

Structural Imperfections in Family Trees: Quantifying the Ancestor Paradox

H. Daniel Wagner

Weizmann Institute of Science, Rehovot 76100, Israel

Email: daniel.wagner@weizmann.ac.il

When close family members marry, the structure of their children's ancestral tree is perturbed as lines of ancestors merge and the number of individuals on their ancestral tree decreases compared to an imaginary 'perfect tree'. This makes the search for ancestors a little easier on the family genealogist –because there are less ancestors to search for–, and provides a plausible explanation for the fact that the ancestral earth population could not have reached the enormous numbers predicted by a simple exponential calculation. Here we introduce a way to quantify the effects of such faults in the family tree of a single individual, and discuss consequences in the context of the so-called 'Ancestor Paradox' (also called 'Diamond Theory'). A numerical coefficient π is introduced, which provides a measure of the amount of ancestral branch merging in that individual's family tree. Simple examples are presented and discussed.

I. Introduction

Many contemporary Jewish individuals and families claim descent from Rabbi Shlomo Yitzhaki, generally known by the acronym Rashi, a famous medieval French rabbi who produced comprehensive commentaries on the Hebrew Bible. Rashi lived between 1040 and 1105 and had three daughters and indeed, many descendants.

Assuming that the length of one generation is 25 years, the distance from today (2012) to Rashi's adult time (1065) is 947 years, or 38 generations. In principle, since all of us have two parents, four grandparents, eight great grandparents, and so on, each living Jewish individual had, back to Rashi's time, $2^{38} = 275$ billion ancestors (275×10^9). In fact, this is the number of names that should appear at the time of Rashi on the family tree of each living person on earth (not only Jews). But obviously, at that time there were not that many people on the whole planet: the number of people on earth in the year 1100 is estimated to be 400 million only¹. This puzzle is known as the Ancestor Paradox (Pears, 1991a, 1991b, 1991c). The solution to this brainteaser is that while this enormous number of ancestors is indeed mathematically correct, a large number

¹ <http://www.prb.org/Articles/2002/HowManyPeopleHaveEverLivedonEarth.aspx>

of them are identical people. Why? Because when for example first cousins marry, their children will have only six great grandparents instead of the theoretical eight: two couples of grandparents are identical, they are the same people. This is illustrated in Figure 1.

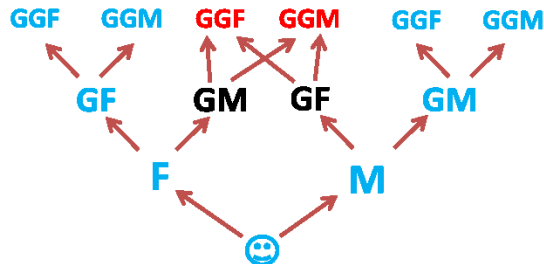


Figure 1: When father (F) and mother (M) are first cousins, one grandfather (GF in black) and one grandmother (GM in black) are siblings, meaning that they have the same parents, who are the great-grandparents (GGF and GGM in red) of the original individual (the blue smiley). As a consequence, one set of great-grandparents is 'lost' and, with it, an entire ancestral set.

The family tree of such a child has "collapsed" from 8 to 6 in the 3th generation back, and the disappearance of that branch has shrunk by 25% the number of ancestors at this and all earlier generations (see Figure 2, where I assumed that parents belong to generation 1, grandparents to generation 2 and so on).

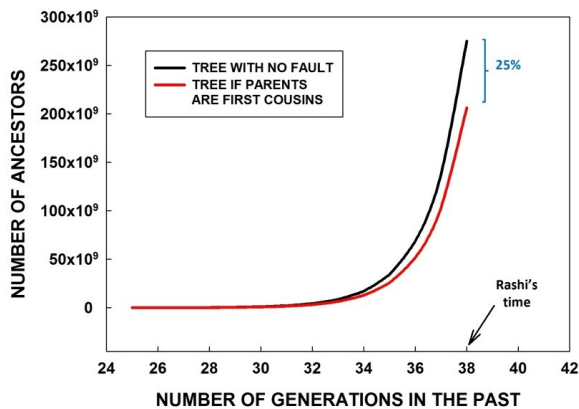


Figure 2: Comparison between the number of ancestors in each generation for a 'perfect' family tree and for the same tree when parents are first cousins (see Figure 1). The difference is 25% starting from generation 3 (not shown). Only shown here for clarity are generations 25 to 38.

