TME Volume 4, Number 2

THE MONTANA MATHEMATICS ENTHUSIAST

[ISSN 1551-3440]
Vol.4, no.2, pp.138-267

AIMS AND SCOPE

The Montana Mathematics Enthusiast is an eclectic journal which focuses on mathematics content, mathematics education research, interdisciplinary issues and pedagogy. The articles appearing in the journal address issues related to mathematical thinking, teaching and learning at **all** levels. The secondary focus includes specific mathematics content and advances in that area, as well as broader political and social issues related to mathematics education. Journal articles cover a wide spectrum of topics such as mathematics content (including advanced mathematics), educational studies related to mathematics, and reports of innovative pedagogical practices with the hope of stimulating dialogue between pre-service and practicing teachers, university educators and mathematicians. The journal is also interested in research based articles as well as historical, philosophical and cross-cultural perspectives on mathematics content, its teaching and learning.

The journal is accessed from 94+ countries and its readers include students of mathematics, future and practicing teachers, university mathematicians, mathematics educators as well as those who pursue mathematics recreationally. The journal exists to create a forum for argumentative and critical positions on mathematics education, and especially welcomes articles which challenge commonly held assumptions about the nature and purpose of mathematics and mathematics education. Reactions or commentaries on previously published articles are welcomed. Manuscripts are to be submitted in electronic format to the editor preferably in APA style. The typical time period from submission to publication (including peer review) is 7-10 months. Please visit the journal website at http://www.montanamath.org/TMME

Indexing Information
EBSCO Products; *Directory of Open Access Journals (DOAJ); MathDI/MathEDUC; Ulrich's Periodicals Directory*

The Montana Mathematics Enthusiast (ISSN 1551-3440)
Vol.4, no.2 (June 2007) pp.138-267

Editorial: New horizons-four years later

Bharath Sriraman, Editor
The University of Montana

Vol.4, no .2 signals the conclusion of this volume and four years of the journal's existence. Although we are still in our infancy in comparison to other journals in the field of mathematics education, there is no doubt that we have carved a niche in a highly competitive business where scholars and readers have numerous choices when it comes to supporting journals. The journals niche is the fact that it attracts scholars/contributors from a variety of domains such as mathematics, critical theory, philosophy, educational psychology, educational philosophy, social justice, teacher education and the history and philosophy of mathematics and science, in addition to practitioners at all levels. The articles that have appeared in the four volumes and the monograph are indicative of this niche. Journals survive or perish depending on the flow of manuscripts. In this respect we have been very lucky with a steady stream of high quality submissions as well as conscientious reviewers. We also continue to receive invitations for indexing, which speaks for the standing of the journal.

Over the last 16 months, we received 86 manuscripts from 26 different countries. The acceptance rate is currently 22-25%. In cases of rejection, we have supported the authors with extensive suggestions for improvement and other avenues for publication. In some cases, we reject manuscripts because the mathematics is too sophisticated and may not be accessible to the average reader. We are not a journal that publishes pure mathematics articles that require research level knowledge within a specific sub-domain of mathematics. Articles that involve mathematics are determined on the basis of whether or not they would appeal to advanced undergraduate/beginning graduate students of mathematics, practicing teachers and those that enjoy mathematics recreationally. In cases, when it is difficult to find reviewers for a particular manuscript, we ask the authors to provide a list of three possible reviewers. I am moving towards a review process which is open and constructive and beneficial to all parties concerned. One of the consequences of the increased flow of manuscripts is the "bottleneck" effect, i.e., reviewed and accepted manuscripts having to wait inline for publication. This problem can be circumvented by increasing the frequency of issues to 3/year. Having said that, starting from vol.5 we will be publishing 3 issues/year [February, June and October]. Among the issues in the pipeline (vol.5,no.3) is focused on statistics education for which manuscripts are still being submitted. Readers are encouraged to submit papers on this topic. We still are consistent with our goal of keeping the transit time from submission to publication to approximately 8 months.

***The Montana Mathematics Enthusiast*, ISSN 1551-3440, Vol. 4, no.2, pp. 138-139**

The readership base of the journal is now approximately 4200 from 94 countries based on statistically sieving uniqueness of IP addresses and average repeated visits. The geographic distribution of readers is as follows: 35%- North America; 18%- Western and Central Europe; 12%-Scandinavia; 12% -Asia; 8% -Middle East; 8%- Australia and NZ; 4%- Africa; 3% -South America. Needless to say we are thriving! The journal recently received another offer from a reputed publishing house for conversion into a print journal in addition to the electronic version. However accepting the conditions would have meant restricting access to paying subscribers, and putting an embargo on when the articles become available online through indexes. This goes against our philosophy of free and open access. So the offer was turned down.

The first monograph of the journal on social justice issues released in January this year was very well received. A telegraphic review appeared in the *Journal for Research in Mathematics Education* in May 2007, and full length reviews are forthcoming in *Mathematical Thinking and Learning*, as well as *ZDM- The International Journal on Mathematics Education*. I thank the reviewers for their gesture of writing these reviews. A limited number of print copies are still available for sale for the cost price ($20). We are hoping to use the proceeds to sustain print monographs on special topics or themes on an on-going basis and are open to suggestions on possible future looking topics.

This journal issue contains articles from well-known scholars as well as new doctoral recipients and those currently working on their doctorates. Again the sheer range of topics covered in this issue represents the true face of the journal. The articles in this issue will appeal to a wide audience: teachers of mathematics, university mathematics educators, philosophers, cognitive psychologists and naturally math enthusiasts.

The journal is happy to extend its support to Sense Publishers in the Netherlands, which publishes affordable and high quality books of interest to the mathematics and science education community as well as to the larger field of education. A new feature of the journal is to inform readers of new and noteworthy books published by Sense and sources for reviews of these books.

Finally, I hope you enjoy this issue and I thank you for your continued support.

Objective Truth versus Human Understanding in Mathematics and in Chess

Olle Häggström
Chalmers University of Technology[1]

Abstract

This paper begins with a review of the collection *18 Unconventional Essays on the Nature of Mathematics* edited by Hersh (2006). Inspired especially by the contribution by Thurston to that collection, I then go on to discuss, by means of a couple of thought experiments involving computer "oracles", the nature of mathematics as a human activity, hopefully providing some balance to to the simplified view (sometimes held by research mathematicians such as myself) of the discipline as purely a quest for objective truth.

Keywords: Philosophy of mathematics, Platonism, computer-assisted proof, thought experiment, chess

1 A stimulating collection of essays

To anyone who has experienced the inescapable force and logical necessity of a mathematical proof, the Platonic existence of numbers and their properties – independently of us humans who think and argue about them – is obvious. The first time I encountered this argument,[2] I felt immediately excited: Yes! That is exactly how it is. Surely this settles the old question of whether new mathematical results are discovered or invented!

My excitement didn't last very long, however, because on second thought I realized that the argument has the same structure as the following proof of God's existence, which I think is flawed: *To anyone who has met God, His existence can no longer be in doubt.* Clearly, if we are to take the question about the ontological status of mathematical concepts seriously, we need to do better than merely to refer to an intuitive feeling shared by many (probably most) mathematicians, and there is a lot to say about the subject.

Of course, a mathematician is not obliged to consider this issue, but if he or she chooses to do so, then a nice entrance gate to the literature on this and other topics in the philosophy of mathematics is the recent collection *18 Unconventional Essays on the Nature of Mathematics*

[1]Mathematical Sciences, Chalmers University of Technology, 412 96 Göteborg, Sweden,
e-mail: olleh@math.chalmers.se

[2]This was in Hacking (1999), although of course the argument does not originate from him, and in fact I even extrapolated a bit from what he writes, because all he says is that those who have *not* had this experience do not realize why a mathematician would be inclined towards Platonism.

The Montana Mathematics Enthusiast, ISSN 1551-3440, Vol. 4, No. 2, pp 140–153

edited by Reuben Hersh (2006). The present paper begins as a review of this collection. Then, inspired by what I have read, I will move on in Sections 2 and 3 to some thoughts of my own concerning objective truth versus human understanding in mathematics – and in chess.

All of the texts in Hersh's collection have previously appeared elsewhere. Most of them are fairly recent, but a few classics are also included, such as Alfréd Rényi's "A Socratic Dialogue on Mathematics", which is a very fitting opening of the collection, as it beautifully states (and even begins to answer) some of the basic questions that the other essays are concerned with, such as: What are mathematical objects? Do they exist? If they do not exist in the same ordinary sense as physical objects, how can it be that they are useful to the real world? And the aforementioned issue concerning new mathematical results: discoveries or inventions?

Although the matematicians, philosophers, and social scientists that contribute to this volume cannot be said to represent a common "school" of philosophy of mathematics, a kind of joint theme nevertheless emerges, namely the emphasis on distinctly human aspects of mathematics. Few of the authors seem to share my Platonic Hunch.[3]

Does it really matter whether Platonism is true or not, other than for the somewhat esoteric purpose of predicting whether or not, once we encounter an extraterrestrial species advanced enough to have come up with (say) radio astronomy[4] and interplanetary travel, they will be familiar with things like prime numbers, the Central Limit Theorem, and the Newton–Raphson method? Timothy Gowers, in his thoughtful contribution "Does Mathematics Need a Philosophy?" that contains many illuminating concrete examples, suggests at first that the answer is no:

> Suppose a paper were published tomorrow that gave a new and very compelling argument for some position in the philosophy of mathematics, and that, most unusually, the argument caused many philosophers to abandon their old beliefs and embrace a whole new -ism. What would be the effect on mathematics? I contend that there would be almost none, that the development would go virtually unnoticed. And basically, the reason is that the questions considered fundamental by philosophers are the strange, external ones that seem to make no difference to the real, internal business of doing mathematics.

Then, however, he goes on to balance this statement by pointing out, in the following pragmatic argument in favor of formalism, that the choice of philosophy may not be devoid of pedagogical consequences:

> If you are too much of a Platonist or logicist, you may well be tempted by the idea that an ordered pair is *really* a funny kind of set [of the form $\{\{x\},\{x,y\}\}$]. And

[3]In speaking of the Platonic Hunch, I am inspired by philosopher Daniel Dennett's talk of the Zombic Hunch (see, e.g., Dennett, 2005, or Blackmore, 2005). When philosophers speak of zombies, they mean imaginary creatures that are exact copies of ourselves, speaking and acting like us, the only exception being that zombies lack subjective experience. The Zombic Hunch is the intuitive inclination, shared by most of us, to think that zombies are possible, at least in principle. Dennett says they are not, granting only that the purely materialistic approach to explaining conciousness (identifying consciousness with certain physical processes in the brain) which he advocates also needs to explain the Zombic Hunch. Likewise, in my mind, any philosophy of mathematics that denies Platonism faces the challenge of explaining the Platonic Hunch.

[4]Or perhaps I should say astronomy based on electromagnetic radiation at wavelengths unavailable directly to their own sensory organs.

> if you teach that to undergraduates, you will confuse them unnecessarily. The same goes for many artificial definitions. What matters about them is the basic properties enjoyed by the objects being defined, and learning to use these fluently and easily means learning appropriate replacement rules rather than grasping the essence of the concept.

Inevitably in a multiple-contributor collection such as Hersh's, the quality of the contributions is somewhat uneven.[5] For instance, Gian-Carlo Rota's "The Pernicious Influence of Mathematics upon Philosophy" contains repeated condemnations of philosophers guilty of "mathematizing" their subject, without exemplifying that practice by even a single name or a single reference. That is poor scholarship. And Rafael Núñez' essay "Do *Real* Numbers Really Move? Language, Thought and Gesture: The Embodied Cognitive Foundations of Mathematics" suffers from the author's limited familiarity with the mathematics he discusses. One example is when he muses on the use of temporally loaded words like *approaches* and *tends* in statements like "the sum $s_n = \sum_{i=0}^{n} a_i$ approaches $s = \sum_{i=0}^{\infty} a_i$ as n tends to infinity" where no time dynamics is involved, without considering the obvious explanation that we tend to think about s as arising by *starting* with a_0, *then* adding a_1, *then* adding a_2, and so on.

When I set out to read *18 Unconventional Essays*, I was particulary curious about what the two contributors – Donald MacKenzie and Andrew Pickering – who, in view of some of their earlier writings,[6] belong firmly to what I perceive as the "other side" (i.e. the side of social-constructivists and postmodernists) in the infamous Science Wars[7], had to offer on the subject of mathematics. It turned out that they both give fairly reasonable and reader-friendly summaries of factual case histories: in MacKenzie's case mainly about formal verification in computer science, and in Pickering's about Sir William Rowan Hamilton's fascinating search for higher-dimensional extensions of complex numbers culminating in 1843 in the discovery of quaternions. MacKenzie reveals his position most clearly when he writes, concerning the existence of several different kinds of logic employed by contemporary computer scientists, that

> [i]t cannot be guaranteed that [this plurality] would remain intact in a situation where there are major financial or political interests in the validity or invalidity of a particular chain of formal reasoning,

thus suppressing the possibility that the inner "logic" (if I may use that word in an informal sense) of the science itself eventually reveals that only one of the logics (and here I revert to the strict meaning of the word) is in the end intellectually – as opposed to commercially or politically – rewarding. As to Pickering, he gives his position away at an early stage of his essay where he explains his theoretical approach. His goal is to understand why, in a subject such as mathematics that deals purely with abstract concepts that the researchers are

[5]The collection would also have gained from a more active editing and a more careful proofreading. Annoying examples abound, such as the reference on p. 57 to the non-existent Figure 1, or the confusing duplication of the figure on p. 261. Jody Azzouni's choice in "How and Why Mathematics is Unique as a Social Practice" to italicize every 20th or so word is likewise annoying, and should have been moderated by an editor.

[6]See, e.g., MacKenzie (1981) or Pickering (1984).

[7]See, e.g., the collection edited by Ashman and Baringer (2001) for a variety of views on these, or Weinberg (2001) for one that I find myself in basic agreement with.

free to define in any way they please, a researcher can ever run into problems. Why indeed. We mathematicians are so used to the fact that not all mathematical objects and definitions are equally fruitful and interesting, and that once the definitions are set we cannot do as we please with the resulting objects, that the question sounds utterly silly – whereas it is easy to imagine that to a sociologist of science working in the postmodern "anything goes" tradition it sounds quite reasonable.

Rényi, as mentioned above, touches upon the question of the "unreasonable effectiveness of mathematics in the natural sciences" that was so elegantly phrased in the classical paper of Eugene Wigner (1960), and a couple of other contributors – Eduard Glas, and Hersh himself – also find reasons to discuss Wigner. I feel, however, that both Glas and Hersh misunderstand the essence of Wigner's question.[8] Glas, in his contribution "Mathematics as Objective Knowledge and as Human Practice", writes that

> [t]he effectiveness of pure mathematics in natural science is miraculous only to a positivist, who cannot imagine how formulas arrived at entirely independently of empirical data can be adequate for the formulation of theories supposedly inferred from empirical data. But once it is recognized that the basic concepts and operations of arithmetic and geometry have been designed originally for the practical purpose of counting and measuring, it is almost trivial that all mathematics based on them remains applicable exactly to the extent that natural phenomena resemble operations in geometry and arithmetic sufficiently to be conceptualized in (man-made) terms of countable and measurable things, and thus to be represented in mathematical language.

Well, yes. But the true miracle – the one that Wigner is concerned with – is that nature is sufficiently ordered in such a way as to "resemble operations in geometry and arithmetic" and admit laws that can be "represented in mathematical language". Once that is the case, it is hardly miraculous that Darwinian evolution followed by human cultural developments equips us with the capacity to grasp these things.

Hersh, in a similar vein, portrays Wigner's miracle as being that of how often pure mathematics turns out to have profound applications in physics when no such application was originally intended. But again, the miracle of nature's amenability to mathematical description wouldn't go away even if it were the case that the mathematics were developed in direct contact with the physical application.

The best essay in Hersh's collection is, in my mind, William Thurston's "On Proof and Progress in Mathematics", which originally appeared as part of a discussion in the *Bulletin of the American Mathematical Society* in 1993–94 concerning the role of rigor and proof versus conjecture and speculation in mathematics, and the relation between mathematics and theoretical physics.[9] Thurston's essay lacks the philosophical ambition of most of the others, and instead provides a series of down-to-earth observations concerning how mathematics is done

[8] It has become a bit of a bad habit of mine to criticize other writers' understanding of Wigner's question; readers who know some Swedish may turn to Häggström (2004) for an earlier instance. In all these cases, it seems that I reveal my Platonic Hunch quite clearly.

[9] This discussion is very much worth rereading in its entirety; see Jaffe and Quinn (1993), Thurston (1994), Atiyah et el. (1994) and Jaffe and Quinn (1994).

and how mathematicians interact, drawing generously on the author's personal experience. What I read as his main point is that the view that "progress made by mathematicians consists of proving theorems" is a harmful simplification, resulting in a malign one-sidedness towards considering *theorem-credits* in the way we evaluate each other's qualifications. Progress in mathematics, according to Thurston, consists of improving our *understanding* of mathematics, and this is a much broader goal than mere theorem-proving. As one illustration to this point, he discusses the famous computer-assisted proof by Kenneth Appel and Wolfgang Haken of the four-color theorem,[10] noting that

> when Appel and Haken completed [their proof] using a massive automated computation, it evoked much controversy. I interpret the controversy as having little to do with doubt people had as to the veracity of the theorem or the correctness of the proof. Rather, it reflected a continuing desire for *human understanding* of a proof, in addition to knowledge that the theorem is true. [emphasis in original]

I wholeheartedly agree with Thurston's views on these matters as he expresses them in his essay. Expanding the research frontiers of mathematics is a collective enterprise, where progress achieved by one mathematician builds upon that of others. What sort of progress, then, can serve to be built upon in this manner? The mere production of true mathematical statements, perhaps accompanied by formal proofs, does not suffice: what a mathematician needs in order to be able to make further progress is to find out *why* these statements are true, in a way that she can truly internalize. We may think of such understanding as a (more or less rudimentary) map of the mathematical terrain, while the true statements and their formal proofs merely provide isolated snapshots, of little use for further exploration.

The remainder of this paper will be devoted to a couple of thought experiments that I hope will serve to illustrate and illuminate these issues. In particular, I hope that even readers who share my deeply felt Platonic Hunch will be convinced that there is a side to what we set out to accomplish in mathematics which is distinctly human. First, in Section 2, I will describe my thought experiment in the setting of the game of chess, where a dichotomy between objective truth and human understanding exists which is similar to that in mathematics, the advantage of the chess setting being that the situation there is more clear-cut than in mathematics and not at all controversial. Then, in Section 3, I will adapt the thought experiment to the setting of mathematics.

As far as physical plausibility is concerned, the thought experiments will go from bad to worse. I don't think that is a problem. Writes Richard Dawkins, in the opening chapter of his book *The Extended Phenotype* (1982): "Thought experiments are not supposed to be realistic. They are supposed to clarify our thinking about reality."

