**Fearless Symmetry**

*Exposing the Hidden Patterns of Numbers*

*A.Ash & R.Gross*

September 19, 2012

Can you make Wiles proof of Fermat's Last Theorem (FLT) accessible to the general public? Obviously not, and for most people they could not care less. Still people intrigued by the FLT make up a large public, at least by the standards of mathematicians, a public to a large extent made up by colleagues in other fields as well as by students of mathematics. Singh has addressed this public very successfully, writing not so much to impart mathematical understanding as to evoke the excitement mathematicians may feel. The authors of the present book has another ambition, even if they want to impress the reader with the fact that mathematics is a subject of discovery and far from a boring closed endeavor. The nature of mathematics they identify, following the lead of Hardy, with the looking for patterns and their explanations. Number theory is a good subject to illustrate this, because so many beautiful patterns can indeed be presented with little prior preparation. Now as authors of popular book on mathematics with the ambition of being honest and present mathematical arguments they face a dilemma. How to avoid becoming yet another text-book. Text-books you read not for fun but to get instructed. Text-books are in fact interactive, presenting exercises that you need to go through to get a hands on experience and to anchor the concepts, while a popular book you expect to read like a novel, as books by Singh and Gleick and others are designed to do. So the standard question is to whom is the book addressed? The obvious answer would be the intelligent layman. In mathematics it does not mean anybody, but somebody who at least have had calculus and hence acquired some minimal amount of mathematical maturity. In practice this means a student who is well versed in high-school algebraic manipulations[1]. Still the authors decide to include even the most basic things such as matrices in an attempt at being self-sufficient. In principle an ignoramus would logically be able to understand the material, but understanding in that sense can mean nothing more than a formal understanding. The only potential reader who could possibly benefit from this, would be a very young math whiz, who by lucky accident comes across the book and with the avidity of a young mind, readily and quickly absorbs new concepts. Because only among the very young can we expect the assumed mixture of ignorance and intelligence. I think that the authors could have done well by accepting the restraints on the readership they assume in their preface (but maybe they had to sell the idea of a wider readership to the publishers) and thus do away with much of the elementary material and the exercises to achieve some minimum of personal intimacy. After all this is not going to be a text-book after all.

[1] Many people are apparently stymied by the most elementary notions in algebra. This is cause for pity bordering contempt, and is deplorable, because it indicates a certain cognitive inability for abstract thinking. Such people are also for ever debarred from getting the barest understanding of natural science.

This does not mean that one should not dwell on elementary material, but it should be understood that the reader already knows it. There is nothing inherently bad being told what you already know, it can in fact if clearly presented be a pleasure[2], just as it can be a pleasure to repeatedly read a fairy-tale you have heard so many times before, or to once again be reminded by some lines of poetry. In fact a popular book like this can most beneficially differ from a textbook by instead of presenting the unfamiliar rehearsing the familiar.

Now to discuss the book in more detail. It begins already in the preface with a howler, by referring to the Dane Abel. True, Abel was born as a subject of the Danish King and died as a subject of the Swedish, but foremost he was Norwegian. In fact he may be the only mathematician in the world nationally known! With the Abel prize established in the early 2000 it is inexcusable not to know his provence in 2006 (the year of the publication of the book). True, later on you encounter him a again, but now as a Norwegian. Most likely the preface was the result of a slip, but there are slips and slips, to refer say to George Washington as a British general bespeaks a more than acceptable absent-mindedness in a student of American History.

The authors make a big thing about representations. I do not think that this is in any way a stumbling block, it may be so in the formal way it is presented, instilling in the reader an expectation of unforeseen problems. There are enough conceptual problems in mathematics as it is for new ones to be introduced. The musings on counting is pleasant enough, and I have myself played with the idea of writing something similar for a truly wide audience, but is it necessary in a book like this, with such constraints on space? True, some more readers may be enticed to follow along, but they will soon be dropped by the wayside anyway. So my advice to the authors would be, 'cut the crap' and go directly to groups. SO(3,$\mathbf{R}$) is indeed a very good example that everybody should be familiar with, but it should be written with the understanding that the reader is already familiar with linear transformations, matrices and orthogonal maps, but may not have realized many elementary and illuminating facts which are normally not included in textbooks. The authors should be on the look-out for such. I find it personally illuminating to look at say orthogonal matrices with integral coefficients. They make up nice, finite subgroups, which can easily be described. This of course gives right away a clue to natural linear representations of the permutations groups, in terms of permutation matrices (which the authors of course note). When it comes to linear algebra I think that more emphasis should be given to the characteristic equation, especially in degree two, because traces and norms play very important roles further on in the book. As to trace, it is interesting to observe that the trace of a permutation matrix is given by the number of fixed points. This is a discrete baby-version of the celebrated Lefschetz-fixed point formula on co-homology, and I think a suggestive one, which is normally not mentioned in textbooks on topology. Absolutely crucial is the notion of conjugation for linear maps, the authors mention this en passant further on in the text, realizing that this is a crucial notion and hoping that the reader is already familiar with it. This is an afterthought and illustrates the problems of popular writing when you are in the watering down process. You are familiar with