When 2nd, 3rd, 4th or 5th cousins marry the same 'collapse' occurs but it is increasingly less intense. The most dramatic family tree collapse occurs when siblings marry (which seems to have been common during some Roman periods, as well as with Egyptian pharaohs and Inca kings), in which case there is a 50% collapse (from 4 to 2) starting at the 2nd generation, the grandparents, and in all previous generations. Examples abound, as follows. (i) In 1964 it was found that a third of the marriages in Andhra Pradesh, southern India, were between first cousins and that almost 12% of marriages were between uncles and their nieces. (ii) In 1875 in England 7.5 % of Jewish marriages were between 1st cousins. (iii) King Alfonso XIII of Spain (1886-1941) had only eight different people as his great great-grandparents instead of the normal sixteen, a 50% collapse of his pedigree at the 4th generation. (iv) It was found that 17 generations back, Prince Charles' family tree exhibits a collapse down to 35 % of the theoretical maximum. (v) It is estimated that, for a specific Amish family about whom a thorough genealogical tree has been compiled, 21.5 % of 627 marriages took place between 2nd cousins or closer.

More examples can be found in the literature (Shoumatoff, 1985). All of these are made more fascinating and complex by the intermarriages that have been taking place within families, isolated regions, close-knit religious communities, small tribes, and so on.

Every person's family tree has most probably the shape of a diamond. One's family tree back a few generations gets wider and wider. But eventually it will start to narrow at each generational level if you go back far enough, and then finally converge to a few ancestors (perhaps even back to humanity's single generic couple, Eve and Adam). This 'diamond scheme' is a likely description of the Ancestor Paradox.

II. Naming the relationships between people on a family tree

A family relationship (or kinship) between two individuals exists when they have one or more common ancestors. The link between family members is sometimes difficult to define. Since some of the examples used below involve more or less distant relatives, here are a few useful definitions (refer to Figure 3):

- (i) Cousin is a loosely used term to designate a relative who has a common ancestor. It is not an accurate definition.
- (ii) First cousin ('cousins germain'² in French, 'primo hermano' in Spanish) – Children of siblings are first cousins: they share grandparents. Children of first cousins are second cousins: they share

² This denomination has nothing to do with Germany. The term exists since the 12th century and comes from the latin '*germanus*' which means: "*issued from the same blood*".

great-grand-parents. Children of second cousins are third cousins, and so on.

- (iii) First cousin once removed - The 'removed' term describes the difference in generations between two people. For example, the children of your first cousins are first cousins once removed. Note that a person may have two sets of first cousins once removed: your first cousin's children are your first cousins once removed (these first cousins are once removed forward a generation); and your parent's first cousins are also your first cousins once removed (these first cousins are once removed back a generation).
- (iv) Cousins from different sides of the family (for example, maternal 1st cousin and a paternal 2d cousin) are not related. They may both be 1st cousins or 2d cousins to you but they have no connection other than that.

These definitions will suffice for the present article.

