



Conference Proceedings Paper – Entropy

The Entropy Conundrum: A Solution Proposal

Rodolfo A. Fiorini^{1,*}

¹ Politecnico di Milano, Department of Electronics, Information and Bioengineering, Milano, Italy; E-Mail: rodolfo.fiorini@polimi.it

* E-Mail: rodolfo.fiorini@polimi.it;
Tel.: +039-02-2399-3350; Fax: +039-02-2399-3360.

Received: 11 September 2014 / Accepted: 11 October 2014 / Published: 3 November 2014

Abstract: In 2004, physicist Mark Newman, along with biologist Michael Lachmann and computer scientist Cristopher Moore, showed that if electromagnetic radiation is used as a transmission medium, the most information-efficient format for a given 1-D signal is indistinguishable from blackbody radiation. Since many natural processes maximize the Gibbs-Boltzmann entropy, they should give rise to spectra indistinguishable from optimally efficient transmission. In 2008, computer scientist C.S. Calude and physicist K. Svozil proved that "Quantum Randomness" is not Turing computable. In 2013, academic scientist R.A. Fiorini confirmed Newman, Lachmann and Moore's result, creating analogous example for 2-D signal (image), as an application of CICT in pattern recognition and image analysis. Paradoxically if you don't know the code used for the message you can't tell the difference between an information-rich message and a random jumble of letters. This is an entropy conundrum to solve. Even the most sophisticated instrumentation system is completely unable to reliably discriminate so called "random noise" from any combinatorially optimized encoded message, which CICT called "deterministic noise". Entropy fundamental concept crosses so many scientific and research areas, but, unfortunately, even across so many different disciplines, scientists have not yet worked out a definitive solution to the fundamental problem of the logical relationship between human experience and knowledge extraction. So, both classic concept of entropy and system random noise should be revisited deeply at theoretical and operational level. A convenient CICT solution proposal will be presented.

Keywords: entropy; statistical mechanics; quantum-classical transition, decoherence, combinatorial optimization; reversibility; quantum mechanics.

PACS Codes: 03.67.-a; 03.67.Ac; 03.67.Bg; 03.67.Hk; 03.67.Pp.

1. Introduction

In 2004, University of Michigan physicist Mark Newman, along with biologist Michael Lachmann and computer scientist Christopher Moore, has extended the pioneering 1940s research of Claude Shannon on communication theory to electromagnetic transmission. They showed that if electromagnetic radiation is used as a transmission medium, the most information-efficient format for a given 1-D signal is indistinguishable from blackbody radiation [1]. Since many natural processes maximize the Gibbs-Boltzmann entropy, they should give rise to spectra indistinguishable from optimally efficient transmission. In 2008, computer scientist C.S. Calude and physicist K. Svozil proved that "Quantum Randomness" (QR) is not Turing computable. While "true randomness" is a mathematical impossibility, the certification by value indefiniteness ensures that the quantum random bits are incomputable in the strongest sense [2,3]. Algorithmic random sequences are incomputable, but the converse implication is false. In 2013, Politecnico di Milano academic scientist R.A. Fiorini confirmed Newman, Lachmann and Moore's result, creating analogous example for 2-D signal (image), as an application of computational information conservation theory (CICT) to pattern recognition and image analysis for biomedical application [4]. Specifically, CICT showed that classical Shannon entropy computation is completely unable to reliably discriminate so called computational "random noise" from any combinatorially optimized encoded message by OECS, called "deterministic noise" (DN) in [4].

Paradoxically if you don't know the code used for the message you can't tell the difference between an information-rich message and a random jumble of letters. This is "the information double-bind" (IDB) problem in current classic information and transmission theory [4].

Shannon's entropy (usually denoted by $H(x)$ or $Sh(x)$) is used as a fast stochastic measure of probabilistic information uncertainty in every information processing system. $H(x)$ is the average unpredictability in a random variable, which is equivalent to its information content. The noise discrimination problem is still faced by the single domain channel transfer function concept (Shannon's noisy channel), starting from classic Shannon's information theory concept [5], and then applying traditional stochastic perturbation computational model under either additive or multiplicative perturbation hypothesis [4].

Even the most advanced and sophisticated contemporary instrumentation system like "Soil Moisture Active Passive" (SMAP) satellite is using it. SMAP is designed to measure the intensity of electromagnetic radiation and equipped with an onboard computing system capable of crunching data equivalent to 192 million samples per second [6]. The new radiometer is expected to be integrated into NASA's SMAP spacecraft, along with a synthetic aperture radar system developed by JPL, and launched in 2014 finally [6]. The classical experimental observation process, even in highly ideal operative controlled condition, like the one achieved in contemporary most sophisticated and advanced experimental laboratories like CERN, can capture just a small fraction only, with misplaced precision, of overall ideally available information from unique experiment [7]. The remaining part is lost and inevitably added to something we call "background noise" or "random noise" usually, in every scientific experimental endeavor. Information and communication theory was introduced by Claude E. Shannon in his seminal 1948 paper "A Mathematical Theory of Communication" [5]. Shannon's entropy provides an absolute limit on the best possible lossless encoding or compression of any communication, assuming that the communication may be represented as a sequence of independent and identically distributed random variables (and this is a strong assumption based on a "fixed universe" model). Mathematically, a sequence of coin flips (fair or not) is an example of a Bernoulli process, and its entropy is given by the binary of coin flips (fair or not) is an example of a Bernoulli process, and its entropy is given by the binary entropy function [5]. The interested

reader in digging deeper details into mathematical theory of entropy and information theory, inference, and learning algorithms, is referred to [8] and [9] respectively. As a matter of fact, biologists measure information in different ways. Neurobiologists and researchers in bioinformatics often measure information using information-theoretic measures such as Shannon's entropy or algorithmic compression on mutual information. Behavioral biologists and evolutionary ecologists more commonly use decision-theoretic measures, such the value of information, which assess the worth of information to a decision maker. It can be shown that these two kinds of measures are intimately related in the context of biological evolution research areas.[10] Therefore, the classic concept of information theory has spread across different disciplines. In communication theory, the transmission of information is the reduction of uncertainty about what signals will come through a channel, from an initial set of all possible signals down to the post-hoc set of signals actually received. In thermodynamics, a decrease in entropy refers to the fold reduction in the number of states that a system can be in. In evolutionary biology, the fitness value of a cue about an uncertain environment refers to the fold increase in the number of surviving lineages made possible by responding to the cue [11]. In quantum information theory, quantum relative entropy is a measure of distinguishability between two quantum states. It is the quantum mechanical analog of relative entropy. But, computing quantum relative entropy is a NP-complete problem [12]. Entropy fundamental concept crosses so many scientific and research areas, but, unfortunately, even across so many different disciplines, scientists have not yet worked out a definitive solution to the fundamental problem of the logical relationship between human experience and knowledge extraction. Epistemic uncertainty sources are still treated by risk analysis approach only, which provides an acceptable cost/benefit ratio to producer/manufacturer, but in some cases it may not represent an optimal solution to end user. In fact, deep epistemic limitations reside in some parts of the areas covered in decision making. More generally, decision theory, based on a "fixed universe" or a model of possible outcomes, ignores and minimizes the effect of events that are "outside model".