2 Truth versus understanding in chess

Chess is played on a board with a certain finite but extremely large number of possible positions. Two players, White and Black, take turns making moves according to certain

[10] See, e.g., Appel and Haken (1977).

rules – moves that alter the position on the board. The rules include in particular two – the three-fold repetition rule and the 50-move rule – that prevent the game from going on indefinitely.[11] As a result, the number of moves of a game is bounded.[12] In certain positions, the rules stipulate that Black is check-mate, and when such a position is reached the game ends and White is declared winner. In others, White is check-mate and Black is declared winner, while in yet others the game is declared drawn. In all other positions, the game goes on.[13]

When a chess player sits down to think about what move to play, her thinking can roughly be described as having two aspects, which we may call *combinatorial* and *abstract*. By the combinatorial aspect, I mean looking ahead into concrete possible sequences, such as:

> If I play my knight to f3, then my opponent can either play his knight to c6 or his pawn to d6. In the former case, I have the choice of playing my bishop to b5 (which will probably be answered by pawn to a6) or maybe to c4, while in the latter case I should probably instead play my pawn to d4.

And so on. This has to be combined with the abstract aspect, which means looking at a position to evaluate it using various general principles. The most basic such principle involves counting the pieces on the board weighted by their values, where one rule-of-thumb says that a knight or a bishop is worth about three pawns, a rook five pawns and a queen nine. This is then supplemented by various corrections, such as penalizing doubled or isolated pawns or rewarding rooks on open files; in general, the stronger a chess player is, the more such corrective factors is she able to account for in her abstract evaluation. The evaluation typically results in a more-or-less nuanced statement such as "due to Black's isolated c-pawn, White has a small advantage which can be expected to endure into the endgame" or "in compensation for the sacrificed pawn, White has a lead in development and pressure against f7, resulting in fairly equal chances for the two players".

Idealizing slightly, the art of playing good chess may be described as consisting of looking combinatorially into the various positions that can result, evaluating each of them abstractly, and choosing the moves that result in the best prospects under worst-case assumptions concerning what the other player does. It is desirable to be able to look many moves ahead, but since the number of possible future positions grows exponentially in how far one looks, some pruning of the tree of variations will then be necessary, and a very important skill is to judge which branches are irrelevant enough to admit pruning without harming the quality of the final decision.

[11] Henceforth, by a "position", I mean to include not only where the chess pieces are located, but also certain other aspects that are physically visible not on the board but only on the obligatory score sheets kept by the two players. The most obviously important such aspect is "who's turn is it to move?", but there are also others pertaining to three-fold repetition and the 50-move rule.

[12] Knowing the rules of the game, it is not hard to check that 5100 moves is an upper bound, and that this bound is close to sharp. That a game goes on for n moves means (following standard chess terminology) that White moves n times and Black moves n or $n-1$ times. In practice, games longer longer than 100 moves occur rarely, and games longer than 200 hardly ever.

[13] I am ignoring the fact that in all these other positions, the game can also end, if one of the players resigns in view of hopeless prospects, or if the two players agree on calling it a draw. This additional complication has little effect on the following discussion.

Chess-playing computer programs use the same two aspects as humans, but with a different balance: computer programs are vastly superior as regards the combinatorial aspect (they can consider many more variations) whereas humans are comparatively better at the abstract aspect. On balance, the strongest computer programs today are at least as strong as the best human chess players.

Sometimes, it is possible to make a definite assesment of a position that leaves no room for doubt or further refinement; such will be the case, for instance, if White works out a procedure by which she can check-mate Black within the next three moves no matter what Black does. In this case it makes sense to say that the position is *objectively* a win for White.

I now claim that *every* position has an objective value, and that there are only three possibilities: either the position is a win for White, a draw, or a win for Black; let us denote these by the numerical values 1, $\frac{1}{2}$ and 0 (thus taking White's point of view). The claim follows by considering the so-called game tree for chess; this is well known, but for the reader's convenience I explain in the Appendix how the argument goes.

Thus, somewhat disappointingly for chess addicts, although the better a chess player becomes the more refined and nuanced evaluations of positions will she be able to make, in the limit as she learns chess perfectly these refinements collapse into just the three values 1, $\frac{1}{2}$ and 0. This feature of the game, although of some interest from a philosophical point of view, is of little or no practical interest to chess players, and is therefore seldom discussed in the chess literature; see, however, Rowson (2005) for an exception.[14]

Today, grandmasters and other advanced level chess players tend to spend large amounts of time and effort on opening preparation before their games and tournaments. The opening is the early part of the game where one can reasonably expect the game to follow *exactly* paths foreseen before the game. There is a huge literature on openings, as well as databases of published games and analyses. An important part of opening preparation, especially on grandmaster level, is the search for improvements, i.e., for better moves than those previously known or published. During the past decade, chess-playing computer programs have become an increasingly widely used tool in this kind of search. Such use of computer programs is perfectly allowed, whereas of course using computers *during* games is not allowed.[15]

Imagine now – and this is the thought experiment I have in mind in the present section – that a computer program producing the objective value of any given position is developed and becomes available to these chess players.[16] It is easy to see that in a position with a given value, at least one of the possible moves in the position leads to a position that achieves this value, while none of the available moves improve on it. Let us assume that in addition to the objective evaluation of the position at hand, the program also informs about exactly which moves in the position are good enough to retain that value; this is in principle no harder than

[14] Rowson (2005) and Nunn (2001, 2002) are the only references on the topic of chess that I cite here. My main source on chess, however, consists of my 26 years (starting at age 13) of experience as a competitive chess player.

[15] The last remark refers to ordinary over-the-board games. For correspondence chess, the situation remains unsettled: some tournaments and organizations allow the use of computers during games, while others rule them out (despite the difficulty of enforcing the rule).

[16] I am assuming that the program is based on the brute-force approach outlined in the Appendix. As explained there, such a program is possible in principle, although a highly unlikely prospect in practice.

the original evaluation.

What will happen? Probably the first thing is that we will find out the answer to a question that pops up every now and then in conversations among chess players: what is the objective value of the initial position? Rowson (2005) agrees confidently with the general consensus among grandmasters and other experts that the value is $\frac{1}{2}$, although strictly speaking the answer is not known.[17] Let us assume, for the sake of the discussion, that this general consensus is correct, so that in other words chess played perfectly is a draw. (For the other two possibilities, my discussion below carries through *mutatis mutandis.*)

Then what? Will chess disappear, as a result of everone realizing that they can no longer win chess games? Not at all.[18] Even in objectively drawn positions, prospects of winning may be good due to the possibility for the opponent to make a mistake and slip into a lost position. After all, we are just humans, and chess is such an extraordinarily complicated game that it seems highly unlikely that even the best of us will be able to learn to play it perfectly.

How about the use of this new computer program – let us call it Orakel – in opening preparation? I predict that after a brief feeling of excitement, chess players would begin to realize that Orakel does not provide them with all the information they are looking for. In each position, Orakel provides a move or a number of moves being objectively the best. But in case of more than one objectively best move, which of these moves should the chess player take to heart in order to play when the position arises in a game? Her goal is to win games against fellow human beings, and from this point of view the different moves that tie for the title “objectively best” are typically far from equally good. She should prefer moves that give as many reasonable-looking options as possible for her opponent to make fatal mistakes, at the same time as making it as easy for herself as possible to avoid such mistakes.

We can even imagine a scenario where she maximizes her practical chances of winning the game by playing a move which is *not* among the objectively best ones, thus deliberately decreasing the objective value of her position in return for setting up a puzzle that will most likely turn out to be too difficult for her opponent to handle. Such play is (apart from our extremely rigorous notion of “objectively best”)[19] well-known and usually looked down upon as “cheap trap-setting”, but I am convinced that examples abound that are sufficiently subtle that even on the highest grandmaster level the best thing to do in practice is something that is objectively sub-optimal.

We can think up various refinements of the three values 1, $\frac{1}{2}$ and 0 to be incorporated in improved versions of Orakel. For instance, when a position is a win for White, it may be considered even better for White the smaller the number of moves is in which White can force check-mate. But neither this refinement, nor any of the variations that come quickly to mind, come anywhere near solving the problem of what is the best shot against a fellow human being. I find it hard to imagine how a definite solution to that problem could possibly be achieved in an automatized manner. For one thing, the best move in terms of maximizing

[17] The other “reasonable” answer is that the value is 1 (White wins), although no proof is known ruling out that the answer is 0 (Black wins). The latter is a weird possibility, implying that the initial position is one of so-called mutual *zugzwang*.

[18] Except possibly for the case of correspondence chess.

[19] Somtimes in the chess literature, one move is pointed out as objectively the best and another as preferable in practice, but with the rigorous terminology employed here both judgements are typically far from objective.

one's chances against a particular opponent will depend on exactly who the opponent is: Is he a beginner, an average club player, or a grandmaster? Was he brought up in the Soviet chess tradition? Does he have an obsession with maintaining a harmonic pawn chain or (at the other end of the spectrum och chess players' tempraments) with launching a direct attack upon the enemy king? All this, and much more, influence what is a good move in practice. In short, human psychology is of crucial importance to a chess player, but Orakel will in itself have little or nothing to offer on this aspect of the game.

I also predict that the specialized literature on chess openings will continue to flourish in the presence of Orakel.[20] The variations exhibited in these books will be supplemented not only by the objective verdicts provided by Orakel, but also by the kind of much more nuanced positional evaluations that today's chess players are used to, which offer valuable advice on how they should play in order to maximize their practical chances. Orakel's objective assessments are, on their own, simply not sufficient for providing human chess players with the understanding that they need.

3 Truth versus understanding in mathematics

The goal for a chess player is to win games against other chess players. The issue of what the goal for a mathematician is is far less clear-cut, but as an aid in thinking more clearly about that, I invite the reader to join me in a variation of the thought experiment of the previous section.[21]

This thought experiment involves Orakel II, a machine that is even more preposterous than Orakel. Orakel II is to number theory what Orakel is to chess. What it does is to answer, simply by brute force search, all questions we may have about the natural numbers. More precisely, for any property of natural numbers, or finite tuplets thereof, expressible in standard arithmetic, Orakel II works its way through all candidate instances and then reports back whether or not an instance was found with the desired property. Some examples of questions we may feed into Orakel II are the following.

(1) Do there exist positive integers x, y, z and $n \geq 3$ such that $x^n + y^n = z^n$?

(2) Does there exist an even number $n \geq 4$ which is not the sum of two primes?

(3) Does there exist a positive integer n such that if we write out the number 2^n in decimal form and read it backwards, then we get an integer power of 5?[22]

[20] A somewhat analogous situation has in fact already occured. If we restrict to chess positions with only at most six peices (including pawns and the two enemy kings) on the board, then databases have been set up which solve this part of the game in the same sense that Orakel solves the entire game. Still, there is a need for books that combine information from the database with pedagogical and humanly understandable explanations. For instance, Nunn (2001, 2002) does exactly this. An anonymous reviewer at Amazon.com asks about Nunn (2001) whether "with the advent of endgame databases, is this book worth buying anymore?", to which I answer, most emphatically, yes!

[21] I am talking about pure mathematics here, where real-world applications – helping engineers construct bridges or assisting geneticists in their search for genes that influence susceptibility to breast cancer – are so remote that they play little or no role in the mathematician's daily work.

[22] For instance, $n = 32$ will not do, because $2^{32} = 65,536$, and $63,556$ is not equal to 5^m for any integer m.

We know since about 1994 that the answer to (1) is no.[23] As to (2), the famous Goldbach conjecture from 1742 says that the answer is no, and this is what everyone still believes to be the case although the search for proof has so far been in vain. Finally, concerning (3), Freeman Dyson speculates in his contribution to Brockman (2006) that the answer is no, but also that this fact is not formally provable within the standard axiomatizations of arithmetic,[24] thus exemplifying Gödel's incompleteness theorem.

In all these examples, there are infinitely many cases to check, so what could I possibly mean when I say that Orakel II will do it for us? Well, Orakel II is a bit like an ordinary computer, except that the speed at which it works varies much more dramatically: it carries out its first operation in 1 second, its second operation in $\frac{1}{2}$ second, its third in $\frac{1}{4}$ second, and so on. After $\sum_{i=0}^{\infty} 2^{-i} = 2$ seconds, Orakel II has carried out infinitely many operations, and stops so we can read off the results it has stored in its memory.[25] The memory is infinite in terms of number of bits, but not in terms of physical size, as the first bit occupies 1 Å, the second $\frac{1}{2}$ Å, the third $\frac{1}{4}$ Å, and so on.

Imagine now that, at the time Orakel II is finally manufactured, we still haven't figured out how to prove (or disprove) Goldbach's conjecture. The machine is advertised as telling us everything we could possibly want to know about number theory, so we set it up to solve Goldbach's conjecture, press the "start" button, and after just over two seconds it tells us that indeed, Goldbach was right. Are we happy? Well, not for long. After a brief feeling of excitement, we realize that this is not the answer we wanted. That Goldbach was right we sort of knew all along, but what we really wanted to know was *why* no even number $n \geq 4$ exists which is not the sum of two primes. Orakel II has told us nothing about this.

OK, next try. We program Orakel II to work its way through all syntactically correct formal proofs within standard arithmetic, and check for each of them whether it is a valid proof of Goldbach's conjecture. If we are out of luck, it will turn out that no such proof exists (so we end up with another witness to Gödel's incompleteness theorem, but no useful information as to why Goldbach's conjecture is true). But suppose we are lucky, and Orakel II does produce a proof. This time, our feeling of excitement is *very* brief, because we quickly notice that the proof is $12,804,771$ steps long. Orakel II assures us that this is the shortest formal proof there is. After a couple of years of trying to comprehend the proof and to translate it into high-level mathematical language, we have managed to understand bits and pieces of it, but as to the overall structure of the proof itself, we are completely bamboozled.

So we give up on that particular proof, and turn to Orakel II again, asking for *other* formal proofs of Goldbach's conjecture. It generously provides us with a long list of such proofs, and after ten years M., one of our most brilliant mathematicians, announces that she, together with two of her assistants, has been able to translate a $19,228,630$-step formal proof into high-level mathematical language, and that the proof as a whole actually makes sense. They are initially met by a bit of skepticism from the mathematical community, but after a few months of patiently describing their work in seminars, they have managed to convince a

[23] Wiles (1995), Taylor and Wiles (1995).

[24] Dyson provides good arguments for believing the first part of his speculation, but fails to do so for the second part.

[25] Each bit in the memory can then take *three* values: 0, 1 and $*$, where $*$ means that the value of the bit at time t failed to converge as $t \to 2$.

couple of dozen of the best experts in the field of the soundness of their outline. The riddle of Goldbach's conjecture is finally considered solved, and M. and her two assistants go on to receive medals, prizes, and all sorts of other recognition for having solved it.

Could we have programmed Orakel II to find this, as it turned out in the end, highly elegant solution to the problem for us? I find that hard to imagine, and even if it were possible, it would have had to involve deep insights into human psychology. It seems that some logical structures are much easier for us to grasp than others, for reasons that are hard-wired into our brains.[26] The extraterrestrial beings we encountered in Section 1 may very well be hard-wired in a different manner, giving them a taste for mathematical proof that differs drastically from ours.[27] While they, too, know about the Central Limit Theorem, it may take us some time to figure out that we and they are talking about the same thing, and it may be that their reaction to our favorite proofs of that result is something along the lines of "Well, yes, that is formally correct, but why do it in such a strange and convoluted manner?" followed by the presentation of an alternative proof that looks utterly weird to us.[28]

These thought experiments may, as acknowledged already at the outset, be a bit on the wild side, but I think they nevertheless help us think clearly about objective truth versus human understanding. Like Orakel's $\{0, \frac{1}{2}, 1\}$-valued verdict of a chess position fails to provide us with the understanding we are looking for, Orakel II's $\{\text{true}, \text{false}\}$-valued verdict of number-theoretic statements is equally insufficient. And not even the formal proofs provided by Orakel II suffice to give us the desired understanding. That human psychology is an important aspect of what is to be considered a satisfactory solution to a mathematical problem seems like an unavoidable conclusion even for a hard-headed Platonist.

Appendix: The game tree argument

Here I describe the game tree representation of chess; for more on game trees and how to analyze them I recommend Berlekamp, Conway and Guy (1982).

The tree representing all possible chess games consists of nodes and directed links, and is built up as follows. Starting with a node v corresponding to the initial position, we create one new node w for every position that can be reached by White's first move, together with a link from v to w. We then continue building the tree inductively: for each node we create an outgoing link corresponding to each possible move, together with a new node.[29] A game of chess now corresponds to a branch in this tree, starting at the initial node v, following outgoing links until a node is reached that lacks outgoing links. Note, crucially, that since the

[26] Pinker (1995) presents plenty of evidence in this direction.

[27] As additional support for this wild-looking speculation, let me simply note that in order to find diverging opinion about what constitutes satifaction and beauty in a mathematical proof, I need not even visit other planets: going out of my office and crossing the corridor will in fact suffice.

[28] What I'm saying here is that Erdős' God is probably a bit anthropomorphic. For readers who do not know what I am talking about here, I quote from the preface of the marvellous book by Aigner and Ziegler (2004): "Paul Erdős liked to talk about The Book, in which God maintains the perfect proof for mathematical theorems, following the dictum of G.H. Hardy that there is no permanent place for ugly mathematics. Erdős also said that you need not believe in God but, as a mathematician, you should believe in The Book."

[29] Thus, positions that can be reached via more than one move order appear as more than one node in the tree.

number of moves in each position is finite, and the number of moves in a game is bounded, the game tree is finite – notwithstanding the number of nodes will be Vast (Very much larger than ASTronomical), to borrow a term from Dennett (1995).

We can now go on to assign values (0, $\frac{1}{2}$ or 1) to the nodes of the tree, in the following manner. First, all nodes without outgoing links correspond to positions where the rules of the game stipulate that the game ends with a specific result: 0, $\frac{1}{2}$ or 1. We then continue as follows. Whenever at least one node in the tree remains to be assigned a value, there is one whose outgoing edges all lead to nodes that already have a value.[30] Let w be such a node, and denote the nodes that its outgoing edges point at by $w_1, \ldots, w_k$. If w corresponds to a position where White is to move, then White, playing perfectly, should choose a move that leads to as favorable a position as possible, corresponding to a node w_i that makes Value(w_i) as large as possible among the available options. Thus,

$$\text{Value}(w) = \max\{\text{Value}(w_1), \ldots, \text{Value}(w_k)\}\,.$$

Similarly, if w corresponds to a position where Black is to move, then

$$\text{Value}(w) = \min\{\text{Value}(w_1), \ldots, \text{Value}(w_k)\}\,.$$

In this way, we eventually assign values to all nodes of the tree, and the game of chess is, in a sense, solved.

Note that the game tree argument not only gives an existence result for the objective value of any chess position, but also provides an algorithm for finding that value. However, the Vast-ness of the game tree prevents us from carrying this out in practice by implementing and running the algorithm on a computer.[31]

Acknowledgement. I am grateful to Torbjörn Lundh for triggering me to tear down parts of the wall that separates my thoughts on mathematics from those on chess.

