Relation of π=3,14... to DNA

Marinos Spiliopoulos[1]

[1]Affiliation not available

September 28, 2020

The contingency of the connection of π (ratio of a circle's circumference to its diameter) with the DNA (genetic material) sounds like impossible. In fact, they are unrelated things, and some may characterize it as incoherence and nonsense. That is, something similar, as if you want to connect the length of a person's shadow with the money he has in his pants. In the case you insist, some may guess the double helix of DNA and its relationship to the cycle. Moreover, circular DNA has also been found in organisms. They use to say that life and its evolution, as well as the history of things, have a spiral shape. Let us get things in the right order: DNA building blocks are called nucleotides and have a specific chemical structure. Depending on the nitrogenous base they contain, they are divided into nucleotides of Adenine (A), Thymine (T), Guanine (G) and Cytosine (C). Their molecular weights are 331, 322, 347 and 307, respectively. Prime are the numbers 331, 347, 307, while 322 is composite. The symbols A, T, G, C, will help us symbolize at the same time the respective nucleotides, the respective nitrogenous bases of the nucleotides and their molecular weights. A mathematical formula that connects these numbers and therefore the molecular weights of DNA nucleotides, is: G = . There is no need for further analysis, as it results from a combination of representations between the nucleotides.

For the record, from the sums or the differences of other representations between the nucleotides, arise the following: A + T + G + C = 1307, A + T = 653, the graph AT − GC = 53, the (G + T) − (A + C) = 31 etc. are all prime numbers. The latter, leads us to look for more things. The following remains applicable:

G + T − A − C = 31

14 G + 6T − 14A − 6C = 314

141 G + 59T − 141A − 59C = 3141

1415G + 585T − 1415A − 585C = 31415

14159G + 5841T − 14159A − 5841C = 314159 etc. Table 1

The relation of each series gives us π = 3,14159... with one, two, three etc. decimal digits multiplied by 10, 100, 1000 etc. respectively. Furthermore, the sum of the coefficients of G and T, respectively of A and C, is 2, 20, 200, 2000, 20000 etc. a special connection of DNA with π . Proportionally, the approximation ofπ by Archimedes, may come to someone's mind. On one hand, there were sides of regular polygons inscribed, whereas here are by proportion nucleotides. We will try to "mathematize" various processes and see how far we can go. Let two consecutive series e.g. ν and $\nu+1$. The coefficients of the nucleotides of the **nth** order are φ and x with a total result of Z_1. Respectively the coefficients and the result of the ν **+1** series are y, ω and Z_2. In particular:

$\varphi G + xT - \varphi A - xC = Z_1$ (1)

$yG + \omega T - yA - \omega C = Z_2$ (2)

Example: for the third row ($\nu = 3$) there is $\varphi = 141$, $x = 59$, $Z_1 = 3141$. For the fourth row ($\nu + 1 = 4$), $y = 1415$, $\omega = 585$, Z2= 31415 (see table 1) etc. From the previous there is $\varphi + x = 2$ [?] $10^{\nu\text{-}1}$ (3) and $y + \omega = 2$ [?] 10^{ν} (4). For example, for $\nu = 2$, $\varphi + x = 20$ and $y + \omega = 200$ etc.

Next: From (1) $\varphi(G - A) + x\,(T - C) = Z_1$ (5) and from (2)

$y\,(G - A) + \omega\,(T - C) = Z_2$ (6).

Since $G - A = 347 - 331 = 16$ and $T - C = 322 - 307 = 15$ the relations (5) and (6) can be written respectively:

$16\varphi + 15x = Z_1$ and $16y + 15\omega = Z_2$. These last relations give us the possibility to connect all six factors of equations (1) and (2) in an equation. Specifically, to equalize the representations e.g. equivalent to 16. That is 16 = and to arrive at the equation

$Z_2\varphi = Z_1 y - 15xy + 15\varphi\omega$ (7).

Additionally, those that equate to 15, and end up in the relationship

$Z_2 x = Z_1\omega - 16\varphi\omega + 16xy$ (8).

From (1) and (2) by division by members easily results the (9).