Statistical domains assume "Small World" (SW) \equiv coin tosses and "Large World" (LW) \equiv Real World. Note that "Measure Theory" is not SW, but LW, thanks to the degrees of freedom it confers. The real world is about incompleteness: incompleteness of understanding, representation, information, etc., what one does when one does not know what's going on, or when there is a non-zero chance of not knowing what's going on. It is based on focus on the unknown, not the production of mathematical certainties based on weak assumptions. Rather it is much better to measure the robustness of the exposure to the unknown, which can be done mathematically through metamodel (a model that examines the effectiveness and reliability of the model), what we call metaprobability, even if the meta-approach to the model is not strictly probabilistic. Nevertheless, under the cover of the discipline of "Extreme Value Theory", even tail events are very opaque computationally and misplaced precision leads to confusion [13].

So, patterns in "natural systems" abound, from the stripes on a zebra to ripples in a river bed. In these systems, the emergent appearance of an ordered state is not unexpected as the outcome of an underlying hidden ordered and folded process. On the other hand, stochastically computed "noisy and disordered" processes such as thermal fluctuations or mechanically randomized scattering generate surprisingly similar patterns in long computer simulations. Since its conception, scientific community has been laying itself in this double-bind situation. Even the most sophisticated instrumentation system is completely unable to reliably discriminate so called "random noise" (RN) from any combinatorially optimized encoded message, called DN [4]. This is an entropic conundrum to solve. So, both classic concept of entropy and system RN should be revisited deeply at

theoretical and operational level. More broadly, researchers need to realize the limits of conventional statistics. They should instead bring into their analysis elements of scientific judgement about the plausibility of a hypothesis and study limitations that are normally banished to the traditional discussion section. In the following section we go to the root of the problem.

2. The Root of the Problem

To understand what the quantum-classical transition really means, consider that our familiar, classical world is an "either/or" kind of place. A compass needle cannot point both north and south at the same time. The quantum world, by contrast, is "both/and" and a magnetic atom model has no trouble at pointing both directions at once. The same is true for other properties such as energy, speed, etc. Generally speaking, they can take on a range of values simultaneously, so that all you can say is that this value has that probability. When that is the case, physicists say that a quantum object is in a "superposition" of states. Thus, one of the key questions in understanding the quantum-classical transition is what happens to the superpositions as you go up that atoms-to-apple scale. Exactly when and how does "both/and" become "either/or"? Over the years, physicists have proposed many answers to that question. But the most favoured one involves a phenomenon known as "decoherence", which was identified and elucidated in the 1970s and 1980s [14]. In simple terms, decoherence is a sort of leaking away of quantum behaviour when a particle interacts with its surroundings. For example, when an atom or molecule collides with those around it, or when light bounces off it. All we are left with is a partial picture of the system, in which only a well-defined subset of macroscopic properties, such as position, are apparent. In effect, different quantum states have very different resistances to decoherence. So only the resistant states will survive when a system interacts with its environment. These robust states are those that feature in classical physics, such as position and its rate of change, which is associated with momentum. In a sense, these are the "fittest" states, which is why Los Alamos scientist Wojciech Zurek and his colleagues call their idea "quantum darwinism" [15,16]. Decoherence also predicts that the quantum-classical transition is not really a matter of size, but of time. The stronger a quantum object's interactions are with its surroundings, the faster decoherence kicks in. So larger objects, which generally have more ways of interacting, decohere almost instantaneously, transforming their quantum character into classical behaviour just as quickly. For instance, if a large molecule could be prepared in a superposition of two positions just 10 ångströms apart, it would decohere because of collisions with the surrounding air molecules in about 10^{-17} seconds. Decoherence is unavoidable to some degree. Even in a perfect vacuum, particles will decohere through interactions with photons in the omnipresent cosmic microwave background. Decoherence offers a way to understand classicality as emergent from within the quantum formalism, and the classical world no longer sits in opposition to quantum mechanics, but is demanded by it.

But is decoherence real? Can we actually see quantum effects leaking away into the environment? The decoherence description shows that there is no abrupt boundary, no critical size, at which quantum behaviour switches to classical. And the blurry boundary itself shifts depending on how it is measured. It is the choice of the measuring apparatus that defines whether a specific object can be represented by a quantum or classical model [17]. The same object can behave as a quantum system in one situation, and as a classical system in another. Extremely high-precision measurements which allow to measure positions to within 10^{-13} metres, close to the width of a proton, are nevertheless potentially feasible by interferometric methods. The important point is that the conditions for an observation of interference effects can be precisely specified and quantified using

theoretical decoherence models. We no longer have to limit ourselves to the assumption of a poorly defined yet fundamental divide between the quantum and the classical realms. But the separation between energy levels is very small, so quantum behaviour will be blurred unless the observed system is kept very cold to eliminate thermal noise. For instance, at temperatures of around 25 millikelvin, researchers at Cornell University in Ithaca, New York, have managed to restrict the vibrations of nanoelectromechanical system (NEMS) [18] to only the 25 lowest energy levels. But that it is not enough. All the remaining excited states have to be "sucked out" by active cooling, akin to the laser cooling used to draw heat from small clusters of ultracold atoms [19,20]. But, as soon as you look for the superposition, you destroy it and the superposition can only be inferred from its effect on another system to which the resonator is coupled. This type of coupling will have a specific signature as coherence will reappear periodically at a rate equal to the vibration period of the beam. This would reveal quantum effects in a system containing around ten billion atoms, much more than is feasible with molecular interference experiments.

