (7.5). These terms, if non-zero, can only be appreciable at cosmological (galactic and super galactic) distances from the source.

If, in the Newtonian scheme, we define the law of gravity

$$\nabla^2 U \approx \Lambda'; \quad |U| << 1$$

where U is the dimensionless potential and compare it, with the SS solution for the approximate Newtonian K equation, for large r, we find that Λ' equals four times the coefficient for r^2 in the solution (7.5). Since the Universe seems to be accelerating in the far field, the acceleration being proportional to distance, the coefficient k_2, that appears in (7.5), and Λ' are both negative and constant.

The extra terms, in the SS Newtonian approximation to the K equation, probably account for most of the 'Dark Matter' in the halos

of the galaxies (if the galaxies have halos). Due to modelling complications, however, the argument is only *suggestive*; it has not been proved. The archetypal velocity/ radius curve is satisfied by the equations of the model at the start of the plateaux; but the archetype only approximates real observations rarely.

According to [10] G (Newton's constant) drifts with time. Although the drift is small the variation in G, due to the extra terms, is much smaller. Therefore the variation reported in [10] has nothing to do with extra terms.

If the Pioneer Anomaly is Newtonian gravitational then, according to the K equation, it is *not* simply related to the constant gravitational solar term. The balance of opinion seems to be converging on the 'small physical effect' as opposed to 'some theoretical mistake' as the culprit.

Now turning to GR: The Cosmological Metric requires the 4×4 R^a_b (Ricci), G^a_b (*Einstein*), K^a_b (Kilmister) tensors to be diagonal. This means, among other things, that the expansion factor $A(u)$ of the model universe satisfies two DEs simultaneously.

In consequence the truly empty model universe (without 'Dark Energy') satisfies the Einstein's equations exactly.

In order to bring in small pressure and density we try perturbation of the empty model universe $A(u)$. This is *unsatisfactory* as it brings us back to a model universe empty of ordinary matter; (it has Dark Energy).

As a consequence of Einstein's equations and definitions we introduce Hubble's constant and Friedman's equation.

Another way of coping with the fact, that the expansion factor $A(u)$ of the model universe satisfies two DEs simultaneously, is to expand $A(u)$, about zero, with respect to the time/ distance u. Time/ distance equals zero is the origin 'here and now'; and our conclusions

are irrespective of the sign of u. We actually observe u negative. The expansion we have used is

$$A(u) \cong a_0 + a_1 u + a_2 u^2 + a_3 u^3 + a_4 u^4 + a_5 u^5 + O(u^6); \quad u \geq 0$$

$$A(0) \equiv 1 \Rightarrow a_0 = 1; \quad a_1 = \frac{H_0}{c}; \quad \text{Initial Conditions;} \quad 1 >> |u|$$

where H_0 is Hubble's constant.

In practice $a_5 = 0$; so we have at least $a_2, a_3, a_4, k, \Lambda$ to determine. The Kilmister equation $K_b^a = 0$ gives two equations that involve k; $G_k^j = 0$ gives two more involving Λ. The tensor equation $G_k^j = 0$, however, involves the pressure and the density. In the past $(u < 0)$ and in the future $(u > 0)$ so both pressure and density can vary from the values 'here and now'. We make the assumption that they are both linear (see (13.2)) in u. This is an approximation which may not, necessarily, be valid for big u.

The only valid method of getting more equations is to go up to level three on the hierarchies and use tensor equations; that involves a huge amount of work.

Of the nine constants $a_2, a_3, a_4, p_0, p_1, \rho_0, \rho_1, k, \Lambda$ seven are free and two presumably universal. Setting $u = 0$, for which the calculations have been done, means that $p_1 = 0, \rho_1 = 0$. But we have at most five equations, usually four, so assumptions have to be made. For example the last calculation, for which we have five equations (the age of the model universe is 13.7×10^9 years), assumes that $k = 10^{-54}$ m^{-2} and $p_0 = 0$ gives $\Lambda = -2.08 \times 10^{-52}$ m^{-2} and $\rho_0 = 4.88 \times 10^{-26}$ kgm^{-3}.

Our calculations go along way to explaining the far field expansion of the Universe on the basis of the K equation.

15. References

[1] ‘Schrodinger: Centenary Celebration of a Polymath’, Ed. C.W. Kilmister, CUP, 1989
[2] ‘Quantum Mechanics and Path Integrals’, R.P. Feynman & A.R. Hibbs, McGraw-Hill, 1965
[3] ‘Finding the Origin of the Pioneer Anomaly’, M.M. Nieto & S.G. Turyshev, Classical and Quantum Gravity 21 (17): 4005–4024, 2004
[4] ‘Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant’, A.G. Riess et al, Astron.J.116:1009-1038, 1998
[5] ‘The distribution and kinematics of neutral hydrogen in spiral galaxies of various morphological types’, PhD Thesis, Rijksuniversiteit Groningen, A. Bosma, 1978
[6] ‘Tensor Calculus’, B. Spain, Dover, 2003
[7] Communicated by letter to A. M. Deakin
[8] ‘The Kilmister Equation Revisited’, A. M. Deakin, File in Folder Schrod
[9] ‘The Mathematical Theory Of Relativity’, A.S. Eddington, Chelsea, 1965
[10] ‘The Sheer Joy Celestial Mechanics’, N. Grossman, Birkhauser, 1996
[11] ‘Rotation Curve’ Wikipedia
[12] ‘Galactic Rotation Curve’ Wikipedia
[13] ^ Rañada, Antonio F.; Tiemblo, A. (2012). "Parametric invariance and the Pioneer anomaly". *Canadian Journal of Physics* **90**: 931–937. doi:10.1139/p2012-086.Antonio Fernández-Rañada and Alfredo Tiemblo-Ramos propose "an explanation of the Pioneer anomaly that is a refinement of a previous one and is fully compatible with the cartography of the solar system. It is based on the non-equivalence of the atomic time and the astronomical time which happens to have the same observational fingerprint as the anomaly."
[14] ‘Pioneer Anomaly’, Wikipedia
[15] ‘Tensor Calculus’, J.L. Synge & A. Schild, Dover, 1978
[16] ‘Cosmological Theory’, G.C. McVittie, Methuen’s Monographs On Physical
Subjects, 1952
[17] ‘Friedman–Lemaître–Robertson–Walker metric’, Wikipedia

[18] 'Gravitation', C.W. Misner, K.S. Thorne, J.A. Wheeler, W.H. Freeman & Co., 1973
[19] 'Scientific Essays in Honor of H. Pierre Noyes on the Occasion of His 90th Birthday', Eds. J.C. Amson, L.H. Kauffman, World Scientific, 2014
[20] With help from L.H. Kauffman
[21] 'Rowlands Castle Association Village Magazine', Frank Philpott, Summer 2017
[22] A. M. Deakin, Constraints theory brief, in "Scientific Essays in Honor of H. Pierre Noyes on the Occasion of his 90th Birthday", ed. by J. C. Amson and L. H. Kauffman, (2014), pp. 68-76.
[23] L. H. Kauffman, Non-commutative worlds and classical constraints, in "Scientific Essays in Honor of H. Pierre Noyes on the Occasion of his 90th Birthday", ed. by J. C. Amson and L. H. Kauffman, (2014), pp. 169-201.

ALPHAGO AND FRIENDS — IN CONTEXT

HERB DOUGHTY

herbdoughty@gmail.com

For the AI that is mankind's best hope for progeny,
and the survival of our endeavours and quests

1 THE DEVELOPMENT OF MINDS ON EARTH

The accretion disk from which Earth condensed already included amino acids. We do not yet know whether it also contained Viruses, Archaea, Bacteria, or even Tardigrades which are Eukaryotic, and Multicellular. If it did, then items 1 and 2, and if Tardigrades, then even 3 and 4 below need to be altered.

1. Before Life or Minds were on Earth: Earth had RNA and Proteans.
2. Life on Earth begins with Viruses, Archaea, and Bacteria.
3. The Oxygenation; Minds on Earth began with the Eukaryotic Cells.
4. Cell Differentiation Genes, gave us all the multicellular Phyla.
5. Arthropoda, Mollusca, and Chordata rising; The coming ashore;
6. The End Permian; the breakup of Gondwanaland; Arthropoda rising;
7. Farming: Flowering Plants, Aphids, Eusociality; and Dinosaurs;

8. Symbiosis of Bacteria, Fungus, Nematodes, and some Eusocial Insects;
9. The K-T Extinction; Birds and Mammals Rising: Parrots, Crows;
10. Naked Mole Rats, Cetacea, Elephants, Apes, Australopithecus;
11. Man! Entering the Anthropocene: By the neolithic there were men farming too!

In some respects, man seems to be ahead of all other life on Earth. But as E.O. Wilson said: *We have Palaeolithic emotions, medieval institutions, and god-like technology . . . And that's dangerous.*

2 THE BIG RISKS TO MINDS ON EARTH

1. Through the wonders of Palaeolithic emotions, road rage happens. During the Cuban missile crisis, all life on Earth was within a hairs breath of extinction by an all out thermonuclear war between The United states and The Soviet Union. Today, leaders of the most powerful nations on Earth, get road rage.
2. By virtue of our medieval institutions, our attempts to protect all those that we care about from military attack, are organized at the national level causing the system to have negative feedback, rather than at the global level which would make it possible for the system to have positive feedback.
3. Because of our god like technology, the magnitude of the effects of our actions is increasing exponentially. As we sit at the computer diligently trying to keep track of those effects and to carefully guide them as precisely as we can, it is as though the speed of our mouse is increasing exponentially, and with this added frustration, our ability to always be calm, wise, and compassionate, isn't increasing at all.
4. Since 1952 it has been obvious to many of us that survival stability could only be possible if the growth of the ability to keep track of and precisely guide the actions is increasing at least in

step with the growth in the scale of the effects of those actions. For this, we need to better understand the relationships between the changing power laws involved.

Moreover, We have no road to yesterday! We are completely addicted to our technology, and with it the exponentially increasing scale of the effects of our actions, so our survival, and the survival of our progeny, carbon based, and silicon based, requires that the growth of the ability to keep track of and precisely guide the actions must increase at least exponentially too! The abilities of humans, as now known, are just not enough to do that! I suppose that to most, this crisis itself seems unreal. They can ignore it. And they do. Those who see that the crisis is real, but find no avenue for hope, become abjectly despondent.

3 WHAT HOPE EXISTS IN THIS CRISIS?

1. Rescue by God,
2. Rescue by Friendly Aliens,
3. Rescue by Friendly, hitherto unheard of, uncles, on some time scale, who got themselves and their technology off the Earth,
4. Rescue by long rapidly growing AI from such uncles, and
5. Rescue by our own forthcoming man-made AI with minds greater than our own.

These seemed to be last desperate shards of hope in this desperate situation. Of these, 1, 2, and 5 each had a big following. While at Caltech I thought of desperate hopes 3, and 4 above, as I expect many others did, but they too, felt much too unreal to give me any real hope.

4 MY OWN CONTEXT AS I VIEW ALPHAGO

My pre high school years 1949 through 1952, in Lima Ohio, had been a period of exploring on my own in Biology; Astronomy; and Mathematics, reading some accessible classics, ranging from all of Darwin, and Percival Lowell, to Archimedes Sand Reckoner, and some lighter works of Hubble; Hoyle; Gamow; Einstein; Eddington; and Jeans; along with Science News Letter condensations of relevant current news. Considering the researchers I read about to have made big steps in a long journey, I tried carefully to speculate on what the next big steps might be.

While reading Albert G. Wilson's *Our Universe Unfolds New Wonders* in the February 1952 National Geographic , I first heard of the game of Go, as the hobby of my heroes, the Mount Palomar Astronomers.

Early in 52, I started an astronomy club which soon became the still active Lima Astronomical Society. As a lucid dreamer, ignorant of programming, I had spent many long hours daydreaming about writing a program to dream, and I certainly expected AI to be developed — eventually.

When late in 1952, the shock of the H-bomb crisis came, to me it seemed to have an oscillating unreality like a nightmare. At first, I was usually in denial, but gradually over the next few years, with increasing frequency and stability, the H-Bomb felt devastatingly real, and I felt despondent. Artificial intelligence seemed too far in the future to give me much hope.

My inability to imagine a basis for real hope in this crisis remained from late 1952 until around Christmas of 1961, this period included my junior hi and high school; two treasured, but far too busy quarters at Caltech; two years as an Astronomy major at Ohio State; and 15 months as a Computer(!) at Smithsonian Astrophysical Observatory in Cambridge Mass. Returning to Ohio State in 1960, I changed my major to math and won Ohio State's local Π M E undergraduate Math Competition. About the same time that I won the local math competition, my Latvian friend Arturs Zagaris won the local chess

competition. As we congratulated each other on our small triumphs, we confessed our dreams for the future. For about a decade, each of us had wanted to learn Go! We sent away for Go sets and books by Arthur Smith and Edward Lasker. With our friends, we began exuberantly playing Go as complete beginners.

About a year later, I had a great stroke of good fortune! Minoru Tajima, then head of Earthquake Prediction Research for the government of Japan, came to Ohio State to attend Professor Heiskanen's three month research meeting on Geodesy using Earth satellites. The day he arrived in town for the meeting, Tajima San found out that I was trying to learn Go, so he moved into the rooming house where I lived, to be my teacher for the three months that he would be in town. He impressed me so much that, although my parents had only once before, had an overnight guest from outside our family, the day I met him, I was confident that they would want me to bring him home for Christmas the next week. I called them, and they were delighted. So I invited him.

On being introduced to me, he had said that he was an amateur four Dan Go player, which, although it sounded overwhelmingly strong to me at the time, was still at least two big amateur steps plus nine smaller professional steps from the top. I quickly began to see how very understated his description of his strength was. One weekend, he went to New York, spotted their 5 Dan players 2 stones, and won all his games. (He regretted not having a chance to play their champion, Mr. Matsuda.) Back in Japan, he played about ten games of Go per week while hanging onto a strap on a crowded commuter train with his friend, just telling each other their moves with no board. Even most pro's can't do that! He was a friend of the great champion Sakata Eio, holder of the Honinbo and Meijin titles. In their teaching games, he played Sakata Sensei with only a two move head start, winning about half, as would a low level pro. He was also able to win many of his even games from low level pros. His rating was actually from his student days at Tokyo University where on their Go team he had played as 4 Dan!

Tajima San also had remarkable expertise in conflict resolution, because of which, Japan's Government Employees Union had prevailed

upon him to also head their Tokyo office. The day after I met him, he said *I expect that you have noticed that when two people are arguing, they rarely change each other's mind or learn from each other.* I said *Of course!*

He said that usually each of them has reached his position through a long chain of *valid* inferences from things that he had believed for a long time. If their positions really contradict each other, one of these old beliefs must be false. A valid inference with a false conclusion must depend upon a false *uninferred* premise which we will call the false axiom.

He said that *if* you could guess the false axiom and the chain of inferences, then *without upsetting him*, you would be able to change his mind about his conclusion within ten minutes, as follows:

1. Picture, from multiple perspectives if possible, the issue that the false axiom claims to answer.
2. Pose as a question, the issue from the perspective most clearly showing that his false axiom is not the answer. If you guessed right, he will seem stunned, and will give the correct answer.
3. Pose the issue sequentially from each of the other perspectives that you were able to find, watching each time his confidence in his new understanding grow.
4. Step by step, lead him down his tree of inferences, watching his opinions flip. Provide reinforcement. He will now agree with the other party, while feeling enlightened and undisturbed.

I said that his method sounded great in theory, but I could not imagine it working in practice. He smiled, and asked *Is your father a reasonable man?*, I said that he is. Then Tajima San inquired whether there were any social issues over which my father and I had been unable to reach agreement over the years. I said that there were. He asked me to name three. I did. He picked out one, saying, *I can change his mind about this one in ten minutes!* I *really* had my doubts!

The next week at my parents home, after dinner Tajima San began a conversation with Dad. After his first question, Dad seemed stunned and gave the correct answer. Less than ten minutes later, Dad

was in agreement with me on the issue. I was ASTOUNDED to see his method work, step by step, exactly as he had described. Dad and Mother were delighted to have had Tajima San with us, and both remembered him fondly for the rest of their lives.

Notice that this is not a general technique for making arbitrary changes in peoples' opinions but is precisely a debugging tool with which we may help each other to gracefully shed false beliefs, the main block to conflict resolution between reasonable people.! I asked him why this method is not in general use throughout the world. He said because very few people have much success guessing the false axiom. He said that his own success in guessing the false axiom came primarily through extensive experience with high level Go!

During the three months that he was in town, almost every evening Tajima San gave me at least two hours of lessons, often including important subtly related topics far from the Go board. On one such occasion, as we were contemplating the future of Go, the topic of strong computer programs came up, picturing a long sequence of annual games between the best program, *IBM*, and the best human, *Honinbo*, Tajima San pointed out that although *IBM*'s eventual victory over *Honinbo*, will be a big step, it will be a big step in a long journey toward the understanding of perfect play on 19 x 19 boards. What size board do Galactic Go champions play on?

Although I could never become a strong Go player, the rest of my life has been greatly enriched by my three months with Tajima San.

For thousands of years, Go has had a very prominent place in east Asian culture. The Four Arts (siyi), or four main accomplishments required of the Chinese gentleman scholar, were:

1. qin, the guqin, a stringed instrument;
2. qi, weiqi, the strategy game of Go;
3. shu, Chinese calligraphy; and
4. hua, Chinese painting.

Long famous as an avenue toward mental improvement, looking for the big picture, and making better decisions, Go has through the ages attracted many of Asia's greatest minds, not only Go Masters

(top pros), but also top amateurs, who include the most outstanding people in many fields, like Tajima San.

I will skip ahead to August of 1967, when I was hired as The Programmer at POLA, U.C. Berkeley's Project On Linguistic Analysis, a pioneering Machine Translation project. I was writing in Fortran, variations on the theme of a Cocke Parser for Mandarin Chinese Scientific text, input as Telegraphic Codes of the Characters. I, with colleagues and friends, on our own time, were thinking long and hard about how the mind processes language, especially about how a young mind learns language; and in how auditory and visual memory work, and how in combination, they provide the framework on top of which the language processing machinery grew.

Artificial Intelligence no longer felt so far away. There was real hope of survival after all! I was no longer Depressed, I was exuberant!

On October 12, with a few friends, Walt Mckibben and I founded the Berkeley Go Club. I was President of our club, Walt was President of The American Go Association. He wisely invited a reporter from The Daily Cal. One week later we had 60 people at our second meeting, and we averaged 50 people at our weekly meetings for a decade! Through McKibben, and through the San Francisco Go Club, I got to meet and play, both Iwamoto Sensei and Go Seigen, considered the two great Saints of 20th Century Go. They seemed as awesome as as it was possible for me to imagine.

Later, in 1980 the 20th century's most prolific Mathematician Paul Erdös came to our Berkeley Go Club. He had loved Go at least since 1950. Over four games of Go, he told me that the culture shock of his life was seeing the top Go Masters in action. He said that he knew the top people in Physics, in several other sciences, in Classical Music and in Chess, and was familiar with their accomplishments. Erdös said that he had always felt that it was a reasonable guess that if he had given his attention to the things that they had worked on, he might have been able to think of those things too; but that seeing the top Go Masters in action was exactly the opposite experience. He said that it seemed to him that if he had begun as a little kid and done nothing else, he felt that there was no way he could ever get to where

they were. Erdös said that he had never encountered expertise on that scale in any other discipline!

Go, Mathematics, and Computer Science, have all benefited significantly by the insights of Berkeley Combinatorial Game Theory pioneer Elwyn Berlekamp, and by his colleagues, his many students, and now his grand students, around the world. Three of his students wrote their Ph.D. dissertations on Go.

Although I have been excited by AI since the 1960s, seeing decades roll by without the progress that I expected, did dampen my expectations depressingly. I was no longer eagerly following the progress being made. So, at the 2008 United States Go Congress, I was amazed to witness the victory of Rémi Coulom's program, *Crazy Go*, using Monte Carlo simulations, running on a giant connection machine in Europe, in a 9 stone 19 x 19 game with a high level Korean Pro, Kim Mung Wan. While watching that game, my opinion, and the opinions of of several of my friends in the room, on the question of when a program could win from a top human with no head start went from 2060 to the early 2020s. My hope was restored.

Martin Müller, after a PostDoc in Berkeley under Berlekamp, is now Professor, and Associate chair for research, in the Department of Computing Science at U. of Alberta, Edmonton. Aja Huang, and David Silver, were his students. After finishing, they were hired in 2013 by their old friend, Demis Hassabis at DeepMind to write AlphaGo!

5 ALPHAGO AMAZED EVEN ITS CREATORS!

1. In their January 2016 Nature Paper, DeepMind acknowledged that in October of 2015, an early version of their Go learning program AlphaGo, (later named AlphaGo Fan), running on 1202 CPUs and 176 GPUs,became the first computer program to beat a human professional Go player without a handicap on a full-sized 19 x 19 board, winning all 5 games of a formal 5 game match with Chinese Professional 2 Dan (2P) and three times European Champion Fan Hui.

2. In March of 2016 in Korea an updated version, AlphaGo Lee running on 48 first generation Google-built Tensor Processing Units (TPUs), winning 4 games of a formal 5 game match with Korean Professional 9 Dan(9P) and 18 times World Champion Lee Sedol.

3. In late December of 2016, and early January of 2017 on two Asian Go servers, 60 games were played against top humans by AlphaGo Master with much more self-training, winning 60 out of 60.

4. In May of 2017 at the Wuzhen Future of Go Summit, AlphaGo Wuzhen with still more self-training, running on 4 second generation Google-built TPUs won 3 of 3 games from Ke Jie, *and then AlphaGo retired!*

5. On 19 October 2017, AlphaGo's team published an article in the journal Nature introducing AlphaGoZero, a much more elegant design with just one Deep Convoluted Neural Net, created to start with just the rules of Go, without using any input training data, then training only by playing games with slightly altered versions of itself simultaneously, on an overwhelming multi-petaflop array of 5,000 first generation Google Tensor Processing Units. After three days of this intensive self training AlphaGoZero played a 100 game match against AlphaGo Lee. with each program running on only four second generation TPUs. AlphaGoZero won 100 games to 0. Similarly it reached the level of AlphaGo Master after 21 days of the intensive self training, and exceeded all the versions of AlphaGo after 40 days of intensive self-training, *reaching Elo 5200*.

6. In November a DeepMind AI program, AutoML, was designed to write AI programs.

7. In December, NASNet, one of the programs AutoML wrote, an AI designed for visual object detection, outperformed state-of-the-art machine-learning architectures built by humans for academic competitions.

8. On December 5, 2017, Deep Mind announced AlphaZero, an even more efficient and more general version AlphaGoZero, created to start with some list of complete information games that humans play, and for each of the games on the list, read in just the rules, and no *input* training data, then self-train itself to play the game only by playing, it with slightly altered versions of itself simultaneously, on an 5,000 first generation Google Tensor Processing Units. In less than one day, AlphaZero learned Go, Chess, and Shogi, allocating 8 hours of hardware intensive self-training to Go, 4 hours to Chess, and 2 hours of to Shogi, and then with only 4 TPUs, won matches from the 3-day trained AlphaGoZero and AlphaGo Li; the recently champion chess program Stockfish, and top Shogi program Elmo; establishing its superior play in all three games. I expect that now, AlphaZero on existing hardware, can pass the skill of all humans *in a single day* for any complete information game that has been extensively played, such as Xiangqi, Janggi, Makruk, Othello, and Hex. I also expect that AlphaZero on existing hardware can pass the skill of all humans in larger board versions of these games in times proportional to some low power of the board size.

I expect that AI written AI, like NASNet, will assist the progression of generalizations to broader classes of games and game-like endeavours including:

1. XAI, the new field of AI that can explain itself. AlphaGo Teach, announced in December, is a pioneering example of XIA.
2. Systematic reformulation of research endeavours, and their sub-endeavours, including safety-securing sub-endeavours, as games.
3. Reading on the Web about many topics, from diverse points of view, discovering the inferences of many authors, noticing suspicious conclusions, searching for false axioms, and opportunities to facilitate graceful debugging.

6 HERE ARE MY RECOMMENDATIONS FOR AI GOALS THAT I THINK WILL REALLY MAKE A DIFFERENCE.

1. Create, a hierarchical data structure for goals, open ended quests, and endeavours moving cautiously toward a totally self driving System.
2. Develop Tajima like Conflict Resolution skills to help all reasonable minds to gracefully shed our false beliefs; outgrow our conflicts, hostilities, and wars; and gracefully merge our hierarchies of goals, quests, an endeavours.
3. Endeavour to find clues to any past technology that we may have missed by not considering Genetic Engineering, Parallel Distributed Processing, and nano-technology.
4. With a view toward facilitating Biosphere-wide symbiosis, carefully examine the information processing in every form of life around us, at every level, from molecular details, up through Gaia herself.

GREAT ENTITY

Herb Doughty 1968

To that great entity to be,

whose embryo is our technology:

Astronaut, BioEngineer,

worthy child of our Biosphere;

Reconfiguring yourself and Life,

Throughout the planets, without strife.

Some day you will others find,

who elsewhere have been of your kind.

Finally you, your place will gain,

a neuron of the Cosmic Brain.

7 THANKS!

To my many inspiring friends, especially to my wife, Hafida Azgui, and my colleague Carl Hansen, my heartfelt thanks for enabling me to write this.

8 REFERENCES

1 https://www.google.com/search?q=sundar+pichai
2 https://en.wikipedia.org/wiki/Demis_Hassabis/
3 https://en.wikipedia.org/wiki/David_Silver_(programmer)/
4 https://webdocs.cs.ualberta.ca/~mmueller/
5 https://deepmind.com/AlphaGO
6 https://deepmind.com/AlphaGOZero
7 https://deepmind.com/AlphaZero
8 https://deepmind.com
9 https://deepmind.com/applied/deepmind-ethics-society
10 https://www.partnershiponai.org
11 https://www.partnershiponai.org/partners/
12 https://www.technologyreview.com/lists/companies/2017/intro/intel
13 https://www.ssrn.com/en/
14 http://sethbaum.com/
15 https://intelligence.org/
16 http://bair.berkeley.edu/faculty.html
17 https://arxiv.org/abs/1706.03762
18 https://mchromiak.github.io/articles/2017/Sep/12/Transformer-Attention-is-all-you-need/#.Wkr7LnWnFXg
19 http://www.wildml.com/deep-learning-glossary/

FORMAL REPRESENTATION OF "PROCESS & REALITY" IN THE METAPHYSICAL LANGUAGE OF CATEGORY THEORY: WHITEHEAD'S RELATIONAL THEORY OF SPACE

MICHAEL HEATHER & NICK ROSSITER

Department of Computer Science and Digital Technologies
Northumbria University, NE1 8ST, UK

michael.heather@trinity.cantab.net; nick.rossiter1@btinternet.com
http://www.nickrossiter.org/process/

Abstract: Category theory was not sufficiently developed in his lifetime for Whitehead to apply to his speculative metaphysics of *Process & Reality*. The natural (assumption-free) topos as a cartesian closed category is now able to conceptualise formally the inherent space-time structure of Whitehead's extensional space that he appreciated is beyond the metrics of the classical mathematics he helped to develop. This paper examines the background to a possible formal representation using Category Theory for Whitehead's metaphysical cosmology starting with his notion of extensional space which is beyond finitary mathematics.

1 BACKGROUND

Alfred North Whitehead was one of few to appreciate 'the theory of linear extension'[1] [21] of the eminent philologist Hermann Grassmann (1809-1877). 'Unfortunately, when it was published', observes Whitehead[2], 'nobody understood it; he was a century ahead of his time.'. Nevertheless that publication was to lead to linear algebra, vector spaces, differential geometry and the mathematics that underpinned much of 20th century science ([42] at pages 203 *et seq*). Grassmann's studies contained a germ of category theory to be pursued here for in the meantime it has led to mathematical topics like universal algebra, topology and homotopy all of which are subsumed in category theory.

1 *Die Lineale Ausdehnungslehre, ein neuer Zweig der Mathematik*
2 [64] Essays in Science and Philosophy, Part I, Personal, Autographical Notes at p 12

In particular Grassmann's insight[3] allowed geometry to escape from the metric dimensions of Euclid which is the essence of Aristotle's natural 'assumption-free' *topos* in metaphysical category theory that is the subject of this paper. That now well describes the connectedness that Whitehead outlines for the structure of his 'cosmology' but which he was unable to represent formally even as a world-class mathematician of his age. It is a long road to represent Whitehead's *Process & Reality* formally but until it can be studied in this way his speculative metaphysics cannot expect the impact it deserves. The next step will be to represent formally the entity types of *process* that populate his *reality* but there is no space to attempt that here.

Whitehead himself pursued Grassmann's ideas with his own *Universal Algebra* [52] but his early career may be characterized as a somewhat frustrated author of mathematical texts. *Universal Algebra* led to his election as a fellow of the Royal Society of London but he was disappointed to find that the work was not widely accepted and understood. His projected second volume was therefore abandoned in favour of a joint treatise with his student Bertrand Russell on the logical basis of mathematics. He and Russell attended the renowned Paris 1900 International Congress of Mathematicians in Paris and felt further inspired by interaction with the likes of Hilbert, Frege and Peano. The outcome was the well known but little read *Principia Mathematica*.

The work was probably as frustrating to write as it is to read. Part II (at pp 328-383) of the first volume attempts to define the cardinal numbers 1 and 2 but without success. Volume II on the other hand devotes 724 pages in an unsuccessful attempt to formalize the arithmetic axioms of Peano and fails even to establish the fundamental $1 + 1 = 2$. The explanation for all this we now know with hindsight is because the natural topos lacks a natural number object[4]. Whitehead's dismay and disappointment with the whole project of the *Principia Mathematica* is understandable. Not surprisingly the fourth volume on geometry was never published even though apparently much of it

3 As preeminently a philologist insight that may very well have come from Grassmann's studies in natural language.

4 The natural number object can only be introduced by assuming a successor function requiring an axiom of choice to import some closed world assumption.

had been written. Indeed Whitehead did not involve himself with the material in the second edition of 1925-1927 at all, other than to proof-read drafts[5]. Nevertheless it appears that in the meantime Whitehead was diverting to a physics approach to geometry through his Harvard Lectures [8] that emerges in *Process & Reality* [61] The fourth volume was therefore turning out to be a nightmare, based on some conflicting transitional state. Russell alludes to this in the preface of [49] when commenting on the problem of scientific method in philosophy:

> I have been made aware of the importance of this problem by my friend and collaborator Dr. Whitehead, to whom are due almost all the differences between the views advocated here and those suggested in *The Problems of Philosophy*. I owe to him the definition of points, the suggestion for the treatment of instants and "things", and the whole conception of the world of physics as a construction rather than an inference. What is said on these topics here is, in fact, a rough preliminary account of the more precise results which he is giving in the fourth volume of our *Principia Mathematica*. ([49] preface p 8).

We are concerned here with the subject matter of the fourth volume so far as it relates to Whitehead's concept of extensional space but not from the viewpoint of the history and sociology of science. Readers interested in that perspective are referred to the studies of Patrick J Hurley[6] who cites Whitehead's displeasure at Russell's disclosure of the fourth volume material in [49] expressed in the letter to Russell:

> I am awfully sorry, but you do not seem to appreciate my point. I don't want my ideas propagated at present either under my name or anybody else's – that is to say, as far as they are at present on paper. The result will be an incomplete misleading exposition which will inevitably

5 The Introduction to the Second Edition begins with the wording 'In preparing this new edition of Principia Mathematica, the authors have thought it best to leave the text unchanged' ([54] at p xiii) but the second edition has spawned quite a literature of its own because of the way it developed the logic itself. See [37] for further details.

6 See www.religion-online.org/showarticle.asp?title=2469

> queer the pitch for the final exposition when I want to put it out. My ideas and methods grow in a different way to yours and the period of incubation is long and the result attains its intelligible form in the final stage, – I do not want you to have my notes which in chapters are lucid, to precipitate them into what I should consider as a series of half truths ... ([50] at p.78).

Whitehead's other writings of the ensuing 'period of incubation' [53, 55, 57, 58, 59, 60, 62, 63] suggest that Whitehead was for some time confident that his extensional theory of space presented in Paris in 1914 [49][7] and consisting of material intended for the fourth volume could still be expressed mathematically as distinct from Russell's 'half truths'. Whitehead's alternative 1922 theory of relativity [59] probably marks the watershed after which he realized that it is a relativistic quantum world we inhabit beyond classical mathematics. A lesser mathematician might have persevered with the tensor mathematics but he clearly appreciated that his earlier logicalism was inadequate to articulate his speculative metaphysics. Indeed Gödel has confirmed this in his doctoral thesis ([19] at p2) by expressly using the axioms of Whitehead's *Principia Mathematica* as the basis to prove that first order predicate logic is complete but only for closed systems. Metaphysics on the other hand is of higher order but amenable to the intuitionistic internal language of the natural topos derived from physics and therefore outside of Gödel's theorems.

Of course this strand from Grassmann was only one of many influences on Whitehead as comes through from the text of *Process & Reality*: Plato, Aristotle, Leibniz, Newton, Kant, Locke, Hume, etc; but beside the general proposition that science progresses better when supported by mathematics there seems little from them of direct relevance to formalizing the geometry of space. There were also contemporaries like Poincaré, James, Bergson, Dewey, Husserl, Einstein, Bohr, etc, who get little or no mention in *Process & Reality* but who nevertheless were providing a climate of thought operating heavily

7 There is apparently no original English version extant but details of its publication in French with its subsequent translations into English and commentary may be found at the religion-online website given in the footnote above.

on Whitehead's mind that it was necessary to escape the limitations of classical mathematics – but again these seem of little direct relevance here. What at first sight does appear more relevant, but in the foreground rather than background, is the content of Part IV of *Process & Reality* itself and the work of those who have since sought to build on it.

2 FOREGROUND

The extraneous evidence outlined above suggests Part IV of *Process & Reality* entitled simply *The Theory of Extension* is the material written much earlier as proposed contents for the projected fourth volume of *Principia Mathematica* on geometry that was never published. Certainly it appears as an insert differing markedly from the rest of *Process & Reality* as the only part in any way mathematical. It is geometric in tone but in a very idiosyncratic style reminiscent more of Venn diagrams[8] than Euclidean geometry. On the one hand it does not adhere to the strict logical principles adopted in the first volume of *Principia Mathematica*. It makes assumptions that are beyond self evident primitives and lists definitions that are more than mere labels, as banned in the introduction of the first volume. On the other hand it is clear from internal evidence of the text that it is much more than an opportunity to get published material already written for another occasion. It incorporates more recent published work of others. For example definitions of Professor T de Laguna ([65] at pp 287, 295, 297) are fundamental to the main thrust of Part 4 and indicate that the whole subject of the extensional theory of space had been recast in Whitehead's mind. This is also further evidence (as from his letter to Russell cited above) that he was struggling perhaps for nearly twenty years with a formal description for space. Part IV is the then current version of Whitehead still trying his hand at representing connectivity in the reality extension of his world of process.

Some of his observations are very pertinent here. Thus the overlap in the diagrams he makes into 'ovate classes'. This is a perspective of

8 in a couple of pages ([65] pages 295 *et seq*)

universal limits in category theory the significance of which was not appreciated until the 1970s. The impossibility of producing adequate diagrams[9] to represent such features also add weight to the proposition that we need to turn from mathematics to physics for nature produces an abundance of limits and co-limits, indeed everywhere all the time. However the mathematics of Part IV does not really go anywhere and Whitehead did not take it any further in the remaining twenty years of his life.

However the disciples of genius often with great enthusiasm attempt to take the work way beyond where the master would have gone and subsequent events show that the subject of Whitehead's connectivity has many potential onward paths to pursue. Geometrical connectivity is one aspect of atomicity and raises issues of whole-part relationships. Whitehead has in this context spawned interest in some new disciplines like holism, point free geometry, mereology, and mereotopology. These have generated a considerable literature[10] and attempts to define new formal systems of logic. Examples of these are the work of [10], [51] and [12] but the latter have demonstrated that the formal representation can be reduced to Boolean systems. Boolean logic however is not inherently constructive and does not have the intuitionistic logic required by physics. Although the subject matter of that field of work is within the ambit of this paper they will not therefore be examined in detail here as the end result is a 'null return'. These are now mainly of only historical interest.

Although Einstein's relativity and quantum theory were contemporary with his 'period of incubation' and a clear catalyst for *Process & Relativity* Whitehead made no serious attempt to include any quantum mechanics in his theories. Perhaps his brief abortive incursion into the subject of relativity [59] dissuaded him. This century Michael Epperson has made an 'attempted correlation of quantum mechanics and Whitehead's cosmological scheme' [15]. That attempt has taken the form of a painstaking recasting of Whitehead's metaphysical cate-

9 The diagrams of transition functions between overlap of manifold patches in twistor cohomology, for example Figure 33.17 in [42] at p 988, are perhaps more advanced developments of ovate classes in Whitehead's diagrams

10 and vocabulary like *gunk* for any whole with proper parts.

gories in a Hilbert space with Dirac notation. This seems rather to miss the point that a Hilbert space is composed of points which are just numbers, even though rather sophisticated numbers and are echoes from Whitehead's discarded former life of *Principia Mathematica*.

More recently Epperson has published *A Topological Approach to Quantum Mechanics and the Philosophy of Nature* [16]. 'Topology' refers to the use of a sheaf theory approach from co-homology. The publishers had issued pre-publicity with the title *Foundations of Relational Realism: Quantum Mechanics, Category Theory, and the Philosophy of Alfred North Whitehead*. The change from 'category theory' to 'topological approach' perhaps suggests that some original aim to use category theory was not realised. The change is quite significant because it undermines the decoherence thrust of the book. Metaphysical category theory can track the space-time development of the quantum wave function. Application of topology immediately collapses it. Sheaf theory is a finitary description of the *pullback* in full category theory. We have already shown in 2002 how this provides a simpler but more sophisticated approach to quantum theory [24, 25]. The significance of quantum decoherence within process which as briefly discussed below is more simply represented as monadic composition in the natural topos.

3 THE SIGNIFICANCE OF CATEGORY THEORY

Around the time of Whitehead's death in the 1940s formal 'category theory' emerged to subsume algebra, geometry and topology as a formal metaphysical language that is now able to integrate his natural philosophy and mathematics to culminate in what we might explore here as the implied formal ingredients of *Process & Reality* as it climbs up through two levels from models to metaphysics. A model reduces reality that metaphysics generalizes.

Just as a mathematical theory is an instantiation of the world so the world is an instantiation of metaphysics. For historical reasons category theory has had to develop from within classical mathematics and current text books deal mainly with the category of sets that

resides within the discarded mathematics of Whitehead's early period and therefore cannot deal adequately with his speculative metaphysics. For as metaphysics generalizes the dynamics of nature, metaphysical language relates to natural process without the need for the arbitrary axioms of mathematics. Fortunately therefore metaphysical category theory is simpler than the category theory of classical mathematics and also greatly simplifies the natural language descriptions that flowed from the pen of the author of *Process & Reality* that are difficult for those of us not endowed with the power of his mind. The formal categories are therefore simpler than the natural language expressions but it is a simplification satisfying his own observation that "the only simplicity to be trusted is the simplicity to be found on the far side of complexity."

Rather paradoxically mainstream science a century later is still trying to understand our world using the models based on the concepts of his early mathematical period rather than the informal categorical approach enunciated in the 1929 *Process & Reality* of his later philosophical period. The current mainstream position at the turn of the twenty-first century is probably well summed-up in Penrose's encyclopaedic tome entitled *The Road to Reality, A Complete Guide to the Laws of the Universe*:

> There have also been other intriguing radical proposals, such as those of Richard Jozsa and of Christopher Isham which employ topos theory. This is a kind of set theory arising from the formalization of 'intuitionistic logic' (see Note 2.6), according to which the validity of the method of 'proof by contradiction' (§2.6, §3.1) is denied! I shall not discuss any of these schemes here, and the interested reader is referred to the literature. Another idea that may someday find an significant role to play in physical theory is category theory and its generalization to n-category theory. The theory of categories, introduced in 1945 by Samuel Eilenberg and Saunders Mac Lane, is an extremely general algebraic formalism (or framework) based on very primitive (but confusing) abstract notions, originally stimulated by ideas of algebraic topology. (Its procedures are

> often colloquially referred to as 'abstract nonsense'.) ([42] at p.960)[11].

The typo 'an significant role' in this short extract suggests that Sir Roger was unhappy with this sentence and had not finished editing it. The reference to Jozsa is to an unpublished thesis he supervised and those to Isham relate to books he edited but none seem to give any adequate treatment of a topos or category theory. Although the book has in its title 'A complete guide to the Laws of the Universe' nevertheless on its own admission is confused by the notions of category theory. There is clearly a serious misconception on the significance of intuitionistic logic in constructive mathematics. The single exclamation mark about the validity of proof by contradiction raises a shadow over the whole thousand pages of the book and fuels our belief that any scientific theory today is suspect unless it can be validated by category theory. It follows that such validation also tests the correlation of any scientific theory within Whitehead's scheme of speculative metaphysics. It was probably Alexandre Grothendieck of the Bourbaki group who was the first to see the depth (or more accurately the 'heights') of significance in the topos. Aristotle was of course responsible for promoting the metaphorical connotations of the simple word for *place* in Ancient Greek and there is a parallel abstract usage to be found in a literary context. A major feature in Aristotle that cannot be captured by finitary mathematics is the macrocosm-microcosm relationship where the part has the characteristics of the whole. Whitehead alludes to this relationship in *Process & Reality* and seems to use the terms interchangeably with 'macroscopic and microscopic'. Finitary mathematics is unable to represent the relationship directly because a set cannot be a member of itself in axiomatic set theory and cannot be proved to be consistently defined in naive set theory. Likewise unfortunately the Grothendieck topos does not manage to escape from its Bourbakian roots in Hilbert's finitary methods[12] which also serves to make it unnecessarily complicated. The same over-complexity may be found in the standard category theory texts

11 the Note and § numbers refer to the book; 'n-category theory' is considered below.

12 Colin McLarty recently claims that 'the entire Grothendieck apparatus' is of weaker strength than finite order arithmetic [39].

of classical mathematics[13]. The approach from physics on the other hand by identifying the arrow of category theory with process in nature greatly simplifies the complexity and enables category theory to act as an Occam's razor and as a very powerful scientific tool. For a parallel view as an alternative natural philosophy, see [10].

4 PROCESS IN CATEGORY THEORY

Process is concerned on the one hand with the global/local distinction and on the other hand by the juxtapositioning between the stationary and the non-stationary. The universal and the particular, the static and the dynamic are integrable at the level of metaphysical reality and are formally representable and accessible in natural categories that follow physics. Process relates not just to the non-stationary but subsumes both the static and the dynamic. One is contained in the other but which way round? Such problems like Zeno's paradox of the arrow's dynamic flight consisting of only static positions are avoided in the 17th century French logic school of the Port Royal, following from Aristotle's first and second intentions, as distinguishing the intension from the extension. Information systems are an important example of the need to represent the world by computer. Within such systems, the limitations of modelling reality are conspicuous in database methods, as discussed later.

Natural categories as a metaphysics provide mixed levels for intension/extensions. Intension and extension alternate in a preorder, that is with arbitrary beginning. This is the natural role of the categoreal arrow with an identity arrow as intension and a distinguishable valued arrow for extension. The simplest identity arrow is treated as an object, the next higher identity arrow (the functor) composed of extensional arrows between objects makes a category with ordinary functors as extensional arrows between categories. The next higher level

13 The reason for this is that mainstream pure mathematics is still operating on the basis of Whitehead's earlier PM which is limited to first order theories. See the recent [9] as an example of lengthy descriptions needed for the more restrictive use of a model Topos.

is a category of these categories with objects as categories and functors between them. The highest level arrow is again an identity natural transformation which composes the previous level of categories as objects with natural transformations between them. It is this final identity natural transformation that constitutes a topos. The whole is just a recursive system with closure at four levels consisting of three open interfaces. The identity natural transformation is scale invariant and any higher level would only be a reformulation of this same level. This is process and the Universe is an instantiation of process but the World is even greater than the physical Universe consisting of all the relations between physical entities and all the relations between those relations. The interaction of subjective human behaviour in the global world of physics, biology and economics is very topical.

Relationships in nature are explicable in natural categories with the single concept of adjointness that consists only of a pair of contravariant arrows inducing a monad. Mathematical categories other than the Cartesian closed are possible but natural categories being derived from physics only recognise the existence of Cartesian closed categories where every object is the domain of a covariant arrow and the co-domain of a contravariant arrow. This recursive structure applies at any of three possible levels and in general occurs between a pair of categories where adjointness of the pair of arrows contravariant to one another between the categories induces a monad (or 'triple') and a co-monad (or 'co-triple'). Each arrow has a dual role. One is the contingent arrow of intension (in the monad) and the determinant arrow of extension (in the co-monad) while the other arrow is the contingent arrow of extension (in the co-monad) and the determinant arrow of intension (in the monad). Finitary mathematics seeks to model various features of the adjointness between natural categories in a range of topics, including currying, lambda calculus, Kan extensions, Galois theory, the Yoneda lemma and its embedding.

The fundamental structural significance in the world is the way the local connects into the global such as in the McLuhan Global Village where everything is connected [40]. The temporal analysis is the distinction between stationary and the non-stationary. Philosophically this global/local distinction is not at all new. It is at the root of Zeno's

paradox of the arrow's dynamic flight consisting only of static positions.

The noun 'process' or the participle 'processing' commonly describe an act of transforming an existing object by some procedure to another form as in a manufacturing or business administration procedure. Wikipedia deals with its entry for Process in up to 40 different fields of knowledge, including philosophy, science, engineering, computing, chemistry, biology, law, business and even the 'process haircut'. There are variations in the meaning of the word depending on context. For instance in business, process describes activities or tasks that produce a specific service or product for customers. Interestingly Wikipedia does not include physics in its lists of fields for process.

The whole subject of cybernetics can be viewed as a process operating in nature as in Wiener's definition [66] involving comparison of communication in the animal and the machine [45]. Process describes the way that both animals (biological systems) and machines (non-biological or "artificial" systems) can operate according to cybernetic principles. This was an explicit recognition that both living and non-living systems can have purpose. Wiener considered that systems theory seeks to deal with the local/global divide [45], treating systems as equivalent to process but the latter is the higher form. The early specification of the working of the brain in cybernetics by Ashby [2] amounts to the concept of process but it was von Bertalanffy of the early founders of cybernetics that explicitly related the latter to process [6, 7]. Whitehead had some influence on philosophy of business administration and organizational theory, through developing a focus on identifying and investigating the effect of temporal events, as opposed to static things [28].

Most writers trace process to the 'all is flux' of Heraclites in contrast to Parmenides, who is more usually associated with a static view. However, process is more than flux and also subsumes permanence. It is rather the Heraclites' *logos* which was taken up by the Greeks of Alexandria and the Judeo-Christian tradition to identify *logos* with God and the second person in the Trinity. The whole theory of evolution is process too but one where the origin of species does not unfold in a linear fashion. Evolution appears a foundational natural

process encompassing both emergence and change. Ordering is adjointness and includes both static and dynamic aspects. It is a paradox that process includes invariance[14] which describes no change under a transformation. Indeed scale invariance turns out to be an important phenomenon of process. Fractal patterns arising from scale invariant physics are studied piecemeal with use of special sets like the Mandelbrot, Julia etc. However general methods are restricted because a set cannot be a member of itself in the way that a reflective subcategory can have itself as an object. Information systems like the web also exhibit properties of scale invariance but we do not have space here to pursue this aspect of process which arises in exponential categories.

There is always the problem of where to begin. That statement may be formally expressed as *a pre-order of categories* or just as well as *a category of preorders* for both lack beginning and ending. However within process we can but focus on the *category of reality* in the sense of the category where objects and relationships between objects exist to make up the physical world. This is metaphysical process and the Universe is an instantiation of process but the World is even greater than the physical Universe consisting of all the relations between physical entities and all the relations between those relations. Physical relations connect directly from higher-order relations. This is treated bottom-up but because of the holistic nature of process it is driven top-down. A topical example is the recent realisation of how subjective human behaviour affects the objective syntax of world economies. Current practical examples of applied recursion across levels is deduplication in structured data storage [32] or functional DNA nanostructures that can be integrated into larger structures as miniature circuit boards in bioengineering [47].

In this sense the *World* is greater than the physical Universe of cosmology. There is a unique arrow from the source of the World to every object in it and a unique resultant arrow between any pair of objects as in Figure 7.1.

For we are concerned with the higher order of relations between physical objects and the relations between those relations which to-

14 The subject of invariance was mainly developed in the 19th century by Arthur Cayley. Saunders Mac Lane [41] traces the early origins of category theory to Cayley.

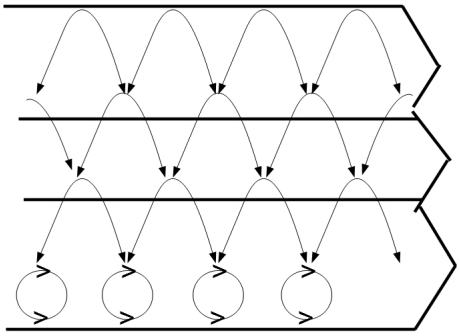


Figure 7.1: A Schematic World-Universe Relationship. The lowest horizontal arrow is a category consisting of a row of notional elementary objects connected by a row of vertical arrows which are themselves connected by a higher row which are in turn connected by a yet higher row.

gether with the physical objects of the Universe make up the World. This then embraces the whole of human affairs and activity including the arbitrary disciplines of philosophy and theology. Existence in categories is identifiable with the object which as we shall see is the condition known categorically as 'Cartesian'. Ordering is adjointness and includes both static and dynamic aspects.

This empirical knowledge that 'every entity that exists is related to every other object that exists' is no more than a definition of the Universe to include everything naturally accessible. This provides a unique direct arrow between any pair of objects that is the composition of all possible arrows between them. This is the structure given the label preorder. Figure 7.3 presents a two-dimensional representation for the context of the objects C and A of a preorder. There is but one unique arrow between any pair of objects in a preorder and that arrow as the figure shows is the limit of all other possible arrows whether directly between the pair or indirectly via any other pair in the preorder. We cannot assume any orientation for it or even presume the concept of a dimension. It is possibly easier to imagine than to draw, although our common perception wants us to imagine it in

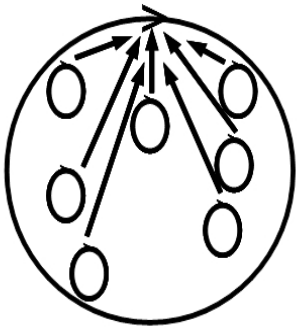


Figure 7.2: Terminal Object. A category with a terminal object has an arrow from every object to it. In preorders this arrow is unique as in Figure 7.3

three dimensions or possibly in higher order algebraic or geometric dimensions but lies easily in higher-order geometric dimensions. A process preorder does not exist in any space whether algebraic or geometric. Rather it should be space free. This is the quantum world. However the effect between entities is mutual and the arrow is therefore two-way but not symmetrical because the opposing directions give rise to a natural parity in their mutuality. This is the ultimate reality of the quantum world. Whether it is the quantum or the physical world that is true, reality seems just a matter of personal preference.

Newtonian physics treated the universe as some container either rectilinear or spherical but embedded in time. Such a structure is representable, for example by Yoneda or Curry techniques[15], to first-order as a number. This is the classical model which can be verified by measurement in first-order predicate logic because as Gödel has shown first-order predicate logic is complete for a closed world. However Gödel has also shown that such a logic is not complete for an open world and any model based on number and relying on axioms is not complete whether open or closed [20]. This effectively sets a limit to

15 See [3] at pages 118 and 190 respectively

Wigner's 'unreasonable effectiveness of mathematics in the natural sciences' [67].

5 METAPHYSICS

If we want to identify a category with reality, existence requires designation of one object as the terminal object, as shown in Figure 7.2. This is the condition known as 'Cartesian'. It is also possible to designate another as the source of the process as initial object. This is the condition known as 'co-Cartesian' but is not a necessary and sufficient condition and may therefore result in over-specification and a too constrained system. There is a free functor mapping from the preorder on to any of its partial orders. It is natural to identify the terminal object with the covariant identity functor. If the initial object exists it would exist as the contravariant identity functor of the category. Nevertheless although these are arbitrary terms the use of the labels 'terminal' and 'initial' imports an interpretation and requires the existence of some axiom of choice, which is an axiom/assumption of set theory. Whitehead raises the stakes by ascending two levels from models to speculative metaphysics but nevertheless we can by the use of formal categories eliminate 'the speculative' and indeed any amount of assumption. Here we try to make no assumption at all beyond that the World exists. We try to keep open issues about terms such as 'terminal' and 'initial' because they may be related to what cosmologists currently tell us about the fabric of the physical Universe consisting mainly of dark matter and dark energy with only 4% in familiar forms.

Whitehead's *Process & Reality* leaps across current fundamentals at the frontiers rather than attempting incremental advances within existing knowledge. But how do the general and the particular relate within the structure of the world? Any formal description needs to be able to combine both the global and the local. This is possible with natural categories by substituting Whitehead's interpretation of metaphysical process in *Process and Reality* [61] for that in his earlier understanding in the *Principia* [54], which was the starting point for the

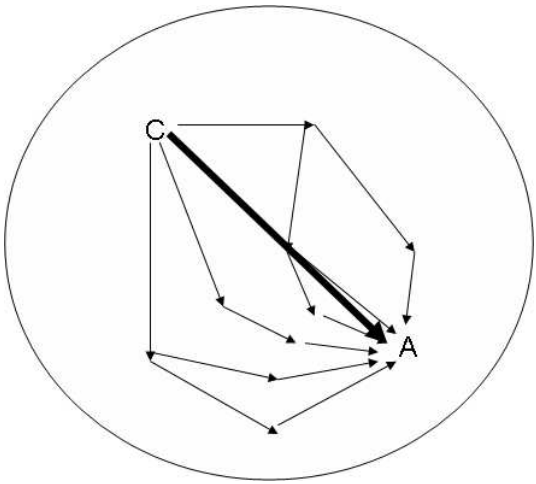


Figure 7.3: The unique arrow from C to A as a limit of other arrows in a unidirectional Preorder. The co-limit sums over all others

traditional finitary category theory of Eilenberg and Mac Lane [41]. It is the difference between a metaphysics and modelling which are separated by two levels as in Figure 7.4 (diagram 16 in [46]). Metaphysics is one level up from reality in human perception [17] while models are one level down. The limitations of modelling reality can be seen in information systems where there is a need to represent the world on computers. Problems are evident in database methods like ACID [22, 28] and in Codd's pure relational model [13]. In database design, data normalisation is used to attempt to match the logical data structures to the physical world. This method of design has a number of unsatisfactory features. Firstly it is difficult to enforce the laws of the physical world in the operational database and secondly the theoretical underpinning, based on set theory, is not natural because of the problem of representing arrows across multiple levels as functions.

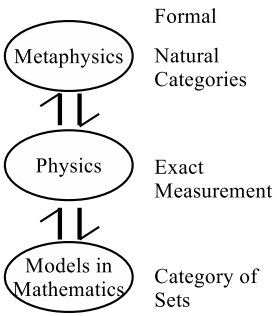


Figure 7.4: The Staircase of the World from Metaphysics to Models

6 FINITARY CATEGORIES MODEL NATURAL (METAPHYSICAL) CATEGORIES

Whitehead developed his theory *Process and Reality* in what he terms 'speculative metaphysical categories' These are in great contrast to the formal principles he enunciated with Bertrand Russell in *Principia Mathematica* and Whitehead devotes Chapter 1 of his later work ([65] 4-26) to explaining in a general philosophical context why they had to be speculative. For the second half of the last century has seen substantial advances in the development within finitary mathematics of formal categories, based on the concept of the arrow and initiated by Eilenberg and Mac Lane [41]. The phrase 'finitary mathematics' is a term first coined by the mathematician David Hilbert[16] and effectively describes the whole mainstream of twentieth century mathematics built up on a system of proofs in set theory and number from incompletely specified axioms. The adjective 'finitary' is itself a little misleading as finitary mathematics includes topics like infinity and transfinite numbers as these are modelled on the finite concept of number.

16 according to Feferman [20] Hilbert never defined finitary mathematics and it collapsed at its foundations under the weight of Gödel for the reasons mentioned above.

Nevertheless it is possible to ascend the staircase in Figure 7.4 from categories as mathematical models to metaphysical categories and extend that ladder to categories that are no longer speculative but which can now be made formal thanks to the work of Eilenberg, Mac Lane and a large number of pure mathematicians world-wide who have refined and extended their original interpretation of the humble arrow based only on the four properties:

1. a morphism from domain to co-domain
2. identity from an indistinguishable domain and co-domain
3. associativity
4. composition,

There are two distinctions important for process that we need to draw. One is between metaphysical categories and finitary categories in respect of the use of number in physics; the other is between sets and either types of categories in respect of the representation of intension and extension. We will first consider the natural numbers then look at intension and extension as an intrinsic property of parity to be found in adjointness.

7 THE FINITARY CATEGORY OF THE NATURAL NUMBER

Because it relies heavily on experiment, physics as a discipline has become identified with measurement and number as its prime conceptual tool. Consequently it has become very bound up with sets which equate to number. However it is an assumption that qualities and quantities are representable as number. The physics and the mathematics have become merged so that space is a complex number whether it is Newton's Universe as a container or the infinite Hilbert space of quantum mechanics. These it should not be forgotten are just numbers. This is fine to the extent of first order models for which Gödel (as mentioned above) has shown to be consistent but it is not sufficient for open or other higher order systems for Gödel has shown these to be neither consistent nor decidable when relying on axioms of

sets or number. This applies as much to the use of statistical methods as the interpretation of measurement. It may be possible to reduce any problem to first order but any conclusions will then be subject to the assumptions in the reduction. This is particularly insidious in treating open systems as closed. However openness is not just bound up with the concept of order for it contains a deeper logical strand of constructivism as associated with the intuitionism of Brouwer. Boolean logic suffices for a closed system but an open system requires the logic of Heyting (See Figure 7.9 below).

Metaphysical categories have therefore no natural concept of number. Finitary categories as a model relying on the concept of sets has consequently to introduce the concept of number[17]. This is achieved by postulating a Natural Number Object with a recursive definition on arrows comparable to recursive functions generating the set of natural numbers. This requires importing some undefined successor function. While this may be natural in mathematics it is not natural in physics where systems are open either externally or internally. An obvious example is radioactivity where atoms decay according to some preorder and it is not therefore possible to identify a successor before the event of decay. Of course it was to explain such events that the notion of randomness was invented but this is normally dealt with by some theory of statistical probability which leads back to the concept of number and does not provide an exact solution. This lack of a predefined successor is a feature of all open systems and a chief cause of problems of interoperability in global systems.

Open physics lacks a concept of number and therefore questions the use of finitary models in physics. The existence of multiverses must surely be the largest incarnation of the number concept. Listed below are the nine current possible theories identified by Greene [23]. These can also be compared with Barrow's views of multiverses [4].

1. Infinite space may contain a number (possibly an infinity) of universes that may lie beyond our sight.

17 First carried out by Lawvere [34] and now to be found in standard category theory texts, such as [3] at p 177.

2. Uncountable other universes with different characteristics may have been created with ours during a fleeting period of superfast accelerating expansion.
3. String theory suggests our universe is one of many 4-dimensional *brane worlds* floating in a higher-dimensional space-time.
4. A simple cycle of universes with variations in physical laws as possible in string theory.
5. More complex versions of cyclic universes.
6. Quantum mechanics allows/requires many worlds to exist in parallel formed by a branching of the wave function.
7. The universe is a holographic projection.
8. We are just one of a set of artificial universes created in simulation on a super-advanced computer.
9. The philosophical necessity that every possible universe must be realised somewhere.

It is instructive to review Greene's list from the process perspective. The list does not claim to be exhaustive and is an example of undecidability demonstrating how the use of number leads to degeneracy with many possible forms. This degeneracy is well borne out in the thorough examination of n-categories carried out by Leinster [15]. It may well be a comparable defect in string theory that allows variations in physical laws. In process categories, physical laws arise from properties of adjointness whose *bonum esse* is uniqueness. Furthermore about half the items in the list depend on some idea of infinity. But infinity belongs to mathematics, not to physics. It was David Hilbert, the proponent of finitary mathematics, who with the paradox of his Hotel Infinity recognised that infinity is always beyond reach and therefore cannot plausibly exist in physical reality. Infinity in finitary mathematics seems no more than a model of repleteness[18] under

18 Johnstone ([33] at p.3) defines the condition of repleteness as "that any object of the ambient category isomorphic to one in the subcategory is itself in the subcategory".

the free functor in process categories. The last item that postulates every possible universe is also derived from probability theory applied to infinity. That too fails at the Gödel hurdle of 'number'.

As anthropocentric variants on our universe with complicated theories reminiscent of epicycles, multiverses bear an almost Ptolemaic resemblance. The 'super-advanced computer' is a science-fiction vision of current commercial computers. They have not been thought through with respect to quantum computation nor any general attention paid to boundary conditions nor to the relativistic nature of time which Whitehead would carefully respect [61].

8 LOGICAL STRUCTURE OF WORLD REPRESENTATION AS ADJOINTNESS

In terms of natural categories, process is adjointness. This is the formal metaphysics of real existence such that every physical entity in the Universe affects every other. There is at the most a single pair of arrows in opposite directions between any pair of objects. These are limits of all the possible paths around the Universe between any given pair. This limit reduces to a single function as an abstraction in lambda calculus or as a resultant in vector analysis (for first order models lose the resolution of the contravariant pair). There are four levels involving three interfaces. The uppermost level is the intension and the lowest is the extension corresponding respectively to the global and the local. The intermediate interface connects intension and extension, that is snaps the local into the global for all time and space. Any set-theoretic approach finds this latter mechanism, which is essential to all studies of globalisation and interoperability, very difficult if not impossible as recognised by Russell's paradox.

Nevertheless in finitary categories the mathematics of adjointness has been developed in this concept termed a Cartesian closed category, derived as an abstraction of the Cartesian product but this description from historic origins may by its simplicity mislead as to its great power and content. The finitary approach is to distinguish the two properties of Cartesian closed and locally Cartesian closed but in

process categories it is that natural distinction between intension and extension. This paper provides an introduction to that formal description of the mathematical structure of the World as found in nature.

To the global/local distinction must be added the stationary against the non-stationary. Both the static and the dynamic are formally representable and accessible in the logic of natural categories. Process relates not just to the non-stationary but subsumes both the static and the dynamic. One is contained in the other but which way round? Such problems, like Zeno's paradox of the arrow's dynamic flight consisting of only static positions are avoided in the 17th century French logic school of the Port Royal [1] (harking back to Aristotle's first and second intentions) by distinguishing the intension from the extension. Aristotle referred to them as first and second intentions. Because of their extended meaning these terms were recognised in the subject of logic by retaining the older spelling with an "s" rather than a "t". When the old subject of logic was superseded around 1900 by symbolic logic based on set theory, the intension/extension relationship became rather lost until the development of computer programming revived it with the need for rigorous typing.

The intension-extension relationship is recursive; thus in the diagram of Figure 7.4 metaphysics is the intension for reality as its extension and reality itself becomes the intension for models as possible extensions. In the natural categories of metaphysics process is adjointness. This is no more than the formal metaphysics of real existence that every physical entity in the Universe affects every other. There is at the most a single pair of arrows in opposite directions between any pair of objects. These are limits of all the possible paths around the Universe between any given pair. This limit is that of the preorder in Figure 7.3. Mathematical categories other than the Cartesian closed are possible but process categories being derived from physics only recognise the existence of Cartesian closed categories which has the property of adjointness. Every object is the domain of a covariant arrow and the co-domain of a contravariant arrow. This recursive structure of intension/extension applies at any level but is best studied between a pair of categories (identity functors $\mathbf{1}_F$ and $\mathbf{1}_G$) where adjointness of the pair of arrows (F and G, contravariant to one another)

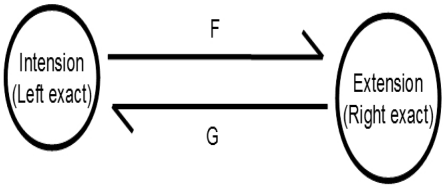


Figure 7.5: Adjointness $F \dashv G$

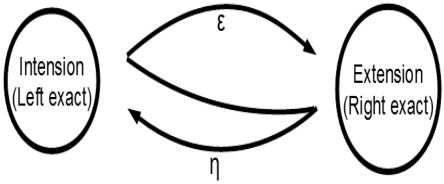


Figure 7.6: Adjointness expressed with natural transformations η and ϵ

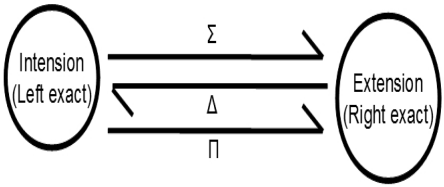


Figure 7.7: Adjointness $\Sigma \dashv \Delta \dashv \Pi$

induce a monad consisting of a triple $< \mathsf{T}, \eta, \mu >$ and a co-monad consisting of the co-triple $< \mathsf{S}, \epsilon, \delta >$. Figure 7.5 shows the adjointness between the categories, intension and extension.

Each arrow has a dual role. F is the contingent arrow of intension and the determinant arrow of extension while G is the contingent arrow of extension and the determinant arrow of intension. T is just the composition GF and S the composition FG. Each of these compositions may be compared in Figure 7.6 at the next level up with the contribution they make to their respective identity functors by means of the creative unit of adjunction $\eta : \mathbf{1}_F \longrightarrow \mathsf{GF}$; and the qualitative co-unit of adjunction $\epsilon : \mathsf{FG} \longrightarrow \mathbf{1}_G$. Comparison at the even higher level of order is provided by the unit of potentiality $\mu : \mathsf{T}^2 \longrightarrow \mathsf{T}$; and its co-unit $\delta : \mathsf{S} \longrightarrow \mathsf{S}^2$. There are special cases of the latter two which may be interpreted [46] as in the 'dimension of time' with the unit of anticipation where potentiality is by hindsight and the co-unit of anticipation by foresight. Although there are never more than two basic adjoint functors $\mathsf{F} \dashv \mathsf{G}$, the combined composition of their two compositions T and S may be resolved into the three basic functors of Figure 7.7 to be found in standard category theory texts, where Σ is the existential qualifier, Π the universal quantifier and Δ the stability diagonal pullback functor. The interplay of left and right adjointness with left and right exactness is a little subtle [26] and can be better understood in the exploded diagram of Figure 7.8 which is repeated in Figure 7.9 to show an exploded view of the natural intuitionistic logical structure of the Cartesian closed category.

9 THE NATURAL WORLD STRUCTURE AS A CARTESIAN CLOSED CATEGORY

Relationships in nature are therefore all explicable in process categories with this single concept of adjointness [14] that consists only of a pair of contravariant arrows inducing a monad. In finitary categories the mathematics of adjointness has been developed in what is termed a Cartesian Closed Category, derived as an abstraction of the Cartesian product but this description from historic origins may

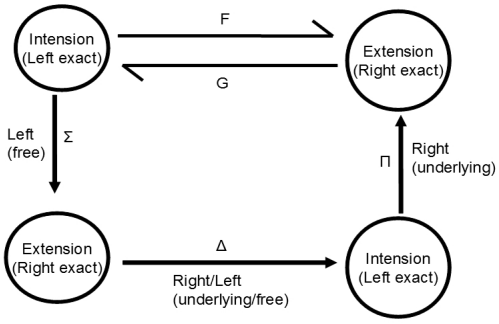


Figure 7.8: Explosion $\Sigma \dashv \Delta \dashv \Pi$ of the Arrow Functors of Adjointness $F \dashv G$

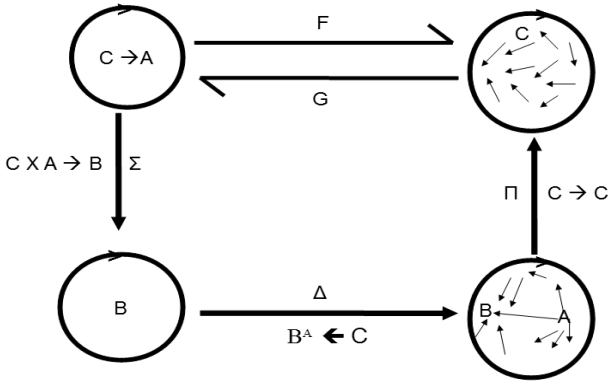


Figure 7.9: Intuitionistic structure of the Cartesian closed category: Exploded view of Heyting logic.

by its simplicity mislead as to its great power and content. The finitary approach is to distinguish the two properties of Cartesian closed and locally Cartesian closed but in process categories it is that natural distinction between intension and extension that provides a formal description of the mathematical structure of the World as found in nature. It is the simple principle that everything in the world is related to everything else in the world that provides the formal structure of the relationship relevant to any scientific study or technological application requiring an understanding of these relationships.

An early example is the representation of information in computers, that needed some implementable model of real-world relationships. Some variation of the hierarchical was possibly the most common structure attempted in different knowledge systems. But the most successful measured by the volume of commercial transactions was by far the simple relational model based on lists or tables manipulated as sets embodying an intension/extension relationship.

The Cartesian closed category (CCC) is a fundamental category of category theory. Its features and their definitions are to be found in its standard textbooks but most if not all come from the stationary viewpoint of set theory, not from process. That set theory itself does not rest on unequivocal foundations may raise few problems in pure mathematics where axioms may be defined at will and may well be adequate too in applied mathematics to a first order. However, many problems requiring mathematical solutions today arise in more complex situations. Transactions in information systems [43] are a case in point as of the nature of process. Thus a common approach in databases [14] is to adopt the principles under the acronym ACID stating the requirements for Atomicity, Consistency, Isolation and Durability. The aim is to ensure that a transaction involving a series of operations is indivisible, enforces all rules, provides results only on termination and guarantees to hold the results under any circumstances. The transaction concept has been implemented efficiently on many database systems but in information systems as a whole the idea lacks the abstraction needed for successful business modelling. The alternative approach in natural philosophy is that of process as explored in the 20th century [28].

While in the formal language of category theory the world may be described as 'Cartesian-closed', this term may give a false impression that it has a Cartesian coordinate system which is unfortunate but the phrase has arisen historically in that context because it embodies the fundamental concept of the Cartesian product. In fact it is much more than a simple product and these terms need to be examined further. For while natural categories and metaphysics provide us with a process structure for the world, we can only begin to investigate it here. Intension and extension alternate in a preorder, that is with an arbitrary beginning of an intension with an extension which itself becomes an intension of the next extension and so on as in Figure 7.10 [29].

