THE FINE STRUCTURE CONSTANT

MICHAEL ATIYAH

For Lily.

1. Historical Introduction

Eureka and Archimedes

The fine-structure constant α is a dimensionless number that is ubiquitous in physics, but has remained an enigma for over a century. Does it have mathematical significance analogous to π? Its numerical value is now known accurately to 12 significant figures but it has no satisfactory mathematical explanation as shown by the following opinions.

Here is what Richard Feynman had to say about α: *Where does α come from; is it related to π, or perhaps to e? Nobody knows, it is one of the great damn mysteries of physics: a magic number that comes to us with no understanding by man. You might say the hand of God wrote the number and we don't know how He pushed his pencil.*

The mathematical statistician I.J.Good argued that a numerological explanation would only be acceptable if it came from a more fundamental theory that provided a Platonic explanation of the value.

By contrast, the real number π was fully understood more than two thousand years ago by Archimedes, who showed how to calculate it to arbitrary accuracy.

In the mid 18th century, Euler discovered the deep relation between π and complex numbers $\mathbb{C}$, embodied in his famous theorem $e^{2\pi i} = 1$. Euler's proofs had to wait for Riemann a century later. In the same period Hamilton discovered the quaternions $\mathbb{H}$ which generalized complex numbers from 2 to 4 dimensions, but were non- commutative. Although physicists embraced the non-commutative group $SU(2)$ of unit quaternions, as opposed to the commutative group $U(1)$ of complex numbers, they had problems with the algebra of non-commutative polynomials. *In particular there appeared to be no quaternionic version of Euler's formula.*

In the early 20th century, von Neumann (and Murray) discovered the algebras that now bear his name, and factors of Type II in which dimensions take all (positive) real values, in contrast with factors of Type I which take just integer values. This was later refined by Alain Connes into a theory of non- commutative geometry [2].

Around the same period, Eddington speculated controversially and inconclusively on the fine-structure constant [13].

In the mid-20th century, Hirzebruch [1] created the formalism of the Todd genus, which revolutionized geometry and topology. Spurred on by this, Atiyah, Bott and Singer created index theory, which was taken up enthusiastically by physicists.

In this paper I will weave all these diverse strands together to provide a rigorous and elegant mathematical model of the fine structure constant α, or rather $1/\alpha$. It will be denoted by the Cyrillic letter ж which I will connect both to π and to e, answering Feynman's plea. It arises from a fundamental Platonic theory as required by Good. This theory is called renormalization and it rests on solid mathematical foundations.

Renormalization is a flow involving change of scale which physicists think of as Energy. Under this flow, numbers get renormalized, and when taken to the limit, π gets renormalized to ж. The direction of the flow depends on the whether numbers increase or decrease and is a matter of convention. The standard convention is that Energy increases so π has to increase to ж, which models $1/\alpha$.

Before proceeding further, I should point out that physicists already have a model for α which fits remarkably well with experimental data. This model rests on Feynman diagrams, but these have shaky foundations and involve herculean computational work. So even Feynman himself, as the quotation above makes clear, wanted to know what this mysterious number α really was. The purpose of this paper is to answer Feynman's question. I will produce a Platonic answer which does not rest on dubious numerology or experimental data.

This is a *mathematical* paper in the spirit of Archimedes, Euler and their modern successors. It provides a mathematical model for the physical world, and can be interpreted as giving a rigorous foundation for the results derived from Feynman diagrams, without the heavy computational work.

However theoretical physicists have struggled to get to 9 decimal places and a Platonic number should be calculable, like π, to arbitrary accuracy. I will therefore indicate how to produce a decimal expansion of ж to, say 12 decimals, to impress any sceptics. This is the one place where I delve not into numerology, but into numerical calculation intrinsic to the theory. Moreover I use a fast algorithm that produces 9 decimal places in 3 tranches of 4 steps. The extension to 12 decimal places probably requires just 5 further steps.

The plan of the paper is as follows. In section 2, I explain the main strategy and use renormalization to define the Todd map T (see 3.4) This gives ж $= T(\pi)$, putting it in Plato's world. This already achieves my main aim of explaining *what α really is.*

For computational purposes, and to relate to Feynman diagrams, I then introduce, in section 7, ч $= T(\gamma)$ where γ is Euler's constant γ . This Cyrillic letter is related to ж, by the formula

$$ч/\gamma = ж/\pi. \tag{1.1}$$

In section 7, I build on this formula showing how it helps with the computation of ж. There is one further step, inspired by Euler, to speed up convergence which makes the computation much simpler. All this is put together in section 8.

2. Statements and Strategy

The Dublin bridge and Hamilton

In this section I will indicate the main strategy and formulate the precise statements which will be proved. The key question is how to extend the Archimedes-Euler approach to π, from the commutative world of $\mathbb{C}$ to the non-commutative world of Hamilton's quaternions $\mathbb{H}$, constructed by *an infinite iteration of exponentials, discrete or continuous*. The key idea goes back to von Neumann as I now explain.

2.1 The von Neumann hyperfinite factor A **of type II** . The hyperfinite type II-1 factor A is unique up to isomorphism. It is constructed by an infinite sequence of iterated exponentials, followed by taking the weak closure in Hilbert space. This process converts type I, with integer dimensions, to type II with real dimensions. A dimension which is formally infinite in type I becomes finite in type II, so I call the process **renormalization**. All factors have an invariant trace mapping the factor to its centre. Any inner automorphism gives an isomorphic but different trace. The comparison between two such automorphisms is what leads to the Todd map T (see 3.4).