The entropy of a quantum system can be considered as a vector of the superposition of many different entropies whose values are conditioned by the observer. Any observer makes the system in superposition collapse to a classical value of entropy, i.e. it fixes classical information by a measurement. When the observer does not interact with a quantum system, the composed system given by the quantum microsystem + observer (and thus, somewhat the universe) is in an entangled state. When the observer interacts with the microsystem under consideration thus becoming an active element, the composed system given by the quantum microsystem + observer is in a pure state and follows the classical thermodynamics defined by one specific entropy and, in correspondence, the measured variable of the system under consideration takes the measured value. Quite recently, theoretical Physicist I. Licata et al. have presented an interesting Fisher-Bohm geometric point of view for quantum information [21]. Given the relation between the number of the system microstates W , function of the macrostates and given the entropies S as Cartesian coordinates, the reference can be changed and from Euclidean coordinates of the entropies, in which the observers are independent, one can move to a non Euclidean space of the parameters (averages, variances ...). In this background it is possible to compute covariant derivatives in the parameter space and we can obtain by entropy the Bohm quantum potential and the quantum effects [21]. The approach suggested in this background leads to the idea that, in the extreme condition of Fisher information, Bohm's quantum potential emerges as an information channel determined by informational lines associated with the vector of the superposed Boltzmann entropies. This approach has the merit that, in a relativistic curved space-time, the informational lines associated with the quantum entropy appear as real intermediaries between gravitational and quantum effects of matter, determining a high coupling between the effects of gravity on geometry and the quantum effects on the geometry of space-time. To gain a better understanding about the operational compromise offered by the information geometry theory (IGT) approach, in the following section we offer an IGT and Non-Euclidean IGT Manifolds synthetic overview as a propedeutic step to section 4.

3. Information Geometry Theory

In 1945, by considering the space of probability distributions, Indian-born mathematician and statistician Calyampudi Radhakrishna Rao (1920-) suggested the differential geometric approach to statistical inference. He used Fisher information matrix in defining the metric, so it was called Fisher-Rao metric. In 1975, American statistician Bradley Efron (1938-) carried the argument a step forward

when he introduced a new affine connection on the parameter space manifold, and thus shed light on the role of the embedding curvature of the statistical model in the relevant space of probability distributions. So, information geometry theory (IGT) emerged from the study of the geometrical structure of a manifold of probability distributions under the criterion of invariance. IGT reached maturity through the work of Shun'ichi Amari (1936-) and other Japanese mathematicians in the 1980s. Amari and Nagaoka's book [22] is cited by most works of the relatively young field due to its broad coverage of significant developments attained using the methods of information geometry up to the year 2000. Many of these developments were previously available in Japanese-language publications only. IGT takes probability distributions for a statistical model as the points of a Riemannian manifold, forming a statistical manifold. The Fisher information metric provides the Riemannian metric. Moreover, a family of dually coupled affine connections are introduced. Mathematically, this is a study of a triple $\{M, g, T\}$, where M is a manifold, g is a Riemannian metric, and T is a third-order symmetric tensor. It largely focuses on typically multivariate, invariant and higher-order asymptotic results in full and curved exponential families, through the use of differential geometry and tensor analysis. Many important families of probability distributions are dually flat Riemannian manifolds. A dually flat manifold possesses a beautiful structure: it has two mutually coupled flat affine connections and two convex functions connected by the Legendre Transformation. It has a canonical divergence, from which all the geometrical structure is derived.

During the last 15 years (1999-2014), IGT has been applied not only to statistical inferences but also to various fields of information sciences where probability plays an important role. The Kullback-Leibler divergence [23-25] for short KLD, in probability distributions is automatically derived from the invariant flat nature. KLD (also called information divergence, information gain, relative entropy) is a non-symmetric measure of the difference between two probability distributions P and Q . Specifically, the KLD of Q from P , denoted $KLD(P||Q)$, is a measure of the information lost when Q is used to approximate P [26]. KLD divergence measures the expected number of extra bits required to code samples from P when using a code based on Q , rather than using a code based on P . Typically P represents the "true" distribution of data, observations, or a precisely calculated theoretical distribution. The measure Q typically represents a theory, model, description, or approximation of P . KLD is a special case of a broader class of divergences called "f-divergences." It was originally introduced by American cryptanalyst and mathematician Solomon Kullback (1907 – 1994) and Richard Leibler (1914-2003) in 1951 as the directed divergence between two distributions. It can be derived from a Bregman divergence (BD) [27], introduced by Russian mathematician Lev M. Bregman (1941-) in 1967. Moreover, the generalized Pythagorean and geodesic projection theorems hold. Conversely, we can define a dually flat Riemannian structure from a convex function. This is derived through the Legendre Transformation and BD connected with a convex function. Therefore, IGT is applicable to convex analysis, even when it is not connected with probability distributions. This widens the applicability of IGT to convex analysis, machine learning, computer vision, Tsallis entropy, economics, game theory, etc. The Fisher-information matrix is used to calculate the covariance matrices associated with maximum-likelihood estimator (MLE). It can also be used in the formulation of test statistics, such as the Wald test. MLE gives a unified approach to estimation, which is well-defined in the case of the normal distribution and many other problems.

However, in some complicated problems, tensorial formulation may become quite cumbersome and computational difficulties do occur: in such problems MLEs are unsuitable or do not exist. In fact, sometime the "probabilistic veil" can be very opaque computationally, and misplaced precision leads to information opacity, fuzzyness, irreversibility, chaos, complexity and confusion.

3.1. Non-Euclidean IGT Manifolds and Riemann Surfaces

Riemann surfaces are named after the 19th century mathematician Bernhard Riemann, who was the first to understand the importance of abstract surfaces, as opposed to surfaces arising concretely in some ambient space. Compact Riemann surfaces are geometric structures constructed by cleverly stretching, bending, and gluing together parts of the complex number plane. A surface becomes a Riemann surface when it is endowed with an additional geometric structure. One can think of this geometric structure as a so-called complex structure, which allows one to do complex analysis on the abstract surface. On a hyperbolic Riemann surface, we don't have straight lines, given the curvature of the surface. Instead, geometers consider "geodesics" when working in a non-Euclidian space. Geodesics are the natural analogue of lines: just as a line segment in a plane is the shortest path between two points, a geodesic on a Riemann surface is the curve with the shortest hyperbolic length between two points on the surface. Because of the compact closed nature of surfaces, geodesics can sometimes loop around the surface and come back to their starting point, and these are called closed geodesics. Closed geodesics that don't intersect themselves while tracing their path around the surface are called simple closed geodesics.