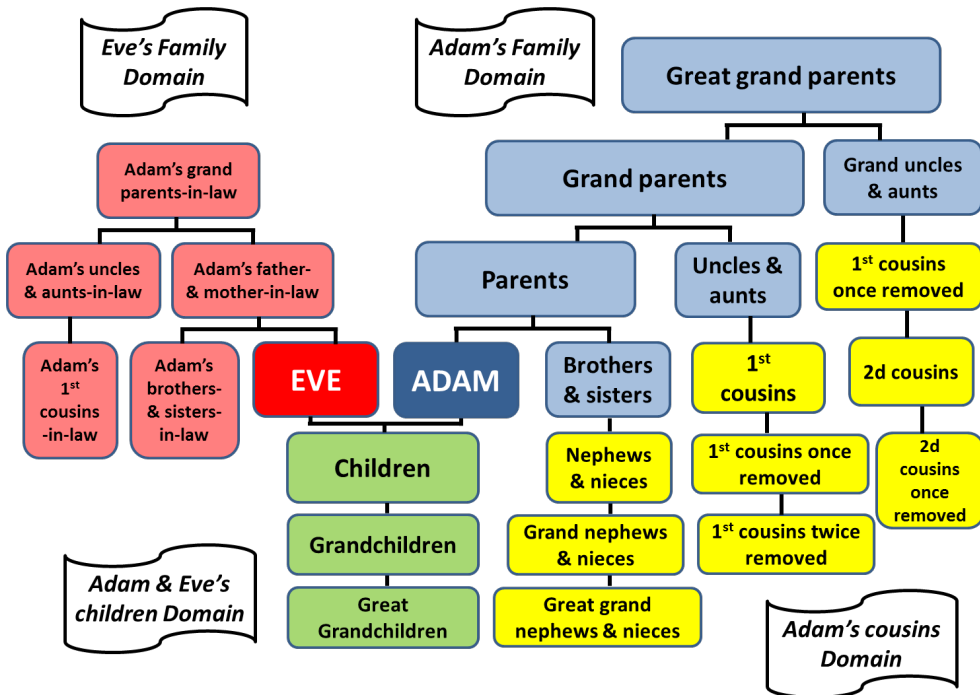


Figure 3: Relationship chart focused on a generic couple ('Eve' and 'Adam') and some of their closest family members.

III. Quantifying the ancestral populations

As already described earlier, matings occasionally arise between close relatives such as first cousins or uncle and niece. Mating between close cousins may be intentional or it may be random such as when it occurs between distant relatives who don't even know they are related. Closer interactions –between father and daughter, or brother and sister – occur rarely but are avoided (and condemned) in most cultures and societies. This is termed inbreeding (or consanguineous mating) and has the consequence that an individual has fewer different ancestors in a certain generation than the maximum possible. Consanguineous matings are frequently made in the breeding of domestic animals (cows, horses).

The degree, intensity or closeness of the relationship is, in general, proportional to the number of ancestors which the two individuals have in common, out of the maximum number they might possibly have in common. Each time cousins marry, duplication occurs in their descendant's ancestral trees. The farther back one traces any individual's genealogy the greater the rate of duplication grows, until eventually there is more cousin intermarriage than input from unrelated individuals. At that point the shape of one's ancestral tree stops expanding and begins to shrink.

It is of interest to quantify the degree of inbreeding, as measured by the extent of reduction in numbers of different ancestors. There are several ways to do this, some of which are complex. Using a few illustrative examples I present here a simple procedure for calculating the number of ancestors at a certain generation in the past. Pointers are subsequently provided for readers interested in more in-depth approaches, as the present paper only serves as an introduction to issues that are, in fact, quite complex.

a. A simple approach to perturbations in the family tree: the π coefficient

Ignoring inbreeding, if we assume that an individual's generation is $g=0$, that his parents' generation is $g=1$, his grandparents' is $g=2$ etc., the number of ancestors at generation g will be given by $N_g = 2^g$. At the next generation in the past, we have

$$N_{g+1} = 2N_g = 2^{g+1} \quad (1)$$

To account for the reducing effect of inbreeding on the number of ancestors at a given generation, the simplest way is to modify the right hand side of equation 1 by including a perturbation coefficient in the form of a numerical factor $\pi \leq 1$. Thus:

$$N_{g+1} = 2\pi N_g \quad (2)$$

If $\pi = 1$ the population is ideal (exponential); If $\pi < 1$, the population is 'perturbed' and decreases by inbreeding; And if $\pi = 0.5$, the number of ancestors remains constant from one generation to the next ($N_{g+1} = N_g$), which is the extreme case of brother-sister mating. Values of π for various intermediate levels of inbreeding can be easily calculated as follows.

- (i) Parents are 1st cousins: as discussed earlier (Fig. 1), grandparents ($g=2$) are siblings, thus there are only six great grandparents ($g=3$) instead of the theoretical eight: at $g=3$, $\pi = 6/8 = 0.75$.
- (ii) Parents are 2d cousins: thus grandparents ($g=2$) are 1st cousins, great grandparents ($g=3$) are siblings, thus there are only fourteen great great grandparents ($g=4$) instead of the theoretical sixteen: at $g=4$, $\pi = 14/16 = 0.875$.
- (iii) Parents are half-siblings: thus **a** and **b** share one parent (**d**), and have only 6 great grandparents instead of 8 (Fig. 4). Therefore, in terms of the value of the perturbation coefficient π , at $g=0$, $\pi = 1$; At $g=1$, $\pi = 1$; At $g=2$, $\pi = 3/4 = 0.75$; And at $g=3$, $\pi = 6/8 = 0.75$. This result is almost similar to the case where parents are 1st cousins, the population decreases by inbreeding starting at the second ancestral generation (instead of the third for 1st cousins).