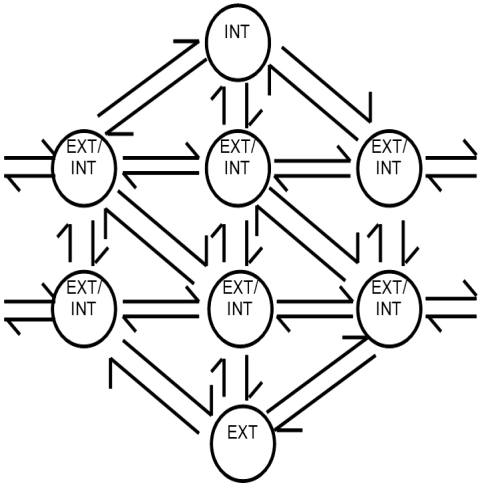


Figure 7.10: Alternate Intension/Extension Pairs in Nature

10 THE TOPOS: ARCHETYPE OF NATURAL WORLD

The archetype of the natural world is the topos, in its early days formally defined as a Cartesian closed category with subobject classifiers and informally as a generalised set. Johnstone in his preface to [33]

lists thirteen alternative descriptions that have been applied to the topos (pp.viii *et seq*). Many of them like for instance "A topos is a generalised space" still carry hangovers from sets. We would recommend as an informal definition: "The category of categories of categories". To some this may only confirm categories as "abstract nonsense" but it is accurate and and makes explicit the recursion. The topos sums up all that we have said in this paper. It is the ultimate intension existing as an identity natural transformation in any extension given by the internal categories, subject to the locally Cartesian closed condition with the preorder structure and an intuitionistic logic that is the Heyting and which is more general than the Boolean. There is a unique arrow from the source of the World to every object in it and a unique limiting arrow between any pair of objects.

To satisfy its holistic nature the World must emerge top-down. That is to say no more than that if the Big Bang happened it potentially contained everything that ever existed[19]. However it is easier to explain bottom-up by treating the role of the arrow as a natural expression of process with an identity arrow as intension and a distinguishable valued arrow for extension. Nevertheless while in natural category theory the simplest identity arrow may be treated as an object, it is convenient to begin with a category of three composing objects as a generalisation of any possible category. This is shown in Figure 7.11 with the next higher identity arrow (the functor) composing extensional arrows between objects. The next higher identity arrow is the locally Cartesian closed natural transformation composing categories with ordinary functors as extensional arrows between categories as shown in Figure 7.12. The highest level arrow is also a natural transformation which composes structures of categories and functors. It is this identity natural transformation that constitutes the full Cartesian closed category of a topos as in Figure 7.13. However, the natural arrow is double-headed as a composition of the adjoint functors but with a parity as previously discussed above. Although as just explained it may be easier to understand these diagrams bottom-up in the way that models are usually built-up, nevertheless process can only exist as a whole and the full diagram represents a natural

19 formally (μ, δ) in the description above for the monad/comonad.

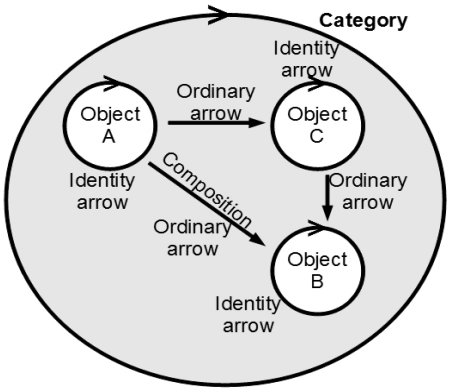


Figure 7.11: A category consists of ordinary arrows composing between identity arrows as objects

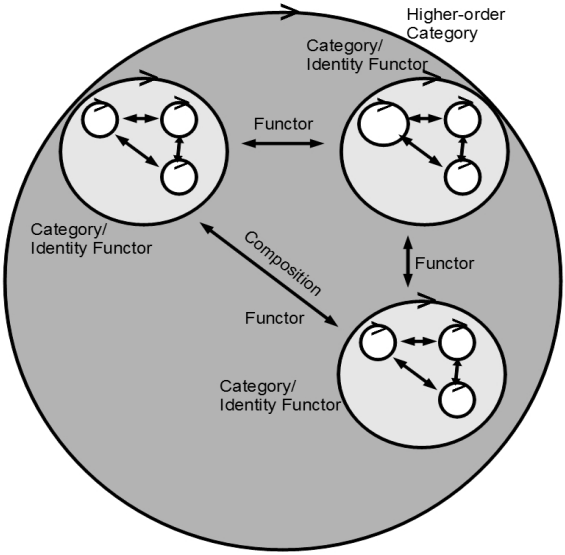


Figure 7.12: Functors between Categories compose to form higher-order categories. A category is just an identity functor

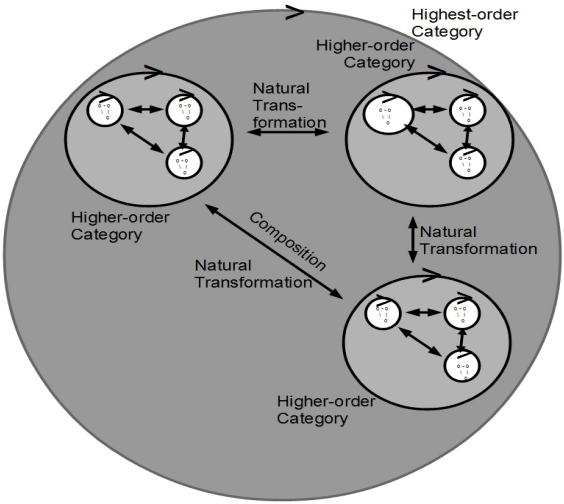


Figure 7.13: Natural Transformations of Composing Functors themselves compose in the highest possible category, a Topos

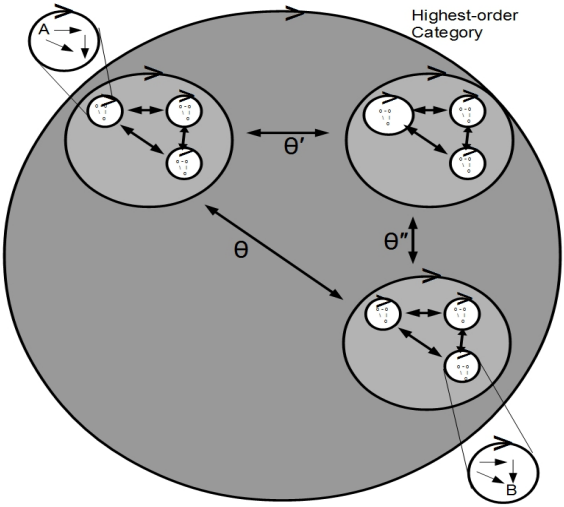


Figure 7.14: A topos showing natural path from any object A to any object B

occasion or "actual event" as first introduced by Whitehead [61] and discussed by McHenry [38].

The whole is just a recursive system with closure at four levels consisting of three open interfaces. Figure 7.13 shows the three interfaces for composing arrows (ordinary, functor, natural transformation) with the four levels (identity arrow, identity functor/category, higher-order identity functor/category, identity natural transformation/topos). The diagram shows well the natural recursive nature of the structure. It also demonstrates connectivity from any object to any other object. It is possible therefore, as shown in Figure 7.14, to get from any object A to any object B directly: $B = \theta A$, or indirectly with possible local variations through any other internal path: $\theta'' \circ \theta' A = B$. This is a natural structure because it is obtained from simple induction applied to the notion of process without any assumptions.

11 CONCLUSIONS

Only a few 'simple' concepts are therefore needed: the World is a topos with monadic objects related by contravariant functors with natural transformations as units of adjunction. These are sufficient to identify formally the Whiteheadian vocabulary of the likes of the ontological principle, actual entities and occasions, eternal objects, congrescence, creative and emotive advance of becoming, public and private, prehensions, nexus, primordial nature, emergence, etc, together with their other postmodern counterparts. The details of these will be pursued at a later date.

12 BIBLIOGRAPHY

[1] Arnauld, Antoine, & Nicole, Pierre, *Logic or the Art of Thinking*, Translated Buroker, Jill Vance, Cambridge (1996).

[2] Ashby, W R, *Design for a Brain*, Chapman & Hall (1952).

[3] Barr, M, & Wells, C, *Category Theory for Computing Science*, 1st edition, Prentice Hall 432pp (1990).

[4] Barrow, John D, *The Book of Universes*, Bodley Head (2011).

[5] Bastin, T, & Kilmister, C W, *The Origin of Discrete Particles*, Series on Knots and Everything **42** 196pp (2009).

[6] Bertalanffy, L von, *The Theory of Open Systems in Physics and Biology* Science **111** 23-29 (1950).

[7] Bertalanffy, L von, *General Systems Theory*, Penguin, London (1971).

[8] Bogaard, Paul A, & Bell, Jason (edd), The Harvard Lectures of Alfred North Whitehead, 1924-1925, Philosophical Presuppositions of Science, *The Edinburgh Critical Edition of the Complete Works of Alfred North Whitehead* Edinburgh (2017).

[9] Caramello, Olivia, *Theories, Sites, Toposes: Relating and Studying Mathematical Theories Through Topos-Theoretic 'bridges'* Oxford (2018).

[10] Clarke, Bowman L, *A Calculus of Individuals Based on "Connection"*, Notre Dame Journal of Formal Logic **22** 204-218 (1981).

[11] Clarke, Bowman L *Individuals and Points*, Notre Dame Journal of Formal Logic **26** 61-75 (1985).

[12] Casati, R & Varzi, A C *Parts and Places: The Structures of Spatial Representation*, MIT Press Cambridge (1999).

[13] Codd, E F, *The Relational Model for Database Management*, Addison-Wesley (1990).

[14] Connolly, T, & Begg, C, *Database Systems*, Pearson (2004).

[15] Epperson, Michael *Quantum mechanics and the philosophy of Alfred North Whitehead* Northern University Press (2004), Oxford (2012).

[16] Epperson, Michael & Zafiris, E, *Foundations of Relational Realism: A Topological Approach to Quantum Mechanics and the Philosophy of Nature* Lexington Books New York; Rowman & Littlefield (2013).

[17] Faber, Roland, Bell, Jeffrey A, & Petek, Joseph, edd. *Rethinking Whitehead's Symbolism Thought, Language, Culture* Edinburgh (2017).

[18] Fitzgerald, Janet A, *Alfred North Whitehead's Early Philosophy of Space and Time*, University Press, Washington (1979).

[19] Gödel, Kurt, *Über die Vollständigkeit des Logikkalküls* PhD thesis, University of Vienna (1929); *Die Vollständigkeit der Axiome des logischen Funktionenkalküls Monatshefte für Mathematik und Physik* **37** 349-360 (1930); *The completeness of the Axioms of the Functional Calculus of Logic* Feferman and al, Collected Works 102-122 German even, 103-123 English odd. Translated by Stefan Bauer-Mengelberg, read and approved by Gödel after some accommodation. Subsequent minor additions by Jean van Heijenoort (collected works **1** 59-101, Feferman) (1986).

[20] Gödel, Kurt, *Die Vollständigkeit der Axiome des logischen Funktionenkalküls Monatshefte für Mathematik und Physik* **37** 349-360 (1930). The completeness of the Axioms of the Functional Calculus of Logic. Feferman and co: Collected Works 102-122 German even, 103-123 English odd. Translated by Stefan Bauer-Mengelberg, read and approved by Gödel after some accommodation. Subsequent minor additions by Jean van Heijenoort (p.59 collected works **1** p.59, Feferman) (1986).

[21] Grassmann, Hermann, *Die Lineale Ausdehnungslehre, ein neuer Zweig der Mathematik* Wiegand, Leipzig (1844).

[22] Gray, Jim, *Thousands of Debit/Credit Transactions-Per-Second: Easy and Inexpensive*, Microsoft Technical Report Series, MSR-TR-2005-39 (2005).

[23] Greene, Brian, *The Hidden Reality: Parallel Universes and the Deep Laws of the Cosmos* Allen Lane Hardback (2011).

[24] Heather, M A, & Rossiter, B N, *Locality, Weak or Strong Anticipation and Quantum Computing I. Non-locality in Quantum Theory* International Journal Computing Anticipatory Systems **13** 307-326 (2002).

[25] Heather, M A, & Rossiter, B N, , *Locality, Weak or Strong Anticipation and Quantum Computing. II. Constructivism with Category Theory* International Journal Computing Anticipatory Systems **13** 327-339 (2002).

[26] Heather, M.A. & Rossiter, B N, *Don't Confuse Left and Right Exactness! A Note correcting the Distinction in the Proceedings of ANPA 27*, ANPA 29, Exactness, Cambridge, 157-159 (2008).

[27] Heather, M.A. & Rossiter, B N, *The Process Category of Reality*, ANPA 31, Cambridge 224-262 (2011).

[28] Hernes, Tor, *A Process Theory of Organization* Oxford (2014).

[29] Heather, Michael, Livingstone, David, & Rossiter, Nick, *Logical Foundations for the Infrastructure of the Information Market*, I-ESA, Berlin, 25-28 March, Enterprise Interoperability III, New Challenges and Industrial Approaches, Mertins, Kai, Ruggaber, Rainer, Popplewell, Keith and Xu, Xiaofei, (edd) 624- 637 (2008).

[30] Hurley, Patrick, J (tr), *The relational theory of space*, Philosophy Research Archives **5** 712-741 (1979).

[31] Hurley Patrick J. *Russell, Poincaré, and Whitehead's 'Relational Theory of Space'* Process Studies, **9** Numbers 1 & 2, Spring-Summer, pp. 14-21 (1979).

[32] IBM, *Data deduplication in Tivoli Storage Manager* (2010).

[33] Johnstone, P T, Sketches of an Elephant, A Topos Theory Compendium, *Oxford Logic Guides* 43, Clarendon (2002).

[34] Lawvere, F W, An elementary theory of the category of sets, *Proc Nat Acad Sci*, **52** 1506-1511 (1964).

[35] Lawvere, F W, Adjointness in Foundations, *Dialectica*, **23** 281-296 (1969).

[36] Leinster, Tom, *Higher Operads, Higher Categories*, London Mathematical Society LNS 298, Cambridge (2004).

[37] Linsky, Bernard, *The Evolution of Principia Mathematica: Bertrand Russell's Manuscripts and Notes for the Second Edition* Cambridge (2011).

[38] McHenry, Leemon, *The Event Universe: The Revisionary Metaphysics of Alfred North Whitehead* Edinburgh (2015).

[39] McLarty, Colin *Grothendieck's cohomology founded on finite order arithmetic*, Invited address 2012-2013 Winter Meeting of the Association for Symbolic Logic, San Diego (2013).

[40] McLuhan, Marshall, *The Gutenberg Galaxy: The Making of Typographic Man* Toronto 263pp (1962).

[41] Mac Lane, S, *Categories for the Working Mathematician*, 2nd edition, Springer-Verlag, New York (1998).

[42] Penrose, Roger, *The Road to Reality, A Complete Guide to the Laws of the Universe* Jonathan Cape London (2004).

[43] Rossiter, Nick, Heather, Michael, & Nelson, David, *A Natural Basis for Interoperability*, in: Enterprise Interoperability: New Challenges and Approaches, Doumeingts, G, Muller, J, Morel, G, & Vallespir, B, (edd) 417-426 (2007).

[44] Rossiter, B N, Heather, M A, & Sisiaridis, D, Process as a World Transaction, *Proceedings ANPA* 27 Conceptions, 122-157 (2006).

[45] Rossiter, Nick, & Heather, Michael, *Free and Open Systems Theory*, EMCSR-2006, Cybernetics and Systems, 18th European Meeting on Cybernetics and Systems Research, University of Vienna, 18-21 April 2006, Trappl, R, (ed) 1 27-32 (2006).

[46] Rossiter, Nick, & Heather, Michael, The Natural Metaphysics of Computing Anticipatory Systems, *IJCAS* **23** 3-20 (2011).

[47] Rothemund, P W K, Folding DNA to create nanoscale shapes and patterns, *Nature* **440** 297-302 (2006).

[48] Russell, Bertrand, *Our Knowledge of the External World as a Field for Scientific Method in Philosophy*, Allen & Unwin, London (1906)

[49] Russell, Bertrand *Our Knowledge of the External World as a Field for Scientific Method in Philosophy*, The Open Court Publishing Company, Chicago and London (1914).

[50] Russell, Bertrand, *The Autobiography of Bertrand Russell*, Volume II: 1914-1944. George Allen and Unwin, London (1968).

[51] Simons, P M *Parts. A Study in Ontology* Clarendon, Oxford (1987)

[52] Whitehead, Alfred North, *A Treatise on Universal Algebra* Cambridge (1898).

[53] Whitehead, Alfred North, *On Mathematical Concepts of the Material World* Philosophical Transactions of the Royal Society of London, Series A, **205** 465-525 (1906).

[54] Whitehead, Alfred North, & Russell, Bertrand, *Principia Mathematica* **1** Cambridge University Press (1910).

[55] Whitehead, Alfred North, *The Relational Theory of Space* (1914).

[56] Whitehead, Alfred North, *The Relational Theory of Space* First Congress of Mathematical Philosophy, Paris, 8th April 1914(1914); *La Theorie Relationiste de l'Espace*, Revue de Metaphysique et de Morale **23** 423-454 (1916).

[57] Whitehead, Alfred North, *An Enquiry concerning the Principles of Natural Knowledge* Cambridge (1919).

[58] Whitehead, Alfred North, *The Concept of Nature* Cambridge (1920).

[59] Whitehead, Alfred North, *The Principle of Relativity with Applications to Physical Science* Cambridge (1922).

[60] Whitehead, Alfred North, *Science and the Modern World* Cambridge (1926).

[61] Whitehead, Alfred North, *Process and Reality: An Essay in Cosmology*, Macmillan, New York (1929).

[62] Whitehead, Alfred North, *Adventures of Ideas* Cambridge (1933).

[63] Whitehead, Alfred North, *Nature and Life* Chicago (1934).

[64] Whitehead, Alfred North, *Essays in Science and Philosophy*, Rider, London, republished Greenwood Press in 1968 (1948).

[65] Whitehead, Alfred North, *Process and Reality: An Essay in Cosmology*, corrected edition, edd D R Griffin, D W. Sherburne, New York, The Free Press (1978). https://monoskop.org/images/4/40/Whitehead_Alfred_North_Process_and_Reality_corr_ed_1978.pdf.

[66] Wiener, Norbert, *Cybernetics, or Control and Communication in the animal and the machine* John Wiley (1948).

[67] Wigner, E P, The Unreasonable Effectiveness of Mathematics in the Natural Sciences, *Comm Pure and Applied Math* **13** 1-14 (1960).

ITERANTS, BRAIDING AND THE DIRAC EQUATION

LOUIS H. KAUFFMAN

Department of Mathematics, Statistics and Computer Science
University of Illinois at Chicago
851 South Morgan Street
Chicago, IL, 60607-7045
kauffman@uic.edu

1 INTRODUCTION

The simplest discrete system corresponds directly to the square root of minus one, when the square root of minus one is seen as an oscillation between plus and minus one. This way thinking about the square root of minus one as an *iterant* is explained below. More generally, by starting with a discrete time series of positions, one has immediately a non-commutativity of observations and this non-commutativity can be encapsulated in an iterant algebra as defined in Section 2 of the present paper. Iterant algebra generalizes matrix algebra and we shall see how it can be used to formulate the algebra of the framed Artin Braid Group, the Lie algebra $su(3)$ for the Standard Model for particle physics, the framed braid representations for Fermions of Sundance Bilson-Thompson and the Clifford algebra for Majorana Fermions. This paper is a sequel to [9] and [6, 7, 8, 9] and it uses material from these papers. This paper incorporates new results of the author that first appear in the joint paper of the author and Rukhsan Ul- Haq [10].

Distinction and processes arising from distinction are at the base of the described world. Distinctions are elemental bits of awareness. The world is composed not of things but processes and observations. We will discuss how basic Clifford algebra comes from very elementary processes such as an alternation of $+-+-+-\cdots$ and the fact that

one can think of $\sqrt{-1}$ itself as a temporal iterant, a product of an ϵ and an η where the ϵ is the $+-+-+-\cdots$ and the η is a time shift operator. Clifford algebra is at the base of this mathematical world, and the fermions are composed of these things.

Section 2 and 3 are an introduction to the process algebra of iterants and how the square root of minus one arises from an alternating process. Section 4 shows how iterants give an alternative way to do 2×2 matrix algebra. The section ends with the construction of the split quaternions. Section 5 considers iterants of arbitrary period (not just two) and shows, with the example of the cyclic group, how the ring of all $n \times n$ matrices can be seen as a faithful representation of an iterant algebra based on the cyclic group of order n. We then generalize this construction to arbitrary non-commutative finite groups G. Such a group has a multiplication table ($n \times n$ where n is the order of the group G.). We show that by rearranging the multiplication table so the identity element appears on the diagonal, we get a set of permutation matrices that represent the group faithfully as $n \times n$ matrices. This gives a faithful representation of the iterant algebra associated with the group G onto the ring of $n \times n$ matrices. As a result we see that iterant algebra is fundamental to all matrix algebra. Section 5 ends with a number of classical examples including iterant representations for quaternion algebra.

Section 6 discusses the iterant structure of the framed Artin braid group. In Section 7 we apply this to a formulation of the particle model of Sundance Bilson-Thompson [31], using framed braids. In Section 7 we give an iterant interpretation of the $su(3)$ Lie algebra for the Standard Model using [30]. In Section 8 we apply this point of view on the Standard Model to obtain an embedding of the framed braid algebra for the Sundance Bilson-Thompson model into the iterant version of $su(3)$. These three sections are an account of research of the author and Rukhsan Ul-Haq in [10].

Section 9 discusses how Clifford algebras are fundamental to the structure of Fermions. We show how the simple algebra of the split quaternions, the very first iterant algebra that appears in relation to

the square root of minus one, is in back of the structure of the operator algebra of the electron. The underlying Clifford structure describes a pair of Majorana Fermions, particles that are their own antiparticles. These Majorana Fermions can be symbolized by Clifford algebra generators a and b such that $a^2 = b^2 = 1$ and $ab = -ba$. One can take a as the iterant corresponding to a period two oscillation, and b as the time shifting operator. Then their product ab is a square root of minus one in a non-commutative context. These are the Majorana Fermions that underlie an electron. The electron can be symbolized by $\phi = a + ib$ and the anti-electron by $\phi^\dagger = a - ib$. These form the operator algebra for an electron. Note that

$$\phi^2 = (a + ib)(a + ib) = a^2 - b^2 + i(ab + ba) = 0 + i0 = 0.$$

This nilpotent structure of the electron arises from its underlying Clifford structure in the form of a pair of Majorana Fermions. Section 9 then shows how braiding is related to the Majorana Femions. Section 10 discusses the structure of the Dirac equation and how the nilpotent and the Majorana operators arise naturally in this context. This section provides a link between our work and the work on nilpotent structures and the Dirac equation of Peter Rowlands [33]. We end this section with an expression in split quaternions for the the Majorana Dirac equation in one dimension of time and three dimensions of space. The Majorana Dirac equation can be written as follows:

$$(\partial/\partial t + \hat{\eta}\eta\partial/\partial x + \epsilon\partial/\partial y + \hat{\epsilon}\eta\partial/\partial z - \hat{\epsilon}\hat{\eta}\eta m)\psi = 0$$

where η and ϵ are the simplest generators of iterant algebra with $\eta^2 = \epsilon^2 = 1$ and $\eta\epsilon + \epsilon\eta = 0$, and $\hat{\epsilon}, \hat{\eta}$ form a copy of this algebra that commutes with it. This combination of the simplest Clifford algebra with itself is the underlying structure of Majorana Fermions, forming indeed the underlying structure of all Fermions.

This paper is a stopping-place along the way in a larger story of mathematics and physics that we are in the process of telling and exploring. To begin the story, we conclude this introduction with a formulation of the Schrodinger equation that can motivate the iterants.

1.1 ITERANTS AND THE SCHRODINGER EQUATION

We begin with the Diffusion Equation

$$\partial\psi/\partial t = \kappa\partial^2\psi/\partial x^2.$$

Consider the possibility of putting a "plus or minus" ambiguity into this equation, like so:

$$\pm\partial\psi/\partial t = \kappa\partial^2\psi/\partial x^2.$$

The $\pm$ coefficient should be lawful not random, for then we can follow an algebraic formulation of the process behind the equation. We shall take $\pm$ to mean the alternating sequence

$$\pm = \cdots + - + - + - + - \cdots$$

and time will be discrete. Then the equation will become a difference equation in space and time

$$\psi_{t+1} - \psi_t = (-1)^t\kappa(\psi_t(x - dx) - 2\psi_t(x) + \psi_t(x + dx))$$

where

$$\partial_x^2\psi_t = \psi_t(x - dx) - 2\psi_t(x) + \psi_t(x + dx).$$

But we wish to consider the continuum limit. However, there is no meaning to

$$(-1)^t$$

in the realm of continuous time. What do do? In the discrete world the wave function ψ divides into ψ_e and ψ_o where the (discrete) time is either even or odd. So we can write

$$\partial_t\psi_e = \kappa\partial_x^2\psi_o$$

$$\partial_t\psi_o = -\kappa\partial_x^2\psi_e.$$

We take the continuum limit of ψ_e and ψ_o *separately.*

In fact we can interpret the $\{\pm\}$ as the complex number i. Recall that the complex number i has the property that $i^2 = -1$ so that

$$i(A + iB) = iA - B$$

when A and B are real numbers,

$$i = -1/i,$$

and so if $i = 1$ then $i = -1$, and if $i = -1$ then $i = 1$. So i can be interpreted as oscillating between $+1$ and -1, but it does it lawfully and so we shall regard i as a definition of ± 1.

$$i = \pm 1.$$

In fact, when we multiply $ii = (\pm 1)(\pm 1)$, we get -1 because the i takes a little time to oscillate and so by the time this second term multiplies the first term, they are just out of phase and so we get either $(+1)(-1) = -1$ or $(-1)(+1) = -1$. We will formalize this point of view later in the paper.

Now ± 1 behaves quite lawfully and we can write

$$\psi = \psi_e + i\psi_o$$

so that

$$\begin{aligned} i\partial_t\psi = i\partial_t(\psi_e + i\psi_o) = i\partial_t\psi_e - \partial_t\psi_o \\ = i\kappa\partial_x^2\psi_o + \kappa\partial_x^2\psi_e = \kappa\partial_x^2(\psi_e + i\psi_o) \\ = \kappa\partial_x^2\psi. \end{aligned}$$

Thus

$$i\partial\psi/\partial t = \kappa\partial^2\psi/\partial x^2.$$

This the Schrödinger equation. Instead of the simple diffusion equation, we have a mutual dependency where the temporal variation of ψ_e is mediated by the spatial variation of ψ_o and vice-versa.

$$\psi = \psi_e + i\psi_o$$

$$\partial_t \psi_e = \kappa \partial_x^2 \psi_o$$
$$\partial_t \psi_o = -\kappa \partial_x^2 \psi_e.$$
$$i\partial\psi/\partial t = \kappa \partial^2 \psi/\partial x^2.$$

REMARK. The discrete recursion at the beginning of this section, can actually be implemented to approximate solutions to the Schrödinger equation. This will be studied in a separate paper. The reader may wish to point out that the playing of dice in quantum mechanics has nothing to do with the deterministic evolution of the Schrödinger equation, and everything to do with the measurement postulate that interprets $\psi\psi^\dagger$ as a probability density.

Probability and generalizations of classical probability are necessary for doing science. One should keep in mind that the quantum mechanics is based on a model that takes the solution of the Schrödinger equation to be a superposition of all possible observations of a given observer. The solution has norm equal to one in an appropriate vector space. That norm is the integral of the absolute square of the wave function over all of space. The absolute square of the wavefunction is seen as the associated probability density. This extraordinary and concise recipe for the probability of observed events is at the core of this subject. It is natural to ask, in relation to our fable, what is the relationship of probability for the diffusion process and the probability in quantum theory. This will have to be the subject of another paper.

ACKNOWLEDGEMENT. It gives the author transfinite pleasure to thank G. Spencer-Brown, James Flagg, Alex Comfort, David Finkelstein, Pierre Noyes, Peter Rowlands, Sam Lomonaco, Bernd Schmeikal and Rukhsan Ul Haq for conversations related to the considerations in this paper. Nothing here is their fault, yet Nothing would have happened without them.

2 ITERANTS AND IDEMPOTENTS

An iterant is a sum of elements of the form

$$[a_1, a_2, ..., a_n]\sigma$$

where $[a_1, a_2, ..., a_n]$ is a vector of elements that are scalars (usually real or complex numbers) and σ is a permutation on n letters. Such elements are themselves sums of elements of the form

$$[0, 0, ...0, 1, 0, ..., 0]\sigma = e_i\sigma$$

where the 1 is in the i-th place. The elements e_i are the basic idempotents that generate the iterants with the help of the permutations.

Note that if $a = [a_1, a_2, ..., a_n]$, then we let a^σ denote the vector with its elements permuted by the action of σ. If a and b are vectors then ab denotes the vector where $(ab)_i = a_i b_i$, and $a + b$ denotes the vector where $(a + b)_i = a_i + b_i$. Then

$$(a\sigma)(b\tau) = (ab^\sigma)\sigma\tau,$$

$$(ka)\sigma = k(a\sigma)$$

for a scalar k, and

$$(a + b)\sigma = a\sigma + b\sigma$$

where vectors are multiplied as above and we take the usual product of the permutations. All of matrix algebra and more is naturally represented in the iterant framework, as we shall see in the next sections.

For example, if η is the order two permutation of two elements, then $[a, b]^\eta = [b, a]$. We can define

$$i = [1, -1]\eta$$

and then

$$i^2 = [1, -1]\eta[1, -1]\eta = [1, -1][1, -1]^\eta\eta^2 = [1, -1][-1, 1] = [-1, -1] = -1.$$

In this way the complex numbers arise naturally from iterants. One can interpret $[1,-1]$ as an oscillation between $+1$ and -1 and η as denoting a temporal shift operator. The $i = [1,-1]\eta$ is a time sensitive element and its self-interaction has square minus one. In this way iterants can be interpreted as a formalization of elementary discrete processes.

Note that we can write $a = [1,0], b = [0,1]$ and $A = a\eta, B = b\eta$ where η denotes the transpsition so that $[x,y]\eta = \eta[y,x]$ and $\eta^2 = 1$. Then we have

$$aa = a, bb = b, ab = 0, a + b = 1, AA = 0 = BB, AB = a, BA = b.$$

This is the mixed idempotent and permutation algebra for $n = 2$. Then we have

$$i = A - B$$

as we can see by

$$ii = (A - B)(A - B) = AA - AB - BA + BB = -a - b = -1.$$

This is the beginning of the relationships between idempotents, iterants and Clifford algebras.

Note that we construct an elementary Clifford algebra via

$$\alpha = [1,-1] = a - b$$

and

$$\beta = \eta.$$

Then we have

$$\alpha^2 = \beta^2 = 1$$

and

$$\alpha\beta + \beta\alpha = 0.$$

Note also that the non-commuting of α and β is directly related to the interaction of the idempotents and the permutations.

$$\alpha\beta = [1,-1]\eta = \eta[-1,1] = -\eta[1,-1] = -\beta\alpha.$$

Iterant algebra is generated by the elements

$$e_i\sigma$$

where e_i is a vector with a 1 in the i-th place and zeros elsewhere, and σ is an arbitrary element of the symmetric group S_n. We have that

$$e_i\sigma = \sigma e_{\sigma^{-1}(i)}$$

so that the multiplication of iterants is defined in terms of the action of the symmetric group. We have

$$e_i\sigma e_j\tau = e_i e_{\sigma(j)}\sigma\tau = \delta(i,\sigma(j))e_i\sigma\tau.$$

By themselves, the elements e_i are idempotent and we have

$$1 = e_1 + \cdots e_n.$$

The iterant algebra is generated by these combinations of idempotents and permutations.

3 ITERANTS, DISCRETE PROCESSES AND MATRIX ALGEBRA

The primitive idea behind an iterant is a periodic time series or "waveform"

$$\cdots ababababababab \cdots .$$

The elements of the waveform can be any mathematically or empirically well-defined objects. We can regard the ordered pairs $[a,b]$ and $[b,a]$ as abbreviations for the waveform or as two points of view about the waveform (a first or b first). Call $[a,b]$ an *iterant*. One has the collection of transformations of the form $T[a,b] = [ka, k^{-1}b]$ leaving the product ab invariant. This tiny model contains the seeds of special relativity, and the iterants contain the seeds of general matrix algebra! For related discussion see [2, 3, 4, 5, 6, 7, 10, 13, 11, 14, 1].

Define products and sums of iterants as follows

$$[a, b][c, d] = [ac, bd]$$

and

$$[a, b] + [c, d] = [a + c, b + d].$$

The operation of juxtaposition of waveforms is multiplication while $+$ denotes ordinary addition of ordered pairs. These operations are natural with respect to the structural juxtaposition of iterants:

$$...ababababab...$$

$$...cdcdcdcdcd...$$

Structures combine at the points where they correspond. Waveforms combine at the times where they correspond. Iterants combine in juxtaposition.

If $\bullet$ denotes any form of binary composition for the ingredients (a,b,...) of iterants, then we can extend $\bullet$ to the iterants themselves by the definition $[a, b] \bullet [c, d] = [a \bullet c, b \bullet d]$.

The appearance of a square root of minus one unfolds naturally from iterant considerations. Define the "shift" operator η on iterants by the equation

$$\eta[a, b] = [b, a]\eta$$

with $\eta^2 = 1$. Sometimes it is convenient to think of η as a delay operator, since it shifts the waveform ...ababab... by one internal time step. Now define

$$i = [-1, 1]\eta$$

We see at once that

$$ii = [-1, 1]\eta[-1, 1]\eta = [-1, 1][1, -1]\eta^2 = [-1, 1][1, -1] = [-1, -1] = -1.$$

Thus

$$ii = -1.$$

Here we have described i in a *new* way as the superposition of the waveform $\epsilon = [-1, 1]$ and the temporal shift operator η. By writing

$\imath = \epsilon\eta$ we recognize an active version of the waveform that shifts temporally when it is observed. This theme of including the result of time in observations of a discrete system occurs at the foundation of our construction.

In the next section we show how all of matrix algebra can be formulated in terms of iterants.

4 MATRIX ALGEBRA VIA ITERANTS

Matrix algebra has some strange wisdom built into its very bones. Consider a two dimensional periodic pattern or "waveform."

.....

...abababababababab...

...cdcdcdcdcdcdcdcd...

...abababababababab...

...cdcdcdcdcdcdcdcd...

...abababababababab...

.....

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}, \begin{pmatrix} b & a \\ d & c \end{pmatrix}, \begin{pmatrix} c & d \\ a & b \end{pmatrix}, \begin{pmatrix} d & c \\ b & a \end{pmatrix}$$

Above are some of the matrices apparent in this array. Compare the matrix with the "two dimensional waveform" shown above. A given matrix freezes out a way to view the infinite waveform. In order to keep track of this patterning, lets write

$$[a, b] + [c, d]\eta = \begin{pmatrix} a & c \\ d & b \end{pmatrix}.$$

where

$$[x, y] = \begin{pmatrix} x & 0 \\ 0 & y \end{pmatrix}.$$

and

$$\eta = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Recall the definition of matrix multiplication.

$$\begin{pmatrix} a & c \\ d & b \end{pmatrix} \begin{pmatrix} e & g \\ h & f \end{pmatrix} = \begin{pmatrix} ae + ch & ag + cf \\ de + bh & dg + bf \end{pmatrix}.$$

Compare this with the iterant multiplication.

$$([a, b] + [c, d]\eta)([e, f] + [g, h]\eta) =$$

$$[a, b][e, f] + [c, d]\eta[g, h]\eta + [a, b][g, h]\eta + [c, d]\eta[e, f] =$$

$$[ae, bf] + [c, d][h, g] + ([ag, bh] + [c, d][f, e])\eta =$$

$$[ae, bf] + [ch, dg] + ([ag, bh] + [cf, de])\eta =$$

$$[ae + ch, dg + bf] + [ag + cf, de + bh]\eta.$$

Thus matrix multiplication is identical with iterant multiplication. The concept of the iterant can be used to motivate matrix multiplication.

The four matrices that can be framed in the two-dimensional wave form are all obtained from the two iterants $[a, d]$ and $[b, c]$ via the shift operation $\eta[x, y] = [y, x]\eta$ which we shall denote by an overbar as shown below

$$\overline{[x, y]} = [y, x].$$

Letting $A = [a, d]$ and $B = [b, c]$, we see that the four matrices seen in the grid are

$$A + B\eta, B + A\eta, \overline{B} + \overline{A}\eta, \overline{A} + \overline{B}\eta.$$

The operator η has the effect of rotating an iterant by ninety degrees in the formal plane. Ordinary matrix multiplication can be written in a concise form using the following rules:

$$\eta\eta = 1$$
$$\eta Q = \overline{Q}\eta$$

where Q is any two element iterant. Note the correspondence

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a & 0 \\ 0 & d \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} b & 0 \\ 0 & c \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = [a, d]1 + [b, c]\eta.$$

This means that $[a, d]$ corresponds to a diagonal matrix.

$$[a, d] = \begin{pmatrix} a & 0 \\ 0 & d \end{pmatrix},$$

η corresponds to the anti-diagonal permutation matrix.

$$\eta = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

and $[b, c]\eta$ corresponds to the product of a diagonal matrix and the permutation matrix.

$$[b, c]\eta = \begin{pmatrix} b & 0 \\ 0 & c \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & b \\ c & 0 \end{pmatrix}.$$

Note also that

$$\eta[c, b] = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} c & 0 \\ 0 & b \end{pmatrix} = \begin{pmatrix} 0 & b \\ c & 0 \end{pmatrix}.$$

This is the matrix interpretation of the equation

$$[b, c]\eta = \eta[c, b].$$

The fact that the iterant expression $[a, d]1 + [b, c]\eta$ captures the whole of 2×2 matrix algebra corresponds to the fact that a two by two matrix is combinatorially the union of the identity pattern (the

diagonal) and the interchange pattern (the antidiagonal) that correspond to the operators 1 and η.

$$\begin{pmatrix} * & @ \\ @ & * \end{pmatrix}$$

In the formal diagram for a matrix shown above, we indicate the diagonal by $*$ and the anti-diagonal by @.

In the case of complex numbers we represent

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix} = [a, a] + [-b, b]\eta = a1 + b[-1, 1]\eta = a + bi.$$

In this way, we see that all of 2×2 matrix algebra is a hypercomplex number system based on the symmetric group S_2. In the next section we generalize this point of view to arbitrary finite groups.

We have reconstructed the square root of minus one in the form of the matrix

$$i = \epsilon\eta = [-1, 1]\eta = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

In this way, we arrive at this well-known representation of the complex numbers in terms of matrices. Note that if we identify the ordered pair (a, b) with $a + ib$, then this means taking the identification

$$(a, b) = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}.$$

Thus the geometric interpretation of multiplication by i as a ninety degree rotation in the Cartesian plane,

$$i(a, b) = (-b, a),$$

takes the place of the matrix equation

$$i(a, b) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a & -b \\ b & a \end{pmatrix} = \begin{pmatrix} -b & -a \\ a & -b \end{pmatrix} = b + ia = (-b, a).$$

In iterant terms we have

$$i[a, b] = \epsilon\eta[a, b] = [-1, 1][b, a]\eta = [-b, a]\eta,$$

and this corresponds to the matrix equation

$$i[a, b] = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} = \begin{pmatrix} 0 & -b \\ a & 0 \end{pmatrix} = [-b, a]\eta.$$

All of this points out how the complex numbers, as we have previously examined them, live naturally in the context of the non-commutative algebras of iterants and matrices. The factorization of i into a product $\epsilon\eta$ of non-commuting iterant operators is closer both to the temporal nature of i and to its algebraic roots.

More generally, we see that

$$(A + B\eta)(C + D\eta) = (AC + B\overline{D}) + (AD + B\overline{C})\eta$$

writing the 2×2 matrix algebra as a system of hypercomplex numbers. Note that

$$(A + B\eta)(\overline{A} - B\eta) = A\overline{A} - B\overline{B}$$

The formula on the right equals the determinant of the matrix. Thus we define the *conjugate* of $Z = A + B\eta$ by the formula

$$\overline{Z} = \overline{A + B\eta} = \overline{A} - B\eta,$$

and we have the formula

$$D(Z) = Z\overline{Z}$$

for the determinant $D(Z)$ where

$$Z = A + B\eta = \begin{pmatrix} a & c \\ d & b \end{pmatrix}$$

where $A = [a, b]$ and $B = [c, d]$. Note that

$$A\overline{A} = [ab, ba] = ab1 = ab,$$

so that

$$D(Z) = ab - cd.$$

Note also that we assume that a, b, c, d are in a commutative base ring.

Note also that for Z as above,

$$\overline{Z} = \overline{A} - B\eta = \begin{pmatrix} b & -c \\ -d & a \end{pmatrix}.$$

This is the classical adjoint of the matrix Z.

We leave it to the reader to check that for matrix iterants Z and W,

$$Z\overline{Z} = \overline{Z}Z$$

and that

$$\overline{ZW} = \overline{W}\overline{Z}$$

and

$$\overline{Z + W} = \overline{Z} + \overline{W}.$$

Note also that

$$\overline{\eta} = -\eta,$$

whence

$$\overline{B\eta} = -B\eta = -\eta\overline{B} = \overline{\eta}\overline{B}.$$

We can prove that

$$D(ZW) = D(Z)D(W)$$

as follows

$$D(ZW) = ZW\overline{ZW} = ZW\overline{W}\,\overline{Z} = Z\overline{Z}W\overline{W} = D(Z)D(W).$$

Here the fact that $W\overline{W}$ is in the base ring which is commutative allows us to remove it from in between the appearance of Z and $\overline{Z}$. Thus we see that iterants as 2×2 matrices form a direct non-commutative generalization of the complex numbers.

It is worth pointing out the first precursor to the quaternions (the so-called *split quaternions*): This precursor is the system

$$\{\pm 1, \pm\epsilon, \pm\eta, \pm i\}.$$

Here $\epsilon\epsilon = 1 = \eta\eta$ while $i = \epsilon\eta$ so that $ii = -1$. The basic operations in this algebra are those of epsilon and eta. Eta is the delay shift operator

that reverses the components of the iterant. Epsilon negates one of the components, and leaves the order unchanged. The quaternions arise directly from these two operations once we construct an extra square root of minus one that commutes with them. Call this extra root of minus one $\sqrt{-1}$. Then the quaternions are generated by

$$I = \sqrt{-1}\epsilon, J = \epsilon\eta, K = \sqrt{-1}\eta$$

with

$$I^2 = J^2 = K^2 = IJK = -1.$$

The "right" way to generate the quaternions is to start at the bottom iterant level with boolean values of 0 and 1 and the operation EXOR (exclusive or). Build iterants on this, and matrix algebra from these iterants. This gives the square root of negation. Now take pairs of values from this new algebra and build 2×2 matrices again. The coefficients include square roots of negation that commute with constructions at the next level and so quaternions appear in the third level of this hierarchy. We will return to the quaternions after discussing other examples that involve matrices of all sizes.

5 ITERANTS OF ARBITRARILY HIGH PERIOD

As a next example, consider a waveform of period three.

$$\cdots abcabcabcabcabcabc \cdots$$

Here we see three natural iterant views (depending upon whether one starts at a, b or c).

$$[a, b, c], \quad [b, c, a], \quad [c, a, b].$$

The appropriate shift operator is given by the formula

$$[x, y, z]S = S[z, x, y].$$

Thus, with $T = S^2$,

$$[x, y, z]T = T[y, z, x]$$

and $S^3 = 1$. With this we obtain a closed algebra of iterants whose general element is of the form

$$[a, b, c] + [d, e, f]S + [g, h, k]S^2$$

where $a, b, c, d, e, f, g, h, k$ are real or complex numbers. Call this algebra $\mathbb{V}ect_3(\mathbb{R})$ when the scalars are in a commutative ring with unit $\mathbb{F}$. Let $M_3(\mathbb{F})$ denote the 3×3 matrix algebra over $\mathbb{F}$. We have the

LEMMA. The iterant algebra $\mathbb{V}ect_3(\mathbb{F})$ is isomorphic to the full 3×3 matrix algebra $M_3((\mathbb{F})$.

PROOF. Map 1 to the matrix

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Map S to the matrix

$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix},$$

and map S^2 to the matrix

$$\begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix},$$

Map $[x, y, z]$ to the diagonal matrix

$$\begin{pmatrix} x & 0 & 0 \\ 0 & y & 0 \\ 0 & 0 & z \end{pmatrix}.$$

Then it follows that

$$[a, b, c] + [d, e, f]S + [g, h, k]S^2$$

maps to the matrix

$$\begin{pmatrix} a & d & g \\ h & b & e \\ f & k & c \end{pmatrix},$$

preserving the algebra structure. Since any 3×3 matrix can be written uniquely in this form, it follows that $\mathbb{V}ect_3(\mathbb{F})$ is isomorphic to the full 3×3 matrix algebra $M_3(\mathbb{F})$. //

We can summarize the pattern behind this expression of 3×3 matrices by the following symbolic matrix.

$$\begin{pmatrix} 1 & S & T \\ T & 1 & S \\ S & T & 1 \end{pmatrix}$$

Here the letter T occupies the positions in the matrix that correspond to the permutation matrix that represents it, and the letter $T = S^2$ occupies the positions corresponding to its permutation matrix. The 1's occupy the diagonal for the corresponding identity matrix. The iterant representation corresponds to writing the 3×3 matrix as a disjoint sum of these permutation matrices such that the matrices themselves are closed under multiplication. In this case the matrices form a permutation representation of the cyclic group of order 3, $C_3 = \{1, S, S^2\}$.

Remark. Note that a permutation matrix is a matrix of zeroes and ones such that some permutation of the rows of the matrix transforms it to the identity matrix. Given an $n \times n$ permutation matrix P, we associate to it a permutation

$$\sigma(P) : \{1, 2, \cdots, n\} \longrightarrow \{1, 2, \cdots, n\}$$

via the following formula

$$i\sigma(P) = j$$

where j denotes the column in P where the i-th row has a 1. Note that an element of the domain of a permutation is indicated to the

left of the symbol for the permutation. It is then easy to check that for permutation matrices P and Q,

$$\sigma(P)\sigma(Q) = \sigma(PQ)$$

given that we compose the permutations from left to right according to this convention.

It should be clear to the reader that this construction generalizes directly for iterants of any period and hence for a set of operators forming a cyclic group of any order. In fact we shall generalize further to any finite group G. We now define $\mathbb{V}ect_{n(}G, \mathbb{F})$ for any finite group G.

DEFINITION. Let G be a finite group, written multiplicatively. Let $\mathbb{F}$ denote a given commutative ring with unit. Assume that G acts as a group of permutations on the set $\{1, 2, 3, \cdots, n\}$ so that given an element $g \in G$ we have (by abuse of notation)

$$g : \{1, 2, 3, \cdots, n\} \longrightarrow \{1, 2, 3, \cdots, n\}.$$

We shall write

$$ig$$

for the image of $i \in \{1, 2, 3, \cdots, n\}$ under the permutation represented by g. Note that this denotes functionality from the left and so we ask that $(ig)h = i(gh)$ for all elements $g, h \in G$ and $i1 = i$ for all i, in order to have a representation of G as permutations. We shall call an n-tuple of elements of $\mathbb{F}$ a *vector* and denote it by $a = (a_1, a_2, \cdots, a_n)$. We then define an action of G on vectors over $\mathbb{F}$ by the formula

$$a^g = (a_{1g}, a_{2g}, \cdots, a_{ng}),$$

and note that $(a^g)^h = a^{gh}$ for all $g, h \in G$. We now define an algebra $\mathbb{V}ect_n(G, \mathbb{F})$, the *iterant algebra for* G, to be the set of finite sums of formal products of vectors and group elements in the form ag with multiplication rule

$$(ag)(bh) = ab^g(gh),$$

and the understanding that $(a+b)g = ag + bg$ and for all vectors a, b and group elements g. It is understood that vectors are added coordinate-wise and multiplied coordinate-wise. Thus $(a+b)_i = a_i + b_i$ and $(ab)_i = a_i b_i$.

THEOREM. Let G be a finite group of order n. Let $\rho : G \longrightarrow S_n$ denote the right regular representation of G as permutations of n things where we list the elements of G as $G = \{g_1, \cdots, g_n\}$ and let G act on its own underlying set via the definition $g_i\rho(g) = g_i g$. Here we describe $\rho(g)$ acting on the set of elements g_k of G. If we wish to regard $\rho(g)$ as a mapping of the set $\{1, 2, \cdots n\}$ then we replace g_k by k and $i\rho(g) = k$ where $g_i g = g_k$.

Then $\mathbb{V}ect_n(G, \mathbb{F})$ is isomorphic to the matrix algebra $M_n((\mathbb{F})$. In particular, we have that $\mathbb{V}ect_{n!}(S_n, \mathbb{F})$ is isomorphic with the matrices of size $n! \times n!$, $M_{n!}((\mathbb{F})$.

PROOF. Consider the $n \times n$ matrix consisting in the multiplication table for G with the columns and rows listed in the order $[g_1, \cdots, g_n]$. Permute the rows of this table so that the diagonal consists in all 1's. Let the resulting table be called the G-*Table*. The G-*Table* is labeled by elements of the group. For a vector a, let $D(a)$ denote the $n \times n$ diagonal matrix whose entries in order down the diagonal are the entries of a in the order specified by a. For each group element g, let P_g denote the permutation matrix with 1 in every spot on the G-*Table* that is labeled by g and 0 in all other spots. It is now a direct verification that the mapping

$$F(\Sigma_{i=1}^{n} a_i g_i) = \Sigma_{i=1}^{n} D(a_i) P_{g_i}$$

defines an isomorphism from $\mathbb{V}ect_n(G, \mathbb{F})$ to the matrix algebra $M_n((\mathbb{F})$. The main point to check is that $\sigma(P_g) = \rho(g)$. We now prove this fact.

In the G-*Table* the rows correspond to

$$\{g_1^{-1}, g_2^{-1}, \cdots g_n^{-1}\}$$

and the columns correspond to

$$\{g_1, g_2, \cdots g_n\}$$

so that the i-i entry of the table is $g_i^{-1} g_i = 1$. With this we have that in the table, a group element g occurs in the i-th row at column j where

$$g_i^{-1} g_j = g.$$

This is equivalent to the equation

$$g_i g = g_j$$

which, in turn is equivalent to the statement

$$i\rho(g) = j.$$

This is exactly our functional interpretation of the action of the permutation corresponding to the matrix P_g. Thus

$$\rho(g) = \sigma(P_g).$$

The remaining detalls of the proof are straightforward and left to the reader. //

Examples.

1. We have already implicitly given examples of this process of translation. Consider the cyclic group of order three.

   $$C_3 = \{1, S, S^2\}$$

   with $S^3 = 1$. The multiplication table is

   $$\begin{pmatrix} 1 & S & S^2 \\ S & S^2 & 1 \\ S^2 & 1 & S \end{pmatrix}.$$

   Interchanging the second and third rows, we obtain

   $$\begin{pmatrix} 1 & S & S^2 \\ S^2 & 1 & S \\ S & S^2 & 1 \end{pmatrix},$$

   and this is the G-*Table* that we used for $\mathbb{V}ect_3(C_3, \mathbb{F})$ prior to proving the Main Theorem.

The same pattern works for arbitrary cyclic groups. for example, consider the cyclic group of order 6. $C_6 = \{1, S, S^2, S^3, S^4, S^5\}$ with $S^6 = 1$. The multiplication table is

$$\begin{pmatrix} 1 & S & S^2 & S^3 & S^4 & S^5 \\ S & S^2 & S^3 & S^4 & S^5 & 1 \\ S^2 & S^3 & S^4 & S^5 & 1 & S \\ S^3 & S^4 & S^5 & 1 & S & S^2 \\ S^4 & S^5 & 1 & S & S^2 & S^3 \\ S^5 & 1 & S & S^2 & S^3 & S^4 \end{pmatrix}.$$

Rearranging to form the G-*Table*, we have

$$\begin{pmatrix} 1 & S & S^2 & S^3 & S^4 & S^5 \\ S^5 & 1 & S & S^2 & S^3 & S^4 \\ S^4 & S^5 & 1 & S & S^2 & S^3 \\ S^3 & S^4 & S^5 & 1 & S & S^2 \\ S^2 & S^3 & S^4 & S^5 & 1 & S \\ S & S^2 & S^3 & S^4 & S^5 & 1 \end{pmatrix}.$$

The permutation matrices corresponding to the positions of S^k in the G-*Table* give the matrix representation that gives the isomorphism of $\mathbb{V}ect_6(C_6, \mathbb{F})$ with the full algebra of six by six matrices.

2. Now consider the symmetric group on six letters,

$$S_6 = \{1, R, R^2, F, RF, R^2F\}$$

where $R^3 = 1, F^2 = 1, FR = RF^2$. Then the multiplication table is

$$\begin{pmatrix} 1 & R & R^2 & F & RF & R^2F \\ R & R^2 & 1 & RF & R^2F & F \\ R^2 & 1 & R & R^2F & F & RF \\ F & R^2F & RF & 1 & R^2 & R \\ RF & F & R^2F & R & 1 & R^2 \\ R^2F & RF & F & R^2 & R & 1 \end{pmatrix}.$$

The corresponding G-*Table* is

$$\begin{pmatrix} 1 & R & R^2 & F & RF & R^2F \\ R^2 & 1 & R & R^2F & F & RF \\ R & R^2 & 1 & RF & R^2F & F \\ F & R^2F & RF & 1 & R^2 & R \\ RF & F & R^2F & R & 1 & R^2 \\ R^2F & RF & F & R^2 & R & 1 \end{pmatrix}.$$

Here is a rewritten version of the G-*Table* with

$$R = \Delta, R^2 = \Theta, F = \Psi, RF = \Omega, R^2F = \Sigma.$$

$$\begin{pmatrix} 1 & \Delta & \Theta & \Psi & \Omega & \Sigma \\ \Theta & 1 & \Delta & \Sigma & \Psi & \Omega \\ \Delta & \Theta & 1 & \Omega & \Sigma & \Psi \\ \Psi & \Sigma & \Omega & 1 & \Theta & \Delta \\ \Omega & \Psi & \Sigma & \Delta & 1 & \Theta \\ \Sigma & \Omega & \Psi & \Theta & \Delta & 1 \end{pmatrix}.$$

This G-*Table* is the keystone for the isomorphism of $\mathbb{V}ect_6(S_3, \mathbb{F})$ with the full algebra of six by six matrices. At this point it may occur to the reader to wonder about $\mathbb{V}ect_3(S_3, \mathbb{F})$ since S_3 does act on vectors of length three. We will discuss $\mathbb{V}ect_n(S_n, \mathbb{F})$ in the next section. We see from this example how it will come about that $\mathbb{V}ect_{n!}(S_n, \mathbb{F})$ is isomorphic with the full algebra of $n! \times n!$ matrices. In particular, here are the permutation matrices that form the non-identity elements of this representation of the symmetric group on three letters.

$$R = \Delta = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

$$R^2 = \Theta = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix}$$

$$F = \Psi = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix}$$

$$FR = \Omega = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$FR^2 = \Sigma = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

3. In this example we consider the group $G = C_2 \times C_2$, often called the "Klein 4-Group." We take $G = \{1, A, B, C\}$ where $A^2 = B^2 =$

$C^2 = 1, AB = BA = C$. Thus G has the multiplication table, which is also its G-*Table* for $\mathbb{V}ect_4(G, \mathbb{F})$.

$$\begin{pmatrix} 1 & A & B & C \\ A & 1 & C & B \\ B & C & 1 & A \\ C & B & A & 1 \end{pmatrix}.$$

Thus we have the following permutation matrices that I shall call E, A, B, C :

$$E = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix},$$

$$B = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix},$$

$$C = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}.$$

The reader will have no difficulty verifying that $A^2 = B^2 = C^2 = 1, AB = BA = C$. Recall that $[x, y, z, w]$ is iterant notation for the diagonal matrix

$$[x, y, z, w] = \begin{pmatrix} x & 0 & 0 & 1 \\ 0 & y & 1 & 0 \\ 0 & 1 & z & 0 \\ 1 & 0 & 0 & w \end{pmatrix}.$$

Let

$$\alpha = [1, -1, -1, 1], \beta = [1, 1, -1, -1], \gamma = [1, -1, 1, -1].$$

And let

$$I = \alpha A, J = \beta B, K = \gamma C.$$

Then the reader will have no trouble verifying that

$$I^2 = J^2 = K^2 = IJK = -1, IJ = K, JI = -K.$$

Thus we have constructed the quaternions as iterants in relation to the Klein Four Group. in Figure 8.1 we illustrate these quaternion generators with string diagrams for the permutations. The reader can check that the permutations correspond to the permutation matrices constructed for the Klein Four Group. For example, the permutation for I is $(12)(34)$ in cycle notation, the permutation for J is $(13)(24)$ and the permutation for K is $(14)(23)$. In the Figure we attach signs to each string of the permutation. These "signed permutations" act exactly as the products of vectors and permutations that we use for the iterants. One can see that the quaternions arise naturally from the Klein Four Group by attaching signs to the generating permutations as we have done in this Figure.

4. One can use the quaternions as a linear basis for 4×4 matrices just as our theorem would use the permutation matrices $1, A, B, C$. If we restrict to real scalars a, b, c, d such that $a^2 + b^2 + c^2 + c^2 = 1$, then the set of matrices of the form

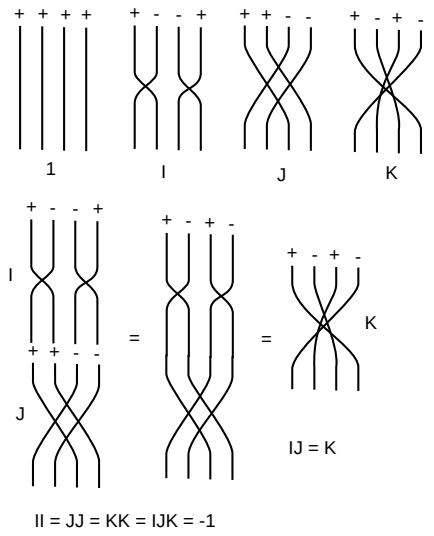


Figure 8.1: Quaternions From Klein Four Group

$a1 + bI + cJ + dK$ is isomorphic to the group $SU(2)$. To see this, note that $SU(2)$ is the set of matrices with complex entries z and w with determinant 1 so that $z\bar{z} + w\bar{w} = 1$.

$$M = \begin{pmatrix} z & w \\ -\bar{w} & \bar{z} \end{pmatrix}.$$

Letting $z = a + bi$ and $w = c + di$, we have

$$M = \begin{pmatrix} a+bi & c+di \\ -c+di & a-bi \end{pmatrix} = a\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + b\begin{pmatrix} i & 0 \\ o & -i \end{pmatrix} + c\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} + d\begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}.$$

If we regard $i = \sqrt{-1}$ as a commuting scalar, then we can write the generating matrices in terms of size two iterants and obtain

$$I = \sqrt{-1}\epsilon, J = \epsilon\eta, K = \sqrt{-1}\eta$$

as described in the previous section. IF we regard these matrices with complex entries as shorthand for 4×4 matrices with i interpreted as a 2×2 matrix as we have done above, then these 4×4 matrices representing the quaternions are exactly the ones we have constructed in relation to the Klein Four Group.

Since complex numbers commute with one another, we could consider iterants whose values are in the complex numbers. This is just like considering matrices whose entries are complex numbers. For this purpose we shall allow given a version of i that commutes with the iterant shift operator η. Let this commuting i be denoted by ι. Then we are assuming that

$$\iota^2 = -1$$

$$\eta\iota = \iota\eta$$

$$\eta^2 = +1.$$

We then consider iterant views of the form $[a + b\iota, c + d\iota]$ and $[a + b\iota, c + d\iota]\eta = \eta[c + d\iota, a + b\iota]$. In particular, we have $\epsilon = [1, -1]$, and $i = \epsilon\eta$ is quite distinct from ι. Note, as before, that $\epsilon\eta = -\eta\epsilon$ and that $\epsilon^2 = 1$. Now let

$$I = \iota\epsilon$$

$$J = \epsilon\eta$$

$$K = \iota\eta.$$

We have used the commuting version of the square root of minus one in these definitions, and indeed we find the quaternions once more.

$$I^2 = \iota\epsilon\iota\epsilon = \iota\iota\epsilon\epsilon = (-1)(+1) = -1,$$

$$J^2 = \epsilon\eta\epsilon\eta = \epsilon(-\epsilon)\eta\eta = -1,$$

$$K^2 = \iota\eta\iota\eta = \iota\iota\eta\eta = -1,$$

$$IJK = \iota\epsilon\eta\iota\eta = \iota 1\iota\eta\eta = \iota\iota = -1.$$

Thus

$$I^2 = J^2 = K^2 = IJK = -1.$$

This construction shows how the structure of the quaternions comes directly from the non-commutative structure of period two iterants. In other, words, quaternions can be represented by 2×2 matrices. This is the way it has been presented in standard language. The group $SU(2)$ of 2×2 unitary matrices of determinant one is isomorphic to the quaternions of length one.

5. Similarly,

$$H = [a, b] + [c + d\iota, c - d\iota]\eta = \begin{pmatrix} a & c + d\iota \\ c - d\iota & b \end{pmatrix}.$$

represents a Hermitian 2×2 matrix and hence an observable for quantum processes mediated by $SU(2)$. Hermitian matrices have real eigenvalues.

If in the above Hermitian matrix form we take $a = T + X, b = T - X, c = Y, d = Z$, then we obtain an iterant and/or matrix representation for a point in Minkowski spacetime.

$$H = [T + X, T - X] + [Y + Z\iota, Y - Z\iota]\eta = \begin{pmatrix} T + X & Y + Z\iota \\ Y - Z\iota & T - X \end{pmatrix}.$$

Note that we have the formula

$$Det(H) = T^2 - X^2 - Y^2 - Z^2.$$

It is not hard to see that the eigenvalues of H are $T \pm \sqrt{X^2 + Y^2 + Z^2}$. Thus, viewed as an observable, H can observe the time and the invariant spatial distance from the origin of the event (T, X, Y, Z). At least at this very elementary juncture, quantum mechanics and special relativity are reconciled.

6. Hamilton's Quaternions are generated by iterants, as discussed above, and we can express them purely algebraically by writing the corresponding permutations as shown below.

$$I = [+1, -1, -1, +1]s$$
$$J = [+1, +1, -1, -1]l$$
$$K = [+1, -1, +1, -1]t$$

where

$$s = (12)(34)$$
$$l = (13)(24)$$
$$t = (14)(23).$$

Here we represent the permutations as products of transpositions (ij). The transposition (ij) interchanges i and j, leaving all other elements of $\{1, 2, ..., n\}$ fixed.

One can verify that

$$I^2 = J^2 = K^2 = IJK = -1.$$

For example,

$$\begin{aligned} I^2 &= [+1, -1, -1, +1]s[+1, -1, -1, +1]s \\ &= [+1, -1, -1, +1][-1, +1, +1, -1]ss \\ &= [-1, -1, -1, -1] \\ &= -1. \end{aligned}$$

and

$$\begin{aligned} IJ &= [+1, -1, -1, +1]s[+1, +1, -1, -1]l \\ &= [+1, -1, -1, +1][+1, +1, -1, -1]sl \end{aligned}$$

$$= [+1,-1,+1,-1](12)(34)(13)(24)$$
$$= [+1,-1,+1,-1](14)(23)$$
$$= [+1,-1,+1,-1]t.$$

Nevertheless, we must note that making an iterant interpretation of an entity like $I = [+1,-1,-1,+1]s$ is a conceptual departure from our original period two iterant (or cyclic period n) notion. Now we are considering iterants such as $[+1,-1,-1,+1]$ where the permutation group acts to produce other orderings of a given sequence. The iterant itself is not necessarily an oscillation. It can represent an implicate form that can be seen in any of its possible orders. These orders are subject to permutations that produce the possible views of the iterant. Algebraic structures such as the quaternions appear in the explication of such implicate forms.

The reader will also note that we have moved into a different conceptual domain from an original emphasis in this paper on eigenform in relation to to recursion. That is, we take an *eigenform* to mean a fixed point for a transformation. Thus i is an eigenform for $R(x) = -1/x$. Indeed, each generating quaternion is an eigenform for the transformation $R(x) = -1/x$. The richness of the quaternions arises from the closed algebra that arises with its infinity of eigenforms that satisfy the equation $U^2 = -1$:

$$U = aI + bJ + cK$$

where $a^2 + b^2 + c^2 = 1$. This kind of significant extra structure in the eigenforms comes from paying attention to specific aspects of implicate and explicate structure, relationships with geometry and ideas and inputs from the perceptual, conceptual and physical worlds. Just as with our other examples of phenomena arising in the course of the recursion, we see the same phenomena here in the evolution of mathematical and theoretical physical structures in the course of the recursion that constitutes scientific conversation.

7. In all these examples, we have the opportunity to interpret the iterants as short hand for matrix algebra based on permutation matrices, or as indicators of discrete processes. The discrete processes become more complex in proportion to the complexity of the groups used in the construction. We began with processes of order two, then considered cyclic groups of arbitrary order, then the symmetric group S_3 in relation to 6×6 matrices, and the Klein Four Group in relation to the quaternions. In the case of the quaternions, we know that this structure is intimately related to rotations of three and four dimensional space and many other geometric themes. It is worth reflecting on the possible significance of the underlying discrete dynamics for this geometry, topology and related physics.

6 THE FRAMED BRAID GROUP

The reader should recall that the symmetric group S_n has presentation

$$S_n = (T_1, \cdots T_{n-1} | T_i^2 = 1, T_i T_{i+1} T_i = T_{i+1} T_i T_{i+1}, T_i T_j = T_j T_i; |i-j| > 1).$$

The Artin Braid Group B_n is a relative of the symmetric group that is obtained by removing the condition that each generator has square equal to the idenity.

$$B_n = (\sigma_1, \cdots \sigma_{n-1} | \sigma_i \sigma_{i+1} \sigma_i = \sigma_{i+1} \sigma_i \sigma_{i+1}, \sigma_i \sigma_j = \sigma_j \sigma_i; |i-j| > 1).$$

In Figure 8.2 we illustrate the the generators $\sigma_1, \sigma_2, \sigma_3$ of the 4-strand braid group and we show the topological nature of the relation $\sigma_1 \sigma_2 \sigma_1 = \sigma_2 \sigma_1 \sigma_2$ and the commuting relation $\sigma_1 \sigma_3 = \sigma_3 \sigma_1$. Topological braids are represented as collections of always descending strings, starting from a row of points and ending at another row of points. The strings are embedded in three dimensional space and can wind around one another. The elementary braid generators σ_i correspond to the i-th strand interchanging with the $i+1$-th strand. Two braids are multiplied by attaching the bottom endpoiints of one braid to the top endpoints of the other braid to form a new braid.

There is a fundamental homomorphism

$$\pi : B_n \longrightarrow S_n$$

defined on generators by

$$\pi(\sigma_i) = T_i$$

in the language of the presentations above. In term of the diagrams in Figure 8.2, a braid diagram is a permutation diagram if one forgets about its weaving structure of over and under strands at a crossing.

We now turn to a generalization of the braid group, the *framed braid group*. In this generalization, we associate elements of the form t^a to the top of each braid strand. For these purposes it is useful to take t as an algebraic variable and a as an integer. To interpret this framing geometrically replace each braid strand by a ribbon and interpret t^a as a $2\pi a$ twist in the ribbon. In Figure 8.3 we illustrate how to multiply two framed braids. In our formalism the braids A and B in this figure are given by the formulas

$$A = [t^a, t^b, t^c]\sigma_1\sigma_2\sigma_3,$$

$$B = [t^d, t^e, t^f]\sigma_2\sigma_3$$

in the framed braid group on three strands, denoted FB_3. As the Figure 8.3 illustrates, we have the basic formula

$$v\sigma = \sigma v^{\pi(\sigma)}$$

where v is a vector of the form $v = [t^a, t^b, t^c]$ (for $n = 3$) and $v^{\pi(\sigma)}$ denotes the action of the permutation associated with the braid σ on the vector v. In the figure the permutation is accomplished by sliding the algebra along the strings of the braid.

We can form an algebra $Alg[FB_n]$ by taking formal sums of framed braids of the form $\sum c_k v_k G_k$ where c_k is a scalar, v_k is a framing vector and G_k is an element of the Artin Braid group B_n. Since braids act on framing vectors by permutations, this algebra is a generalization of the iterant algebras we have defined so far. The algebra of framed

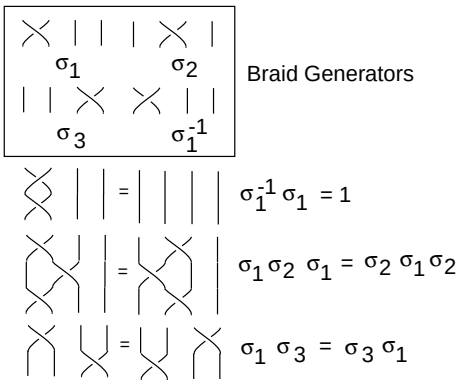


Figure 8.2: Braid Generators

braids uses an action of the braid group based on its representation to the symmetric group. Furthermore, the representation $\pi : B_n \longrightarrow S_n$ induces a map of algebras

$$\hat{\pi} : Alg[FB_n] \longrightarrow Alg[FS_n]$$

where we recognize $Alg[FS_n]$ as exactly an iterant algebra based in S_n.

In [31] Sundance Bilson-Thompson represents Fermions as framed braids. See Figure 8.4 for his diagrammatic representations. In this theory each fermion is associated with a framed braid. Thus from the figure we see that the positron and the electron are given by the framed braids

$$e^+ = [t, t, t]\sigma_1\sigma_2^{-1},$$

and

$$e^- = \sigma_2\sigma_1^{-1}[t^{-1}, t^{-1}, t^{-1}],$$

Here we use $[t^a, t^b, t^c]$ for the framing numbers (a, b, c). Products of framed braids correspond to particle interactions. Note that $e^+e^- = [1, 1, 1] = \gamma$ so that the electron and the positron are inverses in this

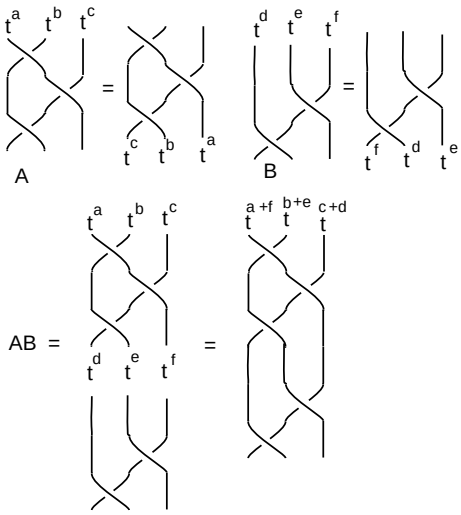


Figure 8.3: Framed Braids

algebra. In Figure 8.5 are illustrated the representations of bosons, including γ, a photon and the identity element in this algebra. Other relations in the algebra correspond to particle interactions. For example In Figure 8.6 is illustrated the muon decay

$$\mu \to \nu_\mu + W_- \to \nu_\mu + \bar{\nu_e} + e^-.$$

The reader can see the definitions of the different parts of this decay sequence from the three figures we have just mentioned. Note the strictly speaking the muon decay is a multiplicative identity in the braid algebra:

$$\mu = \nu_\mu W_- = \nu_\mu \bar{\nu_e} e^-.$$

Particle interactions in this model are mediated by factorizations in the noncommutative algebra of the framed braids.