The interesting hyperbolic Riemann surfaces all have a broad global structure that looks like some number of toruses or donuts glued together. It's possible to take two surfaces made out of the same number of donuts and deform by pulling or stretching on one surface to get the other surface, and so these surfaces are in some way related to each other. The number of donuts a surface is made out of is called the genus g of the surface. While the underlying topological surface remains the same, its geometric shape changes during a deformation. Riemann knew that these deformations depend on " $6g-6$ " parameters or "moduli", meaning that the "moduli space" (MS) of Riemann surfaces of genus g has dimension " $6g-6$." However, this says nothing about the global structure of moduli space, which is extremely complicated and still very mysterious. For instance, when one considers only the closed geodesics that are simple, meaning that they do not intersect themselves, the growth of the number of geodesics of length at most L is no longer exponential in L but is of the order of " L^{6g-6} ." Recently, Iranian mathematician Maryam Mirzakhani et al. showed that in fact the number is asymptotic to " $c L^{6g-6}$ " for large L (going to infinity), where the constant " c " depends on the hyperbolic structure [28,29].

The theory of Riemann surfaces and algebraic geometry are closely linked. Every complex curve is an algebraic curve, meaning that the complex curve, although defined abstractly, can be realized as a curve in a standard ambient space, in which it is the zero set of suitably chosen polynomials. Thus, although a Riemann surface is a priori an analytic object defined in terms of complex analysis on abstract surfaces, it turns out to have an algebraic description in terms of polynomial equations. An alternative but equivalent way of defining a Riemann surface is through the introduction of a geometry that allows one to measure angles, lengths, and areas. The most important such geometry is hyperbolic

geometry (HG), the original example of a non-Euclidean geometry discovered by Bolyai, Gauss, and Lobatchevski. The equivalence between complex algebraic and hyperbolic structures on surfaces is at the root of the rich theory of Riemann surfaces. Sometimes properties of a fixed hyperbolic surface can be better understood by studying the MS that parametrizes all hyperbolic structures on a given topological surface. For example, thinking of Riemann surfaces as algebraic curves leads to the conclusion that MS itself is an algebraic object called an algebraic variety. In addition, MS has a metric whose geodesics are natural to study. These MSs have rich geometries themselves, and arise in natural and important ways in differential, hyperbolic, and algebraic geometry. There are also connections with theoretical physics, topology, and combinatorics. The behavior of geodesics in moduli space is even related to dynamical system on MS. Non-closed geodesics in MS are very erratic and even pathological, and it is hard to obtain any understanding of their structure and how they change when perturbed slightly. However, complex geodesics and their closures in MS are in fact surprisingly regular, rather than irregular or fractal [29]. It turns out that, while complex geodesics are transcendental objects defined in terms of analysis and differential geometry, their closures are algebraic objects defined in terms of polynomials and therefore have certain rigidity properties [28]. Rigidity is a fundamental phenomenon in HG and holomorphic dynamics. Its meaning is that the metric properties of certain manifolds or dynamical systems are determined by their combinatorics. Moreover, this phenomenon is intimately linked to the universality phenomenon, to basic measure-theoretical and topological properties of systems, to the problem of describing typical systems.

An important development in Riemann surface theory has been the discovery of fertile connections between rational billiards, translation surfaces and flows on Teichmüller space and moduli space. In fact, the Teichmüller space is the universal covering orbifold of the Riemann moduli space. A major focus of specific current mathematics research program is to explore this subject and its connections to HG, and the combinatorics of the complex of curves on a surface. So, Riemann surfaces and their MSs bridge several mathematical disciplines like HG, complex analysis, topology, and dynamics, and influences them all in return.

So, the above Non-Euclidean IGT Manifolds and Riemann Surfaces overview suggest us that to grasp a more reliable representation of reality and to get stronger physical and biological system correlates, researchers and scientists need two intelligently articulated hands: both stochastic and combinatorial approaches synergically articulated by natural coupling. We just need to remember the Relativity's father inspiration quote: "We cannot solve our problems with the same thinking we used when we created them." Because we cherish the past as a collective guide to behaviour, the general consensus alters very slowly. Scientists as well as laymen ignore evidence incompatible with their preconceptions. New theories which fail to fit established views are resisted, in the hope that they will prove false or irrelevant; old ones yield to convenience rather than to evidence.

4. Computational Information Conservation Theory

The first attempt to identify basic principles, to synergically articulate computational information conservation theory CICT by natural coupling to IGT, for scientific research and application, has been developing at "Politecnico di Milano" since the end of last century. In 2013, the basic principles on CICT, from discrete system parameter and generator, appeared in literature [30]. CICT defines an arbitrary-scaling

discrete Riemannian manifold uniquely, under hyperbolic geometry (HG) metric, that, for arbitrary finite point accuracy level L going to infinity (exact solution theoretically), is isomorphic to traditional IGT Riemannian manifold. In other words, HG can describe a projective relativistic geometry directly hardwired into elementary arithmetic long division remainder sequences, offering many competitive computational advantages over traditional Euclidean approach. Traditional Number Theory and modern Numeric Analysis use classic mono-directional interpretation (left-to-right, LTR) for \mathcal{Q} Arithmetic single numeric group generator, so information entropy generation cannot be avoided in contemporary computational algorithm and application (external world representation point-of-view). On the contrary, according to CICT, it is quite simple to show information conservation and generator reversibility (right-to-left, RTL), by using basic considerations only (internal world representation point-of-view).

Traditional digital computational resources are unable to capture and to manage not only the full information content of a single Real Number \mathbf{R} , but even Rational Number \mathcal{Q} is managed by information dissipation. In numeric representation of Rational Number \mathcal{Q} , rational proper quotient is represented by infinite repetition of a basic digit cycle, called "reptend" (the repeating decimal part). Let us consider fraction $1/D$, where D in \mathbf{Z} , called Egyptian fraction, with no loss of generality for common fraction. According to CICT, the first repetition of basic digit cycle of max length L corresponds to the first full scale interval where number information can be conserved completely, and is called "Representation Fundamental Domain" (RFD) [30]. Elementary number theory considerations give us worst case RFD word length $L = D - 1$ digits, if and only if 10 (by decimal base representation system, with no loss of generality) is a primitive root modulo D . Otherwise L is a factor of " $D - 1$." If the period of the corresponding repeating decimal to $1/D$ for prime D is equal to " $D - 1$ ", then the repeating decimal part is called "cyclic number" and D can be referred as "primitive number" or solid number (SN) or "full reptend prime" elsewhere. Thus an SN is necessarily prime. It is a sufficient qualification. Conversely a prime number may not be an SN. In classical arithmetic long division algorithm (the one you learn to divide at elementary school), usual dominant result (quotient, \mathcal{Q}) is important, and traditionally minority components (remainders, R) are always discarded. What a waste! In fact, Remainder R_L , at any division computation evolutive stage L , is the fixed multiplicative ratio of a formal power series associated to optimal decimal representations of $1/D$, at increasing arbitrary accuracy levels. In 2013, CICS showed that long arithmetic division minority components (Remainders, R), for long time concealed relational knowledge to their dominant result (Quotient, \mathcal{Q}), not only can always allow quotient regeneration from their remainder information to any arbitrary precision, but even to achieve information conservation and entropy minimization, in systems modeling and post-human cybernetic approaches [31,32]. According to CICT optimized infocentric worldview [4], symmetry properties play a fundamental role and affect word level structures and properties in analogous way to phoneme level and syllable level properties which create "double articulation" in human language, at least [33,34]. In case of figures or image sequences, Italian semiotician Umberto Eco (1932-) argued that "Iconic Language" has a "triple articulation": Iconic Figures, Semes (combinations of Iconic Figures), and Kinemorphs (combination of Semes), like in a classical movie [35,36]. Traditional \mathcal{Q} Arithmetic can be regarded as a highly sophisticated open logic, powerful and flexible optimized "OpeRational" (OR) LTR and RTL formal numeric language of