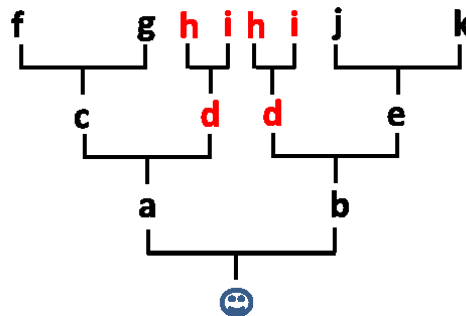


Figure 4: Ancestor tree of half siblings **a** and **b**, who share one parent, **d**. There are only 3 grandparents instead of 4 at $g=2$, and 6 great grandparents instead of 8 at $g=3$.

- (iv) Mother marries her uncle: this is the simplest example of cross-generation marriage and is slightly more difficult to visualize. The situation is depicted first in a descendant tree (Fig. 5, left) where **b** and **c** are the parents of **a** and **d**, and **d** and **e** are the parents of **f**. The corresponding ancestral tree (Fig. 5, right) shows that it is not possible to know that a marriage is of uncle and niece until the great grandparents' generation is reached: the common ancestors are **b** and **c**, but this is not known until the third ancestral generation is reached. At $g=0$, $\pi = 1$; At $g=1$, $\pi = 1$; At $g=2$, $\pi = 1$; but at $g=3$, $\pi = 6/8 = 0.75$ because **b** and **c** already

appear in the tree. In terms of the value of the perturbation coefficient π , this result is similar to the case where parents are 1st cousins, the population decreases by inbreeding starting at the third ancestral generation.

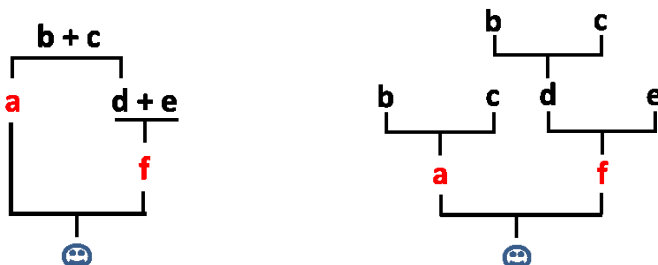


Figure 5: Left: Descendant tree showing that **a** is marrying his niece **f**. Right: Corresponding ancestor tree showing that when **a** and **f** marry, the common ancestors are **b** and **c**. However there is no evidence that these individuals are at all related (that **a** and **f** are uncle and niece) until the common ancestral generation is reached.

Clearly the perturbation coefficient π may vary from one generation to the next depending on the degree of inbreeding at each level. Table 1 shows the results of a simulation for the effect of varying π on the number of ancestors. In this hypothetical example, inbreeding starts at generation 10 and the number of distinct ancestors grows in a much milder-than-exponential fashion. As the perturbation coefficient decreases towards its minimum value (0.5) due to increasing inbreeding in the population (for example, in small isolated populations or exclusive religious communities)³, the number of ancestors stabilizes (generation 28) at one percent only of the perfect tree value.

b. More advanced models

The relevance of mathematical definitions for the coefficients of inbreeding and for the state of relationship (or kinship) between individuals has originally been brought out by R. Pearl (1917a, 1917b), and further discussed by T. Ellinger (1920), S. Wright (1922), and Ohno (1996). Later, instructive discussions were presented by B. Pears in a lighter style (1991a, 1991b, 1991c).

³ In the case of Jews, for example, which today number about 13.5 million individuals, it is estimated that in 1840 and 1940 they amounted to 4.5 million and 16.5 million, respectively. The latter figure represents the largest number of Jews ever living on earth. Source: http://en.wikipedia.org/wiki/Historical_Jewish_population_comparisons.