[30] To see this, start at the initial node v, follow outgoing edges, always choosing to go to a node that has not yet been assigned a value. When this is no longer possible, we stand at a node of the desired kind. (Note that, with this procedure, the initial node v will be the last one to be assigned a value.)

[31] Using a technique known as dynamical programming, the waste of time computing power arising from the fact that each chess position is represented by many nodes in the tree can be avoided, but not even this is enough to make the actual implementation of this "final solution" to chess anywhere near feasible.

The sheer physical implausibility of such implementations makes it tempting to deduce that we will never be able to solve chess in the sense of being able to tell the objective value of any given position. But other approaches leading to such a solution are conceivable. Someone may come up with a procedure for abstract evaluation of any position landing in a $\{0, \frac{1}{2}, 1\}$-valued assesment Eval(w) that agrees with Value(w) at all positions where the rules of the game stipulate that it ends, together with an argument showing that in any position with White (resp. Black) to move, all moves reach positions w_i satisfying $\text{Eval}(w_i) \leq \text{Eval}(w)$ (resp. $\text{Eval}(w_i) \geq \text{Eval}(w)$) with equality for at least one of the possible moves. An induction argument then shows that $\text{Eval}(w) = \text{Value}(w)$ for all positions w. The evaluation procedure might even be simple enough for chess players to learn by heart, thus killing the game. This scenario, although not physically preposterous like the brute force solution, still seems quite unlikely to ever occur in reality. But then again, I might be wrong about this.

References

Aigner, M. and Ziegler, G.M. (2004) *Proofs from the Book* (3rd ed.), Springer, New York.

Appel, K. and Haken, W. (1977) Solution of the four color map problem, *Scientific American* **237**, no. 4, 108–121.

Ashman, K.M. and Baringer, P.S., eds. (2001) *After the Science Wars*, Routledge, London.

Atiyah, M., Borel, A., Chaitin, G.J., Friedan, D., Glimm, J., Gray, J.J., Hirsch, M.W., Mac Lane, S., Mandelbrot, B.B., Ruelle, D., Schwarz, A., Uhlenbeck, K., Thom, R., Witten, E. and Zeeman, C. (1994) Responses to Jaffe and Quinn (1993), *Bull. Amer. Math. Soc.* **30**, 178–207.

Berlekamp, E.R., Conway, J.H. and Guy, R.K. (1982) *Winning Ways for Your Mathematical Plays*, Volumes 1 and 2, Academic Press, London.

Blackmore, S. (2005) *Conversations on Consciousness: What the Best Minds Think about the Brain, Free Will, and What It Means to Be Human*, Oxford University Press, Oxford.

Brockman, J., ed. (2006) *What we Believe but Cannot Prove: Today's Leading Thinkers on Science in the Age of Uncertainty*, Harper Perennial, New York.

Dawkins, R. (1982) *The Extended Phenotype*, W.H. Freeman, Oxford.

Dennett, D.C. (1995) *Darwin's Dangerous Idea*, Simon & Schuster, New York.

Dennett, D.C. (2005) *Sweet Dreams: Philosophical Obstacles to a Science of Consciousness*, MIT Press, Cambridge, MA.

Hacking, I. (1999) *The Social Construction of What?* Harvard University Press, Cambridge, MA.

Häggström, O. (2004) Ett paradigmskifte i matematiken? *Svenska Matematikersamfundets Medlemsutskick*, May 15. Also available at
`http://www.math.chalmers.se/~olleh/vetenskap.html`

Hersh, R., ed. (2006) *18 Unconventional Essays on the Nature of Mathematics*, Springer, New York.

Jaffe, A. and Quinn, F. (1993) "Theoretical mathematics": toward a cultural synthesis of mathematics and theoretical physics, *Bull. Amer. Math. Soc.* **29**, 1–13.

Jaffe, A. and Quinn, F. (1994) Response to Thurston (1994) and Atiyah et al. (1994), *Bull. Amer. Math. Soc.* **30**, 208–211.

MacKenzie, D. (1981) *Statistics in Britain, 1865–1930: The Social Construction of Scientific Knowledge*, Edinburgh University Press, Edinburgh.

Nunn, J. (2001) *Secrets of Minor Piece Endings*, Rowman Littlefield, Lanham, MD.

Nunn, J. (2002) *Secrets of Pawnless Endings* (2nd ed.), Gambit, London.

Pickering, A. (1984) *Constructing Quarks: A Sociological History of Particle Physics*, University of Chicago Press, Chicago.

Pinker, S. (1995) *The Language Instinct*, Harper Perennial, New York.

Rowson, J. (2005) *Chess for Zebras*, Gambit, London.

Taylor, R. and Wiles, A. (1995) Ring-theoretic properties of certain Hecke algebras, *Ann. Math.* **141**, 553–572.

Thurston, W.P. (1994) On proof and progress in mathematics, *Bull. Amer. Math. Soc.* **30**, 161–177.

Weinberg, S. (2001) *Facing Up: Science and its Cultural Adversaries*, Harvard University Press, Cambridge, MA.

Wigner, E.P. (1960) The unreasonable effectiveness of mathematics in the natural sciences, *Comm. Pure Appl. Math.* **13**, 1–14.

Wiles, A. (1995) Modular elliptic curves and Fermat's last theorem, *Ann. Math.* **141**, 443–551.

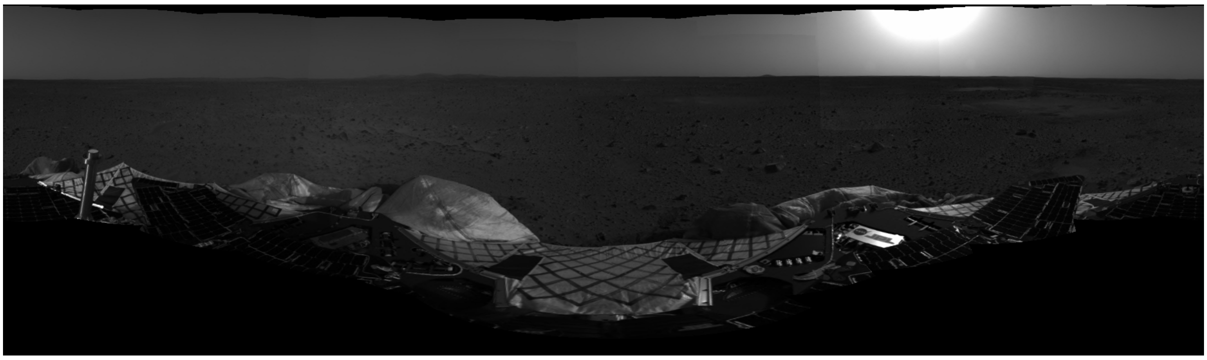
Mars Exploration Rover
Mathematics and People behind the Mission

Uffe Thomas Jankvist & Bjørn Toldbod
Roskilde University*

Abstract

This paper is a selective study of the mathematics and people involved with a specific space mission, namely the Mars Exploration Rover (MER) mission launched in 2003. The specific mathematics of the MER mission are sought uncovered through interviews with applied scientists who worked with different aspects of the mission at the Jet Propulsion Laboratory (JPL). Some of the more specific questions attempted answered through this article for instance concerns if the aerospace industry, exemplified by the MER mission, calls for new developments in mathematics or if it mostly relies on well established theories; how much independence the scientists have in their daily work and in their choosing of solutions to problems and whether or not this variate within different areas of the mission; how ready the aerospace industry is for new solutions and new ideas; and to what extent the economics of the missions play a part in this.

Keywords: Mars Exploration Rover (MER) mission, Jet Propulsion Laboratory (JPL), aerospace industry, applied mathematics, applied scientists.[1]



*IMFUFA, Department of Science, Roskilde University, P.O. Box 260, 4000 Roskilde, Denmark, e-mail: utj@ruc.dk and bjto@ruc.dk

[1] The image on this page is the first picture of the Gusev Crater made public to the press. The picture, which is a mosaic, is taken with the navigation camera onboard the Mars Exploration Rover *Spirit*. `http://nssdc.gsfc.nasa.gov/planetary/mars/mars_exploration_rovers/mera_images.html`

The Montana Mathematics Enthusiast, ISSN 1551-3440, Vol. 4, No. 2, pp. 154-173

> Consider the recent flight to Mars that put a 'laboratory vehicle' on that planet. Whether or not that excites you, you must admit that it is an almost unbelievable technological accomplishment. Now, from start to finish, the Mars shot would have been impossible without a tremendous underlay of mathematics built into chips and software. It would defy the most knowledgeable historian of mathematics to discover and describe all the mathematics that was involved. The public is hardly aware of this; it is not written up in the newspapers. (Davis; 2004)

The adventures of the Mars Exploration Rover Mission truly began on January 4, 2004, when the MER A rover named *Spirit* entered the Martian atmosphere and performed a perfect landing in the Gusev Crater. Later that month, on January 15, the second rover (MER B) known as *Opportunity* landed in the Terra Meridiani. But before these points of time the mission had, of course, been under preparation for years. The place of preparation was the *Jet Propulsion Laboratory* (JPL) in Pasadena, California.

Introduction

In March 2005 we spent a week at NASA's Jet Propulsion Laboratory as part of our joint master thesis[2] at the mathematics department of Roskilde University. The purpose of our stay was to conduct a small investigation of the work related to the Mars Exploration Rover (MER) Mission being performed at JPL and in particular the work relying on the use of mathematics.

The basic idea of this investigation was provided to us by Professor Emeritus Philip J. Davis who in the fall of 2004 implicitly had suggested this in an article from which the introductory quote of this paper is taken (Davis; 2004). Even though we were not "the most knowledgeable historians of mathematics" we still decided to go ahead with an investigation of what mathematics was involved in the MER mission. Unfortunately newspapers were not

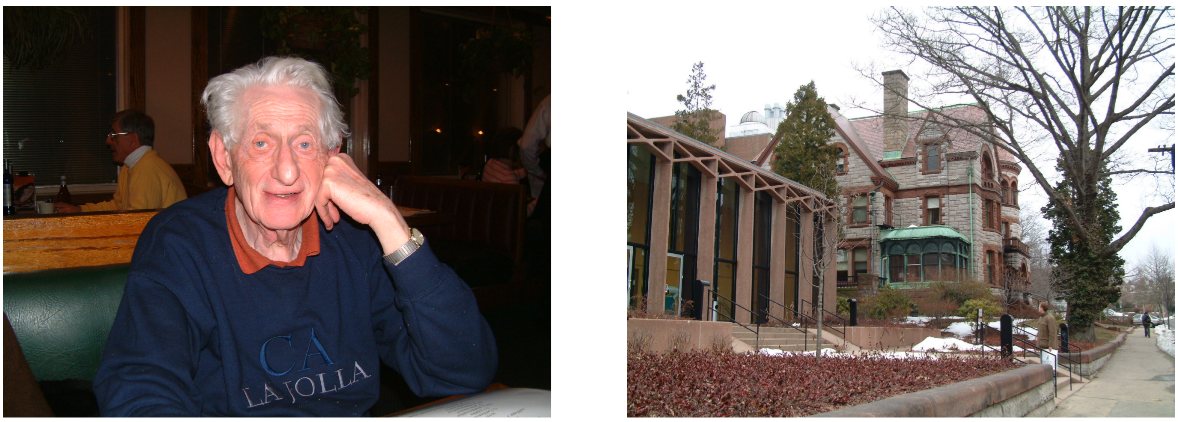


Figure 1 Visiting Brown University. Left: Professor Emeritus Philip J. Davis. Right: The Brown University's Division of Applied Mathematics.

[2] The thesis consists of the texts (Jankvist & Toldbod; 2005a), (Jankvist & Toldbod; 2005b) and (Jankvist & Toldbod; 2005c) and can be found in its original Danish version as IMFUFA-text number 449 at `http://mmf.ruc.dk/imfufatekster/index.htm`

the only place in which this wasn't written up. In fact finding extensive literature on the subject was so difficult that we decided to base the investigation on interviews. Hence the long travel from Roskilde, Denmark to Pasadena, California.

While in the U.S. we decided to take a detour to visit Davis at Brown University in Providence, Rhode Island in order to discuss our pending investigation at JPL. Davis advised us to "be a little bit like journalists rather than teachers of abstracts mathematics" in order to make a more interesting story – perhaps even a story of *general* interest (Davis; 2005). Davis also gave us the following advice:

> I think you want to go also for what you might call the human side of the story. When you interview these people try to get something about their background, what were their experiences before they came to the JPL, what their training is [...] how old they are and how they see their professional future [...] Try to find out what each one does, how much independence they have, how much opportunity is there for them to develop new ideas or new things whether it's in mathematics or whether it's in writing algorithms, software and that kind of thing. [...] are they using packages or are they developing their own stuff, how is this going to fit in to future projects that NASA has? [...] I think you want to make a story here of what it means to work in a space program as a mathematician or a computer scientist or whatever these people are. (Davis; 2005)

To a large extent we have tried to follow the advice of Professor Davis. While at the JPL we tried to get an insight into the employees' personal motivations for working in the aerospace industry as well as an insight into the nature of the mathematical work performed at JPL. Also we looked at what one might call external influences on the daily work, such as deadlines and basic economical limitations. This might also be called the basic *work context*.

In our investigation we have tried to combine the human/daily work approach with a selective study of some of the mathematical problems that arise in a space mission like MER.

Figure 2 Jet Propulsion Laboratory in Pasadena, Californien. Left: JPL seen from above. `http://ipac.jpl.nasa.gov/media_images/jpl_small.jpg` Right: One of JPL's main buildings. `http://www2.jpl.nasa.gov/files//images/browse/jpl18161ac.gif`

As is concluded in the article[3] most of the mathematical problems that arise in a space mission are well known and well described problems (at least from a mathematical point of view). Therefore detailed analysis of the problems will be omitted in the article.

Rather we will use the mathematical problems to illustrate some of our observations regarding how the basic work context of the institution influence the typical approaches taken to solving such mathematical problems.

Both the human approach to JPL and the selective study are sought uncovered by letting the JPL scientist speak for themselves to some extent, i.e. by frequently quoting from our interviews[4]. Before getting to this, however, a small presentation of the Jet Propulsion Laboratory is in order.

The research institution 'the Jet Propulsion Laboratory' is located an out-of-the-way place on the way up in the mountains, surrounded by forrest a little north of the Los Angeles suburb of Pasadena. After about half an hour of driving from downtown Pasadena you reach the area's entrance where you are met every day by uniformed guards. As an outsider you get the firsthand impression that this might be an entrance to a 'secret world of science'. On our first day we were met in the reception by a lady named 'Bobby' who informed us about the safety procedures of the place and then telephoned Dr. William Folkner to announce our arrival. Folkner showed up a few minutes later and took us around the areas of JPL. During our stay we met around a dozen of JPL employees, researchers and scientists, people whom we shall introduce in the following section of this article.

People of JPL

As the main focus of our master thesis was coding theory (which includes both image compression and error correcting codes), we interviewed Dr. Aaron Kiely and Dr. Matthew Klimesh who were deeply involved in the making of ICER – the algorithm used by the rovers for compressing images taken on the surface of Mars, as well as Dr. Jon Hamkins, who told us about the use of error correcting codes for reliable transmission in deep space. We also were fortunate to interview Dr. Mark Maimone who worked with the steering mechanisms of the rovers. Finally we had the pleasure of interviewing Dr. Jacob Matijevic, a mathematicians who had been a long time with JPL and Dr. Miguel San Martin (engineer) with whom we discussed various aspects of the mission.

Dr. William Folkner was the person at JPL with whom we had the most contact and therefore also the person we had most opportunities to interview. We talked to Folkner about the general conditions for working with a space mission and the typical problems that needs solving, as for instance orbit calculations for the spacecrafts. Besides that we also had the opportunity to ask him about his carrier and his view of the people working at JPL. Folkner told us that he himself had taken a Ph.D. in physics from the University of Maryland. After completing his Ph.D. he decided to apply for a job at the JPL to become part of the planetary

[3] A slightly different version of this article in Danish has also appeared in the Nordic mathematical journal *Normat* (Toldbod & Jankvist; 2006). The *hidden* mathematics of the Mars Exploration Rover mission and anticipated consequences of the mathematics being hidden are described in an article in *The Mathematical Intelligencer* (Jankvist & Toldbod; 2007).

[4] Transcriptions of these interviews in their full length along side with the Davis conversation can be found in (Jankvist & Toldbod; 2005c).

exploration. This was in 1988. We asked him if he thought his way to becoming an employee at JPL and his following carrier at the institution was a typical one.

> Most of the people I know have been here a long time or tend to be here a long time. There aren't many people who come here who just want to be here a year or two and then go away again. People who want to do space work tend to *want* to do space work. I think this is the best place in the world to do space work. The people here are all very, very good. They are all very dedicated and they want to make these things fly. (Folkner; 2005)

According to Folkner the strongest motivation for the majority of JPL's employees is the fascination of the missions themselves – a matter we got confirmed several times during our stay. Mark Maimone told us that he wished to stay at JPL as long as possible. He was educated in engineering and had completed a Ph.D. in Pittsburgh, Pennsylvania followed by a post doc in *space robotics* before receiving a position at JPL in 1998. He explained the following about his motives for wanting to work at JPL:

> When I was a kid I saw the Viking missions on Mars and I thought that would be pretty neat, hunched over the terminal looking at these pictures that nobody else would see for a year before it got published. But of course now every picture we take gets published on the Internet the next day so it's not much of an advantage but it's still nice. (Maimone; 2005)

We also discussed the personal motives for wanting to work at JPL with Dr. Aaron Kiely and Dr. Matthew Klimesh. Both of them had studied in Michigan and held Ph.D. degrees. They had come to JPL shortly after completing their studies and both of them wished to stay as long as possible (Kiely; 2005) (Klimesh; 2005).

From what we learned through the interviews a general characterization of JPLs scientists would be people with the highest educational level who join the institution shortly after completing their university studies. They are driven by a desire to be part of the aerospace industry and a passion for planetary exploration. To some extent they were of course also fascinated by the mathematical, physical and engineering problems involved in space exploration, but as a motivating factor this seemed only secondary.

Figure 3 Guided tour at JPL. Left: Uffe and Dr. Albert Haldemann who showed us some of the facilities. Right: Visiting JPL's museum for earlier space missions with Dr. William Folkner.

Work at JPL

One of the first persons we discussed the mathematical aspects of the work at JPL with was Jacob Matijevic. Particularly we discussed the modelling aspects of the work which takes place before the actual mission is set in motion.

A mission like MER is to a large extent about being able to predict how the technology onboard the craft is going to behave in space or in the Martian environment. Once the craft is flying it is impossible to make adjustments which demand more than just a radio signal. Therefore everything must function as expected.