languages, with self-defining consistent words and rules, starting from self-defined elementary generator and relation, based on recursively self-defining atom [30]. For instance, at any LTR computation stage, with remainder knowledge only, it is always possible to regenerate exact quotient and new remainder information at any arbitrary accuracy, with full information conservation. It is like to process tail information to regenerate the associated body information. Thanks to the above properties, the division algorithm can become free from trial and error like in Finite Segment P-adic representation systems, but with no usually associated coding burden. The rich operative scenario offered by combinatorial modular group theory is full of articulated solutions to information processing problems. One of the earliest presentations of a group by generator and relation was given by the Irish mathematician William Rowan Hamilton in 1856, in his Icosian Calculus, a presentation of the icosahedral group [37,38]. Every group has a presentation, and in fact many different presentations. A presentation is often the most compact way of describing the structure of the group. In abstract algebra, the "fundamental theorem of cyclic groups" states that every subgroup of a cyclic group G is cyclic. Moreover, the order k of any subgroup of a cyclic group G of order n is a divisor of n , and for each positive divisor k of n , the group G has exactly one subgroup of order k . This is just the first step to start an intriguing voyage from the concept of "presentation of a group" to the concept of "representation theory" for combinatorial modular group theory [39]. Furthermore, CICT sees rational geometric series as simple recursion sequences in a wider recursive operative framework where all algebraic recursion sequences of any countable higher order include all the lower order ones and they can be optimally mapped to rational number system \mathcal{Q} OR representations and generating functions. For instance, arithmetic progression and Lucas sequences are recursion sequences of the second order. Lucas sequences are certain integer sequences that satisfy Lucas recurrence relation defined by polynomials $Un(P,Q)$ and $Vn(P,Q)$, where Un , Vn are specific polynomials and P , Q are fixed integer coefficients. Any other sequence satisfying this recurrence relation can be represented as a linear combination of the Lucas sequences $Un(P,Q)$ and $Vn(P,Q)$. Famous examples of Lucas sequences include the Fibonacci numbers, Mersenne numbers, Pell numbers, Lucas numbers, Jacobsthal numbers, and a superset of Fermat numbers. CICT is able to fold any recursion sequence of the first order into one digit number D_1 , any recursion sequence of second order into a two digit number D_2 , any recursion sequence of the third order into a three digit number D_3 and so on to higher orders. Then, you can interpret their asymptotic convergence ratios as increasing accuracy approximations to related asymptotic roots from corresponding first, second, third, ..., n -th order equations respectively. Thanks to this brand new knowledge and following this line of generative thinking, it is possible immediately to realize that traditional \mathcal{Q} Arithmetic can be even interpreted, by new eyes, as a highly sophisticated open logic, powerful and flexible LTR and RTL evolutionary, generative, formal numeric language of languages, with self-defining consistent numeric words and rules, starting from elementary generator and relation (you get your specific formal numeric language by just simply choosing your most convenient numeric base to polynomially structure your information). American linguist, philosopher, cognitive scientist, logician, Avram Noam Chomsky's (1928-) Theory of Syntax came after his criticism of probabilistic associative models of word order in sentences by Markov process approaches, in 1957. As a matter of fact, since 1951, the inadequacy of probabilistic LTR models (Markov process) had already been noticed by American psychologist and behaviorist Karl Spencer Lashley (1890-1958),

who anticipated Chomsky's arguments, by observing that probabilities between adjacent words in a sentence have little relation to grammaticality of the string. Ambiguity too provides a strong indication that sentences carry a structure. The treatment of numeric word, generator, relation and language by CICT largely draws its inspiration from many reliable research sources. Between them, the line of research started by Chomsky and French mathematician and Doctor of Western Medicine Marcel-Paul "Marco" Schützenberger (1920-1996) in the early 1960s occupies a singular place. Those fascinating viewpoints invite us to use a mind open logic approach to find new, more convenient solutions to old problems, always!

4.1. CICT First Fundamental LTR-RTL Correspondence Relationship

The CICT fundamental relationship that ties together numeric body information of divergent and convergent monotonic power series in any base (in this case decimal, with no loss of generality), with D ending by digit 9, is given by the following CICT fundamental LTR-RTL correspondence equation [30]:

$$\frac{1}{D} = \sum_{k=0}^{\infty} \frac{1}{10^W} \left(\frac{\bar{D}}{10^W} \right)^k \Rightarrow \dots \Leftarrow \text{Div} \left(\frac{1}{D} \right) = \sum_{k=0}^{\infty} (D+1)^k \quad (1)$$

where \bar{D} is the additive 10^W complement of D , i.e. $\bar{D} = (10^W - D)$, W is the word representation precision length of the denominator D and "Div" means "Divergence of". Further generalizations related to D ending by digit 1 or 3 or 7 are straightforward. Furthermore, When $\bar{D} > D$ the formal power series on the left of eq.(1) can be rescaled mod D , to give multiple convergence paths to $1/D$, but with different "convergence speeds." The total number of allowed convergent paths, as monotonic power series, is given by the corresponding Q_L value, at the considered accuracy level L . So, increasing the level of representation accuracy, the total number of allowed convergent paths to $1/D$, as monotonic power series (as allowed conservative paths), increases accordingly and can be counted exactly, and so on, till maximum machine word length and beyond: like discrete quantum paths denser and denser to one another, towards a never ending "blending quantum continuum," by a top-down perspective. Rational representations are able to capture two different type of information at the same time, modulus (usual quotient information) and associated inner or intrinsic period information which an inner phase can be computed from. So, rational information can be better thought to be isomorphic to vector information rather than to usual scalar one, at least. Furthermore, our knowledge of RFD Q_L and corresponding RFD R_L can allow reversing numeric power convergent sequence to its corresponding numeric power divergent sequence uniquely. Reversing a convergence to a divergence and vice-versa is the basic property to reach information conservation, i.e. information reversibility, as from (1). CICT results have been presented in term of classical power series to show the close relationships to classical and modern control theory approaches for causal continuous-time and discrete-time linear systems. Usually, the continuous Laplace transform is in Cartesian coordinates where the x -axis is the Real axis and the discrete Z-transform is in circular coordinates, where the Rho -axis is mapping the Real axis. By using this approach, it is possible to generate LTR and RTL remainder sequences that show same quotient body information (arbitrary-scale periodic) and specific quotient head and tail information to compute deterministic boundary values, to sustain body periodicity with no information dissipation