Table 1 – Comparison between the theoretical number of ancestors found in an unperturbed ancestor tree and a perturbed tree, with varying degrees of perturbation (π) throughout the generations.

Generations back	Theoretical number of ancestors (perfect tree)	π	Actual number of ancestors (perturbed tree)
0	1.00	1	1.00
1	2.00	1	2.00
2	4.00	1	4.00
3	8.00	1	8.00
4	16.00	1	16.00
5	32.00	1	32.00
6	64.00	1	64.00
7	128.00	1	128.00
8	256.00	1	256.00
9	512.00	1	512.00
10	1,024.00	0.990099	1,013.86
11	2,048.00	0.980392	1,987.96
12	4,096.00	0.970874	3,860.12
13	8,192.00	0.961538	7,423.31
14	16,384.00	0.952381	14,139.65
15	32,768.00	0.943396	26,678.58
16	65,536.00	0.934579	49,866.50
17	131,072.00	0.925926	92,345.37
18	262,144.00	0.917431	169,441.04
19	524,288.00	0.909091	308,074.62
20	1,048,576.00	0.833333	513,457.71
21	2,097,152.00	0.769231	789,934.93
22	4,194,304.00	0.714286	1,128,478.48
23	8,388,608.00	0.666667	1,504,637.97
24	16,777,216.00	0.625	1,880,797.46
25	33,554,432.00	0.588235	2,212,702.90
26	67,108,864.00	0.555556	2,458,558.77
27	134,217,728.00	0.526316	2,587,956.60
28	268,435,456.00	0.5	2,587,956.60

IV. Where family trees and crystals meet

As a materials scientist I find it tempting to draw the attention of the reader to a conceptual resemblance between the structural defects present in the otherwise perfect atomic array of a crystal and those arising in a perfect ancestor tree. Crystalline solids have a periodic crystal structure: atoms or molecules are positioned at repeating fixed distances, determined by the shape and size of a typical unit cell. Every particle has identical surroundings. There are various types of (perfect) crystals –cubic, hexagonal etc- but there is only one type of (perfect) ancestor tree, the repeating triangular array formed by an individual and his two parents. Stated differently but equivalently, every type of crystal has a typical unit cell (the fundamental unit that repeats itself by translation through the crystal) which differs from crystal type to crystal type, but a family tree has only one kind of repetitive unit, namely the child-parents triangle. In crystalline arrays, the ‘dots’ (atoms) are linked together by interatomic forces. In family trees the links between the ‘dots’ (individuals) are genetic links. Finally, a crystalline unit cell is 3-dimensional (crystals expand in all directions) whereas family trees are 2-dimensional only.

What happens when perturbations or defects are introduced in such ideal arrangements? In real materials the presence of defects critically determines many of the electrical and mechanical properties. As an example, without defects semi-conductors –thus, transistors- could not exist. In many cases, the presence of a minute number of structural faults can change a property (such as conductivity, diffusivity, strength etc) by many orders of magnitude. In family trees, if perturbations did not exist in ancestral trees the original human population on earth would theoretically be astronomically large (neglecting epidemics and wars): there would have been zillions of Adams and Eves instead of just one generic couple (according to traditional belief).

In crystals there are various types of defects but I will only mention the simplest ones, those that seem relevant here. A first crystalline defect, the vacancy, is a site in the array which is vacant instead of being occupied. Basically, it is a hole in the structure. The equivalent in a family tree would be the unknown father of a child, but in this case the entire ancestral line of that unknown father would also be missing from the tree structure. Since materials are never 100% pure, atoms of a different type are sometimes present in a regular atomic site in the crystal structure. This is a second type of defect, called a substitutional defect. This atom is not supposed to be anywhere in the crystal, and is thus an impurity. The equivalent in a family tree would be the adoptive father of a child, but in this case the entire ancestral line of that adoptive father would also have to be included in the tree structure (unless the researcher limit his/her tree to purely genetic links). A third type of crystal defect is the most interesting and the analogy in the context of the ancestral family tree is likely the most relevant: the so-called

‘edge dislocation’. An edge dislocation, shown in Figure 6, is caused by the termination of a plane of atoms in the middle of a crystal. Alternatively, an edge dislocation can be viewed as a defect where an extra half-plane of atoms is introduced mid-way through the crystal, distorting nearby planes of atoms. The equivalent in a family tree would be the disappearance of an entire line due to a marriage between cousins, as explained earlier (see Figure 1), or the appearance of an entire line due to a second marriage or to an adoptive family.