2.2 The Hirzebruch Formalism. Hirzebruch [1], following in the footsteps of Euler and Riemann, introduced a formal algebraic process of multiplicative sequences. In such processes he defined exponentials over $\mathbb{Q}$. He showed that any such exponential has a generating function, and he focused on the Todd exponential, whose generating function is the Bernoulli function $\frac{x}{1-e^{-x}}$. The fact that this function is analytic implies that the Hirzebruch process extends from $\mathbb{Q}$ to $\mathbb{R}$, implementing the weak closure of 2.1.

2.3 Hirzebruch and algebras over $\mathbb{Q}$. The relation between 2.1 and 2.2 is best understood as follows. The von Neuman algebra A is defined over $\mathbb{R}$, has $\mathbb{R}$ as centre and weights (4.3) in $\mathbb{R}$. The renormalization flow of weights (4.4) in $\mathbb{R}$ is continuous with weights modeling Energy . The Hirzebruch algebra $A(\mathbb{Q})$ is defined over $\mathbb{Q}$, has centre $\mathbb{Q}$ and weights in $\mathbb{Q}$. The flow of weights in $\mathbb{Q}$ is discrete. The embedding of $\mathbb{Q}$ into $\mathbb{R}$ induces a functor from Hirzebruch algebras to von Neuman algebras which takes $A(\mathbb{Q})$ into A. A and $A(\mathbb{Q})$ are unique up to isomorphism in their respective categories. Changes of weight induce isomorphisms (4.4). Since $\mathbb{R}$ is the weak closure of $\mathbb{Q}$ (by Dedekind sections), A is the weak closure of $A(\mathbb{Q})$. A is constructed (section 3) from a sequence of exponential operations. $A(\mathbb{Q})$ is constructed by Hirzebruch (see section 2.4) from a similar sequence. $A(\mathbb{C})$ is the complexification of $A(\mathbb{R})$.

These constructions are functorial and map $A(\mathbb{Q})$ into A and embed the centre of $A(\mathbb{Q})$ into the centre of A. The two different weights, $r < 1/e < 1/r$, give different central isomorphisms in 2.2 which differ by the automorphism denoted by T (3.4). For $A(\mathbb{Q})$ these inverse weights must be rational. In 4.6 we make the simple choice $1/r = 4$ but, for rapid convergence in section 7, we choose $1/r = 16$.

2.4 Synthesis. Using 2.1 and 2.2, π is renormalized to a positive real number denoted by the Cyrillic letter ж. Euler's formula $e^{2\pi i} = 1$ gets renormalized to the Euler-Hamilton formula $e^{2Жw} = 1$, where $w = \frac{\pi}{ж} i$.

2.5 Conclusion. The number ж is a mathematical candidate for $1/\alpha$. It satisfies the criteria of both Feynman and Good, restores the reputation of Eddington and enhances that of Hirzebruch. More details are set out in sections 3-8.

Finally, this explanation of α should put an end to the anthropic principle, and the mystery of the fine-tuning of the constants of nature. Nobody has ever wondered what the Universe would be like if π were not equal to 3.14159265... Similarly no one should be worried what the Universe would be like if ж were not

137.035999...

2.6. There is one more important constant of nature and that is Newton's constant G (whose ratio to $c^3/\hbar e$ is dimensionless, where e is not Euler's exponential but the charge of the electron). In a subsequent paper [7], I will examine this question by starting with the octonions. The non-associativity of the octonions is a clear sign that this is a much harder problem. This will not surprise followers of Einstein's General Theory of Gravitation. The von Neumann factor this time will be the hyperfinite Type III, which has no preferred rational weight.

3. Details on the hyperfinite factor

The Königsberg bridges and Euler

In this section we will use the hyperfinite factor $A = A(\mathbb{C})$. We recall its definition and basic properties.

3.1. A is the weak closure of the infinite tensor product of $\mathrm{End}(\mathbb{C}^2) = \lim A(n) =$ Clifford algebra of Hilbert space as explained below.

3.2. In 3.1, $A(n)$ is mapped into $A(n+1)$ by putting the identity 1 in the $(n+1)$th place.

3.3. The formal Hilbert space H with coordinates z_j, (for $j = 1, 2, ...$) is given the Hermitian metric in which the nth coordinate is multiplied by 2^{-2n}. This ensures convergence of the trace for the Clifford algebra A of this metric. The weak closure is taken in $H \otimes H$.

3.4. The trace on $\mathrm{End}(\mathbb{C}^2)$ induces a trace on A, with values in $\mathbb{C}$. The restriction to the centre $C(A)$ gives the renormalization of numbers. Note that $C(A)$ is isomorphic to $\mathbb{C}$

by two different isomorphisms t_+ and t_- corresponding to the two eigenvalues of the 2×2 matrix. The Todd map $T : \mathbb{C} \to \mathbb{C}$ is just the change from one to the other: $T = t_-^{-1} \circ t_+$. Note that while traces are linear, the Todd map is highly non-linear.

3.5. Just as $\mathbb{H}$ is the affine part of the right projective line $\text{Proj}(\mathbb{H}^2)$, based on the first quaternionic coordinate, so A is the affine part of $\text{Proj}(A^2)$. The anti-involution $u \to 1/u$ of $\mathbb{H}$, switching left and right quaternions, induces an anti-involution $*$ on $\text{Proj}(A^2)$. This leaves invariant the complex projective line $\text{Proj}(\mathbb{C}(A)^2)$ and the 8 points on it : $1, \pm i, \pm j, \pm k$. The map T above, which is the coordinate description of $*$, is also multiplication by the central element w of 3.6 and so the renormalization of multiplication by i as formulated in 3.7.