Take, for instance, the Mars environment's influence on the instruments onboard Spirit and Opportunity. You have to have a very precise knowledge about how heat and cold distributes inside the rover and how this affects the instruments. To acquire such a knowledge virtual models of the rovers are build in software so that the thermic conditions can be simulated. Such thermic models are typically based on a number of differential equations which are solved within the programs. The work for the JPL employee consist of building the virtual model of the rover. The exact method of solution which the program implements is to some degree subordinate to the JPL scientist as long as it works and isn't too slow.

According to Matijevic (Matijevic; 2005) you also need to have models of how the environment depends of the seasons on Mars to be able to predict the concrete influence on the instruments in the above mentioned models. Such models are partly based on data from the different Mars orbiters and partly on concrete measurements performed on the Martian surface. The correctness of the surface measurements to a large extent depends on how good the description of the instrument's behaviour in the Martian environment is, and can therefore not be guaranteed. By comparing the data from the orbiters with the surface measurements in question a more accurate picture may arise though and this may then be used to modify the models, so that these slowly become better and better. All of this is done in software. Regarding the models of how the seasons affects the Mars environment it is probably fair to compare the work at JPL with the work performed by an institute of meteorology. Matijevic told the following about this work:

> When I first arrived here over twenty years ago there were still efforts to hand-implement certain mathematical models for certain applications. And there were specialist applications here for specialists in the applied mathematical sciences who worked here to make those applications possible. But over time much of that has been incorporated in fairly standard and available simulation and modelling packages – computer packages. Expansions have been introduced slowly over time to these packages and that's basically how the engineers here do their job. Instead of going back to first principles they apply these tools... the foundation theories are from the eighteenth century to a large degree. (Matijevic; 2005)

Hence a lot of work involving modelling and simulation is done at JPL, but all of this is done in software packages. This might lead one to suspect that JPL has its own staff of mathematicians developing such packages, but Matijevic informed us that the packages mostly come from commercial companies.

A few days after our interview with Matijevic we had the opportunity to interview Dr. Miguel San Martin. He told, with great enthusiasm, about the challenges which the scientists must overcome to make the rovers able to figure out their orientation on the Martian surface.

San Martin explained that the navigation on the surface is based on a well known technique which sailors have used for thousands of years on Earth – you look at the Sun. Together with a vector of gravitation, which can be measured by the rover, the position of the Sun on the sky all in all provides the information necessary to figure out the rover's position on the surface. San Martin himself didn't think of this problem as being very mathematical. In fact he claimed that the majority of the mathematics involved in his work was very simple of nature and he concluded:

> The most important is that you have millions of these little, simple things. And that's the trick; to make them all work, and talk to each other and make sure that no parameter is tightened too much or too little. The complexity of the space problem is keeping it simple. (San Martin & Folkner; 2005)

Folkner, who was also present during the interview, elaborated on this view:

> Well, you hit a lot of mathematics in your descriptions, right, because you need to know the positions of the axes of Mars around the Sun as a function of time and you need to know what the orbits around the Sun and the Earth were. There is a lot of mathematics hidden in what you just said. [...] We've worked all that out for us in the tables. So to know where the Sun is now, you just look it up. Somebody had to figure it out the first time. (San Martin & Folkner; 2005)

Folker's answer illustrates why the question of what mathematics is used in MER is difficult to answer. Knowledge that mathematics previously have made accessible can over time become such an integrated part of our conception of the world that we no longer connects it with mathematics. The trajectory of Mars around the Sun as a function of time is a good example of this: Not many will consider looking in a table to see the Sun's position relative to Mars at a given time as being mathematics. The making of such a table on the other hand *is* a mathematical problem. Thus, the mathematics is very much a part of the scientific work at JPL, but is to a large extent disguised as 'common knowledge' and may therefore in a sense be hidden to the scientist.

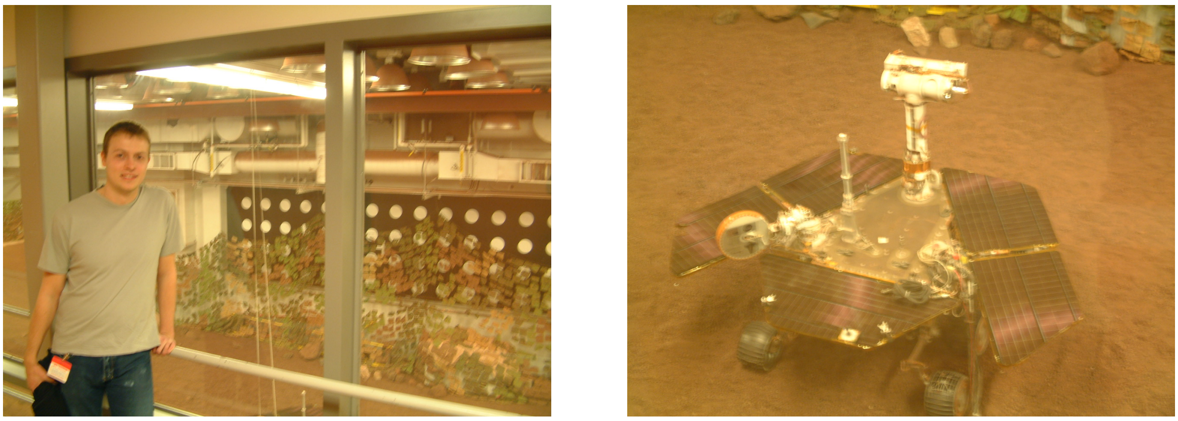


Figure 4 The guided tour takes us by JPL's "sandbox" where rovers are test driven. Left: Bjørn Toldbod in front of the sandbox. Right: A replicate of a Mars Exploration Rover used for test drives at JPL.

The Demand for Reliability

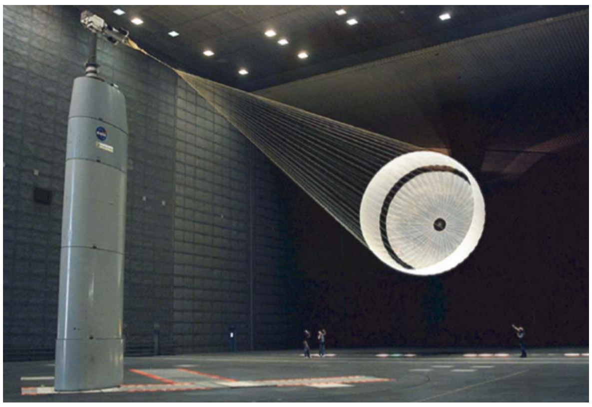
A project like MER sometimes involves around a thousand people at a time. Such a grand scale project of course calls for a huge amount of planning and bringing up to date between different departments and working groups. Besides this an incredibly high reliability of the work performed is demanded. A single mistake in a piece of technology or an algorithm may have serious consequences and in the worst case may result in several years of wasted work for hundreds of people. All of the work being done at JPL is therefore subject to careful development and testing.

The extent to which such planing and testing is done can be hard to imagine. We asked Jacob Matijevic about the development of the parachutes for the rovers, partly because we thought there would be little interaction between the parachutes and other technological devices. In other words, we thought this would be a 'simple' task. Matijevic, however, explained the following about the work with the rovers parachutes:

> We did drop tests. We did wind tunnel tests with the parachutes. But even before this time it was through models of the profiles of these devices that we came up with things like what the entry angles would be, what sorts of release points should we be looking at, as well as designing the algorithm that checks for height above the surface and finding out at which time to deploy the parachute and at which time to fire the rockets for slowing the descent. All of this was based on what we expected to be the environmental profile that the vehicle would see as it came down to the surface. So this was all done in simulation. (Matijevic; 2005)

Reliability is paramount for any mission. If the choice stands between two different approaches to a problem, a space scientist will be most inclined to choose a well known, well tested solution instead of a new and perhaps more efficient solution which has not been thoroughly tested at the planning of the mission.

An area of mathematics in which a lot of new solutions to a problem is being developed all the time is *channel coding*. The *problem* here is reliable communication, and the *solutions* are new error correcting codes. In the last 10-15 years a lot of progress has been made in this



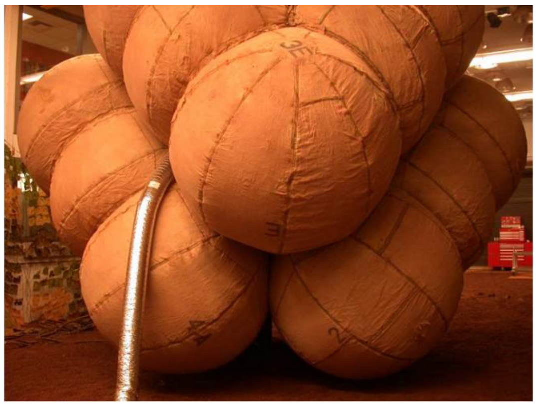


Figure 5 Left: Wind tunnel test of the MER landing module parachute. http://www.nasa.gov/centers/ames/images/content/79641main_picture_2.jpg Right: MER landing module airbags. http://photojournal.jpl.nasa.gov/jpegMod/PIA04999_modest.jpg

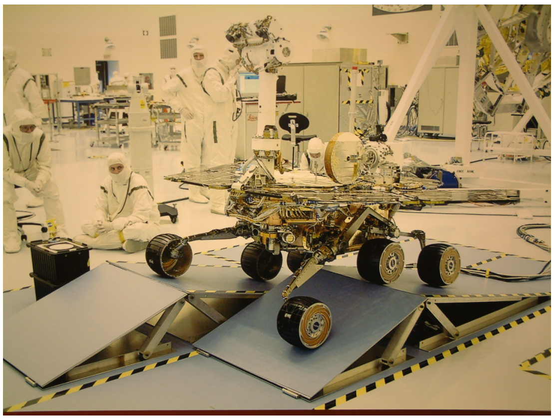
field with the introduction of *turbo codes* which offer a considerable improvement over earlier codes (Berrou et al.; 1993). The demand for reliability, however, has kept the turbo codes out of the space missions, and the codes have only recently been introduced to the missions. Our interview with Dr. Jon Hamkins, one of JPL's leading coding theoreticians, confirmed this for the error correcting codes used in MER:

> The process of flight qualification is very long actually. ... and you know missions that are signing up for a very complicated space craft... they are out to minimize risks so they want stuff that has been flown in previous missions, they don't want something new. It's kind of contrary to the spirit of exploration. They don't want risks even though we are confident that it works... it is a risk to a mission if it hasn't flown before. (Hamkins; 2005)

Mathematics and technology which has been onboard an earlier mission is considered to be safer and therefore makes a more attractive choice. This approach is taken in all aspects of the missions. Of course some development is taking place from mission to mission but only at a pace that makes extensive testing possible. Jacob Matijevic called this "steady progress" (Matijevic; 2005). New ideas which are introduced into the missions will be at least 5-10 years old at launch time, because they must be laid down already when the missions are planned. In the case of channel coding a lot of the mathematics involved has to be implemented in hardware for speed. Generally hardware is much more expensive to replace than software, so the gain of introducing a new error correcting code has to be considerate in order to balance the expense of the substitution.

Deadlines

The scale of the project also means that the work performed by different departments must be completed at specific deadlines. Not surprisingly deadlines may serve as a stop block for the development of new ideas, since it may be difficult to keep a deadline when working with tasks whose solutions are not always well known. It is inconceivable to take on a working



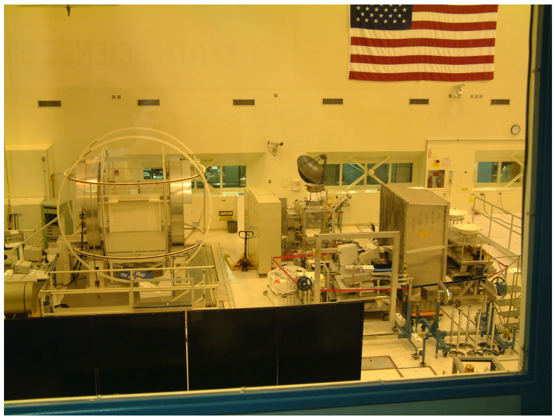


Figure 6 Guided tour at JPL. Left: A JPL photo of a MER rover testing prior to launch. Right: One of JPL's laboratories.

task if you are not sure that there will be enough time to conduct the necessary testing of the work performed. Folkner explained how the constant deadlines affects the composition of JPL employees:

> There are five thousand people here. How many post docs are there? I don't know, maybe a hundred and fifty – something like that. We don't have a lot of graduate students and part of that is because almost everything we do here has to be done on a particular time scale. Very little of our budget is spent doing research where if we get the answer next year or the year after that, it doesn't matter. Everything has to be done on a schedule. So it is not a good training environment for graduate students. Graduate students here are not used much, because the people who could supervise them are busy doing other things. (Folkner; 2005)

Small versus big missions

Besides specific deadlines there is another matter which clearly enhances the tendency to opt out the new and more unsafe: In recent years JPL has gone from a small number of large and expensive missions to a large number of small but cheap missions. For instance the Pathfinder mission of 1996 had a total budget of 265 million dollars[5] whereas the Viking missions of the seventies had a budget of around 8 billion dollars. In this light the Pathfinder mission is most certainly to be considered a 'low cost' mission. The MER mission was more expensive than Pathfinder, but still nowhere near the Viking budget.

If anything from a previous mission can be used again there are huge amounts of money and time to save. Many of the cheaper missions must therefore necessarily rely on reuse from earlier missions. To some extent this reuse issue also applies for the scientists involved in the missions. Folkner explained:

> A problem in doing the smaller missions is that you can only do them with reasonably experienced people, and because they are small you don't have the budget to have an experienced person train an inexperienced person. So JPL is getting older on the average, because we don't have a big mission to afford enough people to have senior and junior people. So JPL is short of junior people. That is not a problem yet, but in five years or ten years it will be a disaster. The management here knows that and is trying to deal with it but it is not often you fix problems until they occur. They are trying to get ahead of that, but it is a difficult thing. Because we are trying to do so many cheap missions we depend on experienced people and we are not budgeting training inexperienced people. It's probably true throughout NASA and the aerospace industry. (Folkner; 2005)

So far we have mostly focused on how external factors influence the mathematics of a space mission. We will now turn to examples of concrete mathematics in the MER mission. Still the purpose of our selection will be to illustrate common features of the mathematics in a space mission.

[5] http://nssdc.gsfc.nasa.gov/database/MasterCatalog?sc=1996-068A

Mathematics in MER

When describing the mathematics used in MER, the scientists at JPL often referred to the different phases of the mission. We shall adopt this approach and begin by presenting these phases.

As mentioned above, before the actual mission begins there is a lot of planning, computer simulation, and testing of equipment taking place. All of these activities belong to what you might call the pre-mission. The actual mission begins with the *launch phase*. When the spacecraft is separated from the launch vehicle the *cruise phase* begins. This phase lasts until 45 days before the craft enters the Martian atmosphere where the *approach phase* begins, a phase under which the trajectory of the craft is constantly adjusted. The *entry, descent and landing phase* (EDL) begins when the craft enters the Martian atmosphere and ends when the landing module's airbags are being retracted (see figure 8). After EDL comes a phase called the *post landing phase*. This phase begins when the solar arrays are unfolded and ends with the rover driving onto the surface – the beginning of the *surface operation phase*. It is in this phase that the rover is driving around, examining the surface, taking and transmitting pictures. In the following we shall refer to these phases of the actual mission.

Miguel San Martin told us about several mathematical problems which are involved in the missions. One of these problems which is always of interest for planetary missions consists of finding out the craft's position in space at a given time – the so-called *Lost in Space* problem. This problem is an always recurring element of the cruise phase. The problem is in essence solved by looking at the stars and trying to identify the stars you are looking at. As it turns out this identification is not at all trivial. According to San Martin you cannot identify a star

Figure 7 The beginning of the launch phase. Left: Launch of Spirit (MER A) on the 10th of June 2003 from Cape Canaveral Air Force Station. `http://www.nasa.gov/lb/missions/solarsystem/merb_covg.html` Right: Launch of Opportunity (MER B) on the 8th of July 2003 also from Cape Canaveral. `http://en.wikipedia.org/wiki/Opportunity_rover_timeline`

by its luminance or position relative to other unidentified stars alone, you have to look at star patterns. San Martin explained:

> It's a pattern recognition problem. The way we do it is using the Sun. That gives us two axes. [...] People have been playing with variations of this since the '60s. Some versions are more clever than others. This one for instance cheated because we used the Sun. 'I'm in three dimensions and I don't know where I am. By looking at the Sun I know two dimensions.' Without using the Sun, this problem is called the 'Lost in Space' problem. (San Martin & Folkner; 2005)

So the problem solved in MER was a simplified version of the Lost in Space problem. The Lost in Space problem was both formulated and solved during the 1960s, i.e. in the beginning of the space exploration era. The reason we have begun this tale by pulling the Lost in Space problem out of the hat, is that it has a quality which is typical for the problems in MER – the solution rests upon *well established* applied mathematics.

The solving of the Lost in Space problem requires a number of measurements performed onboard the spacecraft. This gives rise to another problem.

> In the cruise we are spinning the spacecraft. We use conservation of angular momentum to keep our spacecraft from turning. The most sophisticated piece of software that we have onboard during that time is a Kalman filter. A Kalman filter is a statistical framework, or algorithm perhaps, to mix information from different sensors. So you have a dynamic model – in this case it's a simple rigid body, which you represent mathematically and then you have some sensors which measure where the Sun is. And you have a statistical noisy model. (San Martin & Folkner; 2005)

The Kalman filter makes it possible to combine data from different sensors. Data from these sensors will often be incomplete, according to the behavior of the measuring equipment in the specific situation, or affected by noise of some kind. The Kalman filter gives an estimate of what the measurements would have been had it not been disturbed by all these elements. San Martin commented on the use of the filter:

> It was invented in the '60s ... You can come up with a sequential filter that allows you to optimally combine the information from three things; your dynamic model, your sensors and your star tracking into optimal information about your attitude [position of spacecraft relative to a frame of reference] and the inertial properties of your plant. So that is going on all the time. It's a well known aerospace industry. (San Martin & Folkner; 2005)

The Kalman filter is named after Rudolph Emil Kalman who first published on this matter in 1960 (Kalman; 1960). Thus the Kalman filter is another example of MER's use of a well established mathematical theory also dating back to the '60's.

During our interview with William Folkner we came across another issue that all space missions needs to attend to, that of finding the optimal trajectory between the point of launch and the destination. The trajectory is first corrected during the launch phase by firing rockets on the launch vehicle. During the cruise phase the trajectory is also frequently adjusted and, as mentioned before, especially during the approach phase a lot of adjusting is taking place. The Mars missions always uses the same trajectory, a so called *Hohmann* trajectory. For two bodies (Mars and Earth) circling another body (the Sun) the Hohmann trajectory is the solution of the trajectory problem which calls for the least amount of energy at launch.