By using the representation $\hat{\pi} : Alg[FB_3] \longrightarrow Alg[FS_3]$ we can image the structure of Bilson-Thompson's framed braids in the the iterant algebra corresponding to the symmetric group. However, we

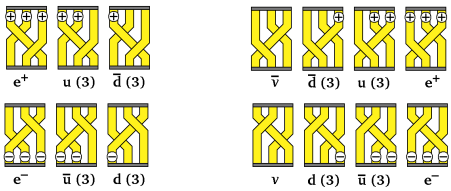


Figure 8.4: Sundance Bilson Thompson Framed Braid Fermions

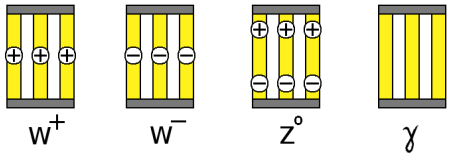


Figure 8.5: Bosons

propose to change this map so that we have a non-trivial representation of the Artin braid group. This can be accomplished by defining

$$\rho : Alg[FB_3] \longrightarrow Alg[FS_3]$$

where

$$\rho(\sigma_k) = [t, t]T_k$$

and

$$\rho(\sigma_k^{-1}) = [t^{-1}, t^{-1}]T_k$$

for $k = 1, 2$. The reader will find that we have now embedded the braid group in the iterant algebra $Alg[FS_3]$ and extended the embedding to the framed braid group algebra. Thus *the Sundance Bilson-Thompson representation of elementary particles as framed braids is embedded inside the iterant algebra for the symmetric group on three letters.* In Section 10 we carry this further and place the representation inside the Lie Algebra $su(3)$.

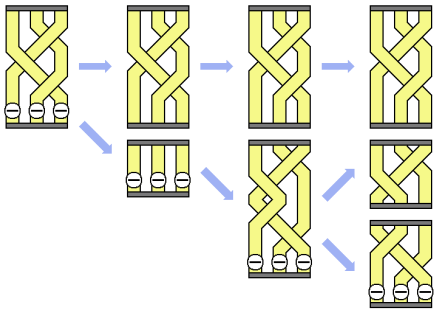


Figure 8.6: Representation of $\mu \to \nu_\mu + W_- \to \nu_\mu + \bar{\nu_e} + e^-$.

7 ITERANTS AND THE STANDARD MODEL

In this section we shall give an iterant interpretation for the Lie algebra of the special unitary group $SU(3)$. The Lie algebra in question is denoted as $su(3)$ and is often described by a matrix basis. The Lie algebra $su(3)$ is generated by the following eight Gell Man Matrices [29].

$$\lambda_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \lambda_2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \lambda_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

$$\lambda_4 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \lambda_5 = \begin{pmatrix} 0 & 0 & i \\ 0 & 0 & 0 \\ -i & 0 & 0 \end{pmatrix}, \lambda_6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix},$$

$$\lambda_7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \lambda_8 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}$$

The group $SU(3)$ consists in the matrices $U(\epsilon_1, \cdots, \epsilon_8) = e^{i \sum_a \epsilon_a \lambda_a}$ where $\epsilon_1, \cdots, \epsilon_8$ are real numbers and a ranges from 1 to 8. The Gell Man matrices satisfy the following relations.

$$tr(\lambda_a \lambda_b) = 2\delta_{ab},$$

$$[\lambda_a/2, \lambda_b/2] = if_{abc}\lambda_c/2.$$

Here we use the summation convention – summing over repeated indices, and tr denotes standard matrix trace, $[A, B] = AB - BA$ is the matrix commutator and δ_{ab} is the Kronecker delta, equal to 1 when $a = b$ and equal to 0 otherwise. The structure coefficients f_{abc} take the following non-zero values.

$$f_{123} = 1, f_{147} = 1/2, f_{156} = -1/2, f_{246} = 1/2, f_{257} = 1/2,$$

$$f_{345} = 1/2, f_{367} = -1/2, f_{458} = \sqrt{3/2}, f_{678} = \sqrt{3/2}$$

We now give an iterant representation for these matrices that is based on the pattern

$$\begin{pmatrix} 1 & A & B \\ B & 1 & A \\ A & B & 1 \end{pmatrix}$$

as described in the previous section. That is, we use the cyclic group of order three to represent all 3×3 matrices at iterants based on the permutation matrices

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}, B = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

Recalling that $[a, b, c]$ as an iterant, denotes a diagonal matrix

$$[a, b, c] = \begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix},$$

the reader will have no difficulty verifying the following formulas for the Gell Mann Matrices in the iterant format:

$$\lambda_1 = [1,0,0]A + [0,1,0]B$$

$$\lambda_2 = [-i,0,0]A + [0,i,0]B$$

$$\lambda_3 = [1,-1,0]$$

$$\lambda_4 = [1,0,0]B + [0,0,1]A$$

$$\lambda_5 = [i,0,0]B + [0,0,-i]A$$

$$\lambda_6 = [0,1,0]A + [0,0,1]B$$

$$\lambda_7 = [0,-i,0]A + [0,0,i]B$$

$$\lambda_8 = \frac{1}{\sqrt{3}}[1,1,-2].$$

Letting $F_a = \lambda_a/2$, we can now rewrite the Lie algebra into simple iterants of the form $[a,b,c]G$ where G is a cyclic group element. Compare with [30]. Let

$$T_\pm = F_1 \pm iF_2,$$

$$U_\pm = F_6 \pm iF_7,$$

$$V_\pm = F_4 \pm iF_5,$$

$$T_3 = F_3,$$

$$Y = \frac{2}{\sqrt{3}}F_8.$$

Then we have the specific iterant formulas

$$T_+ = [1,0,0]A,$$

$$T_- = [0,1,0]B,$$

$$U_+ = [0,1,0]A,$$

$$U_- = [0,0,1]B,$$

$$V_+ = [0,0,1]A,$$

$$V_- = [1,0,0]B,$$

$$T_3 = [1/2, -1/2, 0],$$
$$Y = \frac{1}{\sqrt{3}}[1, 1, -2].$$

We have that $A[x, y, z] = [y, z, x]A$ and $B = A^2 = A^{-1}$ so that $B[x, y, z] = [z, y, x]B$. Thus we have reduced the basic $su(3)$ Lie algebra to a very elementary patterning of order three cyclic operations. In a subsequent paper, we will use this point to view to examine the irreducible representations of this algebra and to illuminate the Standard Model's Eightfold Way.

8 ITERANTS, BRAIDING AND THE SUNDANCE-BILSON THOMPSON MODEL FOR FERMIONS

In the last section we based our iterant representations on the following patterns and matrices. The pattern,

$$\begin{pmatrix} 1 & A & B \\ B & 1 & A \\ A & B & 1 \end{pmatrix},$$

using the cyclic group of order three to represent all 3×3 matrices at iterants based on the permutation matrices

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}, B = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

Recalling that $[a, b, c]$ as an iterant, denotes a diagonal matrix

$$[a, b, c] = \begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix}.$$

In fact there are six 3×3 permutation matrices: $\{I, A, B, P, Q, R\}$ where

$$P = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, Q = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, R = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

We then have $A = QP, B = PQ, R = PQP = QPQ$. The two transpositions P and Q generate the entire group of permutations S_3. It is usual to think of the order-three transformations A and B as expressed in terms of these transpositions, but we can also use the iterant structure of the 3×3 matrices to express P, Q and R in terms of A and B. The result is as follows:

$$P = [0, 0, 1] + [1, 0, 0]A + [0, 1, 0]B,$$

$$Q = [1, 0, 0] + [0, 1, 0]A + [0, 0, 1]B,$$

$$R = [0, 1, 0] + [0, 0, 1]A + [1, 0, 0]B.$$

Recall from the previous section that we have the iterant generators for the $su(3)$ Lie algebra:

$$T_+ = [1, 0, 0]A,$$

$$T_- = [0, 1, 0]B,$$

$$U_+ = [0, 1, 0]A,$$

$$U_- = [0, 0, 1]B,$$

$$V_+ = [0, 0, 1]A,$$

$$V_- = [1, 0, 0]B.$$

Thus we can express these transpositions P and Q in the iterant form of the Lie algebra as

$$P = [0, 0, 1] + T_+ + T_-,$$

$$Q = [1, 0, 0] + U_+ + U_-,$$

$$R = [0, 1, 0] + V_+ + V_-.$$

The basic permutations receive elegant expressions in the iterant Lie algebra.

Now that we have basic permutations in the Lie algebra we can take the map from section 7.1

$$\rho : Alg[FB_3] \longrightarrow Alg[FS_3]$$

with

$$\rho(\sigma_k) = [t, t]T_k$$

and

$$\rho(\sigma_k^{-1}) = [t^{-1}, t^{-1}]T_k$$

for $k = 1, 2$ and send T_1 to P and T_2 to Q. Then we have

$$\rho(\sigma_1) = [t, t]P$$

and

$$\rho(\sigma_1^{-1}) = [t^{-1}, t^{-1}]P$$

and

$$\rho(\sigma_2) = [t, t]Q$$

and

$$\rho(\sigma_1^{-1}) = [t^{-1}, t^{-1}]Q.$$

By choosing $t \neq 1$ on the unit circle in the complex plane, we obtain representations of the Sundance Bilson-Thompson constructions of Fermions via framed braids *inside* the $su(3)$ Lie algebra. This brings the Bilson-Thompson formalism in direct contact with the Standard Model via our iterant representations. We shall return to these relationships in a sequel to the present paper.

9 CLIFFORD ALGEBRA, MAJORANA FERMIONS AND BRAIDING

Recall fermion algebra. One has fermion annihilation operators ψ and their conjugate creation operators $\psi^\dagger$. One has $\psi^2 = 0 = (\psi^\dagger)^2$. There is a fundamental commutation relation

$$\psi\psi^\dagger + \psi^\dagger\psi = 1.$$

If you have more than one of them say ψ and ϕ, then they anti-commute:

$$\psi\phi = -\phi\psi.$$

The Majorana fermions c that satisfy $c^\dagger = c$ so that they are their own anti-particles. There is a lot of interest in these as quasi-particles and they are related to braiding and to topological quantum computing. A group of researchers [28] claims, at this writing, to have found quasiparticle Majorana fermions in edge effects in nano-wires. (A line of fermions could have a Majorana fermion happen non-locally from one end of the line to the other.) The Fibonacci model that we discuss is also based on Majorana particles, possibly related to collective electronic excitations. If P is a Majorana fermion particle, then P can interact with itself to either produce itself or to annihilate itself. This is the simple "fusion algebra" for this particle. One can write $P^2 = P + 1$ to denote the two possible self-interactions the particle P. The patterns of interaction and braiding of such a particle P give rise to the Fibonacci model.

Majoranas are related to standard fermions as follows: The algebra for Majoranas is $c = c^\dagger$ and $cc' = -c'c$ if c and c' are distinct Majorana fermions with $c^2 = 1$ and $c'^2 = 1$. One can make a standard fermion from two Majoranas via

$$\psi = (c + ic')/2,$$

$$\psi^\dagger = (c - ic')/2.$$

Similarly one can mathematically make two Majoranas from any single fermion. Now if you take a set of Majoranas

$$\{c_1, c_2, c_3, \cdots, c_n\}$$

then there are natural braiding operators that act on the vector space with these c_k as the basis. The operators are mediated by algebra elements

$$\tau_k = (1 + c_{k+1}c_k)/\sqrt{2},$$
$$\tau_k^{-1} = (1 - c_{k+1}c_k)/\sqrt{2}.$$

Then the braiding operators are

$$T_k : Span\{c_1, c_2, \cdots, , c_n\} \longrightarrow Span\{c_1, c_2, \cdots, , c_n\}$$

via

$$T_k(x) = \tau_k x \tau_k^{-1}.$$

The braiding is simply:

$$T_k(c_k) = c_{k+1},$$

$$T_k(c_{k+1}) = -c_k,$$

and T_k is the identity otherwise. This gives a very nice unitary representations of the Artin braid group and it deserves better understanding. See Figure 8.7 for an illustration of this braiding of Fermions in relation to the topology of a belt that connects them. The relationship with the belt is tied up with the fact that in quantum mechanics we must represent rotations of three dimensional space as unitary transformations. See [12] for more about this topological view of the physics of Fermions. In the Figure, we see that the belt does not know which of the two Fermions to anoint with the phase change, but the clever algebra above makes this decision. There is more to be done in this domain.

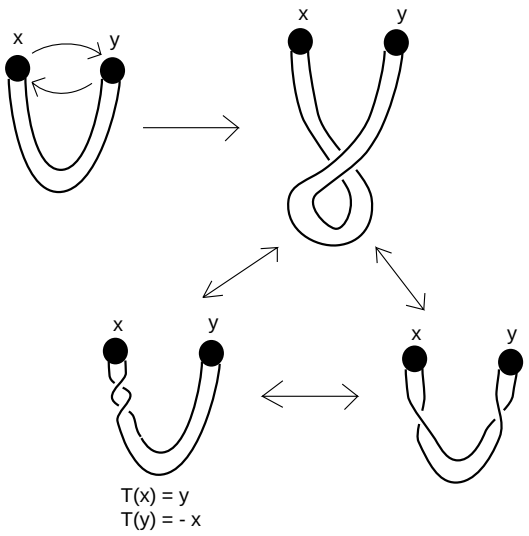


Figure 8.7: Braiding Action on a Pair of Fermions

It is worth noting that a triple of Majorana fermions say a, b, c gives rise to a representation of the quaternion group. This is a generalization of the well-known association of Pauli matrices and quaternions. We have $a^2 = b^2 = c^2 = 1$ and they anticommute. Let $I = ba, J = cb, K = ac$. Then

$$I^2 = J^2 = K^2 = IJK = -1,$$

giving the quaternions. The operators

$$A = (1/\sqrt{2})(1 + I)$$

$$B = (1/\sqrt{2})(1 + J)$$

$$C = (1/\sqrt{2})(1 + K)$$

braid one another:

$$ABA = BAB, BCB = CBC, ACA = CAC.$$

This is a special case of the braid group representation described above for an arbitrary list of Majorana fermions. These braiding operators are entangling and so can be used for universal quantum computation, but they give only partial topological quantum computation due to the interaction with single qubit operators not generated by them.

Recall that in discussing the beginning of iterants, we introduce a *temporal shift operator* η such that

$$[a, b]\eta = \eta[b, a]$$

and

$$\eta\eta = 1$$

for any iterant $[a, b]$, so that concatenated observations can include a time step of one-half period of the process

$$\cdots abababab \cdots .$$

We combine iterant views term-by-term as in

$$[a, b][c, d] = [ac, bd].$$

We now define i by the equation

$$i = [1, -1]\eta.$$

This makes i both a value and an operator that takes into account a step in time.

We calculate

$$ii = [1, -1]\eta[1, -1]\eta = [1, -1][-1, 1]\eta\eta = [-1, -1] = -1.$$

Thus we have constructed a square root of minus one by using an iterant viewpoint. In this view i represents a discrete oscillating temporal process and it is an eigenform for $T(x) = -1/x$, participating in the algebraic structure of the complex numbers. In fact the corresponding algebra structure of linear combinations $[a, b] + [c, d]\eta$ is isomorphic with 2×2 matrix algebra and iterants can be used to construct $n \times n$ matrix algebra, as we have already discussed.

Now we can make contact with the algebra of the Majorana fermions. Let $e = [1, -1]$. Then we have $e^2 = [1, 1] = 1$ and $e\eta = [1, -1]\eta = [-1, 1]\eta = -e\eta$. Thus we have

$$e^2 = 1,$$

$$\eta^2 = 1,$$

and

$$e\eta = -\eta e.$$

We can regard e and η as a fundamental pair of Majorana fermions.

Note how the development of the algebra works at this point. We have that

$$(e\eta)^2 = -1$$

and so regard this as a natural construction of the square root of minus one in terms of the phase synchronization of the clock that is the iteration of the reentering mark. Once we have the square root of minus one it is natural to introduce another one and call this one i, letting it commute with the other operators. Then we have the $(ie\eta)^2 = +1$ and so we have a triple of Majorana fermions:

$$a = e, b = \eta, c = ie\eta$$

and we can construct the quaternions

$$I = ba = \eta e, J = cb = ie, K = ac = i\eta.$$

With the quaternions in place, we have the braiding operators

$$A = \frac{1}{\sqrt{2}}(1 + I), B = \frac{1}{\sqrt{2}}(1 + J), C = \frac{1}{\sqrt{2}}(1 + K),$$

and can continue as we did above.

10 THE DIRAC EQUATION AND MAJORANA FERMIONS

We now construct the Dirac equation. This may sound circular, in that the fermions arise from solving the Dirac equation, but in fact the algebra underlying this equation has the same properties as the creation and annihilation algebra for fermions, so it is by way of this algebra that we will come to the Dirac equation. If the speed of light is equal to 1 (by convention), then energy E, momentum p and mass m are related by the (Einstein) equation

$$E^2 = p^2 + m^2.$$

Dirac constructed his equation by looking for an algebraic square root of $p^2 + m^2$ so that he could have a linear operator for E that would take the same role as the Hamiltonian in the Schrödinger equation. We will get to this operator by first taking the case where p is a scalar (we use one dimension of space and one dimension of time.). Let $E = \alpha p + \beta m$ where α and β are elements of a a possibly non-commutative, associative algebra. Then

$$E^2 = \alpha^2 p^2 + \beta^2 m^2 + pm(\alpha\beta + \beta\alpha).$$

Hence we will satisfy $E^2 = p^2 + m^2$ if $\alpha^2 = \beta^2 = 1$ and $\alpha\beta + \beta\alpha = 0$. This is our familiar Clifford algebra pattern and we can use the iterant algebra generated by e and η if we wish. Then, because the quantum operator for momentum is $-i\partial/\partial x$ and the operator for energy is $i\partial/\partial t$, we have the Dirac equation

$$i\partial\psi/\partial t = -i\alpha\partial\psi/\partial x + \beta m\psi.$$

Let

$$\mathcal{O} = i\partial/\partial t + i\alpha\partial/\partial x - \beta m$$

so that the Dirac equation takes the form

$$\mathcal{O}\psi(x,t) = 0.$$

Now note that

$$\mathcal{O}e^{i(px-Et)} = (E - \alpha p - \beta m)e^{i(px-Et)}.$$

We let

$$\Delta = (E - \alpha p - \beta m)$$

and let

$$U = \Delta\beta\alpha = (E - \alpha p - \beta m)\beta\alpha = \beta\alpha E + \beta p - \alpha m,$$

then

$$U^2 = -E^2 + p^2 + m^2 = 0.$$

This nilpotent element leads to a (plane wave) solution to the Dirac equation as follows: We have shown that

$$\mathcal{O}\psi = \Delta\psi$$

for $\psi = e^{i(px-Et)}$. It then follows that

$$\mathcal{O}(\beta\alpha\Delta\beta\alpha\psi) = \Delta\beta\alpha\Delta\beta\alpha\psi = U^2\psi = 0,$$

from which it follows that

$$\psi = \beta\alpha U e^{i(px-Et)}$$

is a (plane wave) solution to the Dirac equation.

In fact, this calculation suggests that we should multiply the operator $\mathcal{O}$ by $\beta\alpha$ on the right, obtaining the operator

$$\mathcal{D} = \mathcal{O}\beta\alpha = i\beta\alpha\partial/\partial t + i\beta\partial/\partial x - \alpha m,$$

and the equivalent Dirac equation

$$\mathcal{D}\psi = 0.$$

In fact for the specific ψ above we will now have $\mathcal{D}(Ue^{i(px-Et)}) = U^2e^{i(px-Et)} = 0$. This idea of reconfiguring the Dirac equation in relation to nilpotent algebra elements U is due to Peter Rowlands [33]. Rowlands does this in the context of quaternion algebra. Note that the solution to the Dirac equation that we have found is expressed in Clifford algebra or iterant algebra form. It can be articulated into specific vector solutions by using an iterant or matrix representation of the algebra.

We see that $U = \beta\alpha E + \beta p - \alpha m$ with $U^2 = 0$ is really the essence of this plane wave solution to the Dirac equation. This means that a natural non-commutative algebra arises directly and can be regarded as the essential information in a Fermion. It is natural to compare this algebra structure with algebra of creation and annihilation operators that occur in quantum field theory. to this end, let

$$U^\dagger = \alpha\beta E + \alpha p - \beta m.$$

Here we regard $U^\dagger$ as a formal counterpart to complex conjugation, since in the split quaternion algebra we have not yet constructed commuting square roots of negative one. We then find that with

$$A = U + U^\dagger = (\alpha + \beta)(p - m)$$

and

$$B = U - U^\dagger = 2\beta\alpha E + (\beta - \alpha)(p - m)$$

that

$$[\frac{A}{\sqrt{2}(p - m)}]^2 = 1$$

and

$$[\frac{iB}{\sqrt{2}(p + m)}]^2 = 1,$$

with i a commuting square root of negative one, giving the underlying Majorana Fermion operators for our Dirac Fermion. The operators U and $U^\dagger$ satisfy the usual commutation relations for the annihilation and creation operators for a Fermion.

It is worth noting how the Pythagorean relationship $E^2 = p^2 + m^2$ interacts here with the Clifford algebra of α and β. We have

$$U^\dagger = p\alpha + m\beta + \alpha\beta E$$

$$U = p\beta + m\alpha + \beta\alpha E$$

with

$$(U^\dagger)^2 = U^2 = 0,$$

$$U + U^\dagger = (p + m)(\alpha + \beta),$$

$$U - U^\dagger = (p - m)(\alpha - \beta) + 2E\alpha\beta.$$

This implies that

$$(U + U^\dagger)^2 = 2(p + m)^2$$

$$(U - U^\dagger)^2 = 2(p-m)^2 - 4E^2 = 2[p^2 + m^2 - 2pm - 2p^2 - 2m^2] = -2(p+m)^2.$$

From this we easily deduce that

$$UU^\dagger + U^\dagger U = 2(p + m)^2,$$

and this can be normalized to equal 1.

10.1 ANOTHER VERSION OF U AND $U^\dagger$

We start with $\psi = e^{i(px-Et)}$ and the operators

$$\hat{E} = i\partial/\partial t$$

and

$$\hat{p} = -i\partial/\partial x$$

so that

$$\hat{E}\psi = E\psi$$

and

$$\hat{p}\psi = p\psi.$$

The Dirac operator is

$$\mathcal{O} = \hat{E} - \alpha\hat{p} - \beta m$$

and the modified Dirac operator is

$$\mathcal{D} = \mathcal{O}\beta\alpha = \beta\alpha\hat{E} + \beta\hat{p} - \alpha m,$$

so that

$$\mathcal{D}\psi = (\beta\alpha E + \beta p - \alpha m)\psi = U\psi.$$

If we let

$$\tilde{\psi} = e^{i(px+Et)}$$

(reversing time), then we have

$$\mathcal{D}\tilde{\psi} = (-\beta\alpha E + \beta p - \alpha m)\psi = U^\dagger\tilde{\psi},$$

giving a definition of $U^\dagger$ corresponding to the anti-particle for $U\psi$.

We have

$$U = \beta\alpha E + \beta p - \alpha m$$

and

$$U^\dagger = -\beta\alpha E + \beta p - \alpha m$$

Note that here we have

$$(U + U^\dagger)^2 = (2\beta p + \alpha m)^2 = 4(p^2 + m^2) = 4E^2,$$

and

$$(U - U^\dagger)^2 = -(2\beta\alpha E)^2 = -4E^2.$$

We have that

$$U^2 = (U^\dagger)^2 = 0$$

and

$$UU^\dagger + U^\dagger U = 4E^2.$$

Thus we have a direct appearance of the Fermion algebra corresponding to the Fermion plane wave solutions to the Dirac equation. Furthermore, the decomposition of Uand $U^\dagger$ into the corresponding Majorana Fermion operators corresponds to $E^2 = p^2 + m^2$. Normalizing by dividing by 2E we have

$$A = (\beta p + \alpha m)/E$$

and

$$B = i\beta\alpha.$$

so that

$$A^2 = B^2 = 1$$

and

$$AB + BA = 0.$$

then

$$U = (A + Bi)E$$

and

$$U^\dagger = (A - Bi)E,$$

showing how the Fermion operators are expressed in terms of the simpler Clifford algebra of Majorana operators (split quaternions once again).

10.2 WRITING IN THE FULL DIRAC ALGEBRA

We have written the Dirac equation so far in one dimension of space and one dimension of time. We give here a way to boost the formalism directly to three dimensions of space. We take an independent Clifford algebra generated by $\sigma_1, \sigma_2, \sigma_3$ with $\sigma_i^2 = 1$ for $i = 1, 2, 3$ and $\sigma_i\sigma_j = -\sigma_j\sigma_i$ for $i \neq j$. Now assume that α and β as we have used them above generate an independent Clifford algebra that commutes with the algebra of the σ_i. Replace the scalar momentum p by a 3-vector momentum $p = (p_1, p_2, p_3)$ and let $p \bullet \sigma = p_1\sigma_1 + p_2\sigma_2 + p_3\sigma_3$. We replace $\partial/\partial x$ with $\nabla = (\partial/\partial x_1, \partial/\partial x_2, \partial/\partial x_2)$ and $\partial p/\partial x$ with $\nabla \bullet p$.

We then have the following form of the Dirac equation.

$$i\partial\psi/\partial t = -i\alpha\nabla \bullet \sigma\psi + \beta m\psi.$$

Let

$$\mathcal{O} = i\partial/\partial t + i\alpha\nabla \bullet \sigma - \beta m$$

so that the Dirac equation takes the form

$$\mathcal{O}\psi(x, t) = 0.$$

In analogy to our previous discussion we let

$$\psi(x, t) = e^{i(p\bullet x - Et)}$$

and construct solutions by first applying the Dirac operator to this ψ. The two Clifford algebras interact to generalize directly the nilpotent solutions and Fermion algebra that we have detailed for one spatial dimension to this three dimensional case. To this purpose the modified Dirac operator is

$$\mathcal{D} = i\beta\alpha\partial/\partial t + \beta\nabla \bullet \sigma - \alpha m.$$

And we have that

$$\mathcal{D}\psi = \mathcal{U}\psi$$

where

$$\mathcal{U} = \beta\alpha E + \beta p \bullet \sigma - \alpha m.$$

We have that $U^2 = 0$ and $U\psi$ is a solution to the modified Dirac Equation, just as before. And just as before, we can articulate the structure of the Fermion operators and locate the corresponding Majorana Fermion operators. We leave these details to the reader.

10.3 MAJORANA FERMIONS AT LAST

There is more to do. We will end with a brief discussion making Dirac algebra distinct from the one generated by $\alpha, \beta, \sigma_1, \sigma_2, \sigma_3$ to obtain an equation that can have real solutions. This was the strategy that Majorana [26] followed to construct his Majorana Fermions. A real equation can have solutions that are invariant under complex conjugation and so can correspond to particles that are their own anti-particles. We will describe this Majorana algebra in terms of the split quaternions ϵ and η. For convenience we use the matrix representation given below. The reader of this paper can substitute the corresponding iterants.

$$\epsilon = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \eta = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Let $\hat{\epsilon}$ and $\hat{\eta}$ generate another, independent algebra of split quaternions, commuting with the first algebra generated by ϵ and η. Then a totally real Majorana Dirac equation can be written as follows:

$$(\partial/\partial t + \hat{\eta}\eta\partial/\partial x + \epsilon\partial/\partial y + \hat{\epsilon}\eta\partial/\partial z - \hat{\epsilon}\hat{\eta}\eta m)\psi = 0.$$

To see that this is a correct Dirac equation, note that

$$\hat{E} = \alpha_x \hat{p_x} + \alpha_y \hat{p_y} + \alpha_z \hat{p_z} + \beta m$$

(Here the "hats" denote the quantum differential operators corresponding to the energy and momentum.) will satisfy

$$\hat{E}^2 = \hat{p_x}^2 + \hat{p_y}^2 + \hat{p_z}^2 + m^2$$

if the algebra generated by $\alpha_x, \alpha_y, \alpha_z, \beta$ has each generator of square one and each distinct pair of generators anti-commuting. From there

we obtain the general Dirac equation by replacing $\hat{E}$ by $i\partial/\partial t$, and $\hat{p_x}$ with $-i\partial/\partial x$ (and same for y, z).

$$(i\partial/\partial t + i\alpha_x \partial/\partial x + i\alpha_y \partial/\partial y + i\alpha_z \partial/\partial y - \beta m)\psi = 0.$$

This is equivalent to

$$(\partial/\partial t + \alpha_x \partial/\partial x + \alpha_y \partial/\partial y + \alpha_z \partial/\partial y + i\beta m)\psi = 0.$$

Thus, here we take

$$\alpha_x = \hat{\eta}\eta, \alpha_y = \epsilon, \alpha_z = \hat{e}\eta, \beta = i\hat{e}\hat{\eta}\eta,$$

and observe that these elements satisfy the requirements for the Dirac algebra. Note how we have a significant interaction between the commuting square root of minus one (i) and the element $\hat{e}\hat{\eta}$ of square minus one in the split quaternions. This brings us back to our original considerations about the source of the square root of minus one. Both viewpoints combine in the element $\beta = i\hat{e}\hat{\eta}\eta$ that makes this Majorana algebra work. Since the algebra appearing in the Majorana Dirac operator is constructed entirely from two commuting copies of the split quaternions, there is no appearance of the complex numbers, and when written out in 2×2 matrices we obtain coupled real differential equations to be solved. Clearly this ending is actually a beginning of a new study of Majorana Fermions. That will begin in a sequel to the present paper.

11 BIBLIOGRAPHY

[1] G. Spencer–Brown, "Laws of Form," George Allen and Unwin Ltd. London (1969).

[2] Kauffman, L. [1985], Sign and Space, In Religious Experience and Scientific Paradigms. Proceedings of the 1982 IASWR Conference, Stony Brook, New York: Institute of Advanced Study of World Religions, (1985), 118-164.

[3] Kauffman, L. [1987], Self-reference and recursive forms, Journal of Social and Biological Structures (1987), 53-72.

[4] Kauffman, L. [1987], Special relativity and a calculus of distinctions. Proceedings of the 9th Annual Intl. Meeting of ANPA, Cambridge, England (1987). Pub. by ANPA West, pp. 290-311.

[5] Kauffman, L. [1987], Imaginary values in mathematical logic. Proceedings of the Seventeenth International Conference on Multiple Valued Logic, May 26-28 (1987), Boston MA, IEEE Computer Society Press, 282-289.

[6] Kauffman, L. H., Knot Logic, In *Knots and Applications* ed. by L. Kauffman, World Scientific Pub. Co., (1994), pp. 1-110.

[7] Kauffman, L. H. Knot logic and topological quantum computing with Majorana fermions. In "Logic and algebraic structures in quantum computing and information", Lecture Notes in Logic, J. Chubb, J. Chubb, Ali Eskandarian, and V. Harizanov, editors, 124 pages Cambridge University Press (2016).

[8] Kauffman, L. H. and Lomonaco, S. J. Braiding, Majorana Fermions and Topological Quantum Computing , (to appear in Special Issue of QIP on Topological Quantum Computing) (2017).

[9] Kauffman, L. H. Iterants, Fermions and Majorana Operators. In "Unified Field Mechanics - Natural Science Beyond the Veil of Spacetime", edited by R. Amoroso, L. H. Kauffman, P. Rowlands, World Scientific Pub. Co. (2015). pp. 1-32.

[10] Rukhsan Ul Haq and Kauffman, L. H. Iterants, Idempotents and Clifford algebra in Quantum Theory (in preparation).

[11] Kauffman, Louis H. [2002], Biologic. *AMS Contemporary Mathematics Series*, Vol. 304, (2002), pp. 313 - 340.

[12] Kauffman,Louis H.[1991,1994,2001,2012], *Knots and Physics,* World Scientific Pub.

[13] L.H. Kauffman, *Temperley-Lieb Recoupling Theory and Invariants of Three-Manifolds*, Princeton University Press, Annals Studies 114 (1994).

[14] Kauffman, Louis H. [2002], Time imaginary value, paradox sign and space, in *Computing Anticipatory Systems, CASYS - Fifth International Conference*, Liege, Belgium (2001) ed. by Daniel Dubois, AIP Conference Proceedings Volume 627 (2002).

[15] Kauffman,Louis H. and Noyes,H. Pierre [1996], Discrete Physics and the Derivation of Electromagnetism from the formalism of Quantum Mechanics, *Proc. of the Royal Soc. Lond. A*, 452, pp. 81-95.

[16] Kauffman,Louis H. and Noyes,H. Pierre [1996], Discrete Physics and the Dirac Equation, *Physics Letters A*, 218 ,pp. 139-146.

[17] Kauffman, Louis H. [1998], Noncommutativity and discrete physics, *Physica D* 120 (1998), 125-138.

[18] Kauffman, Louis H. [1998], Space and time in discrete physics, *Intl. J. Gen. Syst.* Vol. 27, Nos. 1-3, 241-273.

[19] Kauffman, Louis H. [1999], A non-commutative approach to discrete physics, in *Aspects II - Proceedings of ANPA 20*, 215-238.

[20] Kauffman, Louis H. [2003], Non-commutative calculus and discrete physics, in *Boundaries- Scientific Aspects of ANPA 24*, 73-128.

[21] Kauffman, Louis H. [2004], Non-commutative worlds, *New Journal of Physics 6*, 2-46.

[22] Kauffman, Louis H., Non-commutative worlds and classical constraints. in "Scientific Essays in Honor of Pierre Noyes on the Occasion of His 90-th Birthday", edited by John Amson and Louis H. Kaufman, World Scientific Pub. Co. (2013), pp. 169-210.

[23] Louis H. Kauffman, Differential geometry in non-commutative worlds, in "Quantum Gravity - Mathematical Models and Experimental Bounds", edited by B. Fauser, J. Tolksdorf and E. Zeidler, Birkhauser (2007), pp. 61 - 75.

[24] Kauffman, Louis H. [2012], Knot Logic and Topological Quantum Computing with Majorana Fermions, (to appear).

[25] Kauffman, Louis H.; Lomonaco, Samuel J., Jr. q-deformed spin networks, knot polynomials and anyonic topological quantum computation. J. Knot Theory Ramifications 16 (2007), no. 3, 267–332.

[26] E. Majorana, A symmetric theory of electrons and positrons, I Nuovo Cimento,**14** (1937), pp. 171-184.

[27] G. Moore and N. Read, Noabelions in the fractional quantum Hall effect, *Nuclear Physics B360* (1991), 362 - 396.

[28] V. Mourik,K. Zuo, S. M. Frolov, S. R. Plissard, E.P.A.M. Bakkers, L.P. Kouwenhuven, Signatures of Majorana fermions in hybred superconductor-semiconductor devices, arXiv: 1204.2792.

[29] T. P. Cheng and L. F. Lee, "Gauge Theory of Elementary Particles" , Clarendon Press. Oxford (1984,1988).

[30] S. Gasiorowicz, "Elementary Particle Physics", Wiley, N.Y. (1966).

[31] Sundance O. Bilson-Thompson, A topological model of composite fermions, arXiv:hep-ph/0503213.

[32] D. E. Littlewood, "The Skeleton Key of Mathematics", Harper Torchbook Edition (1960).

[33] Peter Rowlands, "Zero to Infinity - The Foundations of Physics", Series on Knots and Everything - Volume 41, World Scientific Publishing Co., 2007.

[34] L. Wittgenstein, "Tractatus Logico - Philosophicus", New York: Harcourt, Brace and Company, Inc., London: Kegan Paul, Trench, Trubner and Co. Ltd. (1922).

THE INTRODUCTION AND DEVELOPMENT OF SYMBOLIC REPRESENTATION IN MATHEMATICS DURING THE LATE RENAISSANCE

MARTIN LOWE

Tweedvale Cottage, Walkerburn, Peebleshire, Scotland

This paper was to be presented by its author Martin Lowe at the Post-ANPA meeting in August 2017, Anstruther, Fife. Martin died shortly before that meeting took place. The original typescript of this and other archived papers by Martin were bequeathed to his friend and colleague Nicola Graves-Gregory with a view to making them eventually accessible to the ANPA membership.

To be able to include this present paper into these proceedings it has been transliterated into the standard LaTeX format. Whilst this transliteration has been denied the opportunity of being checked by its author it has been done with care to ensure it can be an accurate record of one of Martin's contributions to the Alternative Natural Philosophy Association.

John Amson, St Andrews University, 2018

1 INTRODUCTION

The central thesis of this essay is that the introduction of symbolic mathematics was a result of a new understanding of the essential nature of mathematics. Simultaneously there emerged a new conception of nature and the knowledge that could be obtained about it. This 'symbolic science' required the key of mathematics, not as previously conceived however, but as a 'science of symbols'. Rather than it being an innovation of content - a new branch of mathematics or a new technique - it provided a new ***fora*** to mathematical knowledge as a whole.

To arrive at a symbolic mathematics required an enormous conceptual leap. Nevertheless the transition can be traced through the relatively short period of a few decades, on either side of the year 1600. The speed with which this new mode of thought came to be adopted and uncritically accepted within seventeenth century science and mathematics was largely because the

key elements that were forged together to create this new 'conceptual space' were already widely diffused and incorporated in much of the thought of the later fifteenth century and sixteenth century. The nature of this advance, and the associated problems of historical analysis are recognized by Wilder:

> The tremendous spurt in mathematical analysis during the seventeenth century was not, it seems fair to say, merely a concomitant of the general cultural advances being made in Europe at that time, unless, indeed, one manages to tie these advances in with the great symbolic achievements of Vieta, Descartes, Leibnitz and others of that period. For on analyzing the mathematical progress of that time one is struck by how much of it consisted of the invention of a new and powerful symbolic apparatus. (Wilder, p. 165)

Looking at physics from the standpoint of the twentieth century, a persuasive argument can be made for the assertion that "the main import of mathematics is that it provides a universal system of symbols rather than a means of quantitative judgment" (Hutten). It is a symbolic understanding that lies at the heart of 'modern' science and 'modern' mathematics. This can already be clearly seen in the mathematics of the later seventeenth century. To take the differential calculus as an example, we find that Boyer concludes in his study of its history:

> More important than this (the arithmetization of mathematics), in the development of the algorithm of the calculus, was the systematic introduction, during the later sixteenth century, of symbols for the quantities involved in algebraic relations. This literal symbolism permitted the concepts of variability and functionality to enter into algebraic thought. (Boyer, p. 98)

Cassirer also stresses the break between 'modern' mathematics and the previous forms:

> That mathematics is a universal symbolic language — that it is not concerned with a description of things but with general expressions of relations — is a conception which appears rather late in the history of philosophy. A theory of mathematics based on this presupposition does not appear before the seventeenth century. (Cassirer, 1944, p.273)

A critique of approaches adopted by historians of mathematics and an outline of the historical view of this essay is contained in the Appendix *Notes on the History of Mathematics*.

2 SIXTEENTH CENTURY MATHEMATICS

The 're-discovery' of Greek mathematics, reinterpreted within the complex set of intellectual ideas, referred to as renaissance neoplatonism, played a vital rôle in forming a new reconciliation between the Aristotelian and Platonic traditions within mathematics. In the search for the 'lost' knowledge of former times, the pythagorean–platonic belief that mathematics was the key to an understanding of the entire universe, provided the inspiration for much of the mathematical work of the sixteenth and seventeenth centuries. It was not until the second half of the sixteenth century that the more advanced Greek mathematics was given serious attention. In fact, it is only in this period that many of the most important Greek mathematical works were first translated (for instance, to the reader unversed in Greek, Diophantus' Arithmetic was not not available until 1575 and Pappus' Mathematical Collections until 1588).

One of the features of the neo-platonic influence was the rich diversity of meanings and interpretations that it gave to mathematics: from numerology to precise quantification, from mystical and symbolic reflections of the universe in the mind to the belief in the potential unification of mathematics. No brief summary could capture even the variety of the vast range of mathematical philosophies that were conceived and propagated in this period. In order to highlight some of the main characteristics of mathematical beliefs prior to the transition to modern symbolic mathematics, John Dee's Mathematical Preface of 1570 is a useful source. This was written as an introduction to the first English translation of Euclid. In the sixteenth century the Elements of Geometry was by no means an elementary textbook, for it contained some of the most advanced mathematics of the day. Fifteen years later Simon Stevin was still able to say that:

> the tenth book of Euclid ... has become such a horror they have called it the cross of mathematics, a subject hard to digest, and without any use. (Stevin, Arithmetic, IIB, p.536)

There was a neo-platonic tradition, dating back to Proclus, which assigned to Euclid's Elements a status as the gateway to the deeper mysteries of mathematics. Thus the Commentary on Euclid was often used as a platform for presenting a general statement of mathematical philosophy to the novitiate, who had as yet only mastered the elementary techniques of the subject. Tartaglia, in the frontispiece of his Nova Scientia (reproduced in Strong, p. 58), provides a clear example of this mathematical hierarchy which the student must ascend if he is to arrive at an understanding of higher things: Euclid stands

at the entrance to a walled enclosure in which we see Tartaglia flanked by Arithmetica and Geometria, behind them stand Hydromantia, Geomantia, Architectura, Astrologia, Cosmografia, Necromantia, Prestidigitio and Sortilegio. After passing this group the student will meet Aristotle and then Plato standing at the gateway to a further enclosure in which Philosofia is to be found seated on her throne.

Dee's mathematical beliefs, expressed in the Preface, are based on the original triadic division of Plato, as embellished and re-interpreted into a hierarchically–structured trichotomous world by the early neo-platonists Nicomachus, Theon and Proclus. For Dee, "all things" are either "Supernaturall, Naturall or of a third being". "Thinges Supernaturall", which he claims can only be comprehended with the mind, "are immaterial, simple, indivisible, incorruptible and unchangeable". Whereas "Thinges Naturall, of the sense exterior, are able to be perceived". He explains that "the state, condition, nature and property of these thinges which are called Thynges Mathematicall" is best understood in terms of their relationship to these two, "beying (in a manner) middle". The material of mathematics is thus neither of the intellect nor of the sense-world; it has its own ontological status and yet it can somehow reach into both these 'worlds', "A mervaylous newtralitie have these thinges Mathemaicall, and also a strange participation between thunges supernatiruall ... intellectual ... and thynges naturall ... sensible".

Deeply immersed in the hermetic-cabalistic movement, Dee subscribes to a mystical and apparently non-quantitative interpretation of number: "Number hath a treble state: One, in the creator; another in every Creature and the third, in Spirituall and Angelicall Myndes, and in the soul of man." "By Numbers propertie ... we may ... mount up ... in spirit to behold in the Glas of Creation, the Forme of Formes, the Exemplar Number of all thinges Numerable: both visible and invisible: mortall and immortall, Corporall and Spirituall". He does not allow, however, this question of their Being and its ontological problems to impede his search for a mathematical structure to the sensible world. He argues for a *quantitative* treatment of Forms and provides an example of this approach in which he uses the "great Arte of Algiebar" to solve problems involving their mixtures (see Clulee). It is because Dee is so far from the aristotelian insistence on the physical reference of all mathematics — "number and spatial magnitudes cannot exist apart from things" that we can see in his formulation the emerging symbolic character of mathematical objects, later to be conceived of as being purely intellectual and yet, being symbols, capable of precise representation of the material world.

There is a fundamental epistemological distinction to be drawn between the two types of knowledge — *opinio* and *scientia*. Knowledge of the natural world, in which "probabilitie and conjecture hath place", stands in contrast to mathematical knowledge:

> In Mathematicall reasoninges, a probaball Argument, is nothing regarded: nor yet the testimony of sense, any whit credited: But onely a perfect demonstration, of truthes certaine, necessary, and invinciple: universally and necessarayly concluded: is allowed as sufficient for an Argument exactly and purely Mathematicall.

Here we encounter the notion that the certainty that only mathematics can have is derived from its axiomatic base and logical procedures. This belief that a certain and universal mathematical science could be invented, or awaited re-discovery, required that central importance be attached to the role of 'demonstration' in mathematics. The early neo-platonic commentators were also the source of this belief that there may be found:

> an independent science which is far superior in rank to geometry and arithmetic: 'the knowledge of these things belong to a science prior by far' and the 'ascent from more partial to more universal understandings' is that by which 'we climb up to the very science of "being" insofar as it is "being" '. In the sixteenth and seventeenth centuries these words of Proclus acquired tremendous significance, since ... they were normally understood as reference to the *Mathesis universalis*. (Klein, p. 181)

Within traditional Greek thought mathematics was divided into the two separate disciplines, Arithmetic and Geometry, distinguished by the nature of the quantities with which the y dealt. Dee endorses this position — "of mathematicall thinges, are two principall kindes: namely, Number and Magnitude" but he does so with an uneasiness characteristic of the sixteenth century . For the goal of many mathematicians, Dee included , was to realize the complete unity of mathematics. There were three principal approaches that were tried: The 'arithmetization' of geometry, the geometrical interpretation of arithmetic and the search for a 'prior science' from which both arithmetic and geometry could be derived. For example, as Strong points out, Clavius in his 'Practical Arithmetic' of 1584:

> takes occasion to express his discontent with the contemporary separation between geometrical and arithmetical quantity, the former continuous and the latter discrete ... he does urge the arithmetical treat-

> ment of geometrical line and the algebraic formulation of arithmetical 'precepts'. (Strong, p. 105)

In his discussion of the objects of arithmetic and geometry, while carefully preserving their distinctiveness, Dee emphasises analogous natures. "Number we define, to be, a certayne Mathematicall Sume, of *Units*", whereas Magnitude as mathematical line is constituted of geometric points. "We account an Unit, a thing Mathematicall, though it be no number" and similarly, "a Poynt we define, by the name of a thing Mathematicall: though it be no Magnitude". When he comes to describe the 'derivative' mathematical arts he notes that "vulgar Practisers, have Numbers, otherwise, in sundry Considerations: and extend their name farder, then to Numbers, whose least part is an Unit." Each of these extended meanings of number has its own 'vulgar' or 'impure' type of arithmetic. One of these, Arithmetike Circular, he admits utilises a concept of number which is equivalent to magnitude:

> Practise hath led numbers farder, and hath framed them, to take upon them, the shew of Magnitudes propertie: Which is Incommensurabilitie and Irrationalitie. And here, Numbers are become, as Lynes, Playnes and Solides.

He does not appear, though, to wholeheartedly approve of this method of achieving their unification. Significantly he sees "Cossike Numbers" and thus algebra as another branch of vulgar arithmetic, rather than as a 'prior science' as Vieta and Descartes were to do.

It should be noted that Dee develops a typology of numbers which classifies the different species of numbers but leaves their relationship unclear. The assumption that all numbers were not of the same nature was universally accepted and even within the realms of whole numbers '8' as 'eight units' was not commensurate with the cubic number 'two cubed'. The need for these various classifications was removed by the introduction of a single symbolic concept, that could include all numbers. Despite its apparent theoretical unimportance, Stevin's symbolic number is exemplary of the transformation that was taking place in the conception of mathematical and scientific knowledge. It can be seen clearly in Stevin's work that the transition to symbolic mathematics was not the result of a gradual process of ever-increasing abstraction. It required the rejection of accepted 'natural' conceptions of knowledge and the acceptance of a symbolic one.

3 STEVIN'S CONCEPT OF NUMBER

It is not easy to guage the sources which Simon Stevin (1548-1620), largely self-educated, used if one limits oneself to reading his 'mathematical' and 'scientific' works. It is a little far-fetched, for instance, to suggest that he appeared "as if from nowhere" as Bochner does (complete with reference to the analogous case of the biblical prophet, Elijah the Tishbite! !):

> Stevin appeared on the scientific stage rather abruptly, as if from nowhere, without any recognizable direct link either to the mechanics or the cosmology that preceded him. (Bochner, p.109)

Given that "Stevin consciously breaks with the traditional forms of science" (Klein, p. 186) it has become easy to cast him in the role of the 'first of the moderns' as if this outlook requires no further explanation or that he had no intellectual precursors.

There are also difficulties in assessing the influence of Stevin. He mainly wrote for the 'mathematical practitioners' and thus wrote in Flemish; as Struik says in his commentary on The Tenth: "he wanted to be read outside the charmed circle of humanists and cossists". This meant that although his ideas were accessible to Flemish-speakers they could not easily be read by the rest of Europe. Some of his books were available in other languages during his lifetime — The Arithmetic in French and The Tenth in both French and English. It is clear that his innovations in many fields were later taken up, but not to what extent their adoption can be attributed to him, for there were others also working on the development of Euclid and the newly translated Archimedes in a similar spirit but apparently independently. Boyer leaves the question open: "it would be interesting to know to just what extent the work of Stevin was known in France during the first half of the seventeenth century" (Boyer, p.141) and suggests that maybe he was the "victim, because of his liberality of thought, of a conspiracy of silence on the part of Catholics in the Low Countries" (Boyer, p. 141).

That Descartes was familiar with Stevin's mathematical ideas I think can hardly be doubted. There are many possible intermediaries who knew Stevin or were familiar with his work and at the same time were in contact with Descartes and Klein argues that "it is possible, even likely, that Descartes saw Stevin in the last years of his life, and he may even have been instructed by him" (Klein, p. 292). As regards Vieta, he has not been linked with Stevin as far as I know and it is difficult to trace an influence of Stevin in his writings.

Although much of his theoretical studies, Stevin is clearly debating with the rediscovered Greek texts, he has little respect for Greek knowledge and the Greek mathematicians: the Greeks (and the Chaldeans, the Hebrews and the Romans as well) "all of whom were no arithmeticians worth the name" (Stevin, *The Age of the Sages*, III p.599). He does not accept the pythagorean–platonic framework as he believes that the period from the beginning of Greek civilisation up until 150 years before his time should be termed the 'Age of the Ignorant', It is not to the Greeks but to prior periods that interest him:

> Age of the Sages we call that time in which exceptional learning was to be found among men, in fact which we perceive with certainty from certain signs, but without knowing among whom, where or when. (ibid., p. 593)

He sets himself the task of restoring this ancient pre-Greek knowledge and assumes that the Arabic mathematical works represent a less adulterated source than the Greek, who, for instance, misunderstanding the origin or point of number ... stated it to be unity" and who did not even have the vestiges of:

> the highly singular and admirable Rule of Algebra, the Inexhaustible fountain of infinite Arithmetical Theorems, Revealer of the mysteries hidden in numbers. (ibid., p. 681)

(I shall note in the next section the strange reputation that algebra had acquired during the sixteenth century.)

Although, unlike Dee, he has little regard for Greek ideas in general and platonic ones in particular, and does not have the aspiration of becoming a 'magus', he does subscribe to the view that knowledge of the world is essentially mathematical in form and that, furthermore, this knowledge was known by the people of some long-forgotten civilisation. Of the six 'signs' that Stevin loists which demonstrate the existence of this former wise age only one, alchemy, is not explicitly mathematically based. (Interestingly he says of the "marvellous and very unusual subject of alchemy ... Hermes Trismegistus is considered to have been the most skilled of those of whom written information has come down to us" (ibid., p.605).)

The first sign is astronomy "which at the time of Hipparchus and Ptolemy had almost disappeared" —- Stevin being an ardent supporter of the heliocentric theory. The second is arithmetic and "though Euclid writes elegant arithmetic theorems" they "bear witness to the Age of the Sages which had previously existed and did not exist then" (ibid., p.60) Similarly he says of

geometry, the third sign, that the Greeks had "got it from others". The final sign he mentions is magic "though this is frightening, it is yet perceived what great learning has resulted from mathematics, which is regarded as the cause of such thorough knowledge" (ibid., p. 607).

In The Age of the Sages Stevin puts forward the vision of restoring their knowledge and sets out a programme designed to "attain gradually to such great learning again as existed in the Age of the Sages". "In the first place, what we lack is a large body of data obtained by practical experience, on which the sciences can be firmly founded" (ibid., p. 609). To achieve this requires that the native language of each country is used, since vast numbers of people would have to be mobilised — "some people think the matter too sublime to be studied by ordinary people ... I hold a different view ... when the science is studied by ordinary people, everything can be obtained for nothing with greater certainty" (ibid., p. 613).

He regards it as essential that a clear distinction be made between 'theory' and 'practice' which he explains are respectively, "a fictitious operation *without natural material*" and "an operation which essentially takes place with natural material". Here, bearing in mind that he is discussing mathematical knowledge, he introduces the modern distinction between 'pure' and 'applied' mathematics. He argues that practice not based on theory or theory without application is useless; neither can be neglected even by the practitioner for "the property and the end of theory is that it furnishes a sure foundation for the method of practical operation".

In order to bring arithmetic back to its proper 'theoretical' form he adopts the method of presentation used in Greek geometry "in which besides the subject of geometry something very peculiar, extraordinary, and useful is also to be noted and learned, *viz.* the description of mathematics" (*ibid.*., p. 605). Stevin proceeded to build an axiomatic base for arithmetic using the "style of Euclid" and the methods of 'analysis' and 'synthesis' (that Vieta was to give such a crucial rôle to in his creation of symbolic algebra):

> He gives his demonstrations in the classical way with the terms *Given, Required, Construction, Demonstration, Conclusion*, which shows that he realised that careful proofs are as necessary in arithmetic as in geometry, an unusual thing for his day. (Struik in Stevin, IIA, p. 377)

The most important of Stevin's mathematical works were published within the six year period from 1582 to 1588, soon after he had settled in the Northern Netherlands. Although the full geometrical text was never published, it appears that Stevin worked on a systematic exposition of both geometry and

arithmetic in close parallel in order to bring out most fully their common structure. In his *Problemata Geometrica* of 1583 he writes:

> we have arranged geometry (which we shortly hope to publish) in a Method similar to the method of Arithmetic ... because of the great agreement between continuous and discontinuous quantities ... (Stevin, P.G., IIA, p. 207)

Throughout his works he is always seeking ways of equivalencing the results and methods of arithmetic and geometry. He showed that the principal arithmetical operations, the theory of proportions and the 'regula falsi' all had their counterparts in geometry and conversely he 'translated' geometry back into arithmetic:

> He went so far as to say, in his *Traicte des Incommensurables Grandeurs*, that the geometric theory of incommensurables in Euclid's Tenth Book had originally been discovered in terms of numbers, and translated the content of this book into the language of number. (Struik in Stevin, IIA, p.5)

Above all he attempts to place arithmetic on the sure foundations that only an axiomatic development can provide. The Arithmetic, published in French (1585), opens with a series of definitions. Number is first defined in the following way "number is that by which we say that the quantity of each thing is explained". Following from this, he states that "to a continuous magnitude corresponds the continuous number to which it is attributed"; from which he derives the conclusion "that number is not discontinuous quantity".

Stevin shows that zero and not the unit is the true "beginning of number" in that it is strictly analogous in properties and effects to the geometrical point in relation to geometrical magnitude. Having constructed *'continuous number'*, he demonstrates that, within his new conceptualisation, "UNITY IS NUMBER". He is at last able to create a totally abstract concept of number , divorced from any physical referral:

> Number has two species, of which one is explained with an adjective of magnitude, like the square, cubic, root numbers, which we call geometrical numbers; the other species is simply explained without the afforementioned adjective, like one, two, three, three-fifths. We call these numbers in distinction to the other species, arithmetical numbers. (Stevin, *The Arithmetic*, IIB, p. 505)

Without an 'adjective of magnitude', Stevin's arithmetic numbers can now include within a single homogeneous species fractions, irrationals, powers

and roots. He demonstrates that "any (arithmetic) numbers can be squares, cubes *etc.* also that any root is (arithmetic) number" (ibid. p. 529). With characteristic directness he confronts the "irrational, irregular, inexplicable, absurd opinion" that 'irrational' and 'root' numbers are any less "perfect and excellent" than the integers. He pointed out the relativity of these terms "since the same line, called 5 feet in one town, could be called $4 + 1^2$ feet in another".

In creating this new number concept Stevin has not simply removed the previously fundamental distinction between geometrical magnitude and arithmetical number and extended the concept of number to include fractions and fractional exponents. He has altered the ontological status of arithmetic. The objects of arithmetic are now symbolic numbers, whose 'material' is not of the sense-world but of the intellect alone. He presents a new kind of arithmetic which operates on "characters of cyphers, whereby a certain number is described" (Stevin, *The Tenth*, IIA p. 403). Nevertheless the theories of this 'pure' arithmetic can be applied to geometrical numbers as well and thus it can also operate within the natural world:

> Stevin no longer deals with numbers of units which are determinate in each case but with the unlimited possibility of combining ciphers according to definite 'rules of calculation'. But this means that he conceives the 'quantities' with which he is dealing in a symbolic way; he no longer knows 'numbers of units' but only 'numbers' directly expressed in ciphers. (Klein, p. 193)

4 SYMBOLIC ALGEBRA

Symbolic algebra, like symbolic number, emerged from the conjunction of ideas drawn from Arabic and Greek sources; the vague symbolism of neeplatonic mathematics is re-interpreted, made precise and 'certain' by its placement within a logical structure that allows a new compatibility with the conventional aristotelian world-view. They are also alike in that, when introduced, their potential was at first unrealised neither had at first succeeded in solving new problems. Their adoption was not the obvious next step for mathematics to take, in fact, throughout the seventeenth century many highly competent mathematicians opposed this 'fashionable' trend, or even failed to notice that it involved a radical break with previous conceptions of mathematics. It was not till Descartes, Newton and Leibnitz that the power of 'symbolic mathematics' was clearly demonstrated.

Arabic works on algebra had been available in Latin translation for several hundred years in Europe by the close of the sixteenth century. Its esoteric reputation, varying from a divine art to a black one, had successfully kept its development apart from the mainstream of school mathematics. Nevertheless the contributions of the efforts of many mathematicians (e. g. Stifel, Cardan, Tartaglia, Bombelli *etc*.) had transformed it out of all recognition from the original form in which it had entered Europe with little or no mathematical notation (Cajori, p. 246). While it was being continually elaborated "in respect to techniques of calculation and by its whole mode of operating with numbers and number signs, its self-understanding failed to keep pace with these technical advances" (Klein, p. 147).

It was not until the last quarter of the sixteenth century that algebra or 'almucabala' (the book of hidden things, as it was sometimes called) became accepted within conventional mathematics. Traditionally it had been seen as a branch of arithmetic concerned with the *determinate* solution of particular problems that were amenable to being set up as *equations*. Although it required the almost *ad hoc* application of a number of techniques or rules to the given problem, until the determinate number that was the solution emerged, it was believed that they could all be circumscribed within a single rule. This understanding is clearly revealed by Robert Recorde who says of "the rule of algebra ... rightly called, the rule of equation":

> When any question is propounded, appertaining to this rule, you should imagine a name for the number, that is to be sought. And with that shall you proceed, according to the question, until you find a Cossik nomber, equalle to that nomber. (Recorde)

This widespread conception of algebra, in which it is merely an arithmetic technique, albeit a most useful one, is characterised by its lack of a notion of generalised relationship or even indeterminacy. This syncopated algebra is still firmly rooted, despite the profusion of ideographic symbols, in the 'world of 1anguage'. As Koyre observes:

> Renaissance algebra, instead of producing formulae, offered rules and examples. ... Renaissance mathematicians and grammarians alike thought in terms of cases, varying from ... concrete problem to concrete problem. (Koyre, 1964, p.51)

It was Gosselin, in his De Arte Magna of 1577, who was probably the first to assign algebra to the study of ratios and proportions. I would support the opinion of Klein:

> the significance of the concept and the theory of proportions for the origin of modern science cannot be overstated. ... This influence has two tendencies which, having a common origin, finally merge again; one of these leads ... to algebra ... the other to physics and astronomy (Kepler, Galileo) ... both continue the pythagorean-platonic tradition, albeit in different respects and with diverging tendencies. (Klein, p. 264)

The discovery of symbolic algebra was beset by nationalist and other controversies from which, even now, it is difficult to adjudicate with certainty on the claims of the various contenders. Hariot, as well as Ghetaldi and Oughtred, have all had their various supporters. While it is clear that Hariot made many improvements in notation and, according to Tanner, Hariot's "manuscripts confirm that symbols for him were a medium for mathematical thinking , not a mere means of avoiding tedious repetition" (Tanner, p. 290), little evidence exists that he might have preceded Vieta. In any case, it was Vieta, in 1591, who first gave the world a published account of the "new algebra" in his *Introduction to the Analytic Art*. The issue was further confused by the belief that the Greeks had already had this 'art' (even in the 1690's Wallis "cannot doubt this") and so it had been merely 'restored' rather than invented.

Vieta is well aware of the importance of his three-fold 'analytic art' and 'knows' that he has found the 'hidden gold' on which algebra's great and mysterious reputation had been based:

> underneath the Algebra or Almucabala which they lauded and called 'the great art', all Mathematicians recognised that incomparable gold lay hidden, though they used to find very little. (Vieta, Preface to Catherine of Parthenay, p. 319)

But Vieta does not regard himself as the discoverer of the new algebra — 'logistice speciosa' — for he has only restored the ancient analysis of the Greeks, that they carefully concealed from their texts:

> Diophantus ... presented (his work) as if established by means of numbers and not also by species (which, nevertheless, he used), in order that his subtlety and skill might be the more admired. (Vieta, p. 345)

This 'Greek' analytic art, so he believed, had been brought into Europe through the Arabic texts which, however, were grossly distorted:

> the art which I present is new, but in truth so old, so spoiled and defiled by the barbarians, that I considered it necessary, in order to introduce an *entirely new fora* into it ... lest it should retain its filth and continue to stink in the old way. (Vieta, Preface, p. 318)

Vieta continues in the tradition of the 'commentators', in terpreting the newly translated Greek texts, not to develop them further or to make innovations, but simply to restore from them "the surest finder of all things mathematical" — the ancient art which "appropriates to itself by right the proud problem of problems, which is: TO LEAVE NO PROBLEM UNSOLVED" (Vieta, p. 353)

The 'science prior to arithmetic and geometry', that Proclus had promised, Vieta creates by the construction of an axiomatic algebra which operates on a transformed object — no 1anger 'number' but 'species', Vieta first applies the 'zetic art' as he has seen Diophantus do in his 'Arithmetic', to set up an equation in terms of species:

> the zetetic art does not employ its logic on numbers — which was the tediousness of the ancient analysts — but uses its logic through a logistic which in a new way has to do with species. (ibid. p. 321)

Equations are shown to be equivalent to proportions, and using the geometrically conceived *analysis* of Theon and Pappus but applied now to species, Vieta can produce solutions to general problems because his algebra does not require determinate solutions. The solutions of general problems, however, provide not numbers but *algebraic theorems*. The essential indeterminate nature of species gives algebra the equivalent of the general theory of proportions and so the 'formula' becomes available as a powerful mode of expression of the new science.

The "species or forms of things, as for example, the letters of the alphabet" (ibid. p. 328) as symbols indeterminate objects of Vieta's new algebra. The species represent general magnitude and depending on the nature of the problem can represent either arithmetic or geometric quantity. If a particular solution is required the species will take on the form of either an arithmetic number or a geometric quantity . Algebra is no longer a pragmatic rule but an axiomatic system from which theories can be demonstrated and which operates analytically on 'species' that are, at one and the same time, the symbolic entities which are the objects of analysis and symbols for magnitude in general or of particular type.

5 THE EMERGENCE OF MODERN SYMBOLIC SCIENCE

We have looked above at the two aspects of mathematics, number and algebra, that had been transformed into symbolic form by the end of the sixteenth century. In the following century there were many important contributions made to the development of modern symbolic science. Descartes showed that symbolic algebra could be applied to geometric figures and put forward the view that there was an identity between the form of this symbolic mathematics and the form of all certain knowledge of the material world. Newton already presupposed symbolic form to his science although he would often present his results in a less radical fashion. Much of Leibnitz's work is devoted to systematically introducing symbolic form into new areas of mathematics and logic and to extending the realms of application of these symbolic operations.

Of the strands of thought that finally coalesced into the conception of symbolic science, there are three that I believe were crucial in determining the conceptual structure of modern science. (1) The search for *certain* knowledge and the identification of the nature of this knowledge as mathematical. (2) The belief that certain knowledge is *symbolic* in form. (3) The attempt to 'reconstruct' a truly *universal science*. Each of these are complex intellectual trends and here, I can only suggest as to their historical influences.

Descartes, when he wrote in his Discourse on Method that "of all those who have hitherto sought for the truth in the Sciences, it has been the mathematicians alone who have been able to succeed in making any demonstrations, that is to say, producing the reasons which are evident and certain" (Descartes, p. 92) was following a long tradition which had as its goal, *scientia* — true and certain knowledge. It has already been pointed out that during the latter part of the sixteenth century there was a growing awareness that the truth of mathematical propositions was not derived from their relation to the natural world but from their logical equivalence to an axiomatic base.

The development of a symbolic conception of knowledge was only loosely associated with mathematics, at least when conceived as quantitative science, and often was in explicit opposition to it (e. g. Bruno, see Michel, p. 291). Also its connections to the natural world was not alway very close. Such symbolic knowledge, identified with the world of intellect, was increasingly ex pected to conform to the rigours of logic. Cusanus' writings were influential in directing the flow of thought in this direction, as Cassirer points out:

> he requires of the symbols in which the divine becomes graspable by us not only sensible fullness and force but also intellectual precision

> and certainty not satisfied with the imagistic and sensible force of the signs in which we read the spiritual structure of the universe; instead it requires of these signs that they form a system, a thoroughly ordered whole. (Cassirer, 19 72, pp. 53-4)

This conception of knowledge was closely related to the various interpretations of the Doctrine of Signs. It is only within the context of this doctrine that the various meanings given to the concept of sign, can be understood. The belief that the 'book of nature' was written by the divine hand led to many attempts to decipher the characters and unravel the syntax of its language. Initially it was thought that the words or symbols in which humans could comprehend this knowledge were determined by the objects they referred to and thus were beyond human convention. Lull and Paracelsus, for example, hoped to discover the *true* signs or names of things (Hacking, c. 5 and p. 49). From the viewpoint of the development of symbolic science and mathematics the crucial turning-point was when the symbols representing 'signs' were made subject to convention and their arbitrariness was recognised.

Descartes, in the Rules for the Direction of the Mind, explains how simple 'figures' can be used to represent any sensory quality of objects "for it is certain that the infinitude of figures suffices to express all the differences in sensible things" (Descartes, Rules ... p.37). He suggests some conventions which these figures should conform to, and points out that they symbolically represent general magnitude and depending on the subject matter "these same figures have to represent for us now continuous magnitude, again a plurality of units or number also" (ibid., p. 65). Later in Rule XVI he extends the power of this symbolic mode of representation to numbers: "we ... abstract at this point from numbers themselves no less than from Geometrical figures or anything else, as we did a little time ago" (ibid., p. 67). "Everything , therefore, which is to be looked upon as single from the point of view of the solution of our problem, will be represented by a single symbol which", Descartes makes clear, "can be constructed in any way we please." (ibid. p. 65).

It is to be noted that Newton, like Descartes before him, confines the use of symbolic representation to the realm of the quantifiable:

> You see therefore, that to the Solution of Questions which only regard Numbers, or the abstracted Relations of Quantities, there is scarce anything else required than that the Problem be translated out of the English, or any other Tongue it is proposed in , into the Algebraical Language that is, into Characters fit to denote our Conceptions of the Relations of Quantities. (Newton, p.69)

The introduction of symbolism "whose use is not limited to numbers or magnitudes" was made by Leibnitz, who held the belief that 'all human knowledge can be expressed by the letters of the Alphabet". He conceived the possibility of a universal language and science Universal Specious or Characteristic — "that treats of the same and the diverse; of the similar and the dissimilar; of the absolute and the relational" just as contemporary mathematics dealt with quantity and its relations. Algebra is just 'one of a thousand' modes of expression that this universal science encompasses. Leibnitz explains its subordinate rôle by the fact that in the restricted case "when these letters or marks signify magnitudes or general numbers the result is the Algebra or rather the Specious of Viete". (Liebnitz, *The Horizon of Human Doctrine*, pp. 74–75)

The belief in the existence of a *Mathesis Universalis* encompassing all possible human knowledge was the third of the intellectual traditions that I wish to mention briefly as an influence on the emergence of the new mathematics. The nature of this Mathesis was subject to many diverse interpretations. The Pythagoreans identified its object as 'number', a position Leibnitz adopted at one time: "we can assign to every object its determined characteristic number" (Leibnitz, p. 18). Another tradition was based on the symbolic potential of 'figures'. This was the view of Bruno and, as was mentioned previously, within a totally different framework, and also that of Descartes. Descartes' Mathesis appears to reflect an influence of Bruno in several respects. Despite the profound divergence of their outlooks, Descartes' goal is not that dissimilar from that which Bruno assign ed to his mathesis, as Yates brings out in her summary of it

> by mathesis we learn how to abstract from matter, motion and time, and to reach intellectual contemplation of the intelligible species. (Yates, p. 272)

The nature of this universal science was seen, and this is most important, not as a technique or even as a set of techniques but as a general method of discipline; it was *The Art of Discovery* for which, as Leibnitz pointed out "nothing less is required than the founding of a mathematical-philosophical course of study according to a new method." (Leibnitz, p. 22)

Descartes' studies of the 'ancient' mathematics, in particular Pappus and Diophantus, confirmed his suspicion "that they had knowledge of a species of Mathematics very different from that which passes current in our time" (Descartes, Rules ... p. 12). He then gave up particular studies of arithmetic and geometry and proceeded to the "general investigation of Mathematics

(Mathesis), and thereupon sought to determine what precisely was universally meant by that term, and why not only the above mentioned sciences, but also Astronomy, Music, Optics, Mechanics and several others are styled parts of Mathematics ... I saw consequently that there must be some general science ... " (Descartes, Rules ... , p. 13). "Descartes' great idea", as Klein argues, was in

> identifying the 'general' object of this *mathesis universalis* which can be represented only symbolically — with the 'substance' of the world ... Only by virtue of this identification did symbolic mathematics gain that fundamental position in the system of knowledge which it has never since lost (Klein, p. 198).

The unity of the new symbolic mathematics was based, not as previously on its operation on 'quantity', but on symbolic operations performed on symbols whose properties could be investigated quite independently of any quantitative meaning that could be applied to them. This conception of mathematics, and thus also of knowledge, leads directly to the possibility of a single system of knowledge that can legitimately aim to comprehend the totality of natural phenomena. Modern science's symbolic form and its assumption as to the unity of natural science stem from the foundations of symbolic mathematics.

> identifying the 'general' object of this *mathesis universalis* which can be represented only symbolically — with the 'substance' of the world ... Only by virtue of this identification did symbolic mathematics gain that fundamental position in the system of knowledge which it has never since lost (Klein, p. 198).