3.6. There is a unique element w with $T(w) = i$, while the identity element 1 has $T(1) = 1$. $T(\pi)$ is denoted by ж.

3.7. T is the map that maps Euler's formula to the Euler-Hamilton formula, as in 2.4.

4. Weights and weak closure

The magic of Bernoulli

4.1. If we use powers of say 3 or 5 in section 3, instead of powers of 2, we will get different algebras, just as rational numbers with powers of 2 or 3 or 5 in the denominator are not isomorphic.

4.2. But, if we pass to the weak closures, to form the von Neumann algebras, all become isomorphic, since the rationals with denominators powers of 2 or 3 or 5 are dense in the reals. That is why A is unique (up to isomorphism).

4.3. The inverse of the chosen integer, such as 2,3, or 5, which was used in section 3, is called the weight. It is a point in the open interval (0,1).

4.4. Rescaling by $x \to \frac{x}{1-e^{-x}}$, on the inverses, leads to the flow of weights and establishes concretely the isomorphism of A based on different weights. See section 6 for more on the Bernoulli function.

4.5. As explained in 2.7, for the octonions the von Neuman algebra is of type III. The flow of weights is then more subtle and will be examined in [7].

4.6. Renormalization, ignoring the weight, is conformal, depending only on the conformal structure of $\mathbb{C}$. The weight of the complex number $z = re^{i\theta}$ is $\ln(r)$, which lies in $(0, 1)$ provided $r < e^{-1}$. Any such choice of weight gives the unit disc in $\mathbb{C}$ the hyperbolic metric of curvature -1. Since $e < 4$, the simplest (inverse) integer weight is $r = 4$. We could make this choice but, as we will see in section 7, it is better to take $r = 16$.

4.7. From its definition, involving the operation $x \to 2^x$, the Todd map of section 3 is, at each stage, exponential (taking addition into multiplication) but we need to examine what happens when we iterate and pass to the weak closure. This problem was essentially dealt with in the formalism of multiplicative sequences, developed by Hirzebruch, without reference to von Neumann algebras. This will be explained in the next section.

5. The Hirzebruch formalism

The Todd genus and Hirzebruch

5.1. For the case of inverse-integer weights, Hirzebruch formalized the notion of exponential maps and found the most general solutions [1]-Chapter 3. His motivation came from the behaviour of certain manifold invariants, in algebraic geometry, under taking Cartesian products. But the problem was purely formal and applied to all manifolds, giving topological invariants [especially after Atiyah-Singer index theory]. He showed that the basic example was the Todd genus, named after J.A. Todd, whose generating function, due to Bernoulli, is $\frac{x}{1-e^{-x}}$, already exploited by Euler. It is crucial that this function, which underpins Hirzebruch's formalism, is analytic in the closed interval $[0, 1/2]$ (see section 6). This ensures that the normalized trace of 4.7 extends to the weak closure A. Following Hirzebruch, we have named it after Todd and denoted it by T . Hirzebruch's formalism gave spectacular explicit formulae and led to his early fame.

5.2. Recall that Archimedes calculated the circumference of a circle as the upper bound of inscribed regular N-gons where $N = 2^n$, while the same limit was reached from above by circumscribed N-gons Of course, Archimedes knew that squares could be replaced by regular polygons with any number of sides (and that they need not be regular). This is the antecedent of A being essentially independent of the weight.

5.3. Archimedes also deduced the formula πr^2 for the area of a circular disc of radius r, from his *tombstone* result. This was his great theorem (which he wanted inscribed on his tombstone), that a sphere has the same area, slice by slice, as a cylinder. In modern parlance they are symplectically equivalent.

5.4. In the complex plane the unit disc and the whole plane are conformally distinct, which is why the hyperbolic plane differs metrically from the Euclidean plane. But, as the radius of the disc tends to infinity, the curvature of the hyperbolic plane tends to zero and we recover the Euclidean plane. All geometric formulae in the Euclidean plane are limits of formulae in the hyperbolic plane, and it is sometimes easier to get Euclidean formulae this way. An example is the famous theorem of Gauss for the area of a spherical triangle, which follows directly from the tombstone theorem of Archimedes described in 5.3.

5.5. As mentioned at the beginning of section 2, when $\mathbb{C}$ is replaced by $\mathbb{C}^2 = \mathbb{H}$, this has to be renormalized because $\mathbb{H}$ is not commutative. It will then be viewed as the limit of the renormalized hyperbolic 4-ball, as will be explained in detail in section 7.

6. Analytic functions

The genius of von Neumann

This section is a digression about the role of analytic functions in mathematical models.

6.1. Let m be a positive real number. An analytic function $f(m, r)$ of the real variable r in the closed interval $[0, 1/m]$, extends to an analytic function $F(m, z)$ of the complex variable $z = e^{imr\theta}$, with branch points at $r = 0$ and $r = 1/m$. By Euler's formula it is then periodic in r with period $2\pi/m$, so that there are branch points for all $r = k/m$ with $k \in \mathbb{Z}$.

6.2. There is a symmetry (duality) about the mid-point $r = 1/2m$ which corresponds to inverting z or changing the sign of theta.

6.3. The Bernoulli function $B(z) = \frac{z}{1-1/z}$ is the mid-point $(m = 2)$ value of $B(m, z)$. It leads to the function $b(2, r) = \frac{e^r}{1-e^{-r}}$.

6.4. The Hirzebruch formalism establishes analyticity based on the Bernoulli functions of 6.3. For Hirzebruch, m is the Chern class of a line-bundle. The weight of the von Neumann factor is just a point in the open interval $(0, 1)$. The flow of weights just corresponds to moving the base-point along this interval. For integer m, it is a a discrete flow by jumps.