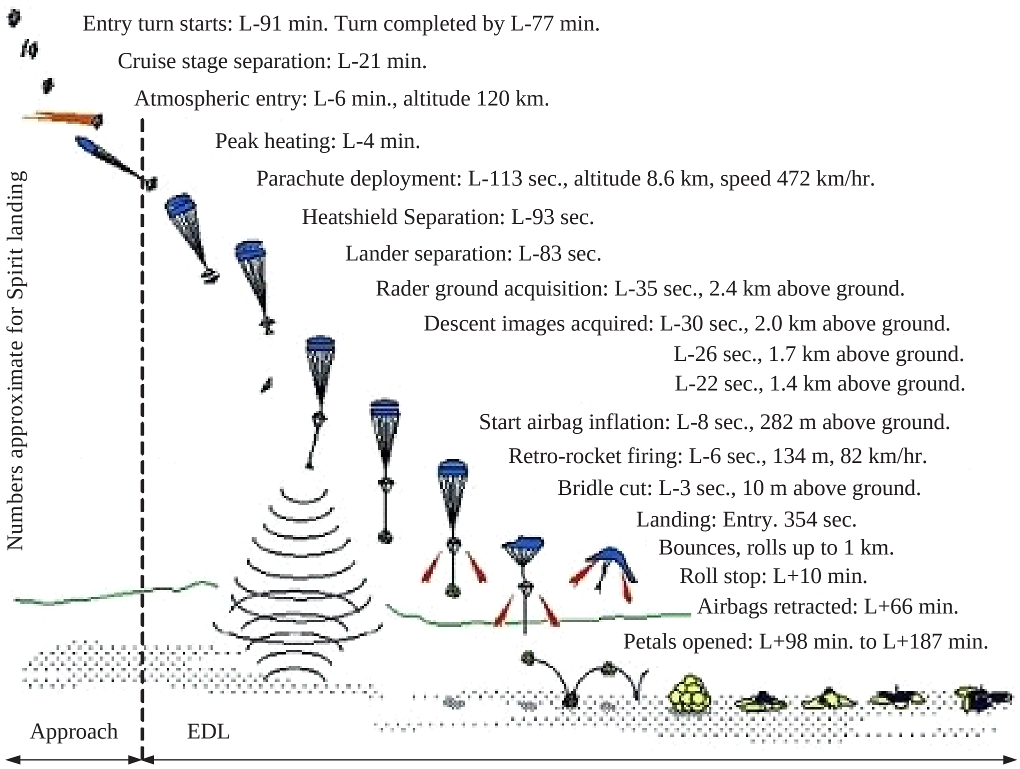


Figure 8 Entry, descent and landing. In the text above 'L' stands for landing. (JPL; 2002)

Because the movements of Mars and Earth around the Sun are not located in the same plane, two optimal solutions (a type 1 and a type 2) exist which calls for the least amount of energy. We asked Folkner if it would be reasonable to call the trajectory problem for going to Mars a standard exercise:

> Yes, for Mars it is a very standard exercise. For the other planets it tends to be more complicated because you'll trade flybys of other planets for angular momentum against the mission operations time. [...] For going to Mars it always comes down to: 'Do you want to do type 1 or type 2?' (Folkner; 2005)

The Hohmann trajectory was discovered by the German engineer, Walter Hohmann in 1925[6] and thus serves as another example of how the aerospace industry relies on well established mathematics. A description of the more complicated trajectories used for flights to other planets of which Folkner speaks can be found in (Marsden & Ross; 2006).

The mathematical problems and disciplines which we have described above are relatively isolated. If you are to mention a larger and more complete theory that played a significant role in MER, *control theory* might be suitable. We asked Mark Maimone about the amount of control theory used for steering the rovers:

> We have to drive the wheels so there is software that controls the motors there and all the instruments have motors and have to be controlled so there is some amount of control theory being applied there. [And] there are a lot of motors,

[6] See for instance `http://vesuvius.jsc.nasa.gov/er/seh/know2.html`

> there are motors that drive the wheels, that spin the cameras, that control the arm and when we landed we had a lot of motors that were used simply to get up off the lander; stand up, spread out the wheels, spread out the legs, open up the solar panels, pull up the mast and do all those things. (Maimone; 2005)

From our interview with Matijevic we could understand that the steering might not be so mathematically complex as one might think:

> Mainly what we're taking advantage of to at least create the driving pattern, is the fact that each of the individual six wheels is controllable at least for steering and so we can create in essence any kind of arc condition. Because of the individual control we can modulate the speed of the motor turns. That gives us a means for being able to accommodate surface interactions between the wheels and the terrain. The foundation is actually fairly simple. We're using in each case a simple proportional integral derivative; the control algorithm is very linear. (Matijevic; 2005)

From Matijevic's description it seems that most of the control theory used was also well established mathematics that has been known and applied for many years. During our interview with San Martin, he suggested that more advanced control theory might be found in the EDL phase. However, we did not have the time to investigate this any further.

As mentioned early in this article the focus of our study was partly another coherent mathematical theory called *channel coding*, which deals with the notion of *reliable communication*.

The signals transmitted to and from Mars are subject to interference during their travel through deep space. Such interferences of a binary signal may result in bits becoming altered. The communication between Earth and the rovers not being reliable is of course not acceptable, just imagine what consequences this might have for the adjustments of the trajectory. The problem is solved by way of channel codes. The use of coding theory makes it possible to correct altered bits in a message, hence the codes are also called *error correcting codes*. The coding system used in MER depends on two different codes used in combination. Jon Hamkins explained:

> The majority of the missions flying now are concatenated. So the data comes in and is Reed-Solomon encoded, then it goes through a block interleaver and then it's convolutionally encoded. (Hamkins; 2005)

Thus MER's coding system consisted of two combined, or concatenated, error correcting codes; a Reed-Solomon code and a convolutional code. Reed-Solomon codes are algebraic codes whose code symbols comes from a Galois Field. Convolutional codes are another type of codes which are not as mathematically well understood as the Reed-Solomon codes, but on the other hand are very efficient and therefore very often used in technology. Convolutional codes are excellent for correcting single bit errors, the kind of errors which most often occurs from interference in deep space. Unfortunately the decoding of convolutional codes[7] often results in a run of consecutive errors, so-called burst errors. Fortunately Reed-Solomon codes are excellent in correcting exactly burst errors, hence the concatenated system. The reason for first encoding the data with the Reed-Solomon code and then the convolutional code is that the decoding procedure must be the reverse of the encoding procedure. Block interleaving is

[7] JPL uses the so-called *Viterbi algorithm* in their convolutional decoder.

Figure 9 The Mars rover *Spirit* moves its robot arm over a Martian stone in order to take a series of pictures with its microscope camera. `http://marsrovers.jpl.nasa.gov/gallery/press/spirit`

a technique used to ensure that the burst errors from the convolutional decoding are no more severe than what the Reed-Solomon decoder can handle.

Reed-Solomon codes are named after Irving S. Reed and Gustave Solomon who introduced these in 1960 (Reed & Solomon; 1960). The convolutional codes are due to Peter Elias who published on this matter in 1954 (Elias; 1954). The concatenated system was originally described and tested in the beginning of the '70's (J. P. Oddenwalder et al.; 1972) and used for the first time during the Voyager 2 mission in 1985.

Coding theory was used in the majority of the phases in the actual mission. Every message between Earth and the rovers, whether they were in flight or on the surface of Mars, were subject to the coding system described above. Adjustments of the trajectory was not the only communication occurring during the cruise phase for which reliable communication was of paramount importance: Software uploads of the rover's steering systems took place both during the cruise phase and the surface phase (Maimone; 2005). On the surface another mathematical theory which is also part of coding theory came into play, namely that of *image compression*.

One of the main purposes of the MER mission was to take pictures of Mars. Before these pictures were transmitted to Earth they had to be compressed. The image compression

technique primarily used in MER is called *The ICER Progressive Wavelet Image Compression*, in short just ICER, and was developed at JPL by Kiely and Klimesh. The word 'progressive' refers to *progressive fidelity compression*. In such a compression a low quality approximation of the picture is first transmitted. Afterwards bits are transmitted in such a way that the quality of the picture are gradually improved. When all bits are transmitted the reconstructed picture equals the original picture. By stopping the transmission before it is complete lossy compression can be obtained. In this way ICER supports lossless as well as lossy compression even though it was entirely used for lossy compression. For lossless compression MER relied on the commercial compression algorithm *LOCO*. The ICER algorithm, like many other image compression techniques, overall consists of three stages; preprocessing, modelling and entropy encoding of data. Kiely explained:

> We got data coming in, an image or whatever it is, and then some sort of preprocessing stage for example a wavelet transform plus quantization or a discrete cosine transform or something. The goal is that it doesn't perform any compression and in fact it is often a lossy process, it might throw out some of the data but the idea is to process the data in a way that makes it more receptive to compression through the entropy encoder. The entropy encoder is sort of the engine. Given some sort of probabilistic model of the source it compresses data or represents it in a more efficient way through something like a variable length code. That is sort of the big picture of what is going on. So for example for ICER what is going on is mostly a probabilistic transform. For LOCO it is in essential trying to project a probability distribution on the next pixel that it is about to encode based on what it has seen in the nearby neighbors. (Kiely; 2005)

ICER uses a wavelet transform that closely but not exactly resembles a *Haar*-transform, a context model (also known as a Markov model) and the majority of the entropy codes used by the entropy encoder are the so-called *Golomb codes* (Kiely & Klimesh; 2003). The LOCO algorithm is a bit different from ICER since it do not have a preprocessing stage which is typical for lossless compression techniques. It does use context modelling and it uses both Golomb codes and *Huffman codes* (Weinberger et al.; 1996). Alfréd Haar's transform dates back to 1910 (Haar; 1910). Golomb codes are due to Solomon W. Golomb who described these codes in 1966 (Golomb; 1966). Huffman codes are due to David A. Huffman who discovered these codes in the beginning of the 1950s while still a student at MIT (Huffman; 1952).

From our point of view the most interesting aspect of ICER is that this compressor for the first time introduced wavelets into an image compressor for space applications. Wavelets has only recently found its way into commercial standards like the JPEG 2000 compressor, so this is an area of the mission where a surprisingly new approach to a problem is taken. The reason for this is probably that using a new image compressor is not so much of a risk because it is implemented in software rather than hardware.

The above presentation of mathematics in MER is of course merely scratching the surface. Discovering every little piece of mathematics put to use in the MER mission probably is an impossible task to undertake despite being "the most knowledgeable historians of mathematics" or not.

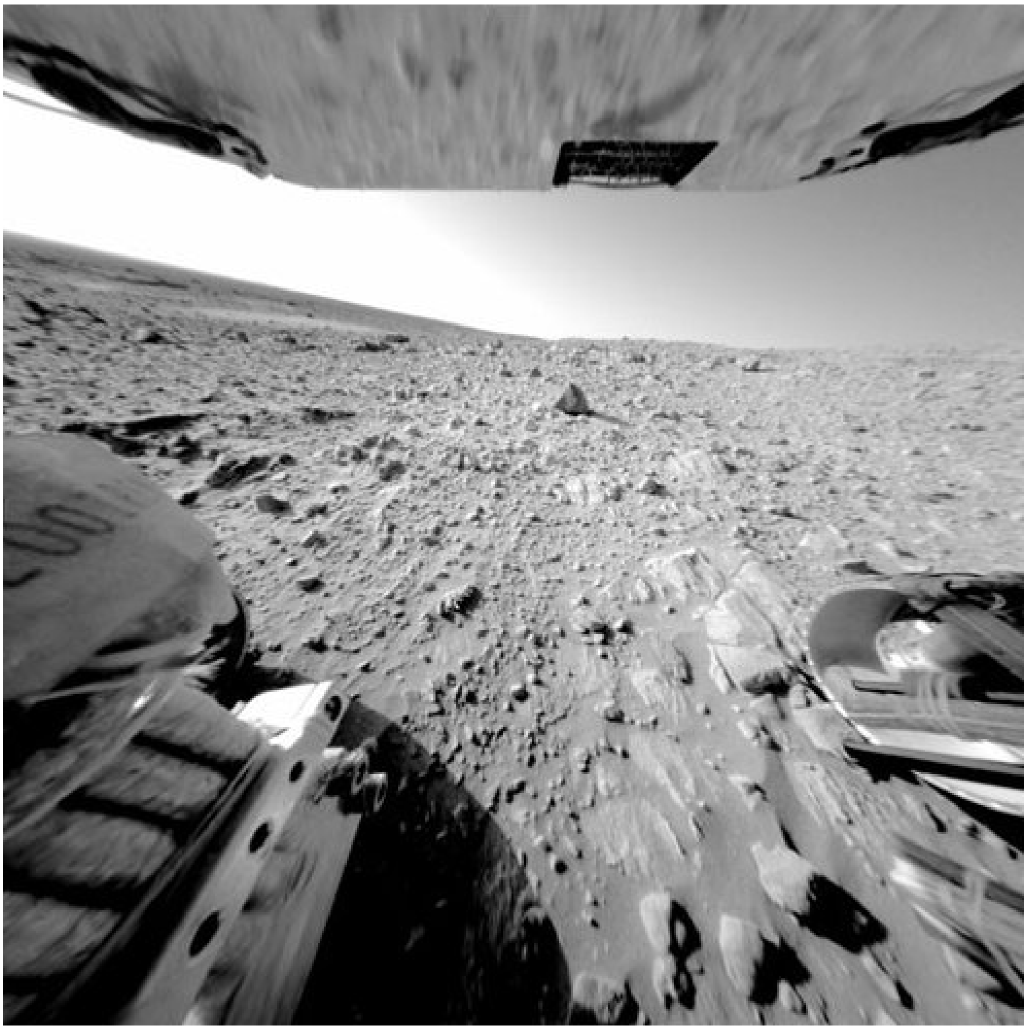


Figure 10 A picture taken from Spirit's rear camera. The rover's left rear wheel can be seen to the right in the picture. `http://marsrovers.jpl.nasa.gov/gallery/press/spirit`

Conclusions

From the mathematics we did discover and investigate we found that the majority of the theories are well known and well established theories of mathematics, like for instance the Hohmann trajectory, Kalman filters, Reed-Solomon codes and convolutional codes. Also in the image compression both Golomb codes and Huffman codes as well as the Markov models are examples of well established mathematics. The newest piece of mathematics we came across was the wavelets in ICER's preprocessing. The mathematical theory of wavelets only dates back to the beginning of the 1990's and is due to Ingrid Daubechies (Daubechies; 1992). However, wavelets has been around in applied mathematics for a long time and as we mentioned the specific transform of ICER is inspired by the Haar-transform from 1910.

Besides the use of all these 'advanced' mathematical theories are of course an enormous amount of more basic mathematics. So Philip Davis was certainly right when he said that "the Mars shot would have been impossible without a tremendous underlay of mathematics".

Folkner said it himself when we discussed the MER mission:

> There is mathematics in everything. There is control theory, aerodynamics, orbital dynamics, Newtonian gravity, bodies going around the Sun. We use general relativity, that's mathematics of physics [...] Linear algebra is a field of mathematics we use all the time. Matrices. That's in the control theory all the time. There is Riemannian geometry in the general relativity. Calculus. (San Martin & Folkner; 2005)

Thus, a mission like MER is relying on a vast amount of mathematics. However, the mathematics is often hidden. Not surprisingly the mathematics is hidden to the public as most mathematics in our society is. But more surprisingly, it is to some extent also hidden to the employees of JPL. The main reason for this is that a lot of the work done at JPL involves mathematics embedded in commercial software packages. Another reason is that some of the mathematics is such an integrated part of a mission, that it is not thought of as mathematics (we mentioned earlier looking up planetary positions in a table). Finally, classification contributes to hiding the mathematics – something which we also encountered with a couple of times during our stay at JPL. (For a further discussion of the hidden mathematics in MER, see (Jankvist & Toldbod; 2007).) The hiding of the mathematics in the mission is bound to make it more difficult to discover the mathematics involved.

Due to the extreme nature of a Mars mission one might think that this would call for 'extreme' mathematics, mathematics that would have to be developed for the sole purpose of this mission. This, however, does not seem to be the case. We did not come across any contributions to basic research in mathematics as a result of the MER mission through our investigation. The role of JPL seems to be another, namely that of the consumer of already developed mathematics – applied mathematics.

Due to this the majority of the scientists of JPL are applied scientists, but with a wide variety of educational backgrounds; engineering, computer science, physics, mathematics and so on. The thing these people have in common is that they all have very high levels of education (Ph.D. degrees). For JPL the educational background does not necessarily seem to be the main focus, rather it is a question of whether or not an employee can solve the task that needs solving. Klimesh told:

> I didn't start working on data compression until I came here. Actually that is not so unusual in engineering... [...] A lot of the mathematics is the same and when they hire someone, what they really want is someone who is good a solving problems. (Klimesh; 2005)

All the people we talked to seemed to have come to JPL more or less immediately after having completed their studies. And all of them seemed driven by the desire to work within the aerospace industri and none seemed to want to leave JPL again.

The work performed at JPL are subject to a number of conditions all with consequences for the work. The lower budgets of the smaller missions has the effect that the JPL staff is composed by many senior scientists and only few juniors, a matter which in part is also conditioned by the smaller missions' dependence on experienced personnel. The lower budgets also makes it attractive to reuse existing solutions to problems from mission to mission, a matter which is not likely to promote the implementation of new ideas. Also the very strict timetables and deadlines may make it almost impossible to pursue new ideas and solutions to old problems. The long and tortuous process of flight qualification, including the extensive

testing of all equipment also enhances this tendency. However, it seems that for some areas of a space mission new ideas are more welcome than in other areas. The image compression scheme used in MER was the newly developed ICER compressor. ICER could more easily be introduced to the mission, because 'only' software needed to be replaced. But on the overall it seems that new ideas must always be weighed against the effort needed to implement them. We shall end our conclusions with a quote by William Folkner in which he nails this point exactly:

> Everything is a cost-benefit analysis. The whole space system is a cost-benefit analysis. (Folkner; 2005)

Acknowledgements

We want to thank all the employees at JPL who set aside their duties to talk to us, especially Bill Folkner who arranged most of our meetings. Also we would like to thank Phil Davis for taking such an extraordinary interest in our investigation and for commenting on this article. Thank you also to Tinne Hoff Kjeldsen, Anders Madsen and Bernhelm Booß-Bavnbek (all at Roskilde University) for making suggestions and commenting on the original thesis. And thank you to Jørgen Larsen for LaTeX assistance all along the way.

Bibliography

Berrou, C., Glavieux, A. & Thitimajshima, P. (1993). Near Shannon Limit Error-Correction Coding and Decoding: Turbo Codes, *Proceedings of the IEEE International Conference on Communications*, IEEE, Piscataway, NJ. Conference held in Geneva, Switzerland.

Daubechies, I. (1992). *Ten Lectures on Wavelets*, CBMS-NSF Regional Conference Series In Applied Mathematics, Capital City Press, Montpelier, Vermont.

Davis, P. J. (2004). A Letter to Christina of Denmark, *EMS* pp. 21–24.

Davis, P. J. (2005). Interview with Professor Emeritus Philip J. Davis March 6th. Recorded at Brown University, Providence.