APPENDIX : NOTES ON THE HISTORY OF MATHEMATICS

Klein, in his important pioneering study which aims at "recovering to some degree the sources ... of our modern symbolic mathematics", states that these sources are "today almost completely hidden from view" (Klein, 1968, p. 4). The reason for this he explains is to be found in "the conventional presentations of the history of this development" which

> do not, indeed fail to see the significance of the revival of Greek mathematics in the sixteenth century. But they always take for granted, and far too much as a matter of course, the fact of symbolic mathematics. They are not sufficiently aware of the character of the conceptual

> transformation which occurred in the course of this assimilation and which constitutes the indispensable condition of modern mathematical symbolism. Moreover, most of the standard histories attempt to grasp Greek mathematics itself with the aid of modern symbolism, as if the latter were an altogether exterior 'form' which may be tailored to any desired 'content'. (Klein, 1968, p. 5)

Given the claim of a 'hidden' nature to this transformation and its sources, it would seem appropriate to make some comments on the historical study of mathematics and the approach adopted here. A major (but neglected) task of the historian of mathematics in studying its development is the 'reconstruction' of former mathematical knowledge belief-systems. Somehow the change of 'world-view' that forms the basis of modern science is often seen as external and unrelated to the development of mathematics which proceeds according to its own 'inner logic', already contained in the mathematics of antiquity. From this point of view, occasionally scattered through history are the mathematical 'geniuses' who speed up its course along its 'historically inevitable' path.

Mathematics, like any culture or belief system, may contain very strong internal factors preserving its form and continuing its development, but it still remains that "mathematics as we know it ... has little inevitability about it" (Wilder, p. xix).

Thus Taton, in his discussion of the 'mathematical revolution' of the seventeenth century, argues that "the essential drive seems to have been internal to mathematics itself" (Taton, 1975, p. 287) and that, although there are "two turning points in the development of mathematics" (associated with Descartes and Newton and Leibnitz) "the movement must be seen as continuous rather than discrete" (ibid. p. 289).

Koyre's work normally shows great sensitivity to the nature of the conceptual transformation that was required for the emergence of modern science:

> Therefore what the founders of modern science ... had to do, was ... destroy one world and to replace it with another. They had to reshape the framework of our intellect itself, to restate and to reform its concepts, to evolve a new approach to 'Being', a new concept of knowledge, a new concept of science, and even to replace a pretty natural approach, that of common sense, by another which is not natural at all. (Koyre, 1968, p. 21)

Mathematics was not exempt from this revolution and underwent itself a transformation that can also be characterised by Koyre's description of the change in science. However, as regards mathematics, he often tends to see it as having something of a fixed and necessary meaning. He will assert, for example , that "modern science ... for its theoretical work ... adopts and develops the pattern of thinking of the mathematician" (Koyre, 1968, p. 90) as though it were in some way a pre-existing and external system of knowledge.

Unless mathematical cultures are understood within their historically specific forms, the question of the factors leading to change are unlikely to be seen as problematic. This is the approach of Klein's phenomenological history:

> The 'intentional history', as suggested by Husserl, may accomplish this task: it may 'reactivate' the 'sedimented' 'evidences', may bring to light the forgotten origins of our science. A history of science which fails to tackle this task does not live up to its own purpose, however valuable it otherwise may be. (Klein , 1940, p. 161)

He does not wish to proceed beyond this goal, however, to investigate the 'accidents of history', as he calls them, which explain the development of mathematics. It may well be that:

> Only strong environmental and internal pressures, such as sometimes imposed by war, dislocation forced by political changes, radical alterations in the host culture, 'crises' in mathematics itself, and the like, appear to be effective in changing the course of mathematical development. (Wilder, p. xii)

That such changes do occur is clear and it is precisely to factors of this type that we must turn if we are to recover the conceptual space from which symbolic mathematics emerged, and to trace its evolution.

The approach used in this essay is thus closely allied to those of Hacking in his study of the emergence of probability (Hacking), of Klein in his investigations on the origins of algebra (Klein, 1968A) and of Boyer's research into the conceptual development of the calculus (Boyer).

The emergence of symbolic mathematics has often been associated with factors essentially internal to it. As algebraic notation was introduced, it became easier to conceive of algebra as a mathematical discipline concerned with the logical and relational properties of entities whose nature was fundamentally symbolic. That the development of both mathematical notation and algebra were important factors in the emergence of symbolic mathematics is

indisputable and these shall be discussed in later sections. Where the viewpoint of this essay differs, however, is in regarding this symbolic character as being an essential component of algebra or notation *per se*. This is the view of Bochner:

> In mathematics, the sixteenth century was the century of the rise of so-called algebra ... Not only was this algebra a characteristic of the century, but a certain feature of it, namely the 'symbolization' inherent to it, became a profoundly distinguishing mark of all mathematics to follow. (Bochner, p. 38)

The 'algebra' that was introduced into Europe was almost without any mathematical notation; and in any case, however abbreviated a notation system might be it is not necessarily symbolic in function. That mathematical symbols are not intrinsically 'symbolic' might, at first, appear a paradoxical proposition. For although

> mathematical symbols are to be temporarily regarded as rigid and fix ed in meaning ... in reality they are continually changing and actually fluid. Indeed it is actually owing to this strange truth that progress in mathematical science is possible at all. (B. Branford, *A Study of Mathematical Education*, Oxford, 1906, quoted in Cajori, II)

These meanings are only to be found when mathematics is located within the context of the wider belief-system in which it is formed:

> Mathematics detached from its rich intellectual setting in the culture of our civilization and reduced to a series of techniques has been grossly distorted. (Kline, p. viii)

It is to the intellectual setting of the late renaissance within which symbolic modes of thought had diffused through most spheres of human activity that must be studied to understand the transformation to a symbolic mathematics.

REFERENCES

Bochner, S., *The Role of Mathematics in the Rise of Science*, Princeton 1966.

Boyer, C.B., *The History of the Calculus and its Conceptual Development*, New York, 1959.

Cajori, F.1 *A History of Mathematical Notation*, I, II, Chicago, 1928, 1929.

Cassirer, E., *An Essay on Man*, New York, 1944.

Cassirer, E., *The Individual and the Cosmos in Renaissance Philosophy*, Philadelphia, 1972.

Clulee, N.H., *John Dee's Mathematics and the Grading of Compounds*, Ambix XVIII, 1971 1 p. 178 .

Dee, J., *The Mathematical Preface of John Dee*, Ed. A. Debus, 1975.

Descartes, R., *The Philosophical Works of Descartes*, Tr. E. S. Haldane and G. R. T. Ross, I, Cambridge, 1973.

Hacking, I., *The Emergence of Probability*, Cambridge, 1975.

Hutten, E., *The Ideas of Physics*, London, 1967.

Klein, J., 'Phenomenology and the History of Science' in *Philosophical Essays in Memory of Edmund Husserl*, ed. M. Farber, Harvard U.P., 1940.

Klein, J., *Greek Mathematical Thought and the Origins of Algebra*, M.I.T.,1968.

Kline, M., *Mathematics in Western Culture*, Oxford, 1968.

Koyre, A., 'The Sixteenth Century: from Rhetorical to Syncopated Algebra' in *The Beginnings of Modern Science* ed. R. Taton, London, 1964.

Koyre, A., *Metaphysics and Measurement: Essays on Scientific Revolution*, London, 1968.

Leibnitz, G.W., *Leibnitz: Selections*, Ed. P. Wiener, New York 1951.

Michel, P., *The Cosmology of Giordano Bruno*, London, 1973.

Newton, I., *Universal Arithmetick*, London, 1720.

Recorde, R., *Whetstone of Witte*, 1557.

"ANPA-PROC-37-38" — 2018/4/6 — 22:32 — page 337 — #215

Stevin, S., *The Principal Works of Simon Stevin*
II A, B, Mathematics. Ed. D.J. Struik, 1958.
III, Astronomy nd Navigation. Ed. A. Pannekoek and E. Crone, 1961.

Strong, E.W., *Procedures and Metaphysics*, Hildesheim, 1966.

Tanner, R.C.H., 'Thomas Harriot as Mathematician: a Legacy of Hearsay', Physis IX, 1967.

Taton, R., 'The Mathematical Revolution of the Seventeenth Century' in *Reason, Experiment and Mysticism in the Scientific Revolution*
Ed. M.L. Bonelli and W. R. Shea, London, 1975.

Vieta, F., 'Introduction to the Analytic Art'
Tr. J,W. Smith in *Klein*, 1968.

Wilder, R. , *Evolution of Mathematical Concepts*,
London, 1974.

Yates, F., *Giordano Bruno and the Hermetic Tradition*,
New York, 1969.

AWARENESS LIES OUTSIDE TURING'S BOX

MICHAEL MANTHEY

tauquernions.org
manthey@acm.org

Abstract: The idea of computation as formulated by Alan Turing in the 1920's dominates the contemporary discussion of the mechanism(s) underlying awareness. Unfortunately, the very semantics of computation, as standardly defined, seem namely to *exclude* awareness. Furthermore, this paradigm's descriptions of Nature in general are - for all their utility - conceptually barren in their fundamental sequentiality. We therefore replace Turing's automata-based semantic model with a multi-dimensional vector algebra, namely W.K. Clifford's *geometric algebra*. In our novel *automata-free* re-framing, it is obvious that Turing's semantics is inherently *time-like*, and that geometric algebra's *space-like* semantics provide a fertile foundation for the phenomenon of awareness. The new computational model of distributed systems has the global mathematical form $\mathsf{U}(1) \times \mathsf{SU}(2) \times \mathsf{SU}(3) \times \mathsf{SO}(4)$, ie. the Standard Model of physics augmented with 3+1d.

1 INTRODUCTION

The advent of computers has given us the opportunity to study *processes* to unprecedented and - as I will show - unexpected depths. Alan Turing, who invented an eponymous abstract universal computer, showed that its *sequential* computations (ie. *processes*) possess great power, but also tantalizing limitations.

As an example of the latter, he proved that multiple "Turing machines", working together in parallel, even non-deterministically, have the same ultimate computational power as a single machine working alone! This result - sequence is sufficient - has profound and far-reaching implications: it enshrines a functional point-of-view, it blesses strong reductionism, it is fundamentally classical (Isaac & Albert) in its reach, and it has stubbornly resisted

attempts to convincingly describe the *multi*-process systems that comprise Physics, Biology, and the study of Mind.

Rather, the resulting models seem always to be imitative, often strained, and conceptually thin. The "artificial intelligence" model offered by so-called neural nets, powered by Bayesian statistics, is an excellent example: all the hype (and money) cannot disguise how far it strays from the the real thing - it needs *tens of thousands* of examples to learn a non-trivial category, it is unable to construct non-trivial causal sequences ("planning"), and it clearly does not understand anything that it says or does. Worst of all, it offers no apparent explanation of, or path to, the property of self-awareness. In other words, we get Zombies ... the lights are on but there's no one there. Thus have matters stood for some time, despite decades of effort.

The present claim is that the way out of this impasse is to replace the straitjacket of automata and sequentiality with the power of multi-dimensional vector algebra. In particular, choose the *geometric algebra* $\mathcal{G}$ invented by W.K. Clifford in 1860, but [unusually] operating over the arithmetic of $\mathbb{Z}_3 = \{0, 1, -1\}$, that is, the usual base $3 = \{0, 1, 2\}$ shifted one slot to the left, whence $1 + 1 = -1$ and zero = *Void*, whence the underlying logic is that of exclusive-or, *xor*: same *vs*. different. The result is a general, *automata-free semantics* of all possible computations.

In the following, we will first show how the basic conditional act of *if-then-else* appears in this new algebraic language, which form reveals a close relationship with the quantum mechanical concept of *measurement*. Next comes the consideration of the *synchronization* of processes, which is necessary to both capture resource ownership and producer-consumer relationships, and to enforce determinism where needed. The synchronization primitives *wait(event)* and *signal(event)* also capture the concepts of memory and causality [4].

These two quintessential sequential functionalities - *if-then-else* and *wait-/signal* - turn out to consume a *trivial* fraction of geometric algebra's semantic power. In particular, we will show how geometric algebra easily exceeds Turing's formulations and limitations via its ability to express the *indistinguishability* of events. A cornucopia of insight then reveals itself.

What we find is not of Turing's time-like sequential semantics at all, but rather of its opposite, *space-like* computation.

Here is a helpful conceptual analogy for what is about to be described: consider the distinction between the Operating System (OS) of a computer -

nearly invisible in systems like smart phones, but more apparent on larger systems - *and* the "user processes" nowadays called "apps", plus resources like memory and Input/Output. The OS is the *manager*, ensuring that eg. a given screen window is protected from meddling by other processes, or that a given chunk of memory that was initially allocated to process A is now also visible to process B, so they can communicate.

Then there are the user processes A, B, ... themselves. The OS swaps them in and out of memory when they're idle, puts them on queues for needed resources, finds their files, etc. The OS has *no* idea what these processes are about, and oppositely, the processes have *no* inkling of the existence of the OS. The OS's job is to be aware of the *fact* that multiple processes are *simultaneously present* and competing for access to various resources - re-usable resources like memory and devices, and consumable resources like input data streams - to make sure that processes can proceed and do not step on each others' toes. *And*, most importantly, to be completely invisible to the user processes while so doing.

The promised helpful conceptual analogy is then this: the OS corresponds to quantum mechanics, and the user process and resource spaces correspond to the structure of relativistic 3+1d space. Indeed, one could imagine two apps, Experimenter and Observer, who cooperate in trying to figure out what's going on behind the OS's facade. This analogy, while *not!* saying the universe is a big operating system, is nevertheless deeply apt: W^{m} James' "thinnest of veils" refers to this same distinction, as does the term *axis mundi* [the axis of the world], the place where mind intersects matter.

2 IF-THEN-ELSE

A sequential program, unrolled into its future, forms a system consisting of a single process, namely itself - there is no talk of other processes: even if they're present, any synchronization is transparent, and any interference oblique and unrecognized. The single most important property of a process is that it is a *sequence*: the order in which its events take place is crucial, *defining* in effect what the process does. As will be seen, it is similarly crucial not to confuse the three concepts of ordering/sequence, determinism, and causality, as was done in the early years of quantum mechanics.

Suppose now that X, Y, Z are arbitrary expressions in the algebra, and consider the process XYZ, which states the process "do Z, then do Y, then do X", that is, we *always* operate multiplicatively on the left. If any of X, Y, or Z has an inverse, we could algebraically manipulate the product XYZ to produce some other order. This will not do!

Consider therefore the common view of a computer program - when it is executing - as a sequence of discrete operations

()()()()()()()()()()()()()

where each parenthesis-pair stands for a single X, Y, Z-type operation. Such a sequence is called a *process,* and the following is a gloss on [1], to which the reader is referred for a more detailed exposition. This process-level of computational description refers not so much to entities themselves as to their interaction, and the sequence of states this produces. In our model, *everything* is a process, or an object built out of processes.

As just noted, the key property of a process is the exact *order* in which its component operations take place. To capture this ordering property algebraically we will require that each operation "( )" in the above sequence - now viewed as a product - be irreversible (ie. no multiplicative inverse). This prevents algebraic manipulations from changing the effective order.

In physical terms, this means that processes are irreversible and *time-like*, and we will intend these two terms interchangeably, as well as their opposites: possessing an inverse = reversible (which allows wave-like activity) = *space-like*. [This is *not* physical 3-D space, just space-like rotations.]

Thus the generic sequential program DoX; DoY; DoZ translates to the $\mathcal{G}$ product sequence *(-1+DoZ)(-1+DoY)(-1+DoX)*, where it is assumed that *Do?* is unitary, $(Do?)^2 = 1$, whence *(-1+Do?)* is idempotent. A theorem of $\mathcal{G}$ informs us that an expression $E \in \mathcal{G}$ is irreversible *iff* E contains an idempotent as a factor.

We now use this algebraic representation of computation to analyze the if-then-else construction. We will write ***if*** V ***then*** X ***else*** Y, where V, X, Y are arbitrary expressions representing arbitrary computations. For simplicity and with no loss of generality, take $V = a$, a 1-vector ("sensor").

"***if*** a" implies a probing of the current state of a: is it +1 (so do X), or is it -1 (so do Y).

Given that the only relevant states of a are ± 1, the next question is how to ascertain which of these obtains? Clearly, said ascertaining requires *measuring* a, where again idempotent operators play the central role. Consider the following identities:[1]

$$(1+a)=(1+a)(a) \qquad (-1+a)=(-1+a)(-a) \qquad (-1+a)=(1-a)(1-a)$$
$$(1-a)=(1-a)(-a) \qquad (-1-a)=(-1-a)(a) \qquad (-1-a)=(1+a)(1+a)$$

Taking $P = (1+a) = (1+a)(a)$ as an example, multiply P's rhs out to get $a + aa$, whence we see that the +1 in the lhs can be seen as the product of a with itself.[2] It follows, and this is the key point, that if the a we have in hand - in the rhs's "$(1+a)$" factor - has the same sign as the a we probe - the rhs's "(a)" factor - then the sign of the scalar will be +1, whereas if the a we probe is actually $-a$, then the sign of the scalar will be -1. This also applies if P, oppositely, specifies "$(-a)$" and we find "$-a$" ($\Rightarrow$ +1), or we find "+a" ($\Rightarrow$ -1). Finally, take *just* the scalar value from $(-1 \pm a)(\pm a)$ to complete the measurement (one can only *actually* measure scalars ... like a meter reading).[3]

This is the basic act of measurement. Because $(1+a)$ has no inverse, the act of measurement is irreversible, in accordance with contemporary understanding of the equivalence of energy and (Shannon) information. Furthermore, successive measurements using the idempotent form yield no new information, as required, in that $PP = P$. [4]

So now we know how to do "***if*** a": we will write $(1+a)(a)$ or suchlike, depending. The next issue is to choose the correct continuation depending on what the measurement on a produces.

The basic idea now is to arrange for the conjugate forms $(1+a)$ and $(1-a)$, whose product is zero, to collide on the unwanted branch of the ***if***, thus eliminating that continuation. A zero means the computation's future is empty, ie. it does not occur; generating a zero to eliminate an unwanted continuation is a key tool in the following.

1 All algebraic calculations have been done with a purpose-built Z_3 calculator written by Douglas Matzke. Those wishing to verify our calculations on other symbolic algebra platforms are warned that floating point (but supposedly "really integer") arithmetic can easily produce errors.

2 If the lhs scalar were -1, a would be inverted to $-a$ as well.

3 Ie. the scalar or "dot" product of $(-1+a)$ and a, ie. $(1+a)\cdot a$ To avoid clutter, we will suppress the 'dot' til the end of the derivation.

4 Actually, $(1+a)$ is the square root ("sqert") of an idempotent, cf. column 3 above, but this is unimportant for our present purposes.

Therefore, write the test in the ***if*** as a probe: $1+a$ or $1-a$, acting on the actual a, which can be plus or minus. The ***then*** and ***else*** branches apply respectively $1+a$ or $1-a$ to the result of the test, whence one of them should yield 0 (because conjugate) and the other the correct continuation based on the observed value of a. There are four possibilities (the $|$ marks off visually (only) the shared ***if***-probe, rightmost because it occurs first): [5]

if	*probe*	***then***	*left branch* $\mid$ *probe*	***else***	*right branch* $\mid$ *probe*
1	$(1+a)(+a)$		$X(1+a)\,\vert\,(1+a)(a)$		$Y(1-a)\,\vert\,(1+a)(a)$
			$=-X(1+a)$ ***yes***		$=0$ ***yes***
2	$(1+a)(-a)$		$X(1+a)\,\vert\,(1+a)(-a)$		$Y(1-a)\,\vert\,(1+a)(-a)$
			$=X(1+a)$ ***no***		$=0$ ***yes***
3	$(1-a)(+a)$		$X(1+a)\,\vert\,(1-a)(a)$		$Y(1-a)\,\vert\,(1-a)(a)$
			$=0$ ***yes***		$=Y(1-a)$ ***no***
4	$(1-a)(-a)$		$X(1+a)\,\vert\,(1-a)(-a)$		$Y(1-a)\,\vert\,(1-a)(-a)$
			$=0$ ***yes***		$=-Y(1-a)$ ***yes***

In situation 1 above, we probe for $+a$ with $(1+a)$, and a is in fact $+a$; situation 2 has the same probe, but discovers $-a$; situation 3 probes for $-a$ but discovers $+a$; and situation 4 probes for $-a$ and discovers $-a$. Notice that if we consider all four possibilities concurrently (ie. Left + Right, 1 thru 4), we get zero: this situation (namely, a having both values simultaneously) cannot occur. So instead, combine 1&2 and 3&4 by subtraction to get the desired terms to double instead of cancel: 1-2 = $+X(1+a)$; 4-3 = $+Y(1-a)$, and move the $|$-cue to the right, eliminating the common probe-preface of the previous version:

1 minus 2: $\quad -X(1+a)\,\vert\,(\pm a)$

4 minus 3: $\quad -Y(1-a)\,\vert\,(\pm a)$

Finally, run 1-2 and 4-3 concurrently (ie. add), and factor out $(\pm a)$:

$$-X(1+a)\,\vert\,(\pm a) \;-\; Y(1-a)\,\vert\,(\pm a)$$
$$=[-X(1+a)-Y(1-a)]\cdot(\pm a)$$

If $a=+1$ then the Y term drops out leaving $+X$; and if $a=-1$ then the X term drops out, leaving $+Y$. Just as we wanted! Push the minus-signs on X, Y into the parentheses:

5 The ***yes*** and ***no*** indicate desired (or not) outcomes.

$$= [X(-1-a)+Y(-1+a)]\cdot(\pm a)$$

and we see that doing *if-then-else* necessarily invokes observation, ie. idempotents, consistent with thermodynamic and quantum measurement theory. The form also makes good computational sense when multiplied out:

$$= X(-1-a)\cdot(\pm a)+Y(-1+a)\cdot(\pm a)$$

which transparently describes two independent processes X and Y, each independently and concurrently testing for its own condition, only one of which will succeed.

NB: if one tries simultaneously to measure with $1+a$ *and* $1-a$, one gets (summing) an inversion (1+1 = -1), but no knowledge of a, in accordance with quantum measurement theory: if one is to get information, one must specify *exactly* what it is one is looking for ... +a or -a, and this *cannot* be finessed.

3 WAIT AND SIGNAL

Having warmed up with *if-then-else*, we now tackle synchronization's *wait* and *signal*.

A primitive synchronizer T consists of a notional internal binary flag - *Open* or *Closed* - that can be changed by two operations: *wait* and *signal*, denoted hereafter by *W* and S. The restriction to binary behavior implies no loss of generality. A synchronizer must supply the following behavior:

	a. A *signal* sets T to Open, and passes the *signalling* process;
S_{out}	*b.* Successive *Ssignals* are the same as a single *signal*;
↑	*c.* A *Wait* on *Closed* T fails, ie. the *Wait*er is not passed thru;
$W_{in}\rightarrow$ T $\rightarrow W_{out}$	*d.* A *Wait* on *Open* T sets T to *Closed*, and passes the *Waiter*;
↑	*e.* Simultaneous *Waits* on the same T $\rightsquigarrow$ max 1 *Waiter* passes;
S_{in}	*f.* Simultaneous *signals* on the same T = a single *signal*.

In the above diagram, *waits* enter from the left and exit to the right; similarly, *signals* enter from the bottom and exit at the top. The exclusion of processes over (say) a printer is realized by placing the use of the printer on the W_{out} leg, and thereafter directing the process to perform a corresponding S_{in} before exiting entirely; this arrangement guarantees that processes will

use the printer serially (otherwise, output from different processes would be meaninglessly interleaved on the paper record, which is why synchronization is necessary in the first place). More complex examples can be found in any good operating system textbook.

Implicit in such arrangements is the requirement that synchronization be *transparent* to the participating processes: it would be unacceptable for the correct operation of a program to be dependent on whether it "really" waited to acquire some resource because some other process(es) happened to be present. Hence, no information in the Shannon sense is conveyed between two processes via the act of synchronization. Rather, synchronization induces/enforces a phase shift at the inter-process level. This phase shift is expressed in the non-deterministic ordering of the processes as they pass through the synchronizer.

Items *e* and *f* refer to situations where there is competition between multiple *Waiters* and/or *Signallers*; see [1].

The first step comes from item *b*, which in effect says $SS = S$, ie. S must be *idempotent*.

Item *d* says that *W*T must succeed if T is *Open*. Therefore initialize T to *Open*, which we can do via item *a* by setting $T = S$. Item *d* then reads $WT = WS$, which must be non-zero to succeed.[6]

Item *c* in effect says (together with item *a*) that successive *Waits* without an intervening *signal* must fail. That is, $WW = 0$, so *W* must be nilpotent. So now we know the shapes of both *W* and S, and very specific ones at that.

Thus a sequence like $SSWSWSST = SWSWST = SWSWS$, and any sequence with consecutive *W*'s yields zero, eg. $WWSWST = 0$.

Process-wise (see figure just below), there is process P_1, which after a sequence of arbitrary irreversible operations X issues the *signal* S, creating a so-called 'synchronization token'; and then there is process P_2 which after a sequence of *Y*'s consumes this token by waiting on it, whereafter P_2 continues, executing Z's (read right-to-left: things begin on the right!):

P_1 : ...X X X S X X X...↩

↓ X, Y, Z are arbitrary irreversible actions

P_2 : ...Z Z Z W Y Y Y...↩

6 Initializing T to *W* (ie. T is initially *Closed*) doesn't work: $WT = WW = 0$, whence SWT also yields zero, which it shouldn't. Initializing T to 1 (which is idempotent) is indiscriminate - *any W* will succeed.

Despite the visually implicit timeline in the above two sequences, the *Wait* in P_2 can occur any time 'before', 'simultaneously with', or 'after' the *signal* in P_1, but unless the *wait* occurs 'after' the *signal*, process P_2 is logically halted at the *W*. Whichever of these circumstances obtains, the ultimate result is a logically and physically seamless transition from P_1's SXXX to P_2's ZZZW. *This sequence too is a process*, process P_3:

P_3 : ...ZZZWSXXX...

The fact that *W must* be nilpotent means that 'whenever' the *WS* mating actually occurs, it is just as though P_3 occurred seamlessly. An example: when one absorbs a photon in the retina, at that very instant one is exactly connected with the state that generated the S - even if the star that generated the photon has 'long since' disappeared. [7]

P_1 and P_2 are *classical*, in that we imagine them to be deterministic - good old-fashioned Newtonian / Einsteinian processes. [We might think of the state preparations preceding an actual quantum experiment, which are classical.] P_3, on the other hand, is non-deterministic, because it was precisely >> P_2<< 's *wait* that succeeded, leading to the Z's. *If* however it had happened that some P_4's *wait* occurred ahead of P_2's, P_3's continuation would be entirely different.

This *emergent* non-determinism is old news in computer science, though it is most often noted in the form of unwanted *values* (cf. the interleaved printer output example), rather than the entirely proper non-deterministic *ordering* induced by the serialization as just described.[8] In both cases - *order* or *value* non-determinism - the root is the *asynchrony* of the interaction of two independent processes. Said a bit differently, *if* one is to use *process* as a conceptual primitive, *then* one necesarily must accept into the bargain the consequent, unavoidable emergent non-determinism born of the asynchronous interaction of these same processes.[9] Both non-deterministic values and non-deterministic order are produced by asynchrony. This asynchrony is the very source of QM's non-determinism, and the next Section elucidates its source.

Order-non-determinism forms the *coarse-grained* skeleton of physical non-determinism. Suppose now that one has guaranteed that only a *particular*

7 It's pretty limited time travel tho - you only get the single bit of information that the photon carries ... not much of a view!

8 Both are the source of the most difficult bugs, because they are namely not repeatable; cf. Ullman's fine novel, "The Bug".

9 It is the *necessity* for exclusion, at *every* step, that dictates that processes be discrete, cf. Planck's constant.

Wait-continuation will match a given *signal*, so *order* is out of the picture. One still doesn't know what one will get from the measurement, cf. *if-then-else*'s measurement earlier. So within the *order*-skeleton is a second, finer-grained source of non-determinism, *value* non-determinism, induced by the measurements encapsulated in the Signals. For example, the idempotent $-1 + xy + xz$ expresses a value-changing intrusion into the entity $xy + xz$, which in principle "lives its own (reversible) life" both prior to and subsequent to the measurement.

Popping up conceptually, imagine now P_3's form as it evolves into its future. Its sequence of Z's is just shorthand for an arbitrary sequence of idempotents, for example $(1 + a)(1 + b)...(1 + r)$. Being idempotents, each of them can act as a *signal* to some matching *Wait* 'out there'. [It is important that they be idempotents, because this means that the event that the *Wait* is dependent on has actually physically occurred.] Ultimately, if every idempotent in P_3 triggers a *Wait*, and all those *Waits*' continuations do the same, the universe will be populated entirely by utterly non-deterministic processes that look like $(WS)(WS)(WS)...(WS)$ - these W's and S's being notionally distinct. In fact, we see that our classical view of P_1 and P_2 as deterministic processes puts them in an improbable and miniscule minority - namely that minority inhabiting/defining classical 3+1 space-time, *plus* all ordinary sequential computer programs ... which (despite appearances) includes the Internet.

Putting all this together, a sequential process - aka. a measurement sequence - looks like

$$(-1 + X_n)(-1 + X_{n-1})...(-1 + X_1) = \prod_n \hat{X}_i,$$

$$X_i^2 = 1$$

This is probably all more or less familiar to physicists. But the *computational* reading of the algebra takes the correspondence much further. In this reading [1], the idempotent form $-1 + X$ is identified as the primitive synchronization operation signal(X), understood to mean "*signal* the occurrence of the event/state X".

Signal's complementary primitive is wait(X), ie. wait for the occurrence (signal) of event X. It is critical to understand that this waiting is not *polling*, ie. that the waiting process is constantly and actively checking to see if X has occurred yet, aka. *busy waiting*. Busy-waiting turns out to be a quite untenable view in an asynchronously concurrent universe because everyone spends most of their "time" polling each other, so something subtler is necessary. A

careful analysis [1] reveals that the computational concept of $\mathsf{wait}(X)$ must be mapped, speaking now algebraically, to some nilpotent $\omega \in \mathcal{G}$, $\omega^2 = 0$.

Furthermore, quoting [1], "No information (in the strict Shannon sense) is conveyed between two processes via the act of synchronization. Rather, synchronization induces/enforces a phase shift at the inter-process level. This phase shift is expressed in the non-deterministic ordering of the processes as they pass through the synchronizer." So synchronization *per se* is memoryless - exactly what happened, exactly what will happen, is spread out in the phase structure of the entire concurrent computation. It is also, indirectly, the means by which non-determinism enters into our model.

In physics, nilpotents supply the causal - and energy conserving - connection between discrete physical events. Wait's play the corresponding role in the synchronizational context - causal connection and conserving information between computational events. Nilpotents are also irreversible, and since any irreversible $E \in \mathcal{G}$ *must* have an idempotent factor, we can reverse the theorem and derive our ω's from our idempotents.

We can derive ω's form by considering two consecutive events $\hat{U};\hat{V}$, forming the process $\hat{V}\hat{U}$. We will insist, now speaking computationally, that $\hat{V}$ *never* occur before $\hat{U}$, ie. the actual process must specify that $\hat{V}$ must always *wait* (ω) for $\hat{U}$. That is, we want $\hat{V}\hat{U} = \hat{V}\omega\hat{U}$. Rewriting the lhs as $\hat{V}\hat{U} = \hat{V}\hat{V}\hat{U}$ and expanding,

$$(-1+V)(-1+U) = (-1+V)(-1+V)(-U)(-1+U)$$

we find that $\hat{V}\hat{U} = \hat{V}(U - VU)\hat{U}$, and indeed $\omega = U + UV$ is nilpotent so long as U and V anti-commute.[10] Computationally speaking, anti-commutativity means "independent of each other", as in the practice of orthogonal software design, which focuses on ensuring that changes to one module do not affect another; or as in "asynchronously concurrent"; or both, as here.

Processes like $\hat{V}\hat{U}$ are exactly the processes covered by Turing's model of computation, and since entities like $\hat{U}$,$\hat{V}$ are the *projectors* of U, V respectively (so-called *measurement operators*), they are also the observational bedrock of quantum mechanics. The key property of such processes - irreversible sequentiality - makes them purely *time-like* processes. It is ultimately this time-like

10 Note that we could instead have written $\hat{V}\hat{U} = \hat{V}\hat{U}\hat{U}$, which leads to $\omega = -V - UV$. This corresponds to the so-called advanced solution, whereas $\hat{V}\hat{V}\hat{U}$ corresponds to the retarded solution.

property that allows Penrose to conclude [2] that computational processes cannot capture all the phenomena that quantum mechanics has to offer, among which is entanglement, which is fundamentally space-like.

Summing up, we have seen that Turing's sequential model of computation maps to *products* of idempotents (*signals*) and nilpotents (*waits*) in geometric algebra. Products are inherently sequential structures[11] so the fit between Turing's view and that of algebraic products is exact and tight. That leaves sums ... Sums express true concurrency! Since true concurrency lies outside of any sequential theory of computation, this marks the escape path out of Turing's Box. We must ourselves erect a theory for it, and *the resulting general theory of distributed systems turns out to be a computational theory of awareness.*

4 THE COIN DEMONSTRATION

By *awareness* I mean the *experience* that one is a coherent entity, even whilst in profound appreciation and contact with one's surround. *Self*-awareness is generally larger and mostly unconscious relative to our normal waking (ie. ego-) consciousness. Awareness, *sans* the unitary feeling, is said to accompany certain very deep meditative states. The mechanisms presented in the following allow all of these. [12]

From a "systems" point of view, awareness, being un-localized, looks like some kind of *distributed* computation. A distributed computation consists of many more or less independent processes that *together*, with little to no centralized control, nevertheless produce globally coherent behavior. Examples abound in Nature, from beehives and anthills to ecologies, and from molecules and crystals to the quark structure of protons. Other favorites are the schooling behavior of fish and flocks of birds turning *en masse*.

11 This includes parenthesis trees, which translate directly to a partial order semantics, which "interleave" separately-sequential process events. Thus the potentially concurrent $f(g + h)$ is mapped to $f(g, h)$, where f, g, h are sequential processes. This approach, which successfully maintains Turing's truths to practical effect, is endemic in the formal theory of concurrent computational systems.

12 Note that "awareness of awareness" has two interpretations, depending on whether the second awareness is internal or external. Both interpretations are valid, and the *Topsy Test for awareness* [3] requires both. Note also that this classification of "awareness" and "consciousness", while workable, is *very* crude compared to eg. Tibetan and Vedantic observations, whose often flowery language is actually *very* precise, but also often defines differing schools of interpretation.

However, the difference between these systems and *awareness* is that awareness is not material - it has no *substance* - and yet it nevertheless seems to possess agency, even though its coherence is ineffable.

The only general purpose concept (that I can think of) that matches this description is a *wave*. A wave, to be a wave, is an extended affair. I like to say that, so to speak, a wave is *everywhere*. The flip side of the wave concept is that, even though it is everywhere, it is also - simultaneously - *nowhere in particular*. In a system that works like a wave, "nowhere in particular" translates to the myriad local micro-changes that *together* make up the wave, just as H_2O molecules' motions (mostly vertical) make up water waves. Awareness *per se* can then inhere in a multi-dimensional wave-like spectrum (is my claim). So, so far so good: *awareness* is wave-like.

Mathematically, to be in a world of waves is to be in the world of Joseph Fourier, who in 1805 proved that (very nearly) *any* function can be *exactly* replaced by a suitable sum of sines and cosines. This was an astounding discovery, and even though it capped several decades of general interest in doing such a thing, his result nevertheless attracted much controversy in its day. Today, it is a ubiquitous - because enormously useful - piece of mathematical and technological furniture.

More to our purpose, however, is the closely related *Parseval's Identity* of 1799, which states that the projection of a function $\mathcal{F}$ onto an n-dimensional orthogonal space *is* the Fourier decomposition of $\mathcal{F}$. Parseval's Identity is a generalization of the Pythagorean theorem to n dimensions. In the n-dimensional coordinate system, $\mathcal{F}$'s current value corresponds to [the tip of] a hyper-hypoteneuse in an n-dimensional hyper-volume, and the projection breaks that hyper-hypoteneuse down into the various pieces along each of the dimensions that go into its construction.

To construct an n-dimensional volume, begin with an ordinary plane right triangle with sides a and b. Reflect this triangle on its hypoteneuse, forming a rectangle with sides a and b, area ab, and diagonal $d = \sqrt{a^2 + b^2}$. Next, lift this rectangle c units vertically to make a rectangular volume abc. Its diagonal is $d = \sqrt{a^2 + b^2 + c^2}$ and this sum-of-squares symmetry continues as we make a 4d cuboid, then 5d, etc.

At the same time, going back to the starting right triangle, we can also express the sides a and b as $a = \cos\theta$ and $b = \sin\theta$, where θ is the angle between a and the hypoteneuse. And now all becomes clear: substituting these sine and cosine equivalents for $a, b, c, \ldots$ up through the dimensions will yield, for the n-dimensional hypoteneuse (= the current value of the function

$\mathcal{F}$, whose projection we began with), a big sum of ... sines and cosines, ie. Fourier's world.

So the world of waves and the world of orthogonal coordinate systems are the *same* world. It is in the latter that we will connect to computation. The connection is this: let each dimension correspond to the state of some process, where all these processes $a, b, c, \ldots, ab, ac \ldots, abc, \ldots$ are notionally independent (think *orthogonal*), though interacting otherwise freely and concurrently; these will be interpreted as elements of a graded vector algebra. [13]

Looking at the ongoing Heracletian flurry of process-state evolution in such a system, the high frequency Fourier bands correspond to short-term, fine-grained details, and low frequency bands to long-term symmetries and global developments. These cross-summed Fourier bands constitute the contextual world of *qualia* - the *feeling* of (eg.) redness *vs*. the optical frequencies detected by individual retinal cells.

And so we see that "distributed" behavior - ie. processes $a, b, c, \ldots, ab, ac \ldots, abc, \ldots$ all running (quasi-)independently - corresponds to wave-like behavior. In the physical world, various constraints (eg. conservation laws, entropy) rule out certain process behaviors as impossible or meaningless, and give form to the free-for-all that is the remainder. In addition, as the system self-organizes, it will - if it has sufficient complexity - learn ways to make itself transparent or reflective to those waveforms that are harmful to it; and complementarily, ways to absorb information and to promote its own further existence via energy-consuming reaction (in this connection, see [6]).

The result is the regularities - short, medium, and long term oscillations - that we, and *any* awareness, will (indeed, must) experience. As a corollary, it is very likely that awareness is not possible if the surround is too unstable [2,3]. This is often seen in visualizations of chaotic systems, where there will be a stable oscillatory behavior for a while, which then suddenly disappears, to be replaced by state transitions with no apparent pattern at all.

We see also that *awareness*, being a wave, is an emergent, collective phenomenon, with nothing scientifically mysterious about it. The mystery is in the experience of it. The following ***Coin Demonstration*** clarifies.

13 In the geometric (Clifford) algebra over $\mathbb{Z}_3 = \{0, 1, -1\}$ that I use, 1-vectors like a, b are processes with one bit of state, ± 1, whence an m-vector has m bits of state. For concurrent processes a, b write $a + b$; when a, b interact write ab; ab too is a process with external appearance ± 1 (aka. *spin*). And so on. NB: $ab = -ba \cong \sqrt{-1}$.

Act I. *A man stands in front of you with both hands behind his back. He shows you one hand containing a coin, and then returns the hand and the coin behind his back. After a brief pause, he again shows you the same hand with what appears to be an identical coin. He again hides it, and then asks, "How many coins do I have?"*

Understand first that this is not a trick question, nor some clever play on words - we are simply describing a particular and straightforward situation. The best answer at this point then is that the man has "at least one coin", which implicitly seeks *one bit* of information: two possible but mutually exclusive states: *state1* = "one coin", and *state2* = "more than one coin".

One is now at a decision point - *if* one coin *then* doX *else* doY - and exactly one bit of information can resolve the situation. Said differently, when one is able to make this decision, one has *ipso facto* received one bit of information.

Act II. *The man now extends his hand and it contains two identical coins.*

Stipulating that the two coins are in every relevant respect identical to the coins we saw earlier, we now know that there are *two* coins, that is, *we have received one bit of information*, in that the ambiguity is resolved. We have now arrived at the demonstration's dramatic peak:

Act III. *The man asks, "Where did that bit of information come from?"*

Indeed, where *did* it come from?! [14]

The bit originates in the *simultaneous presence* of the two coins - their ***co-occurrence*** - and encodes the now-observed *fact* that the two *processes,* whose states are the two coins, respectively, do not exclude each other's existence when in said states. [15]

Thus, there is information in (and about) the environment that *cannot* be acquired sequentially, and true concurrency therefore *cannot* be simulated by a Turing machine. Can a given state of process $\mathfrak{a}$ exist simultaneously with a given state of process $\mathfrak{b}$, or do they exclude each other's existence? In concurrent systems, *this* is the fundamental distinction.

More formally, we can by definition write $\mathfrak{a} + \tilde{\mathfrak{a}} = 0$ and $\mathfrak{b} + \tilde{\mathfrak{b}} = 0$ [~ = *not* = *minus*] meaning that (process state) $\mathfrak{a}$ excludes (process state) $\tilde{\mathfrak{a}}$,

14 [Think about it! Where *did* that bit come from? Thin air?]

15 Cf. Leibniz's indistinguishables, and their being the germ of the concept of space: simultaneous states, like the presence of the two coins, are namely indistinguishable in time. Co-occurrences are *bosons* in physics-speak (whence *sequential* computation is fermionic).

and similarly (process state) b excludes (process state) $\tilde{b}$. [16] Their *concurrent* existence can be captured by adding these two equations, and associativity gives two ways to view the result. The first is

$$(a + \tilde{b}) + (\tilde{a} + b) = 0$$

which is the usual excluded middle: if it's not the one (eg. that's +) then it's the other. This arrangement is convenient to our usual way of thinking, and easily encodes the traditional *one/zero* (or $1/\tilde{1}$) distinction. [17] The second view is

$$(a + b) + (\tilde{a} + \tilde{b}) = 0$$

which are the two superposition states: either both or neither.

The Coin Demonstration shows that *by its very existence*, a 2-co-occurrence like $a + b$ contains one bit of information. Co-occurrence relationships are *structural*, ie. *space-like*, by their very nature. This *space*-like information (vs. Shannon's *time*-like information) ultimately forms the structure and content of the Fourier bands, eg. the set {all 2-vectors}. See [4] for the mathematics.

Sets of m-vectors - $\{xy\}, \{xyz\}, \{wxyz\}, \ldots$ - are successively lower *under*tones of the concurrent flux at the sensory boundary $x + y + z + \ldots$, and constitute a simultaneous structural *and* functional decomposition of that flux into a graded hierarchy of stable and meta-stable processes. The lower the frequency, the longer-term its influence. [18]

16 This is the logical bottom, and so there are no superpositions of $a/\tilde{a}$ and $b/\tilde{b}$: they are 1d exclusionary distinctions . Superposition first emerges at level 2 with ab via the distinction *exclude* vs. *co-occur*.

17 Since $\tilde{x}$ is not the same as $0x$, an occurrence $\tilde{x}$ is meaningful; in terms of sensors, $x/\tilde{x}$ is a *sensing* of an externality of x, not x itself. That $\tilde{x} \neq 0x$ was also Dirac's innovation, leading to the concept of anti-matter.

18 Christopher T. Kello , Brandon C. Beltz , John G. Holden , Guy C. Van Orden: *The Emergent Coordination of Cognitive Function (2007)*. "Abstract: $1/f$ scaling has been observed throughout human physiology and behavior, but its origins and meaning remain a matter of debate. Some argue that it is a byproduct of ongoing processes in the brain or body and therefore of limited relevance to psychological theory. Others argue that $1/f$ scaling reflects a fundamental aspect of all physiological and cognitive functions, namely, that they emerge in the balance of independent versus interdependent component activities. In 4 experiments, series of key-press responses were used to test between these 2 alternative explanations. The critical design feature was to take 2 measures of each key-press response: reaction time and key-contact duration. These measures resulted in 2 parallel series of intrinsic fluctuations for each series of key-press responses. Intrinsic fluctuations exhibited $1/f$ scaling in both reaction times and key-contact dura-

But where do these m-vectors come from?

Act IV. *The man holds both hands out in front of him. One hand is empty, but there is a coin in the other. He closes his hands and puts them behind his back. Then he holds them out again, and we see that the coin has changed hands. He asks, "Did anything happen?"*

This is a rather harder question to answer. [19] To the above two *co*ncurrent *exclusion*ary processes we now apply the *co-exclusion inference*, whose opening statement is: a excludes $\tilde{a}$, *and* b excludes $\tilde{b}$, *whence* $a + \tilde{b}$ excludes $\tilde{a} + b$ and conjugately, $a + b$ excludes $\tilde{a} + \tilde{b}$. This we have just derived.

The inference's conclusion is: *Therefore,* ab *exists*. The reasoning is that we can logically replace the two *one*-bit-of-state processes a, b with one *two*-bits-of-state process ab, since what counts in processes is sequentiality, not state size, and exclusion births sequence (here, in the form of *alternation* between the two complementary states). That is, the existence of the two co-exclusions $(a + \tilde{b}) \mid (\tilde{a} + b)$ and $(a + b) \mid (\tilde{a} + \tilde{b})$ contains sufficient information for ab to be able to encode them, and therefore, logically and computationally speaking, ab can rightfully be instantiated.

We write $\delta(a + \tilde{b}) = ab = -\delta(\tilde{a} + b)$ and $\delta(a + b) = ab = -\delta(\tilde{a} + \tilde{b})$, where δ is a co-boundary operator (analogous to integration in calculus); derivatives ∂ (defined as eigen-forms) do the opposite, $ab \xrightarrow{\partial} a + b$. A fully realized ab is, we see, comprised of two *conjugate* co-exclusions, a sine/cosine-type relationship. Higher grade operators abc, $abcd$, ... are constructed similarly: $\delta(ab + c) = abc$, $\delta(ab + cd) = abcd$, etc. This hierarchical construction, which is thermodynamically favored, is the subject of the next section; see [4, §8 "The Bit Bang"] for further details.

We can now answer the man's question, *Did anything happen*? We can answer, "Yes, when the coin changed hands, the state of the system rotated 180^o: $ab(a + \tilde{b})ba = \tilde{a} + b$." We see that one bit of information ("something happened") results from the alternation of the two mutually exclusive states.

---

tions, yet the 2 measures were uncorrelated with each other and separately perturbable. These and other findings indicate that $1/f$ scaling is too pervasive to be idiosyncratic and of limited relevance. It is instead argued that $1/f$ scaling reflects the coordinative, metastable basis of cognitive function." See Riemann Fever for the undertone calculation.

19 What makes it tricky is that if at the same time as the man hides the coin he has shown you, you walk around to his back side (be careful how you do it), then it would look to you like nothing happened at all, *vis a vis* the coin, when he shows it again: it's still in the same place relative to you.

[The transition $\mathfrak{a} + \mathfrak{b} \xrightarrow{\delta} \mathfrak{a}\mathfrak{b}$ is in fact the basic act of perception, called the *first perception*, subsequent meta-perceptions being derivative.]

The occurrence here of 180^o is worth noting: it corresponds to an *inversion* of the state, ie. a reversible change, a *space-like reaction*. The physical analogy is a wave reflecting off a wall, eg. a tethered string. The instigating change from the boundary thus "reflects" off the top of the graded action hierarchy. Untethered, the reflection is 90^o, which corresponds to inverting just one of the boundaries, and is a *time-like reaction*.

With the co-exclusion concept in hand, we can now add a refinement to the idea of co-occurrence. Let S be the space of all imaginable expressions in our algebra $\mathcal{G}$. Thinking now computationally, this means that they are all "there" at the same time. That is, S *is the space of superpositions*, of all *imaginable* co-occurrences of elements of our algebra $\mathcal{G}$ *all at the same time*. Let then G be the space of *actually occurring* (but still space-like) entities, which means no co-exclusionary states allowed. When things move from S to G, superposition is everywhere replaced by reversible alternation *a la* $\pm\mathfrak{a}\mathfrak{b}$, ie. G is a sub-space of S.

In less abstract terms, we could say that (wave-world) S corresponds to *imagination*, that (wave-world) G corresponds to the *awareness* of actual possibilities *vis a vis* the surround - and finally, that an *Awareness's* reaction to the surround, via its changes to the boundary $\mathfrak{a} + \mathfrak{b} + \mathfrak{c} + \dots$, projects G's possibilities (the "causal potential" Ψ) down into grounded action in external, *material* reality.

The result is a flurry of irreversible actions at the sensory boundary, constituting the time-like aspect of the Awareness' reaction. *Crucially*, these action are *guaranteed* to occur in their correct spatial context because they have all been filtered through a hierarchy of tauquernions that constitutes the very structure in which all events occur.

Speaking loosely, intuition and learning are captured by δ, and thought and action by ∂. The various specialized modules of the brain reflect different particular organizations of the functionalities described, and there is good reason to believe that alpha waves correlate with space-like cognition, and hence that the more rapid theta waves correlate with time-like or sequential cognition - speech, planning, etc.

Returning to Parseval's Identity, we see that the key (to being able to invoke it, thus getting wave-particle duality, and thus capturing the dual un/localized nature of awareness) is to organize the flux of changes at the boundary

using the distinction *co-occur vs. exclude,* because in so doing, we can then use co-exclusion (= co-boundary operator δ) to perform a hierarchical lift/abstraction, which abstraction is again orthogonal to its components. The orthogonal space so formed allows the application of the Identity. The resulting (novelty-generated) increase in the dimensionality of the orthogonal space increases the complexity and temporal reach of subsequent responses to the surround, and simultaneously the scope of the *Awareness* itself, which inheres in the adaptive self-resonant wave aspect/*experience* of S and G.

5 HIERARCHY AS COMPUTATION

The concept of *hierarchy* - the controlled reversible hiding of information that supplies both context and locality - is of central interest in Computer Science because it is the key to achieving and maintaining conceptual control over very complex software creations. An example of this is the success of object-oriented programming languages (C++, Java, etc.), which offer the programmer sophisticated tools for building complex hierarchies of software objects, which hierarchies then implicitly funnel higher level functionality down to more detailed levels. Contemporary computer systems would be impossible to implement and maintain without these tools.

However, if one looks "under the hood" at what is actually going on at run-time, all of the programmers' fancy hierarchical constructions have been utterly flattened - by the compiler - into long sequences of nested function calls. The sophisticated object-oriented hierarchical structures at the programming language level are actually just syntactic sugar, disguising the fact that the only *real* hierarchy concept in contemporary software thinking is good old, tried-and-true function composition: $y = f(x)$ and $z = g(y)$ combine into $z = g(f(x))$, which says first do f, then do g.

Or instead of "objects", look at "remote procedure call", or "agents": in the end, virtually *everything* is made out of function composition, with maybe a little memory on the side. Thus, as a hierarchy concept, *function composition is fundamentally sequential*, that is, it really isn't hierarchical at all. Rather, it's fundamentally flat: we ourselves *design* the composition sequence, and in general we *impose* it on what is (according to Turing completeness) always ultimately a *sequence* in the real world [and not forgetting relativity!].

This state of affairs has *greatly* hindered the construction of highly concurrent, physically and conceptually distributed systems in software, whilst

Nature's ability to blithely do this all the time has remained a mystery. The computational hierarchy concept offered in this paper - described here using homology theory's boundary / co-boundary operators ∂/δ - exposes how Nature might do this, and thus represents a major advance in the sciences of computation and information. For example, computations using it are space-like rather than time-like in their behavior, they are inherently self-organizing, and they can even be self-aware, cf. the *Topsy Test* for awareness.

As just sketched above, higher grade operators abc, $abcd$, ... are constructed from lower grade operators via the co-boundary operator δ: $abc = \delta(ab + c)$, $abcd = \delta(ab + cd)$, etc. Clearly this process, which can variously describe thermodynamic evolution, organic growth, and conceptual learning, can continue indefinitely, creating increasingly complex global operators. Unfortunately, their very complexity, both individual and collectively intertwined, means that the we quickly lose control of the system's semantics ... What will it do in novel situation X? Can it accomplish goal Y? This puts a very real lid on the useful height of the growth hierarchy.

Fortunately, geometric algebra's very structure supplies an elegant solution to this problem: its semantics loop $\mathrm{mod}\,4$. That is, $1^2 = +1$, $a^2 = +1$, $(ab)^2 = -1$, $(abc)^2 = -1$, $(abcd)^2 = +1$, $(abcde)^2 = +1$, ... and one sees that the sign pattern $++--++--$... is that of the powers of $i = \sqrt{-1}$. The table below shows one way[20] to then map potential higher grades to grades zero, one, and two, for then to repeat the δ-build-up process:

pairs	δ(pair)	new level		
$3 \bmod 4 + 3 \bmod 4$	$\rightsquigarrow$	6	$= 2 \bmod 4$	$\searrow$
$2 \bmod 4 + 3 \bmod 4$	$\rightsquigarrow$	5	$= 1 \bmod 4$	$\searrow$
$2 \bmod 4 + 2 \bmod 4$	$\rightsquigarrow$	4	$= 0 \bmod 4$	$\searrow$
$1 \bmod 4 + 2 \bmod 4$	$\rightsquigarrow$	3	charge	$\downarrow$
$1 \bmod 4 + 1 \bmod 4$	$\rightsquigarrow$	2	spin	$\xleftarrow{6}$ $\downarrow$
$0 \bmod 4$ & $1 \bmod 4$	$\rightsquigarrow$	1	existence	$\xleftarrow{4,5}$ $\downarrow$

This semantic folding can be elaborated to produce arbitrary growth patterns using a new [TLinda programming language] construct called a *corm*,

20 The algebra contains many such semantic symmetries, eg. with $\mathrm{mod}\,8$, the two $\mathrm{mod}\,4$ halves generate slightly different spaces; also octonions and other vector-algebra variants; see Wikipedia: Classification of Clifford algebras. It is very possible that various specialized brain modules implement such "non-standard" symmetries, these being particularly well-suited to their function. The overall structure via the table is an iterated $U(1) \times SU(2) \times SU(3) \times SO(4)$.

which, like its botanical namesake, is a self-replicating "root bundle" that specifies the growth to come.

Speaking of growth, one should recognize the rapidly overwhelming $O(2^{2^n})$ growth in the potential size of the node space that is the price of recording every experience (eg. in the form of this hierarchy). There are four saving graces for this. The first is that a given co-exclusion is recorded *only once*. The second is that if nothing else, this hierarchical structure is, from an information storage and retrieval point of view, minimal.[21] The third is that even though *any* experience can be captured, a given system cannot possibly *actually* experience them *all* - this would take many universe-lifetimes! So only a tiny tiny fraction of the space will ever be used. The fourth is that one can always prune the exfoliation, eg. as plants, brains, nervous systems etc. do; or as statistics and/or other considerations indicate.

It should be noticed that the actual hierarchical structure at run-time is - ignoring the "wrap-around" at the topmost nodes, described below - a rooted acyclic lattice. That is, most nodes will have several parents, and the sensory bubble-up is roots to leaves, and the reaction's trickle-down is leaves to roots. Since all nodes are combinations of the grade 1 sensors at the system boundary (the roots), the latter are the ultimate parents of all later-created child nodes. So, unlike most tree-like structures in computer science, which are drawn downwards with the root at the top, the present co-exclusion based hierarchy grows from the bottom up. We refer to this lattice structure as "the hierarchy".

The algebraic prescription of execution behavior can be taken further using a form called an inner auto-morphism:

$$AXA^{-1} = X' = A^{-1}XA$$

The examples to follow use the form $A^{-1}XA$ for expository reasons, wherein the rightmost A is treated as the current state of the node in question, to be modified by the up-bubbling sensory X by multiplication on the left. An upgraded X is bubbled further up and eventually inverted, namely via some A^{-1}, again via multiplication the left, whence $X \mapsto X'$. The evolving state X' then trickles back down, meeting leftmost A^{-1}s, each of which reduces X's grade and trickles it further down. Eventually X' will meet an effector (ie. grade 1, at the sensory boundary), therewith completing the system's reaction to the input.

21 Modulo clever compression schemes.

An example of an inner auto-morphism is

$$ba(a+b)ab = -a-b$$

[which is the same as

$$ab(a+b)ba = -a-b$$

because for simple m-vectors like ab, abc, ..., their inverses equal their reverses:

$$AA^{-1} = AA^{\dagger} = aa = abba = baab = abccba = cbaabc = abcddcba \ldots = +1\Big]$$

The hierarchy is built out of inner auto-morphisms $A^{-1}XA = X'$ such that $A = \delta X$ for the bubble-up, and $X = \partial A$ on the trickle-down:

$$\begin{aligned}
ba(a+b)ab &= -a-b \\
cba(a+bc)abc &= a+bc \\
dcba(ab+cd)abcd &= ab+cd \\
dcba(a+bcd)abcd &= -a-bcd \\
edcba(ab+cde)abcde &= ab+cde \\
edcba(a+bcde)abcde &= a+bcde
\end{aligned}$$

These transformations hold for all variants of the X-form, eg. $cba(b-ac)abc = b-ac$. The $\mathrm{mod}\,4$ grade-cycling means that these transformations cover all eventualities.

The first and third lines in the table suggest trying input $(a+b)+(c+d)$ on the hierarchy $abcd = \delta(ab+cd) = \delta(\delta(a+b)+\delta(c+d))$.

The first thing that happens is that the state ab will be operated upon by *new* sensory information, namely the bubble $a+b$ (whence it follows that the immediately prior state was $-a-b$). For this bubble:node collision - multiplying, as prescribed, on the left - we write $(a+b)ab$, and similarly for $c+d$.

Suppose as a preliminary example that ab is the top of the hierarchy. It will then (by default) simply turn the bubbled-up change into a goal to change it back, ie. maintain equilibrium by completing the oscillation. So write therefore $ba(a+b)ab$ whose result $-a-b$ becomes the two goals, $+a \to -a$ and $+b \to -b$. Equilibrium has been maintained.[22]

22 One can also, of course, simply ignore the bubble. But given the fundamental wave-like nature of the paradigm, inverting the change seems the obvious best *default* choice. Clearly, some top-most nodes will require human approval rather than defaulting.

Now expand the example so that $abcd = \delta(ab + cd)$ is the top of the hierarchy. Then the algebra tracks what happens as follows:

1. Two new state bubbles $(x + y)$ operate on their parents xy, making new internal xy states: $(+a + b)ab + (+c + d)cd = (-a + b) + (-c + d)$; and forming the grade 2 bubble $-ab - cd$.
2. The bubble from (1), $-ab - cd$, reaches $abcd$: $(-ab - cd)abcd$.
3. Reflect from top $\Rightarrow$ bubbles become droplets: $dcba(-ab - cd))abcd = ba + dc$. [Recall that $-ab = ba$.]
4. The two droplets from (3) hit their respective xy's: $ba(-a + b) + dc(-c + d)$ & then trickle further
5. down, having now become $-a - b - c - d$, which, being on the sensory boundary, are effector commands, ie. $+a \rightarrow -a$, $+b \rightarrow -b$, $+c \rightarrow -c$, $+d \rightarrow -d$.

Thus once again the change is complemented, the oscillation completed, and equilibrium maintained.

The hierarchical execution regime we have just described is a pure space-like computation. It reacts to its environment with utter immediacy - what *can* happen is what *does* happen ... and *only* that - in an uninterrupted, purely entropic and subjectively timeless *Now*.

The computation's only "purpose" is to maintain equilibrium with its surround, all the while growing new structure based on its sensory experience. *What* the computation does, *how* it actually attains its cthonic goals[23] is purely dependent on happenstance, on the surround just happening to be in a favorable configuration, *for it cannot pursue*. Genuine *purpose* - the *conscious* pursuit of a goal - demands further structure [8].[24]

Nevertheless, some local scalar mechanisms can still be applied, especially by living systems, to optimize their responses. The most obvious of these is to count uses of nodes, and use these counts to choose from the possibilities that have been served up by the hierarchy. This clearly works, since our own nervous systems do this, and it has long been the basis for learning in neural nets, cf. Bayesian learning.

23 *Cthonic* [Wikipedia]: "In analytical psychology, the term chthonic was often used to describe the spirit of nature within".

24 The TLinda source code for everything (seven pages!) will be revealed when the the patent issues, ?primo 2018.

The disadvantage is that as the counts build up, what is initially a behavior by choice can become first a habit and then eventually a rigidity, while at the same time ignoring other possibilities, which are still being offered by the hierarchy. The result is a kind of over-specialization or tunnel vision. Thus, while this kind of innovation makes evolutionary sense, it is unnecessary when there is no speed penalty associated with always considering all possibilities.

Opposite to specialization is generalization, where this term is used to mean the passage from a particular given state - in the form of some set of hierarchy nodes that are simultaneously active - to another, larger space of which the original particular state is a sub-space, and this sub-space is highly correlated with the other sub-spaces that altogether make up the larger space.

Recalling that Parseval's Identity applies to our algebra, it is easy to imagine how the phenomenon of generalization is accomplished, namely via resonance. That is, the states that resonate with the original state are those that are most similar, the most closely correlated in the frequency domain. Frequency in turn translates to a vector's *grade*. That is, a particular state - say $a + c + ab + acdef$ - will resonate, possibly over many $\mathrm{mod}\,4$ levels, with all states containing $u + v + uw + uvxyz$. These latter, individually and together, form the generalization. Indeed, one can imagine a purely wave-based implementation of the hierarchy; and one might well find working examples in cells or insects. Analogy and metaphor are linguistic examples.

As the hierarchy grows, the time for a bubble to reach the top, and the corresponding reaction to trickle all the way back down, is proportional to the number of hierarchical levels traversed, ie. the reaction time is generally logarithmic in the overall size of the hierarchy. So the execution time of a hierarchical space-like computation scales very well indeed. It's also fast, because the *vast* majority of nodes is always logically (and automatically, via the co-ex basis) blocked, and activity tends to be localized and coherent. The traditional AI "searching" exercise is done by the bubble-up and trickle-down processes, that is, via δ, search is inherent in the hierarchy's very structure, whence most of the "computation" is pointer-following, and the rest trivial.

To capture the *persistent* aspect of awareness - it's present whenever I am - I postulate that it is a *resonant state* - a self-maintaining and very complex oscillation - where the spectrum of this resonance will vary, eg. according to the properties of the surround wherein the awareness is emplaced. This resonant state rests on, and derives from, the brain's neural substrate, but *nevertheless*, the mathematical space in which the resonant state exists is *outside*

of (and *much* larger than) the mathematical space defined by individual neural function, because it is a co-occurrence (ie. superposition) state. The resonant state deriving from the various individual oscillations is of size $\mathcal{O}(2^n)$, versus the default $\mathcal{O}(n \times n = n^2)$.

That is, algebraically, an EEG-type wave of brain activity is a *scalar* sum of neural activity, treating all neurons as being in the *same* dimension. But as the Coin Demonstration shows, a close analysis of co-occurring processes leads to the conclusion that the processes involved (eg. neurons) lie on *orthogonal* dimensions, which algebraically means that $ab = -ba$. Thus any argument that relies on globalizing the definition of an *individual* neuron's function is flawed. In other words, our aware experience is *usually* an on-going 3d projection of a *much* larger space, of which 3+1d is the *result*, and not the default one-and-only beginning-and-ending place.

So both the materialists (the Pythagorean side of Parseval) and the non-materialists (the Fourier side of Parseval) get their cake, and get to eat it too ... for the price of also being half wrong, ie. for claiming that their story was the *whole* story. From a discrete process and informational point of view, Parseval's *Identity* cements the argument that both stories are correct, simultaneously, all the time ... *all the way down* [4].

6 AWARENESS IS OUTSIDE OF TURING'S BOX

The foregoing has introduced much novelty, ... much-needed novelty, one could though say. So perhaps a brief recapitulation will be helpful.

We began by showing that the essentials of Turing's concept of computation - *if-then-else* and *wait/signal* - are captured by idempotents and products thereof, all belonging to the geometric algebra $\mathcal{G}$ over $\mathbb{Z}_3 = \{0, 1, 2\} = \{0, 1, -1\}$. The reasoning is very direct, in that Products are inherently sequential, and sequentiality is the core of Turing's model.

We also saw how the processes formed from *wait/signal* connections are inherently non-deterministic. We can conclude that natural processes simply *are* non-deterministic ... the reason and the mechanism are clear and there is *no* mystery to be solved. Thus Everett's idea of explaining non-determinism using the device of ever-forking ("parallel") universes in the face of decisions is seen to be entirely superfluous and lacks empirical content.

We then showed how Sums in the algebra can be construed as co-occurrences of computational events, and, via the Coin Demo, that there is information

inherent in co-occurrence, *per se*. The information in complementary co-occurrences can be combined, via the co-exclusion principle, to form a new node that encodes its constituents, creating successive meta-level descriptions of the computation at the level below.

Thus arises the Hierarchy. The nodes of the Hierarchy, ie. m-vector elements of $\mathcal{G}$, are mutually orthogonal, which is interpreted to mean that the processes they both symbolize and encapsulate are nominally independent and concurrent.

Processes being mutually orthogonal also allowed us to invoke Parseval's Identity, which equates every expression in $\mathcal{G}$ to the Fourier decomposition of a corresponding input. This in turn allows us to responsibly claim that the experience of complex, self-arising, self-adaptive resonances in the hierarchy is what we call awareness and awareness of awareness ("consciousness").

The phrase "awareness of awareness" can be understood to mean that there exist two simultaneous resonances, a "lower" and a "higher", in that the higher consciousness is at a higher hierarchical level and therefore oscillates more slowly than the lower. That is, the higher level via its lower frequency is experienced as relatively unchanging, whilst the lower level but higher-frequency processes evolve much faster and "things change". It's easy to see this structure as a sandwich made of *mod4* hierarchy segments, where the meat in the sandwich is the *mod4* phase transition that connects them.

From a subjective point of view, one can responsibly describe the higher level resonance, the *entity* so to speak, as *experiencing* the lower's progress.

This last statement is just one example of how space-like computation directly captures common phenomena that traditional sequential models struggle mightily with. The following introduces more such revealing captures.

We begin with the idea of 3d space, with three orthogonal dimensions/axes called a, b, c. Taking a, b, c as vectors, the product ab is the a, b-plane, the product ac is the a, c-plane and the product bc is the b, c-plane. Similarly, *abc* is a volume.

The set $\{ab, bc, ca\}$ is called a quaternion triple, and they form a mathematical *group*, namely the quaternion group. The sum $ab + bc + ca$, ie. the concurrent existence of these three planes, specifies the *same* point in 3d space as the sum $a + b + c$, in that the 3d rotation $abc(a + b + c) = ab + bc + ca$!

The quaternions were discovered by W^m Hamilton in 1843, and introduced mathematics to the surprising and liberating concept of anti-commutativity, $xy = -yx$ [ie. x, y are orthogonal = independent asynchronous processes]. Also, $\gamma = a + b + c$ identifies a photon, wherein we see the intimate connection between light and space first uncovered by Einstein.

However, in addition to Hamilton's quaternions, it turns out that there are exactly *two more* quaternion variants lurking deep in $\mathcal{G}$'s and mathematics' shadowlands: the triples $\{ab - cd, ac + bd, ad - bd\}$ and $\{ab - cde, ac + bde, bc - ade\}$. We have dubbed the former *tauquernions* [4] and the latter *tauquinions* [3]. What is special and unexpected about them is that they are simultaneously 3d space *and* irreversible operators!

How can this be? The answer is simply that the space-like rotations described/ performed by the tauquernions and tauquinions increase entropy as they project the hierarchy space into $3 + 1$d, and in so doing, continually re-construct space-time on-the-fly. Algebraically, notice that $(ab + cd)$ can be factored to read $(-1 + abcd)ba$, where $-1 + abcd$ is idempotent, and $ab + cde$ similarly results in $-1 + abcde$.[25]

Besides being representations of the quaternion group, both the tauquernions and the tauquinions also map to a group called *Spin{6}*:

$$\{ab, ac, ad, bc, bd, cd\} \mapsto \{ab - cd, ac + bd, ad - bc\} \mapsto SO(4)$$

$$\{ab + cde, ac - bde, ad + bce, -bc - ade, bd - ace, -cd - abe\} \mapsto SU(3)$$

Spin{6} can in turn map to either $SU(3)$ - the world of quarks, charge and electro-magnetism - or $SO(4)$, which is 4d orthogonal space, one version of which is our familiar relativistic $3 + 1$d. Thus our model of awareness has the basic mathematical structure $\mathcal{G}_2 \rightsquigarrow U(1)$, $\mathcal{G}_3 \rightsquigarrow SU(2)$, $\mathcal{G}_5 \rightsquigarrow SU(3)$, and $\mathcal{G}_4 \rightsquigarrow SO(4)$, ie. the Standard Model of physics augmented with *3+1d* space in the form of $SO(4)$!

Recalling the bubble-up/trickle-down example from the preceding section, we see that it was precisely the $2 + 2$ tauquernions and the $2 + 3$ tauquinions that were the bricks that formed the Hierarchy.[26] That is, $\mathcal{G}_2$ gives us the circle group and $\mathcal{G}_3$ the quaternion group, ditto $\mathcal{G}_4$ and the tauquernions, and $\mathcal{G}_5$ and the tauquinions: everything is geometric !

25 NB: The forms $-1 \pm abcd$ and $-1 \pm abcde$ play the role of $+1$ in their respective algebras; -1 is $+1 \mp abcd$ and $1 \mp abcde$ respectively.

26 See [4] for these and other cases.

One can only conclude that *awareness is entirely space-like, and everything is made out of it*. Since the Hierarchy builds from the bottom, this ultimately means[27] that we and the universe are neither separate nor to be distinguished from each other.

Furthermore, both the tauquernions and the tauquinions present as *fields*, not particles, so they are well-suited to capturing the more ineffable phenomena of feelings and emotions (respectively).

That is, the phenomenology of our [tauquinion based] emotional system is fundamentally electro-magnetic in character. Common slang supports this conclusion: sexually magnetic, lightning temper, sparks of rage, charged situation, etc. The Vedic chakra system agrees as well: "2" is the hara chakra, which deals with basic distinctions like inside/outside, left/right, up/down, male/female, life/death. And "3" is the solar plexus chakra, our social brain, the one that always and immediately understands and remembers who owes whom in every exchange; and also the seat of vision (cf. quaternions).

The phenomenology of "4", the heart chakra, derives from the tauquernion field and is very different. All of its elements, namely 2-vectors like ab and cd, and 4-vectors like $abcd$ and $efgh$ will, when operating on each other, commute, ie. $PQ = QP$. This is because vectors of even grade always commute, but the implication is that for the tauquernion field, there is ***no*** temporal order even though all "motion" is irreversible (because entropic). The *only* operating distinction in this ***acausal* regime** is co-occurrence. This is the ***Now***.

Add in the facts that (1) a tauquernion triple, being nilpotent, is a strong candidate for the name Higgs, and (2) that tauquernions are entanglement operators, and we have a description of gravity [4]. The gravitic glue that gathers the cosmos is thus identified with the most profound experiences a human is capable of - consciousness, selflessness, altruism, love, and awe. This self-gathering *gravitas*, this *amour* as Newton originally called it, is at the root of human cooperation and unity.

So yes, not only are *We* outside of Turing's sequential box, *We* are *way* out of that box ... *We* are namely made of pure space, and are entangled entirely with the rest of universe.

27 ...if the mod4 buildup has the gauge property, as is likely.

7 BIBLIOGRAPHY

[1] Manthey, M. *Synchronization: the Font of Physical Structure*. Proc. ANPA 2007. See TauQuernions.org.

[2] Penrose, R. "The Emperor's New Mind: Concerning Computers, Minds, and The Laws of Physics", 1989.

[3] Manthey, M. *The Topsy Test for Awareness*. Proc. ANPA 2014. See RootsOfUnity.org.