6.5. If we truncate the infinite tensor product of section 3 at a finite level $A(n)$, the only real numbers we meet are rational and our formulae are algebraic. However, they are not in a closed system, so we get untidy approximations (also known as edge effects). But we know from the theory of von Neumann factors, that these approximations will, in the limit, yield the beautiful formulae of complex analysis, involving transcendental numbers like π. Moreover, there is a perfect duality in the limit as explained in 6.2.

6.6. There is a General Principle about mathematical models which should be borne in mind: C^∞ functions describe physics, Analytic functions describe mathematics. Physics has uncertainty (freewill), Mathematics has certainty (determinism). This is modulo Gödel, and logical difficulties in the foundations of mathematics, centering on the Axiom of Choice or proof by contradiction. But most physicists, with their feet on the ground, rightly do not care about such niceties.

6.7. Application of this Principle connects physical particles to their mathematical idealization. There is a mathematical parameter m, where m^2 is a model for mass, as in 6.1 above. If $m = 0$ we are on the light cone, both mathematically and physically (though we should distinguish between the light cone and its dual in which m is replaced by $1/m$). For $m = 0$, the theory is mathematically conformal and physically Diracian. Physically, small m represents a weak gravitational force, pervading the entire Universe. The number m is positive, reflecting the fact that gravitation is an attractive force. But as explained in section 7 of [8] changing the sign of m interchanges Newtonian gravitational attraction for Coulombic electrical repulsion.

6.8. If we only consider a finite number of decimal places, then we are in a truncated version of the algebras $A(n)$ as in 3.2, and the issue of C^∞ versus analytic is left open. From a rational approximation one cannot tell whether the subsequent digits, when found, will provide the decimal expansion of an analytic or just a differentiable function. This is why such subtleties can, as noted in 6.6, be ignored in practice.

6.9. In the truncated version, renormalization is only realized approximately, with untidiness at the edges. This untidiness disappears in the limit. The fine-structure constant is a mathematical idealization from the conformal world, but we can get arbitrary good approximations by using truncated versions. The number of terms needed in the truncation will depend on the number of terms required for the desired approximations.

7. Calculation of ж and Ч

The charisma of Bott

The key formula (1.1), relating ж and Ч, follows from the formula

$$2\text{Ч} = \lim_{n\to\infty} \sum_{j=1}^{j=n} 2^{-j}(1 - \int_{1/j}^{j} \log_2 x \; dx) \tag{7.1}$$

Since the terms in (7.1) are all positive and the sum is bounded by a convergent series, Ч is like γ the limit of a monotone increasing bounded sequence. For γ, the integral analogous to that in (7.1) is twice the integral from 1 to ∞, and the same is then true for the renormalization Ч of γ explaining the factor 2 in (7.1). (7.1) also shows why we could replace e by 2 and ln by $\log_2$. The proof of (7.1) follows from the mimicry principle of 7.6 below. To use (7.1) for computation, we need to specify the initial data, something which will be done in section 8. The numerical verification that ж agrees with $1/\alpha$ to all decimal places, so far calculated, follows from the numerics of section 8.

This comes in three steps, the first involving the sum and integral of the formulae (1.1) and (7.1) as with γ. But, as Euler discovered, the convergence in this process is too slow for effective computation.

So, the second step makes n jump by leaps of $2^4 = 16$ and speeds up the process. Looking from infinity, this comes from *slowing down*. This is a subtle point that will be explained in section 8. The third step is to calculate the first 16, which provide the base for the subsequent leaps of 16.

This third step is the one which leads to the number 137, as first proposed by Eddington [13]. There are now various ways of arriving at Eddington's number, all by pure algebra, which appear in several different papers [5] and [9]. The simplest is

$$137 = 1 + 8 + 128 = 2^0 + 2^3 + 2^7. \tag{7.2}$$

Eddington first proposed $136 = 8 + 128$, based correctly on Clifford algebras as in [4], but he had difficulty justifying the additional 1 to get 137. The reason he had this difficulty is that the group $SO(3)$ of rotations in $\mathbb{R}^3$ is not simply connected. In fact ж (our improvement on Eddington) is related, not to a particular orthogonal group, but to the stable orthogonal group $SO(N)$ for large N, It is not $\pi_1(SO(3)) = \mathbb{Z}/2$,which is relevant for ж, but the stable homotopy group $\pi_{1+k}(SO(3+k)) = 0$, for $k > 1$ and not congruent to 3 mod 4. The sequence of three powers of 2 that is needed to justify 137, can then be read off from the tables in [4], the algebra behind the Bott periodicity theorems [12]. The stable homotopy groups are periodic with period 8 or semi-periodic with half-period 4 if we switch orthogonal and symplectic.

Ironically, Eddington was later laughed out of court, when 137 was found to need a long string of corrections. In fact these corrections just arise from the iterative process that defines ж, so Eddington's two *mistakes* cancel each other out. Stability helps the tricky initial stage. A wobbly start acquires stability from subsequent motion in a very precise sense as on a bicycle.

This somewhat lengthy digression has fully restored the reputation of Eddington as asserted in this section above, showing that his *gut instinct* was essentially correct.

The Clifford algebra of a quadratic form Q differs subtly from the Clifford algebra of $-Q$, which is why we should take jumps of $2^4 = 16$, rather than $2^3 = 8$. This guarantees that our approximations to Ж are monotonically increasing and not oscillating. In terms of series this is the difference between a series of positive terms, such as $\sum \frac{1}{n^2}$, and a series of terms which alternate in signs, such as $\sum \frac{(-1)^n}{n}$.