[2] However elegance is a must, there are few things more exasperating than to be presented a clumsy presentation of things you already know. If you do not know it, you may not be aware of the clumsiness.

the standard arguments and you realize how many of those may not be familiar to the reader, and you take them out, or at least make a brief apology. A popular text brought about from the top by watering down and removing technicalities, does in the end run the risk of being as incomprehensible to the layman as to the expert. A popular text should be written from the bottom, by rethinking the material from a new perspective, in practice meaning that of the frog. Now to return to the characteristic equation, admittedly the trace is only one coefficient, but if you know the traces of the first powers, you can reconstruct the characteristic equation. I think that it is also unavoidable to talk about character tables, they explain why matrix representations are so useful, far more useful than representations by permutations. Now group representations are not part of the introductory courses, and many students who may be familiar with group theory may not have yet encountered it[3]. I think it can be worth a couple of pages to present the yoga of representation theory, how you can actually formally in simple cases work out the characters of an irreducible representation without having any clue on how to actually exhibit it. This is somewhat magical. Of course you do not need it all, but it builds understanding by stimulating questions. The problem with a mathematical proof, or at least one of its problems, is its economy. Only what is strictly necessary goes into it. This is a triumph of compelling logic, but devastating to understanding. Too often a proof is in the nature of a verification or confirmation, too seldom an exploration of ideas. If you really understand a few key concepts and some relations between them, a proof often falls into place by itself. In a book that is not supposed to be a systematic textbook, the authors are given a lot of latitude. The formal definition of an abstract group should of course be given, the so called axioms of a group, but it should be made clear that this does not mean an extension of the concept, that in fact by representation we can present any abstract group as a subgroup of a very concrete group[4]. The point of an abstraction is economy and the ability to single out key features which may be obscured by too much extraneous detail. The notion of conjugacy classes is fundamental to group theory and it is important to see this manifested for matrices groups (matrices corresponding to different basis) and symmetric groups (the cycle type of a permutation gives its conjugacy class). Now matrices over finite fields combine in an intriguing way linear algebra with elementary combinatorics, particular permutations. The authors consider the case of $GL(2, F_2)$ (or in fact $SL(2, F_2)$ or even $PSL(2, F_2)$ as identical groups), but fail to point out at an early stage that this is actually $S_3$. This cute fact I think every mathematician should know[5]. The conjugacy structure of the groups $GL(2, F_q)$ (and related groups) can be done by linear algebra over finite fields. To work this out would take too much space and it would not be necessary for the future development of the text, but it is important that the readers appetite is stimulated and that he becomes aware of the possibility of so doing. Also the structure of finite fields should probably be high-lighted early on. The fact that

[3] I did not become familiar with it until I was a professor, in retrospect this is incredible. I was once told by McKay that Barry Mazur was unfamiliar with character tables, I find that very hard to believe. On the other hand we should always be open to the counter-intuitive.

[4] Just as a ring is just a subring of an Endomorphism ring of of a $Z$algebra.

[5] Incredible enough $GL(2, F_2)$ can be embedded in $PSL(2, \mathbf{C})$ the Moebius transformations, important for going from $\lambda$ to $j$ in the theory of elliptic curves.