Elias, P. (1954). Error-Free Coding, *IRE Transactions on Information Theory* **4**: 29–37.

Folkner, W. (2005). Interview with Doctor William Folkner March 17th. Recorded at JPL, Pasadena.

Golomb, S. W. (1966). Run-Length Encodings, *IEEE Transactions on Information Theory* **IT-12**: 399–401.

Haar, A. (1910). Zur Theorie der orthogonalen Funktionensysteme, *Mathematische Annalen.* **69**: 331–371.

Hamkins, J. (2005). Interview with Doctor Jon Hamkins March 14th. Recorded at JPL, Pasadena.

Huffman, D. A. (1952). A Method for the Construction of Minimum-Redundancy Codes, *Proceedings of the IRE* **40**: 1098–1101.

J. P. Oddenwalder et al. (1972). Hybrid Coding Systems Study Final Report, *NASA Document NASA-CR-114486* .

Jankvist, U. T. & Toldbod, B. (2005a). *Matematikken bag Mars-missionen – En empirisk undersøgelse af matematikken i MER med fokus på kildekodning og kanalkodning*, Master's thesis, Roskilde University. Tekster fra IMFUFA, nr. 449a.

Jankvist, U. T. & Toldbod, B. (2005b). *Matematikken bag Mars-missionen – Indførelse i den grundlæggende teori for kildekodning og kanalkodning i MER*, Master's thesis, Roskilde University. Tekster fra IMFUFA, nr. 449b.

Jankvist, U. T. & Toldbod, B. (2005c). *Matematikken bag Mars-missionen – Transskriberede interviews fra DTU, Brown University, MIT og JPL*, Master's thesis, Roskilde University. Tekster fra IMFUFA, nr. 449c.

Jankvist, U. T. & Toldbod, B. (2007). The hidden mathematics of the Mars Exploration Rover mission, *The Mathematical Intelligencer* **29**(1).

JPL (2002). Mars Exploration Rover (MER) Project Mission Plan.

Kalman, R. E. (1960). A New Approach to Linear Filtering and Prediction Problems, *Transactions of the ASME – Journal of Basic Engineering* **82**(Series D): 35–45.

Kiely, A. (2005). Interview with Doctor Aaron Kiely March 14th. Recorded at JPL, Pasadena.

Kiely, A. & Klimesh, M. (2003). The ICER Progressive Wavelet Image Compressor, *IPN Progress Report* **42**(155): 1–46.
URL: `http://ipnpr.jpl.nasa.gov/tmo/progress\_report/42-155/155J.pdf`

Klimesh, M. (2005). Interview with Doctor Matthew Klimesh March 14th. Recorded at JPL, Pasadena.

Maimone, M. W. (2005). Interview with Doctor Mark W. Maimone March 17th. Recorded at JPL, Pasadena.

Marsden, J. E. & Ross, S. D. (2006). New Methods in Celestial Mechanics and Mission Design, *Bulletin of the American Mathematical Society* **43**: 43–73.

Matijevic, J. (2005). Interview with Doctor Jacob Matijevic March 14th. Recorded at JPL, Pasadena.

McEliece, R. J. & Swanson, L. (1994). Reed-Solomon Codes and the Exploration of the Solar System, *in* S. B. Wicker & V. K. Bhargava (eds), *Reed-Solomon Codes and Their Applications*, IEEE Press, New York, pp. 25–40. Chapter 3.

Reed, I. S. & Solomon, G. (1960). Polynomial Codes over Certain Finite Fields, *Journal of the Society for Industrial and Applied Mathematics* **8**: 300–304.

San Martin, M. & Folkner, W. (2005). Interview with Doctor Miguel san Martin & Doctor William Folkner March 14th. Recorded at JPL, Pasadena.

Toldbod, B. & Jankvist, U. T. (2006). Reportage fra en Mars-mission, *Normat* **54**(3): 101–115 and 144.

Weinberger, M. J., Seroussi, G. & Sapiro, G. (1996). LOCO-I: A Low Complexity, Context-Based, Lossless Image Compression Algorithm, *Proceedings of the IEEE Data Compression Conference* pp. 1–10.
URL: `http://www.hpl.hp.com/loco/dcc96copy.pdf`

Wicker, S. B. (1998). Deep Space Applications, *in* V. S. Pless & W. C. Huffman (eds), *Handbook of Coding Theory Volume II*, Elsevier Science B.V., Amsterdam, pp. 2121–2169. Chapter 25.

The Philosophy of Mathematics, Values and Keralese Mathematics

Paul Ernest[1]
University of Exeter, UK

Abstract

This paper explores the philosophical significance of the Keralese and Indian subcontinent contribution to history of mathematics. Identifying the most accurate genesis and trajectory of mathematical ideas in history that current knowledge allows should be the goal of every history of mathematics, and is consistent with any philosophy of mathematics. I argue for the need of a broader conceptualization of philosophy of than the traditional emphasis on scholastic enquiries into epistemology and ontology. For such an emphasis has been associated, though I add need not necessarily be so, with an ideological position that devalues non-European contributions to history of mathematics. The philosophy of mathematics needs to be broad enough to recognise the salient features of the discipline it reflects upon, namely mathematics.

Keywords: Non-european roots of mathematics; Keralese mathematics; Philosophy of mathematics; mathematics and values; history of mathematics.

1. What is the Business of the Philosophy of Mathematics?

Traditionally, in Western philosophy, mathematical knowledge has been understood as universal and absolute knowledge, whose epistemological status sets it above all other forms of knowledge. The traditional foundationalist schools of formalism, logicism and intuitionism sought to establish the absolute validity of mathematical knowledge by erecting foundational systems. Although modern philosophy of mathematics has in part moved away from this dogma of absolutism, it is still very influential, and needs to be critiqued. So I wish to begin by summarising some of the arguments against Absolutism, as this position has been termed (Ernest 1991, 1998).

My argument is that the claim of the absolute validity for mathematical knowledge cannot be sustained. The primary basis for this claim is that mathematical knowledge rests on certain and necessary proofs. But proof in mathematics assumes the truth, correctness, or consistency of an underlying axiom set, and of logical rules and axioms or postulates. The truth of this basis cannot

[1] Author's address: University of Exeter, School of Education, Exeter EX1 2LU, UK, E-mail: PErnest@ex.ac.uk; Homepage: http://www.ex.ac.uk/~PErnest/

be established on pain of creating a vicious circle (Lakatos 1962). Overall the correctness or consistency of mathematical theories and truths cannot be established in non-trivial cases (Gödel 1931).

Thus mathematical proof can be taken as absolutely correct only if certain unjustified assumptions made. First, it must be assumed that absolute standards of rigour are attained. But there are no grounds for assuming this (Tymoczko 1986). Second, it must be assumed that any proof can be made perfectly rigorous. But virtually all accepted mathematical proofs are informal proofs, and there are no grounds for assuming that such a transformation can be made (Lakatos 1978). Third, it must be assumed that the checking of rigorous proofs for correctness is possible. But checking is already deeply problematic, and the further formalizing of informal proofs will lengthen them and make checking practically impossible (MacKenzie 1993)

A final but inescapably telling argument will suffice to show that absolute rigour is an unattainable ideal. The argument is well-known. Mathematical proof as an epistemological warrant depends on the assumed safety of axiomatic systems and proof in mathematics. But Gödel's (1931) second incompleteness theorem means that consistency and hence establishing the correctness and safety of mathematical systems is indemonstrable. We can never be sure mathematics theories are safe, and hence we cannot claim their correctness, let alone their necessity or certainty. These arguments are necessarily compressed here, but are treated fully elsewhere (e.g., Ernest 1991, 1998). So the claim of absolute validity for mathematical knowledge is unjustified.

The past two decades has seen a growing acceptance of the weakness of absolutist accounts of mathematical knowledge and of the impossibility in establishing knowledge claims absolutely. In particular the 'maverick' tradition, to use Kitcher and Aspray's (1988) phrase, in the philosophy of mathematics questions the absolute status of mathematical knowledge and suggest that a reconceptualisation of philosophy of mathematics is needed (Davis and Hersh 1980, Lakatos 1976, Tymoczko 1986, Kitcher 1984, Ernest 1997). The main claim of the 'maverick' tradition is that mathematical knowledge is fallible. In addition, the narrow academic focus of the philosophy of mathematics on foundationist epistemology or on Platonistic ontology to the exclusion of the history and practice of mathematics, is viewed by many as misguided, and by some as damaging.

2. Reconceptualizing the Philosophy of Mathematics

Although a widespread goal of traditional philosophies of mathematics is to reconstruct mathematics in a vain foundationalist quest for certainty, but a number of philosophers of mathematics agree this goal is inappropriate. "To confuse description and programme - to confuse 'is' with 'ought to be' or 'should be' - is just as harmful in the philosophy of mathematics as elsewhere." (Körner 1960: 12), and "the job of the philosopher of mathematics is to describe and explain mathematics, not to reform it." (Maddy 1990: 28). Lakatos, in a characteristically witty and forceful way which paraphrases Kant indicates the direction that a reconceptualised philosophy of mathematics should follow. "The history of mathematics, lacking the guidance of

philosophy has become blind, while the philosophy of mathematics turning its back on the...history of mathematics, has become empty" (1976: 2).

Building on these and other suggestions it might be expected that an adequate philosophy of mathematics should account for a number of aspects of mathematics including the following:

1. **Epistemology**: Mathematical knowledge; its character, genesis and justification, with special attention to the role of proof
2. **Theories**: Mathematical theories, both constructive and structural: their character and development, and issues of appraisal and evaluation
3. **Ontology**: The objects of mathematics: their character, origins and relationship with the language of mathematics, the issue of Platonism
4. **Methodology and History:** Mathematical practice: its character, and the mathematical activities of mathematicians, in the present and past
5. **Applications and Values**: Applications of mathematics; its relationship with science, technology, other areas of knowledge and values
6. **Individual Knowledge and Learning**: The learning of mathematics: its character and role in the onward transmission of mathematical knowledge, and in the creativity of individual mathematicians (Ernest 1998)

Items 1 and 3 include the traditional epistemological and ontological focuses of the philosophy of mathematics, broadened to add a concern with the genesis of mathematical knowledge and objects of mathematics, as well as with language. Item 2 adds a concern with the form that mathematical knowledge usually takes: mathematical theories. Items 4 and 5 go beyond the traditional boundaries by admitting the applications of mathematics and human mathematical practice as legitimate philosophical concerns, as well as its relations with other areas of human knowledge and values. Item 6 adds a concern with how mathematics is transmitted onwards from one generation to the next, and in particular, how it is learnt by individuals, and the dialectical relation between individuals and existing knowledge in creativity.

The legitimacy of these extended concerns arises from the need to consider the relationship between mathematics and its corporeal agents, i.e., human beings. They are required to accommodate what on the face of it is the simple and clear task of the philosophy of mathematics, namely to give an account of mathematics.

3. Challenging Epistemological Assumptions and Values

The challenge to the traditional philosophy of mathematics to broaden its epistemological goal, as indicated above, raises some critical issues. In particular, if providing ironclad foundations to mathematical knowledge and mathematical truth is not the main purpose of philosophy of mathematics, has this fixation distorted philosophical accounts of mathematics and what is deemed valuable or significant in mathematics? To what extent is the philosophical emphasis on mathematical proof and deductive theories justified? I want to argue that the emphasis on mathematics as made up of rigorous deductive theories is excessive, and this focus in fact existed

for only two periods totaling possibly less than ten percent of the overall history of mathematics as a systematic discipline, and then only in the West.[2]

The first of these two periods was the ancient Greek phase in the history of mathematics which reached its high point in the formulation of Euclid's *Elements*, a systematic exposition of deductive geometry and other topics. The second period is the modern era encompassing the past two hundred years or so. This second period was first signaled by Descartes' modernist epistemology, with its call to systematize all knowledge after the model of geometry in Euclid's *Elements*. However, fortunately, his injunction was not applied in the practices of mathematicians for the next two hundred years, which was instead a period of great creativity and invention in the West. Only in the 19th century did the newly professionalized mathematicians turn their attention to the foundations of mathematical knowledge and systematize it into axiomatic mathematical theories. The contributions of Boole, Weierstrass, Dedekind, Cantor, Peano, Hilbert, Frege, Russell and others in this enterprise up to the time of Bourbaki are well known.

I am not claiming that all or even most mathematical work was foundational during these two exceptional periods. But the foundational work is what caught the attention of philosophers of mathematics, and in the spirit of Cartesian modernism has become the epistemological focus of modern philosophy of mathematics, as well as the touchstone for what is deemed to be of epistemologically valuable. I do not want to detract from either the magnificence of the achievement in the foundational work carried out by mathematicians and logicians, nor from the pressing nature of the problems that made attention to it so vital in the early part of the 20th century. Nevertheless, the legacy of this attention has been to overvalue the philosophical significance of axiomatic mathematics at the expense of other dimensions of mathematics. Two underemphasized dimensions of mathematics are calculation and problem solving. All three of these aspects of mathematics involve deductive reasoning, but axiomatic mathematics is valued above the others as the supreme achievement of mathematics.

There is another feature shared by the two historical periods that emphasised axiomatic mathematics, namely a purist ideology involving the philosophical dismissal or rejection of the significance of practical mathematics. The antipathy of the ancient Greek philosophers to practical matters including numeration and calculation is well known. This aspect of mathematics was termed 'logistic' and regarded as the business of slaves or lesser beings. In the modern era, calculation and practical mathematics have been viewed as mathematically trivial and philosophically uninteresting. The fact that philosophers have been concerned with ontology and the nature of the mathematical objects has engendered little or no interest in the symbolism of mathematics, or calculations and transformations that convert one mathematical object (or rather its name, a term) into another. Such a view is typified by Platonism, which concerns itself primarily with mathematical truths and objects. These are presumed to exist in an unearthly and idealized world beyond that which we inhabit as fleshy and social human beings, such as Popper's (1979) objective World 3.

---

[2] I take the beginning of disciplinary mathematics to be around 2500-3000 BCE, following Høyrup (1980) and (1994).

Of course at the same time as these modern developments were taking place applied mathematics and theoretical or mathematical physics were making great strides, but this was not considered to be of interest to philosophers of mathematics (however much interest it was to philosophers of science), because of their purist ideology. Even in British public schools, during the late Victorian era, mathematics was taught in with ungraduated rulers because graduations implied measurement and practical applications, which was looked down upon for the future professional classes and rulers of the country. (Admittedly some of the rationale was that Euclid's geometry only requires a straight-edge and a pair of compasses as drawing instruments).

What I have described here (in order to critique it) is an ideological perspective that elevates some aspects of mathematics above others, but typically does not acknowledge that it is based on a set of values, a set of choices and preferences to which no necessity or logical compulsion is attached. Furthermore, it appears that such values have only been prominent during a small part of the history of mathematics.

In order to strengthen my critique of these values I want to point out that mathematical proof, the cornerstone of axiomatic mathematics, and calculation in mathematics, are formally very close in structure and character. In Ernest (forthcoming) I have argued that mathematical topic areas (e.g., number and calculation) can be interpreted as being made up semiotic systems, each comprising (1) a set of signs, (2) rules of sign production and transformation, and (3) an underpinning (informal) meaning structure. Such signs include atomic, i.e., basic, signs and a range of composite signs comprising molecular constellations of atomic signs. These signs may be alphanumeric (made up of numerals or letters) or figural (e.g., geometric figures) or include both (e.g., figures with labels and the types of inference employed). The use of semiotic systems is primarily that of sign production in the pursuit of some goal (e.g., solving a problem, making a calculation, producing a proof for a theorem). I want to claim that most recorded mathematical activity concerns the production of sequences of signs (within a semiotic system). Typically these are transformations of an initial composite sign ($\boldsymbol{S}_{\mathbf{1}}$), resulting after a finite number ($\boldsymbol{n}$) of transformations, in a terminal sign ($\boldsymbol{S}_{\mathbf{n+1}}$), satisfying the requirements of the activity. This can be represented by the sequence: $\boldsymbol{S}_{\mathbf{1}} \rightarrow \boldsymbol{S}_{\mathbf{2}} \rightarrow \boldsymbol{S}_{\mathbf{3}} \rightarrow \dots \rightarrow \boldsymbol{S}_{\mathbf{n+1}}$. Each transformation (represented by $\rightarrow$) constitutes the application of one of the rules of the semiotic system to the sign, resulting in the derivation of the next sign in the sequence. More accurately these should be represented by $\rightarrow_{\mathbf{i}}$, with i = 1, …, n, since each transformation in the sequence is potentially different.

My claim is that this formal (semiotic) system describes most mathematical domains and activities. If the initial sign is the statement of a problem, the sequence represents the derivation of a solution to the problem. I will not dwell on this case as there are many complications involved in problem solving, such as the use of multiple representations, branching solution attempts[3], etc. and some of the transformations (such as interpreting an initial problem formulation and constructing a problem representation) are neither easily made explicit nor fully formalizable. Furthermore, there is no simple characterization of the relationship between the transformational rules and the underlying informal meaning structure, for the transformations are partly structure preserving morphisms, and partly calculational.

[3] Clearly branching derivations can occur in virtually all mathematical processes or activities including those involved in problem solving, deriving mathematical proofs, and mathematical calculations. However, they are mostly eliminated from the transcriptions of successfully completed activities.

More significantly in the present context, such transformational sequences can represent a deductive proof for a theorem. In this case it consists of a sequence of sentences, each of which is derived from its predecessors by the deductive rules of the system (including the introduction of axioms or other assumptions). The final sign in the sequence is the theorem proved. The meaning structure underpinning the rules of proof is based on the principle of the preservation of the truth value of sentences in each deductive step, and hence along the length of the proof sequence (which is why axioms can be inserted, and why proofs 'work', i.e., do what they are designed to do.)[4]

In the case of a calculation, the initial sign is usually a compound term. Subsequent terms are derived by calculational rules and typically each is a simplification in some sense of its predecessor. The final term in a calculation is the simplified numerical 'answer' to the problem. With the introduction of algebra and other functions and operations such as trigonometrical functions, the answer may instead be a simplified but non-numerical term (i.e., a function). Thus calculations are sequences of terms, each derived from predecessors by the rules of the system. The meaning structure underpinning the rules of calculation is based on the principle of the preservation of numerical value.[5]

Thus there is a strong analogy between the semiotic systems based on calculation and those based on deductive proof. The transformations of terms and sentences are based on the underlying principles of value preservation, namely numerical value or truth value, respectively, as I have demonstrated. In addition, calculation concerns terms and proof concerns sentences (or formulas), and both of these are defined similarly. Terms (sentences) are defined recursively as follows. An atomic term (sentence) consists of a constant or a variable (an n-place predicate applied to constants or variables, respectively). A compound term (sentence) is defined as the result of applying a function or operation (a logical connective or quantifier, respectively) to one or more terms (sentences, respectively), to make a new term (sentence, respectively). Thus structurally terms and sentences are very similar, defined analogously by induction.