The value of ж(n) that emerges from these calculations varies with the arithmetic mod 16 of the starting point. To eliminate this variation, we should start with 0 mod 16, so that we get unbroken blocks. Musicians, following Pythagoras, will notice the close analogy with octaves in both major and minor keys. Our starting point, to avoid dissonance, should be the key of C major. In 2.3, Hirzebruch's approach was explained as the arithmetic version of that of von Neumann. The algebra A(Q) being a refinement of the von Neuman algebra A. This clarifies the sections that follow.

7.3. Hirzebruch showed that a multiplicative sequence was determined by its value on line-bundles, which he called its generating function. Applied to A, as the limit of $A(n)$, this corresponds to restricting to the first factor $A(1)$.

7.4. As explained in 5.1, Hirzebruch showed that the Bernoulli function is this generating function. The exponential term indicates, as in 4.6, the role of the hyperbolic metric of curvature -1. It is better therefore to replace x by a multiple mx, so we can take the limit when m tends to zero (something Hirzebruch also did). This amounts to rescaling the weights of 4.3 and 4.4.

7.5. For reasons that will become apparent in 9.4, in the renormalized hyperfinite world of the algebra A, the quaternions $\mathbb{H}$ replace the complex numbers $\mathbb{C}$, and the unit sphere in $\mathbb{H}$ replaces the circle, ж replaces π and the central element w replaces i. In general, renormalization involves a choice of basis but, in the centre, the choice of basis is immaterial. That is why in 1.1 and 7.1, the computation of ж is related to the computation of the Euler-Hamilton constant Ч and mimics the relation between π and γ. We will turn this mimicry into a Principle in 7.6 below. Having chosen the weight $1/4$ in 4.6 we are then over the rationals $\mathbb{Q}$. Hirzebruch's work leads to analyticity which self- propagates: an analytic function of an analytic function is analytic and its derivatives are analytic. The theory of (hyperfinite) von Neumann algebras tells us that the choice of weight was irrelevant.

7.6. The Mimicry Principle then asserts that any analytic formula about real numbers implies the same analytic formula about their (hyperfinite) renormalizations. For example, renormalization preserves order, differentiation and ultimately analyticity.

7.7. Archimedes got π from an increasing bounded sequence. The mimicry principle gives us ж from a similar sequence as claimed in 2.4. The two formulae (1.1) and (7.1) now follow by mimicry from the corresponding formulae for π and γ. More precisely, each truncated version, for fixed n is analytic. The mimicry is then applied before we pass to the limit.

7.8. The next section on numerics will carry out the computations and check the answers at every level n. This provides a proof analogous to those of Archimedes and Euler and is sufficient for physicists. But, as mentioned in 6.6, mathematical logicians since Gödel have raised the bar for the notion of proof. It is probable that such computations will never fail, i.e a counterexample will never be found, but that it is not possible to prove this without additional axioms.

7.9. The logical issues raised in 7.8 appear in a different form in the size of the steps needed for the effective computation of further decimals in the expansion of ж. I asserted, for example, that 16 should be sufficient to produce 9 figure accuracy. There is no algorithm that will confirm this fact, though general theory will guarantee that some (unknown) large number would be sufficient. But someone doing the calculation is likely to stumble on a sufficient number, probably finding that 16 is enough. But 12 figure accuracy might well require 32. So, the logical qualms that hard-nosed physicists ignore at the conceptual level, come back to haunt them at the computational level. Gödel can only be ignored at your peril.

8. Double Limits

Cartwright and Littlewood

8.1 Archimedes, Euler and single limits. Archimedes defined and calculated π as a limit of increasing approximations $\pi(n)$. These have simple geometric interpretations in terms of isosceles triangles of angle $2\pi/n$. This was elegantly reformulated in terms of complex numbers by Euler through his famous formula based on taking the unit circle as

$e^{2\pi i\theta}$. The sequences $\pi(n)$ converge slowly but, as Euler discovered, they converge much faster if n is a power of 2.

8.2 Hamilton and ж. When $\mathbb{C}$ is replaced by $\mathbb{H}$ we have described a renormalization process which, by the principle of mimicry, replaces Euler's formula by the Euler-Hamilton formula in which π is renormalized to ж.

The Archimedes sequence $\pi(n)$ is now renormalized to a Hamilton sequence ж(n), but there is an important difference between the sequence of Archimedes and that of Hamilton. For the former, n is just a positive integer while, for the latter, n is a pair of coupled integers (n, p) with p odd and $|p| < n$, (but p can be negative) . Note that p will appear in an exponential $e^{\pi ip/n}$ which is a primitive $2n$-th root of unity. The single limit of Archimedes, where n tends to infinity, is now replaced by a double limit with both n and p tending to infinity.

The mimicry principle controls p, after which we just take the single limit over n. This gives an explicit sequence converging to ж. But again mimicry tells us that the convergence is very slow. We are free to choose any cofinal subsequence, so to speed things up we will take $n = 2^{s+1}$. We then put $j = 2^s - p$. We carry out the numerics in the next section.