(full information conservation and reversibility) [30]. Traditional rational number system \mathcal{Q} properties allow to compute evolutive irreducible co-domain for every computational operative domain used. Then, all computational information usually lost by using traditional computational approach can be captured and recovered by a corresponding complementary co-domain, step-by-step. Then co-domain information can be used to correct any computed result, achieving computational information conservation. CICT is offering an operational optimized infocentric worldview [4]. CICT fundamental relation (1) allows to focus our attention on combinatorially optimized number pattern generated by LTR or RTL phased generators and by convergent or divergent power series with no further arbitrary constraints on elementary generator and relation. Thanks to subgroup interplay and intrinsic phase specification through polycyclic relations in each SN remainder sequence, word inner generator combinatorial structure can be arranged for "pairing" and "fixed point" properties for digit group with the same word length. As a matter of fact, those properties ("pairing" and "fixed point") are just the operational manifestation of universal categorical irreducible dichotomy hard-wired into integer digit and digit group themselves (i.e. "evenness" and "oddness") and to higher level structures (i.e. "correspondence" and "incidence"). Actually, since space is limited, the discussion here will not be extended further to the subgroup interplay of the family group and polycyclic groups. We refer the interested reader to good general references on polycyclic groups [40,41].

5. A CICT Solution Proposal

Usual knowledge on significant figures of a number teaches that any 0 digit that comes before the first nonzero digit (leading zeros) can be omitted in a number string in positional notation representation system [42]. When leading zeros occupy the most significant digits of an integer, they could be left blank or omitted for the same numeric value [43]. Therefore, the usual decimal notation of integers does not use leading zeros except for the zero itself, which would be denoted as an empty string otherwise [44]. However, in decimal fractions between 0 and 1, the leading zeros digits between the decimal point and the first nonzero digit are necessary for conveying the magnitude of a number and cannot be omitted [42]. Let us introduce a convenient LTR symbolic compression operator as $SCO \equiv \langle M \mid DS \rangle$, where DS is a finite digit string of length L and M is the number of times DS is repeated to get our unfolded digit string in full (e.g. $(4 \mid 1) \equiv 1111$ or $(2 \mid 123) \equiv 123123$). Usual symbolic string operations can be applied to SCO. Then, we can write usual rational number OpeRational Representation (OR) corresponding to their Symbolic Representation (SR) as [30]:

$$\begin{aligned}
 Q_1 &= \frac{1}{D_1} = \frac{1}{9} = 0.1111111111\dots \\
 Q_2 &= \frac{1}{D_2} = \frac{1}{99} = 0.010101010101010101\dots \\
 Q_3 &= \frac{1}{D_3} = \frac{1}{999} = 0.001001001001001001001001001\dots
 \end{aligned} \tag{2}$$

in a more compact RFD Q_L format as:

$$\begin{aligned}
Q1 &= \frac{1}{D1} = \frac{1}{9} \equiv 0. < \infty 11 > \\
Q2 &= \frac{1}{D2} = \frac{1}{99} \equiv 0. < \infty 101 > \\
Q3 &= \frac{1}{D3} = \frac{1}{999} \equiv 0. < \infty 1001 >
\end{aligned} \tag{3}$$

In the same way, we can write:

$$\begin{aligned}
Q4 &= \frac{1}{D4} = \frac{1}{7} = 0.142857142857142857142857142857... \text{ as} \\
Q4 &= \frac{1}{D4} = \frac{1}{7} \equiv 0. < \infty 1142857 >
\end{aligned} \tag{4}$$

On the other hand, we have:

$$\begin{aligned}
Q5 &= \frac{1}{D5} = \frac{1}{142857} = 0.000007000007000007000007000007... \text{ written as} \\
Q5 &= \frac{1}{D5} = \frac{1}{142857} \equiv 0. < \infty 1000007 > \quad \text{either (first compressed form)} \\
Q5 &= \frac{1}{D5} = \frac{1}{142857} \equiv 0. < \infty I(< 510 \times 117 >) > \quad \text{or (second compressed form)}.
\end{aligned} \tag{5}$$

Now, we can realize that $Q4$ RFD is related by $Q5$ RFD and vice-versa by period precision length $L = 6$. So, to conserve the full information content of rational correspondence between $Q4$ and $Q5$, we realize that we have to take into account not only the usual $Q4$ and $Q5$ modulus information, but even their related period precision length information $L = 6$ (external world representation phase). As far as it concerns $D5$ it comes almost automatically from previous example (see (5)), but the same it is not true for $D4$ (see (4)) because we wrote digit 7 only as denominator, without its five leading zeros, according to traditional knowledge on significant figures of a number. In that way, we loose the rational correspondence intrinsic period information (coherence) which an inner relative phase for each RTL string generator can be computed from (i.e. from their optimized exponential cyclic sequences (OECS) of R_L [30]). With no coherent inner phase information, we get entropy generation and dissipation. In fact, misplaced precision leads to information opacity, fuzzyness, irreversibility, chaos, complexity and confusion. Therefore, rational information can be better thought to be isomorphic to vector information rather than to usual scalar one, at least. Now, from (5) second compressed form, it is immediate to verify the following phase relations:

$$\begin{aligned}
QQ1 &= \frac{1}{DD1} = \frac{1}{142857} \equiv 0. < \infty I(< 510 \times 117 >) > \\
QQ2 &= \frac{1}{DD2} = \frac{1}{142857142857} \equiv 0. < \infty I(< 1110 \times 117 >) > \\
QQ3 &= \frac{1}{DD3} = \frac{1}{142857142857142857} \equiv 0. < \infty I(< 1710 \times 117 >) > \\
&\dots \quad \dots \\
QQN &= \frac{1}{DDN} = \frac{1}{< N1142857 >} \equiv 0. < \infty I(< 5 + 6(N-1)10 \times 117 >) > \quad \text{for } N=1,2,3,\dots,\infty
\end{aligned} \tag{6}$$