The sequential and rule-based nature of calculation is something that precedes the development of the deductive proof of theorems by at least two thousand years. My contention is that without the long and ancient tradition of rule following in sequences of calculations, and the entrenchment of the grammatical and value preservational features noted above, the development of proof would not be possible. As I have indicated here, there is a striking analogy between calculations and deductive proofs of theorems, rarely if ever remarked upon, that puts into question the claimed superiority of proof.

---

[4] Note that I have not distinguished between the two analogous forms of proof which employ equivalence or deductive consequence as the transformational relationship at each step. In the latter the truth value derived is greater than or equal to is precedent value, in the former it is equal to it. But in each case (in bivalent logic), since the initial truth value in the sequence must be 1 the whole sequence of truth values including the final term, the theorem proved, is 1.

[5] The preservation of one of the four inequality relations along the sequence is possible variation, where an upper or lower bound on the value of the term is determined

Furthermore, proof and calculation are formally equivalent, in modern foundational terms. Calculations utilize the term as a basic unit of meaning (and as that which is transformed), whereas deductive proofs use the sentence (including formulas or open sentences) as a basic unit. However, there are equivalence transformations between calculations and proofs. A calculation sequence of the form $t_1, t_2, t_3, \ldots, t_n$, where each t_i ($1 \leq i \leq n$) is a term, can be represented as a deductive proof of the form $t_1=t_2$, $t_2=t_3$, $t_3=t_4$, …, $t_{n-1}=t_n$ in which each identity asserts that numerical values of adjacent terms are preserved identically in the calculation. By an extended or repeated application of the transitivity of identity ($x=y$ & $y=z \rightarrow x=z$, for all x, y & z), $t_1=t_n$ is derived, thus equating the initial term of the calculation and the terminal term, the 'answer'.

Likewise, a deductive proof of the form $S_1, S_2, S_3, \ldots, S_n$, can be represented as a series of terms. These are the values of the truth value function *f* defined on numerical representations of true and false sentences to give the values 1 and 0, respectively. For a valid proof these values must be $f(S_1 \rightarrow S_2) = f(S_2 \rightarrow S_3) = \ldots = f(S_{n-1} \rightarrow S_n) = 1$. The formal details are messy and omitted here (see Gödel 1931 for the introduction of arithmetization of logic, and Kleene 1952) but the principle is both simple and sound. It is well known that *f* is a morphism mapping $<S, \rightarrow>$ onto a Boolean algebra $<f(S), \leq>$.[6]

The very strong analogy and structural similarities between proof and calculation, including their inter-convertibility, challenges the preconception often manifested in philosophical and historical accounts of mathematics that proof is somehow intellectually superior to calculation in mathematics.[7] Looking in detail at the technical and structural aspects of proof and calculation reveals that they cannot be so easily attributed to different epistemological domains as is often claimed. It is not defensible to say that proof alone in mathematics pertains to the true, good, beautiful, to wisdom, 'high-mindedness' and the transcendent dimensions of being, and that calculation is only technical and mechanical, pertaining to the utilitarian, practical, applied, and mundane; the lowly dimensions of existence. Such assertions are part of an ideological position incorporating a set of values that overvalues pure proof-based mathematics as having epistemological significance, and undervalues calculation and applied mathematics as having only practical significance; going back to the social divisions of ancient Greek society, as noted above, and the prejudices and ideology to which it gave rise.

This preconception or prejudice is used as the basis for asserting that the contributions of some cultures and civilizations are intellectually superior to others in history of mathematics. It also undervalues the solving of problems, calculations and other local applications of deduction in mathematics (including proof, see Joseph 1994). Thus the mathematics of ancient Egypt, Mesopotamia and India, as well as other countries outside of the Greco-European tradition, is viewed as inferior and immature. Part of the argument is that only cultures that produce axiomatic proof in mathematics achieve the highest levels of abstract intellectual achievement.

---

[6] Technically the truth value function *f* can simply be defined on the domain of sentences under a given interpretation provided that there is an effective procedure for determining whether each sentence is true or false (thus giving values 1 or 0, respectively) under the given interpretation of the underlying theory or formal language.

[7] Joseph (1991) is among the few to note the importance of algorithms and calculation in the history of mathematics and to note their almost universal devaluation by other commentators.

I have argued that philosophical dispositions and values have underpinned a prejudice against ascribing value to certain forms of mathematical activity. In particular, that axiomatic systems are greatly valued over less systematic forms of deduction including problem solving, calculation and unsystematized proofs. Furthermore, this prejudice also maintains the contrast between and overvalues any form of proof, including unsystematized and unaxiomatized proofs over any form of calculation or problem solving.

These two levels of prejudice, these two value-based distinctions and preferences are frequently elided in the history and philosophy of mathematics. Thus the contributions of the ancient Greeks of the Euclidean type, and the modern focus on axiomatics of the past two centuries are seen to characterize the superior forms of thought of what is purported to be a Greco-European tradition. Furthermore, the unsystematized and unaxiomatized proofs and methods characterizing the official European history of mathematics from the late-Renaissance to the beginning on the Nineteenth century are seen as also reflecting the superior methods and concepts and higher forms of thought of the modern European tradition in their nascent phase, whose superiority and value is demonstrated in the subsequent flowering of the axiomatic tradition in Europe.

Through this elision, there arises the discounting of the proof-based contribution of cultures and civilizations outside of the 'Greco-European' tradition. Thus although there is a tradition of convincing demonstration or proofs, known as Upapatis, originating around two millennia ago in India, these proofs are discounted as intellectually inferior (Joseph 1994). Admittedly there are significant differences between the ancient Greek and the early Indian concepts of proof. Joseph 2000) has convincingly argued that ancient Indian mathematics was at least partially shaped by linguistic and grammatical conceptions of knowledge, based on the contributions of Panini; whereas ancient Greek mathematics was shaped by developments in philosophical thinking. So there are differences in the epistemological basis for different forms of proof that have emerged in different cultures and civilizations. However, the current challenges to the philosophy of mathematics discussed in the beginning of this paper, legitimate challenges to the traditional univocal and absolute conceptions of mathematics, knowledge and proof. From the perspective of the new fallibilist or social constructivist philosophies of mathematics, there is no ultimate or uniquely correct form of proof. Rather the forms of proof accepted within any culture or civilization during any epoch are a function of the historically contingent conceptual history and epistemological preconceptions that emerge and are accepted by the relevant geographico-historical communities of scholars. So there is no basis for elevating certain cultural forms of proof and demoting others on epistemological grounds alone. Each must be judged within the cultural contexts of its geographico-historical location.

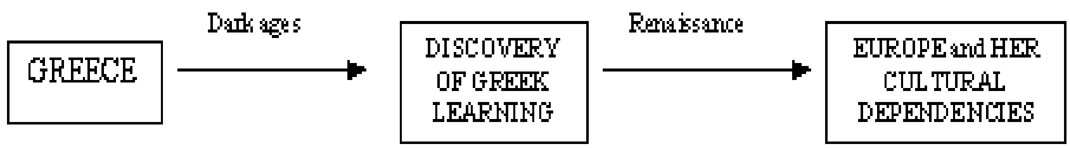
4. Eurocentrism in the History and Philosophy of Mathematics

The above discussion raises the question of why informal and unsystematized proofs and demonstrations that occur in the mathematical histories of certain cultures are valued more than those of others. Why, for example, are the unsystematized proofs, methods and results of post-Renaissance European mathematics regarded as superior to antecedent developments in Kerala of comparable character? To answer this it is necessary to turn to another dimension of ideological prejudice at work in the history and philosophy of mathematics. This is eurocentrism,

the racist bias that claims that the European 'mind' and its cultural products are superior to those of other peoples and races. Thus Bernal (1987) has argued that during the past two hundred year or so, ancient Greece has been 'talked up' as the starting point of modern European thought, and the 'Afroasiatic roots of Classical Civilisation' have been neglected, discarded and denied.

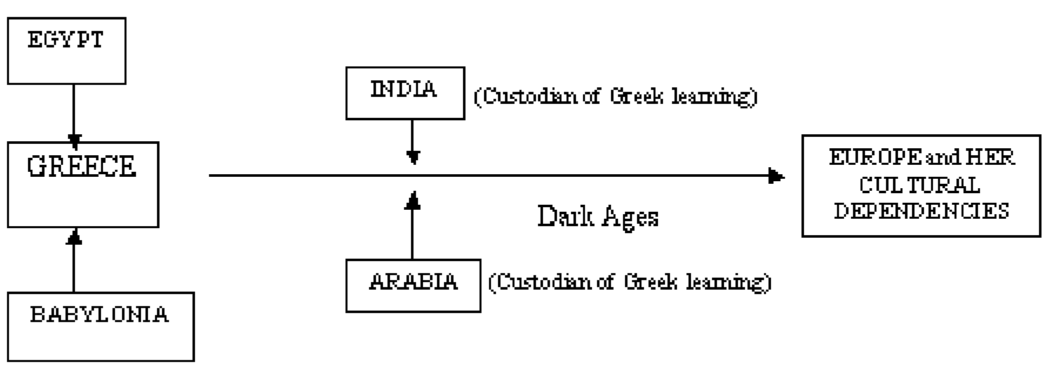
Against this backdrop it is not surprising that that mathematics has been seen as the product of European mathematicians. However, there is now a widespread literature supporting the thesis that mathematics has been misrepresented in a eurocentric way, including Almeida and Joseph (2004), Joseph (2000), Powell and Frankenstein (1997) and Pearce (undated). A common feature of eurocentric histories of mathematics is to claim that it was primarily the invention of the ancient Greeks. Their period ended almost 2000 years ago, which was followed by the 'dark ages' of around 1000 years until the European renaissance triggered by the rediscovery of Greek learning led to modern scientific and mathematical work in Europe (and its cultural dependencies). This trajectory is illustrated in Figure 1.

Figure 1: Eurocentric chronology of mathematics history (from Pearce, undated).



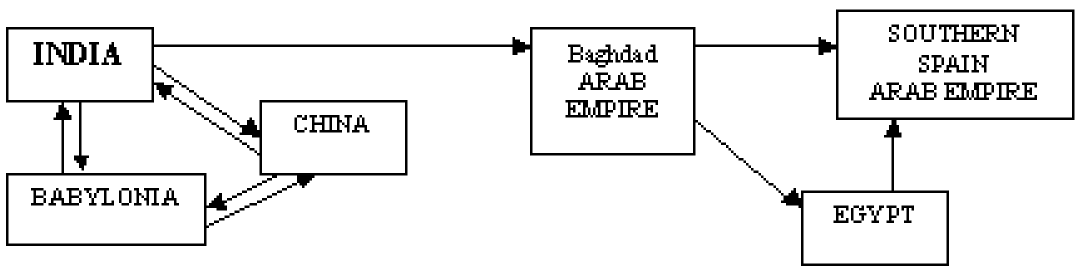
Some accounts have acknowledged the impact of lower level Egyptian and Babylonian mathematics on ancient Greek developments, as well as the later minor contributions of Indian and Arabic mathematicians (often seen primarily as custodians of Greek knowledge) on the history of mathematics in Europe (i.e., *The* history of mathematics). This is shown in the Modified Eurocentric model (figure 2).

Figure 2: Modified Eurocentric model (from Pearce, undated)



Pearce, Joseph and others go on to argue that in the so-called 'dark ages' and beyond, from 5th - 15th centuries, a great deal of mathematical work continued. Further the relationships between different regions and countries was complex and multidirectional and "A variety of mathematical activity and exchange between a number of cultural areas went on while Europe was in a deep slumber." (Joseph, 2000: 9). In figure 3 I reproduce Pearce's diagram of interrelationships in the development Non-European mathematics during the dark ages.

Figure 3: Non-European mathematics during the dark ages (from Pearce, undated)



Thus out of ignorance or prejudice arising from ideologically based values and preconceptions, eurocentric histories of mathematics, neglect the 'Non-European roots of mathematics' (to quote the subtitle of Joseph, 2000). There is a small but growing impact of such critical ideas in the history and philosophy of mathematics as indicated here. However, in my view, there is still an under-emphasis on the vital role of pre-Hellenic civilizations in providing the conceptual basis for modern mathematics through calculation, problem solving, etc.

5. Mathematics of the Indian Subcontinent and Kerala

One of the major casualties in the Eurocentric view of mathematics has been the ignoring or undervaluing of the contributions to mathematics of the Indian subcontinent. The long presence of deductive proofs in mathematics from this region has already been noted (Joseph 1994, 2000). Although the invention of zero by mathematicians of the Indian subcontinent has long been acknowledged, the significance of this as the lynchpin of the decimal place value system is often underestimated. Rotman (1987) presents a view of this innovation that puts its significance as reaching far beyond mathematics, right at the heart of European cultural and intellectual development in the Renaissance and early modern times. Pearce (undated) argues the Indian development of decimal numeration together with the place value system is the most remarkable development in the history of mathematics, as well as being one of the foremost intellectual productions in the overall history of humankind. I have indicated above how both philosophically and in the published histories of mathematics, calculation and numeration have traditionally been downplayed as epiphenomena of what is perceived to be the much more important Platonic conception of number. This is a misrepresentation of the intellectual significance of these developments without which the modern conceptions of number (including its computerization, with all of the applications this brings) would not be possible.

In the history of mathematics in the Indian subcontinent, much attention has been given to very large numbers, including powers of ten up to near 50. Whether these were contributors to or results of the development of the decimal place value system is for historians to say. Likewise it is tempting to speculate as to whether the extension of the decimal place value system into decimal fractions helped in the conceptualisation and formulation of the remarkable series expansions developed in Kerala. Although there is no unequivocal historical basis for this, it is convincingly claimed that floating point numbers were used by Kerala mathematicians to investigate the convergence of series (Almeida et al. 2001).

This brings me to one of the most remarkable and most neglected episodes in the history of mathematics, and the focus of this conference. This is the fact that Keralese mathematicians discovered and elaborated a large number of infinite series expansions and contributed much of the basis for the calculus, which is traditionally attributed to 17th and 18th century European mathematicians. Furthermore, this is not a case of simultaneous discovery in Kerala, for the work in Kerala took place two centuries before that in Europe.

Pearce (undated), Joseph (2000) and others attribute to Madhava of Sangamagramma (c. 1340 - 1425), the Keralese mathematician-astronomer, the important step of moving on from the finite procedures of ancient mathematics to treat their limit, the passage to infinity, the essence of modern classical analysis. He is also thought to have discovered numerous infinite series expansions of trigonometric and root terms, as well as for π, for which he calculated the value up to 13 (some say 17) decimal places (Pearce, undated). These inestimably important results anticipate some of the discoveries attributed to or named after the great mathematicians Gregory, Maclaurin, Taylor, Wallis, Newton, Leibniz and Euler.

Joseph (2000: 293) claims that “We may consider Madhava to have been the founder of mathematical analysis. Some of his discoveries in this field show him to have possessed extraordinary intuition”. Almeida *et al*. (2001) have argued that Keralese contributions as a whole anticipate developments in Western Europe by several centuries in work on infinite series for numerical integration results.

In addition, these results are very possibly not just the anticipations of unacknowledged genius in the Indian subcontinent, and as such a very remarkable case of independent discovery. There is the very real possibility that these Keralese discoveries were transmitted to Europe by Jesuit missionaries and ‘appropriated’ by European mathematicians as their own (Almeida and Joseph 2004). The arguments for this transmission and appropriation are very persuasive, if not yet established with certainty. Certainly the mathematicians of Renaissance Europe are known to have been secretive about their methods and knowledge, and if they had ‘purloined’ the foundational results of calculus from Kerala would conceal and deny their origins.

As a non- historian of mathematics, I find this new recognition of the major Keralese and Indian subcontinent contributions to the history of mathematics remarkable. The fact that traditional histories of mathematics fail to acknowledge these and other non-European contributions is partly due to ignorance, for until recently it was difficult to find proper sources on this in standard texts. But there is much more to this, as there have been some reports of the anticipations in the literature for almost two centuries which have been disparaged or ignored.

Instead there are two sets of entwined ideological presuppositions that have led to this denial and blindness. The first is the epistemological prejudice towards a certain style of mathematics, namely the axiomatic theories and purist ideas discussed above as well as favoring proofs over calculation al and applied mathematics. Through the lenses of these modern prejudices the historical contributions of non-Eurocentric traditions has been minimized and trivialized. The second set of ideological presuppositions is more sinister. This is the racial prejudice of Eurocentrism. Namely, that only the 'Western mind' (i.e., the Caucasian or European) is capable of the pure thought and insights required in the highest forms of mathematics. Thus the contributions of African, Asian, Indian subcontinent, and Oriental peoples is discounted and minimized, because by the presupposed 'very inferior nature' of these peoples they are incapable of the high levels of thought involved. Hence any results that contradict these prejudices is *ab initio* incorrect. Thus such discoveries are minimized as intellectually inferior, or doubted and attributed to the transmission and copying of ideas from West to East, or in the last resort, challenged with regard to their chronology.

6. Conclusion

So what is the philosophical significance of the Keralese and Indian subcontinent contribution to history of mathematics? Identifying the most accurate genesis and trajectory of mathematical ideas in history that current knowledge allows should be the goal of every history of mathematics, and is consistent with any philosophy of mathematics. However, I have argued that a broader conceptualization of philosophy of mathematics is needed than the traditional emphasis on scholastic enquiries into epistemology and ontology. For such an emphasis has been associated, though I add need not necessarily be so, with an ideological position that devalues non-European contributions to history of mathematics. The philosophy of mathematics needs to be broad enough to recognise the salient features of the discipline it reflects upon, namely mathematics. As Lakatos (1976) indicated in the quote given above, the philosophy of mathematics has become empty by ignoring the history of mathematics.

It is no little charge to claim that the history and philosophy of mathematics have in effect become infused with error and a racist ideology, through the implicit and unacknowledged values and prejudices. Elsewhere, as well as above, I have argued that it is the business of mathematics and the philosophy of mathematics to take the issue of values seriously (Ernest 1998), and it is no longer enough to claim that these are outside of its proper subject matter. After all, ethics is just another branch of philosophy and I can see no grounds for its *a priori* exclusion. All human activities, however rarefied and abstruse are part of the vast cultural project of humankind, and as such none has the right to claim exemption from awareness of values and social responsibility (provided that this is not used as an excuse to limit freedom of thought and critique).

Endnote: This paper was delivered at an International Workshop on "Medieval Kerala Mathematics: its historical relevance and the possibility of its transmission to Europe", held at Kovalam, Kerala, India, Dec 15 and 16, 2005

References

Almeida, D. F. and Joseph, G. G. (2004) Eurocentrism in the History of Mathematics: the Case of the Kerala School, *Race & Class*, Vol. 45, No. 4, 2004: 5-59.

Almeida, D. F., John, J. K. and Zadorozhnyy, A. (2001) Keralese Mathematics: Its Possible Transmission to Europe and the Consequential Educational Implications, *Journal of Natural Geometry* Vol. **20:** 77-104.