8.3 The numerics. The explicit formula needed to calculate ж to arbitrary accuracy is (8.11) at the end of this section. The rest of 8.3 explains why (8.11) follows from the definition of ж in Section 2. Define sequences $v(j)$, $t(j) \in \mathbb{C}$, $k(j) \in \mathbb{Z}$, for $j \geq 0$ (all logs are taken to be base 2 in this section):

$$\log v(j+1) - v(j) = \log k(j) \tag{8.1}$$

$$t(j) = v(j+4) \tag{8.2}$$

$$\log k(j+1) = k(j), \quad k(j+1) = 2^{k(j)} \tag{8.3}$$

Use initial conditions $k(0) = 0$, $v(0) = i$, where i is chosen in preference to $-i$. This seems harmless and it is, if only made once as for $\mathbb{C}$. But we have to iterate it a large number of times (say n). This means we have 2^n choices or 2^{n-1} if we want to preserve signs (orientation). For $n = 1$, i.e. for $\mathbb{C}$ itself, this means the choice is unique. But for large n the choice is enormous and reflects the large number of orderings of n non-commuting variables. However inner automorphisms leave an invariant trace unaltered so the many choices have no effect on the traces which define ж. What is important is that each $v(j)$ is a primitive $2n$th root of unity.

For any choices, equations (8.1)-(8.3) lead to the formula

$$\log v(j) = v(j-1) + k(j-2) \quad \text{for } j < n \tag{8.4}$$

with a similar formula for $t(j)$. Next define the sequence of products ж$(n) = v(0)v(1)...v(j)$ and the two cofinal, sequences whose limits are independent of all choices:

$$\text{ж} = \lim_{j\to\infty} \text{ж}(j) \tag{8.5}$$

$$ж = \lim_{j \to \infty} ж(j+4) \tag{8.6}$$

Formulae (8.1)-(8.3) show that the inductive step from $j-1$ to j is exponential, so that the sum in (8.4) becomes a product at the next step. The initial conditions of (8.3) seem to run into trouble when we put $j = 1$, since $j - 2$ is then negative and formally not defined. But since we are interested in the limits in (8.5) and (8.6) we can ignore the first term in (8.5), and (8.6) is well into the "stable range". The formulae (8.1)-(8.3) define the Todd sequence in the Hirzebruch formalism based on the weight 4. As we have seen, Hirzebruch's formalism extends to the weak closure of the von Neumann factor. This leads to (8.4) defining the real number ж as the limit of finite products ж(j). Formula (8.6) shifts by $2^4 = 16$ and leads to rapid convergence and effective computation (but see 7.10 for a caveat).

We now come to the subtle point alluded to after (7.1). For the sequences (8.5) and (8.6) to converge, the iteration has to be made at the slowest possible speed t, so that the product expansion

$$\prod_{i=1}^{n} \{1 + ta(i)\} \tag{8.7}$$

is arbitrarily close to the sum

$$1 + t \sum_{i=1}^{n} a(i) \tag{8.8}$$

up to terms of a fixed degree t^m, provided t (viewed as time) is sufficiently small. Focusing on the limit as $n \to \infty$ of (8.7) or (8.8) involves the inverse or dual point of view expressed in 6.2 and interchanges fast and slow.

A good analogy is provided by a scientific rocket being sent say to Mars. It has to be launched into space with a high velocity but, to ensure a soft landing on Mars, it has to descend very slowly. Readers who prefer to avoid metaphors or analogies can consult Hirzebruch [1], where he evaluates a formal power series of cohomology classes on a complex n-manifold, and notes that all terms above dimension n give zero in a stable manner. This is precisely similar to the way the product (8.7) is equivalent to the sum (8.8) . The only difference is that for Hirzebruch the parameter t is a 2-dimensional cohomology class, whereas in (8.7) and (8.8), t is interpreted as time. This was typical of Hirzebruch's magic. This skillful interpretation of parameters, with t becoming time, not only smells of Relativity, but also acquires a computer reality when we compare the speed of algorithms. Our fast algorithm, following Euler, means that the algorithm in question gives the answer on your computer in a much smaller number of steps.

Note that the limits in (8.5) and (8.6) appear to involve only the positive integers, but careful examination of the integral term in (7.1) shows that it involves both j and $1/j$, a multiplicative symmetry that converts to the additive symmetry between n and $-n$.

The Eddington number of 137 simply comes from the first few steps. The emergence of the Eddington number has been discussed above.

For rockets to Mars in the previous paragraphs, we have to think of rockets that plan to return to Earth, with the outward and return journeys being symmetrical. The very soft landing is obviously needed to avoid damage to the rocket!

We can describe what we are doing in the following way. Given any number 2^n, we can factor it as a product of two numbers $2^{n(0)}2^{n(1)}$ where $n = n(0) + n(1)$. As n gets larger, we keep $n(0)$ fixed, say $n(0) = 4$, and let $n(1)$ get larger. This describes our chosen algorithm and explains the shift by 4 with $t(n) = v(n + 4)$. This will give the correct 12 digits. When we increase n, to improve on the approximations ж(n) we will have to increase $n(0)$ and $n(1)$, but we cannot be sure of their optimal values. However, since our sequences are monotonic increasing, we can adopt the stopping rule : stop one step before the product (8.7) exceeds the sum (8.8). This can be formalized in terms of the Bernoulli numbers B_k^n of higher order which, as explained below, are essentially Hirzebruch's Todd polynomials.

The notation automatically chooses the initial conditions for the blocks of 16 described in section 1. Starting from the Eddington number 137, explicit calculations as indicated above and explained below now give the value

$$\text{ж} = 137.035999... \tag{8.9}$$

This agrees with current values and more calculations will predict subsequent decimals, confirming the expectations of 2.4.