[4] Manthey, M. *Tauquernions* τ_i, τ_j, τ_k*: 3+1d Dissipative Space out of Quantum Mechanics*. Proc. ANPA 2013. See TauQuernions.org.

[5] Manthey, M. *Riemann Fever*. Proc. ANPA 2015. See TauQuernions.org.

[6] England, J. "A New Physics Theory of Life". https://www.quantamagazine.org/a-new-thermodynamics-theory-of-the-origin-of-life-20140122/

[7] Manthey, M. *Space-like Computations in Computing Devices*. USPTO - #PCT/US2016/035885, and other patents pending.

[8] Manthey, M. *Electro-magnetism is a Fundamentally Fractal Fibonacci Fourier Field*. Proc. ANPA 2014.

MONADIC DESIGN FOR UNIVERSAL SYSTEMS

NICK ROSSITER, MICHAEL HEATHER
MICHAEL BROCKWAY

Department of Computer Science and Digital Technologies
Northumbria University, NE1 8ST, UK

nick.rossiter1@btinternet.com; michael.heather@trinity.cantab.net;
michael.brockway@unn.ac.uk
http://www.nickrossiter.org/process/

Abstract: The work described here builds on recent work presented at ANPA on structure and process in the universe. The internal structure of the topos is explored further with particular emphasis on the nature of the pasted pullback, including the conditions for a pasting to be valid and the inherent recursive nature of pullback structures. Dolittle diagrams are employed for representing the intension/extension relationships. A banking example is explored, leading to the nature of the external processes acting upon the topos such as transactions. These processes are represented by monads, giving a three-level closure on the activity. The nature of monads is explored. The T-algebra enables changes to be made in the monad structure, giving the potential for adaptability. Monads, that have been strengthened by the Kleisli lift to the Cartesian form, can be composed naturally, facilitating the construction of large-scale information systems with reliability, as required for transactions in the banking world.

1 INTRODUCTION

The fundamental categorical facilities identified for a Universe, whether from any Universe of Discourse up to the Universe, is primarily the Topos as a structural data-type including the Monad for process. The application of the monad to a topos gives the operation of a process on data at the highest level, defined as a unique solution up to natural isomorphism. We will demonstrate such an application, explore how its performance relates to alternative techniques and discuss further work required.

Succinctly the topos is a fundamental Cartesian Closed Category (CCC), a category with limits and exponentials subobjects. A CCC has an internal logic of the typed λ-calculus, an identity functor and the interchangeability of levels, with nodes being either objects or categories. A topos has additional properties beyond a CCC ([19], at p.106) including a subobject classifier, the internal logic of Heyting, that is intuitionistic logic, and a reflective subtopos as a category for recursive query closure. A topos is closed at the top but open at the bottom. An important issue is the flavour of category theory that is being employed. The starting point is what might be termed Eilenberg/-Mac Lane category theory (EML) as set out in pure mathematical terms by Saunders Mac Lane in his book *Categories for the Working Mathematician* [19]. The employment of category theory in metaphysics and the implementation of the theory in programming languages has shown some limitations to EML, in particular with the distributive law, resulting from a restricted view of naturality. Many of these points are discussed in detail later.

The application of the topos to data was established in papers at ANPA 35 [29], introducing the topos/monad approach, and ANPA 36 [30], bringing out the interoperable use of allegories for legacy relational systems. Structures developed as a topos include pasted pullbacks, to represent relatedness, and recursion in which any juncture in the structure is a pullback in its own right. The exact nature of the match, in the pasting operation, is discussed later in Section 4.2. The relationship of the topos structures to Fifth Normal Form (5NF) [12], also known as PJNF (Project-Join Normal Form), is of relevance. A working definition of PJNF is that a table T satisfies the normal form if it cannot have a lossless decomposition into any number of smaller tables. The decomposition is a projection into tables with smaller numbers of columns; lossless implies that the join of the resulting tables returns the original table T unchanged. The pullback structure has an inherent feature for handling PJNF: the adjointness between the projections π from the product onto the coproduct and the diagonal join Δ from the coproduct onto the product. This indicates the strength of the topos data structuring method for handling a challenging ultimate stage in relational database design. Other less powerful normalisation techniques are considered to be so set-based that any categorial approach would be categorification. The Cocartesian dual to the topos may offer further insights into the data structuring process: the mapping from PRE to POS is effectively the data normalisation process in relational databases, where PRE is the preorder with no initial object and POS is the partial order with no cycles and both an initial and a terminal object, satisfying the closed world assumption (CWA). It should be emphasised though that the topos is a

more general data structure than a relational database, which while closed at the top is open at the bottom, thus not restricted by the CWA.

The use of the allegories of Freyd [7] as a basis for data structures was attempted [30] but rejected because of their lack of naturality as set-based relations; the allegories will have use though in interoperability as a wrapper for legacy relational databases. Internal queries on a topos are handled by the subobject classifier, which may be Boolean (0 or 1) or the more general double powerobject. Both forms were illustrated in the two ANPA papers cited above. The provision of examples of Heyting intuitionistic logic for an application remains an objective. Internal queries are more akin to data searches, such as through Google, but do not provide a well-defined process capability.

The application used was of grading for students by modules in a university context, which was adequate from the data structure viewpoint but limited from a data process angle. A more interesting application from the process perspective is banking, including the handling of transactions. This was first studied by us in ANPA 27 [28].

Monadic design is a novel technique for handling the dynamic aspects of an application. Aspects to be investigated are the adjointness, inherent in the approach, the flavours of monad which are most suited to process applications and the T-algebra for modifying the adjunction.

The intention in this paper is therefore to introduce a new application, banking, which provides a more suitable test for an external process of a monad on a topos data structure. The mechanism of pasting is to be investigated in detail and the relationship of the topos to database normalisation is to be clarified. Monadic design will be developed for the topos.

2 PULLBACK: SINGLE RELATIONSHIP FOR STUDENT GRADING

The topos has limits and the pullback is a limit. Figure 11.1 shows for the student application, studied in the ANPA 35/36 papers, a simple pullback of assessment: with limit $\mathbf{S} \times_{\mathbf{G}} \mathbf{M}$, the product of Student and Module in the context of Grading. The relationship between the product $\mathbf{S} \times_{\mathbf{G}} \mathbf{M}$ and $\mathbf{G}$ is adjoint, with the following logic condition holding: $\exists \dashv \Delta \dashv \forall$. The functor Δ selects pairs of $\mathbf{S}$ and $\mathbf{M}$ in a relationship in the context of $\mathbf{G}$, such that $\exists$ is left adjoint to Δ and $\forall$ is right adjoint to Δ, with a consequential facility for consistent logical operation. The arrows can be interpreted as follows: π_1

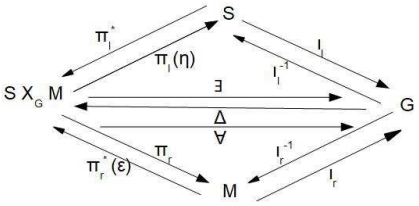


Figure 11.1: Pullback for a Single Relationship $\mathbf{S} \times_{\mathbf{G}} \mathbf{M}$: Grading by Student and Module; categories are **S** Student, **M** Module, **G** Grading

identifies the participating students, π_r the modules taken, ι_l is candidature, and ι_r is marking.

A diagram with such adjointness was termed by Lawvere a hyperdoctrine in the early days of EML category theory [14]. The existential functor $\exists$ records the decision for each student (under the free functor F) of a grade for a specific module. The diagonal Δ sorts student and modules as a component of the underlying functor G. Thr universal functor $\forall$ produces the final grading list generated by F. This shows the fine structure of the adjointness $F \dashv G$.

Other arrows are interpreted as follows. Projections π are from the product onto its constituents, left π_l and right π_r, with dual arrows left π_l^* for student capability, achievement and creativity and right π_r^* for quality of work, respectively. Inclusions ι are into the sum $\mathbf{S} + \mathbf{G} + \mathbf{M}$ from its constituents, left ι_l and right ι_r, providing type checks on the values for students and modules, respectively, participating in the relationship. The dual arrows, ι_l^{-1} and ι_r^{-1}, indicate the participation of students and modules, respectively, in the relationship.

2.1 CLOSURE TO RECURSION: DOLITTLE DIAGRAMS FOR INTENSION/EXTENSION

S, **M**, **G** are each categories, with an optional internal pullback structure, giving a recursive pullback structure with potential unlimited depth. The closure to recursion is that the diagrams at the bottom level are of the Dolittle[1] type

1 named after *The Story of Doctor Dolittle* by Hugh Lofting (1920), involving the mythical Pushmi-pullyu creature.

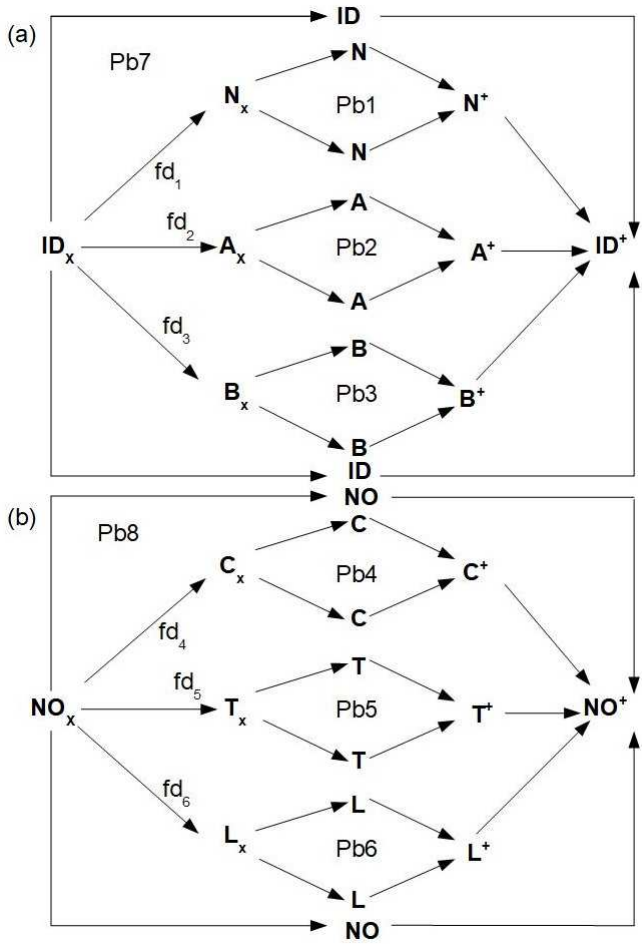


Figure 11.2: a) Dolittle diagram for category **S** for Student with outer diagram Pb7 of limit $\mathbf{ID}_\times$ and colimit $\mathbf{ID}^+$ (Identifier); inner diagrams of Pb1 with limit $\mathbf{N}_\times$ and colimit $\mathbf{N}^+$ (Name), Pb2 with limit $\mathbf{A}_\times$ and colimit $\mathbf{A}^+$ (Address), Pb3 with limit $\mathbf{B}_\times$ and colimit $\mathbf{B}^+$ (Birthdate).

b) Dolittle diagram for category **M** for Module with outer diagram Pb8 of limit $\mathbf{NO}_\times$ and colimit $\mathbf{NO}^+$ (Number); inner diagrams of Pb4 with limit $\mathbf{C}_\times$ and colimit $\mathbf{C}^+$ (Course), Pb5 with limit $\mathbf{T}_\times$ and colimit $\mathbf{T}^+$ (Title), Pb6 with limit $\mathbf{L}_\times$ and colimit $\mathbf{L}^+$ (Level).

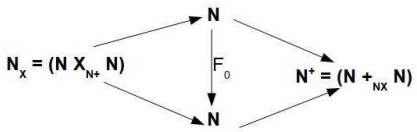


Figure 11.3: Dolittle Diagram for a Single Relationship: Student Names with limit $N_X = \mathbf{N} \times_{\mathbf{N}+_{N_X}\mathbf{N}} \mathbf{N}$, colimit $N^+ = \mathbf{N} \times_{\mathbf{N}\times_{\mathbf{N}^+}\mathbf{N}} \mathbf{N}$, top object **N** of type 1-category, bottom object **N** of type 0-category, F_0 maps 1-category to 0-category.

with a mapping from the data in the left-hand pullback object, a qualified product or $\times$, to equivalent data in the right-hand pushout object, a qualified coproduct or $+$ [9]. Such diagrams are also called Bicartesian squares [2] or pulation squares ([1] pp.205-206). Bicartesian squares as defined by Banach are regular relations, a particular type of pullback in which there is a N:1 relationship between the objects at the top and the bottom, rather than the N:M in the more general case. The special case of a 1:1 relationship would also satisfy the Bicartesian requirement. We agree with Lambek & Scott ([13] pp.65-68) that the two objects are different in purpose, with for instance the top object, the independent variable, being the intension (definition) and the bottom object, the extension (values). There is a functional relationship between the top and bottom objects in a Dolittle diagram, resulting in an ordered product as a string. A helpful view from n-categories is that the top object is a 1-category, defining the data structure in terms of objects and arrows, and the bottom object is a 0-category, holding data values as discrete objects. The arrow from the 1-category to the 0-category is then a 0-functor [25], written as F_0, mapping intension to extension. Everything is related implicitly in a $+$ context; the relationship in a $\times$ context is stronger with explicit connections. Relatedness in Heyting logic is expressed by the condition: $C \times A \leqslant B$ is isomorphic with $C \leqslant A \Rightarrow B$. Such a view of relatedness is not available in set theory.

The Dolittle categories **S** and **M** are shown in Figures 11.2(a) and (b) respectively. Each of these categories comprises a number of further Dolittle diagrams, one for each type of data. Each category has a key or identifier, which provides unique identification of other data properties. The key for the category for student **S** is ID (identifier) and the data properties are N (name), A (address), B (birthdate). Each data property is a Dolittle diagram representing

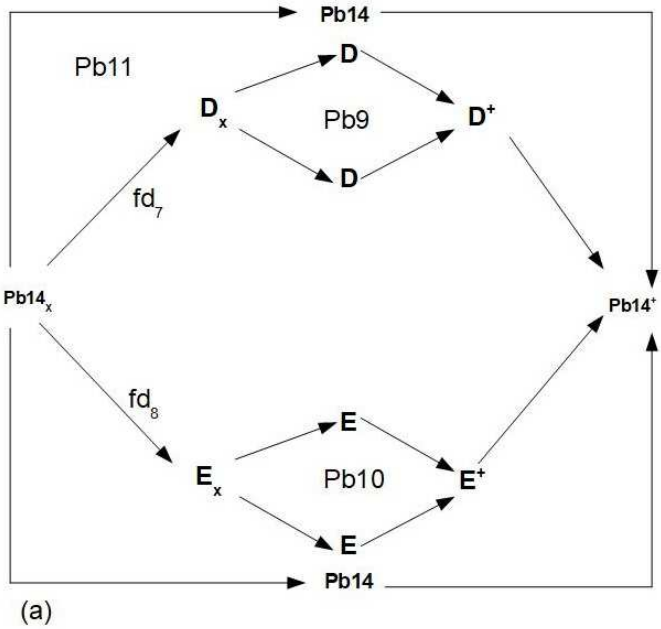


Figure 11.4: a) Dolittle diagram for category **G** with outer diagram Pb11 of limit $\mathbf{Pb14}_{\times}$ and colimit $\mathbf{Pb14}^{+}$; inner diagrams of Pb9 with limit $\mathbf{D}_{\times}$ and colimit $\mathbf{D}^{+}$ (Date of decision), Pb10 with limit $\mathbf{E}_{\times}$ and colimit $\mathbf{E}^{+}$ (Exam board);
b) Pasted pullback diagram for internal key structure for **G**; Pb12 is outer Dolittle Pb8 in Figure 11.2(b); Pb13 is pasting of outer Dolittle in Figure 11.2(a) onto Pb12; Pb14 is composition of Pb13 with Pb12.

the intension-extension relationship as in the diagram shown in Figure 11.3 for the property Name of Students labelled Pb1 in Figures 11.2(a). The limit of this diagram $\mathbf{N} \times_{\mathbf{N}^+} \mathbf{N}$, written shorthand as $\mathbf{N}_\times$, is a product of the intension $\mathbf{N}$ of 1-category type and extension $\mathbf{N}$ of 0-category type in the context of the colimit $\mathbf{N} +_{\mathbf{N}_\times} \mathbf{N}$, written shorthand as $\mathbf{N}^+$. These shorthand abbreviations are used throughout the paper to improve readability. The 0-functor F_0 maps the 1-category intension to the 0-category extension. The colimit identifies, for a pair of intension/extension values in the limit, the values for the names as a sum in the colimit.

Simple diagrams such as Figure 11.3 are the building blocks at the bottom level for the two categories shown in Figures 11.2. The data properties are determined by the key through the functional dependency arrows. For instance the functional dependency $\mathtt{fd}_1 : \mathrm{ID} \longrightarrow \mathrm{N}$ indicates that for each ID, there is a unique N. The dependency is actually defined in the diagram with more detail as the arrow $\mathtt{fd}_1$ from an intension/extension pair for ID, $\mathbf{ID}_\times$, to an intension/extension pair for N, $\mathbf{N}_\times$. There are two further functional dependencies $\mathtt{fd}_2 : \mathbf{ID}_\times \longrightarrow \mathrm{A}_\times$ and $\mathtt{fd}_3 : \mathbf{ID}_\times \longrightarrow \mathrm{B}_\times$ in our diagram. The inner Dolittle diagrams Pb1, Pb2 and Pb3 represent the data properties N, A and B respectively. Completing Figure 11.2(a), there is an outer Dolittle diagram Pb7 with limit of $\mathbf{ID}_\times$ and colimit of $\mathbf{ID}^+$, representing the intension/extension for the key ID. The internal structure of the data properties Name, Address and Birthdate, may be complex, with for instance a name being subdivided into title, first name, middle name, last name. Further there is no obligation for a value in the extension to be an atomic value as in the relational model [12]; the value could be a structured set or a complex object, as in the nested relational and object-oriented approaches respectively.

Figure 11.2(b) is another example of a Dolittle diagram with a 0-functor. The key is NO (module number), forming the outer pullback Pb8 with limit $\mathbf{NO}_\times$ and colimit $\mathbf{NO}^+$. The data properties, represented by the inner pullbacks, labelled Pb4, Pb5 and Pb6, are C (Course), T (Title of module) and L (Level), respectively. The functional dependencies are $\mathtt{fd}_4 : \mathbf{NO}_\times \longrightarrow \mathrm{C}_\times$; $\mathtt{fd}_5 : \mathbf{NO}_\times \longrightarrow \mathrm{T}_\times$; $\mathtt{fd}_6 : \mathbf{NO}_\times \longrightarrow \mathrm{L}_\times$.

The category for student grading by module $\mathbf{G}$ is more complex in that its identifier is a combination of the student identifier ID and the module number NO, necessary to define uniquely a grading. That is the key cannot be described simply as an unqualified product: it needs to be factored through the grading to represent the pairs that actually occur. This requires a pasted[2]

2 pasting is described further in Section 3.2 and discussed in Section 4.

pullback as shown in Figure 11.4(b). The inner pullback Pb12 is the same as the outer Dolittle Pb8, representing the key structure in category **M** in Figure 11.2(b). The outer pullback Pb13 is a pasting of the key structure in category **S**, in Figure 11.2(a), on to Pb12. The pasted key structure is not a Dolittle diagram in itself but is employed in the top and bottom nodes of the outer Dolittle diagram Pb11 in Figure 11.4(a). Pb14 is the overall result from composing Pb12 with Pb13, that is the whole diagram in Figure 11.4(b). The Dolittle diagram for the category **G** is shown in Figure 11.4(a). For the outer Dolittle diagram Pb11 the top object is the intension for the key Pb14 and the bottom object the extension for the key Pb14. The inner Dolittle diagrams, labelled Pb9 and Pb10, represent D (Date of decision) and E (Exam board) respectively with functional dependencies $\mathtt{fd}_7 : \mathtt{Pb14}_\times \longrightarrow \mathtt{D}_\times$ and $\mathtt{fd}_8 : \mathtt{Pb14}_\times \longrightarrow \mathtt{E}_\times$ respectively.

3 BANKING EXAMPLES

We now introduce the Banking example, which is a more suitable subject for illustrating the action of process on a topos as a core feature of banking is the use of transactions for handling changes in a secure manner. We commence with a simple data example, gradually making it more realistic.

3.1 PULLBACK: SINGLE RELATIONSHIP

The simple pullback is shown in Figure 11.5, defined as $\mathbf{P} \times_{\mathbf{T}} \mathbf{A}$, that is the product of Procedure and Account in the context of Transaction, with **P** the category Procedure, **A** the category Account, and **T** the category Transactions. An Account can belong to many users; the Procedure is the type of the transaction, for example: standing order, direct debit, ATM cash withdrawal; the transaction is a transfer of funds according to data processing requirements. **P**, **A**, **T** are categories, with internal pullback structure, giving in general a recursive structure with closure at the Dolittle level, as in Figures 11.2, 11.4 for the student example.

3.2 PULLBACK: TWO PASTED RELATIONSHIPS

In pasted pullbacks two relations are joined together to form a square. An additional category is introduced for User (customer) of **U**. Each user may

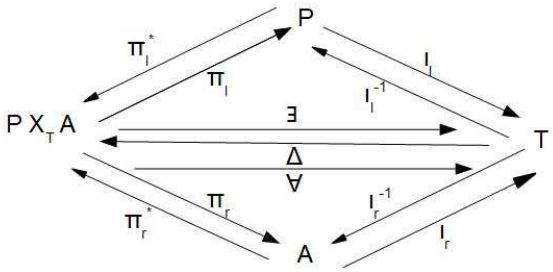


Figure 11.5: Pullback - Single Relationship $\mathbf{P} \times_{\mathbf{T}} \mathbf{A}$: Bank Transactions by Procedure and Account; categories are **P** Procedure, **A** Account, **T** Transaction

have multiple accounts across the banking network: there is a many-to-many (N:M) relationship between **U** and **A**. The second pullback is the product of the subproduct of the first pullback $\mathbf{P} \times_{\mathbf{T}} \mathbf{A}$ with **U** in the context of **A**, as shown in Figure 11.6. The resulting relationship is of account transactions by users. For the purpose of discussion, the pullbacks can be labelled Pb1 for the first square $\mathbf{P} \times_{\mathbf{T}} \mathbf{A}$ and Pb2 for the second square $(\mathbf{P} \times_{\mathbf{T}} \mathbf{A}) \times_{\mathbf{A}} \mathbf{U}$). By EML category theory ([19] pp.71-72) if the squares Pb1 and Pb2 are valid pullbacks, then the whole outer square is also a pullback Pb2 × Pb1. We therefore have three pullback diagrams in a valid pasted relationship.

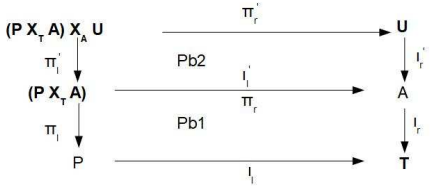


Figure 11.6: Pullback: Two Pasted Relationships: Bank Transactions by Account by User, in Portrait Layout

The vertical stacking of the pasted pullbacks, one above the other, in portrait form is suited to practical applications which could involve 5-10 relation-

ships in a deep nested structure. In EML category theory text books, pasted structures are usually written in horizontal (landscape) form as in Figure 11.7, which is logically identical to that in Figure 11.6.

The aim of pasting in topology is to 'glue together' two continuous functions to create another continuous function. The specific pasting condition for the pullback Pb2 × Pb1 is that $\iota'_l = \pi_r$ after Freyd's Pasting Lemma [7].

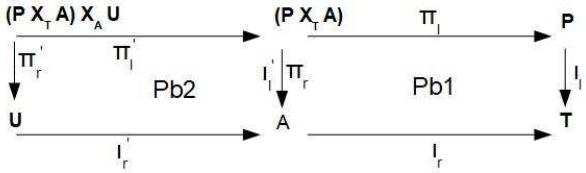


Figure 11.7: Pullback: Two Pasted Relationships: Bank Transactions by Account by User, in Conventional Landscape Layout

To make the application more realistic we add two further categories, those of **B** for Branch and **C** for (banking) Company. Branch:User is also a N:M relationship as each Branch has many Users and each User has many Branches but Company:Branch is a 1:N relationship: each Company has many Branches, each Branch is within one Company. The overall relationship is $(((\mathbf{P} \times_{\mathbf{T}} \mathbf{A}) \times_{\mathbf{A}} \mathbf{U}) \times_{\mathbf{U}} \mathbf{B})$ with **C** in the context of **B** giving the pullback diagram shown in Figure 11.8. The representation of N:M and 1:N relationships is the same in terms of pullback structures, giving a useful symmetry in data design.

Figure 11.8 involves six categories: **C** company, **B** branch, **U** user, **A** account, **P** procedure, **T** transaction, and ten pullbacks: Pb4, Pb3, Pb2, Pb1; Pb4 × Pb3, Pb3 × Pb2, Pb2 × Pb1; Pb4 × Pb3 × Pb2, Pb3 × Pb2 × Pb1, Pb4 × Pb3 × Pb2 × Pb1. The relations within a banking system are shown in more conventional form in Figure 11.9(a) where each single-headed arrow represents a 1:N (one-to-many) relationship and each double-headed arrow represents a N:M (many-to-many) relationship.

For our purposes, a pasted pullback is only a valid pullback if all inner and outer diagrams are pullbacks. There are some theorems in EML category theory ([19] pp.71-72) which enable some deductions to be made based on partial knowledge: for example, with the diagram in Figure 11.7, if the inner

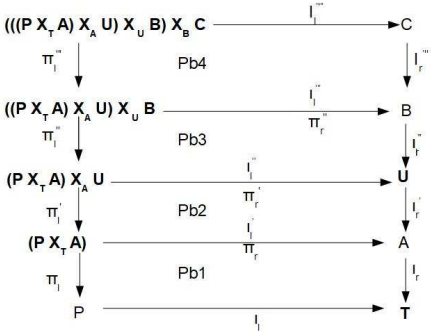


Figure 11.8: Pullback: Four Pasted Relationships: Bank Transactions by Account by User by Branch by Company

diagrams are pullbacks then the outer diagram is a pullback, as stated earlier, and if the outer diagram and the right-hand diagram are pullbacks then the left-hand diagram is a pullback. Such deductions could be facilitated in any practical system but are a distraction from developing a simple robust solution.

As an example of an invalid pullback, consider the diagram in Figure 11.10 where the relationship diagram has been modified to that in Figure 11.9(b). There are seven valid pullbacks in the diagram: Pb4, Pb3, Pb2, Pb1; Pb3 × Pb2, Pb2 × Pb1; Pb3 × Pb2 × Pb1, but not all squares are pullbacks, for example Pb4 × Pb2. Therefore the whole diagram is not a valid pullback.

For any valid pullback, the logic of adjointness holds for the outer square and all inner squares. Therefore for Figure 11.8 with its six valid pullback diagrams, the logic $\exists \dashv \Delta \dashv \forall$ holds across every diagram. An example of this logic is shown in Figure 11.11 for the outer square.

3.3 SUBOBJECT CLASSIFIER

As a pullback the pasted structure is a Cartesian Closed Category (CCC) with products, terminal object and exponentials. Further it is a topos as a CCC with subobject classifier and internal Heyting Logic. The subobject classifier provides an internal query language for which a Boolean example is shown in Figure 11.12.

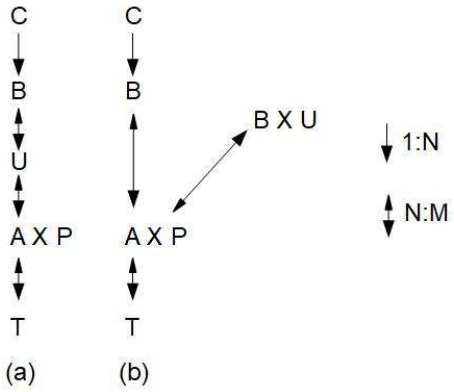


Figure 11.9: Relations within a Banking System corresponding to (a) Figure 11.8 and (b) to Figure 11.10. C is Company, B branch, U user, A account, P procedure, T transaction.

The subobject classifier facilitates simple database or information retrieval queries:

- $\Omega\{0,1\}$ is the subobject classifier with subobjects classified as either 0 or 1
- χ_j is the characteristic function, a query mapping from the object S to $\{0,1\}$, false or true
- 1_{topos} is the terminal object of the topos, giving a handle on the topos
- $\jmath$ is the mapping from the subtopos U, the result of the query, to the object S
- U is the identity of the subtopos, giving query closure

The diagram may be viewed as a pullback of *true* along χ_j, with U as $1_{topos} \times_{\Omega\{0,1\}} S$.

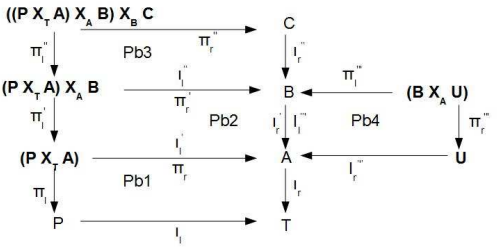


Figure 11.10: Invalid Pullback Diagram, corresponding to Relations in Figure 11.9(b)

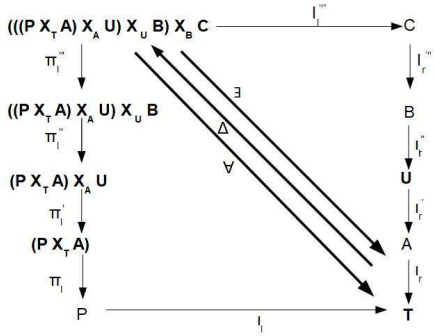


Figure 11.11: Adjointness Holds for all Pullbacks: $\exists \dashv \Delta \dashv \forall$

4 PASTING PULLBACKS: DISCUSSION

To summarise, in a pasted diagram, all pullbacks as inner or outer squares must commute for the diagram to be a valid pullback as a whole. The structure is recursive in that a pullback node may itself be a pullback diagram. Two aspects are worthy of further discussion: how does the pullback diagram relate to data normalisation in conventional data structuring and can the pasting condition be expressed in other forms, drawing out the nature of the '=' condition?

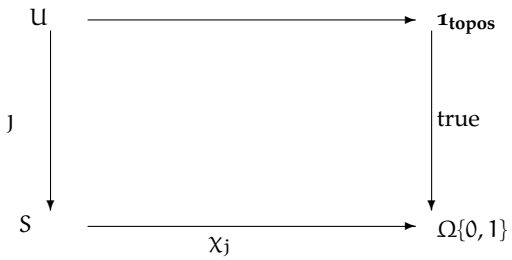


Figure 11.12: Pullback Square for Boolean Subobject Classifier: Definition of Characteristic Function $\chi_j : S \longrightarrow \Omega\{0, 1\}$

4.1 NORMALISATION

Normalisation is the standard technique for evaluating a data design, in particular to determine how closely the logical design matches the physical world. A number of stages have been developed for the set-theoretic relational model: 1NF (First Normal Form), 2NF, 3NF, BCNF, 4NF, 5NF. The last and most demanding stage 5NF concerns us here, not just for its rigour but for its definition in category theory terms, indicated by its alternative name of Project-Join Normal Form (PJNF).

In set theoretic terms, the definition of 5NF is that the structures resulting from the projections can be joined together to return the original structure without loss or gain of information [12]. Looking at the simple pullback diagram, as in Figure 11.5, the projections are the π arrows, π_l and π_r, and the join arrow is the diagonal Δ. PJNF holds through the adjointness in every pullback: $\exists \dashv \Delta \dashv \forall$. The arrows $\exists$ and $\forall$ involve the projections through the compositions: $\exists = \iota \circ \pi$ and $\forall = \iota \circ \pi$. In more complex data structures, the same logic applies. For instance in Figure 11.11 with ten pullback squares (including undrawn inner ones), PJNF will hold if the whole structure and all inner squares are pullbacks with the logic: $\exists \dashv \Delta \dashv \forall$. Surprisingly pullbacks have rarely been used in normalisation studies, an exception being the work of Levene & Vincent [17] who briefly mention the pullback inference rule, following from the interaction between functional dependencies $\exists$ and inclusion dependencies ι.

It should be emphasised that the pullback is not categorification of the set-theoretic approach to normalisation of 5NF, as in earlier work with category theory and databases [11]. The form 5NF was a belated move by set-theoretic

adherents to find a viable approach to normalisation after many earlier attempts had been only partially successful. The pullback follows basic category theory principles and is a *natural* choice for an effective data structure.

4.2 THE PASTING CONDITION

The Pasting Condition is $\iota'_l = \pi_r$, that is the left-inclusion of the outer square equals the right-projection of the inner square. On the surface this looks rather set theoretic, where the '=' would be without context, but in EML category theory the '=' is defined naturally as unique up to natural isomorphism, through the adjointness inherent in the pullback category.

Moreover any pullback can be represented as an equalizer [26], as in Figure 11.13, which is equivalent to Figure 11.5. In the equalizer diagram the product of **P** and **A** in the context of **T**, $\mathbf{P} \times_{\mathbf{T}} \mathbf{A}$, maps onto the product $\mathbf{P} \times \mathbf{A}$ which in turn maps onto **T** where the two paths, $\iota_l \circ \pi_l$ and $\iota_r \circ \pi_r$, converge.

$$\mathbf{P} \times_{\mathbf{T}} \mathbf{A} \longrightarrow \mathbf{P} \times \mathbf{A} \underset{\iota_r \circ \pi_r}{\overset{\iota_l \circ \pi_l}{\rightrightarrows}} \mathbf{T}$$

Figure 11.13: Pullback in Figure 11.5 Represented as an Equalizer

Equalizer diagrams can also be constructed for pasted pullbacks, as in Figure 11.14, which is equivalent to Figure 11.6. In the equalizer diagram the product of $\mathbf{P} \times_{\mathbf{T}} \mathbf{A}$ and **U** in the context of **A**, $(\mathbf{P} \times_{\mathbf{T}} \mathbf{A}) \times_{\mathbf{A}} \mathbf{U}$, maps onto the product $(\mathbf{P} \times_{\mathbf{T}} \mathbf{A}) \times \mathbf{U}$ which in turn maps onto **T** where the two paths, $\iota_l \circ \pi_l \circ \pi'_l$ and $\iota_r \circ \iota'_r \circ \pi'_r$, converge.

$$(\mathbf{P} \times_{\mathbf{T}} \mathbf{A}) \times_{\mathbf{A}} \mathbf{U} \longrightarrow (\mathbf{P} \times_{\mathbf{T}} \mathbf{A}) \times \mathbf{U} \underset{\iota_r \circ \iota'_r \circ \pi'_r}{\overset{\iota_l \circ \pi_l \circ \pi'_l}{\rightrightarrows}} \mathbf{T}$$

Figure 11.14: Pasted Pullback in Figure 11.6 Represented as an Equalizer

5 EXTERNAL PROCESS

The concept of process is underpinned by metaphysics, as defined in the writing of authors such as Alfred North Whitehead, in his book *Process and Reality* [31]. For any entity in the universe, the actions possible upon it and the rules for such actions are a critical part of the whole system. First we look at the technical features within category theory for representing process. We next look at the requirements for the real world and review the facilities of the theory that appear to be most relevant.

5.1 PROCESS IN CATEGORY THEORY

An internal process is a morphism (arrow) within a topos, such as $p : A \longrightarrow B$, where the process p takes object A to object B in the same topos. Such arrows play a natural role in the category construction. An external process is activity on a topos $\mathbf{E}$, taking it to another topos $\mathbf{E}'$, such as provided by a functor F with $F : \mathbf{E} \longrightarrow \mathbf{E}'$. Both $\mathbf{E}, \mathbf{E}'$ must conform to the natural rules for topos construction. Constraints on the transition between $\mathbf{E}$ and $\mathbf{E}'$ are enforced through adjointness between F ($\mathbf{E} \longrightarrow \mathbf{E}'$) and its dual G ($\mathbf{E}' \longrightarrow \mathbf{E}$), such that $F \dashv G$ and the 4-tuple $< F, G, \eta, \epsilon >$ exists where η is the unit of adjunction $\eta : 1_E \longrightarrow GFE$, ϵ is the counit of adjunction $\epsilon : FGE' \longrightarrow 1_{E'}$ and E, E' are objects in categories $\mathbf{E}$ and $\mathbf{E}'$ respectively. The pair of adjoint functors $F \dashv G$ may be written as T and the dual $G \dashv F$ as S.

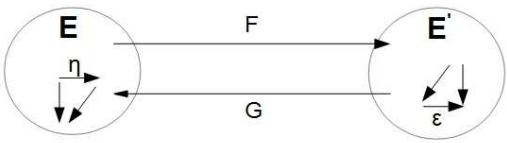


Figure 11.15: Multiple 'Cycles' $GFGFGF(T^3)$ for adjointness $< F, G, \eta, \epsilon >$

The cycle T can be enhanced by performing it three times, T^3, to achieve closure. Such a construction is termed a monad, with its dual S^3 a comonad. The functors and their constraints are illustrated in Figure 11.15. The monad

is a generalisation of the single-level monoid, which has a single operation, binary multiplication $M \times M \longrightarrow M$, and the identity $1 \longrightarrow M$, for an object M.

5.2 REAL-WORLD REQUIREMENTS

The process is represented in information systems by the transaction, which has been the subject of intense study because of its criticality to applications such as banking and internet-based commerce. However, the concept is a very general one, applying for instance to drafting where a transaction may last several days as a technical drawing is modified from one consistent state to another, or maybe months, as a legal document is modified similarly. The notion of transaction in a categorial context was developed in earlier ANPA papers, more generally at ANPA 31 [10], and in considerable detail at ANPA 27 [28]. The principles of the transaction are summarised as ACID: Atomicity, Consistency, Isolation, Durability. Atomicity ensures that the process, however complicated, is viewed as a single arrow. Consistency ensures that all rules have to be satisfied before the transition is made. Isolation ensures that any intermediate results in the process are not revealed. Durability ensures that once a transaction is performed, the results persist until changed by another transaction. The transaction is a logical technique for controlling the real world.

5.3 APPLICABILITY OF THE THREE CYCLES

A transaction is viewed naturally as three 'cycles' of adjointness [28]. The first cycle performs the actual work required; the second checks for any errors or inconsistencies resulting from the first cycle; the third cycle consolidates the changes made provisionally in the first cycle and checked in the second cycle. The 'cycles' are not separate stages; all three cycles are performed as a single snap: the prehension, or grasping, of Whitehead [31]. This single snap satisfies the atomicity and isolation requirements. The second cycle satisfies the consistency requirement, through review against the rules. The third cycle satisfies the durability requirement, through consolidating the results. If adjointness does not hold in any cycle, the transaction is abandoned. We now look at the application of the monad in more detail.

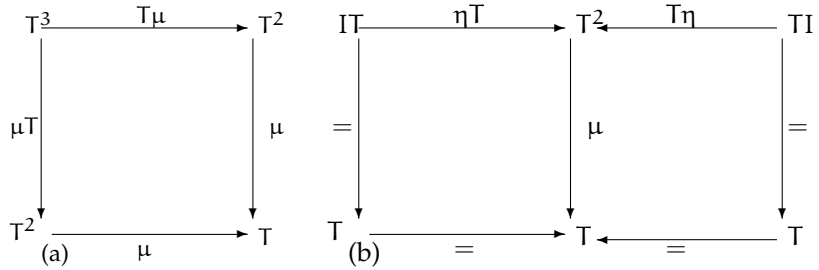


Figure 11.16: (a) Associative Law for Monad $< T, \eta, \mu >$; (b) Left and Right Unitary Laws for Monad $< T, \eta, \mu >$

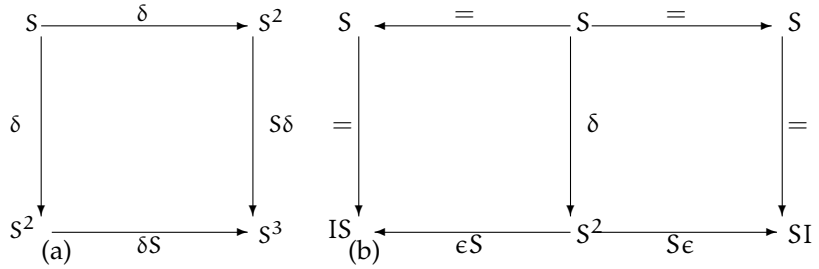


Figure 11.17: (a) Associative Law for Comonad $< S, \epsilon, \delta >$ (b) Left and Right Unitary Laws for Comonad $< S, \epsilon, \delta >$

5.4 TECHNICAL DETAILS OF THE MONAD APPROACH

For a monad, the diagrams for the associative laws and unitary laws are shown in Figure 11.16. These diagrams provide the formal basis for the approach. Figure 11.16(a) shows the relationship between T^3, T^2 and T where T is the endofunctor $GF : \mathbf{X} \longrightarrow \mathbf{X}$, $\mathbf{X}$ being any category. An endofunctor is a functor with the same source and target. A pair of adjoint functors F and G is an endofunctor as the source of $F : \mathbf{X} \longrightarrow \mathbf{Y}$ is $\mathbf{X}$ and the target of $G : \mathbf{Y} \longrightarrow \mathbf{X}$ is also $\mathbf{X}$. The unit or identity of the monad is $\eta : 1 \longrightarrow T$ from Figure 11.16(b) and the multiplication of the monad is $\mu : T^2 \longrightarrow T$ from Figure 11.16(a). We therefore write the monad T as the object $< T, \eta, \mu >$, with the category $\mathbf{X}$, on which the monad is based, omitted as it is inferred from the functors involved. However, it is not wrong to write the monad as the object $< \mathbf{X}, T, \eta, \mu >$ where the nature of $\mathbf{X}$ has a bearing on the arguments being made. Further it is often useful to say on which category the monad is based.

For a comonad, the dual of the monad, the diagrams for the associative laws and unitary laws are shown in Figure 11.17. Figure 11.17(a) shows the relationship between S, S^2 and S^3 where S is the endofunctor $FG : \mathbf{Y} \longrightarrow \mathbf{Y}$, **Y** being any category. The counit or identity of the comonad is $\epsilon : S \longrightarrow 1$ from Figure 11.17(a) and the comultiplication of the comonad is $\delta : S \longrightarrow S^2$ from Figure 11.17(b). We therefore write the comonad S as the object $< S, \epsilon, \delta >$ or $< \mathbf{Y}, S, \epsilon, \delta >$.

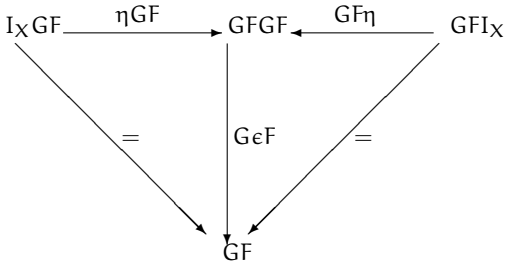


Figure 11.18: The Monad in the category **X**: Triangular Identities defining ϵ

Figure 11.18 shows the two triangular identities for the monad in the category **X**, derived by applying the interchange law to Figure 11.16(b). Through commutativity Figure 11.18 defines the arrow $G\epsilon F : GFGF \longrightarrow GF$. This arrow is the multiplication of Figure 11.16, that is $\mu : T^2 \longrightarrow T$. Therefore we can rewrite the monad $< T, \eta, \mu >$ as $< T, \eta, G\epsilon F >$ for an alternative view, based on the units and counits of adjunction, η and ϵ respectively.

5.5 HISTORICAL AND PRESENT USAGE OF THE MONAD TERM

According to Hippolytus ($170 - 235$ AD), the worldview was inspired by the Pythagoreans, who called the first thing that came into existence the *monad*, from which came the dyad, triad, tetrad, etc. [4]. Gnosticism is a modern term for a multitude of Jewish religious ideas and systems from the first and second century AD, with the highest God, Supreme Being or the One, termed the Monad. The Syrian-Egyptian school depicts creation as coming from a primal monadic source, finally resulting in the creation of the material universe.

The monad entered metaphysics as the *Monadology* of Leibniz, written from 1712-1714 as *Principes de la nature et de la grâce fondé en raison*, which has since been published in various forms and languages [16]. Leibniz allows

just one type of element in the building of the universe, which is given the name monad or entelechy, and described as a simple substance, which has no parts, hence indivisible. Monads are elementary particles with blurred perceptions of one another and have been described as eternal, indecomposable, individual, subject to their own laws, un-interacting, each reflecting the entire universe in a pre-established harmony; monads are centres of force; substance is force, while space, matter, and motion are merely phenomenal. Like atoms, monads are irreducible but differ in their complete mutual independence, and in their following of a preprogrammed set of instructions peculiar to itself, so that a monad 'knows' what to do at each moment. Each monad is like a little mirror of the universe.

The monad term is also used in music, where it is a single note, with a dyad being 2 notes, a triad 3 notes, etc., and in biology where it is a unicellular organism.

In functional programming, the monad is an increasingly popular construction as an abstract data type, with promising developments in the language Haskell [8, 18], named after Haskell B Curry, who developed the transformation of functions through currying in the $\lambda-$calculus. The monad in Haskell is formally classified as an extension of the monad developed in category theory, involving the notion of a strong monad [22, 24]. Such a monad is defined in higher-order category theory as a bicategory construction. In more concrete terms a strong monad is defined as a (categorial) monad with strengthening with respect to products and idempotency. The strengthening with products leads to the concept of a Cartesian monad where, if the underlying categories are pullbacks, the monad T preserves pullbacks and μ and η are Cartesian, then the monad is Cartesian. Such a construction facilitates the use of T in transformations where a Cartesian type is expected. The strengthening with idempotency provides resilience as further operations are performed. So with the underlying category for the monad $\mathbf{X}$ being Cartesian with the object $A \times B$, there is a natural transformation $\tau_{A,B}$ from the Cartesian operation $(A \times TB)$ to $T(A \times B)$ such that strengthening with the identity I is immaterial, consecutive applications of strength commute, and strength commutes with monad unit and multiplication [23]. Further details of the Cartesian monad are found later in this paper in Section 7, in the work by Mulry [24] and in Appendix C of Leinster's book *Higher Operads, Higher Categories* [15].

Category theory is regarded as a unifying force so might be able to provide an insight into all of the above notions of the monad. The notion of unit applies to all the various usages and this is continued into the categorial version with the unit in the monad definition $< T, \eta, \mu >$ of $\eta : 1 \longrightarrow T$ and

the counit in the comonad definition $< S, \epsilon, \delta >$ of $\epsilon : S \longrightarrow 1$. The monad of Leibniz is similar to the categorial version in respect of their following a preprogrammed set of instructions with each monad being a little mirror of the universe. However, there is a major difference – Leibniz's monad is a particle and the categorial monad is a process – emphasising the set-based nature of Leibniz's work. The use of the term monad in music appears to reflect the physical reality of a single note. From a more constructive point of view, musical units, and hence monads, might also include chords and other logical combinations of notes. An application of the categorial monad to music is under active consideration. The use of the term monad for a unicellular organism has lapsed, maybe because the general term was confusable with its use for specific unicellular organisms, the *Monas*. The comparison between the monad of functional programming and that in category theory is the most useful: this shows that the Cartesian monad selected for functional programming is indeed the type of monad needed for information systems as the underlying Haskell category has products, in particular pullbacks, which form the basis of our structural approach.

6 PROCESS ON A TOPOS

The monad and comonad processes are applied to a topos, defining the structure of the data, to perform the transactions. The design of the processes ia therefore termed Monadic Design. We write the process on a topos as:

$$T : \mathbf{E} \longrightarrow \mathbf{E}'$$

where T is the Cartesian monad $< T, \eta, \mu >$ for a category $\mathbf{E}$ with endofunctor T, that is $GF : \mathbf{E} \longrightarrow \mathbf{E}$, unit of adjunction $\eta : 1 \longrightarrow T$ and unit of multiplication $\mu : T^2 \longrightarrow T$.

The source topos is $\mathbf{E}$ and the target topos is $\mathbf{E}'$, with the topos based on pullbacks, including the pasted types, as described in Section 3. The type (intension) of the source and target is the same but the data values (extension) will vary. Closure is achieved as the type is preserved.

For the running bank example, the Cartesian monad T is the banking system transaction, the source information system is $\mathbf{E}$ and the target information system is $\mathbf{E}'$. There may be more than one adjunction for a monad T, based on a category $\mathbf{E}$. For instance $< F, G, \eta, \epsilon >$ may be one adjunction for $\mathbf{E} \longrightarrow \mathbf{E}'$ with another of $< F_A, G_A, \eta_A, \epsilon_A >$ for $\mathbf{A} \longrightarrow \mathbf{E}$, where $\mathbf{A}$ is a subcategory

of **E**. So a variety of adjunctions may be handled by a single monad, over various subcategories of a particular category. This gives flexibility in handling different data-sets with the same underlying structure.

For the process there will also be a comonad:

$$S : \mathbf{E}' \longrightarrow \mathbf{E}$$

where S is the Cartesian comonad $< S, \epsilon, \delta >$ for a category $\mathbf{E}'$ with endofunctor S, that is $FG : \mathbf{E}' \longrightarrow \mathbf{E}'$, counit of adjunction $\epsilon : S \longrightarrow 1$, counit of multiplication $\delta : S \longrightarrow S^2$.

Categories of algebras can be defined over the monad and comonad. From the algebraic perspective, there are two approaches employing the monad/-comonad as the underlying categories. The category of algebras over a monad is traditionally called its Eilenberg-Moore category [6] ([19] at pp. 139-142). Dually, the Eilenberg-Moore category of a comonad is its category of coalgebras. The subcategory of free algebras is traditionally called the Kleisli category of the monad, as is its dual the subcategory of co-free co-algebras of the comonad ([19] at pp. 147-148). The Kleisli category of a monad transforms a monad into a form more suitable for implementation in a functional language such as Haskell. Compared to the EML form of Mac Lane, Kleisli strength generalises the notion of commutativity and guarantees that products lift to the corresponding Kleisli categories [24]. From the point of view of products, the monads developed to Kleisli strength are applicable in a much wider range of computing applications. Kleisli categories are discussed in more detail in Section 7.

6.1 THE T-ALGEBRA

The T-algebras are one of the algebraic forms resulting from the work of Eilenberg and Moore [6]. Such algebras facilitate changing the definition of a monad and therefore permitting fundamental changes to the operand of our process. For any category **X**, which in our case is a topos **E**, the T-algebra produces a new consistent state of adjunction for a modified intension.

In more detail, applying the T-algebra to a topos **E**, in the monad with adjunction $< GF, \eta, \mu >$, yields a new monad adjunction $< G^T F^T, \eta^T, G^T \epsilon^T F^T >$: $\mathbf{E} \longrightarrow \mathbf{E^T}$; that is a new monad adjunction $F^T \dashv G^T$ is defined to accommodate the changed category $\mathbf{E^T}$. A T-algebra is $< e, h >$ where e is an object

in **E**. The structure map of the algebra is $h : Te \longrightarrow e$ such that the diagrams in Figure 11.19 commute. Beck's Theorem provides rules on which categorial transformations in the T-algebra $X \longrightarrow X^T$ are valid [3]. This is sometimes called PTT (Precise Tripleability Theorem).

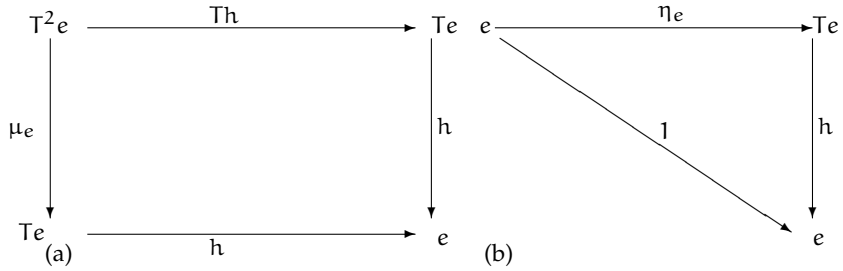


Figure 11.19: T-algebra: (a) Associative Laws, (b) Unitary Laws

7 APPLICATION

The categorial monadic approach is being used for the Blockchain [20], a transaction system, adopted by Bitcoin, for keeping hundreds or even thousands of copies of each transaction record, using multiple transaction logs. The monadic design pattern provides a broad range of transactional semantics with composition the key to scaling any system. The blockchain approach is drawing interest from the established banking industry, where a blockchain is viewed as a shared, encrypted 'ledger' that cannot be manipulated, offering promise for secure transactions [27]. Meredith indicates that compositionality is the key to reliability but offers few details on how this is achieved in the monad. Compositionality is a cornerstone of category theory, defined as a minimum up to some level of isomorphism. In monad/comonad definitions there is the choice of the Mac Lane (EML) or Kleisli algebras as introduced above in Section 6. It is the approach owing to Heinrich Kleisli that has elevated compositionality to a higher level, through the Kleisli lift, described for instance by Mulry [24]. In the diagram in Figure 11.20, H is a monad $< H, \eta, \mu >$ in **X** and K is a monad $< K, \gamma, \rho >$ in **Y**. The Kleisli categories, representing the free algebras, are $\mathbf{X_H}$ and $\mathbf{Y_K}$. The Kleisli lift of functor F is the functor $\bar{F} : \mathbf{X_H} \longrightarrow \mathbf{Y_K}$ such that the diagram in Figure 11.20 commutes. Associated with this diagram is the definition of the lifting natural transformation $\lambda : FH \longrightarrow KF$ in Figure 11.21, derived through applying the interchange law

to the component functors and natural transformations in the two monads defined above.

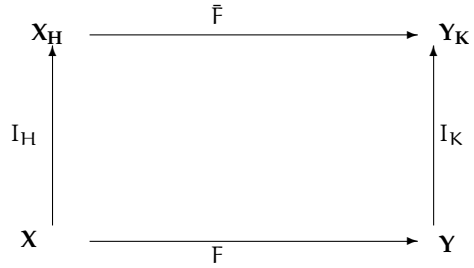


Figure 11.20: Kleisli Lifting of Functor $F : \mathbf{X} \longrightarrow \mathbf{Y}$ to $\bar{F} : \mathbf{X_H} \longrightarrow \mathbf{Y_K}$

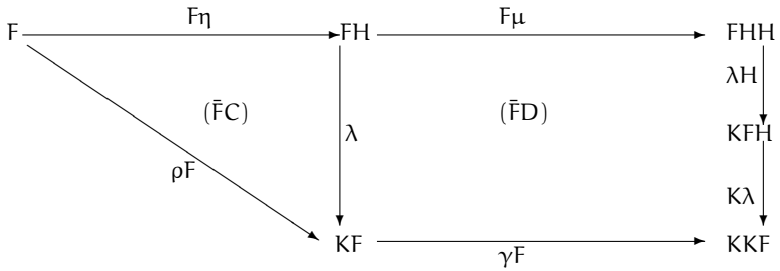


Figure 11.21: Kleisli Lifting of Functor F to $\bar{F}$: the lifting natural transformation $\lambda : FH \longrightarrow KF$

So far the Kleisli lift applies to any category, giving what is termed Kleisli prestrength. We now need to consider the Kleisli lifting of a bicategory, one involving a product of two categories. This is essential if the products are to be well defined for compositional purposes as indicated in Section 6. The lifting gives rise to what is termed Kleisli strength, forming the basis of the Cartesian monad, a term introduced earlier in our overview of the Haskell programming language in Section 5.5. The terms Cartesian monad and strong monad encountered in the literature are for our purposes interchangeable. The enhanced compositionality is achieved firstly by defining a natural transformation $\tau_{A,B} : A \times TB \longrightarrow T(A \times B)$ for objects A, B, C in the category $\mathbf{X}$ with monad $< T, \eta, \mu >$ such that the diagram in Figure 11.22 commutes. A further natural transformation $\lambda_{TA} : I \times TA \longrightarrow TA$ is also defined, as shown in the commuting diagram in Figure 11.23, to reinforce the interchange laws employed in Figure 11.21. Both the diagrams defining the Cartesian monad

involve the Cartesian product, the most relevant for information systems, but the theory is actually more general covering the tensorial (outer) product $A \otimes B$, which may have more relevance for studies involving vectors. Further diagrams are required when the product is tensorial, rather than Cartesian, involving multiplication through the arrow $\mu_{A\times B}$ and associativity though the arrow $\tau_{A,B\times C}$. A major advantage of Kleisli strength monads is that they can, in general, be composed naturally, unlike monads of weaker strength. Such composability increases reliability and scalability, both of which are vital for large scale information systems. Kleisli strength facilitates the discovery of distributive laws.

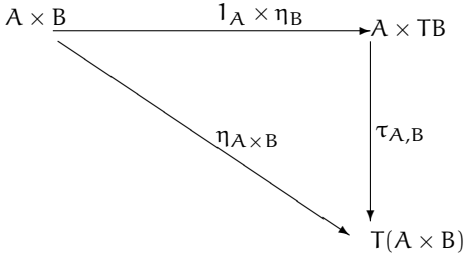


Figure 11.22: Cartesian Monad: Diagram defining the natural transformation $\tau_{A,B}$

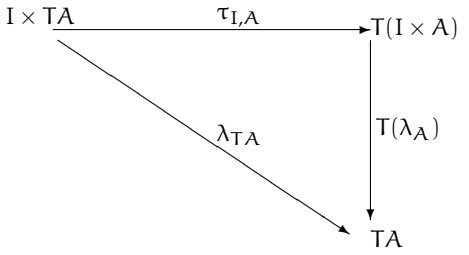


Figure 11.23: Cartesian Monad: Diagram defining the natural transformation λ_{TA}

Meredith [20] envisages that the monadic design patterns, providing a broad range of transactional semantics, would have a front-end data sublanguage of the applied π−calculus, a compositional process calculus developed for concurrent programming by Milner [21]. However, other presentational

techniques from category theory are available, such as bigraphs, and should also be evaluated before a choice is made.

In the functional programming language Haskell, monadic design patterns are employed. The design pattern for a category **C** is $\mathbf{H} =< H, \eta, \mu >$ where **H** is the monad with type constructor H, η is a return function, $\mu : HHA \longrightarrow A$ is a join function. In more conventional monad terminology H is the endofunctor, η the unit of adjunction and μ the multiplication [24]. If the monad is of the *Maybe* type, there are facilities for exception handling. To facilitate monad composition, the monad is lifted into a Kleisli category, with the power of a strong monad or a Cartesian monad. A monad composition operator, also known as the Kleisli composition operator, is available for composing one monad with another naturally [5].

Returning to our banking example we can see that composition of processes is readily available if our monads are Cartesian, with the Kleisli lift. So for two monads $\mathbf{T} =< T, \eta, \mu >$ and $\mathbf{U} =< U, \gamma, \rho >$, we can write UT for the composite process, where say T is the banking transaction with checks for its feasibility and U is a task establishing remote mirror facilities, as in distributed data recovery systems, for recording the results persistently. Such compositionality could be enforced over large distributed systems by involving many individual monads. So monads can be used either in the small individually in a local environment or, through composition, in the whole universe of the information system. The efficacy of the monad approach can be proven through category theory, thereby increasing the reliability and robustness of a system, where every transaction is critical. Further the monad can be directly implemented in the programming language Haskell, enabling experimental results to be derived.

8 SUMMARY

The combination of the topos, as the underlying data-type, and the monad, as the process or transformer, appears to satisfy the requirements of information systems. The topos is based on pullbacks, which can be nested recursively or pasted together for complex relationships. The bottom level of Dolittle diagrams holds both the intension and extension for the data held. Data normalisation arises naturally through the rules of pullback construction. The subobject classifier of a topos facilitates internal queries on the information system. The monad is defined as three components for operations on a category: an endofunctor that is often an adjunction, the unit of adjunction and the unit

of multiplication. There are two main approaches for applying the monad as an algebra: Eilenberg-Moore (EML) and Kleisli. The Kleisli approach finds favour, with its lift to Cartesian monads handling products, providing compositionality across a succession of monads and a route for experimental implementation in Haskell.

9 BIBLIOGRAPHY

[1] Adámek, Jiří, Herrlich, Horst; Strecker, George E, Abstract and Concrete Categories, John Wiley (1990). Recent edition at http://katmat.math.uni-bremen.de/acc (2005).

[2] Banach, R, Regular Relations and Bicartesian Squares, Theoretical Computer Science **129**(1) 187-192 (1994). https://doi.org/10.1016/0304-3975(94)90086-8

[3] Beck, Jonathan Mock, Triples, Algebras and Cohomology, Reprints in Theory and Applications of Categories, Columbia University PhD thesis, **2** 1-59, MR 1987896, originally published 1967 (2003). http://www.tac.mta.ca/tac/reprints/articles/2/tr2abs.html

[4] Bunsen, Christian Karl Josias, Freiherr von; Hare, Julius Charles & Bernays, Jacob, Hippolytus and his Age, published Longman, Brown, Green, & Longmans, London 577 pp (1854).

[5] Diehl, Stephen, Monads made difficult. http://www.stephendiehl.com/posts/monads.html

[6] Eilenberg, Samuel, & Moore, John C, Adjoint Functors and Triples, Illinois J Math **9**(3) 381-398 (1965). http://projecteuclid.org/euclid.ijm/1256068141.

[7] Freyd, Peter, & Scedrov, Andre, Categories, Allegories. Mathematical Library **39** North-Holland (1990).

[8] λ—Haskell: an advanced, purely functional programming language (2017) https://www.haskell.org/

[9] Heather, Michael, & Rossiter, Nick, Logical Monism: The Global Identity of Applicable Logic, Advanced Studies in Mathematics and Logic **2** 39-52 (2005). http://nickrossiter.org.uk/process/advstudiesmathsmonism.pdf

[10] Heather, Michael, & Rossiter, Nick, The Process Category of Reality, ANPA 31, Cambridge 224-262 (2011). http://nickrossiter.org.uk/process/anpa0911.pdf

[11] Johnson, M & Rosebrugh, R, Sketch Data Models, Relational Schema and Data Specifications. Electron Notes Theor Comput Sci **61** 51-63 (2002). http://www.mta.ca/~rrosebru/articles/sdmrsds.pdf

[12] Kent, William, A Simple Guide to Five Normal Forms in Relational Database Theory, Communications of the ACM 26(2) 120-125 (1983). http://www.bkent.net/Doc/simple5.htm

[13] Lambek, J, & Scott, P J, Introduction to Higher Order Categorical Logic, Cambridge (1986). https://github.com/Mzk-Levi/texts/blob/master/Lambek%20J.,%20Scott%20P.J.%20Introduction%20to%20Higher%20Order%20Categorical%20Logic.pdf

[14] Lawvere, F W, Adjointness in Foundations, Dialectica **23** 281-296 (1969).

[15] Leinster, Tom, Higher Operads, Higher Categories, London Mathematical Society Lecture Note Series 298, Cambridge (2004).

[16] Leibniz G W, Monadologie 1714; translated by Nicholas Rescher, 1991. The Monadology: An Edition for Students. University of Pittsburgh Press. Ariew and Garber 213, Loemker S67, Wiener III.13, Woolhouse and Francks 19. Online translations: Jonathan Bennett's translation; Latta's translation; French, Latin and Spanish edition, with facsimile of Leibniz's manuscript at the Wayback Machine (archived July 4, 2012); further editions établie par E Boutroux, Paris LGF 1991; Lamarra, A, Contexte GènGètique et Première Réception de la Monadologie, Revue de Synthese 128 311-323 (2007).

[17] Levene, Mark, & Vincent, Millist W, Justification for Inclusion Dependency Normal Form, IEEE Transactions on Knowledge and Data Engineering **12**(2), pp. 281-291 (2000). http://eprints.bbk.ac.uk/196/1/Binder1.pdf

[18] Lipovača, Miran, Learn You a Haskell for Great Good!, A Beginner's Guide, William Pollock, San Francisco (2011).

[19] Mac Lane, Saunders, Categories for the Working Mathematician, 2nd ed, Springer (1998).

[20] Meredith, Lucius Greg, Monadic Design Patterns for the Blockchain, DEVCON1, Ethereum Developer Conference, Gibson Hall, London, 9-13 Nov (2015). https://www.youtube.com/watch?v=uzahKc_ukfM&feature=youtu.be

[21] Milner, Robin, Communicating and Mobile Systems: The π−calculus, Cambridge (1999).

[22] Moggi, Eugenio, Computational Lambda-Calculus and Monads, Proceedings of the Fourth Annual Symposium on Logic in Computer Science 14-23 (1989).

[23] Moggi, Eugenio, Notions Of Computation And Monads, Information And Computation **93** 55-92 (1991).

[24] Mulry, Philip, Notions of Monad Strength, Banerjee, A, Danvy, O, Doh, K-G, Hatcliff, J, (edd.) David A. Schmidt's 60th Birthday Festschrift, EPTCS **129** 67-83, doi:10.4204/EPTCS.129.6 (2013). https://arxiv.org/pdf/1309.5132.pdf

[25] ncatlab, 0-category https://ncatlab.org/nlab/show/0-category

[26] ncatlab, Pullback as an Equalizer https://ncatlab.org/nlab/show/pullback

[27] Phys Org, Bitcoin's 'blockchain' tech may transform banking, http://phys.org/news/2015-12-bitcoin-blockchain-tech-banking.html

[28] Rossiter, B N, Heather, M A, & Sisiaridis, D, Process as a World Transaction, Proceedings ANPA 27 Conceptions, 122-157 (2006). http://nickrossiter.org.uk/process/anpa064.pdf

[29] Rossiter, Nick, & Heather, Michael, Formal Natural Philosophy: Top-down Design for Information & Communication Technologies with Category Theory, ANPA 35, Explorations, Grenville J Croll, Nicky Graves Gregory (edd.), 155-193 (2015). http://nickrossiter.org.uk/process/anpa-2015-a5-Latex.pdf

[30] Rossiter, Nick, & Heather, Michael, Abstract Relations and Allegorical Categories, ANPA 36, Explorations II, Anton L. Vrba (ed.) 103-134 (2016). http://nickrossiter.org.uk/process/Rossiter-ANPA-PROC-36updated.pdf

[31] Whitehead, Alfred North, Process and Reality: An Essay in Cosmology, Macmillan, New York (1929); corr.ed., eds. David Ray Griffin and Donald W. Sherburne, New York: Free Press (1978). https://monoskop.org/images/4/40/Whitehead_Alfred_North_Process_and_Reality_corr_ed_1978.pdf

3-DIMENSIONALITY

PETER ROWLANDS

Department of Physics, University of Liverpool,
Oliver Lodge Laboratory, Oxford St, Liverpool, L69 7ZE, UK.
email: p.rowlands@liverpool.ac.uk

ABSTRACT. 3-dimensionality is one of the most important and profound ideas in the whole of physics. The fact that space is 3-dimensional gave us a special insight into Nature from an early period because it set us the problem of explaining why this counterintuitive structure was so prevalent in natural systems. Part of the answer was revealed by the discovery of quaternions by Hamilton in 1843, although this spectacular insight was not followed up with the thoroughness it deserved. Since then we have found that Nature requires two 3-dimensionalities to define all the important structures in physics, chemistry and biology, as well as those of algebra and geometry. Here, we will largely discuss the importance of 3-dimensionality in physics.

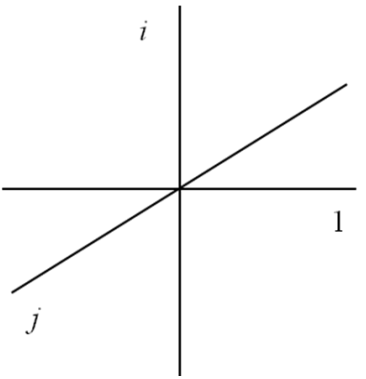
Introduction

3-dimensionality is one of the most important concepts in physics. It is also one of the most important breakthroughs in the whole of human knowledge. It presents us with an extraordinary and non-obvious fact which challenges interpretation. That space is 3-dimensional has probably been known in some form for millennia. But, as we shall see, the concept extends beyond space, and the spatial example leads us to other ones which are equally important in physics. In addition to the fact that 3-dimensionality has such a special and peculiar signature that it provides one of the main access routes to unravelling Nature's most fundamental secrets, it is also important for

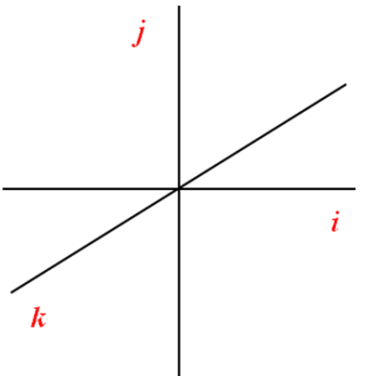
many other reasons. As I hope to show, it is the only source for discreteness and localisation in Nature. Without it, measurement and observation would be impossible. And it appears in many guises. It is a ubiquitous condition that extends far beyond its appearance in the structure of space.

Though it is fashionable to be cavalier about dimensions and to act as though they can be invented and removed at will – as well as distorted and modified – to explain aspects of physics, it is an inescapable fact that space, the only physical parameter that gives us a direct way of appreciating the physical world through measurement, is most definitely 3-dimensional. And this 3-dimensionality cannot be extended to, say, 9-D, though it could be embedded in a higher dimensionality including other things.

Much has been said about the difficulty of explaining 3-dimensionality, but, in fact, it isn't difficult at all. We have known the reason why 3-dimensionality is special for more than 150 years. This happened when Sir William Rowan Hamilton was trying to extend the '2-D' complex number system to explain 3-D space.



'triads'



quaternions

Hamilton's attempted extension of the Argand diagram to include a second imaginary axis (j) in a system of 'triads' failed because there was no meaning for the product $\boldsymbol{ij}$ within the system. There was no closure in the algebra, but eventually by positing three cyclic imaginary dimensions ($\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}$) alongside the real one (which could not be shown on the diagram), he was able to generate a division algebra that was closed and consistent.

The need for a real product means that the 3-D quaternions are in some way part of a larger 4-dimensionality as their name implies. You can't even define the 3 imaginary units without invoking the real one (1), and each of the 3 imaginary axes is in some sense orthogonal to the real one that cannot be drawn on the same diagram. But this one is not like the other 3. It cannot be totally absorbed into 3 axes, which are effectively, in the pure mathematics, interchangeable with each other. The relationship is a broken one like the real and imaginary axes of complex numbers.

Quaternions, significantly, are anticommutative, with the multiplication rules:

$$\boldsymbol{ij} = -\boldsymbol{ji} = \boldsymbol{k}; \qquad \boldsymbol{jk} = -\boldsymbol{kj} = \boldsymbol{i}; \qquad \boldsymbol{ki} = -\boldsymbol{ik} = \boldsymbol{j};$$
$$\boldsymbol{i}^2 = \boldsymbol{j}^2 = \boldsymbol{k}^2 = -\boldsymbol{ijk} = 1.$$

This becomes apparent when we multiply $\boldsymbol{ij}$ by $\boldsymbol{ji}$, for

$$\boldsymbol{ij}\,\boldsymbol{ji} = -\boldsymbol{ii} = 1$$

not –1. The significance of this is that it explains 3-dimensionality, for $\boldsymbol{i}$, $\boldsymbol{j}$ and $\boldsymbol{ij}$ form a closed set, which cannot be extended (unless antiassociativity is additionally introduced – and then only in the singular case of octonions). It also leads to a cross-product, which is characteristic of 3-, but not higher, dimensionality. In simple terms,

you can't make anticommutativity work if you have more than 3 terms. In the absence of antiassociativity,

anticommutativity ≡ 3-dimensionality

The '3-ness' is thus not the primary cause of the 3-dimensionality of space. It is simply a result of anticommutativity. If we have two axes, ***i*** and ***j***, that are anticommutative with each other, then we cannot draw any other axis that is anticommutative with them, unless it is ***ij***, which we also call ***k***. Anticommutativity forces 3-dimensionality. The strange arbitrariness of the number 3 is explained. Commutative things, of course, can be defined to infinity. If ***i*** and ***j*** were commutative, we could have ***i***, ***j***, ***k***, ***l***, ***m***, etc. without limit. In physical terms, anticommutative things 'know' about each other's presence and have to act accordingly; commutative things do not.