By studying the initial table of series with the arithmetic coefficients, other interesting things are being noticed, too, by which **a specialist in number theory and beyond** , would be pleasantly surprised. Let us take for instance the numerical differences of the coefficients (e.g. those corresponding to nucleotides G and T), of two consecutive series and compare them. We see that the difference of the coefficients in the $\nu + 1$ series, is equal to ten times the difference of the corresponding coefficients in the ν series, plus twice the digit of the units of Z_2 or respectively of the coefficient of G of the $\nu + 1$ series. Generally, $y - \omega = 10\,(\varphi - x) + 2\alpha$ (10), where α is the digit of the units we mentioned.

Example between third and fourth series ($\nu = 3$): $141 - 59 = 82$ $1415 - 585 = 830$

$830 = 10 \cdot 82 + 2 \cdot 5$ etc. It easily follows from the relation (10) and (4) $y = 10^{\nu} + 5(\varphi - x) + \alpha$ (11). For example,$\nu = 4$ $y = 10^4 + 5\,(1415 - 585) + 9 = y = 14159$ (coefficient G of $\nu + 1 = 5$ series) etc. It also holds for $y = 10\varphi + \alpha$ and $Z_2 = 10Z_1 + \alpha$. Moreover, interesting is the difference between the crossed products of the coefficients, in two consecutive series. To make a long story short, the general relation $xy - \varphi\omega = 2 \cdot 10^{\nu-1}\alpha$ (12) is valid.

An example between the third and the fourth series: $141 \cdot 585 = 82485$,

$59 \cdot 1415 = 83485$ $83485 - 82485 = 1000$ $1000 = 2 \cdot 10^2 \cdot 5$ etc. Other mathematical relations can also be found and after all a question can arise in everyone: If we know φ, x, Z_1(previous series) and the relationship between the coefficients of the nucleotides of the next series (as well as of other mathematical relations), can we calculate the Z_2 or the α; Namely the digits of the units Z_2 and consequently the next decimal digit of π ; No matter how many attempts I made, no matter how many combinations of procedures and mathematical formulas I used, it always turned out for Z_2 and α a representation that led to zero. The factors were contradictory and resulted in zero results. This forced me for a long time not to deal with the issue (to leave it in the drawer of my office). At one point I thought (the thought matured), that some others, based on the ideas of the essay, can find other better and more interesting things and why not some bio-computational approach of digits of π . On the other hand, why should the essay have only this goal? After all, mathematicians have found mathematical formulas and special algorithms, to calculate through the use of computers, trillions of digits of π , and this has no end. The essay should not be based solely on the calculation of decimal places of π , but to show the beauty of mathematical relations in DNA, in relation to π and where else it arises, etc. Unfortunately, what we use to say in many cases was true for me: "I couldn't see the woods for the trees". I missed something, that is self-evident for every mathematician: "**mathematicians do not do what they do, because it is necessary and no one knows in advance if something is needed or not** ". In the end, every cloud has a silver lining. Along the way I found various things. Specifically, at the beginning of the essay we saw some representations of nucleotides, which in relation to their molecular weights give prime numbers. We can easily and "artificially" give DNA chains with a total molecular weight of prime numbers. Example: from the part of a DNA chain that contains scattered but in total 29A, 59T, 131G, 107C, results a molecular weight of 106903, which is the prime number. The numbers 29, 59, 131, 107 have been chosen to be prime numbers in this example, while the complementary part of the chain with a set of nucleotides 29T, 59A, 131C, 107G has a molecular weight of 106213, which is also a prime number. These sums result from the addition of nucleotides in free form. If we remove the water molecules (which has a molecular weight of 18) that are removed when the nucleotides are connected together by phosphodiester chemical bonds, in the second chain (is it transcribed?) has a molecular weight of 100363 which is also a prime number. One might also wonder what is happening in real conditions. Do the first numbers define properties there? Do they leave their mark? What happens, for example, to the genes, to their regions (introns) that do not encode amino acids, to the regions of DNA (which are the most) and do not encode proteins, where they encode, etc. Do they play a role in their presence or absence? Let it be researched. Of course, the most important thing in DNA is the sequence of nucleotides, because it determines the order of amino acids in proteins based on the genetic code. However, the fact that the prime numbers "have a finger in every pie" can make someone wonder whether there is a deeper connection between the evolution of genomic information (DNA) through natural selection, mutations, etc. and prime numbers. In addition, in the mathematical formula (mentioned at the beginning) and it connects the molecular weights of the nucleotides, if we put in the place of C and G some prime number, we calculate the product $A{\cdot}T$. If we put (in proportion of the nucleotides) in A a prime number and in T a composite that expresses their number, the sum of the numbers as in DNA is very often a prime number. Example: C=3, G=7 we find $A \cdot T = 54$. We put $A = 3$ and therefore T is 18. The sum $3 + 7 + 3 + 18 = 31$ (prime). If $A = 2$ and $T = 27$ does not result a prime one. The mathematical formula is to function as a "machine" for the production of prime numbers. Let the prime one be investigated in relation to the other nucleic acid, the RNA, which is also composed of nucleotides. RNA is connected to the DNA from which it is derived, as well as to the proteins produced by the encoded information. Who knows whether this perspective will let us study Covid-19 in a better way? We continue with something else amazing. The first epigenetic lesion (mutation) found in DNA was the cytosine methylation. Methylation therein means the addition of a methyl group, i.e. a small chemical molecule consisting of one carbon atom and three hydrogen atoms with a total molecular weight of 15. The cytosine nucleotide gains molecular weight $307 + 15 = 322$, equal to that of thymine. From the prime number of cytosine, we come to the composite number of thymine. And here a "confusion" arises. This mutation in cytosine causes various phenomena (inactivation of genes, or their activation above normal, depending on the number of methylations, etc.) that can result in various diseases (cancers, etc.). We see the cytosine mutation to occur on one hand at the chemical level of its structure which is the most important, but also in the equation of the molecular weight of its nucleotide with that of the nucleotide of another chemical base (thymine) of DNA. Apart from the chemical change and the change in the numbers in some proportion, they play their role at least in the specific event. Shall we say something that may seem strange to many? Can further or basic mathematics cure, in the future, diseases if we associate some features with them, through some mathematical model? Biomathematics (connection of mathematics and biology) something similar to biochemistry and biophysics (connection of chemistry and physics with biology respectively), have offered and will offer more in the future, to the study of life in living organisms. Biomathematics (biostatistics is a subset of them), in addition to the huge contribution to science, can also offer career prospect to mathematicians. Let us conclude the essay with something important that was found along the way. In Table 1 (in the various series), we saw the numerical coefficients of the nucleotides and the relationships, developed between them. Taking into account the quotients (ratios) of these coefficients of each series (see table 2), we observe the following:

$1/1 = 1$

$14/6 = 2{,}33333333\ldots$

$141/59 = 2{,}389830508\ldots$

1415/ 585 = 2,418803419...

14159/ 5841 = 2,424071221...

141592/ 58408 = 2,424188467...

....

2,42422678468...

Table 2

With the help of PC and for a large number of divisions (the computer "endured" up to 307 digits of $\boldsymbol{\pi}$) the ratio of coefficients (e.g. $\varkappa$ and λ) after the 12th digit, tends steadily to a very specific number, the 2,42422678468... At first glance, this number seems random. And I wish to protest very strongly at this. This number is very close to the three seconds of the number Φ (golden ratio). That is $\frac{3}{2}\,\Phi = 2{,}427050983...$ A connection of DNA other than $\boldsymbol{\pi}$ and $\boldsymbol{\Phi} = 1{,}618033988...$ And as we colloquially say in Greek, a study of DNA "in pi and fi". That means very rapidly. Nonetheless, the approximate formula $\boldsymbol{\pi}_{\Phi}$ = between π and Φ, is well-known to mathematicians. We continue: the bio-computer of π can work even approximately with the system of equations: $\varkappa + \lambda = 2 \cdot 10^{\nu-1}$ and $\frac{k}{l} \approx \frac{3}{2}F$ or $\frac{k}{l} = 2{,}42422678468...$ Its calculation (the digit of its units) determines the ninth decimal digit of $\boldsymbol{\pi}$.

After all these, what could we assume as a conclusion? Is the ultimate goal of nature to achieve the perfect circle or the perfect spiral? We know that the logarithmic spiral is connected to Φ. Probably when this is achieved (case of DNA formation), life is created. Life in the form of DNA loves the perfect circle, the perfect spiral and does not reject the golden ratio of things. Mutations that occur continuously and in very large numbers and that lead to the removal of the stereochemical structure of DNA (functionality and the information it contains) from the perfection of the cycle, nature itself rejects them. It allows only those that meet certain conditions (maintaining shape, functionality and information) to survive, and which are generally legitimate and can support life.