The double limit process was elegantly treated by Hirzebruch, following the earlier work of Norlund on Bernoulli polynomials,see section 1.8 of [1]. These are defined by the Hirzebruch mutiplicative sequence over a polynomial ring $Q[y]$, whose generating function is

$$Q(y, x) = x + \frac{y + 1}{e^{x(y+1)} - 1}.$$

The Bernoulli polynomials of higher order B_k^n, defined by Norlund in 1924, are essentially the Todd polynomials related by

$$T_k(c_1, .., c_k) = \frac{(-1)^k}{k!} B_k^n(\gamma_1, \ldots, \gamma_n), \quad \text{for } k \leq n \tag{8.10}$$

where the c_i are regarded as the elementary symmetric functions in $\gamma_1, ., \gamma_n$. If we put all $\gamma_i = 1$, so that c_i is the binomial coefficient $\binom{n+1}{i}$, we get the Bernoulli numbers of higher order B_k^n: for $k = 1$, we recover the usual Bernoulli numbers.

We take j and n to be as in 8.2, so that $e^{\pi ij/n}$ is a primitive $2n$-th root of unity. Our explicit limit formula for ж is now the double limit

$$\text{ж} = \lim_{\substack{n\to\infty \\ j\to\infty}} 2^{-2n} B_{k(j)}^n \tag{8.11}$$

recall from (8.3) that $k(j+1) = 2^{k(j)}$ for $j > 0$ and, from (8.2), that $e^{\pi ik/n}$ is a primitive $2n$th root of unity.

For each n, the limit over j is independent of the choices of primitive root, and this will be seen in the explicit formula below. To get accuracy to 12 decimal places, we start anew with $n(0) = 5$ and $ж(1) = 137.035$ as starting point. This will give the required accuracy with very high probability. The uncertainty is not in the final value of ж , but in the rate of convergence of the approximation $ж(n+5)$.

The verification that the rules, given in the first part of this section, yield the value of ж given in (2.5) to 9 decimal places, follows by mimicry from the corresponding formula for π.

The same will be true for 12 decimal places, and the last 3 digits will emerge from mimicry of a more accurate approximation to π. It remains as an exercise to write these down and, as mentioned, there are uncertainties about the rate of convergence.

9. Further Comments

The natural philosophy of Maxwell

Having explained the fine-structure constant α and its inverse $1/\alpha$ in mathematical terms, I will now indicate many other ways in which similar ideas have been used in my publications. I will then move on, and explain various physical contexts in which these mathematical ideas seem useful. While such agreement between mathematics and physics can never be proved, it can become increasingly plausible. Maxwell clearly enunciated this meta-physical attitude 150 years ago [11].

9.1. The iteration required to move from Type I to Type II factors in section 2 is very closely related to the iteration in my proof of the Feit- Thompson Theorem [6]. We should now distinguish between geometry and arithmetic (as will be explained in section 10). In [2] the iterative construction was geometric, the ground field being fixed as $\mathbb{Q}(\rho)$. For a finite group this iteration stops after a finite number of steps, but for suitable infinite discrete groups, the iteration need not stop and could lead to the Type II factor A. Such matters are related to Burnside type issues and they will be discussed in my forthcoming paper [10] with A.Zuk.

The arithmetic version of the algebraic structure in [6] built from the group of order 1 is an infinite tower of extensions of $\mathbb{Q}(\rho)$ leading to Ж.

9.2. More general Type II factors can arise from an ergodic group action on a finite measure space as realized by von Neumann.

9.3. As mentioned in 2.6, Newton's constant G will, in a forthcoming paper [7], be treated in a similar fashion to α, but involving a Type III factor, based on the octonions rather than the quaternions.

Although motivated by physical constants (α and G), the comments 9.1, 9.2 and 9.3 were mathematical. I now turn to physical interpretations of the mathematical models. Such interpretations provide a tentative dictionary between mathematics and physics. Any dictionary is constantly being modified and never bridges the entire gap between the two languages and cultures. That is why it is called meta-physics, a topic I will pursue in subsequent papers. So, with this preamble, here is an initial description of the dictionary.

9.4. The 4 real division algebras, the reals $\mathbb{R}$, the complex numbers $\mathbb{C}$, the quaternions $\mathbb{H}$ and the octonions $\mathbb{O}$ translate into the 4 basic forces. The electro-weak forces come from $\mathbb{R}$ and $\mathbb{C}$, while the strong force comes from $\mathbb{H}$, and gravity comes from $\mathbb{O}$. The dictionary, for the first 3, can be expressed in terms of the ascending sequence of compact groups $U(1) \subset U(2) \subset U(3)$. The 4-th term of the sequence, arising from $\mathbb{O}$, cannot be a group, because $\mathbb{O}$ is not associative. One option is to use the group of fractions $f(t)/g(t)$, where t is a formal symbol and f and g are power series over $\mathbb{C}$ with constant term 1.

9.5. The last group in 9.4 is not finite-dimensional, but it can be approximated by truncating the power series to polynomials. These approximations are finite-dimensional but are not groups. The untidy effects of the truncation were explained in 6.9.

9.6. Since 7.6 was used to understand α, and since 9.5 arose from the octonions, this indicates that gravity is playing a role, which seems surprising since α is related only to electro-magnetism. The explanation is that, while not necessary for the definition of ж in section 2, the best way to compute α is to add parameters and then study carefully coupled limits as they tend to zero, as done in section 7 and pioneered by Gauss. Although this is a purely mathematical process, its physical interpretation suggests that experimental calculations to determine approximations to α should be made in a weak gravitational field. But this is precisely what a lab is. There is no way entirely to avoid the gravitational force of other matter.

9.7. In 6.8, I distinguished between differentiable functions and analytic functions, explaining that the latter were not physical. In [8] the Atiyah-Sutcliffe determinants were analytic, so that their real and imaginary parts determined each other. This translates back to saying that energy and entropy determine each other: true in the model but not in the physical world. However, looking at any finite part of the decimal expansion of a number does not tell you whether you are in a differentiable or an analytic situation. This explains why all the fundamental equations of physics, including Maxwell's equations and its non-linear generalizations, are analytic with the possible exception of Einstein's equation of General Relativity.