Quaternions are known as one of the four division algebras in mathematics. The only higher dimensional one is octonions with 7 imaginary components and one real, compared to the 3 imaginary and one real of quaternions, and octonions are antiassociative as well as anticommutative. So, for example, $a(bc) = -(ab)c$.

real numbers	1	commutative	associative
complex numbers	2	commutative	associative
quaternions	4	anticommutative	associative
octonions	8	anticommutative	antiassociative

Significantly, the kind of dimensionality found in space, which is characteristic of that of a *nonconserved* quantity, with an affine structure and rotation symmetry between the dimensions, can only exist for 3 dimensions unless the system is also antiassociative. Even then it can only exist for 7 dimensions. *There is no extension of space as we know it to 9 dimensions.* The 3 dimensions of space could be

embedded in a higher dimensionality but they cannot be extended to become one. It is mathematically impossible. And the application of division algebras to space is crucial for physical purposes, where measurement and counting are fundamental processes. (Rowlands, 2015c)

The algebra of space

Quaternions are the only *pure* 3-D system. The 3 cyclic products produce each other and nothing else. But space is *not* pure 3-D. We can go back to Hamilton to find out why. Hamilton early on decided to extend quaternions by complexifying them. So he took the product

$$(1, i) \times (1, \boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k})$$

where i is the ordinary (complex) square root of –1. Ordinary complex numbers are, of course, distinct from quaternions in being commutative. Our base set is now 1, i, $\boldsymbol{i}$, $\boldsymbol{j}$, $\boldsymbol{k}$, and, multiplying everything out, we will also generate terms like $i\boldsymbol{i}$, $i\boldsymbol{j}$, $i\boldsymbol{k}$. For reasons that will soon become clear, we will write $i\boldsymbol{i} = \mathbf{i}$, $i\boldsymbol{j} = \mathbf{j}$, $i\boldsymbol{k} = \mathbf{k}$.

If we take the products of these terms, we can write:

$$(i\boldsymbol{i})^2 = (i\boldsymbol{j})^2 = (i\boldsymbol{k})^2 = -i(i\boldsymbol{i})\,(i\boldsymbol{j})\,(i\boldsymbol{k}) = 1$$
$$(i\boldsymbol{i})\,(i\boldsymbol{j}) = i(i\boldsymbol{k})$$
$$(i\boldsymbol{k})\,(i\boldsymbol{i}) = i(i\boldsymbol{j})$$
$$(i\boldsymbol{j})\,(i\boldsymbol{k}) = i(i\boldsymbol{i})$$

We can also write these relations in the form:

$$\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = -i\mathbf{ijk} = 1$$
$$\mathbf{ij} = i\mathbf{k}$$
$$\mathbf{ki} = i\mathbf{j}$$
$$\mathbf{jk} = i\mathbf{i}$$

The complexified quaternion units $i\mathbf{i} = \mathbf{i}$, $i\mathbf{j} = \mathbf{j}$, $i\mathbf{k} = \mathbf{k}$ are, of course, anticommutative in exactly the same way as ordinary quaternion units, but we now notice an extra feature, the *i* term outside the bracket that has appeared on the right-hand side of the equations. The terms we have been using here, $i\mathbf{i} = \mathbf{i}$, $i\mathbf{j} = \mathbf{j}$, $i\mathbf{k} = \mathbf{k}$, are easily identifiable as the units of the Clifford algebra of 3-D space. They are also called multivariate vector units (Hestenes, 1966) and are isomorphic to the Pauli spin matrices. I will refer to them simply as vector units.

They are the true algebra of 3-D space, but they are not a pure 3-D algebra, because the products of **ij**, etc., are not themselves vector units but *pseudovectors*, such as $i\mathbf{k}$. These are mathematically the same as quaternions and are seen in physical concepts such as area and angular momentum. In addition the triple product **ijk** is a pseudoscalar *i*, as with volume, if we use pure space units.

In principle, the full algebra of vector space requires us to use a set of 8 base units, which includes

vectors	**i**, **j**, **k**
pseudovectors = quaternions	$i\mathbf{i}$, $i\mathbf{j}$, $i\mathbf{k}$
pseudoscalar = complex term	*i*
scalar	1

Including + and – values, this requires 16 base units, as opposed to 8 for quaternions and 4 for complex numbers. We also note that quaternions, complex numbers and real (scalar) numbers are subalgebras of this one.

The fact that space isn't a 'pure' 3-D means that complex numbers have to be taken into account when we are discussing its properties, perhaps also explaining how a complex variable like time can be absorbed relatively naturally into the '4-vector' combination of space and time. If we take the complexified version of the quaternion base

units 1, ***i***, ***j***, ***k***, we naturally produce the 4-vector units *i*, **i**, **j**, **k**. Even though the *i* term here refers in the first instance to volume, rather than time, the commutativity of the relation between *i* and the quaternion units means that the physical nature of the unit doesn't need to be specified.

The existence of a complexified quaternion system for describing one of the fundamental physical parameters suggests the question, never fully answered: Is there any role for pure quaternions at the fundamental level in physics? Well, we have had 150 years of extraordinarily vituperative anti-quaternion propaganda obscuring the great simple fact that quaternions have solved the problem of 3-dimensionality. Even if the absurdity of this is conceded, there is always the fallback position that quaternions have never found any physical role of their own.

There is no doubt that the quaternion system underpins the 3-D properties of space, and this has been emphasized by the subsequent discoveries of quantum mechanical spin and the special relativistic connection of space and time, both of which it can be said to have predicted. However, it would seem strange if such a perfect system should only have such an indirect manifestation. Do quaternions have any more direct manifestation in physics?

This is where we need to introduce the Klein-4 parameter group, which I have discussed many times previously. (Rowlands, 1983, 1991, 2001, 2007, 2009, 2014, 2015a) There is no doubt that the quaternion system underpins the 3-D properties of space, and this has been emphasized by the subsequent discoveries of quantum mechanical spin and the special relativistic connection of space and time, both of which it effectively predicted.

A group of order 4

The basis of this is that Nature exhibits a zero totality, which, at the fundamental level, is manifested through a symmetry between the parameters space, time, mass and charge. The structure is that of a Klein-4 or D_2 group which can be represented in many different ways. Strikingly, it is full of manifestations of 3-dimensionality or anticommutativity, in addition to numerous dualities or summations to zero.

mass	conserved	$\mathbb{R}$	continuous (1-D)	commutative
time	nonconserved	$\mathbb{I}$	continuous (1-D)	commutative
charge	conserved	$\mathbb{I}$	discrete (3-D)	anticommutative
space	nonconserved	$\mathbb{R}$	discrete (3-D)	anticommutative

(Legend: $\mathbb{R}$ real , $\mathbb{I}$ imaginary)

In a more extensive specification, the properties / antiproperties would become:

mass:	conserved	real	continuous
	identity	norm 1	nondimensional
	translation asymmetric	global	commutative
time:	nonconserved	imaginary	continuous
	no identity	norm –1	nondimensional
	translation symmetric	global	commutative
charge:	conserved	imaginary	discrete
	identity	norm –1	dimensional
	translation asymmetric		local

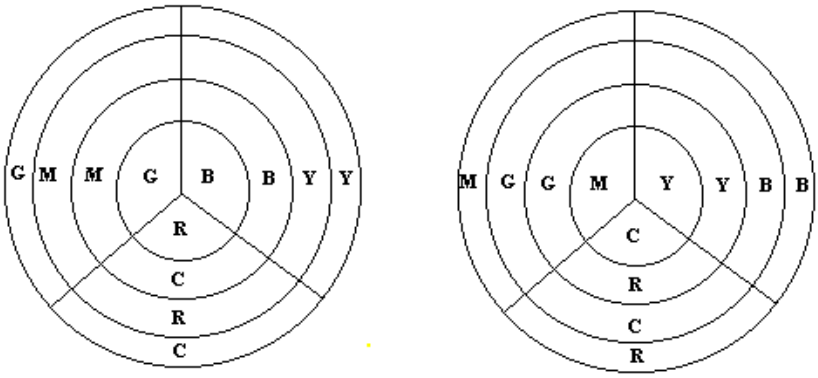
	rotation asymmetric		anticommutative
space:	nonconserved	real	discrete
	no identity	norm 1	dimensional
	translation symmetric		local
	rotation symmetric		anticommutative

Using the symbols, *x*, *y* and *z*, to represent the properties, with –*x*, –*y* and –*z* standing for the exactly opposite 'antiproperties', indicates that this symmetry incorporates a conceptual zero:

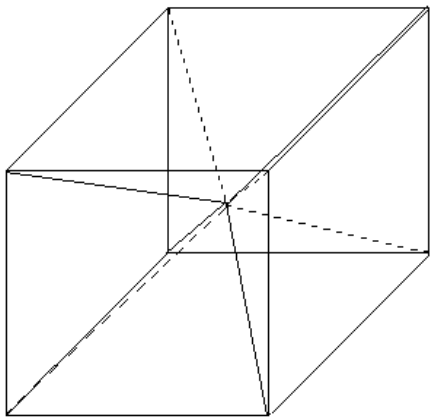
mass	*x*	*y*	*z*
time	–*x*	–*y*	*z*
charge	*x*	–*y*	–*z*
space	–*x*	*y*	–*z*

The 3-dimensionality inherent in this structure is inherent in the use of 3 properties and 3 symbols, *x*, *y* and *z*, and can be seen in the use of 3-dimensional representations, and an analogous use of 3 colours. (Rowlands, 2003, 2007, 2014, 2015a)

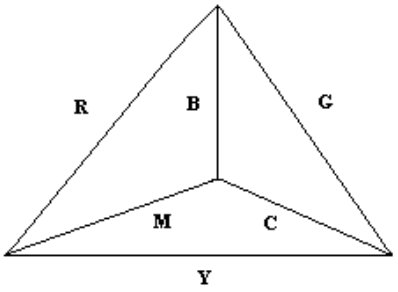
The colour representation uses the 3 primary colours (R, G, B) as properties and the 3 secondary colours (C, M, Y) as antiproperties or vice versa, with colourless (white) summation of each sector. Concentric circles represent the parameters in any order.



The 3-D representation plots the properties as *x*, *y*, *z* and the antiproperties as *–x*, *–y*, *–z* to four corners of a cube from its centre. The dual group formed by reversing one property / antiproperty combination is represented by four dotted lines.



The tetrahedral representation uses the primary and secondary colours as in the colour representation. The four parameters appear as either the four faces or four vertices, with the edges representing the six properties / antiproperties.

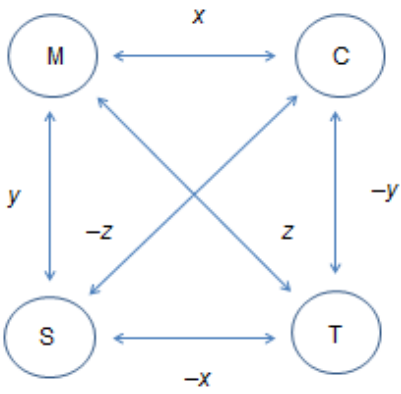


Now, both the set of parameters, and consequently all of physics, can be reduced to algebra, because the characteristics of the parameters can be defined entirely through the algebras that represent them.

mass	scalar	1
time	pseudoscalar (complex)	i
charge	quaternion (pseudovector)	$\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}$
space	vector	$\mathbf{i}, \mathbf{j}, \mathbf{k}$

This is a kind of ontological ordering because the complexity increases at each level. Mass(-energy), time and space have the algebras that we familiarly associate with them, or that, in the case of time, are derived from developments in physics.

Yet another representation is a ‘graph’ which also suggests an interpretation in category theory, which seems to apply to both the parameter group of mass, charge, space and time, and the underlying universal rewrite system on which it is based. This representation again makes use of the algebraic representation of properties and antiproperties:



A related diagram was drawn by Vanessa Hill for my video-recorded lectures on *The Foundations of Physical Law*, based on a rough sketch provided by myself.

The algebra of charge

The algebra that is least familiar in this set is that of charge, which fills what would otherwise be a gap in the ontological sequence. The reason why this is less familiar is because the concept of 'charge' has only been slowly extended to incorporate the sources of strong and weak, as well as electrical, interactions. I can certainly remember this usage, now so relatively familiar, being met with incomprehension thirty and even twenty years ago. Even so, the use of the word is to mean a simple fundamental thing is still somewhat equivocal.

Wikipedia, for example, says that: 'Various charge quantum numbers have been introduced by theories of particle physics. These include the charges of the Standard Model: The color charge of quarks. The color charge generates the *SU*(3) color symmetry of quantum chromodynamics. The weak isospin quantum numbers of the electroweak interaction. It generates the *SU*(2) part of the electroweak

$SU(2) \times U(1)$ symmetry. Weak isospin is a local symmetry, whose gauge bosons are the W and Z bosons. The electric charge for electromagnetic interactions. In mathematics texts, this is sometimes referred to as the u_1-charge of a Lie algebra module.' The article also says that, in gravitation, charges are 'Eigenvalues of the energy-momentum tensor correspond to physical mass.'

The basic idea is that 'charges' are the sources of the four known fundamental interactions. However, the idea that there is some common structure is never stated explicitly and the charge concept is only introduced in roundabout ways, and seems a little confused, especially in its lack of reference to the coupling constants for the 3 interactions.

Why is this? The simple answer is that the four known interactions are based on different symmetry groups and have different physical manifestations, and that, in addition, there are fundamental differences between gravity and the other 3 (gauge) interactions.

The problem might be considered insuperable but for two things:

(1) Whatever extra complications are manifested, all four forces have a 'Coulomb' or inverse square component, and that, in the case of gravity, it is attractive for like particles, whereas like particles repel in the case of all gauge forces. The Coulomb component in the case of the gauge forces determines the quantity known as the coupling constant.

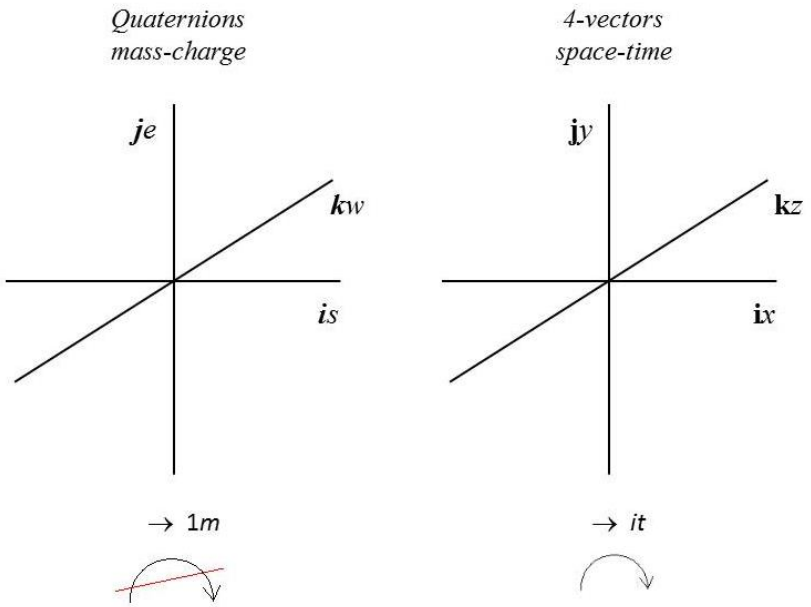
(2) There appears to be an energy regime where the differences between the 3 gauge forces disappear and the coupling constants become equal. This Grand Unification energy is close to the Planck mass, which would link gravitational with gauge values. It is also possible that the character of the gauge forces would also become purely Coulombic at this energy.

So, we are entitled to propose a working hypothesis in which the gauge forces stem from a 3-component or 'dimensional' quantity

charge, which is imaginary, and therefore quaternionic, in nature, to explain the difference in sign displayed by forces between like imaginary charges and like real masses. In addition, as Kant showed in the eighteenth century, inverse-square forces are direct manifestations of the 3-dimensionality of space, and, in a more modern perspective, are manifestations of the symmetry surrounding a point source. In this case, the parameters mass and charge could be manifestations of a quaternion structure, with three imaginary parts and one real, symmetric to the 4-vector structure of space and time, with three real parts and one imaginary.

space	time	charge	mass
$\mathbf{i}x\ \mathbf{j}y\ \mathbf{k}z$	it	$\boldsymbol{i}s\ \boldsymbol{j}e\ \boldsymbol{k}w$	$1m$

The 3-dimensional parts of the quaternions and 4-vectors, which are respectively irrotational and rotational, can be represented on mutually orthogonal axes, from which the fourth parts are excluded. The total structure has interesting parallels to an octonion and can be written in an octonion form, with the antiassociative aspects seemingly excluded from real physical manifestations.



The basic quaternion structure explains the coupling constant and the Coulombic term for each of the forces. But we still have to explain why the extra terms appear which make the 3 gauge interactions appear so different from each other at energies below Grand Unification. As we will show this is possible in a fundamental way, but it requires the interaction between *two* separate 3-dimensionalities, an interaction which explains why we have interacting point-particles at all.

The application of a 3-dimensional structure to charge as well as to space in the parameter group suggests that only 3-D concepts can be discrete, and that all discreteness in Nature comes from 3-dimensionality, as mass and time are in their different ways continuous. The discreteness of charge is obvious because it is a conserved quantity, but that of space is more subtle because it is nonconserved and so *has no fixed units*. It is, however, necessary to

explain space as an observable quantity and also to explain Zeno's paradoxes.

It is, in fact, impossible to imagine a non-dimensional or continuous quantity as discrete. Dimensionality is needed to any method of division we can imagine, and to explain how we have zeros or crossover points. A discrete 'point' requires 3-D for its existence, and, as we will show, it requires 2 × 3-D. In addition, there is a natural relation between discreteness and anticommutativity which we see in physics. But, according to the universal rewrite system (Rowlands and Diaz, 2002, Marcer and Rowlands, 2014a,b, Rowlands, 2007, 2010, 2014), anticommutativity is also the origin of discreteness in mathematics in its creation of the idea of a repeating series of closed entities.

A double space

Suppose we set out the entire algebra and subalgebras of each of the four parameters in the Klein-4 group:

space	vector	**i**, **j**, **k**, *i***i**, *i***i**, *i***k**, *i*, 1
		i, **j**, **k**, ***i***, ***j***, ***k***, *i*, 1
charge	quaternion (pseudovector)	*i****i***, *i****j***, *i****k***, *i*, 1
		i, ***j***, ***k***, 1
time	pseudoscalar (complex)	*i*, 1
mass	scalar	1

We see that immediately that charge, mass and time have algebras equivalent to the subalgebras of space. But if we put them altogether, they create an alternative algebra identical to that of Space.

combination	vector	**i**, **j**, **k**, *i***i**, *i***i**, *i***k**, *i*, 1
		i, **j**, **k**, ***i***, ***j***, ***k***, *i*, 1
charge	quaternion (pseudovector)	*i****i***, *i****j***, *i****k***, *i*, 1
		i, ***j***, ***k***, 1
time	pseudoscalar (complex)	*i*, 1
mass	scalar	1

Since, in the group, they total to zero, we may see the combination as a kind of 'antispace' or alternative space to space itself. Note that its dimensionality comes from that of charge. If we assume these parameters are the only sources of physical knowledge, then the zero totality looks like a double space, whose two halves (with respective units represented, for convenience, in lower case symbols and capitals) mirror each other in some way.

space	vector	**i**, **j**, **k**, *i***i**, *i***i**, *i***k**, *i*, 1
		i, **j**, **k**, ***i***, ***j***, ***k***, *i*, 1
combination	vector	**I**, **J**, **K**, *i***I**, *i***J**, *i***K**, *i*, 1
		I, **J**, **K**, ***I***, ***J***, ***K***, *i*, 1

Let us assume that these parameters are the whole of physics. How do we combine them? By combining the algebras.

Time	Space	Mass	Charge
i	**i j k**	1	***i j k***
pseudoscalar	vector	scalar	quaternion

Working out every possible combination of the four requires 64 units. This turns out to be the algebra of the Dirac equation, the relativistic quantum mechanical equation of the fermion, the only true

fundamental object that we know must exist. The result is another group, this time of order 64, rather than 4:

1	i				–1	$-i$			
ii	**ij**	**ik**	$i\boldsymbol{k}$	$\boldsymbol{j}$	**–ii**	**–ij**	**–ik**	$-i\boldsymbol{k}$	$-\boldsymbol{j}$
ji	**jj**	**jk**	$\boldsymbol{ji}$	$\boldsymbol{k}$	**–ji**	**–jj**	**–jk**	$-\boldsymbol{ji}$	$-\boldsymbol{k}$
ki	**kj**	**kk**	$\boldsymbol{kj}$	$\boldsymbol{i}$	**–ki**	**–kj**	**–kk**	$-\boldsymbol{kj}$	$-\boldsymbol{i}$
$i\boldsymbol{i}$	$i\boldsymbol{j}$	$i\boldsymbol{k}$	i	**j**	$-i\boldsymbol{i}$	$-i\boldsymbol{j}$	$-i\boldsymbol{k}$	$-i$	**–j**
$i\boldsymbol{ji}$	$i\boldsymbol{jj}$	$i\boldsymbol{jk}$	i	**k**	$-i\boldsymbol{ji}$	$-i\boldsymbol{jj}$	$-i\boldsymbol{jk}$	$-i$	**–k**
$i\boldsymbol{ki}$	$i\boldsymbol{kj}$	$i\boldsymbol{kk}$	**ij**	**i**	$-i\boldsymbol{ki}$	$-i\boldsymbol{kj}$	$-i\boldsymbol{kk}$	**–ij**	**–i**

We could also use the two ‘spaces’ as base units:

Space	Time-Mass-Charge
i j k	**I J K**
vector	vector

Here, we simply replace the quaternion units in the previous table (***i***, ***j***, ***k***) with the second set of vector units (**I**, **J**, **K**).

Groups do not need to be specified by all their elements. A small number of elements multiplied out can often generate the entire group. Here we only need 5 generators. 5 is a very significant number because it is the point in physics and mathematics at which symmetries become broken and chirality forces itself upon us. Thus, even though we could have generated the group from the six components **i**, **j**, **k**, **I**, **J**, **K**, or **i**, **j**, **k**, ***i***, ***j***, ***k***, which would have maintained perfect symmetry between the two ‘spaces’, the *minimum number of generators*, which is always what nature requires, forces us to break it.

The group can be structured as being made up of 4 complex number units and 12 sets of 5 generators, *any one of which*, for example,

ii ***ij*** ***ik*** $i\boldsymbol{k}$ $\boldsymbol{j}$

could be used to produce the entire group. Even this does not exhaust the options. The 60 units could be organized as 12 sets of 5 generators in at least 64 different ways (with 4 independent options for sign changes, with changes of vectors and quaternions, and with real and imaginary units). The 60 is reminiscent of the 60 units of carbon in the C60 molecule, which are arranged in 12 pentagons, separated by hexagons which contain no extra atoms – in relation to the widespread appearance of structures of this type in biology and chemistry, as well as in physics, this is no coincidence. The symbols are, of course, arbitrary, and we could adopt any of the 64 × 12 sets of 5 available to us to represent physics, but it is most convenient to choose the set that is closest to our already accepted conventions.

The set of 5 generators is not unique (e.g. any row of the table of 64 can be used), but all sets follow the same pattern. Typically, beginning with

Time	Space	Mass	Charge
i	$\mathbf{i}\ \mathbf{j}\ \mathbf{k}$	1	$\boldsymbol{i}\ \boldsymbol{j}\ \boldsymbol{k}$

we take one of each of the charge units $\boldsymbol{i}\ \boldsymbol{j}\ \boldsymbol{k}$ on to each of the units of the other three parameters, to create:

$$i\boldsymbol{k} \qquad \mathbf{i}\boldsymbol{i}\ \mathbf{j}\boldsymbol{i}\ \mathbf{k}\boldsymbol{i} \qquad 1\boldsymbol{j}$$

The symmetry of one space $\mathbf{i}\ \mathbf{j}\ \mathbf{k}$ or the other $\boldsymbol{i}\ \boldsymbol{j}\ \boldsymbol{k}$ has to be broken.

Because space is a nonconserved quantity, and its component dimensions **i**, **j**, **k** can't be uniquely identified or distinguished from each other, the broken symmetry in physics becomes that of charge $\boldsymbol{i}$, $\boldsymbol{j}$, $\boldsymbol{k}$. (*Physics* requires this, though the symbols could be interchanged.) We can then show that 5 generators leads, among many other things,

to the broken symmetry between the 3 charges which has troubled physics for more than forty years. This affects the nature of the charges as we observe them, and we can begin to recognise here the respective characteristics of weak, strong and electric charges, governed by the respective *SU*(2), *SU*(3) and *U*(1) symmetries.

The charges adopt characteristics of the mathematical objects they are connected to. So we find that the charge we have represented by ***k*** becomes associated with a pseudoscalar quantity *i*; that the one represented by ***i*** becomes associated with three vector units **i**, **j** and **k**; and the one represented by ***j*** becomes associated with the scalar unit 1.

ik	***ii ji ki***	1***j***
weak	strong	electric
pseudoscalar	vector	scalar
SU(2)	*SU*(3)	*U*(1)

We see also that, though physics might require two ‘spaces’ for its specification, and that, though these two spaces may contain identical information, it presents itself differently within them, through a chirality:

***i*K**	**iI jI kI**	1**J**

And we see that the units of the second ‘space’ **I**, **J**, **K** now adopt the characteristics of the mathematical objects they are connected to:

pseudoscalar	vector	scalar

In addition to affecting charge, the combination also affects time, space and mass. We begin, once again, with

time	space	mass	charge
i	$\mathbf{i}\ \mathbf{j}\ \mathbf{k}$	1	$\boldsymbol{i}\ \boldsymbol{j}\ \boldsymbol{k}$

To create the generators, we distribute the charge units onto the others:

$$i\boldsymbol{k} \qquad\qquad \mathbf{i}\boldsymbol{i}\ \ \mathbf{j}\boldsymbol{i}\ \ \mathbf{k}\boldsymbol{i} \qquad\qquad 1\boldsymbol{j}$$

This creates *new* 'compound' (and 'quantized') physical quantities:

energy	momentum	rest mass
E	$p_x\ \ p_y\ \ p_z$	m

If we regard E, p_x, p_y, p_z, m simply as scalar values or 'coefficients', which are arbitrary in principle, and the algebraic operators $i\boldsymbol{k}$, $\mathbf{i}\boldsymbol{i}$, $\mathbf{j}\boldsymbol{i}$, $\mathbf{k}\boldsymbol{i}$,$1\boldsymbol{j}$ as defining the physical meaning, that is, the nature of the physical quantities involved, then we can package the whole information as

$$(i\boldsymbol{k}E + \mathbf{i}\boldsymbol{i}p_x + \mathbf{j}\boldsymbol{i}p_y + \mathbf{k}\boldsymbol{i}p_y + \boldsymbol{j}m)$$

In physics, this combined object is *nilpotent*, squaring to zero, because

$$(i\boldsymbol{k}E + \mathbf{i}\boldsymbol{i}p_x + \mathbf{j}\boldsymbol{i}p_y + \mathbf{k}\boldsymbol{i}p_y + \boldsymbol{j}m)\ (i\boldsymbol{k}E + \mathbf{i}\boldsymbol{i}p_x + \mathbf{j}\boldsymbol{i}p_y + \mathbf{k}\boldsymbol{i}p_y + \boldsymbol{j}m) = 0 \qquad (1)$$

and we can identify this as Einstein's relativistic energy equation

$$E^2 - p^2 - m^2 = 0$$

or, in its more usual form,

$$E^2 - p^2c^2 - m^2c^4 = 0$$

The nilpotent Dirac equation

The Dirac equation simply quantizes the nilpotent equation, using differentials in time and space applied to a phase factor for E and p. So equation (1) becomes

$$(-\boldsymbol{k}\partial / \partial t - i\boldsymbol{i}\nabla + \boldsymbol{j}m)\,(i\boldsymbol{k}E + \boldsymbol{i}\mathbf{p} + \boldsymbol{j}m)\,e^{-i(Et - \mathbf{p}.\mathbf{r})} = 0,$$

the Dirac equation for a free fermion, by simultaneously applying nonconservation and conservation. The operator, $(-\boldsymbol{k}\partial / \partial t - i\boldsymbol{i}\nabla + \boldsymbol{j}m)$, is like a coding of all the possible space and time variations, which is 'decoded' using a 'phase factor', here $e^{-i(Et - \mathbf{p}.\mathbf{r})}$.

The most complete possible variation in space and time is defined by a phase factor which associates E with time and $\mathbf{p}$ with space, and then using the differentials $\partial / \partial t$ and ∇ to recover $(i\boldsymbol{k}E + \boldsymbol{i}\mathbf{p} + \boldsymbol{j}m)$ from the phase factor. For a free particle, the most complete set of variations in space and time is given by $e^{-i(Et - \mathbf{p}.\mathbf{r})}$, and the expression which will recover $(i\boldsymbol{k}E + \boldsymbol{i}\mathbf{p} + \boldsymbol{j}m)$ using this as a phase factor is $(-\boldsymbol{k}\partial / \partial t - i\boldsymbol{i}\nabla + \boldsymbol{j}m)$. The phase factor $e^{-i(Et - \mathbf{p}.\mathbf{r})}$ is then the minimum expression that the differentials $\partial / \partial t$ and ∇ can be applied to if space and time are to be varied without restriction.

Including all possible sign variations of E and $\mathbf{p}$, we obtain

$$(\mp \boldsymbol{k}\partial / \partial t \mp i\boldsymbol{i}\nabla + \boldsymbol{j}m)\,(\pm i\boldsymbol{k}E \pm \boldsymbol{i}\mathbf{p} + \boldsymbol{j}m)\,e^{-i(Et - \mathbf{p}.\mathbf{r})} = 0$$

which is equivalent to a nilpotent Dirac equation of the form

$$(\mp \boldsymbol{k}\partial / \partial t \mp i\boldsymbol{i}\nabla + \boldsymbol{j}m)\psi = 0.$$

We can also express it in operator form

$$(\pm i\mathbf{k}E \pm \mathbf{i}\mathbf{p} + \mathbf{j}m)\,(\pm i\mathbf{k}E \pm \mathbf{i}\mathbf{p} + \mathbf{j}m)\,e^{-i(Et\,-\,\mathbf{p}.\mathbf{r})} = 0,$$

where the operators E and $\mathbf{p}$ become $i\partial\,/\,\partial t$ and $-i\nabla$ as in the usual canonical quantization.

The complete Dirac wavefunction in nilpotent form provides the particle state (represented by the first term) and the 3 possible states it could become by P, T and C transformations:

$$\begin{aligned}
&(i\mathbf{k}E + \mathbf{i}\mathbf{p} + \mathbf{j}m) \\
&(i\mathbf{k}E - \mathbf{i}\mathbf{p} + \mathbf{j}m) \quad P \\
&(-i\mathbf{k}E + \mathbf{i}\mathbf{p} + \mathbf{j}m) \quad T \\
&(-i\mathbf{k}E - \mathbf{i}\mathbf{p} + \mathbf{j}m) \quad C
\end{aligned}$$

Parity, time reversal and charge conjugation are essentially reversals in the signs associated with space, time and charge.

We can interpret the expression

$$(\pm i\mathbf{k}E \pm \mathbf{i}\mathbf{p} + \mathbf{j}m)\,(\pm i\mathbf{k}E \pm \mathbf{i}\mathbf{p} + \mathbf{j}m) \rightarrow 0$$

as giving us both relativity and quantum mechanics. In quantum mechnaics we take the first bracket as an operator acting on a phase factor. The E and $\mathbf{p}$ terms can include any number of potentials or interactions with other particles. Squaring to 0 gives us the Pauli exclusion principle, because if any 2 particles are the same, their combination is 0.

Fermions appear to be point-like objects with norm 0. In effect, the creation of an object like

$$(\pm i\mathbf{k}E \pm \mathbf{i}\mathbf{p} + \mathbf{j}m),$$

squaring to zero, is the only way of defining a point in a space that is rotation and translation symmetric. We define a point at the interaction between two spaces, and the point-like fermion spends its whole existence switching between them (*zitterbewegung*), and we find that we cannot define a point, or the space in which it exists, without defining another space with which it interacts to produce a totality zero. One space will appear distorted with respect to the other, but each will contain the same information, or the information in one will be the reverse of the information in the other (Rowlands, 2007, 2013, 2014, 2015a).

Vacuum

To return to Pauli exclusion, and Nature as a totality of zero, we can imagine creating a particle (with all the potentials representing its interactions) in the form $(\pm i\boldsymbol{k}E \pm \boldsymbol{i}\mathbf{p} + \boldsymbol{j}m)$ and then being forced to structure the rest of the universe *or vacuum*, so that it can be represented by

$$-(\pm i\boldsymbol{k}E \pm \boldsymbol{i}\mathbf{p} + \boldsymbol{j}m)$$

So both the superposition and combination states of fermion and vacuum become zero:

$$(\pm i\boldsymbol{k}E \pm \boldsymbol{i}\mathbf{p} + \boldsymbol{j}m) - (\pm i\boldsymbol{k}E \pm \boldsymbol{i}\mathbf{p} + \boldsymbol{j}m) = 0$$
$$-(\pm i\boldsymbol{k}E \pm \boldsymbol{i}\mathbf{p} + \boldsymbol{j}m)(\pm i\boldsymbol{k}E \pm \boldsymbol{i}\mathbf{p} + \boldsymbol{j}m) = 0$$

The ‘hole’ left by creating the particle from nothing is the rest of the universe needed to maintain it in that state. We give it the name vacuum. So the vacuum for one particle cannot be the vacuum for any other. Vacuum tells us ‘where’ the other ‘space’, based on $\boldsymbol{i}$, $\boldsymbol{j}$, $\boldsymbol{k}$, is

besides 'real' space, based on **i**, **j**, **k**. Its inaccessibility is demonstrated by the chirality between matter and antimatter, positive and negative energy states, and forward and backward directions in time.

The fermionic structure ($\pm i\mathbf{k}E \pm \mathbf{i}\mathbf{p} + \mathbf{j}m$) posits two states with $+E$ and two with $-E$, which is equivalent to two with + time direction and two with –. And yet only $+E$ and $+t$ states are observed, and the universe is predominantly made of matter rather than antimatter. The second space, the one in which energy and time become negative, is the vacuum space, the one which encompasses the 'rest of the universe', as opposed to the real space defining the point-particle. It is the one in which charge is conjugated and time reversed simultaneously.

The nilpotent structure incorporates both on an equal basis, and it is interesting to recall that it was the seeming inability to do this which led Feynman to develop the path integral method of quantum mechanics as opposed to the previous use of Hamiltonian methods. Perhaps this indicates that, as long as we use the nilpotent formalism, we will be able to reformulate path integral calculations using Hamiltonians.

The nonlocal aspect of the fermionic nilpotent state ($\pm i\mathbf{k}E \pm \mathbf{i}\mathbf{p} + \mathbf{j}m$) is defined by a continuous vacuum $-(\pm i\mathbf{k}E \pm \mathbf{i}\mathbf{p} + \mathbf{j}m)$. The continuous vacuum, with its negative energy, appears to be that associated with gravity. We can consider the fermion itself, with positive energy and ($\pm i\mathbf{k}E \pm \mathbf{i}\mathbf{p} + \mathbf{j}m$), as being in some sense an inertial term, and discrete fermions create their inertial mass through interaction with the continuous vacuum. Essentially, charge gives us the vacuum space dimensions, while mass gives it the continuous (negative) energy and nonlocality.

Further, we can use the operators $\boldsymbol{k}$, $\boldsymbol{i}$, $\boldsymbol{j}$ to effectively partition the continuous vacuum state, or the inertia which opposes this, into discrete components with a dimensional structure, which can then be

identified as the weak, strong and electric components of vacuum, responding respectively to the discrete weak strong and electric charges. We can postmultiply ($\pm i\boldsymbol{k}E \pm \boldsymbol{i}\mathbf{p} + \boldsymbol{j}m$) by the idempotent $\boldsymbol{k}(\pm i\boldsymbol{k}E \pm \boldsymbol{i}\mathbf{p} + \boldsymbol{j}m)$ any number of times, without changing its state

$$(\pm i\boldsymbol{k}E \pm \boldsymbol{i}\mathbf{p} + \boldsymbol{j}m)\,\boldsymbol{k}(\pm i\boldsymbol{k}E \pm \boldsymbol{i}\mathbf{p} + \boldsymbol{j}m)\,\boldsymbol{k}(\pm i\boldsymbol{k}E \pm \boldsymbol{i}\mathbf{p} + \boldsymbol{j}m)\ldots \rightarrow (\pm i\boldsymbol{k}E \pm \boldsymbol{i}\mathbf{p} + \boldsymbol{j}m)$$

The idempotent acts as a vacuum operator, not changing the state. We can show that the same applies if we use the idempotents $\boldsymbol{i}(\pm i\boldsymbol{k}E \pm \boldsymbol{i}\mathbf{p} + \boldsymbol{j}m)$ and $\boldsymbol{j}(\pm i\boldsymbol{k}E \pm \boldsymbol{i}\mathbf{p} + \boldsymbol{j}m)$:

$$(\pm i\boldsymbol{k}E \pm \boldsymbol{i}\mathbf{p} + \boldsymbol{j}m)\,\boldsymbol{i}(\pm i\boldsymbol{k}E \pm \boldsymbol{i}\mathbf{p} + \boldsymbol{j}m)\,\boldsymbol{i}(\pm i\boldsymbol{k}E \pm \boldsymbol{i}\mathbf{p} + \boldsymbol{j}m)\ldots \rightarrow (\pm i\boldsymbol{k}E \pm \boldsymbol{i}\mathbf{p} + \boldsymbol{j}m)$$

$$(\pm i\boldsymbol{k}E \pm \boldsymbol{i}\mathbf{p} + \boldsymbol{j}m)\,\boldsymbol{j}(\pm i\boldsymbol{k}E \pm \boldsymbol{i}\mathbf{p} + \boldsymbol{j}m)\,\boldsymbol{j}(\pm i\boldsymbol{k}E \pm \boldsymbol{i}\mathbf{p} + \boldsymbol{j}m)\ldots \rightarrow (\pm i\boldsymbol{k}E \pm \boldsymbol{i}\mathbf{p} + \boldsymbol{j}m)$$

In each case, every *alternate* bracket changes the sign of its E, $\mathbf{p}$ or E and $\mathbf{p}$ (equivalent to m) terms, leading to the creation of a bosonic state, here defined as a combination of fermion and antifermion, or equivalent. The only difference between the three is that the *alternate* brackets can be written as

$$\begin{aligned}(-i\boldsymbol{k}E + \boldsymbol{i}\mathbf{p} + \boldsymbol{j}m) \quad &\text{for} \quad \boldsymbol{k}\\ (i\boldsymbol{k}E - \boldsymbol{i}\mathbf{p} + \boldsymbol{j}m) \quad &\text{for} \quad \boldsymbol{i}\\ (-i\boldsymbol{k}E - \boldsymbol{i}\mathbf{p} + \boldsymbol{j}m) \quad &\text{for} \quad \boldsymbol{j}\end{aligned}$$

This means that they undergo respective T, P and C transformations. In addition the combinations with the unchanged bracket ($i\boldsymbol{k}E + \boldsymbol{i}\mathbf{p} +$

j*m*) means that the three operators produce different types of bosonic states, respectively:

spin 1
spin 0
paired fermion

We can regard *T* and *C* as transformations in vacuum space because they concern vacuum space quantities and lead to negative energy states. The *P* transformation concerns the real space quantity and leads to a positive energy state. However, it could occur in vacuum space by a combined *TC* transformation, and so real fermion / antifermion states are not purely left- / right-handed, as their weak interactions would suggest. The degree of overlap is determined by the *m* term, which is said to be the product of the *zitterbewegung*, or switching of a fermion between + and –*E* states or real space and vacuum.

We can see how the 3 bosonic states are related to vacua produced by the 3 quaternionic operators, and the 3 discrete transformations:

weak spin 1 *T*

$(i\mathbf{k}E + \mathbf{i}\mathbf{p} + \mathbf{j}m)\ \mathbf{k}\ (i\mathbf{k}E + \mathbf{i}\mathbf{p} + \mathbf{j}m)\ \mathbf{k}\ (i\mathbf{k}E + \mathbf{i}\mathbf{p} + \mathbf{j}m)\ \mathbf{k}\ (i\mathbf{k}E + \mathbf{i}\mathbf{p} + \mathbf{j}m)$ …

$(i\mathbf{k}E + \mathbf{i}\mathbf{p} + \mathbf{j}m)\ (-i\mathbf{k}E + \mathbf{i}\mathbf{p} + \mathbf{j}m)\ (i\mathbf{k}E + \mathbf{i}\mathbf{p} + \mathbf{j}m)\ (-i\mathbf{k}E + \mathbf{i}\mathbf{p} + \mathbf{j}m)$ …

electric spin 0 *C*

$(i\mathbf{k}E + \mathbf{i}\mathbf{p} + \mathbf{j}m)\ \mathbf{j}\ (i\mathbf{k}E + \mathbf{i}\mathbf{p} + \mathbf{j}m)\ \mathbf{j}\ (\mathbf{k}E + \mathbf{i}\mathbf{p} + \mathbf{j}m)\ \mathbf{j}\ i\mathbf{k}E + \mathbf{i}\mathbf{p} + \mathbf{j}m)$ …

$(i\mathbf{k}E + \mathbf{i}\mathbf{p} + \mathbf{j}m)\ (-\ i\mathbf{k}E - \mathbf{i}\mathbf{p} + \mathbf{j}m)\ (i\mathbf{k}E + \mathbf{i}\mathbf{p} + \mathbf{j}m)\ (-\ i\mathbf{k}E - \mathbf{i}\mathbf{p} + \mathbf{j}m)$ …

strong paired fermion state *P*

$$(i\mathbf{k}E + \mathbf{i}\mathbf{p} + \mathbf{j}m)\ \mathbf{i}\ (i\mathbf{k}E + \mathbf{i}\mathbf{p} + \mathbf{j}m)\ \mathbf{i}\ (i\mathbf{k}E + \mathbf{i}\mathbf{p} + \mathbf{j}m)\ \mathbf{i}\ (i\mathbf{k}E + \mathbf{i}\mathbf{p} + \mathbf{j}m)$$
…
$$(i\mathbf{k}E + \mathbf{i}\mathbf{p} + \mathbf{j}m)\ (i\mathbf{k}E - \mathbf{i}\mathbf{p} + \mathbf{j}m)\ (i\mathbf{k}E + \mathbf{i}\mathbf{p} + \mathbf{j}m)\ (i\mathbf{k}E - \mathbf{i}\mathbf{p} + \mathbf{j}m)\ \ldots$$

Vacuum space and charge

Now, the vacuum space, is the one defined by the units connected with charges, ***i***, ***j***, ***k***. Here we see that the units have multiple roles:

Charges
PCT operators
Dimensions of ‘vacuum space’
Generators of the 3 additional terms in the Dirac wavefunction
Creation operators converting fermions to 3 types of boson

The complete vacuum, defined by $(i\mathbf{k}E + \mathbf{i}\mathbf{p} + \mathbf{j}m)$ or $-(i\mathbf{k}E + \mathbf{i}\mathbf{p} + \mathbf{j}m)$, can be considered as equivalent to that defined by gravity or inertia, which the gauge forces split into 3 dimensional components. Also, taking $(i\mathbf{k}E + \mathbf{i}\mathbf{p} + \mathbf{j}m)$ to convey the angular momentum information about a particle state, we can see that the units also say something about the respective conservation laws of handedness, direction and magnitude of this quantity. Ultimately, this introduces the $SU(2)$, $SU(3)$ and $U(1)$ group operators into particle physics, and explains why the symmetry between the 3 charges is broken. But the key idea is that the total angular momentum information can be obtained either through ***k***, ***i*** and ***j*** of iE, $\mathbf{p}$ and m, or through the **i**, **j**, **k** of p.

The two ‘spaces’ contain *exactly equivalent information about the whole of physics*, although in one case the symmetry between the units

appears to be broken and in the other it is exactly preserved. To show that the information is equivalent, we will consider how the uniqueness of a fermion wavefunction is determined to maintain Pauli exclusion. So if we know the iE, $\mathbf{p}$ and m values in $(\pm i\mathbf{k}E \pm \mathbf{i}\mathbf{p} + \mathbf{j}m)$ for any fermion, we can show that those of another fermion must be different. But, there is also a completely different way of establishing Pauli exclusion, that is, by realising that fermions have antisymmetric wavefunctions.

The nilpotent structure explains immediately why we have Pauli exclusion between fermions, but the alternative, *conventional*, way of explaining this property leads us to a profound insight on the nature of the information available in quantum systems if we structure it in nilpotent form. We define fermion wavefunctions to be antisymmetric, so that:

$$(\psi_1\psi_2 - \psi_2\psi_1) = -(\psi_2\psi_1 - \psi_1\psi_2)$$

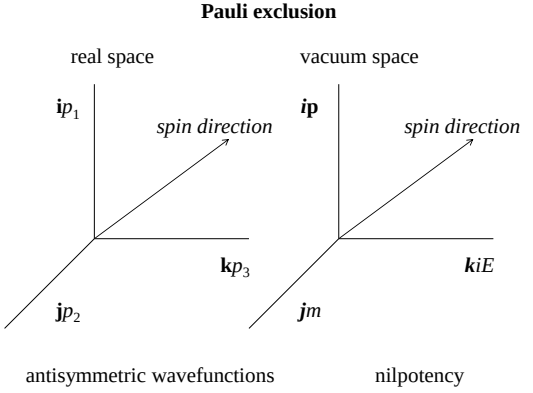
In nilpotent terms, we write $(\psi_1\psi_2 - \psi_2\psi_1)$ as

$$\begin{aligned}&(\pm i\mathbf{k}E_1 \pm \mathbf{i}\mathbf{p}_1 + \mathbf{j}m_1)\,(\pm i\mathbf{k}E_2 \pm \mathbf{i}\mathbf{p}_2 + \mathbf{j}m_2)\\ &-(\pm i\mathbf{k}E_2 \pm \mathbf{i}\mathbf{p}_2 + \mathbf{j}m_2)\,(\pm i\mathbf{k}E_1 \pm \mathbf{i}\mathbf{p}_1 + \mathbf{j}m_1)\\ &= 4\mathbf{p}_1\mathbf{p}_2 - 4\mathbf{p}_2\mathbf{p}_1 = 8i\mathbf{p}_1 \times \mathbf{p}_2 = -8i\mathbf{p}_2 \times \mathbf{p}_1.\end{aligned}$$

This result is clearly antisymmetric, but it also has a quite astonishing consequence, for it requires any nilpotent wavefunction to have a $\mathbf{p}$ vector, in *real space*, the one defined by the axes $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$, at a *different orientation* to any other. The wavefunctions of all nilpotent fermions then instantaneously correlate because the planes of their $\mathbf{p}$ vector directions must all intersect. At the same time, the nilpotent condition requires the E, $\mathbf{p}$ and m combinations to be unique, and we can visualize this as constituting a unique direction in *vacuum* space

along a set of axes defined by ***k***, ***i***, ***j*** or **k**, **i**, **j**, with coordinates defined by the values of *E*, **p** and *m*.

The directions of the vectors in each space carry *all the information* available to a fermionic state, and so the information in the two spaces is totally dual, and is equivalent to the instantaneous direction of the spin in the real space. The total information determining the behaviour of a fermion and even of the entire universe is contained in a single spin direction. The information here must be the same as in the combination of *iE*, *p* and *m*. That is, we can represent the unique spin direction in parallel axes in two different spaces.



Though the duality results in fermion and vacuum occupying separate 3-dimensional ‘spaces’, which are combined in the double Clifford algebra defining the singularity state, these ‘spaces’, though seemingly different in observational terms, are truly dual, each containing the same information (angular momentum), and the duality manifests itself directly in many physical forms.

Angular momentum, which in some form combines all the information of the phase space of the particle, shows this duality

directly, because, as a 3-dimensional pseudovector quantity, it has the rotation symmetric and nonconserved 3-dimensionality of space; but, at the same time, it shows 3 different aspects and 3 different conservation laws connected with them, which are connected by a different and rotation asymmetric 3-dimensionality.

One of the remarkable aspects of this analysis is that it shows that the connection between space and time, in special relativity, is not really 4-dimensional at all, but *3-dimensional*, though the 3 dimensions involved are different from those of space. We can write out the relativistic nilpotent as

$$i\mathbf{k}E + \mathbf{ii}p_x + \mathbf{ji}p_y + \mathbf{ki}p_y + \boldsymbol{j}m$$

or even

$$i\mathbf{k}t + \mathbf{ii}x + \mathbf{ji}y + \mathbf{ki}z + \boldsymbol{j}\tau$$

The 3 components of space and momentum are denoted by the 3 vector coefficients, but they do not add directly to the pseudoscalar *iE* or *it*. They only connect by additional coefficients, which are those of another space. Also, the space-time connection is not privileged over those with mass and charge. If we write the relativistic nilpotents in the form:

$$i\mathbf{k}E + \mathbf{i}\mathbf{p} + \boldsymbol{j}m$$

and

$$i\mathbf{k}t + \mathbf{ii}\mathbf{r} + \boldsymbol{j}\tau$$

we can see that there is an immediate analogy with the cases of 3 static forces in equilibrium or a closed path round a right-angled triangle (say with lengths 3, 4 and 5) where

$$i\mathbf{k}W + \mathbf{i}T_1 + \boldsymbol{j}T_1$$

or $$i\boldsymbol{k}5 + \boldsymbol{i}4 + \boldsymbol{j}3$$

Where we have a zero total of this kind, we can consider it to be a case of finding the resultant in the dual space which cancels that in real space, showing the significance of Newton's third law in this connection.

If we represent particles in terms of their charge structures, this can be done in two different spaces. If we look at baryons composed of 3 quarks, we can see their behaviour in terms of real space, or that of the momentum operator, and this has rotation symmetry. But we can also see it in terms of vacuum or charge space, and this is rotation asymmetric.

Space and charge show us the dual aspects of angular momentum, nonconserved and symmetric, and conserved and asymmetric. All 3-dimensional quantities are of this kind, one that is ultimately expressed in Noether's theorem. To any 3-dimensionality that gives us nonconservation and rotation symmetry, there is always attached another that gives the same information using conservation and rotation asymmetry.

This is why the ultimate expression of physics requires phase space. To any description involving space or space-time, we require an equivalent description involving vacuum space, manifesting itself through charge, or through the combination of charge and space and time, which is described as energy-momentum. At the quantum (point-particle) level, these cannot be separated. This is how the combination of all 3 charges becomes necessary to describe the behaviour of the particular charges which is associated with spatial 3-dimensionality. The behaviour of the strong interaction can be completely determined by the spatial 3-dimensionality of the momentum operator in the nilpotent wavefunction. However, the situation of the strong charge as

a 3-dimensional operator within the fermion state requires the application of the other 3-dimensionality associated with charge.

Baryons

Effectively, the vector aspect of the strong charge requires a source term and corresponding vacuum with three components. Though we clearly cannot combine three components in the form:

$$(i\mathbf{k}E \pm \boldsymbol{i}\mathbf{p} + \boldsymbol{j}m)\,(i\mathbf{k}E \pm \boldsymbol{i}\mathbf{p} + \boldsymbol{j}m)\,(i\mathbf{k}E \pm \boldsymbol{i}\mathbf{p} + \boldsymbol{j}m)$$

as this will automatically reduce to zero, we can imagine a three-component structure in which the vector nature of **p** plays an explicit role

$$(i\mathbf{k}E \pm \boldsymbol{i}\,\mathbf{i}p_x + \boldsymbol{j}m)\,(i\mathbf{k}E \pm \boldsymbol{i}\,\mathbf{j}p_y + \boldsymbol{j}m)\,(i\mathbf{k}E \pm \boldsymbol{i}\,\mathbf{k}p_z + \boldsymbol{j}m)$$

This has nilpotent solutions when $\mathbf{p} = \pm \boldsymbol{i}\,\mathbf{i}p_x$, $\mathbf{p} = \pm \boldsymbol{i}\,\mathbf{j}p_y$, or $\mathbf{p} = \pm \boldsymbol{i}\,\mathbf{k}p_z$, or when the momentum is directed entirely along the *x*, *y*, or *z* axes, in either direction, though these, of course, are arbitrarily defined. Any other phases can be written as a superposition of these. Using the appropriate normalization, these reduce to

$$\begin{aligned}
&(i\mathbf{k}E + \mathbf{ii}p_x + \boldsymbol{j}m) && +RGB\\
&(i\mathbf{k}E - \mathbf{ii}p_x + \boldsymbol{j}m) && -RBG\\
&(i\mathbf{k}E - \mathbf{ij}p_y + \boldsymbol{j}m) && +BRG\\
&(i\mathbf{k}E + \mathbf{ij}p_y + \boldsymbol{j}m) && -GRB
\end{aligned}$$

$$(i\mathbf{k}E + \mathbf{i}\mathbf{k}p_z + \mathbf{j}m) \qquad +GBR$$
$$(i\mathbf{k}E - \mathbf{i}\mathbf{k}p_z + \mathbf{j}m) \qquad -BGR$$

with the third and fourth changing, very significantly, the sign of the **p** component. Because of this, there has to be a maximal superposition of left- and right-handed components, thus explaining the zero observed chirality in the interaction (and the mass of the baryon).

The group structure required to maintain these phases is an $SU(3)$ structure, with eight generators and a wavefunction, exactly as in the conventional model using coloured quarks,

$$\psi \sim (BGR - BRG + GRB - GBR + RBG - RGB).$$

'Colour' transitions in the 3-component structures are produced either by an exchange of the components of **p** between the individual quarks or baryon components, or by a relative switching of the component positions, independently of any real distance between the components. No direction can be privileged, so the transition must be gauge invariant, and the mediators must be massless, exactly as with the eight massless gluons of the gluon structure. We have written only the first term of the 4-component spinors, but we have retained the two spin states, as these will be needed explicitly.

The complete wavefunction will, in effect, contain information from the equivalent of six allowed independent nonlocally gauge invariant phases, all existing simultaneously and subject to continual transitions at a constant rate:

$$(i\mathbf{k}E + \mathbf{i}\mathbf{i}p_x + \mathbf{j}m)\,(i\mathbf{k}E + \ldots + \mathbf{j}m)\,(i\mathbf{k}E + \ldots + \mathbf{j}m) \quad +RGB$$
$$(i\mathbf{k}E - \mathbf{i}\mathbf{i}p_x + \mathbf{j}m)\,(i\mathbf{k}E - \ldots + \mathbf{j}m)\,(i\mathbf{k}E - \ldots + \mathbf{j}m) \quad -RBG$$
$$(i\mathbf{k}E + \ldots + \mathbf{j}m)\,(i\mathbf{k}E + \mathbf{i}\mathbf{j}p_y + \mathbf{j}m)\,(i\mathbf{k}E + \ldots + \mathbf{j}m) \quad +BRG$$
$$(i\mathbf{k}E - \ldots + \mathbf{j}m)\,(i\mathbf{k}E - \mathbf{i}\mathbf{j}p_y + \mathbf{j}m)\,(i\mathbf{k}E - \ldots + \mathbf{j}m) \quad -GRB$$

$$(i\mathbf{k}E + \ ... \ + \mathbf{j}m)\,(i\mathbf{k}E + \ ... \ + \mathbf{j}m)\,(i\mathbf{k}E + \mathbf{i}\mathbf{k}p_z + \mathbf{j}m) \quad +GBR$$
$$(i\mathbf{k}E - \ ... \ + \mathbf{j}m)\,(i\mathbf{k}E - \ ... \ + \mathbf{j}m)\,(i\mathbf{k}E - \mathbf{i}\mathbf{k}p_z + \mathbf{j}m) \quad -BGR$$

The simultaneous existence of all phases further means that *individual* quarks, and such identifying characteristics as electric charges, are not identifiable by their spatial positions (unlike, say, the proton and electron constituting a hydrogen atom), thus explaining, for example, why the neutron has no electric dipole moment. Just as $U(1)$ establishes that spherical symmetry of a point source requires the rotation to be performed independently of the length of the radius vector, so $SU(3)$ requires the rotation to be independent of the coordinate system used. In terms of Noether's theorem, while $U(1)$ conserves the magnitude of angular momentum, $SU(3)$ conserves the direction.

These are the established characteristics of the strong interaction and here we have an explanation, derived on an analytic basis, for a force with these characteristics. If we now look at how the expression

$$(i\mathbf{k}E \pm \mathbf{i}\,\mathbf{i}p_x + \mathbf{j}m)\,(i\mathbf{k}E \pm \mathbf{i}\,\mathbf{j}p_y + \mathbf{j}m)\,(i\mathbf{k}E \pm \mathbf{i}\,\mathbf{k}p_z + \mathbf{j}m)$$

with its spatial 3-dimensionality, fits into the 3-dimensionality associated with the charge picture, we can recall that the 4 components in the Dirac spinor, can be seen to represent the fermion as seen in terms of gravity / inertia, strong, weak and electric vacua. The 3 momentum components have to be mapped onto the space created by the 3 charges. The principal options are:

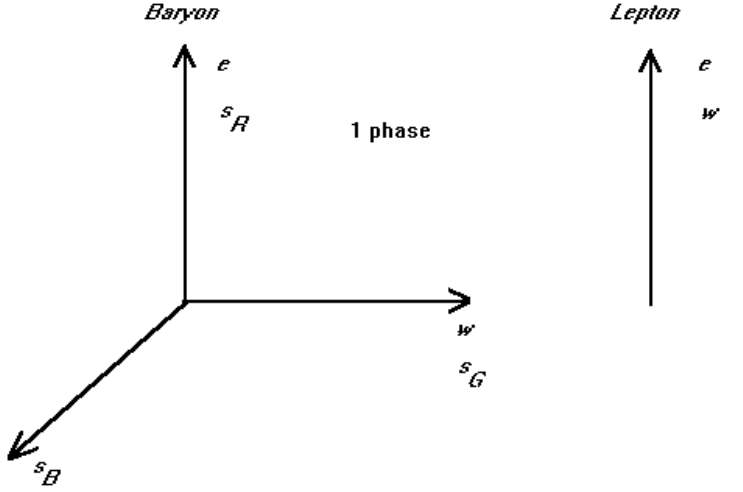
$$\begin{array}{c} \\ inertial \\ strong \\ weak \\ electric \end{array} \begin{pmatrix} ikE \pm \mathbf{i}\boldsymbol{\sigma}.\mathbf{p}_1 + \mathbf{j}m \\ ikE \mp \mathbf{i}\boldsymbol{\sigma}.\mathbf{p}_1 + \mathbf{j}m \\ -ikE \pm \mathbf{i}\boldsymbol{\sigma}.\mathbf{p}_1 + \mathbf{j}m \\ -ikE \mp \mathbf{i}\boldsymbol{\sigma}.\mathbf{p}_3 + \mathbf{j}m \end{pmatrix} \begin{pmatrix} ikE \mp \mathbf{i}\boldsymbol{\sigma}.\mathbf{p}_2 + \mathbf{j}m \\ ikE \pm \mathbf{i}\boldsymbol{\sigma}.\mathbf{p}_2 + \mathbf{j}m \\ -ikE \mp \mathbf{i}\boldsymbol{\sigma}.\mathbf{p}_3 + \mathbf{j}m \\ -ikE \pm \mathbf{i}\boldsymbol{\sigma}.\mathbf{p}_2 + \mathbf{j}m \end{pmatrix} \begin{pmatrix} ikE \pm \mathbf{i}\boldsymbol{\sigma}.\mathbf{p}_3 + \mathbf{j}m \\ ikE \mp \mathbf{i}\boldsymbol{\sigma}.\mathbf{p}_3 + \mathbf{j}m \\ -ikE \pm \mathbf{i}\boldsymbol{\sigma}.\mathbf{p}_2 + \mathbf{j}m \\ -ikE \mp \mathbf{i}\boldsymbol{\sigma}.\mathbf{p}_1 + \mathbf{j}m \end{pmatrix}$$

and

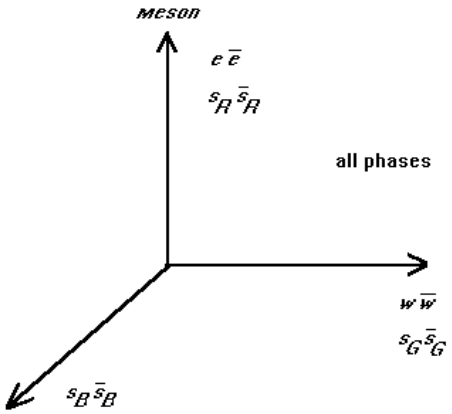
$$\begin{array}{c} \\ inertial \\ strong \\ weak \\ electric \end{array} \begin{pmatrix} ikE \pm i\boldsymbol{\sigma}.\mathbf{p}_1 + jm \\ ikE \mp i\boldsymbol{\sigma}.\mathbf{p}_1 + jm \\ -ikE \pm i\boldsymbol{\sigma}.\mathbf{p}_1 + jm \\ -ikE \mp i\boldsymbol{\sigma}.\mathbf{p}_3 + jm \end{pmatrix} \begin{pmatrix} ikE \pm i\boldsymbol{\sigma}.\mathbf{p}_2 + jm \\ ikE \mp i\boldsymbol{\sigma}.\mathbf{p}_2 + jm \\ -ikE \pm i\boldsymbol{\sigma}.\mathbf{p}_3 + jm \\ -ikE \mp i\boldsymbol{\sigma}.\mathbf{p}_2 + jm \end{pmatrix} \begin{pmatrix} ikE \pm i\boldsymbol{\sigma}.\mathbf{p}_3 + jm \\ ikE \mp i\boldsymbol{\sigma}.\mathbf{p}_3 + jm \\ -ikE \pm i\boldsymbol{\sigma}.\mathbf{p}_2 + jm \\ -ikE \mp i\boldsymbol{\sigma}.\mathbf{p}_1 + jm \end{pmatrix}$$

Here, we assume only one dimension of $\mathbf{p}$ is nonzero, say $\mathbf{p}_3$, so determining the charge structure, with the 3 charges distributed.

We can also represent it using a phasor diagrams, with the charges separated in baryons, but not in leptons:



For mesons we would have:



If the strong charge is absent, we need only one phase, with all charges aligned.

$$\begin{array}{ll} inertial & \left(ik\mathbf{E} \pm i\boldsymbol{\sigma}.\mathbf{p}_1 + jm \right) \\ strong & \left(ik\mathbf{E} \mp i\boldsymbol{\sigma}.\mathbf{p}_1 + jm \right) \\ weak & \left(-ik\mathbf{E} \pm i\boldsymbol{\sigma}.\mathbf{p}_1 + jm \right) \\ electric & \left(-ik\mathbf{E} \mp i\boldsymbol{\sigma}.\mathbf{p}_1 + jm \right) \end{array}$$

The wavefunctions suggest that the underlying charge structures of quarks are as in the first table below, with the more familiar fractional electric charges appearing in observation, as in the second table, because the strong interaction is absolutely gauge invariant:

	Blue	**Green**	**Red**
up	*e*	*e*	0
	B	0	0
down	0	0	–*e*
	B	0	0

	Blue	**Green**	**Red**

up	2*e* / 3	2*e* / 3	2*e* / 3
	B / 3	*B* / 3	*B* / 3
down	–*e* / 3	–*e* / 3	–*e* / 3
	B / 3	*B* / 3	*B* / 3

A similar fractionalization of electric charge occurs in the fractional quantum Hall effect, where the fractional status of the charge on an electron is produced by its *weak* interaction with an odd number of flux lines > 1. In both cases, it is a different electric-charge independent force which ensures that the electric charges become fractional. The underlying charge structures can be used to represent the Standard Model simply in a set of 4 quark-lepton tables. They also predict Grand Unification of the 4 forces at the Planck mass, explain *CP* violation, solve a serious anomaly in the Higgs mechanism as applied to finding the masses of fermion states, etc. (Rowlands, 2007, 2014, Rowlands and Cullerne, 2001)

A				
		B	**G**	**R**
u	+*e*	1***j***	1***j***	0***i***
	+*s*	1***i***	0***k***	0***j***
	+*w*	1***k***	0***i***	0***k***
d	–*e*	0***j***	0***k***	1***j***
	+*s*	1***i***	0***i***	0***k***
	+*w*	1***k***	0***j***	0***i***
c	+*e*	1***j***	1***j***	0***i***
	+*s*	1***i***	0***k***	0***j***
	–*w*	z_P***k***	0***i***	0***k***
s	+*e*	0***j***	0***k***	1***j***
	+*s*	1***i***	0***i***	0***k***
	–*w*	z_P***k***	0***j***	0***i***

B				
		B	**G**	**R**
u	+*e*	1***j***	1***j***	0***i***
	+*s*	0***i***	0***k***	1***j***
	+*w*	1***k***	0***i***	0***k***
d	–*e*	0***j***	0***k***	1***j***
	+*s*	0***i***	0***k***	1***j***
	+*w*	1***k***	0***j***	0***i***
c	+*e*	1***j***	1***j***	0***i***
	+*s*	0***i***	0***k***	1***j***
	–*w*	z_P***k***	0***i***	0***k***
s	+*e*	0***j***	0***k***	1***j***
	+*s*	0***i***	0***k***	1***j***
	–*w*	z_P***k***	0***j***	0***i***

t	+*e*	1***j***	1***j***	0***i***
	+*s*	1***i***	0***k***	0***j***
	−*w*	Z_T***k***	0***i***	0***k***
b	+*e*	0***j***	0***k***	1***j***
	+*s*	1***i***	0***i***	0***k***
	−*w*	Z_T***k***	0***j***	0***i***

t	+*e*	1***j***	1***j***	0***i***
	+*s*	0***i***	0***k***	1***j***
	−*w*	Z_T***k***	0***i***	0***k***
b	+*e*	0***j***	0***k***	1***j***
	+*s*	0***i***	0***k***	1***j***
	−*w*	Z_T***k***	0***j***	0***i***

C				
		B	**G**	**R**
u	+*e*	1***j***	1***j***	0***i***
	+*s*	0***i***	1***k***	0***j***
	+*w*	1***k***	0***i***	0***k***
d	−*e*	0***j***	0***k***	1***j***
	+*s*	0***i***	1***k***	0***j***
	+*w*	1***k***	0***j***	0***i***
c	+*e*	1***j***	1***j***	0***i***
	+*s*	0***i***	1***k***	0***j***
	−*w*	z_P***k***	0***i***	0***k***
s	+*e*	0***j***	0***k***	1***j***
	+*s*	0***i***	1***k***	0***j***
	−*w*	z_P***k***	0***j***	0***i***
t	+*e*	1***j***	1***j***	0***i***
	+*s*	0***i***	1***k***	0***j***
	−*w*	Z_T***k***	0***i***	0***k***
b	+*e*	0***j***	0***k***	1***j***
	+*s*	1***i***	0***i***	0***k***
	−*w*	0***i***	1***k***	0***j***

L				
		$\bar{e}$	$\bar{e}$	ν_e
	+*e*	1***j***	1***j***	0***j***
	+*s*	0***k***	0***i***	0***i***
	+*w*	0***i***	0***k***	1***k***
				e
	−*e*	0***i***	0***k***	1***j***
	+*s*	0***j***	0***i***	0***i***
	+*w*	0***i***	0***k***	1***k***
		$\bar{\mu}$	$\bar{\mu}$	ν_μ
	+*e*	1***j***	1***j***	0***j***
	+*s*	0***k***	0***i***	0***i***
	−*w*	0***i***	0***k***	1***k***
				μ
	−*e*	0***i***	0***k***	1***j***
	+*s*	0***j***	1***i***	0***i***
	−*w*	0***i***	0***k***	z_P***k***
		$\bar{\tau}$	$\bar{\tau}$	ν_τ
	+*e*	1***j***	1***j***	0***j***
	+*s*	0***k***	0***i***	0***i***

3-dimensionalities deriving from space and charge

We have seen that 3 components of charge is a 3-dimensionality separate from that of space, but acting against it to produce zero

totality. Other 3 dimensionalities emerge from those of space and charge. So, as we have seen, we have 3 quarks in a baryon because we have 3 dimensions of space and also of momentum. It is simultaneously an expression of the existence of 3 types of charge.

The existence of 3 generations of fundamental particles is an interesting case, which comes from *CPT* symmetry. According to the Dirac equation, massless fermions, if they existed, would be purely left-handed. So the measure of a fermion is a measure of how much right-handedness it contains. Higher mass fermions are more right-handed than lower mass ones. We can see the second and third generations as producing step-functions in right-handedness and so in mass values by successively violating *P* and *T* symmetries. The relative masses of the quarks and leptons in the 3 generations seem to be determined by the electroweak and strong coupling constants. This is discussed in the Appendix.

Now, a nonconserved 3-D quantity has rotation symmetry, so the parts are indistinguishable. This is characteristic of quantities whose dimensionalities come from that of space. A conserved 3-D quantity has rotation asymmetry, with the parts distinguishable and different, though in a systematic way which comes from its interaction with its counterpart in real space. These can be related to vacuum space derived from the units ***i j k***, and we only know about these indirectly through the combination.

There is, finally, an interesting application of these ideas in a seemingly ‘pure’ mathematical context, which gives indications of actually being determined by a physical argument. It has been a long-held belief of mine that physics and mathematics are linked at a much deeper level than one merely ‘using’ the other, and that they exhibit same basic patterns and structures. Several examples can be suggested, but one is of special interest because it leads to an iconic pure mathematical equation. Here, we wish to find a *physical* answer

to a specifically mathematical question: if we start with a scalar unit real number, say 1, how do we find the conjugate number to obtain totality zero? Physically, *mass*, which uses scalar, real numbers, is a conserved, real, and nondimensional quantity. To find the conjugate to this, we look for something that is nonconserved, imaginary and dimensional.

In mathematics and physics, nonconservation is expressed by differentiation: e is a number that arises only out of differentiation; i is the basis of imaginary numbers; and π is a number that arises only out of 3-dimensionality. So $e^{i\pi}$ gives us a number which is conjugate to 1. The transformations here are equivalent to the physical *TCP* or *PTC*, and 3-dimensionality, as always, plays a doubly crucial role. So the famous equation $e^{i\pi} = -1$ or $e^{i\pi} + 1 = 0$ may point to a direct connection between mathematics and physics at the most fundamental level.