the invertible elements form a cyclic group, and hence that there are generators, is crucial. This is not so hard to prove (one really only needs to make the observation that over a field $n$ roots give a polynomial of degree $n$ so that $x^n - 1$ has at most $n$ roots from which we can leave the rest to the reader). To give a hint of a proof is better than to say that it can be proved or it is hard to prove. Many proofs are omitted in the text, and that is a service for the reader who does not have to go through some logical details, but one should know something of what one is missing. Some of them are in fact easy and within the scope of the reader, others are far more involved and would necessitate long digressions. It is useful to know what is what. Understanding that the group is cyclic makes sense of so many things for finite fields. As to quadratic residues, it makes sense to see them as a subgroup of index two, and also maybe to point out that when written consecutively (rather than in exponential form) they seem to be haphazardly distributed, yet they satisfy many striking features when translated. Thus finite fields give rise to many intricate combinatorial configurations. This is not essential to the proof by Wiles, if this is the ultimate goal, but neither is the digression on Latin Squares. Also you can ask when is a number a cube, and realize that if $3|(q-1)$ then every third element is a cube, otherwise every element is a cube. In the latter case the group homomorphism $x \to x^3$ is 1:1 and hence $x^2+x+1$ has no roots, i.e. is irreducible. Do the same thing for $n$th powers and you see that if $n|(q-1)$ every $n$th element is a $n$-power and hence that there are $n$ solutions to $x^n = 1$ meaning that the cyclotomic splits into linear factors. By taking field extensions you can soup things up, and in fact note that if $q$ is a generator for $F_p^*$ then the $q$-cyclotomic polynomial is irreducible in $F_p$ an conversely[6]. Some of this is actually used further on in the book. Also the fact that $x^2+1$ has a solution iff $p = 1(4)$ follows from the fact of cyclicity[7] and thus this fact is also within reach of the reader. And finally I think it is close to criminally negligent not to mention the fact of the Frobenius $x \to x^p$ being an automorphism, in fact an automorphism par excellence and in fact we have a very easy description of the Galois group of a finite field. People may know this, but still it cannot hurt to be reminded of this magical fact, especially as it is going to play a very important role later.

As to Galois theory I think it is a mistake to discuss $\overline{\mathbf{Q}}$ from the start. In practice what we need is to understand the local picture, namely galois groups of finite extensions of $\mathbf{Q}$. That all those local data fit together into a whole, which can be codified into the general Galois group, may be nice but not so useful after all. The group is easy enough to understand, but the fact that it is impossible to write down explicit elements in it, except the identity and conjugation is food for thought. Of course it has to be mentioned, but it can be more profitably discussed after an understanding of traditional Galois theory has been achieved. And here we have a dilemma, similar to that of group representations, how many readers are familiar with Galois theory? To learn it takes a textbook and usually a few weeks (as with representation theory) but unlike many other theories it can be

[6] Or did I permute the primes $p, q$ anyone given those hints could work it all out. I discovered this well-known fact when I was giving a course on elementary number theory at UCLA in the fall of 1985. I was quite proud of it, as I had never thought of anything similar before, and I did bring finite extensions of the finite fields into play, exploiting the fact of the cyclicity of their invertible elements

[7] Of course the case $p = 2$ is an exception.

encoded very succinctly in the main theorem of Galois theory and its main object - the Galois group can easily be defined. Admittedly this only gives a formal understanding of what it is about, on the other hand you cannot expect more and at least in a formal sense you need not know more. The connections to the solution of equations is historically very important, but really as far as Galois theory is used an almost irrelevant digression. (The authors realize that they are last starting to write a textbook on Galois theory when they get to some nitty-gritty examples.) In the case of Galois theory I am tempted to make a few elementary remarks as how it relates to linear algebra. Given a finite field extension $L : K$ (of degree $n$) then for each $\alpha \in L$ the map $x \to \alpha x$ is linear, hence it has a matrix representation which is not so hard to figure out. It also gives a representation of $L$ into the endomorphism ring of $L$ considered as a $n$-dimensional vector space over $K$. In fact $L$ is the largest field you can find in $M(n, K)$ and $L^*$ is the largest commutative subgroup of $GL(n, K)$ but this is a digression. Anyway this kind of representation is well-known in the case of the complex over the reals (or the Quaternions over the reals or the complexes) and in this way we note that the polynomial that $\alpha$ satisfies is in fact the characteristic polynomial of $\alpha$ as a linear map. (The additive of trace and multiplicativeness of norm then becomes obvious, the letter just being the multiplicativeness of the determinant with which the reader must assume to be familiar). Furthermore an automorphism $\sigma$ is also a linear map, represented by some matrix, its action on the elements of $L$ considered as linear maps, should be by conjugation, and hence we get the multiplicativeness automatically, and the Galois group as a finite group of elements which conjugates $L$ into itself. Whether this is is really useful or not is another matter, but it is I think natural to think in those terms, and especially as it gives ways of representing the Galois group linearly.

Now given a Galois group of a polynomial with integral coefficients, it is natural to consider this reduced mod various primes. Is there any way of relating the Galois group of the former to the very simple Galois group over a prime field? This is a very natural question, the answer may be a bit more involved, but at least would naturally bring in the Frobenius. The way it is done by the authors is of course very economical but seems ad hoc and pulled out of a hat. What does it come from? Where do things come from, this is a perennial question when you do mathematics, and textbooks are not usually very good at that, they just throw things at you, with the tacit assumption that in due time you will understand what it is all about. A book like this which is not meant to be a textbook but to help the reader to see the forests and not just the trees, which is usually the first effect of a textbook.