9.8. Notice that, while we kept e fixed and renormalized π, we could have renormalized both, with the coupling provided by the Euler- Hamilton equation. The renormalized e can be written as e^{ϵ} and the Euler-Hamilton equation becomes $e^{\epsilon Ж w} = 1$. The coupling is

$\pi\epsilon = t/u$, so that $\pi\epsilon/k = t/ku$ for any integer k (where k, t, u are as in section 8).This fully answers Feynman's question. Both e and π are involved, but we have chosen to keep e and just renormalize π. More symmetrically, α is modelled by the ratio of the homogeneous variables $\epsilon/$ж. In affine coordinates, where we put $\epsilon = 1$, we recover our previous formula of $1/$ж.

9.9. As Euler discovered, π is ubiquitous in commutative mathematics. Based on the mimicry principle, I therefore expect ж to be ubiquitous in non- commutative (but associative) mathematics. This has implications for combinatorics with, for example, $n!$ replaced by $p(n)$ the number of partitions of n. I plan to investigate such matters in a subsequent paper.

9.10. The 4 basic forces modeled in 9.4 are those of the standard model, and α is studied in that context. However, our approach can easily be extended beyond the standard model and there would then be corrections to follow. The latest experimental data [14] points in that direction and has yet to be explored.

9.11. The Hirzebruch formalism, applied to the Riemann-Roch Theorem and, more generally, to the index theorem,leads to a Type II index Theorem. Its physical interpretation is a formula for the Witten index $\mathrm{Tr}(-1)^F$. This will be explained elsewhere.

10. Conclusion

I would just like to understand the electron. - Albert Einstein

The interpretation of the fine structure constant alpha that I have offered, opens the door to a new view of physics, which I will now try to describe in broad outline.

In essence von Neumann provided a framework for physics (both classical and quantum) based on the real number field $\mathbb{R}$. This enabled geometric foundations for physics to be laid. In parallel Hirzebruch provided a framework based on the rational number field $\mathbb{Q}$, leading to arithmetic foundations. Together von Neumann and Hirzebruch laid the foundations of the subject of Arithmetic Physics.

Thirty years ago in [16] I explained Manin's vision about a classical bridge between arithmetic and physics. I then went on to speculate that there should be a quantum version of Manin's bridge stretching from quantum field theory to Langlands . I now believe that this can be built on the mimicry principle that has emerged from the study of alpha [17].

So Einstein was right. Understanding the quantum electron requires understanding alpha. This then opens the door to all else.

But there are limitations to human understanding, so we have to settle for what we can understand, and here we need the collective wisdom from all branches of science.

This paper can be read at different levels by different scientists.Experimental physicists will be satisfied by several more decimal places for α, correctly predicted. Theoretical physicists may want formal Platonic explanation, as Feynman and Good required. Mathematicians would like to see rigorous proofs, as provided by von Neumann algebras. Logicians may question the foundations related to the Axiom of Choice. But, to plagiarise Abraham Lincoln: *one can never please all the people all the time.*

11. Acknowledgements

I am grateful to my many physics friends for educating me. But I owe special thanks to Gerard 't Hooft, who shared his critical insight with me in Singapore in January, and challenged me to explain α. My inspiration was Euler, whose ideas have been expounded with great clarity by William Dunham [15]. I am also indebted to the Leverhulme Trust and Trinity College, Cambridge for their support.

References

[1] F.Hirzebruch *Topological methods in algebraic geometry (with appendices by R.L.E.Schwarzenberger, and A.Borel).* Springer 1966
[2] A.Connes *Non-Commutative Geometry.* Academic Press 1994.
[3] H.Freudenthal *Lie groups in the foundations of geometry.* Advances in Math. 1 (1964), 145-190.
[4] M.F.Atiyah, R.Bott and A.Shapiro *Clifford Modules.* Topology 3 (1964) 3-38.
[5] M.F.Atiyah *Characters and cohomology of finite groups.* Publ.Math. IHES 1961, 247-89.
[6] M.F.Atiyah *Groups of Odd Order.* (submitted to Journal of Geometry and Physics).
[7] M.F.Atiyah *Newton's Constant G.* (forthcoming).
[8] M.F.Atiyah and J.Malkoun *The Relativistic Geometry and Dynamics of Electrons.* Foundations of Physics 2018.
[9] M.F.Atiyah and J.Berndt *The Kervaire invariant and the magic square I.* (submitted to Journal of Topology).
[10] M.F. Atiyah and A.Zuk *In preparation.*
[11] P.M.Harman *The Natural Philosophy of James Clerk Maxwell.* CUP 1998.
[12] R.Bott *The Stable Homotopy of Classical groups.* Proc.Nat.Acad.Sci.USA 48 (1957) 933-935.
[13] A.S.Eddington *Fundamental Theory.* CUP 1948.
[14] R.H.Parker, C.Yu, W. Zhong, B.Estey, H.Muller *Measurement of the fine structure constant as a test of the standard model.* Science 360, 191 -195, 2018
[15] W.Duncan *Euler, The Master of us all.* the Mathematical Association of America 1999
[16] M.F.Atiyah *Commentary on Manin's Manuscript New Dimensions in Geometry.* Lecture notes in mathematics, 1111 (Proc.of 25th Mathematics Arbeitstagung, Bonn 1984), Springer- Verlag (1985), 103-9
[17] M.F.Atiyah *Arithmetic Physics.* Proceedings of ICM Rio de Janeiro 2018