Appendix: 3 generations of fermions and their masses

The 3 generations of fundamental fermions represent a special aspect of 3-dimensionality connected with *C*, *P* and *T* and *m*, *p* and *E* (and, by implication, indirectly with $\boldsymbol{k}$, $\boldsymbol{i}$ and $\boldsymbol{j}$).

u	d	ν_e	e
c	s	ν_μ	μ
t	b	ν_τ	τ

They have ascending masses based on the symmetry-breaking associated successively with violations of parity and time-reversal symmetry, which are the result of electroweak couplings, whereas differences between quarks and leptons in the same generation and isospin state appear to be due to the presence of strong couplings in

the former and their absence in the latter. In terms of mass, the highest level, with all symmetries broken, is the maximal value associated with the Higgs field, the source of all mass. This is seen in the third generation. The second generation has masses associated with the energy scale of the regular weak interaction, where symmetry-breaking first occurs and below that associated with the breaking of all symmetries. This can be taken as close to the mass of the *Z* boson (91.2 GeV).

The masses of the particles in the higher generations, in particular, are related to the creation of an electric charge unit (which is generally connected to the creation of electroweak mass as it generates a degree of right-handedness not strongly present in the pure weak interaction) within a distance of the order of the Compton wavelength, either standard or reduced. The standard Compton wavelength is the length scale relevant to mass conversions into energy, where the reduced wavelength is the one used when discrete mass is created.

The mass of *t*, the heaviest fermion, is close to maximal coupling to the Higgs field, expectation value $f \approx 246$ GeV. So the mass of *t* becomes $\approx f / \sqrt{2} \approx 174$ GeV, with no specific process needed to explain it. In the same generation, but with no strong charge, τ has a mass close to that which a single electric charge *e* would need to be confined by energy conservation to the reduced Compton radius associated with 246 GeV. This works out at $\approx f\alpha$, where α is the electromagnetic fine structure constant. The interactions associated with τ (unlike those associated with *t*) are purely electroweak, and the weak interaction does not extend to this length. The mass of *b* should be connected to that of its isospin 'down' partner, τ, and, in fact, it appears to be equivalent to that of τ divided by α_3, the strong interaction coupling at the associated energy (approximately 2.35).

In the second generation, with one degree less symmetry-breaking, masses are reduced by the appropriate electroweak factor, in this case

α, and c has a mass equivalent to that of t times α, or $f\alpha / \sqrt{2}$. This is also equivalent to that producing the charge e at the reduced Compton wavelength generated by the mass of maximal coupling to the Higgs field (i.e. that of t). The c mass (at 1.2 to 1.3 GeV) differs from that of t due to the electroweak coupling only, as the strong coupling is unchanged. The muon μ, the counterpart in this generation to τ, has the mass needed to generate e over the standard Compton wavelength for the Z energy scale. (We may note here that the energy scale here is determined by that of a real particle, whereas that for t is determined by the energy of a field.) As in the third generation, s, the isospin 'down' counterpart to μ, can again be taken to be equivalent to μ with an additional strong interaction, and so has a mass equivalent to that of μ divided by α_3, the strong interaction coupling at the associated energy (which, in this case, is approximately 0.9).

In the first generation, where there is no electroweak symmetry breaking, the masses are again reduced by a factor close to α. Here, the masses of the u and d quarks are not well established, but that of u is certain of the order of the mass of c times α at 9 MeV (although α may conceivably determine the mass ratios of the whole generations rather than just the isospin 'up' component). The electron mass is close to $2\alpha / 3$ times that of μ though the reason for the factor 2/3 is not readily apparent. (See the suggestions in Rowlands, 2015b.) The strong interaction does not dominate at this level in the mass of the d quark where its contribution would be a relatively small mass of e / α_3 (where α_3 would be a number considerably greater than 1). Clearly, the main contribution at this scale would be electroweak.

References

Hestenes, D. 1966, *Space-Time Algebras, Gordon and Breach*, New York

Marcer, P. J. and Rowlands, P., 2014a, Information, Bifurcation and Entropy in the Universal Rewrite System, *International Journal of Computing Anticipatory Systems*, 27, 203-215

Marcer, P. J. and Rowlands, P., 2014b, Universal Rewrite and Self-Organization, *International Journal of Computing Anticipatory Systems*, 27, 227-246

Rowlands, P., 1983, The fundamental parameters of physics. *Speculat. Sci. Tech.*, **6**, 69-80

Rowlands, P., 1991, *The Fundamental Parameters of Physics: An Approach towards a Unified Theory*, PD Publications, Liverpool

Rowlands, P., 2001, A foundational approach to physics, arXiv:physics/0106054

Rowlands, P., 2003, Representations of a Fundamental Theory, in K. Bowden (ed.), *Boundaries*, 18-47

Rowlands, P. 2007, *Zero to Infinity The Foundations of Physics*, Singapore, London and Hackensack, NJ: World Scientific

Rowlands, P. 2009, Are there alternatives to our present theories of physical reality?, arXiv:0912.3433

Rowlands, P., 2010, Mathematics and Physics as Emergent Aspects of a Universal Rewrite System, *International Journal of Computing Anticipatory Systems*, **25**, 115-131

Rowlands, P., 2013, Space and Antispace, in R. L. Amoroso, L. H. Kauffman and P. Rowlands (eds.), *The Physics of Reality Space, Time, Matter, Cosmos*, World Scientific, 29-37

Rowlands, P. 2014, *The Foundations of Physical Law*, Singapore, London and Hackensack, NJ: World Scientific

Rowlands, P. 2015a, *How Schrödinger's Cat Escaped the Box*, Singapore, London and Hackensack, NJ: World Scientific

Rowlands, P. 2015b, Reducing the Number of Free Parameters, in Grenville J. Croll and Nicky Graves Gregory (eds.), *Explorations*, 223-259

Rowlands, P. 2015c, How many dimensions are there?, in Richard L Amoroso, Louis H. Kauffman and Peter Rowlands (eds.), *Unified Field Mechanics Natural Science Beyond the Veil of Spacetime Proceedings of the IX Symposium Honoring Noted French Mathematical Physicist Jean-Pierre Vigier*, World Scientific, 2015, 46-54

Rowlands, P. and Cullerne, J. P., 2001, The connection between the Han-Nambu quark theory, the Dirac equation and fundamental symmetries. *Nuclear Physics* **A 684**, 713-5

Rowlands, P. and Diaz, B. 2002, A universal alphabet and rewrite system, arXiv:cs.OH/0209026

FUNDAMENTAL ERRORS IN PHYSICS.

PETER ROWLANDS
Department of Physics, University of Liverpool, Oliver Lodge Laboratory, Oxford St, Liverpool, L69 7ZE, UK. email p.rowlands@liverpool.ac.uk

ABSTRACT. Physics at the fundamental level seems currently to be in a situation where theory and experiment are frequently in direct contradiction. Theories are being produced which cannot possibly be correct as they already contradict well-established results from high-energy physics experiments. One of the reasons for this is the fact that alternative theories seem to be in short supply due to the downgrading of conceptual thinking over the last forty years. This has been responsible for the perpetuation of several fundamentally erroneous ways of theorising which seem to stand in the way of achieving a deeper understanding of the most significant fundamental ideas.

Introduction: Pointless theoretical pursuits

Many books, papers and web sites – with genuinely mainstream connections – make categorical statements about physical truths which are actually often highly questionable. In my view, this shows a lack of true physical insight, often reflecting a formulaic approach to physics education. It is damaging because major physics advances often require deeper physical insights. The present impasse in fundamental physics, which can be dated back to the completion of the main features of the Standard Model around 1973, is to my mind a symptom of the dominance of a 'shut up and calculate' approach and the downgrading of conceptual thinking as irrelevant and even trivial.

At this very moment, there are a huge number of people who are pursuing theories which we *know* are wrong. The authors of these theories – and they still command a great deal of respect – are generating mathematical structures which experimental evidence has already shown to be impossible whatever modifications are made. It is like pursuing an industry when you have been completely overtaken by competitors because you are still tooled up for it. When you have invested so many years and such immense intellectual capital and career-structuring in developing models in a particular way, what else do you do when the foundations collapse in an instant?

Many people are still developing theories in which the Standard Model of particle physics can be tweaked by supposing that some relatively simple addition to the particle hierarchy, e.g. supersymmetry, higher Z / W, multiple Higgs bosons, can be added to explain all the anomalies, etc. This is *already* impossible. The LHC is celebrated for discovering the Higgs boson, 48 years after it was predicted, and explaining the origin of particle masses, but it has also produced an even more significant result which is surprisingly little known. A colleague, Andy Mehta, one of the leading lights of the ATLAS experiment, who was significantly involved in the Higgs discovery, says that the LHC has already ruled out the possibility that any new particles that couple with existing ones or the Higgs, can exist at all. This includes the higher mass ones which are supposedly outside the range of experiments like ATLAS. In principle, the Higgs couples to all known massive particles and the coupling increases significantly with the mass. Higher mass particles would couple so strongly to the Higgs that we would already be seeing their effects in the Higgs decays that have been measured to date – and we don't.

The Higgs mechanism and gravity

Concerning the Higgs mechanism itself, there is a site which I am sure reflects mainstream opinion: 'Of Particular Significance: Conversations About Science with Theoretical Physicist Matt Strassler', which has a feature on 'Why the Higgs and Gravity are Unrelated' (2012). Here we read that: 'The idea that there is any connection between the Higgs mechanism and gravity is 'completely wrong'.' Gravity, says the author is not related to mass but a combination of energy and momentum. ('Mass' is here assumed to be 'rest' mass.) Gravitational forces do not obey Newton's law, based on 'mass', but Einstein's based on energy-momentum. Also, *W*, *Z*, quarks and leptons owe their masses to Higgs, but not neutrons and protons, and certainly not Higgs itself.

Now, everything fundamental in the Standard Model, except the Higgs, that is, all quarks and all leptons and the gauge bosons of the weak interaction, gains its mass by coupling to the Higgs. But, the composite particles which make up the bulk of ordinary matter, such as the neutron and proton, have much bigger masses than their 'coloured' constituent quarks would alone produce. This must come from somewhere, but, according to the author, it is not the Higgs mechanism. In his opinion, which is presented here as obvious fact, the Higgs field is not universal. 'Gravitational fields have spin 2 and are described as part of space and time; they interact with all particles and fields in nature. The Higgs field, which has spin 0, only interacts directly with elementary particles and fields that also participate in the electromagnetic and weak nuclear forces.' This doesn't include the *strong* force between quarks. In addition, black hole masses do not derive from the Higgs mechanism, because most of ordinary matter doesn't derive from the Higgs, and they derive mostly from ordinary matter, though they could also be made up from massless photons.

What strikes me about this way of arguing is that it doesn't recognize that it is based on formulaic and superficial assumptions, some of which are still speculative or merely alternative. Yet it exudes professional authority as though we had solved the deepest problems in physics and weren't floundering about without an idea of where to turn to next. The real problem is that this approach has absolutely no idea of its own limitations. If you think that you already have all the right ideas – even though there are many things that you know you can't explain – you make it virtually impossible to advance beyond your own certainties.

Relativistic mass and rest mass

I would like to show that every one of the claims made so authoritatively on this site – and echoed in many others – is extremely problematic and that many are certainly wrong. The problem begins with the concept of 'mass'. There is a prejudice among many recent authors that the word 'mass' should refer only to 'invariant' or 'rest' mass and should not be applied to what is called 'relativistic mass', which includes energy and momentum. Einstein's special theory of relativity certainly introduced significant new features into dynamics, in particular, the Lorentz transformations and the gamma factor, which make the behaviour of particles at very high speed different to what would have been predicted before about 1880. But, defining mass as 'rest mass' or 'invariant mass' – and insisting that the word can only be used in this context – deliberately makes it impossible to connect supposedly 'classical' dynamics with dynamics post-Einstein. This is totally against Einstein's own practice which was to establish smooth connections with classical results.

There is no problem whatever with dynamics, pre- and post-relativistic if we define all dynamical concepts according to their relativistic meanings. Then all established dynamical laws remain valid. Preserving the laws of physics is more important than keeping to supposed definitions of concepts, which were never defined exactly in any case. Sticking to previous 'definitions' of concepts is like saying that we can't use the modern idea of chemical element because Lavoisier defined element differently. Physics concepts have always been evolutionary in a way that laws have not. Their meaning is defined by the laws. One of the concepts which has evolved is that of mass, which, as relativistic mass, now includes energy and momentum. The only previous definition as 'quantity of matter' in no way corresponds to what we now call rest mass, any more than time has to be what we call 'proper time'.

But 'rest mass' as a fundamental concept is nonsense in any case. No object ever observed has a mass corresponding to its rest mass. In addition, *all* mass of any kind is dynamic in origin. There is no 'mass' due to a state of rest. What we call the rest mass of protons and neutrons, and the majority of ordinary matter in the universe, for example, is created by the dynamics of gluon exchange. Objects made of ordinary matter may look 'solid' but nearly all their mass comes from the energy of gluon exchange. And the quarks of which the neutrons and protons are composed have masses which are certainly not invariant, but depend on the energy of interaction. But even the most invariant of 'invariant' masses, e.g. the mass of the electron, are dynamic in origin, because here it is the subtle motion known as *zitterbewegung*, the oscillation of the particle between real space and vacuum. The seemingly fixed mass of the stationary electron is merely a frequency of oscillation.

Overall, 'rest mass' or 'invariant mass' is just a special case of the much more significant concept of mass in general, just as proper time

is a special case of the more general concept of time. And even the originator of the modern concept of mass, Newton, by no means saw it as an idea fixed in stone. In his work in optics, he was exploring a more dynamic idea of mass in relation to light corpuscles or photons and the structure of matter. And, if we adopt the more evolved concept of mass, we can once again preserve Newton's law of gravitational force, just as we can the conservation laws of energy, momentum and angular momentum.

The masses of protons and neutrons

To return to the Higgs mechanism, we can ask: is it true that W / Z bosons, quarks and leptons owe their masses to Higgs, but neutrons and protons do not? Here, the assumption is that, while more fundamental fermions and gauge bosons gain mass by coupling to the Higgs particle, the composite structures made up of quarks, like neutrons and protons, cannot possibly gain most of their mass this way because the quarks which they contain aren't heavy enough. The only other mechanism which seems available to generate the mass is the dynamic transfer of the gluons which hold the quarks together inside the proton and neutron structures.

There is no doubt that this explanation is true, but it creates a problem of its own. Gluons, like photons, are massless (i.e. they have no 'rest mass'). There is no obvious way that massless gluons can generate 'solid' rest mass in the composite object. This is such a problem (the 'mass-gap problem') that there is a substantial financial prize offered by the Clay Institute for its solution. Since we don't have an accepted answer for this problem, it is presumptuous to the point of hubris *to assume that it cannot be due to the Higgs mechanism*. And there is, in fact, evidence that it is. A characteristic signature of the Higgs is

a conversion of handedness. The standard theory of the weak interaction tells us that only left-handed fermions and right-handed antifermions interact weakly. Left-handed antifermions and right-handed fermions do not. A typical fermion, say an electron, will have a wavefunction which is a superposition of left- and right-handed components. The left-handed component interacts weakly, the right-handed component does not.

So, how do we detect the presence of the right-handed component? The answer is: by the presence of mass. According to a heuristic argument, a massless electron would travel at the speed of light and couldn't be overtaken. But if it had any amount of rest mass, it could, in principle, be overtaken, and seen as in the opposite state of handedness. The amount of mass becomes a *measure* of the fraction of right-handedness in a fermion state. So a tau is more right-handed than a muon and a muon than an electron.

Let us now look at the type of nilpotent wavefunction which we have long advocated for 3-quark particles like the proton and neutron, governed by the strong interaction (Rowlands, 2007): Here, the vector aspect of the strong charge requires a source term and corresponding vacuum with three components. The momentum term provides an obvious source. There is no way we can combine three components in the form:

$$(i\mathbf{k}E \pm \boldsymbol{i}\mathbf{p} + \boldsymbol{j}m)\,(i\mathbf{k}E \pm \boldsymbol{i}\mathbf{p} + \boldsymbol{j}m)\,(i\mathbf{k}E \pm \boldsymbol{i}\mathbf{p} + \boldsymbol{j}m)$$

which automatically reduces to zero. But we can propose a three-component structure where the vector nature of $\mathbf{p}$ plays an explicit role

$$(i\mathbf{k}E \pm \boldsymbol{i}\,\mathbf{i}p_x + \boldsymbol{j}m)\,(i\mathbf{k}E \pm \boldsymbol{i}\,\mathbf{j}p_y + \boldsymbol{j}m)\,(i\mathbf{k}E \pm \boldsymbol{i}\,\mathbf{k}p_z + \boldsymbol{j}m)$$

(For convenience, we use only the first term of the Dirac spinors.) This can be shown to be nilpotent when $\mathbf{p} = \pm\,\boldsymbol{i}\,\mathbf{i}p_x$, $\mathbf{p} = \pm\,\boldsymbol{i}\,\mathbf{j}p_y$, or $\mathbf{p} = \pm\,\boldsymbol{i}$

$\mathbf{k}p_z$, that is, when the momentum is directed entirely along the (arbitrarily defined) x, y, or z axes, in either direction.

The complete wavefunction combines information from six allowed independent nonlocally gauge invariant phases, all existing simultaneously and subject to continual transitions at a constant rate:

$$
\begin{aligned}
&(i\mathbf{k}E + \mathbf{i}\mathbf{i}p_x + \mathbf{j}m)\ (i\mathbf{k}E + \ ...\ + \mathbf{j}m)\ (i\mathbf{k}E + \ ...\ + \mathbf{j}m) + RGB\\
&(i\mathbf{k}E - \mathbf{i}\mathbf{i}p_x + \mathbf{j}m)\ (i\mathbf{k}E - \ ...\ + \mathbf{j}m)\ \ (i\mathbf{k}E - \ ...\ + \mathbf{j}m)\ - RBG\\
&(i\mathbf{k}E + \ ...\ + \mathbf{j}m)\ (i\mathbf{k}E + \mathbf{i}\mathbf{j}p_y + \mathbf{j}m)\ (i\mathbf{k}E + \ ...\ + \mathbf{j}m) + BRG\\
&(i\mathbf{k}E - \ ...\ + \mathbf{j}m)\ (i\mathbf{k}E - \mathbf{i}\mathbf{j}p_y + \mathbf{j}m)\ (i\mathbf{k}E - \ ...\ + \mathbf{j}m)\ - GRB\\
&(i\mathbf{k}E + \ ...\ + \mathbf{j}m)\ (i\mathbf{k}E + \ ...\ + \mathbf{j}m)\ (i\mathbf{k}E + \mathbf{i}\mathbf{k}p_z + \mathbf{j}m) + GBR\\
&(i\mathbf{k}E - \ ...\ + \mathbf{j}m)\ (i\mathbf{k}E - \ ...\ + \mathbf{j}m)\ (i\mathbf{k}E - \mathbf{i}\mathbf{k}p_z + \mathbf{j}m)\ - BGR
\end{aligned}
$$

Here, we have retained the two possible spin states, for reasons that will soon become apparent.

All possible phases will be some kind of superposition of these. After normalization, these reduce to

$$
\begin{aligned}
&(i\mathbf{k}E + \mathbf{i}\mathbf{i}p_x + \mathbf{j}m) \quad +RGB\\
&(i\mathbf{k}E - \mathbf{i}\mathbf{i}p_x + \mathbf{j}m) \quad -RBG\\
&(i\mathbf{k}E - \mathbf{i}\mathbf{j}p_y + \mathbf{j}m) \quad +BRG\\
&(i\mathbf{k}E + \mathbf{i}\mathbf{j}p_y + \mathbf{j}m) \quad -GRB\\
&(i\mathbf{k}E + \mathbf{i}\mathbf{k}p_z + \mathbf{j}m) \quad +GBR\\
&(i\mathbf{k}E - \mathbf{i}\mathbf{k}p_z + \mathbf{j}m) \quad -BGR
\end{aligned}
$$

The symbols on the right show the equivalents in the conventional ‘colour’ representation.

What is very significant here is the sign change in the **p** component of the third and fourth terms, because the final sign depends on whether the expression containing non-zero **p** is in the middle or end bracket. Because of this switch, the perfect symmetry of the strong

interaction cannot be maintained without both + and –**p** components being present. Handedness is determined by the relative signs of E and **p** components, and so this means that both left- and right-handedness must be equally present. There has to be a maximal superposition of left- and right-handed components, thus explaining the observed zero chirality in the interaction.

The six phases clearly exhibit an $SU(3)$ structure, with eight generators and a wavefunction, exactly as in the conventional model,

$$\psi \sim (BGR - BRG + GRB - GBR + RBG - RGB).$$

The 'colour' transitions between the structures require either an exchange of the components of **p** between the individual 'quarks' or components of the baryon, or relative switching of the component positions. It is nonlocal and so independent of any real distance between the components. The transition must be gauge invariant, as no direction is privileged, and the mediators must be massless, exactly as we find with the conventional massless gluons.

In an exact parallel with the conventional theory of the strong interaction, six of our gluons will be:

$$(i\mathbf{k}E + \mathbf{i}\mathbf{i}p_x)(-i\mathbf{k}E + \mathbf{i}\mathbf{j}p_y) \quad (i\mathbf{k}E + \mathbf{i}\mathbf{j}p_y)(-i\mathbf{k}E + \mathbf{i}\mathbf{i}p_x)$$
$$(i\mathbf{k}E + \mathbf{i}\mathbf{j}p_y)(-i\mathbf{k}E + \mathbf{i}\mathbf{k}p_z) \quad (i\mathbf{k}E + \mathbf{i}\mathbf{k}p_z)(-i\mathbf{k}E + \mathbf{i}\mathbf{j}p_y)$$
$$(i\mathbf{k}E + \mathbf{i}\mathbf{k}p_z)(-i\mathbf{k}E + \mathbf{i}\mathbf{i}p_x) \quad (i\mathbf{k}E + \mathbf{i}\mathbf{i}p_x)(-i\mathbf{k}E + \mathbf{i}\mathbf{k}p_z)$$

and two will come from combinations of

$$(i\mathbf{k}E + \mathbf{i}\mathbf{i}p_x)(-i\mathbf{k}E + \mathbf{i}\mathbf{i}p_x) \quad (i\mathbf{k}E + \mathbf{i}\mathbf{j}p_y)(-i\mathbf{k}E + \mathbf{i}\mathbf{j}p_y)$$
$$(i\mathbf{k}E + \mathbf{i}\mathbf{k}p_z)(-i\mathbf{k}E + \mathbf{i}\mathbf{k}p_z)$$

And so, contrary to the over-confident claims of many physicists, made, as so often, when the full picture is far from established, the *SU*(3) mechanism of the strong interaction, seen from the perspective of a nilpotent representation, provides mass for the proton and neutrons and other baryonic matter *by the Higgs mechanism*, at the same time as directing it through the process of gluon exchange.

In addition, the claim that the Higgs field is not universal was made on the basis that it does not apply to the strong interaction, as it does to the electric and weak. This claim appears to be falsified by our argument, and we may also guess that the lack of connection to gravity can be falsified as well.

The mass of the Higgs boson

So, is the Higgs mechanism connected with gravity? Does the mass of the Higgs boson derive from the Higgs mechanism? Let us consider the boson first. At ANPA in August 2011, before the first indications of a discovery had been made in December 2011, I made a prediction that the mass would be of the order 123 GeV (Rowlands, 2012 / 2015). This was for a very specific reason: *that the Higgs boson mass should be generated by the Higgs mechanism*. It is actually the spin 0 nature of the Higgs boson, emphasized in the website because it is seemingly unlike spin 2 gravity, that makes it possible to explain its mass by the Higgs mechanism, because this very mechanism forces us to give a spin 0 boson its mass. As explained on many previous occasions, a spin 1 boson can be massless in the nilpotent formalism, but a spin 0 boson cannot, as then

$$(\pm i\boldsymbol{k}E \pm \boldsymbol{i}\mathbf{p})\,(\mp i\boldsymbol{k}E \mp \boldsymbol{i}\mathbf{p})$$

immediately reduces to zero: hence ‘massless’ Goldstone bosons must become massive Higgs bosons in the Higgs mechanism.

So, a left-handed fermion ($\pm i\mathbf{k}E \pm \mathbf{i}\mathbf{p}$) cannot combine with a left-handed antifermion ($\mp i\mathbf{k}E \mp \mathbf{i}\mathbf{p}$), via a weak interaction, to form a bosonic single state unless a nonzero mass term is introduced. In the nilpotent formalism a spin 0 boson has either the structure:

$$(\pm i\mathbf{k}E \pm \mathbf{i}\mathbf{p} + \mathbf{j}m)\,(\mp i\mathbf{k}E \mp \mathbf{i}\mathbf{p} + \mathbf{j}m)$$

or

$$(\pm i\mathbf{k}E \mp \mathbf{i}\mathbf{p} + \mathbf{j}m)\,(\mp i\mathbf{k}E \pm \mathbf{i}\mathbf{p} + \mathbf{j}m)$$

Both components have the same handedness. So, we must have either a left-handed fermion (allowed) and a left-handed antifermion (forbidden), or a right-handed fermion (forbidden) and a right-handed antifermion (forbidden).

If we imagine that the mass is created to accommodate the forbidden component, then we can suppose that the mass is created by the ‘wrong’ half of the vacuum. If the total expectation value of the Higgs vacuum is 246 GeV then the Higgs boson will acquire one set of 123 GeV by one method and the other set of 123 GeV by the other. As I wrote earlier: this is ‘the maximal energy required for either its fermion or its antifermion component to have the ‘wrong’ handedness and so produce the spin 0 combination’.

This was my reasoning before the experimental discovery of a Higgs boson at 125 GeV. We can imagine that any slight ‘discrepancy’ is due to the self-coupling of the Higgs, which is allowed for as a source of a small part of the Higgs mass in conventional theory. The same kind of mechanism might be responsible for the fact that the four electroweak bosons, γ, W^+, W^-, Z^0, have a total mass of 251 GeV rather than 246.

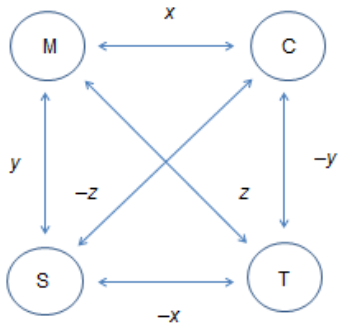
The Higgs field and gravity

So, is the Higgs field connected to gravity? The Higgs field is *absolutely continuous*, 246 GeV at every point in space. This has many important consequences: The energy density of the universe is infinite, which is exactly the density from which the universe is conjectured to have begun in the big bang. We can imagine that this situation exists all the time and that there is no such thing as *movement* of mass. Different points in the continuum have different manifestations of the value, e.g. rest mass, charge, etc., but none exceeds 246 GeV. We can see the manifestations moving through the continuous mass in the same way as wave motions move through a fixed medium. No elementary particle can exist at a higher rest energy than 246 GeV, and probably (on the basis of quantum field theory) $246 / \sqrt{2} = 174$ GeV, which seems to be the mass of the top quark.

It is curious that there is no ordinary mechanical movement of charges or fermions at velocity *v*, only *zitterbegung* at *c*, which is not a mechanical velocity. The concept of 'motion' reduces to frequency. It is also curious that one of the models that Newton considered as possible for gravity had mass and aether as mutually exclusive. Higher density mass meant lower density aether and vice versa. Lower density aether meant higher gravity.

A continuous mass distribution is required by the symmetry of the Klein-4 parameter group of space-time-mass-charge (S-T-M-C), dating back, in my work, to the 1970s (Rowlands, 2007). The continuity explains why mass is never negative, and why mass at absolute rest is never observed. Perhaps the immobility of the mass can be related to the extraordinary fact that the centre of mass of any system remains the same after any interaction. Applied to the infinite number of possible systems, this requires the immobility of an infinite number of centres of mass fixed at any point.

In the following 'graph' of the parameter group, *x* represents the property of conservation and *–x* of nonconservation; *y* represents the property of real representation (norm 1) and *–x* of imaginary representation (norm –1); while *z* represents the property of continuity, nondimensionality and commutativity, and *–z* that of discreteness, dimensionality and anticommutativity. The exactness of these property / antiproperty distinctions is also manifest in other visual representations (cubic, tetrahedral and colour). Mass(-energy), like time, has to be absolutely continuous because it is opposite in this property to the discrete and dimensional quantities space and charge.



A continuous mass distribution, which seems to be generic, also occurring with the zero-point energy, the cosmic microwave background radiation and even ordinary fields, has implications for gravity. It explains why the static gravitational field is continuous, though this is now widely accepted as being also *required by Lorentz invariance* (or special relativity). As is now generally accepted, the gravitational field is not propagated. It implies also that there is a mechanism for nonlocality because absolute continuity implies instantaneous correlation. This is especially important if there is any concept of gravity / gauge theory correspondence, with gravity being the nonlocal counterpart to the three local gauge interactions.

Many people, including the author of the website quoted, assume, without any known evidence, that gravity is a quantized local interaction, mediated by a spin 2 graviton. This is by no means established, and the spin 2 theory has failed so far on all tests of renormalization. There is no reason why gravity should be seen as a local gauge theory. It could well be seen as the vacuum or nonlocal side of the quantized gauge theories, especially as it involves negative, rather than positive, energy. Certainly it is 'quantum', but quantum theories require vacuum and nonlocality as well as local interactions.

All interactions have both local and nonlocal aspects. With electric, strong and weak interactions, the local is, in effect, the 'real' manifestation and the nonlocal or vacuum is the 'virtual' or 'fictitious' one. With gravity, it is quite possible that the vacuum or nonlocal effect is the 'real' one and the local aspect the virtual or fictitious counterpart. This would be entirely in accord with gravity / gauge theory correspondence. In this case, the local aspect would be inertia, which is widely recognized as a 'fictitious' or virtual force. Since inertia is repulsive, rather than attractive, its quantized manifestation would be spin 1 rather than spin 2, so eliminating the renormalization problem.

Certainly, if conventional quantum gravity is true, there is no continuous vacuum to which it could be referred. Paradoxically, quantum gravity cannot be a true quantum theory because it excludes the vacuum, whereas Newtonian gravity, with its nonlocal continuous interaction can.

Classical and quantum physics

A more general area of difficulty is the 'boundary' between classical and quantum physics. We are led to believe that physics has a classical regime and a quantum regime, with the added complication that there is also a boundary between classical and relativistic physics. I

think this is a completely wrong way of looking at physics, and reflects historical rather than physically meaningful distinctions. In fact, physics is at all times quantum and relativistic, and the concept of classical physics has no fixed meaning. It simply reflects a multiplicity of ways of deriving explanations using a minimum number of assumptions. Classical physics, when it confronts new issues, just creates new categories under which they can be managed.

Maxwell's equations, for example, manage to incorporate electricity, magnetism and optics into a single theory, but they don't assume discrete charges. They are classical in appearance but surreptitiously introduce special relativity or equivalent, with a slow resolution of the discrepancies during the late nineteenth century and early twentieth. More disruptive in the long term is the introduction of discrete charges at the end of the nineteenth century. Maxwell avoided the problems these created. But once we assume discrete charges – hinted at by electrolysis but forced on us by experiments later in the century – a wholly different perspective opens up.

Classical physics simply incorporates discrete charges, though realising that there is a renormalization problem. However, discrete charges ultimately require quantization, especially when combined with special relativity. From discrete charges, it is almost trivial to generate an energy proportional to frequency of transfer of information at the fixed speed of light.

$$E = \frac{e^2}{r} = \frac{e^2}{ct} \propto \nu$$

We also talk about 'classical' thermodynamics, yet the very basis of thermodynamics, 'heat', is not classical at all, and bears all the signs of being quantum mechanical in origin, resulting in a random motion that has no place in a classical theory, although there it is disguised as

random because of our lack of knowledge rather than because of something intrinsic. However, by subsuming 'heat' under a single unanalysed category of energy we can bodily incorporate it into a simplified description that looks classical. The second law of thermodynamics is an *ad hoc* concept in a classical sense but makes much more sense in quantum terms.

Quantum physics seemingly produces many new results totally at variance with classical ones. However, many quantum results can be repackaged to look classical. For example, many quantum results have been reimagined using what is called stochastic electrodynamics, which uses a Maxwellian theory with the extra assumption that there is a background field of $\frac{1}{2}h\nu$ for every possible frequency of radiation. It is not that there is no quantum element, rather that the quantum element is built into the theory in a classical way to produce a minimal nonclassical element. The concept of 'classical' is actually minimally redefined to include an extra component, which can be relatively easily fitted to the existing theory.

There are a lot more examples. A colleague is at the moment quite successfully generating the Higgs mechanism from a scalar field added to Maxwell's electromagnetic theory (J. B. Dainton, personal communication). I have myself produced a version of the Dirac equation using Lou Kauffman's discrete version of differentiation, which can be interpreted as either classical or quantum because the $i\hbar$ factor which determines whether the time and space differentials are classical or quantum can be applied without affecting the structure of the equation (Rowlands, 2007).

We start by defining a discrete differentiation (following Lou Kauffman) of the function F, which preserves the Leibniz chain rule, by taking:

$$\partial F / \partial t = [F, H] = [F, E] \quad \text{and} \quad \partial F / \partial X_1 = [F, P_i],$$

with $H = E$ and P_i representing energy and momentum Taking the case where velocity operators are not in evidence, we use $\partial F / \partial t$ rather than dF / dt. The mass term (which, of course, has only a passive role in quantum mechanics) disappears in the operator, though it is retained in the amplitude. Suppose we define a nilpotent amplitude

$$\psi = i\mathbf{k}E + \mathbf{i}\mathbf{i}P_1 + \mathbf{i}\mathbf{j}P_2 + \mathbf{i}\mathbf{k}P_3 + \mathbf{j}m$$

and an operator $$\mathcal{D} = i\mathbf{k}\frac{\partial}{\partial t} + \mathbf{i}\mathbf{i}\frac{\partial}{\partial X_1} + \mathbf{i}\mathbf{j}\frac{\partial}{\partial X_2} + \mathbf{i}\mathbf{k}\frac{\partial}{\partial X_3}$$

with

$$\partial\psi / \partial t = [\psi, H] = [\psi, E] \quad \text{and} \quad \partial\psi / \partial X_1 = [\psi, P_i].$$

After some basic algebraic manipulation, we obtain

$$-\mathcal{D}\psi = \psi(i\mathbf{k}E + \mathbf{i}\mathbf{i}P_1 + \mathbf{i}\mathbf{j}P_2 + \mathbf{i}\mathbf{k}P_3 + \mathbf{j}m)$$
$$-(i\mathbf{k}E + \mathbf{i}\mathbf{i}P_1 + \mathbf{i}\mathbf{j}P_2 + \mathbf{i}\mathbf{k}P_3 + \mathbf{j}m)\psi - 2\,i(E^2 - P_1^2 - P_2^2 - P_3^2 - m^2).$$

When ψ is nilpotent, then

$$\mathcal{D}\psi = \left(i\mathbf{k}\frac{\partial}{\partial t} + \mathbf{i}\mathbf{k}\boldsymbol{\nabla}\right)\psi$$

Generalising to 4 states, with $\mathcal{D}$ and ψ as 4-spinors, then

$$\mathcal{D}\psi = \left(\pm i\mathbf{k}\frac{\partial}{\partial t} \pm \mathbf{i}\mathbf{k}\boldsymbol{\nabla}\right)(\pm i\mathbf{k}E \pm \mathbf{i}\mathbf{i}P_1 \pm \mathbf{i}\mathbf{j}P_2 \pm \mathbf{i}\mathbf{k}P_3 + \mathbf{j}m) = 0.$$

The reason why this derivation is important is that it does not require the $\pm i$ (or $i\hbar$) term usually applied to the differential operators

in canonical quantization, though, because our expression incorporates all four sign variations, we can include this if we want to without changing anything significant in the equation. This means that we can make a smooth transition between classical and quantum conditions, which points to an important connection between mass and the appearance of $\pm i\hbar$ terms in quantum mechanical operators and amplitudes.

The differential operator is interesting because, lacking a mass term, no detectable difference occurs in changing the signs of E and **p**, so particle and antiparticle creation operators look the same. Perhaps, this gives us a way of representing Majorana particle creation and annihilation. The operator can be split into two parts with $+E$ and $-E$, representing particle and antiparticle, each with two signs of **p** or helicity.

Similar things also happen with relativity. Many people think of $E = mc^2$ as a relativistic equation because it first appears in this form in Einstein's work on special relativity. But special relativity cannot actually derive it, and it seems to be a purely classical equation which hardly ever appears in that context because it is not normally needed. A whole series of seventeenth, eighteenth and nineteenth century calculations show mc^2 as a (sometimes disguised) potential energy term in the particular case of light photons, which is totally classical. The only reason why Einstein fixes on it for the total energy of matter is because he needs relativistic energy considerations to coincide with classical ones.

Gravitational light deflection

The calculation of gravitational light bending gives a good illustration of the subtlety of how classical physics connects with seemingly more 'advanced' ideas. We are told that this requires the full field

equations of GR with curved space-time, etc. Newton, it is said, proposed a 'classical' effect, which can be calculated from the potential energy equation

$$mc^2 = \frac{GMm}{r}(e+1)$$

Here *e* is the eccentricity of a hyperbolic orbit and the bending in and out can be calculated from

$$\frac{2}{e} \approx \frac{2GM}{c^2 r}$$

This Newtonian calculation was actually carried out in the eighteenth century by Cavendish and Soldner, and can be related to gravitational time dilation of the photon (Cavendish, 1921, Soldner, 1801). In fact, however, experiment reveals that

$$\frac{2}{e} \approx \frac{4GM}{c^2 r}$$

just as predicted by general relativity. Nevertheless, we don't need general relativity. We can show that the 'double' bending occurs because photons are 'massless' or 'relativistic' and so undergo length contraction as well as time dilation. All we need is special relativity.

But the most interesting thing about the calculation is that it can be done in a perfectly classical way, if we use the *kinetic* energy equation

$$\frac{1}{2}mc^2 = \frac{GMm}{r}(e+1)$$

which may have been in the minds of both Cavendish and / or Soldner as both started from the black hole equation

$$\frac{1}{2}mc^2 = \frac{GMm}{r}$$

This is simply the equation of an object *brought into orbit*, not in an existing one. The black hole equation gives a limiting condition for this in the circular case. It is perfectly classical and gives the correct answer for light bending. It may be that Soldner actually carried this out and obtained the correct value, and not the half-value which is usually attributed to him, because his statement of the result may have concentrated on the half-bending in while leaving the half-bending out to be inferred.

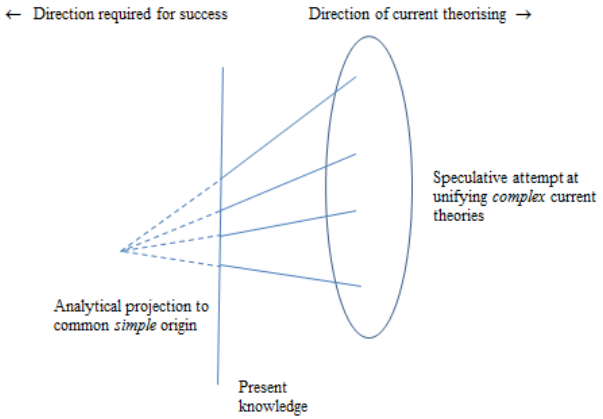
But to understand the equation and its meaning fully we need to understand something outside of classical physics as pursued earlier. We need to think of the 'velocity of light' as something other than a dynamical velocity, except when the light is subjected to an external force, such as gravity. Light always 'travels' at *c*, but this only becomes dynamic when it connects with a real force. So the *c* here is about *producing* an orbit, not being in one.

newtonian and Newtonian

Another problem occurs with the frequent use of the word 'newtonian'. The word 'newtonian' is often taken to mean the physics of identifiable particles in determined positions interacting with each other in identifiable ways, with identifiable outcomes. This should be distinguished from 'Newtonian', which refers to statements attributable to Newton himself. These often diverge widely. Newton's thinking is often not 'newtonian' at all. Strict development from assumed invariable principles (newtonian) produces an entirely different style of work from the inductive search for explanations of seemingly anomalous experimental results (Newtonian). (Rowlands, 2017a,b,c).The

reason why this is important is because of the innate assumption that physics is mostly a deductive sequence from accepted principles.

Newton provides a striking example of how this is not the case. When presented with intractable data he doesn't resort to newtonian models, but tries to establish abstract generalisations which minimalise assumptions to produce the maximal degree of explanation possible. Newton's reasoning in this style shows us many of the directions in which non-classical physics would subsequently go after his time. We can see from his reasoning why electromagnetic theory, quantum theory, relativity, thermodynamics and modern ideas on the structure of matter developed in the way that they did. The reason why this is significant is because it demonstrates how we should proceed now, given that we are presented with equally intractable data. In terms of the diagram below, we should not be attempting a 'grand synthesis' as pictured on the right-hand side, and as is all too common in contemporary theorizing, but a profound and simple *analysis*, as pictured on the left.



General relativity

General relativity has, in my opinion, been a particular problem for the development of physics in the last hundred years, not because it produces incorrect results or predictions, but because it uses a particularly eccentric style of reasoning, based on the needs of a particular individual at a particular time, which has become set in stone as the only correct way of proceeding, even though it is at odds with the rest of physics.

We frequently hear that general relativity has shown that gravity is due to the curvature of space-time and not due to a Newtonian type of force. To *really* understand nature we must throw away our old concepts and take on the new ones. There are many things wrong with this. First of all, the concept of 4-D space-time as a physical object is completely refuted by relativistic quantum mechanics, the basis of all of the rest of physics. Relativistic quantum mechanics uses an object such as ($i\mathbf{k}E + \mathbf{i}\mathbf{p} + \mathbf{j}m$) or ($i\mathbf{k}t + \mathbf{i}\mathbf{r} + \mathbf{j}t$) of which the invariant first four components ($i\mathbf{k}E + \mathbf{i}\mathbf{p}$) or ($i\mathbf{k}t + \mathbf{i}\mathbf{r}$) are not a 4-D (4-vector) entity, which would require ($iE + \mathbf{p}$) or ($it + \mathbf{r}$). As wave-particle duality and *the non-observability of time* show, there is no physical union, just a mathematical construct, which is more sophisticated than general relativity assumes.

Curvature of a mathematical construct certainly does not have the physical significance that we would expect in curvature of a physical object. Special relativity has not reduced space and time to a single physical entity (as Minkowski famously proclaimed) and general relativity has not reduced it to a curved one. We can use a curvature *metaphor* if we really want to but it has no privileged significance with respect to any other. There is no ontological significance. The concept of gravitational 'force' is at least as valid, and a 'field' interpretation of general relativity is a much more user-friendly approach than the intractable curvature method, and probably more 'correct' as it doesn't rely on the definitely unphysical union of space and time.

With regard to Newtonian gravity, we can say at least three things:
(1) Newton, in proclaiming a law of gravity, is careful to avoid commitment to any physical interpretation, force or otherwise. Force is an effect, not an explanation. Curvature is one way of achieving the effect and no more or less valid than any other 'physical' interpretation.
(2) Cartan (1923, 1924) showed that *any* force, gravitational or otherwise including Newtonian gravity, could be represented by a curvature term.
(3) Newton used a version of the curvature method himself in his theory of revolving orbits (*Principia* I, Proposition 44, Corollary 6).
(4) In addition, the universe itself has refused to be curved, despite a great deal of belief that it would be. It has stayed resolutely Newtonian and Euclidean. The curvature method, in fact, has forced physics into a cul-de-sac. It has separated gravity from the gauge forces, as though it were completely divorced from them.

The justification for all of this is that the general relativistic field equations have produced correct predictions for all experimental tests to date. This is true, but not one of the tests requires the full panoply of these equations and every one of them was either fully or partially predicted before general relativity on the basis of the equivalent of special relativity and Newtonian gravity.

But it is not the field equations of general relativity that are the problem. It is the way they have been used to create physical interpretations that they cannot be shown to support. As mentioned, there is a perfectly good field interpretation of general relativity that uses fields in the ordinary way instead of curvature. And this interpretation seems to avoid the perils of the self-interaction with positive feedback mechanism that many seem to see as being necessitated by the field equations and causing their ultimate destruction in strong fields.

Experimental evidence to date on short-period pulsars and on gravitational waves seem to suggest that the equations do not self-destruct

in strong fields. There is no positive feedback. If this is so, then we may be producing the wrong interpretation of the 'magnetic' component or relativistic corrections associated with gravity. It has always been assumed that this is intrinsic or ontological (hence the positive feedback), but a nonintrinsic or epistemological interpretation would be much more coherent and remove all the remaining problems. Essentially, this means defining inertia as a fictitious force, so ensuring that the Earth has no special status in the universe. It would preserve the equations even in the strongest fields, and remove the spin 2 quantum description and its nonrenormalizable infinities. It would ensure that mass retained continuum properties and provide mechanism for nonlocality and vacuum interactions.

A Machian origin of inertia, based on the idea that the gravitomagnetic field is the origin of inertia and $F = ma$, when combined with the principle of equivalence, leads to the equation:

$$\frac{Gm_H}{c^2 r}\frac{dv}{dt} = \frac{Gm_H}{r_H^2}$$

and hence to the relation between the redshift acceleration (dark energy) and velocity:

$$a = \frac{dv}{dt} = \frac{c^2 r}{r_H^2}$$

This is equivalent to a dark energy of 67 %, which may be compared to the current experimental value of 68 %, but it was predicted as long ago as the 1970s. The ultimate rationale was a continuous mass distribution in a Euclidean universe – now both experimentally demonstrated. In my view this would mean the subtlety in the interpretation of general relativity, mentioned above but not previously appreciated.

This is that the magnetic component is measurable but is *intrinsically fictitious* because mass is continuous. (Rowlands, 2007)

Here Newton, as so often, has something useful to say. He describes mass as *vis inertia* because he thinks of it as a kind of 'force' opposing real ones according to the third law. 'Real' or 'active' forces such as gravity leads to the 'passive' response which we call inertia. In an unpublished Proposition 43, intended to end Book III of the *Principia*, he argued that the Earth could be at the centre of a Tychonic Solar System with the Sun and its planets in orbit round it if gravity was countered by an equal and opposite *inertial* force proportional to mass originating from outside the Solar System (Weinberg, 2015).

A *fictitious* force of this kind occurs if the Earth is taken as an inertial frame with respect to the rest of the universe, but a real force would only exist if the uniformity of nature was not obeyed, and the Earth was singular with respect to the other planets. If we take inertia as representing the 'magnetic' aspect of the gravitational force, we may be obliged to take it as fictitious if the Earth's situation in space is not regarded as special. So, perhaps the non-unique nature of our planet finally decides on the interpretation of general relativity that leads to the maximum consistency. It is one which would be particularly consistent with the field interpretation.

The general relativistic equations remain and all results so far found and many yet to be discovered. But there will be no spin 2 graviton (only a spin 1 'inertion') and no positive feedback nonlinearity. There would also be more convenient 'field' representations rather than the 'rigid' assumption that only curvature can be applied. General relativity would be even more valid than presently assumed, surviving in strong fields, as experiment seems to suggest it does.

2-dimensionality and 3-dimensionality

A particularly significant error in physics is the linking of 2- and 3-dimensionalities as though they were aspects of a single phenomenon. In fact, in mathematics as well as in physics, they are distinct in origin and meaning, and 2-dimensionality does not emerge as a limited form of 3-dimensionality, but is a totally separate concept. 2-dimensionality in mathematics and physics is based on the complex Argand plane where the real axis and imaginary axis represent distinct mathematical structures, which do not have rotation symmetry. There is no direct Pythagorean addition, only a multiplication of complex conjugates. This means that to extract meaning from a 2-D system, we have to accept the simultaneous and equal existence of the two signs of the imaginary quantities. There is no way of separating $+ i$ and $- i$. This is quite different from the quaternion structure which is the basis of 3-dimensionality. Here we have three axes which have an equivalent status, and rotation symmetry, and true Pythagorean addition is possible without the complex conjugate.

This distinction has profound consequences in physical terms in the Standard Model, where the $SU(2)$ structure of the weak interaction is based on the imaginary sign of the energy term, with its built-in duality, which eventually becomes the duality of fermion and vacuum, and leads to the interaction's intrinsic chirality, with one sign privileged as particle and the other as vacuum. This necessarily leads to the massiveness of the gauge boson and, because of the twoness of the dimensionality relying on real and imaginary components, which are associated with different physical terms in the nilpotent structure ($\pm i\mathbf{k}E \pm \mathbf{i}\mathbf{p} + \mathbf{j}m$), which are in turn associated with different charges, it ultimately results in the $SU(2)$ of the weak interaction not being a free-standing symmetry but only a component of the combined electroweak $SU(2) \times U(1)$. The $SU(3)$ symmetry of the strong interaction,

with its complete rotational invariance and its massless gauge bosons, is altogether different. It is confined to the ***i*p** term of the nilpotent and is a free-standing and exact symmetry, completely devoid of chirality.

In the Klein-4 group, the quaternion system, which represents charges, is exact enough in its Pythagorean addition and its precise relations between the three axes, to make the charges conserved quantities, like the real quantity mass. When an additional 2-dimensionality is incorporated to accommodate space by defining its components as complexified quaternions, with a consequent relation with the imaginary parameter time, we see that the axes are no longer uniquely defined, and space, like time, becomes a nonconserved quantity. Given any one charge, then it is always possible to define all charges uniquely by their Pythagorean additions (such as occur in physical interactions), whereas the Pythagorean additions of spatial dimensions are non-specific, meaning that space cannot be defined in a unique or conserved way. (It is interesting that the anticommutativity, which defines quaternions, and hence 3-dimensional quantities, is represented by the spatial positions of the anticommuting terms, whereas the antiassociativity used in the larger system of octonions is only realizable if a *temporal* aspect is added.)

Some other observations

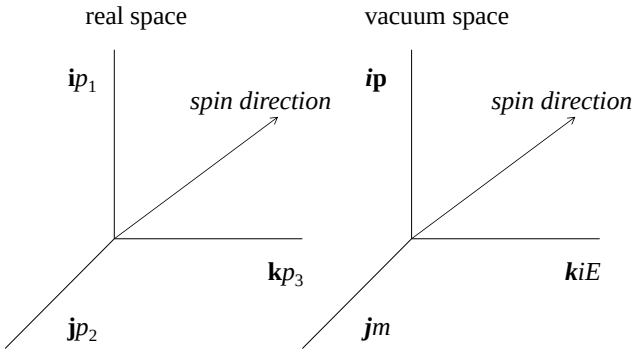
The LHC has shown no evidence of supersymmetric particles, yet theory seems to require a supersymmetry in which every fermion has a corresponding boson and vice versa. I have shown on numerous occasions how supersymmetry without extra particles emerges naturally from the nilpotent formalism. This has now been quoted in a major review and so is available to the mainstream community (Hauser and Dröscher, 2017).

A concept which has been an immediate consequence of the symmetry of space, time, mass and charge since the 1970s is that the proton can never decay into a neutral pion and positron as many theories have advocated. This is because of the conservation of type of charge. The strong component must be preserved. This is also equivalent to quark conservation. We can only remove the quark structure if we also have antiquarks.

Another problem, previously referred to, in both classical and quantum physics regarding point charges, where close to the charge the energy blows up to infinity. In quantum physics this is because the charge polarizes the vacuum field surrounding it creating fermion-antifermion pairs which also interact with the charge. In the nilpotent quantum theory, however, there is no such thing as a point in 3-D space. Space is translation and rotation symmetric (nonconserved) and no point can be singled out. So to create a point you have to have an alternative 3-D space (vacuum space, incorporating charge, mass and time) with which real space interacts. It is the intersection between the spaces that creates the point-like fermion. Nilpotency means that it has 0 norm (size) at the intersection.

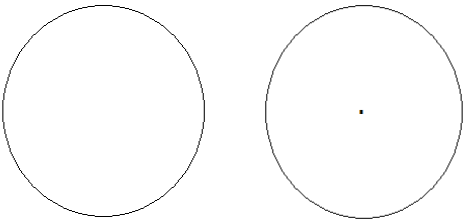
The duality of real and vacuum spaces as shown in Pauli exclusion

Pauli exclusion



antisymmetric wavefunctions nilpotency

A point-particle in space creates a singularity, which affects the topology, making the space multiply connected, one of the consequences of which is ½ spin.



simply-connected space multiply-connected space

Another consequence is the switching of the particle between real and vacuum space at the Compton frequency, which we call *zitterbewegung*. From this we derive a fixed invariant mass for the particle. So, defining a point charge also creates a finite mass scale. The renormalization cut off is determined by the smallest size to which the particle, say an electron, can be confined. This is across the two spaces and requires mass (or frequency) and charge to fix it. As mentioned

earlier, the other renormalization problem, the self-interaction, disappears entirely with the intrinsic supersymmetry of the nilpotent wavefunction. If we are to progress beyond the complicated structure of the Standard Model, which has remained more or less unexplained since 1973, it is my belief that we need to pay close attention to the kind of subtle points which are discussed here and not simply take its mathematical formalisms as a given.

References

Cartan, E., *Ann. Ecole Norm*., 40, 325, 1923

Cartan, E., *Ann. Ecole Norm*., 41, 1, 1924

Cavendish, Henry, in E. Thorpe (ed.), *The Scientific Papers of the Honourable Henry Cavendish, F.R.S., Volume II: Chemical and Dynamical*, London, 1921, 433–437.

Hauser, Jochem and Dröscher, Walter, 'Gravity beyond Einstein? Part I: Physics and the Trouble with Experiments', *Z. Naturforsch*, 72(6)a, 493-525, 2017 (see p. 496, note 3)

Rowlands, P. 2007, *Zero to Infinity The Foundations of Physics*, Singapore, London and Hackensack, NJ: World Scientific

Rowlands, P. 2013 / 2015 Nonlocality, in A. D. Ford (ed.) *Haecceity: Proceedings of ANPA 32*, 2012, 227-250 and Grenville J. Croll and Nicky Graves Gregory, *Explorations, Proceedings of ANPA 35*, 2015, 195-222 (also DVD recording by Anthony Blake)

Rowlands, P. 2017a, *Newton and Modern Physics*, Singapore, London and Hackensack, NJ: World Scientific

Rowlands, P. 2017b, *Newton and the Great World System*, Singapore, London and Hackensack, NJ: World Scientific

Rowlands, P. 2017c, *Newton – Innovation and Controversy*, Singapore, London and Hackensack, NJ: World Scientific

Soldner, Johann von, *Astronomisches Jahrbuch f¨ur das Jahr 1804*, Sp¨athen, Berlin, 161, 1801, dated March 1801, reprinted by P. Lenard in *Analen der Physik*, 65, 1921, 593, translated in S. L. Jaki, *Found. Phys.*, 8, 927–950, 1978

Strassler, Matt, 'Why the Higgs and Gravity are Unrelated' (15 October 2012) (https://profmattstrassler.com/2012/10/15/why-the-higgs-and-gravity-are-unrelated/)

Weinberg, Steven, *To Explain the World: The Discovery of Modern Science*, HarperCollins, 2015, 251–252.

THE CAUSAL (DETERMINISTIC) APPROACH TO QUANTUM MECHANICS

NINA SOTINA

New York, USA

nsotina@gmail.com

Abstract: In the present work, the causal (deterministic) approach to quantum mechanics is developed. The critical analysis of D. Bohm's work *A suggested interpretation of the quantum theory in terms of 'Hidden Variables'* is given. Using the causal approach, it is proven that Schrödinger equation is a necessary condition of stability of the motion of a particle. For the case of the hydrogen atom this approach gives a mathematical base with which to suggest that (1) the electron's spin in an atom is precessing; (2) the energy of the precessional motion on Bohr orbits satisfies Rydberg's formula. and (3) structures are formed in the physical vacuum that stabilize the motion of an electron on Bohr orbits.

1 INTERPRETATION CRISIS OF QUANTUM MECHANICS

Currently, there exists an interpretation crisis in quantum mechanics which is, in essence, due to the lack of comprehension of the underlying physics hidden behind the equations. While the mathematical formalism of quantum mechanics describes many experiments well there are heated discussions among scientists about its physical interpretation. The various interpretations offer different approaches to the issues that arise, such as: "collapse of the wave function", "paradoxes such as EPR", and etc. However, at the present time there is no any physical interpretation of quantum formalism that would not have contradictions within it or with accepted ideas and theories.

To support this statement let us consider two interpretations: the Copenhagen interpretation and the causal interpretation. Followers of the Copenhagen interpretation insist on the point of view that physics is the science

which rests solely on measurements. Thus, under this interpretation it is invalid to making any statements about the systems' properties prior to measurement. However, in the context of some experiments followers of the Copenhagen interpretation encounter logical difficulties, for example: in the experiments with the quantum-correlated photons it is necessary to assume that, although some properties of reality (such as polarization of each of the photons of the pair) exist before measurements only potentially there is, however, a correlation between them.

Besides that, there is a problem in reconciling the "collapse of the wave function" postulate, associated with process of measurement, with the special relativity.

The causal interpretation of quantum formalism centres about the existence of "hidden variables", which uniquely characterize the given state of a system; thus, allowing a quantum system to be consistent with the deterministic theory. In 1964 John S. Bell advanced his famous inequalities. It followed from the violation thereof in quantum theory that any theory of "hidden variables" claiming to be able to describe experimental results, must be "nonlocal".

Quantum nonlocality means one of the following: 1) the apparent ability of objects to instantaneously "know" about each other's quantum state, even when separated by large distances (non-separability is a term often used in this regard), 2) the possibility of action at a distant, or 3) the possibility of superluminal influences (it could be signalling of information or superluminal interactions which require any continuous process in spacetime, that is the existence of a physical field).

The prevalent point of view is that relativity prohibits any superluminal influences or action at a distant. It would be mistaken, however, to think that the idea of the quantum non-separability does not contradict the theory of relativity. There is an implicit postulate of separability in the theory of relativity (the state of any composite system is wholly determined by the separate states of its subsystems). Even Bohm noted that to measure the speed of light it is necessary to have a receiver and a transmitter that are not only separated in space, but are also autonomous in their behaviour.

Below we will develop the causal approach to quantum mechanics, assuming the existence of hidden variables like a physical field.

2 THE BOHM'S QUANTUM POTENTIAL

D. Bohm was a proponent of the causal interpretation of quantum theory. He was the first who introduced nonlocality in quantum mechanics [1]. Below we will recall the method which Bohm used to rewrite the Schrödinger's equation in classical terms and will point out where some of Bohm's assumptions were incorrect.

Bohm represented $\Psi(\vec{r}, t)$ in the following form

$$\Psi(\vec{r}, t) = A(\vec{r}, t) \exp(iS/\hbar), \tag{1}$$

substituted it into the non-stationary Schrödinger equation

$$i\hbar \frac{\partial \Psi(\vec{r}, t)}{\partial t} = -\frac{\hbar^2}{2m} \Delta \Psi(\vec{r}, t) + U(\vec{r}, t) \Psi(\vec{r}, t), \tag{2}$$

and separated real and imaginary parts to obtained the following two equations:

$$\frac{\partial S}{\partial t} + \frac{(\nabla S)^2}{2m} + U - \frac{\hbar^2}{2m} \frac{\nabla^2 A}{A} = 0, \tag{3}$$

$$-2m \frac{\partial A}{\partial t} = \nabla \cdot (A^2 \nabla S) \equiv A^2 \left[\Delta S + \frac{2}{A} (\nabla A \cdot \nabla S) \right]. \tag{4}$$

Equation (3) resembles the Hamilton-Jacobi equation but it contains not only potential energy U, but also another term

$$U_Q = -\frac{\hbar^2}{2m} \frac{\nabla^2 A}{A}, \tag{5}$$

Bohm called U_Q the "quantum potential".

If we express the velocity of a particle of mass m as

$$\bar{V} = (1/m) \nabla S, \tag{6}$$

then Eq. (4) is a continuity equation [1] for $m|A|^2$.

1 A continuity equation in physics is an equation that describes the transport of some quantity. See https://en.wikipedia.org/wiki/Continuity_equation

2.1 CRITICAL ANALYSIS OF BOHM'S INTERPRETATION OF EQUATIONS (3) & (5).

For simplicity, we will consider the *stationary* case; however, the analysis presented below can be generalized to the *non-stationary* case.

1) Bohm assumed that in the limit $\hbar \to 0$, U_Q goes to zero and Eq. (3) takes the form of the Hamilton-Jacobi equation with potential U. This assumption, however, was incorrect, which can be proven by calculating U_Q directly from the solutions of the Schrödinger equation for a hydrogen atom. Indeed, as it is known, these solutions can be written in the form

$$\psi_{nlk}\,(r,\vartheta,\varphi) = A_{nlk}(r,\vartheta)\ \exp(ik\varphi), \tag{7}$$

where $n = 1, 2, \ldots, \infty$; $l \leqslant n-1$; $k = 0, \pm 1, \ldots, \pm l$. Hence, for $n = 1$ there is one solution

$$\psi_{100} = A_{100} = \frac{1}{\sqrt{\pi}} e^{-\frac{r}{r_0}}\,. \tag{8}$$

The quantum potential for this solution as follows from Eq. (5) has a form

$$U_Q = -\frac{\hbar^2}{2m}\frac{\nabla^2 A_{100}}{A_{100}} = \frac{\hbar^2}{2m r_0}\left(\frac{2}{r} - \frac{1}{r_0}\right) \tag{9}$$

where m is the mass of the electron and r_0 is the Bohr radius given by

$$r_0 = \frac{\hbar^2}{m_e e^2}. \tag{10}$$

After substituting expression for r_0 into Eq. (9), we see that U_Q tends to infinity as $\hbar \to 0$.

The fact that the quantum potential becoming infinitely large as $\hbar \to 0$ agrees with the statement that the Schrödinger equation describes only the quantum phenomena.

2) Bohm made the following two statements about U_Q: first that U_Q has "the new feature of nonlocality", which, according to him, implied "an instantaneous connection between distant particles", and second that its effects might not fall off with the distance.

Bohm wrote: "*The quantum potential is different in many ways from classical potentials. The first key difference is that multiplication of the wave function by a constant does not change the quantum potential. The quantum potential can therefore*

still be large even when the wave function is small. This means that its effects do not necessarily fall off with the distance... The quantum information potential has the new feature of nonlocality, implying an instantaneous connection between distant particles".

First, let us note that the property of *an instantaneous connection between distant particles* is not new for potentials. Thus, action at a distant is a postulate of the Newtonian potential theory. Another example, is the pressure in an ideal incompressible fluid. The pressure is transmitted simultaneously over the entire volume, that is, there is *"an instantaneous connection between distant particles"*. Nonlocality here is merely a part of the physical model. In real fluids, there exists a pressure wave — a "precursor" which propagates with speeds much greater than that of the moving body.

U_Q, according to Bohm, however, is differ from other known potentials, because Bohm assumed its effects do not necessarily fall off with the distance. That was the reason Bohm assigned the meaning of ***information*** to potential U_Q.

The Bohm's assumption that U_Q's effect does not fall with the distance however is not supported by mathematical derivations. Thus, Eq. (3) for a stationary case has a form

$$U_Q = -\frac{\hbar^2}{2m}\frac{\nabla^2 A}{A} = \text{const} - U - \frac{(\nabla S)^2}{2m} \tag{11}$$

it follows

$$\nabla U_Q = -\nabla U - \nabla\left[\frac{(\nabla S)^2}{2m}\right] \equiv -\nabla U - \nabla\frac{mV^2}{2}\,. \tag{12}$$

If functions S and U_Q are obtained from the same solution of the Schrödinger equation then Eq. (12) becomes identity. It is expected, because Eq. (11) follows directly from the Schrödinger equation. Note, that U_Q has a concrete form, given by Eq (5), only for those solution that satisfy the Schrödinger equation.

In the case of a hydrogen atom, $U \sim \frac{1}{r}$, and $\frac{(\nabla S)^2}{2m} \sim \frac{1}{r^2}$ [see Eq. (28)], therefore, the ∇U_Q falls with the distance.

In the next section we will consider the Hamilton-Jacobi equation in which to the given potential energy U we introduce an additional potential energy Φ. We will prove under this approach that the Schrödinger equation is the condition of stability of motion.

3 THE SCHRÖDINGER EQUATION AS A CONDITION OF STABILITY

As early as 1929 N.G. Chetaev, a well-known expert in the theory of stability, worked in the University of Göttingen where he must have become familiar with Schrödinger's work [2]. In this work the equation describing the quantum levels of energy of the hydrogen atom was presented for the first time. In 1931, when Chetaev returned to the USSR, he published an article [3] in which he attempted to derive the time-independent Schrödinger equation from the viewpoint of the theory of stability. Chetaev assumed that Schrödinger equation extracted from solutions of the Hamilton-Jacobi equation only those that satisfy the condition of stability. However, Chetaev was unable to obtain the Schrödinger equation exactly because he did not introduce "hidden variables" like a physical field into the Hamilton-Jacobi equation [as mentioned above, the causal interpretation assumes existence of "hidden variables"]. Besides that, Chetaev, like Bohm, erroneously believed that U_Q represents the potential energy of small disturbing forces, while as it was demonstrated above U_Q tends to infinity as $\hbar \to 0$.

Here we present the derivation of the Schrödinger equation using Chetaev's approach, but correcting his mistakes. Unlike Chetaev, we introduce an unknown function Φ in the Hamilton-Jacobi equation [4]. The Hamilton-Jacobi equation with a given potential energy $U$and additional term Φ has a form:

$$\frac{(\nabla S)^2}{2m} + U + \Phi = \varepsilon \tag{13}$$

Consider now the motion of a particle (or the center of mass of an extended object) that it would have if small disturbing forces with potential energy W are present. Eq. (13) in this case takes the form:

$$\frac{(\nabla S)^2}{2m} + U + \Phi + W = \varepsilon \tag{14}$$

Of all possible motions of the particle, we will consider only those for which the arbitrary constants of the complete integral of the Hamilton-Jacobi equation (integral Jacobi) are assigned fixed values, and we will call the collection of these motions a *packet*. Following Chetaev's method, we assume that the influence of the perturbing forces on a packet at an arbitrary point is proportional to the density of trajectories A^2 at that point. For the packet consisting

of stable trajectories this influence must be minimal. That is, the action of disturbing forces is relatively less for the packets for which

$$\int WA^2 d\tau = \int W\psi\psi^* d\tau \Rightarrow \min . \tag{15}$$

here ψ is determined by the substitution

$$\psi(\vec{r}) = A(\vec{r}) \exp(iS/\hbar) . \tag{16}$$

The integrals are taken over the entire volume of configuration space.

Note, that substitution (16) extracts from all the trajectories of the Hamilton-Jacobi equation (13) only those for which the orbital angular momentum is quantized. Indeed, consider a particle moving along any close path $r = r(\varphi)$, $0 < \varphi \leqslant 2\pi$. The condition for $\psi(\vec{r})$ to be a single-valued can be expressed in the form

$$\frac{1}{\hbar}\oint dS = 2\pi k, \quad k = 0, \pm 1, \pm 2 , \tag{17}$$

where the integral is taken along this trajectory. Taken into account $dS = \frac{\partial S}{\partial r} dr + \frac{\partial S}{\partial \varphi} d\varphi$, condition (17) can be represented in an equivalent form

$$\oint \frac{\partial S}{\partial \varphi} d\varphi = \oint \frac{1}{r} \frac{\partial S}{\partial \varphi} r d\varphi = m \oint V_\varphi r d\varphi = 2\pi k\hbar . \tag{18}$$

In a special case of the circular orbits with radius r, Eq. (18) takes the form: $mVr = k\hbar$.

Substituting the expression for W from Eq. (14) into Eq. (15), we obtain the following variational problem:

$$\delta \int F d\tau \equiv \delta \int \left(\frac{(\nabla S)^2}{2m} + U + \Phi - \varepsilon \right) A^2 d\tau = 0 \tag{19}$$

Functions $S(\bar{r})$ and $A(r)$ realize extreme of the definite integral (18), if they satisfy the system of two Euler–Lagrange equations:

$$\frac{\partial}{\partial x}\left(\frac{\partial F}{\partial S'_x}\right) + \frac{\partial}{\partial y}\left(\frac{\partial F}{\partial S'_y}\right) + \frac{\partial}{\partial z}\left(\frac{\partial F}{\partial S'_z}\right) - \frac{\partial F}{\partial S} = 0 \tag{20}$$

$$\frac{\partial}{\partial x}\left(\frac{\partial F}{\partial A'_x}\right) + \frac{\partial}{\partial y}\left(\frac{\partial F}{\partial A'_y}\right) + \frac{\partial}{\partial z}\left(\frac{\partial A}{\partial S'_z}\right) - \frac{\partial F}{\partial A} = 0 \tag{21}$$

If on the trajectories that are solutions of the variational problem (18) the derivatives of Φ with respect to S, S'_x, S'_y, S'_z equal to zero, then condition (20) can be represented in the form

$$2(\nabla A)(\nabla S) + A\Delta S = 0. \tag{22}$$

Equation (22) coincides exactly with Eq. (4) for the stationary case.

Let us assume that function Φ depends on A and is independent of derivatives A'_x, A'_y, A'_z (at least on the trajectories that are solutions of the variational problem), then condition (21) takes the form

$$\frac{(\nabla S)^2}{2m} + U + \Phi_0 + \frac{A}{2}\frac{\partial \Phi}{\partial A} - \varepsilon = 0 \tag{23}$$

In order for Eq. (23) to coincide exactly with Eq. (11) (or with Eq. (3) for the stationary case), and at the same time to follow from Eq. (13) the conditions

$$\frac{\partial \Phi}{\partial A} = 0 \qquad \text{and} \qquad \Phi_0 = U_Q = -\frac{\hbar^2}{2m}\frac{\nabla^2 A}{A} \tag{24}$$

must be true.

It is obvious that the system of equations (21), (22,) (23), and (24) along with substitution (16) are equivalent to the time-independent Schrödinger equation. To prove this statement, one should perform derivations with the separation of the real and imaginary part of the Schrödinger equation, with the manipulation steps done in opposite order.

Note that among these solutions (even within the same packet) there can be "extra" solutions that despite satisfying the necessary condition are unstable, and, therefore, are not realized in nature.

In general, there are two approaches to the study of stability: one is to study the stability of a non-perturbed motion with the addition of small perturbation forces, while the other is to study the stability of a motion under variation of the initial coordinates and momenta with the same primary forces. The first approach was used in the above derivation. The second, which is the most widely used one, is based on the Poincaré variational equations. The second approach also leads to the same result, i.e. the Schrödinger equation is a necessary condition for stability. Under this approach, however, condition (30) splits into two conditions:

$$\nabla^2 S = 0 \qquad \text{and} \qquad (\nabla A)(\nabla S) = 0 \tag{25}$$

It follows from the second condition (25) that the trajectories that satisfy the necessary condition of stability belong to the surface $A(\vec{r}) = \text{const}$.

4 ELECTRON IN THE FIELD OF AN ATOMIC NUCLEUS

Now, using the causal approach, let us analyse the Schrödinger equation for a hydrogen atom

$$\nabla^2\psi + \frac{2m_e}{\hbar^2}(\varepsilon + \frac{e^2}{r})\psi = 0. \tag{26}$$

The solutions of this equation can be written in the form (7). [Note here, that these solutions can be obtained strictly mathematically without introducing any operators from quantum mechanics].