Therefore, we can write the following relation:

$$CQ4 = \frac{1}{CD4} \equiv \frac{1}{\langle \infty I (\langle \infty I 0 \rangle \langle 117 \rangle) \rangle} \equiv 0. \langle \infty I 142857 \rangle \quad (7)$$

According to our SCO approach, the coherent representation $CD4$ emerges out of an LTR infinity of symbolic structured infinite length sequences in (7). By this point of view, natural numbers, according to human common knowledge, appear as just the righthmost approximated part of those sequences. So, $CD4$ in (7) is the correct coherent relation representation of traditional scalar modulus $D4$ in (4) as denominator, while scalar modulus $D4$ in (4) can be interpreted as the decoherented relation representation of $CD4$ denominator in (7). Leading zeros in positional notation representation system for CICT Q Arithmetic do count effectively, and can model the quantum-classical system transition quite efficiently. Finally, our knowledge of RFD Q_L and corresponding RFD R_L can allow reversing LTR numeric power convergent sequence to its corresponding RTL numeric power divergent sequence uniquely (see (1)) [30]. Reversing a convergent sequence into a divergent one and vice-versa is the fundamental property to reach information conservation, i.e. information reversibility. Eventually, OECS have strong connection even to classic DFT algorithmic structure for discrete data, Number-Theoretic Transform (NTT), Laplace and Mellin Transforms [4]. Coherent precision correspondence leads to transparency, ordering, reversibility, kosmos, simplicity, clarity, and, as you saw from previous discussion, to algorithmic quantum uncomputability on real macroscopic machines [12].

6. Conclusions

The final result is CICT new awareness of a hyperbolic framework of coded heterogeneous hyperbolic structures, underlying the familiar Euclidean surface representation system. CICT emerged from the study of the geometrical structure of a discrete manifold of ordered hyperbolic substructures, coded by formal power series, under the criterion of evolutive structural invariance at arbitrary precision. It defines an arbitrary-scaling discrete Riemannian manifold uniquely, under HG metric, that, for arbitrary finite point accuracy L going to infinity (exact solution theoretically), is isomorphic to traditional IGT Riemannian manifold. In other words, HG can describe a projective relativistic geometry directly hardwired into elementary arithmetic long division remainder sequences, offering many competitive computational advantages over traditional Euclidean approach. It turns out that, while free generator exponentially growing sequences can be divergent or convergent, their closures can be defined in terms of polynomials. Furthermore, combinatorially optimized exponential cyclic sequences (OECS) have strong connection even to classic DFT algorithmic structure for discrete data, Number-Theoretic Transform (NTT), Laplace and Mellin Transforms [4]. In this way, even simple traditional scalar moduli can emerge out as an approximation from sequences of phased generators. CICT can help to reach a unified vision to many current physics and biophysics problems and to find their optimized solutions quite easily. Expected impacts are multifarious and quite articulated at different system scale level. One of the first practical result was that usual elementary arithmetic long division remainder sequences can be even interpreted as combinatorially optimized coding sequences for hyperbolic geometric structures, as point on a discrete Riemannian manifold, under HG metric, indistinguishable from traditional random noise sources by classical Shannon entropy, and contemporary most advanced instrumentation, as discussed in Section 1. Specifically, CICT showed that classical Shannon entropy computation is completely unable to reliably discriminate so called

computational "random noise" (RN) from any combinatorially optimized encoded message by OECS, called "deterministic noise" (DN) in [4]. As a matter of fact, for any free generator, CICT can provide us with an "ecoco-domain" multiscale evolutive structured family of sequences that can be used for checking for the presence of a specific generator in laboratory or system "background noise" [4]. Following CICT approach, it is possible even to extend the classic Shannon entropy concept to arrive to a stronger and specific "Coherent Shannon entropy" (CSE) approach. Second result was to realize that classical experimental observation process, even in highly ideal operative controlled condition, like the one achieved in contemporary most sophisticated and advanced experimental laboratories like CERN, can capture just a small fraction only, with misplaced precision, of overall ideally available information from unique experiment. The remaining part is lost and inevitably added to something we call "background noise" or "random noise" usually, in every scientific experimental endeavor. CICT can help us to develop strategies to gather much more reliable experimental information from single experimentation and to conserve overall system information [4]. In this way, coherent representation precision leads to information conservation and clarity. Specifically, high reliability organization (HRO), mission critical project (MCP) system, very low technological risk (VLTR) and crisis management (CM) system will be highly benefitted mostly by these new techniques. The latest CICT claim has been that the "external" world real system physical manifestation properties and related human perception are HG representation based, while Euclidean approximated locally. Furthermore, the fundamental play of human information observation interaction with an external world representation is related by the different manifestation and representation properties of a unique fundamental computational information structuring principle: the Kelvin Transform (KT). KT is key to efficient information representation, structuring "external world representation" information to an "internal world representation" and vice-versa by inversive geometry.

Conflicts of Interest

The author declares no conflict of interest.

References and Notes

1. Lachmann, M.; Newman, M.; Moore, C. The physical limits of communication or Why any sufficiently advanced technology is indistinguishable from noise? *American Journal of Physics* **2004**, Volume 72, 10, 1290–1293.
2. Calude, C. *Information and Randomness-An Algorithmic Perspective*, 2nd edn.; Springer: Berlin, 2002.
3. Calude, C.; Svozil, K. Quantum Randomness and Value Indefiniteness. *Advanced Science Letters* **2008**, 1, 165–168.
4. Fiorini, R.A. How Random is Your Tomographic Noise? A Number Theoretic Transform (NTT) Approach. *Fundamenta Informaticae* **2014**, Volume 135, 3-4, 1-36.
5. Shannon, C.E. A Mathematical Theory of Communication. *Bell System Technical Journal* July/October **1948**, Volume 27, 3, 379–423.
6. Park, J.; Johnson, J.T.; Majurec, N.; Niamsuwan, N.; Piepmeier, J.R.; Mohammed, P.N.; Ruf, C.S.; Misra, S.; Yueh, S.H.; Dinardo, S.J. Airborne L-Band Radio Frequency Interference Observations From the