The book becomes rapidly more technical at the end and one feels that the authors are throwing up their hands in despair. In fact the last few chapters are not so much addressed to the reader as much as the authors who are struggling to understand in their own terms what Wiles proof is all about. This is fair, especially when it comes to the end. This is standard practice in giving math lectures, you should start out slow so everyone in the audience can latch on to something, and in the process you hope that everyone will get something to take home with them, it may not be what the lecturer intended to convey, but that does not really matter. If you get something at least, it does not really matter if you get lost in the end. Somehow or not, your motivation for what motivates the lecturer may have been slightly enhanced and the next time you go to a similar lecture, you

have at least an illusion of a better understanding recognizing things from the past. The command and understanding of a mathematical concept, not even the most elementary, is never completed. Understanding grows with every application of the concept in a new setting. It is a common experience of mathematicians that they often realize that they have not really understood concepts they have taken for granted when getting confused in a mathematical argument. Confusion in mathematics is a godsend, it usually provides an opportunity for illumination.

Now at the end let me confine myself to a few isolated remarks:

First one should perhaps really emphasize the analogy of torsion-points on an elliptic curve with roots of unity on a circle. It is more than a mere analogue in fact. The point is that while the latter just comes in one version, all circles are the same, the latter comes in many, because elliptic curves differ. (In fact this is exploited by modern algorithms of factorization.). Also for every elliptic curve defined over a finite field the points over the field make up a finite group, and one of a very simple structure. (In particular all the points are torsion-points!). It is easy to find the 2-torsion points, as they are the zero of the Weierstrass cubic. And we get three cases depending on one, two or three roots, given subgroups of order 1,2 and 4 respectively, as order of subgroups divide the order of the group, we get simple criteria for when the number of rational points is even or not. The arguments that the authors provide are a bit involved and I guess essentially reprove Lagrange's theorem.

It seems a bit perverse to give a technical definition of ramification without having recalled for the reader the archetypical cases of $\mathbf{Z}[\mathbf{i}]$ and $\mathbf{Z}[\rho]$, the Gaussian and Eisenstein integers respectively.

The notion of $\mathbf{Q_p}$ is a very natural one. The authors are obviously burning with desire to digress, but by circumstances reduced to the intermittent footnote. The essentials could be conveyed in half a page. It comes up very natural in connection of going backwards from reduction modulo $p$. If for each $p^m$ we can solve something, what about in the end? We get an infinite expansion, as we are used from long division in elementary school. The principle is very similar. Such things may or may no be properly explained in a textbook, but usually not. The concept is thrown into your face and you have to take it or leave it. In the end if you stay around in the subject you will take it. If a p-adic expansion is periodic is it then rational? This is a natural question which I do not think is usually framed in an elementary introduction. This is one of those things that may enhance a readers appreciation, even if he will not be able to go the whole way to the end of the book.

Modular forms and q-expansions are very summarily treated. This is a pity, because then much of the magic is lost. Modular forms easily takes up a whole book by itself, yet I think two or three pages may be plenty to intrigue the ignorant reader. It seems very hard to come up with non-trivial examples, and the once we get seem to have dropped down from heaven. In the book some such examples are given, as the infinite product. That you would get such forms from the data of number of points on reduction on elliptic curves from different $p$ seems just incredible. One should establish that modular forms have an independent history and that those particular connections came later. It shows the uncanny interconnectedness of mathematics, which is truly one of the features of mathematics the

authors would like to convey.

When it comes to co-homology, why not mention the Lefschetz fixed-point formula and how it relates to the expression $1 + q - \#E(F_q)$ It is easy to show that the average number of points on an finite elliptic curve is $1+q$ by considering a cubic and its associated quadratic twist. Furthermore in the Weierstrass equation we have two solutions or none with a probability of one half each (except for the obvious cases of its zeroes), however by the Hesse bound things cannot go out of hand, as it would do in a real stochastic situation. Why this is so is something of a mystery.

I think most readers will have trouble with the last chapters. Until then a knowledgable reader could just race ahead and rapidly read on, confident that when the presentation occasionally gets tedious he can just skim and skip confident that he at least will miss out on nothing. This is not true at the end, which calls for another kind of skipping. A skipping and skimming of despair. As noted the last chapters are written for the benefit of the authors rather than the readers. This does not mean that this is not true for the elementary parts, any presentation is beneficial to the understanding of the one who does it, but in the end there is a struggle and a lack of confidence which is not there on the more elementary material. But as noted, this is inevitable. I think it is important that the book has a distant goal, even if it would be beyond the reach of most of its potential readers.