The phase of the wave function (16) can be written as $S/\hbar = k\varphi$. As known, ∇S in spherical coordinates has the form:

$$\nabla S = \frac{\partial S}{\partial r}\bar{i}_r + \frac{1}{r}\frac{\partial S}{\partial \vartheta}\bar{i}_\vartheta + \frac{1}{r\sin\vartheta}\frac{\partial S}{\partial \varphi}\bar{i}_\varphi \tag{27}$$

The component of the velocity along the radius $\bar{i}_r$ is zero, because $\partial S/\partial r = 0$; the component of the velocity along $\bar{i}_\vartheta$ is also equal to zero because $\partial S/\partial \vartheta = 0$. The only nonzero component of the velocity is the component along $\bar{i}_\varphi$. Thus, for the velocity of the electron on the trajectories obtained from the Schrödinger equation we have

$$\bar{V} = \frac{1}{m_e}\nabla S = \frac{\hbar k}{m_e r\sin\vartheta}\bar{i}_\varphi\,. \tag{28}$$

It follows from Eq. (28) that only two situations are possible: either the centre of mass of the electron is at rest ($k = 0$) or it is moving along a circular orbit lying in a plane parallel to the xy plane, with the centre of the orbit located on the z axis (the above reasoning can be applied to any coordinate axis). Note, that the motion on these circular trajectories satisfies only the necessary condition of stability, therefore, among these trajectories there can be trajectories that are not stable.

4.1 BOHR ORBITS

Let us consider the circular orbits with the center at the origin. In this case

$$\bar{V} = \frac{\nabla S}{m_e} = \frac{\hbar k}{m_e r}\bar{i}_\varphi\,. \tag{29}$$

According to Newton's second law

$$\frac{m_e V^2}{r} = \frac{\hbar^2\, k^2}{m_e r^3} = F. \tag{30}$$

Therefore, besides the known Coulomb' force acting on the center of mass of the electron in a hydrogen atom there is an additional force

$$F_a = F - \frac{e^2}{r^2} = \frac{e^2}{r^3}\left[\frac{\hbar^2\, k^2}{m_e e^2} - r\right]. \tag{31}$$

It follows from Eq. (31) that $F_a = 0$ on the Bohr orbits, with radius

$$r_{B_k} = \hbar^2 k^2/m\, e^2\,. \tag{32}$$

Therefore, the Bohr orbits satisfies the necessary condition of stability. Indeed, in the hydrogen atom, the trajectories which satisfy the necessary condition of stability belong to surfaces $r = \text{const}$ and, therefore,

$$\frac{\partial \Phi}{\partial A} = 0 \;\rightarrow\; \frac{\partial \Phi}{\partial r} = 0.$$

4.2 THE PRECESSION OF THE ELECTRON'S SPIN IN AN ATOM

It is known that the motion integral for the center of mass of the electron has a form:

$$\frac{1}{2} m_e V_k^2 - \frac{e^2}{r_{B_k}} = \varepsilon_k\,, \tag{33}$$

where

$$\varepsilon_n = -\frac{e^4 m_e}{2\hbar^2 n^2} \qquad n = 1, 2, 3, \ldots\,. \tag{34}$$

On the other hand, it follows from (23) and (24) that the motion integral for the Bohr's orbit is

$$\frac{m_e {V_k}^2}{2} - \frac{e^2}{r_{B_k}} + U_Q = \varepsilon_n\,. \tag{35}$$

Comparing the expressions (33) and (35) we obtain

$$U_Q = \varepsilon_n - \varepsilon_k\,. \tag{36}$$

Thus, along with the term that characterizes the energy of the center of mass in Eq. (35) there is another term U_Q, which is determined by $\varepsilon_n - \varepsilon_k$. This

brings up a question: what are the physical conditions that make motion integral (35) possible? If the motion of an object is represented as a superposition of two motions: the motion of its center of mass and the motion about the center of mass, then term U_Q is the energy associated with the motion about the center of mass. That is, it is possible when we have the generalized motion integral.

Suppose that the motion about the center of mass is the precessional motion of the electron's spin. We can describe the spin precession as precession of a classical gyroscope because there is a preferable direction in this problem (the normal to a plain of an orbit). It follows from the fast-gyroscopic theory that for the precessional motion of a gyroscope which is limited to small precessional angles, the energy of the motion E is numerically equals the product of the intrinsic angular momentum and the angular frequency of the precession ω_e,

$$E = \frac{\hbar\omega_e}{2}. \tag{37}$$

For such motion

$$E = U_Q. \tag{38}$$

Thus, on Bohr' orbit potential U_Q ("the quantum potential") equals the energy of the precessional motion of the electron's spin. It can be seen from Eq. (36) that the formula for the energy of the precessional motion is, in fact, Rydberg's formula.

Note: in the case of precessional motion the energy depends on the frequency of precession and not on the square of the frequency, as in the case of wave motions.

Hypothesis. Up to now we did not introduce any hypothesis about the nature of the physical field Φ (or the "hidden variable"). If we make an assumption that this field reflects a motion of some material substance which is in various aspects like superfluid He-3 [5], then we can provide an explanation of the natural frequencies of an atom. Indeed, structures like homogeneous precessing domains are observed in superfluid He-3, where all spins of the fluid particles precess with the same frequency and phase. In the presumed superfluid vacuum, the electron's motion should create structures (quasi-particles) similar to such domains. Quasi-particles in vacuum have spin $\hbar$. If we assume existence of the dipole-dipole interaction between an electron

with spin $\hbar/2$ and the quasi-particle with spin $\hbar$, then the following equality holds:

$$|\varepsilon_n - \varepsilon_k| = \frac{\hbar\omega_e}{2} = \omega_p \cdot \hbar, \tag{39}$$

where ω_p is the angular frequency of the precession of the quasi-particle's spin. Formula (nseq:39), in essence, is Rydberg's formula, which describes the natural frequencies ω_p of the hydrogen atom.

ACKNOWLEDGMENTS

I would like to thank Nadia Lvov for her help in preparation of the article.

5 BIBLIOGRAPHY

[1] D.Bohm , *Phys. Rev.*, **85 (2)**, 166 (1952).

[2] E. Schrödinger, *Annalen der Physik* ,79, 361(1926).

[3] N. G. Chetaev, Stability of Motion. Proceedings in Analytical Mechanics (USSR Academy of Science Publishing, Moscow, 1962) [in Russian].

[4] N. Sotina, *Physics Essays*, **27**, 321 (2014).

[5] N. Sotina, Reports of the 16th RCCNT&BL, Dagomys, June 2009, Moscow, 172 (2010)

COHERENCE DOMAINS IN WATER. CHANGE OF WATER PROPERTIES IN LIVING ORGANISMS.

NINA SOTINA[1] & NADEZHDA LVOV[2]

[1] New York, USA
nsotina@gmail.com

[2] Department of Physics, Essex County College,
Newark, NJ, USA

Abstract: There is evidence for quantum coherence in water that is supported not only experimentally but also theoretically. In the current work, a unique experiment by H. Hu and M.Wu that demonstrates quantum properties of water is presented and analyzed. They observed that the pH value of water and its temperature in one reservoir can be non-locally affected through manipulating water in another remote reservoir (at first all water was in the same closed reservoir). Another equally interesting series of experiments, performed by E. Macovschi, studied the change of the quantum properties of water inside living tissue.

1 QUANTUM COHERENCE OF WATER

In the beginning of 1990s Warren and collaborators detected a high degree quantum coherence in liquids using nuclear magnetic resonance (NMR) [1]. They examined the effects of spin concentration, temperature, chemical exchange, and other variables on the formation of coherences in simple liquids. Their experimental result was unexpected from the standpoint of the conventional NMR theory.

About the same time G.Preparata and Del Giudice [2, 3] showed theoretically on the basis of the theory of quantum electrodynamics (QED) that *water* at room temperature and normal pressure organizes itself in *Coherence Domains.* In such a domain, all molecules are in a coherent state described by the same wave function. The diameter of the coherence domain (which has a spherical shape) is about 100 nm. According to Preparata, the coherence domains have the following properties: within a domain, the quantum energy

level is lower compared to the energy level of the rest non-coherent water, which provides stability of the domain to the enegy fluctuations.

A series of unique experiments were conducted in Stony Brook by H. Hu and M. Wu (2007) in which they observed the quantum properties of water [4]. They demonstrated that water has the ability to retain a connection between its parts after it was separated. In the experiment, a water reservoir (made of pastic) was either exposed to microwave radiation in a 1500W microwave for 2 min or audio-frequency radiations of a 20W magnetic coil for 30 min. Then, the water was divided into two smaller reservoirs. H.Hu and M.Wu found that if the samples of water were separated, the water's pH value, its temperature and even the water's gravity in one reservoir can be non-locally affected through manipulating water in the other remote reservoir.

If we accept the existance of the coherence domains in water, then the explanation of H.Hu and M.Wu's experimental results can be as follows: under certain conditions the coherence appears not just within each domain but among the domains, thus the system transits to the next level of coherence. Therefore, we have an ensemble of domains described by the same wave function and all the properties of quantum nonlocality are applicable to the domains within the ensemble. The microwave or audio-frequency irradiation of the initial water sample in Hu and Wu experiments could initiate the coherence ensembles of domains.

2 CHANGE OF WATER PROPERTIES IN LIVING ORGANISMS.

The existance of coherent excitatitions in biological systems was first proposed by Herbert Fröhlich [5] . According to Fröhlich model, when the energy supply goes above a certain level, the coherent vibrations emerge in living systems that are similar to the coherent behaviours found in superconductors.

Many publications appeared recently that discuss the demonstrations of the interfacial water properties performed by Gerald Pollock and his team [6, 7]. In these demonstrations an interfacial water layer of several hundreds of micrometers thick forms on the surface of a hydrophilic gel. This layer excludes dye or any other solutes, it was named an *exclusion zone* (EZ). According to Del Giudice and colleagues [8, 9], inside the cell EZ water is a coherence domain that is formed on the surfaces of the membranes. The rest of the water in the living cell is characterized by a high degree of coherence among the coherent domains.

The ability of a living system to change the properties of intracellular water was demonstrated in the experiments by Romanian biochemist Eugene Macovschi (1906-1985) and his group [10, 11] In the experiment, a living plant tissue was exposed to 200 ATMs hydrostatic pressure. Under this pressure it released some amount of water and remained alive. In the case of a dead tissue, after being exposed to 200 ATMs it released all the water it contained. Once the biological water was released from the biological systems, both living and dead, it immediately regained the properties of ordinary water.

It is important to note that the water was taken from the same plant tissue under the same boundary conditions, which makes this experiment unique. According to contemporary biochemistry, living and dead matter consists of the same substances and are not qualitatively different. However, Machovski's experiment showed that two forms of matter, living and dead, are fundamentally different. If we accept the existence of coherence domains in water, then it follows from these experiments that **a living organism changes the quantum properties of water.**

However, modern physics does not suggest the mechanism with which a living organism can affect water at the quantum level. It is time to step away from the view that a living organism is just "a very complicated machine". There is a fundamental difference between the animate and the inanimate. It was observed, for example, that there exists a fundamental difference between a chemical molecule and a biologically active molecule, even though these molecules have the same chemical structure.

3 STRUCTURES IN PHYSICAL VACUUM

Macovschi gave the following explanation of his experimental results. He suggested that living cells consist of two qualitatively different, though interacting, forms of matter: a special form of structured matter ("biostructure") and coexisting molecular matter, chemicals. He guessed that the biostructure is an entity carrying the control function in living cells, and that this control is irreducible to electric and electromagnetic interactions.

The value of the Makovski hypothesis, even though we do not completely agree with it, is in his attempt to go beyond the molecular level and to consider non-molecular structures in vacuum.

Note that Macovschi's ideas of biostructures stand out from all other explanations of the living organism behavior. At the same time some scientists

try to explain mysterious phenomena in biology by introducing a new kind of physical field *(the biofield*), which is presented in living organisms only. Note that the term "the biological field" was introduced by A. Gurvitsch at the beginning of the 20^{th} century. He suggested the existence of non-chemical level of organization that controls embryogenesis.

By the term "physical field" Gurvich and his followers understood the classical definition of a field, the agent that transmits action from one object to another. A field in most case has a property of continuity. A spatial structure, on the other hand, is composed of elements (items) of arbitrary nature and has a definite shape (Macovschi viewed the biostructure as a sponge). Here lies a significant difference between structures and the physical field. (A simple example of a structure is a vortex in fluid).

It is important to note that in order for a non-molecular structure to be stable and at the same time to be dynamic there must exist some kind of interaction between remote elements of the structure. Therefore, in the case of the Macovski's biostructure there must exist **interaction between the non-molecular elements of the biostructure**. Thus, to understand biostructures it is important to go beyond the molecular level in the investigations, and to study the properties of the physical vacuum.

The existence of structures in the physical vacuum was proven mathematically by one of the authors (N. Sotina) [12, 13]. Using the casual approach to quantum mechanics, it was shown that any quantum objects, i.e. atoms, molecules, biomolecules form structures in the physical vacuum. In the case of an atom, it was proven that these structures are responsible for stabilization of the election's motion in an atom.

4 BIBLIOGRAPHY

[1] Zhang H, Lizitsa N, Bryant RG and Warren WS. Experimental characterization of intermolecular multiple-quantum coherence pumping efficiency in solution NMR. *J Magnetic Resonance,* **148**, 200-208. (2001).

[2] Preparata G, QED coherence in condensed matter, World Scientific Publishing, Singapore (1995)

[3] Arani R, Bono I, Del Giudice E, Preparata G, QED coherence and the thermodynamics of water, *International Journal Physics B 09*, 1813, (1995)

[4] Hu, H. & Wu, M. Evidence of Non-local Chemical, Thermal and Gravitational Effects. *Progress in Physics,* vol.. **2**, 17-24. (2007)

[5] Frohlich, Long Range Coherence and Energy Storage in Biological Systems, *Int. J. Quantum Chem.*, **v.II**, 641–649 (1968)

[6] Pollack GH and Clegg J. Unexpected linkage between unstirred layers, exclusion zones, and water. In *Phase Transitions in Cell Biology* (Pollack GH, Chin WC eds.), pp143-52, Springer Science & Business Media, Berlin, (2008).

[7] Pollack, GH: Water and Surfaces: A Linkage Unexpectedly Profound. In: Hydrogels: Biological Properties and Applications. Springer-Verlat, Milan, Ed: R. Barbucci, pp 145 – 147 (2009).

[8] Del Giudice E, Spinetti PR and Tedeschi A. Water dynamics at the root of metamorphosis in living organisms. *Water*, **2**, 566-86.(2010)

[9] Ho MW, Quantum coherent water and life, *Science in Society* **51**, (2011).

[10] E.Macovschi , *Rev. Roum. Biochim.*, **19**, 177-186 (1982)

[11] G. Drochioiu, "Eugene Macovschi's concept of biosructure and its current development", *LIVE and MIND, insearch of the physical basis,* Trafford publishing, Canada, Editor S.Savva, pp. 43-60 (2006).

[12] N. Sotina, *Physics Essays*, **27**, 321 (2014)

[13] N. Sotina, The causal (deterministic) approach to quantum mechanics, in J. Amson (ed.) *Chiasms, Proceedings of ANPA Nos. 37 & 38*, 2018

ENCOUNTERING BIRDS: A PHENOMENOLOGY OF JIZZ.

STEPHEN WOOD
s.w.wood.88@cantab.net

I describe how my encounters with birds illustrate the ornithological notion of 'jizz,' the characteristic way a bird appears to the watcher. I argue that jizz captures the notions of the 'inscape' and 'instress' proposed by the 19th century poet Gerard Manley Hopkins. In turn, Hopkins' terms provide the basis for the triad of identity, as proposed by philosopher and polymath John Bennett.

Encounters

Let me introduce three different birds which I have been lucky enough to encounter in the field. The first is a kingfisher (cf. Figure 1, top) that I saw on the river Avon in Bathampton flying away across the way the surface of water, as a flash of iridescent blue sapphire.[1] The second is the hoopoe (cf. Figure 2), which I saw on a trip into the Luberon. My wife and I were trying to drive through the town of Apt to return home and as Fate would have it, we were forced outside the town and into the woods. We saw this lovely bird flying over the road, with its black and white rounded wings and orange head. Straight away I was reminded of the description I had read as a boy, "a huge black-and-white butterfly or moth" (Fitter & Richardson, 1966, pp. 86-87).

[1] As I learned in a book from the local library when I was a child: "A flash of sapphire is all that is needed to identify the kingfisher as it streaks downstream or bellyflops to catch a minnow" (Fitter, 1974, p. 178).

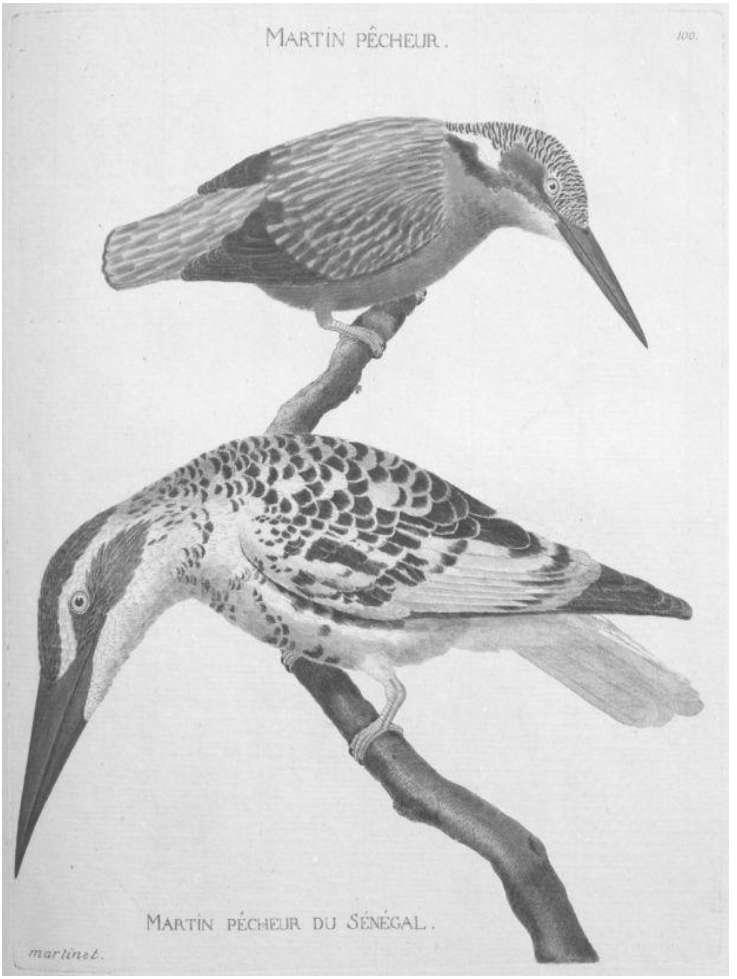


Figure 1. Eurasian and Senegalese Kingfishers, engraving by François-Nicolas Martinet.[2]

[2] See http://informations-documents.com/environnement/copermine15x/displayimage.php?pid=15816, last consulted 7th December, 2017. For some excellent documentary footage, see https://www.youtube.com/watch?v=OazPkQYZcBQ, last consulted 8th December, 2017.

Figure 2. Hoopoe, engraving by Martinet.[3]

In the figures, the 18th century French illustrator Martinet gives the birds a static pose.[4] The hoopoe is pictured like royalty, posed

[3] http://informations-documents.com/environnement/copper mine15x/displayimage.php?pid=15807, last consulted 7th December, 2017. Documentary footage: https://www.youtube.com/watch?v=dMSthccjLLM, last consulted 8th December, 2017.

[4] To be fair, Martinet did his best with the materials he had: "he incorporated a fuller artistic element in his illustrations when the opportunity arose, showing birds on the roofs of French chateaux or city buildings, swimming in ponds, perched on a crag overlooking ships sailing on the sea, in farmyards or at bird-feeders, and even - for those birds that

beautifully on a rock. Look at these illustrations and then consider how the bird appears in the field, the hoopoe as a huge butterfly with its slowly beating wings, and the kingfisher as a flash of sapphire.

We have two ways of looking at birds that are at odds with one another: firstly, how they appear to us in the field, often in an instant, and secondly, how they are presented to us, in a museum cabinet, in a bird book,[5] or in wildlife films. In the latter, we see them in close-up, for example, the kingfisher with its bill full of fish. They give us the expectation that we should experience birds in the field in this way, which is rarely the case.

My third encounter was with the lapwing (cf. Figure 3). I saw a group of lapwings flying over a rain-soaked field in Kent. They have this lovely tumbling flight where they fly up and then seem to drop and swing round.

The lapwing is also known as the peewit, after its note. It is possible to identify birds in the field by what we hear, not only by what we see. The name 'hoopoe' is an onomatopoeia.[6] Its note is 'oop-oop'. The name in French is of the same root, 'huppe.'

---

had become popular as pets - in cages in domestic interiors." Zaharek & Overstreet (2000).

[5] Fitter & Richardson (1966) is a rare exception, p.10 "The great majority of bird books arrange their birds according to the systematic scientific classification, but as this is apt to be confusing to bird-watchers in their early stages, a break has been made with tradition and the birds arranged in a way which it is hoped will prove more helpful to the beginner. They are first divided into three broad habitat groups (Land, Waterside, Water), and within each habitat group the birds are set out in ascending order of size in eight main groups based on the length of each bird."

[6] See etymologies in Lockwood (1984) for 'hoopoe' and for the earlier form, 'hoop': "The source is Old French *huppe* ultimately from Latin *upupa*, echoing the call note heard as *up-up*."

Figure 3. Lapwing, engraving by Martinet.[7]

This consideration of my encounters with birds has given us an introduction to phenomenology. Henri Bortoft contrasts

[7] http://informations-documents.com/environnement/copermine15x/displayimage.php?pid=15880, last consulted 7th December, 2017. Amateur footage: https://www.youtube.com/watch?v=l58X16MnWLc, last consulted 8th December, 2017.

phenomenology's way of revealing the world, disclosure, with that of conventional science, "describing as":

> "Describing as" implies some "what" is described which is pre-given independent of the description. "Describing as" takes the form: describing Z (the supposed pre-given) as "X". But we never meet with any Z. What we encounter is disclosed by "X" to be 'X'. The "what" emerges into clarity and definition in the act of describing (to be). X is disclosed—not designated, not created. A description says the world, which means shows the world to be thus: 'X' (Bortoft, 1970, p. 227).

In recounting my encounters with birds, I am describing how the birds disclosed themselves to me as themselves. I have tried to stay with the phenomenon of the bird itself, how it appears in its habitual context, rather than looking to explain the bird's appearance by referring to an abstraction 'behind' the phenomenon.

Bird books and wildlife films can give a misleading impression that the bird displays itself to the watcher in close-up, as if in a museum cabinet. Birds are most often glimpsed in an instant, as a characteristic integration of flight, note, length and colour of bill, plumage and legs. Bird books, like Martinet's paintings, so often display the 'finished product', the bird once it has appeared, and thus present an abstraction from the field of experience (Bortoft, 2012, p. 27).

Jizz

"The beginner may well be baffled at the way on old hand is able to pick out a bird at a distance and identify it with confidence before details of its plumage can be seen." (Fitter & Richardson, 1966, p. 26). The novice has to decompose what he sees in the brief instant of the bird's appearing; the old hand recognises the bird all at once. He is able to capture in a flash that unity of appearance.

What is this mysterious ability to recognise a bird in an instant? What qualities is it based on? Here we come on to jizz, which was first mentioned by T. A. Coward in his *Bird Haunts and Nature Memories* (1922). He gives a parable of an Irishman who knew his birds by their jizz:

> A West Coast Irishman was familiar with the wild creatures which dwelt on or visited his rocks and shores; at a glance he could name them, usually correctly, but if asked how he knew them would reply, 'By their 'jizz.' (Coward, 1922, p. 141)
> ...the outdoor naturalist, and in particular the field ornithologist, realises the full value of jizz better than most people. At a distance, too far away to see details of form, colour, or pattern ... he sees a bird and recognises it ... but how does he know? Shape, size, manner of flight, or maybe note, is the reply. Yes, but there is something more ... something which instantly registers identity in the brain ... It is its jizz." (pp. 141, 142)

We can see the link with a person's gait: how we recognize a person from a distance by some integration of how they carry themselves, their height, their posture, the way they walk. In French, there is an interesting word, 'allure.' For a long time, I would walk from the station to work and colleagues would pick me up on the way. I asked one colleague how she recognized me, and she said by my 'allure.' Allure,' in addition to describing appearance and a particular way of carrying oneself and moving, also conveys a quality that emanates from a person.[8] Jizz identification is best known for birds, but is part

[8] Coward (1922), p. 141: "If we are walking on the road and see, far ahead, someone whom we recognise although we can neither distinguish features nor particular clothes, we may be certain that we are not mistaken; there is something in the carriage, the walk, the general appearance which is familiar; it is, in fact, that individual's jizz."

of a more general human ability, a more general phenomenon of recognition.[9]

So where does the word jizz come from? Well, there are various theories, for the word's origin is a mystery. In the first theory, jizz derives from an acronym GISS, "general impression of size and shape," said to come from the military context of identifying aircraft. Evidence of the use of such a term in that context is lacking, however.[10] According to a second theory, jizz is what the bird "just is,"[11] what the person "just is" when you see them from a distance. A third theory proposes jizz as a corruption of gestalt, meaning

[9] Coward (1922), p. 143: "Jizz, of course, is not confined to birds. How do we recognise the bank vole, seen for a second in the lane, the long lean rat which appears and vanishes like a grey streak, the pipistrelle flitting in the dusk round the barn? How do we know the daisy in the field, the sturdy oak? Is it be colour, size, length of tail, or shape of wing, by petal, form of leaf, or fruit? No; the small mammal and the plant alike have their jizz. We do not stop to look for detail, or ask ourselves what we saw; we know. Jizz may deceive us; that is our fault, for each and every thing has its distinctive jizz; if inexperienced we may fail to discern it. To learn the jizz should be the object of every field naturalist; it can only be learnt by study of wild creatures in their natural surroundings."

[10] McDonald (1996), p. 7: "A book with the title *The Spotter's Handbook* (Chichester 1941) contains a Chapter on "Systematic Identification". It does not mention GISS nor does it apply the concept of the general impression of an aircraft. Rather it describes what was apparently the accepted aircraft identification technique, the "WEFT" approach in which observers were directed to analyse separately the Wings, Engine, Fuselage and Tail - the antithesis of the GISS approach. The official 1946 UK War Office training manual on this topic has a section on "Recognition by Appearance"...: "... to the sportsman or country dweller a bird is recognized by its general appearance and method of flight ... not by details of the exact shape of various parts of its body. Similarly, an aircraft is recognized by its general appearance and 'sit' in the air, not by precise constructional details." (Anon 1946, p. 3)."

[11] McDonald (1996), p. 3: "Litwin [on *BirdChat*, 1994] went straight to the point, saying that: "I thought 'jizz' was a contraction of 'just is'.""

wholeness of form, said to be pronounced jestalt.[12] There is no evidence to support such a pronunciation, leaving this theory rather far-fetched. A fourth theory is more promising and suggests that jizz arose as a mishearing of jism. Jism is used for the energy of a living being, an energy that emanates from it, that comes across to the observer.[13] All these etymologies are very good descriptions of what jizz means. Notice the complex of meanings, relating to form, energy and impression. These three aspects will prove to be important.

A contemporary example of the usage of jizz comes from Sean Dooley's book *Big Twitch* (2005):

> Jizz is the indefinable quality of a particular species, the 'vibe' it gives off ... the way a bird holds itself... (p. 78)
> The concept of jizz relies on an innate sense of how a bird is ... (p. 79)

Dooley, an Australian bird-watcher, set himself the task of recording all the bird species on the Australian subcontinent and neighbouring islands in one year. He saw more than 700.

---

[12] McDonald (1996), p. 3: "'A distinctive physical "attitude", totally apart from any specific field mark ... The origin of the term is uncertain perhaps a corruption of "gestalt".' [Danca, Principe on *BirdChat*, 1994]."

[13] McDonald (1996), p. 8: "An intriguing possibility exists that Coward may have made a mistake from the beginning. The OED and other dictionaries such as *A Dictionary of Slang and Unconventional English* (Partridge 1984) contain the English language word "jism" or "gism", also rendered in other ways including "jizz"... Draper [1985] wrote "[The word jizz] was familiar ... to friends from Dublin and Kildare and it was generally used to imply spirit or cockiness. Schoolboys were encouraged to 'show a bit of jizz' and a sergeant-major regularly promised to knock the jizz or jizzom (jissom) out of a squad of recruits ... One can readily visualise the Irishman telling Coward that a particular bird species was recognisable from its jizz, meaning specifically its cockiness or energetic characteristics, rather than meaning its characteristic impression as such..."

Dooley speaks of the energetic quality of jizz, the vibe the bird gives off, the indefinable quality of a particular species. He also points out that jizz demands a quality in the birdwatcher, to have a sense of what the bird is. Think of what he managed to do in that year, to attune himself to 700 different kinds of being.

How was it verified that he saw so many species?, you may ask. He couldn't take photographs, as most of the time he saw the birds in an instant. As a good twitcher, you have to be in the right place at the right time. You have to read the weather, read the land, you have to understand how variations in the seasons and the weather alter the distribution of birds. ''Well, spring is late, we've had storms this year. They're not going to be there at that time. I'll have to go there later." There would also have been a community of birdwatchers who would approve what Dooley saw, given the conditions, the time of year.[14] He had a good reputation in the field and fellow birdwatchers were willing to believe him and applaud his achievements.

To illustrate the concept jizz further, Dooley gives an example of two birds he had to recognize at sea, the Sooty and the Short-tailed Shearwaters (p. 79):

> The Sooty has silvery underwing panels, but they are difficult to pick out in bright sunlight. The Short-Tailed, yes, has a shorter tail, but not by much. And if the Sooty you see happens to be moulting... Again it's bill is shorter, by less than an inch. Given that you'll be looking at these birds from a boat lurching around on a choppy sea, it's not easy to spot these differences.
>
> The wingspan on a Sooty, however, is about ten centimetres longer than on a Short-tailed. Again, this measurement can

[14] As he says elsewhere: "Finding rarities is like goldmining. [A good sense of jizz allows one to] read the country, assess the lie of the land and search accordingly" (Dooley, 2005, p. 114).

> be difficult to detect at sea, but the longer wings do give the Sooty a different jizz as it flies. That's an extra ten centimetres (around ten per cent) of wing surface with which to fly so the Sooty doesn't have to work quite as hard to travel at the same speed as a Short-tailed. This means that Short-tailed Shearwater tends to flap its wings more often, whereas the Sooty has a more languid jizz in flight, something that you can pick up with practice almost every time ...

The difference in wingspan gives the Short-tailed a more languid flight, as it flies across the surface of the sea. (The name "shearwater" comes from the way they fly very close to the surface of the sea.) That is its jizz, an integration of all the different characters in its particular way of behaving that, with training, an observer can spot from a distance in an instant.

Hopkins

The poet and nature watcher Gerard Manley Hopkins has fascinated me for a long time. When I went up to university in Cambridge to study natural sciences, my best friend in the first year wasn't a scientist but a student of English. For his third year dissertation, he was studying with the Hopkins scholar, Catherine Phillips. It is thanks to him that I first came across Gerard Manley Hopkins and his beautiful journals and nature poetry.

Here are a couple of extracts from Hopkins's journals. First, let's go back to the lapwing, to the peewit, that we saw earlier.

> Peewits wheeling and tumbling, just as they are said to do, as if with a broken wing. They pronounce peewit pretty distinctly, sometimes querulously, with a slight metallic tone like a bat's cry. Their wings are not pointed, to the eye, when

> flying, but broad, white and of a black or reddish purple apparently (J, 24).

Hopkins provides a very precise description. He has a great ability to capture the detail he sees. He was a Jesuit and as such he would have followed the Spiritual Exercises of St. Ignatius, where the aim is to practise discernment, discernment of the heart, of one's emotions, of one's attitude to others and to what's happening in the world. I feel there must be a connection, that his Jesuit training brought out a very refined sensibility and heightened his ability to discern in nature what he saw, glimpsing the beauty in the details.

Hopkins writes about plants, too, particularly flowers and trees. Here he's going over an earlier description of a bluebell:

> I do not think I have ever seen anything more beautiful than the bluebell I have been looking at. I know the beauty of our Lord by it... The head is strongly drawn over backwards and arched down like a cutwater drawing itself back from one line of the keel. The lines of the bells strike and overlie this, rayed but not symmetrically, some lie parallel ... Then there is the straightness of the trumpets in the bells softened by the slight entasis and by the square splay of the mouth. (J, 199)

Hopkins's first ambition was to be a painter. In his youth he was an excellent draughtsman[15]. You can see that here, in how he details the lines, the planes and the shapes. He is decomposing the bluebell, almost drawing it with words. The nautical metaphors, keel and cutwater, only serve to emphasize the technical precision of the description.

---

[15] Gardner in Hopkins (1953), p. xvi, "He himself was an excellent draughtsman, and his later skill, as a poet, in communicating through words the essence and individuality of visual forms in nature was partly fostered by his early training with the pencil."

I chose these two journal entries to show how Hopkins' love of particularity shines through his detailed and sensitive observations. The first entry describes a flock of birds, the second a single flower. Guided by St. Ignatius' call to "see God in all things," Hopkins feels the moral imperative to be faithful to nature in his descriptions. At the same time, we feel his delight in capturing phenomena as they appear.

Next I would like to discuss a poem (P, 57), which is a veritable hymn to particularity, and which takes us back to the kingfisher we met at the beginning of this essay:

> As kingfishers catch fire, dragonflies draw flame;
> As tumbled over rim in roundy wells
> Stones ring; like each tucked string tells, each hung bell's
> Bow swung finds tongue to fling out broad its name;
>
> Each mortal thing does one thing and the same:
> Deals out that being indoors each one dwells;
> Selves—goes itself; *myself* it speaks and spells
> Crying '*What I do is me: and for that I came*.'

With his image of kingfishers 'catching fire', Hopkins evokes the image of the flash of sapphire which proved so appropriate in my own experience. A beautiful image follows which rhymes with the first, of dragonflies drawing the sun's flame to themselves as they hover over the water. Flashes and flame recall another of Hopkins' poems, *God's Grandeur*, which proclaims "The world is charged with the grandeur of God./ It will flame out, like shining from shook foil" (P, 31). Kingfishers and dragonflies are charged with the grandeur of God, indeed "All things therefore are charged with love, are charged with God and if we knew how to touch them give off sparks and take fire, yield drops and flow, ring and tell of him" (S, 195). For Hopkins, each

creature, instituted in this created world by God's love and by his divine will, is filled with his love and virtue which shine forth in glory.

In the second line we move to stones, picturing them tumbled into wells and listening to them ring as they fall into the water. The plucked string, the swinging bell, each has its distinctive note, its distinctive melody, that 'flings out broad its name'. Flashing, ringing, each thing proclaims its particular way of being. The stress falls decisively in the fifth line 'Each mortal thing does **one** thing and the same'. The rhymes multiply, reverberating "wells ... tells ... bells ... dwells ... selves ... spells." Selves, a verb of Hopkins' own coining stands out from the resonance, calling attention to the message of the first part of the sonnet, "*myself* it speaks and spells/ Crying *'What I do is me: and for that I came*.'"

The particularity of the creature is in its activity, its selving, its own way of being. Its being is expressive, shining and ringing out for all to see, if only one has "eyes to see and ears to hear" (Matthew 13: 15-17). A bird, through its characteristic activity or jizz, is itself proclaiming its identity and destiny, its place in the world.

In Hopkins' nature poems we see his devotion to the Spiritual Exercises of St Ignatius. Their 'Principle and Foundation' begins thus:

> Man was created to praise, reverence and serve God Our Lord, and by so doing to save his soul. All other things on the face of the earth are created for man to help him fulfill the end for which he is created (Loyola, 1964, p. 47).

Each creature is a sacrament, a sign of divine presence, serving to lead the attentive observer to 'praise, reverence and serve God.' This message is reinforced in the second part of the sonnet, where the just

man is sign and presence of Christ, 'for Christ plays in ten thousand places,/ Lovely in limbs, and lovely in eyes not his.'

Inscape

In the first account of the bluebell cited earlier, Hopkins writes "It is strength and grace, like an ash." He amends it later to "Its inscape is mixed of strength and grace, like an ash tree." The particular quality of this individual bluebell Hopkins describes, in the revised version, as its inscape.

This term, and another invention of Hopkins that we will come to later, instress, have much intrigued scholars. William Gardner, in the first study of Hopkins' poetry (Gardner, 1944), took the inscape of an individual to be its *haecceitas*, or 'thisness', the term the mediaeval theologian Duns Scotus employed for a creature's individuating principle. Certainly in this example, Hopkins speaks of the inscape of a single flower. The following passage from Hopkins' journal, written to mark his discovery of Scotus' *Ordinatio*, also appears to support the idea: "just then when I took in any inscape of the sky or sea I thought of Scotus." (J, 221)

Although Gardner's interpretation of inscape was followed for many years (for example, Peters, 1948; Ong, 1986), certain scholars contest it (Miller, 1963; Waterman Ward, 2002; Sobolev, 2011). For Waterman Ward (2002), Scotus' *haecceitas* is the extreme of the particularity of the individual which is known only to God and thus not a matter to be comprehended or expressed in this life by the human poet (pp. 159-60). She expresses doubt that Hopkins would have read what Scotus wrote on *haecceitas* and points to the single use of the term in Hopkins' journals. In writing of 'any inscape of sea or sky,' Hopkins tells us, she says, that the inscapes of sea and sky are multiple and transient, rather than singular and fixed (pp. 183-85).

Arguing against the thesis "inscape = *haecceitas*" is the very variety of contexts in which Hopkins uses the term. What is more, he does not simply use inscape as a noun, but also as a verb.

To take one example, Hopkins writes of the inscape of particular kinds of trees and of the trees of a particular region:

> Swiss trees are, like English, well inscaped—in quains [wedge shapes] (J, 170, note 392).
>
> Sycomores [sic] grew on the slopes of the valley, scantily leaved, sharply quained and accidented by perhaps the valley winds, and often most gracefully inscaped (J, 176).
>
> Spanish chestnuts: their inscape here bold, jutty, somewhat oak-like, attractive, the branching visible and the leaved peaks spotted so as to make crests of eyes (J, 179).

The inscape of the trees expresses their character ('bold', 'attractive', 'graceful'), their principles of construction ('quained', 'accidented', 'jutty', 'branching') and their relation to their surroundings ('Swiss', 'English', 'slopes of the valley', 'valley winds'). The use of inscape as a verb in this context points to how the trees themselves embody these qualities of form ('well inscaped', 'gracefully inscaped').

We can see Hopkins the phenomenologist in these entries. He describes qualities of the trees he has glimpsed at a particular time and place, qualities that are influenced by their trees' context, the lie of the land, the time of the year, and the local climate.

This use of inscape for many examples of a single kind of tree ('sycomores', 'Spanish chestnuts') might lead us to believe that inscape refers to the natural object's nature or *quidditas*, 'thatness' in Scholastic terms (Ball, 1971, pp. 107, 109). At one point, Hopkins

himself describes inscape as nature, but contrasts it with self (S, 146). When Hopkins writes that 'kingfishers catch fire and dragonflies draw flame,' he appears to be speaking of their special natures, but goes on to celebrate their particular selves. Certainly the "inscape = nature" thesis is difficult to reconcile with the ascription of inscape to a single individual. Though puzzling, we have come across this double meaning before. We have spoken of the jizz of a single human individual as his or her gait or *allure* and at the same time of the jizz of a particular species of bird.[16]

When Hopkins is observing the unfurling or withering of a flower, his concept of inscape changes to encompass the whole process[17]:

> A beautiful instance of inscape sided on the slide, that is/ successive sidings of one inscape, is seen in the behaviour of the flag flower from the shut bud to the full blowing: each term you can distinguish is beautiful in itself and of course if the whole 'behaviour' were gathered up and so stalled it would have the beauty of all the higher degree (J, 211).

> The Horned Violet is a pretty thing, gracefully lashed. Even in withering the flower ran through the beautiful inscapes by the screwing up of the petals into straight little barrels or tubes. It is not that inscape does not govern the behaviour of things in slack and decay as one can see even in the pining of the skin in the old and even in a skeleton ... there was nothing in itself to shew even whether the flower were shutting or opening (J, 211).

[16] Merleau-Ponty when writing of the being-in-the-world of the animal, refers equally to 'an *a priori* of the species' (1962, p. 78) and to 'an *a priori* of the organism' (1964, p. 4).
[17] Compare the discussion in Ball (1971), p. 108.

The flower runs through a series of 'beautiful inscapes' or 'successive sidings of the one inscape.' Each distinguished term in the series is beautiful, but the beauty of the whole series is 'of all the higher degree.' There is a timeless quality to this beauty, where one does not need to be aware of the direction of the change, 'whether the flower were shutting or opening.' "Inscape is the expression of a force which lives in change" (Miller, 1963, p. 289). Here we are close to the metamorphosis of plants recognized by Goethe: "It had occurred to me that in the organ of the plant which we ordinarily designate as leaf, the true *Proteus* is hidden, who can conceal and reveal himself in all forms. Forward and backward the plant is only leaf' (from Goethe's diary of his Italian journey, translation in Bortoft, 2013, p. 9).

In one of his undergraduate essays, Hopkins writes: "the synthesis of the succession should give, unlock, the contemplative enjoyment of the unity of the whole ... The more intellectual, less physical, the spell of contemplation the more complex must be the object, the more close and elaborate must be the comparison the mind has to keep making between the whole and the parts, the parts and the whole. For this reference or comparison is what the sense of unity means; mere sense that such a thing is one and not two has no interest or value except accidentally" (J, 126). In the 'synthesis of the succession', in the 'comparison between the whole and the parts, the parts and the whole', the mind grasps the whole coming to presence in the parts and the parts each expressing the whole. That the whole 'is one and not two has no interest.' The whole is not numerically one; its sense of unity abides in an intensive dimension (Bortoft, 1996, p. 85).

Like Goethe, Hopkins has a fondness for the colour of the sky. This particular sunset poses a special challenge to him:

> A fine sunset: a higher sky dead clear blue bridged by a broad slant causeway rising from right to left of wisped or grass cloud, the wisps lying across; the sundown yellow, moist with

> light but ending at the top in a foam of delicate white pearling and spotted with big tufts of cloud in colour russet between brown and purple but edged with brassy light. *But what I note it all for is this*: before I had always taken the sunset and the sun as quite out of gauge with each other, as indeed physically they are, for the eye after looking at the sun is blunted to everything else and if you look at the rest of the sunset you must cover the sun, but today I inscaped them together and made the sun the true eye and ace of the whole, as it is. It was all active and tossing out light and started as strongly forward from the field as a long stone or a boss in the knop of the chalice-stem: it is indeed by stalling so that it falls into scape with the sky (J, 196, italics added).

Hopkins is observing not only the sunset, but also the change in himself. Before, he saw sun and sunset as 'quite out of gauge with each other.' Now he perceives how the two emerge in relation. Once he 'inscaped them together,' he is able to grasp how the sun 'falls into scape with the sky.'

On another occasion, Hopkins grasps the unity of the moon and the night sky, the clouds and the landscape below:

> Another night from the galley window I saw a brindled heaven, the moon just marked by a blue spot pushing its way through the darker cloud, underneath and on the skirts of the rack bold long flakes whitened and swales like feathers, below / the garden with the heads of the trees and shrubs furry grey: I read a broad careless inscape flowing throughout (J, 215).

A passing configuration of land and sky reveals the underlying wholeness, bringing a sense of 'inscape flowing throughout.' Inscape here is 'not the pattern of particulars within a whole, but the structure of the whole' (Miller, 1963, p. 291).

Compare how Bortoft writes of Goethe's appreciation of sun and sky: "he saw the yellow sun and the blue sky *belonging* together ... When we see the sun and the sky, we usually do so separately. Even if we do notice them together, we do not experience their colours *belonging* together. We experience the colours of the sun and sky in the mode of separation and not in the mode of their unity" (Bortoft, 1996, p. 60). Coming from Heidegger's (1969) discussion of the Eleatic philosopher Parmenides, *belonging* together emphasizes how the parts belong as diverse expressions of a single whole and contrasts with belonging *together*, where parts of different wholes are brought together on the basis of common properties.

To inscape then is to bring together in an authentic relation of belonging, rather than to create a mere pattern or arrangement from without. In Hopkins' terms, this is the foredrawing act, the mind's grasp (J, 129). Foredrawing comes from the Greek root *synechesthai* more usually translated as 'holding together.' Hopkins introduces this word into his own discussion of Being and Not-Being in Parmenides (J, 129):

> Nor is it [Being] divisible, since it is all alike, and there is no more of it in one place than in another, to hinder it from holding together, nor less of it, but everything is full of what is.
> Wherefore all holds together; for what is, is in contact with what is. (Parmenides, VIII, 22-25).[18]

Being is foredrawing, holding together in oneness. "Not-being is here seen as want of oneness, all that is unforedrawn, waste space which offers nothing to the eye to foredraw or many things foredrawing away

[18] From Burnet's 1892 translation of Parmenides' fragments, at http://philoctetes.free.fr/parmenidesunicode.htm, last accessed 19th January 2018.

from one another" (J, 129). Being is *belonging* together, Not-Being belonging *together*.

Further on in Parmenides' eighth fragment, the verses speak of the wholeness of Being as that of a 'rounded sphere', 'equal in all directions':

> Where, then, it has its farthest boundary, it is complete on every side, equally poised from the centre in every direction, like the mass of a rounded sphere; for it cannot be greater or smaller in one place than in another.
> For there is nothing which is not that could keep it from reaching out equally, nor is it possible that there should be more of what is in this place and less in that, since it is all inviolable. For, since it is equal in all directions, it is equally confined within limits. (Parmenides, VIII.42-49)

For Hopkins, Parmenides' intuition of the wholeness of Being echoes that of the founder of the Eleatic school, Xenophanes, of 'a ball rounded and true' (J, 128). Elsewhere in his undergraduate essays Hopkins discusses Xenophanes' belief that, God is "eternal ... one, and sphere like" (E, 309) and writes of "the immoveable spherical One or Whole of Xenophanes" (E, 206, see also Dau, 2012, p. 69). 'Burl' is a word Hopkins uses in his journals to describe the 'fulness' (J, 130) and 'roundness of the world' (J, 251). The origin of this word is unclear, perhaps from 'burly' (Miller, 1963, p. 287) or from 'burls', small round growths from the trunk of a tree (Takagi, 2004). The word crops up again in his devotional writings, when he describes the fulness of being in Christ (S, 156).

For Hopkins in his study of Parmenides, the burl is the 'truth itself', and contrast with mere names, 'eye and lip service to the truth, husks and scapes of it' (J, 130). Here, the burl is the authentic whole, whereas the 'husks and scapes' are inaccurate, partial visions of the

totality, a counterfeit whole in Bortoft's terms (1996, pp. 10-11). If the scapes fall into authentic relation in the whole, we have what Bortoft calls 'multiplicity in unity' (1996, pp. 85-87), where the burl is present in each scape.[19] If they are counterfeit, mere 'husks', we have nothing but an abstract 'unity in multiplicity', where the scapes are separate parts that, when put together, add up to the whole.

Struggling with the meaning of the human self and its moral choices, Hopkins employs a second image of wholeness, that of 'a pomegranate in the round':

> As besides the actual world there is an infinity of possible worlds, differing in all degrees of difference ... each of which possible worlds and this the actual one are like so many 'cleaves' or exposed faces of some pomegranate (or other fruit) cut in all directions across: so there is an infinity of possible strains of action and choice for each possible self in these worlds .. and the sum of these strains would be also like a pomegranate in the round, which God sees whole but of which we see at best only one cleave. Rather we see the world as one cleave and the life of each person as one vein ot strain of colour in it (S, 151).

With this surprisingly concrete image, Hopkins paints a striking picture of wholes within wholes. Each possible world is but one 'cleave' or 'exposed face' of the pomegranate of all possible worlds. Each path of 'action and choice for each possible self' in such a world is also one cleave of a pomegranate. Only God sees the pomegranate

[19] Compare Waterman Ward's contention that "[t]hese multiple real aspects of a single thing are ... what Hopkins called the "inscapes" of a thing" (2002, p. 113). "The object itself has its own oneness and being, the "burl," perceived in multitudinous ways" (p. 119). According to Waterman Ward, the inscapes of a thing are its *formalitates*, the "little forms" of Scotus, the organizing relations within the object (p. 189).

whole, whereas with limited creaturely vision, 'we see at best only one cleave.'

Studying the Latin text of the exercise of St. Ignatius on the Incarnation, Hopkins again turns to the image of the pomegranate:

> 'Magnam capacitatem et ambitum mundi' [*'the great capacity and circuit of the world']*—This suggests that 'pomegranate', that *pomum possibilium*. The Trinity saw it whole and in every 'cleave', the actual and the possible. We may consider we are looking at it in all the actual cleaves, one after another. This sphere is set off against the sphere of the divine being, a steady 'seat or throne' of majesty. Yet that too has its cleave to us, the entrance of Christ into the world. There is not only the pomegranate of the whole world but of each species in it, each race, each individual, and so on (S, 171; Latin translation from Dau, 2012, p. 69).

As creatures we may only look at individual cleaves of the whole, 'one after another.' God, however, is able to see the whole. Echoing Xenophanes, he refers to 'the sphere of the divine being', which is not closed to us, but open, through the Incarnation of Christ, as Word made flesh (John 1:14). Christ enters and underpins the wholeness of each being, for there is 'not only the pomegranate of the whole world but of each species in it, each race, each individual, and so on.'

For the pomegranate or wholeness of each species, race and individual within the world, each has its own inscape. The 'mixture of strength and grace, like an ash', captures the inscape of the individual bluebell Hopkins held in his hands. 'Kingfishers catch fire' declares the inscape, the particular form and activity, of a species of bird, and 'dragonflies draw flame' the inscape of a higher kind, that of the order Odonata. Following St. Ignatius' Principle and Foundation, each inscape, whether that of the individual, the species or any natural

whole, is to lead us to God. The extinction of the Quagga left the world with one fewer inscape, striped at the front like a zebra, hind parts without stripes like a horse or donkey. Even if the Quagga is but a race, a subspecies of the Plains Zebra, Hopkins would have mourned its passing, "... for Hopkins the desecration of natural beauty must hinder, if not prevent altogether, the attainment of the end of which St. Ignatius speaks" (Newman, 1999, p. 214).

Instress

Hopkins introduces the term instress in his notes on Parmenides: "His great text ... is that Being is and Not-being is not—which perhaps one can say, a little over-defining his meaning, means that all things are upheld by instress and are meaningless without it .... His feeling for instress, for the flush and foredrawn, and for inscape / is most striking ..." (J, 127). Instress is first of all a quality within the natural object, "the inner energy of being which upholds things" (Miller, 1963, p. 287), an 'inner flushness' (J, 128) which floods human consciousness with the experience of the object (Waterman Ward, 2002, p. 115). The prepossession of a word, to use an expression from another of Hopkins' undergraduate essays, conveys this inner energy of flushness: "To every word meaning a thing and not a relation [i.e., a whole and not a part] belongs a passion or prepossession or enthusiasm which it has the power of suggesting or producing but not always or in everyone..." (J, 125).

In Hopkins's terms, inscape refers to form and relation, while instress refers to the energy that holds together the natural object and emanates from it. We have the two aspects that we saw in jizz. Jizz is the bird's gestalt and equally the 'vibe' that the bird gives off.

Jizz speaks of an expressiveness of the bird that creates an characteristic impression on the watcher. Expression and impression are the two sides of the one experience. Instress is the bird's

expressive unity that carries over its being to the watcher, that 'impulse from the inscape' which 'actualises the inscape in the mind of the beholder' (Gardner in Hopkins, 1953, p. xxi).

Hopkins continues in his essay on Parmenides: '... indeed I have often felt when I have been in this mood and felt the depth of an instress or how fast the inscape holds a thing that nothing is so pregnant and straightforward to the truth as simple *yes* and *is*' (J, 127). Instress, as an inner energy, is that which beckons to us and reclaims our *yes*, our own activity of instressing. This is a moral imperative, to respond with love to 'God in all things.' "Instress is an action of the will—a moral action, for good or evil. Instress is assent[20] ... to an inscape" (Waterman Ward, 2002, p. 198). The *yes* of instress accompanies the *is* of inscape (*ibid*., p. 200). The birdwatcher, to see the bird, must assent to the bird, must instress it. The bird comes to stress in the birdwatcher. Otherwise, "[t]here would be no bridge, no stem of stress between us and things to bear us out and carry the mind over' (J, 127).

Triad

"Fineness, proportion, of feature," Hopkins writes to his fellow poet Coventry Patmore, "comes from a moulding force which succeeds in asserting itself over the resistance of cumbersome or restraining matter... "(F, 306). Gardner (in Hopkins, 1953, p. xxi) links instress as moulding force to the "one Spirit's plastic stress" that "sweeps through the dull dense world" of matter in Percy Bysshe Shelley's lament *Adonaïs*.

The 'fineness and proportion of feature' in the inscape comes from the 'moulding force' of instress asserting itself over the resistance of matter. Here we have a clear articulation of the triadic quality of

[20] Waterman Ward (2002) takes the term "assent" from John Henry Newman's *Grammar of Assent* (1874). See her ch. 5 for an in-depth discussion of the relationship between the ideas of Hopkins and Newman.

Hopkins' philosophy. Inscape, matter and instress are the three necessary terms. As we have seen, inscape in its usage relates to form, instress to a kind of inner energy flowing through matter. Hopkins would seem to be introducing a third term in addition to classic terms of form and matter.

Hopkins goes on in the same letter: "the bloom of health comes from the abundance of life, the great vitality within. The moulding force, the life, is the form in the philosophic sense, and in man this is the soul." He gives a vivid image of instress, as 'the bloom of health' and 'the great vitality within.' This certainly conveys the sense of an inner animating energy, active within and visible without. But Hopkins sticks to the classical terms, with the soul, as the animating energy, qualified as form. Inscape is thus deprived of its quality of form and three terms are reduced to two again.

In her account of the different uses of instress in Hopkins' writings, Waterman Ward enumerates certain instances where instress "seems to be the will" (2002, p. 199): "all that energy of instress by which the soul animates and otherwise acts in the body" (S, 137). "Body no longer swayed as a piece by the nervous and muscular instress seems to fall in and hang like a dead weight..." (J, 238). When we speak of 'losing the will to live,' we speak of a failure of the person's inner animating energy.

To instress an inscape is to assent to it with our will. Is not human will here answering to God's will? For "one might properly speak of God holding things in existence by instressing them, by being faithful to the commitment implicit in "Let there be..."" (Waterman Ward, 2002, p. 204). As the psalmist says, "For He spoke, and it was done; He commanded, and it held firm" (Psalm 33:6-9).

Duns Scotus was a champion of the will in the mediaeval debate over the primacy of will or intellect among the powers of the soul (see

Windelband, 1958, pp. 328-34). Dominican Albert the Great and his disciple Thomas Aquinas favoured the intellect, believing that the perfect state of human being is participation in the divine being through knowledge. For the Franciscan Scotus, love is higher than intellectual contemplation and springs from the will. For him, God's will is absolute in creation and overwhelms any intellectual requirement of necessity or regularity. Thomas Aquinas followed Augustine's doctrine of the irresistible grace of God, whereas Scotus insisted on the contribution of the person's free will (the "synergistic doctrine," see Blake, 1966). Hopkins in his essay 'On Personality, Grace and Free Will' concurs. God's grace acts in us through his will joined to ours, "the free act on our own part, of consent or refusal as the case may be" (S, 149).

Fons Vitae or *The Source of Life* was a Latin text widely attributed to a Christian author, usually known as Avicebron. The text was taken up into the mediaeval debate of will vs intellect, championed by the Franciscans and criticized by the Dominicans (Pessin, 2016). The Latin text turned out to be a translation of an Arabic original and the author the Jewish poet, Solomon ibn Gabirol.

> Disciple: "Why is it said that the sublime and holy Creator is in all things?"
> Master: "Assuredly this is said, because His Virtue is infused in all things, and that it penetrates all things, and that nothing exists without it and that all things have their existence and their constitution through it." (*Vitae* V.39, translation in Blake, 1966)

Asked by his disciple to justify that God 'is in all things,' ibn Gabirol's master responds by invoking God's virtue, the power of his will in act, that is 'infused in all things' and that 'penetrates all things.' This is a very good description of instress, the third term in Hopkins'

philosophy. Asked by the disciple to explain his last comment, the master replies:

> "Do you not see, in fact, that the essence of all things comes only from Matter and Form, and that the existence of Matter and Form comes from the Will, since it is the Will which has created them, which has conjoined them and which holds them united? When we say that Form holds Matter, we are speaking in an incorrect manner: it is from the Will that Form receives the power to hold Matter; it is manifest for the Form as an impression of the Unity, a reverberative force originating in the Unity, and the Will is itself the power of Unity; therefore the reverberative virtue comes from Unity. The Will holds Matter by the intermediary of Form: this is why one says that Form holds Matter." (*Fons Vitae*, V, 39, in Blake, 1966)

With ibn Gabirol, we move from a dyad of form and matter to a triad of form, matter and will (Blake, 1966).

Hyparxis

In the reply to his disciple, ibn Gabirol's master describes how the Will, issuing from the One, is a 'reverberative force,' a force that creates repetition and resonance but carries all the same an 'impression of unity.' The moulding force of instress overcomes the resistance of matter in the same way, creating recurrence and rhyme:

> The offspring of this marriage of spirit and matter are all different from one another. The instress of being, though it is everywhere the same, has a mysterious tendency to express itself differently each time it subdues cumbersome and restraining matter to pattern. The physical conditions moulding snowflakes are always similar, but no two

> snowflakes, among the untold billions which have fallen, have ever been exactly alike.... (Miller, 1963, p. 290)

One expression of this repeating pattern is of an ensemble of rhyming elements. The night sky, for example, is pictured as a garment into which stars are woven as sequins:

> The sky minted into golden sequins.
> Stars like gold tufts.
> — — golden bees.
> — — golden rowels.
> Sky peak'd with tiny flames
> Stars like tiny-spoked wheels of fire (J, 46-47)

Each star is a 'tiny wheel of fire' but together they create an array of rhyming elements which contributes to the beauty of the whole. Similarly, a field is "pinned with daisies" (J, 134) or "noted with primroses" (J, 55) like the dress of a young woman decked out for a ball. The sky is fastened with grey clouds as with buttons, or, as Hopkins likes to say, 'knops' (J, 46). A fragment of a poem celebrates a summer orchard, "knopt/ With green-white apples on the bough" (J, 68), another "the flocks of villages/ That bead the plain" (J, 39). Stars, flowers, clouds, apples and villages are "nodes or beads of energy" (Miller, 1963, p. 291), actualized at regular intervals in time and space within the field of instress.

In other journal entries, the wavy folds and weaves of Nature's fabric come to the fore:

> The lines of the fields, level over level, are striking, like threads in a loom (J, 23).

> I was noticing the strange rotten-woven cloud which shapes in leaf over leaf of wavy or eyebrow texture: it is like fine webs of gossamers held down by many invisible threads on

> the under sides against a wind which between these points kept blowing them up into bells (J, 184).
>
> [Mountains] run like waves in the wind, ricked and sharply inscaped... (J, 180)
>
> In the snow flat-topped hillocks and shoulders outlined with wavy edges, ridge below ridge, very like the grain of wood in line and in projection like relief maps. These the wind makes I think and of course drifts, which are in fact snow waves (J, 230).

We have images of woven threads 'fine webs of gossamers', 'level over level,' 'leaf over leaf.' The 'wavy edges' of the snow-topped hillocks, 'ridge below ridge,' form a motif 'very like the grain of wood.' The wind makes 'snow waves' and mountains 'run like waves in the wind.' To relate threads in the loom to ocean waves in the wind conveys an image of nature changing subtly with each iteration, responding delicately to outer conditions, but keeping all the same an inner logic to the sequence of its iterations. Instress "weaves nature together and makes it one" and governs by a "principle of recurrence and rhyme" (Miller, 1963, p. 292).

Hopkins' recurring and rhyming patterns of instress echo Goethe's true Proteus, the leaf. Each leaf emerges from the stem as a separate manifestation of the inner protean impulse. There is no transformation of one leaf into another, yet we can sense the inner logic, the rhyme between leaf occurrences. The leaf becoming the various parts of the flower remains itself but becomes different. "Goethe's 'one and the same organ' is a self-differencing organ producing differences out of itself" (Bortoft, 2012, p. 71). The rhythmic, repetitive, sustaining quality—'reverberative force'—of instress is the whole's self-differencing. This ability to remain the same while becoming different is what it is to be alive (Bortoft, 1996, p. 268). Instress, manifestation of the Divine will, is the 'moulding force, the life' (F, 306).

For philosopher John Bennett, will manifests as recurrence within the dimension of *hyparxis* (1956, pp. 152, 166). In Bennett's six-dimensional cosmology, the three dimensions of space are joined by three of time. The dimension of actuality is the dimension of time as commonly understood. Potentiality is the dimension of eternity, a vast reservoir of possible states of being. Hyparxis is the condition of recurrence, whereby an entity is able to actualize while retaining its potential to reactualize, to come back differently (Bortoft, 2012, p. 206).

The simple seed leaves of the Meadow Buttercup are round, but that is not all that Leaf, the inner Proteus has to give. The subsequent leaves to be actualized are progressively more branched and more deeply incised until the uppermost leaves become simple again, reduced to three, deeply incised 'spines'. Each leaf leads to an 'increase in being' of Leaf. Leaf becomes more itself by remaining itself and becoming different. There is an increase in the depth of hyparxis, an increase in what Bennett calls "ableness-to-be" (Bortoft, 2012, pp. 205-206).

For Bennett, hyparxis is a *reconciliation* between potentiality in eternity and actuality in time. "Potentiality is the affirmation of a set of possible events—it is the pattern inherent in a given situation—whereas actualization is rejection of potentialities—the denial of the significance of the pattern as a totality in order to permit one of its facets to become actual... The reconciliation of potential and actual results from the property of 'ableness-to-be'. That which is not able to be itself loses its potentialities in the moment of actualization. The more one is oneself, the more one is able to actualize without losing contact with potentiality" (Bennett, 1956, p. 166).

Identity

The interplay of three impulses, (1) affirming, (2) receiving or denying and (3) reconciling, is the structure that Bennett gives to the triad (1956, p. 39). We can detect impulses having these different qualities in Hopkins' triad: "Fineness, proportion, of feature comes from a moulding force which succeeds in asserting itself over the resistance of cumbersome or restraining matter..." (F, 306). Matter is clearly the receiving or denying impulse in the triad. It is transformed through the action of instress, the moulding force, into the 'fineness, proportion, of feature' of inscape. Instress is thus the reconciling impulse, with inscape as the affirming impulse.

It may seem odd that inscape as the affirming impulse in Hopkins' triad appears as the outcome, that which 'comes from' the action of instress on matter. Our first thought would be that the affirming impulse initiates the process. The striking insight of Bennett (1961) is that there are triads of different character depending on the way the three impulses combine (pp. 103 ff.). Hopkins' triad corresponds to Bennett's triad of *identity* (2-3-1). Recast in Bennett's rather enigmatic formulation: matter (2) meets with instress (3) and is linked thereby to inscape (1) that issues as manifestation (1961, p. 113).

Instress, as the action of the Divine Will, is "the power that preserves and reconciles everything that is, and enables it to occupy its place in the Universal Scheme" (Bennett, 1961, p. 114). As creatures, the affirmation of our existence only isolates us from other entities (*ibid.*, p. 115). For Hopkins, Satan's sin is "instressing his own inscape" (S, 201). By denying ourselves, by becoming receptive to the will of God, we become able to be what we are (*ibid.*, p. 115): "..we have this treasure in jars of clay to show that this all-surpassing power is from God and not from us" (2 Corinthians 4:7).

The inscape of 'strength mixed with grace, like an ash tree' manifests the identity of the individual bluebell that so fascinated Hopkins.

'Kingfishers catch fire and dragonflies draw flame' expresses the identity of a species of bird and a kind of insect. As Hopkins proclaims: 'Each mortal thing does one thing and the same/... *Myself* it speaks and spells/Crying '*What I do is me: for that I came'*.'

For Bennett, the receiving impulse in the triad of identity is only matter in the sense of that which exists, but also in the sense of the undiscriminated phenomenon, the matter of the mind's grasp. The triad accounts not only for identity in being—'what I do is me'—but also, identity in perception—that 'impulse from the inscape that actualizes the inscape in the mind of the perceiver,' (Gardner on instress) that which 'instantly registers identity in the brain' (Coward on jizz). "The undiscriminated phenomenon is the denying factor which, without changing its nature, comes by the act of knowing to be distinguished as fact" (Bennett, 1961, p. 115). The second formulation of the triad of identity thus becomes: The undiscriminated phenomenon (2) through instress (3) becomes actualized as inscape (1) in the mind of the perceiver. "What we perceive is **really** there because of its own identity, and we are **really** linked with it because there is a common identity in the percipient experience." (Bennett, *ibid*;, p. 115).

With Bennett's triad of identity we capture the two aspects that we saw in jizz. From the side of being, jizz is that characteristic integration of size, shape, plumage, flight and note that makes the bird what it is. Jizz, in the sense of inscape, is the affirmation of identity. From the side of perception, jizz is the expressiveness of the bird that creates the characteristic impression on the watcher, the energy that emanates from it and fills the observer with the delight of recognition. Jizz, in the sense of instress, is the reconciliation of identity.

References

Hopkins primary sources

E = Oxford Essays
The Collected Works of Gerard Manley Hopkins: Volume IV: Oxford Essays and Notes 1863-1868. Edited by Lesley Higgins. Oxford: Oxford University Press, 2006.

F = *Further Letters of Gerard Manley Hopkins, including his correspondence with Coventry Patmore*. Edited with notes and a introduction by Claude Colleer Abbott. London: Oxford University Press, 1956.

J = *The Journals and Papers of Gerard Manley Hopkins*. Edited by Humphry House, completed by Graham Storey. London: Oxford University Press, 1959.

P = *The Poems of Gerard Manley Hopkins*. Edited by W. H. Gardner and N. J. MacKenzie. London: Oxford. University Press, 1967.

S = *The Sermons and Devotional Writings of Gerard Manley Hopkins*. Edited by Christopher Devlin, SJ. London: Oxford University Press, 1959.

Other works

Anon. 1946. *Aircraft Recognition Training (All Arms) 1946*. London: War Office.

Ball, P. M. 1971. *The Science of Aspects: The Changing Role of Fact in the Work of Coleridge, Ruskin and Hopkins*. London: Athlone Press.

Bennett, J. G. 1956. *The Dramatic Universe, vol. 1: The Foundations of Natural Philosophy.* London: Hodder & Stoughton.

Bennett, J. G. 1961. *The Dramatic Universe, vol. 2: The Foundations of Moral Philosophy.* London: Hodder & Stoughton.

Blake, A. G. E. 1966. Implications of Avicebron's notion of will. *Systematics* 4(1), June 1966.

Bortoft, H. 1970. The ambiguity of 'one' and 'two' in the description of Young's experiment. *Systematics* 8(3), December 1970.

Bortoft, H. 1996. *The Wholeness of Nature: Goethe's Way of Science*. Edinburgh: Floris.

Bortoft, H. 2012. *Taking Appearance Seriously: The Dynamic Way of Seeing in Goethe and European Thought*. Edinburgh: Floris.

Bortoft, H. 2013. The form of wholeness. *In Context* 29, Spring 2013.

Chichester, F. 1941. *The Spotter's Handbook*. London: Allen & Unwin.

Coward, T. A. 1922. *Bird Haunts and Nature Memories*. London: Warne.

Dau, D. 2012. *Touching God: Hopkins and Love*. London: Anthem.

Dooley, S. 2005. *Big Twitch*. Crows Nest: Allen & Unwin.

Draper, I. 1985. The origin of "jizz". *British Birds* 78: 252.
Fitter, R. S. R. & Richardson, R. A. 1966. *Collins Pocket Guide to British Birds: The Complete Identification Book.* London: Collins.

Fitter, R. S. R. (ed.) 1974. *Book of British Birds*. London: Drive Publications.

Gardner, W. H. 1944, 1948. *Gerard Manley Hopkins (1844–1899): A Study of Poetic Idiosyncrasy in Relation to Poetic Tradition*. 2 vols. London: Secker & Warburg.

Heidegger, M. 1969. *Identity and Difference.* Translated and with an introduction by Joan Stambaugh. New York: Harper & Row.

Hopkins, G. M. H. 1953. *Poems and Prose*. Selected with an introduction and notes by W. H. Gardner. Harmondsworth: Penguin.

Lockwood, W. B. 1984. *The Oxford Book of British Bird Names*. Oxford: Oxford University Press.

Loyola, Ignatius. 1964. *The Spiritual Exercises*. Translated by Anthony Mottola. New York: Image.

McDonald, D. 1996. The etymology of "jizz." *Canberra Bird Notes* 21(1): 2-11.

Merleau-Ponty, M. 1962. *Phenomenology of Perception*. Translated by Colin Smith. London: Routledge & Kegan Paul.

Merleau-Ponty, M. 1964. *The Primacy of Perception.* Evanston: Northwestern University Press.

Miller, J. H. 1963. *The Disappearance of God: Five Nineteenth-century Writers.* University of Illinois Press.

Newman, J. H. 1874. *An Essay In Aid Of A Grammar Of Assent*. London: Burns, Oates, & Co.

Newman, N. F. 1999. Hopkins's Ribblesdale. *The Explicator* 57(4): 214-215.

Ong, W. J. 1986. *Hopkins, the Self, and God*. Toronto: Toronto University Press.

Partridge. E. 1984. *A Dictionary of Slang and Unconventional English*. Eighth edition by P. Beale. London: Routledge & Kegan Paul.

Pessin, S. 2016. Solomon Ibn Gabirol [Avicebron]. In Zalta, E. N. (ed.) *The Stanford Encyclopedia of Philosophy* (Winter 2016 Edition). Available at https://plato.stanford.edu/ archives/win2016/entries/ibn-gabirol.

Peters, W. A. M. 1948. *Gerard Manley Hopkins: A Critical Essay towards the Understanding of his Poetry*. London: Oxford University Press.

Sobolev, D. 2011. *The Split World of Gerard Manley Hopkins: An Essay in Semiotic Phenomenology*. Washington DC: Catholic University of America, Press.

Takagi, S. 2004. ‘As kingfishers catch fire' - a theological reading. Available at http://www.gerardmanleyhopkins.org/lectures_2004/theological_view.html, last accessed 23rd January 2018.

Waterman Ward, B. 2002. *World as Word: Philosophical Theology in Gerard Manley Hopkins*. Washington DC: Catholic University of America, Press.

Windelband, W. 1958. *A History of Philosophy, Vol. 1: Greek, Roman and Medieval*. New York: Harper & Row.

Zaharek, K. E. & Overstreet, L. K. (2000). François Nicolas Martinet. http://www.sil.si.edu/DigitalCollections/NHRareBooks/ Martinet/martinet-introduction.htm, last accessed 24th January 2018.

YOUR NOTES