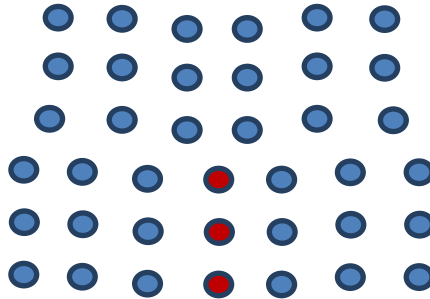


Figure 6: Edge dislocation in a crystal

The effect of defects in an otherwise perfect crystal, and of perturbations in an ideal structure of a family tree, may be felt up to a certain distance away from the defect. This distance defines the ‘strength’ of a perturbation. The accumulation of defects in a crystal leads to radical changes in the physical properties of the crystal as a whole. Similarly, the accumulation of perturbing factors in a family tree leads to strong variations in the growth pattern of the array, and eventually to the diamond-shape macroscopic tree of humans.

V. Conclusions

We have presented a simple way to quantify the effects of faults in family trees. A numerical parameter, the perturbation coefficient π , was introduced to provide a measure of the amount of ancestral branch inbreeding in that individual's family tree. Simple examples are presented and briefly discussed. An analogy with the occurrence of faults in crystalline structures in the field of materials science was suggested. The present article is introductory in nature, in the spirit of this booklet. References to more advanced quantitative discussions and theories are included for the interested reader. It is hoped to publish a more comprehensive paper on this subject in the future.

REFERENCES

- Ellinger, T. 1920. On the Numerical Expression of the Degree of Inbreeding and Relationship in a Pedigree. *The American Naturalist* 54 (635), 540.
- Ohno, S. 1996. The Malthusian parameter of ascents: What prevents the exponential increase of one's ancestors? *Proc. Nat. Acad. Sci. USA* 93, 15276.
- Pearl, R. 1917a. Studies on Inbreeding. VII. *The American Naturalist* 51 (609), 545.
- Pearl, R. 1917b. Studies on Inbreeding. VIII. *The American Naturalist* 51 (610), 636.
- Pears, B. 1991a. Pears, B. Our Ancestors, Conceptions, Misconceptions and a Paradox. <http://www.bpears.org.uk/Misc/AncestorParadox/index.html>.
- Pears, B. 1991b. Pears, B. The Ancestor Paradox Revisited. <http://www.bpears.org.uk/Misc/AncestorParadox/ancpdxrv.html>.
- Pears, B. 1991c. Pears, B. The Ancestor Paradox Yet Again. <http://www.bpears.org.uk/Misc/AncestorParadox/ancpdxya.html>.
- Shoumatoff, A. 1985. *The Mountain of Names*. Simon & Schuster.
- Wright, S. 1922. Coefficients of Inbreeding and Relationship. *The American Naturalist* 56 (645), 330.

BIOGRAPHY

Daniel Wagner holds the Livio Norzi Professorial Chair in Materials Science at the Weizmann Institute of Science in Rehovot, Israel. He is the author of 230 scientific papers, 7 book chapters, and 30 genealogical papers. Wagner has researched his Polish roots since 1995, and is a member of the Israel Genealogy Research Society (IGRA). He was a Co-Chairman of the 24th International Conference of Jewish Genealogy held in Jerusalem in 2004. He is a member of the JRI-Poland Board, the Coordinator of the Grodzisk Mazowiecki Archive Project, the Zdunska Wola JRI Shtetl CO-OP Coordinator and Town Leader, and until recently the Chairman of the Organization of Former Residents of Zdunska Wola in Israel. In September 2001 he and Kamila Klauzinska initiated the "Photographic Census Project" in the Jewish cemetery of Zdunska Wola, which took 7 years to complete. Wagner was subsequently elected Honorable Citizen of Zdunska Wola. A similar project was recently completed by him in the cemetery of Zyrardow.

Wagner was the recipient of the 2009 Yakir (Fellow) Award from the Israel Genealogical Society. He is a member of the Management Committee and of the Academic Committee of the International Institute for Jewish Genealogy and the Paul Jacobi Center.