I have suggested a lot of material to put in, and that is easy to do. Most likely much of what I suggested has no doubt appeared in earlier drafts of the book, and the authors have been forced to make severe cuts in order not to make the bulk of the book forbidding both to the reader and the publisher alike. In making cuts the writers have been advised not to touch the more elementary material, but instead to slim the more advanced. As I argued initially that may not have been the wisest choice, the impact of the book may have been maximized if it had been directed to the middle ground. Ignoring the ignorant as well as the expert (although a well-written and elegant presentation is bound to appeal to the expert, even if he does know everything from the start, and in fact even if he does, he may get some inspiration nevertheless) and concentrate on the colleague in another field or the serious math student who may have learned the botany but not yet figured out how things fit together ecologically. But then of course the book would not have warranted the sympathetic attention of a publisher and given its rather fancy appearance with a potential of appearing on the shelves of general bookstores, and instead been relegated to yet another specialized tract. The fact that a colleague gets inspired from reading it to suggest how he could have done it much better, is actually an acknowledgment. Any mathematical presentation can of course be improved, just as any number can be added one to. But this only shows that mathematics is steadily improving, and by inviting improvement the text shows its worth. An indifferent text does not inspire any such measures, such a one you only want to forget. So in this sense at least the book is unforgettable. The authors express a pious hope that there will be a second edition, I fear that the publishers will not press for this. That is a pity.

Finally the book is written with a measure of urbanity not usually brought to a mathematical text[8]. At least one of the authors has an interest in philosophy and a

[8] An exception would be that of Miles Reid who usually indulges in the sarcastic comment and the

concomitant predilection for the philosophical aside. This makes for pleasant reading and I also personally find this important. It is a pity though that at the end the authors feel so stressed that they no longer can allow themselves the initial pedestrian pace. It is very important to explain why you do something in mathematics as showing the details of what you should do. If you understand why you want to do something, you usually are able to do it. This of course presupposes that you give a mathematical reason compatible with your level. To say that you should solve the Riemann Hypothesis so you can win a million dollar does not get you any closer. In fact, which is seldom if ever pointed out, there is actually no harder way to earn a million dollars than through the Millennium prizes, and this holds also for mathematicians, no matter how clever.

Are the last chapters totally incomprehensible? Of course not, the general strategy is pointed out by the authors. That the existence of an integral solution to the Fermat equation would lead to the existence of an elliptic curve who would have certain strange properties namely the existence of certain modular forms that cannot exist. In fact if sufficiently high-minded one may think that the strategy is obvious and that the only thing that remains is to solve a few technical lemmas. There are easy to come up with similar strategies for the solution of the Riemann Hypothesis, but do they work? The proof is in the pudding, and what may appear as a technical breakthrough at times, is usually something quite more profound. Just one remark. Is it clear from the presentation that the proof breaks down for the exponents $1, 2$? This may be a good exercise for the reader to find out. If he cannot find out he has either understood the whole thing imperfectly, or the authors have not been honest enough.

One thing is clear, it is the road that matters not the goal. The FLT has no consequences for mathematics, it has just been a brainteaser for a few hundred years, and it has in the process given arise to much interesting mathematics. Somewhat ironically more in the past than now, in spite of Wiles spectacular achievement. In a sense the relation with the Fermat equation was something of a quirk, once this connection was noted, the Fermat equation could safely be forgotten. It served no intrinsic mathematical purpose. But mathematicians are human beings, they are motivated as well as others by extra-mathematical motives. The sweet rewards of glory are no strangers to them. Once Wiles realized that a successful resolution of some conjectures that were floating around in his chosen field of expertise would be achieved, the reward would not be a million dollars, but something more elusive, namely a glory which would propel his achievement from being just one of specialized concern to one which would reverberate in the entire mathematical community and even beyond. Wiles could easily have earned a million dollars in a less strenuous and time-consuming way, but to achieve the kind of fame he did, would only be possible for him through that work. The fact that the FLT was the first unsolved mathematical problem he had encountered[9] adds to his story a poetic touch.

Finally whether other variations of the Fermat equations will be solved is I think of less interest. The same methods may be applied and then the achievement becomes one of technical bravado. Of course one may never know what a search may turn up, but the

---

illuminating aside

[9] This might be true for most people, in my case it was the question of the infinitude of primetwins that became an eye-opener to me.

natural problems are associated with elliptic curves and modular representation, obscure diophantine equations may turn up as a spin-off, but they play little if any role in the actual mathematics.

September 20, 2012 **Ulf Persson:** *Prof.em, Chalmers U.of Tech., Göteborg Sweden* ulfp@chalmers.se