- SMAPVEX08 Campaign and Associated Flights. *IEEE Trans. on Geoscience and Remote Sensing* **2011**, Volume 49, 12, 3359–3370.
7. C. E. U. European Laboratory for Particle Physics, Conseil Européen pour la Recherche Nucléaire. <http://home.web.cern.ch/about/updates/2013/11/atlas-sees-higgs-boson-decay-fermions> (accessed on 5th October 2014).
 8. Martin, N.; England, J., *Mathematical Theory of Entropy*, Cambridge University Press:Cambridge, 2011.
 9. MacKay, D. *Information Theory, Inference, and Learning Algorithms*, Cambridge University Press: Cambridge, 2003.
 10. Bergstrom, C. T.; Lachmann, M. Shannon information and biological fitness. *IEEE information theory workshop* **2004**, 50–54.
 11. Donaldson-Matasci, M. C.; Bergstrom, C. T.; Lachmann, M. Shannon information and biological fitness. *Oikos* **2010**, 119, 219–230.
 12. Huang, Y. Computing quantum discord is NP-complete. *New Journal of Physics* 21 March **2014**, Volume 16, 3, 033027.
 13. Taleb, N.N. *Silent Risk*. Preliminary Book Draft For Error Detection, September, 2014.
 14. Schlosshauer, M. *Decoherence and the Quantum-to-Classical Transition*, Springer:Heidelberg/Berlin, 2007.
 15. Olivier, H.; Poulin, D.; Zurek, W.H. Objective properties from subjective quantum states: Environment as a witness. *Phys.Rev.Lett.* **2004**, 93, 220401.
 16. Blume-kohoul, R.; Zurek, W.H. Quantum Darwinism: Entanglement, Branches, and the Emergent Classicality of Redundantly Stored Quantum Information. *Phys. Rev. A.* **2006**, 73, 062310.
 17. Arndt, M.; Nairz, O.; Vos-Andreae, J.; Keller, C.; van der Zouw, G.; Zeilinger, A. Wave–particle duality of C60 molecules. *Nature* October **1999**, 401, 680–682.
 18. LaHaye, M.D.; Buu, O.; Camarota, B.; Schwab, K.C. Approaching the Quantum Limit of a Nanomechanical Resonator. *Science* **2004**, 304, 74–77.
 19. Wineland, D.J.; Dehmelt H. Proposed $10^{14}\Delta\nu < \nu$ laser fluorescence spectroscopy on Ti^+ mono-ion oscillator. *Bull Am Phys Soc.* **1975**, 20,637.
 20. Hänsch, T.W.; Shawlow, A.L. Cooling of gases by laser radiation. (1975). *Optics Communications* **1975**, Volume 13, 1, 68–69.
 21. Licata, I.; Fisaletti, D. A Fisher-Bohm Geometry for Quantum Information. *EJTP* **2014**, Volume 11, 31, 1–18.
 22. Amari, S. *Methods of Information Geometry*, Oxford University Press: Oxford, 2000. (Amari S., Nagaoka H. *Methods of Information Geometry*, originally published in Japanese in 1993 and published in English in 2000 with the American Mathematical Society, AMS).
 23. Kullback, S. Leibler, R.A., On Information and Sufficiency. *Annals of Mathematical Statistics* **1951**, Volume 22, 1, 79–86.
 24. Kullback, S. *Information theory and statistics*, John Wiley and Sons: New York, 1959.
 25. Kullback, S. Letter to the Editor: The Kullback–Leibler distance. *The American Statistician* **1987**, Volume 41, 4,340–341.
 26. Burnham, K.P.; Anderson, D.R. *Model Selection and Multi-Model Inference: A Practical Information-Theoretic Approach*, Springer:Berlin/Heidelberg, 2002.

27. Bregman, L. M. The relaxation method of finding the common points of convex sets and its application to the solution of problems in convex programming. *USSR Computational Mathematics and Mathematical Physics*, 1967, Volume 7, 3, 200–217.
28. Eskin, A.; Mirzakhani, M. Invariant and stationary measures for the $SL_2(\mathbb{R})$ action on moduli space. Preprint 2013 (see arXiv:1302.3320).
29. Eskin, A.; Mirzakhani, M.; Mohammadi, A. Isolation, equidistribution, and orbit closures for the $SL_2(\mathbb{R})$ action on moduli space. Preprint 2013 (see arXiv:1305.3015).
30. Fiorini, R.A.; Laguteta, G. Discrete Tomography Data Footprint Reduction by Information Conservation. *Fundamenta Informaticae*, 2013, Volume 125, 3-4, 261–272.
31. Fiorini, R.A.; Santacroce, G. Economic Competitivity in Healthcare Safety Management by Biomedical Cybernetics ALS. *Proc. International Symposium, The Economic Crisis: Time For A Paradigm Shift - Towards a Systems Approach*, Universitat de València, January 24-25, 2013.
32. Fiorini, R.A.; Santacroce, G. Systems Science and Biomedical Cybernetics for Healthcare Safety Management. *Proc. International Symposium, The Economic Crisis: Time For A Paradigm Shift - Towards a Systems Approach*, Universitat de València, January 24-25, 2013.
33. Hockett, C. *A Course in Modern Linguistics*, Macmillan: New York, 1958.
34. Hockett, C. The Origin of Speech. *Scientific American* 1960, 203, 89–96.
35. Eco, U. *A Theory of Semiotics*, Macmillan: London, 1976.
36. Stam, R. *Film Theory*, Blackwell: Oxford, 2000.
37. Chandler, B.M. The history of combinatorial group theory: A case study in the history of ideas. In *Studies in the history of mathematics and physical sciences* , 1st ed., Springer: Berlin/Heidelberg, 1982, 387–407.
38. Hamilton, W.R. Memorandum respecting a new System of Roots of Unity. *Philosophical Magazine* 1856, 12, 446.
39. Curtis, C.W. *Pioneers of Representation Theory: Frobenius, Burnside, Schur, and Brauer, History of Mathematics*, American Mathematical Society: Providence, 2003.
40. Segal, D. *Polycyclic groups*, Cambridge University Press: Cambridge, 1983.
41. Wehrfritz, BAF. *Group and Ring Theoretic Properties of Polycyclic Groups*, Springer-Verlag: London, 2009.
42. Miller, Jane E. *The Chicago Guide to Writing about Numbers*, University of Chicago Press: Chicago, 2008, pp.276.
43. Zegarelli, M. *Basic Math and Pre-Algebra Workbook For Dummies*, (2nd ed.), John Wiley & Sons: New York, 2014, p.8.
44. Flynt, C. *Tcl/Tk: A Developer's Guide*, Elsevier:Amsterdam, 2012, pp.752–753.