Bernal, M. (1987) *Black Athena, The Afroasiatic roots of Classical Civilisation*, Vol. 1, London: Free Association Books.

Davis, P. J. and Hersh, R. (1980) *The Mathematical Experience*, Boston: Birkhauser.

Ernest, P. (forthcoming) 'A Semiotic Perspective of Mathematical Activity ', *Educational Studies in Mathematics*, Special issue on The Teaching and Learning of Mathematics: Semiotic and Epistemological Perspectives, Guest Editors: Adalira Sáenz-Ludlow, Norma Presmeg.

Ernest, P. (1991) *The Philosophy of Mathematics Education*, London, Falmer Press.

Ernest, P. (1998) *Social Constructivism as a Philosophy of Mathematics*, Albany, New York: SUNY Press.

Ernest, P. (forthcoming) A Semiotic Perspective of Mathematical Activity, *Educational Studies in Mathematics*, Accepted for publication.

Gödel, K. (1931) translation in Heijenoort, J. van Ed. (1967) *From Frege to Gödel: A Source Book in Mathematical Logic*, Cambridge, Massachusetts: Harvard University Press, 592-617.

Høyrup, J. (1980) Influences of Institutionalized Mathematics Teaching on the Development and Organisation of Mathematical Thought in the Pre-Modern Period, Bielefeld, extract reprinted in Fauvel, J. and Gray, J., Eds, *The History of Mathematics: A Reader*, London, Macmillan, 1987, 43-45.

Høyrup, J. (1994) *In Measure, Number, and Weight,* New York: SUNY Press.

Joseph, G. G. (1994) Different Ways of Knowing: Contrasting Styles of Argument in Indian and Greek Mathematical Traditions, in P. Ernest, Ed. *Mathematics, Education and Philosophy*, Falmer Press, London, 1994: 194-207.

Joseph, G. G. (2000). *The Crest of the Peacock, non-European roots of Mathematics*. Princeton and Oxford: Princeton University Press.

Kitcher, P. (1984) *The Nature of Mathematical Knowledge*, Oxford: Oxford University Press.

Kitcher, P. and Aspray, W. (1988) An opinionated introduction, in Aspray, W. and Kitcher, P. Eds. (1988) *History and Philosophy of Modern Mathematics*, Minneapolis: University of Minnesota Press, 3-57.

Körner, S. (1960) *The Philosophy of Mathematics*, London: Hutchinson.

Lakatos, I. (1962) Infinite Regress and the Foundations of Mathematics, *Aristotelian Society Proceedings*, Supplementary Volume No. 36, 155-184 (revised version in Lakatos, 1978).

Lakatos, I. (1976) *Proofs and Refutations*, Cambridge: Cambridge University Press.

Lakatos, I. (1978) *The Methodology of Scientific Research Programmes* (Philosophical Papers Volume 1), Cambridge: Cambridge University Press.

MacKenzie, D. (1993) Negotiating Arithmetic, Constructing Proof: The Sociology of Mathematics and Information Technology, *Social Studies of Science*, Vol. 23, 37-65.

Maddy, P. (1990) *Realism in Mathematics*, Oxford: Clarendon Press.

Pearce, I. G. (undated) Indian Mathematics: Redressing the balance, web material consulted on 18 November 2005 at <http://www-history.mcs.st-andrews.ac.uk/history/Projects/Pearce/index.html>

Popper, K. R. (1979) *Objective Knowledge* (Revised Edition), Oxford: Oxford University Press.

Powell, A. B. and Frankenstein, M. (1997) *Ethnomathematics: Challenging Eurocentrism in Mathematics*, Albany, New York: SUNY Press.

Rotman, B. (1987) *Signifying Nothing: The Semiotics of Zero*, London: Routledge.

Tymoczko, T. Ed. (1986) *New Directions in the Philosophy of Mathematics*, Boston: Birkhauser.

A Hexagon Result and its Generalization via Proof

Michael de Villiers[1]
University of KwaZulu-Natal, South Africa

Abstract

This paper presents the discovery of a hexagon result on Geometer's Sketchpad and its generalization via proof for any 2n-gon. The result is : If ABCDEF is a hexagon with opposite sides parallel (not necessarily equal), then the respective centroids G, H, I, J, K and L of triangles ABC, BCD, CDE, DEF, EFA and FAB, form a hexagon with opposite sides both equal and parallel.

Keywords: Discovery; Geometer's sketchpad; Hexagons; Generalization; Mathematical experimentation; 2n-gons; Proof

"The object of mathematical rigour has been only to sanction and legitimatize the conquests of intuition." – Jacques Hadamard about 1900 (in Kline, 1980:318)

1. Introduction

The above quote represents a fairly common myth, namely, that mathematics is mainly a product of intuition and experimentation, and that the only role of proof is to sanction these empirical discoveries. In the majority of textbooks at high school and university, the purpose of proof in mathematics is still presented almost exclusively as that of *verification*; i.e. only as a means of obtaining certainty and to eliminate doubt. Quite often the approach followed is to allow students to experimentally discover the results, and then to try and cast a little doubt on the process of experimentation as a general means of validation. Proof is then presented as a means of "*making absolutely sure*".

However, proving is not just about making sure. Particularly, given the very high level of conviction one can nowadays obtain through many different computer programs, proof may instead serve the purpose of a logical *explanation* of *why* a certain result is true (see De Villiers, 2003). Moreover, since a proof often provides valuable insight into why a result is true, it often immediately enables one to generalise or vary the result in different ways. Usually this happens during the "*looking back*" or "*reflective*" stage of Polya's famous model of problem solving

[1] University of KwaZulu-Natal, School of Science, Mathematics & Technology Education, Edgewood Campus, Ashwood 3605, South Africa. E-mail: profmd@mweb.co.za

(Polya 1945), and illustrates what I have called a "*discovery*" function of proof that is seldom emphasised in textbooks or teaching.

The purpose of this article is to give one example of a recent problem I worked on that illustrates this "*discovery*" function very well. The example should be well within reach of talented high school or under-graduate students, and can also be a good training problem for a Mathematics Olympiad. Other examples of this "*discovery*" function are given in De Villiers (1997 & 2003). *Sketchpad 4* sketches in zipped format (Winzip) of the problem and its generalisation can be downloaded directly from: http://mysite.mweb.co.za/residents/profmd/hexcentroids.zip

2. The Problem

I was recently exploring some properties of hexagons with opposite sides parallel with the aid of *Sketchpad* and discovered the following interesting result:

If *ABCDEF* is a hexagon with opposite sides parallel (not necessarily equal), then the respective centroids *G*, *H*, *I*, *J*, *K* and *L* of triangles *ABC*, *BCD*, *CDE*, *DEF*, *EFA* and *FAB*, form a hexagon with opposite sides both equal and parallel (see Figure 1).

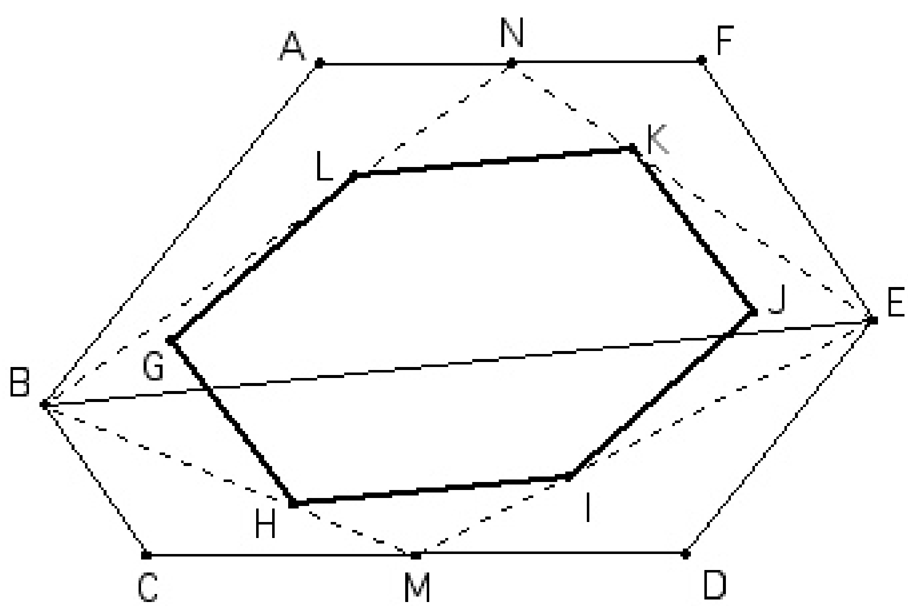


Figure 1

2.1. Proof

A problem like this may at first glance look quite challenging. Where does one start? However, a useful problem solving strategy is to find a way of relating or reducing the problem to results that are well known. One way of doing this is to start making some constructions by adding points and lines. Though this is no guarantee, and one may have to spend a little time experimenting, it allows one to get a better grip on the problem.

In this case, by drawing the diagonal *BE*, midpoints *N* and *M* respectively of *AF and CD*, and the medians *BN*, *BM*, *EN* and *EM*, a proof immediately pops out. For example, since *L* and *K*

are centroids, we have $NL = \frac{1}{3}NB$ and $NK = \frac{1}{3}NE$. From a well known high school theorem, it therefore follows in triangle *NBE* that *LK* // *BE* and $LK = \frac{1}{3}BE$. Similarly, it follows that *HI* // *BE* and $HI = \frac{1}{3}BE$. Thus, *LK* is equal and parallel to *HI*. In the same way, the two other pairs of opposite sides of *GHIJKL* can be shown to be equal and parallel, and completes the proof.

2.2. Looking back

Looking back over this proof, one can immediately see that nowhere is the result dependent on *ABCDEF* having opposite sides parallel. Thus, the result immediately generalises to ANY hexagon, i.e. the centroids of ANY hexagon form a hexagon with opposite sides equal and parallel!

2.3. Comment

Unfortunately the typical textbook or classroom teacher or lecturer is likely to just present the final hexagon generalisation and its proof above, thus missing an excellent pedagogical opportunity for teaching learners or students not only the value of "*looking back*", but also that proof has a very useful "*discovery*" function.

3. Further generalization

It seems natural to ask: Can the result perhaps be generalised further to perhaps other even sided polygons, for example, octagons, decagons, etc.?

Maybe just on the basis of intuition, one will perhaps try looking and testing with dynamic geometry software whether the centroids of triangles, *ABC*, *BCD*, etc. of an octagon *ABCDEFGH* also form an octagon with opposite sides parallel and equal. And perhaps at this point, readers should pause and first try it for themselves?

Unfortunately it does not work, as the reader would've found out by checking. So is it just a case of a result that just works for hexagons, or is there more to it?

Not on the basis of first experimenting, but on the basis of my knowledge of a related theorem and its proof (theorem given further down), I immediately anticipated that the hexagon result should further generalise to an octagon *ABCDEFGH* where the centroids of the 8 quadrilaterals *ABCD*, *BCDE*, *CDEF*, etc. form an octagon with opposite sides equal and parallel (see Figure 2). So this is a far more advanced example of the "*discovery*" function of a proof where one anticipates a result on the basis of related results and proof techniques, but another example nonetheless!

In fact, the result holds generally for a 2*n*-gon, $A_1A_2A_3...A_{2n}$ ($n \geq 2$), that the centroids of the *n*-gons, $A_1A_2A_3...A_n$, $A_2A_3A_4...A_{n+1}$, etc. sub-dividing it, form a 2*n*-gon with opposite sides equal and parallel. (Note that in the trivial case of a quadrilateral, the centroids of the *n*-gon become the centroids of the sides, and we obtain the Varignon parallelogram. So this result is really a generalisation of the Varignon parallelogram result).

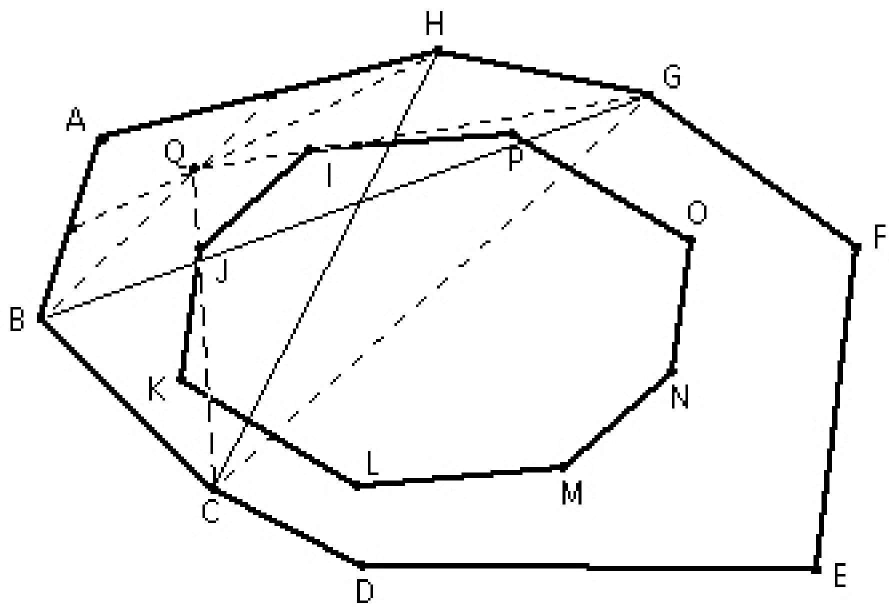


Figure 2

3.1. Proof

The general result depends on the following general theorem referred to above, and given and proved in De Villiers (1999) as well as Yaglom (1968): "Given a *n*-gon $A_1A_2A_3...A_n$ ($n > 3$)…, then the centroids of the (*n*-1)-gons, $A_1A_2A_3...A_{n-1}$, $A_2A_3A_4...A_n$, etc. that subdivide it, form a *n*-gon similar to the original *n*-gon with a scale factor of $\frac{1}{n-1}$, while the centre of similarity is the centroid of the original *n*-gon."

For example, for the given octagon in Figure 2, draw diagonal *CG*. Then from the above theorem $CJ = 3JQ$ because *J* is the centroid of quadrilateral *ABCH* and *Q* is the centroid of triangle *ABH*. Similarly, $GI = 3IQ$. Thus, JI // = $\frac{1}{3}CG$. Similarly, MN //= $\frac{1}{3}$ CG. Thus, JI //= MN. In the same way, the other pairs of opposite sides can be shown to be parallel and equal. It is also obvious that in exactly the same way using the above-mentioned theorem, the result can be proved for a decagon, duodecagon, etc.

3.2. Corollary

Due to the half-turn symmetry of the "inner" 2*n*-gons formed by the centroids, it also immediately follows that the diagonals connecting opposite vertices are concurrent at the centroid of the original 2*n*-gon.

Though this hexagon result and its generalisation are probably not original, I've not yet seen them in the literature available to me. However, I believe this interesting result can be used in much the same way as presented here to give students some appreciation for the discovery function of proof.

More generally, this example shows that mathematics is not just discovered via experimentation (or just deduction for that matter), but often involves a symbiotic interaction between the two processes as argued in De Villiers (2004). For example, sometimes experimentation leads to new results, but proving them can sometimes lead to further avenues of research and discoveries.

References

De Villiers, M. (1997). The Role of Proof in Investigative, Computer-based Geometry: Some personal reflections. Chapter in Schattschneider, D. & King, J. *Geometry Turned On!* Washington: MAA, pp. 15-24.

De Villiers, M. (1999). Generalizations involving maltitudes. *International Journal for Mathematical Education in Science & Technology*, 30(4), pp. 541-548.

De Villiers, M. (2003) *Rethinking Proof with Sketchpad 4*. Emeryville, CA: Key Curriculum Press.

De Villiers, M. (2004). The role and function of quasi-empirical methods in mathematics, *Canadian Journal of Science, Mathematics & Technology Education*, 4(3), Nov, pp. 397-418. (A copy of this article can be downloaded from: http://mysite.mweb.co.za/residents/profmd/experiment2.pdf

Kline, M. (1980). *Mathematics: The loss of certainty*. Oxford: Oxford University Press.

Polya, G. (1945) *How to solve it.* Princeton: Princeton University Press.

Yaglom, I.M., 1968, *Geometric Transformations II*, (Washington, DC: MAA), pp.24; 108-109.

Endnotes

One of my students recently found the hexagon centroid result listed at the Wolfram MathWorld site under the heading Centroid Hexagon at http://mathworld.wolfram.com/CentroidHexagon.html , but no proof is given nor mention of any further generalization is made. However, it is possible that the one of the references at this site contains a proof of the hexagon result & perhaps even the above-mentioned generalization to any 2n-gon. See:

Cadwell, J. H.(1966). *Topics in Recreational Mathematics.* Cambridge, England: Cambridge University Press, 1966.

Wells, D. (1991). *The Penguin Dictionary of Curious and Interesting Geometry.* London: Penguin, pp. 53-54.

On the Solution to Octic Equations

Raghavendra G. Kulkarni[1]
Senior Member IEEE, Deputy General Manager, HMC division, Bharat Electronics

Abstract
We present a novel decomposition method to decompose an eighth-degree polynomial equation, into its two constituent fourth-degree polynomials, as factors, leading to its solution. The salient feature of the octic equation solved here is that, the sum of its four roots being equal to the sum of the remaining four roots. We derive the condition to be satisfied by coefficients so that the given octic is solvable by the proposed method.

Key words: Octic equation; polynomials; factors; roots; coefficients; solvable polynomial equation; polynomial decomposition; solvable octic.

2000 Mathematics Subject Classification: 12D05: Polynomials (Factorization)

1. Introduction

It is well known from the works of Ruffini, Abel and Galois that the general polynomial equations of degree higher than the fourth cannot be solved in radicals [1 – 4]. This does not mean that there is no algebraic solution to these equations. The algebraic solutions to the general quintic have been obtained with symbolic coefficients. Hermite solved the Bring-Jerrard quintic using elliptic functions, and Klein gave a solution to the principal quintic using hypergeometric functions [5, 6]. The general sextic can be solved in terms of Kampe de Feriet functions, and a restricted class of sextics can be solved in terms generalized hypergeometric functions in one variable, using Klein's approach to solving the quintic equation [7]. There is not much literature on the solution to polynomial equations beyond sixth-degree, except in the form of numerical methods. We hope this paper will fill this gap to some extent.

In this paper we present a novel technique of decomposing a solvable octic equation into constituent fourth-degree polynomial factors, thereby facilitating the determination of all of its roots. The salient feature of the octic solved here is that, the sum of its four roots is equal to the sum of its remaining four roots. The condition required to be satisfied by the coefficients, in

[1] Dr. Raghavendra G. Kulkarni, Senior Member IEEE, Deputy General Manager, HMC division, Bharat Electronics Jalahalli Post, Bangalore-560013, INDIA. Phone: +91-80-22195270, Fax: +91-80-28382738,
Email: rgkulkarni@ieee.org

order that the given octic is solvable in radicals with the proposed technique, is derived. At the end of the paper we solve some numerical examples illustrating the applicability of this method.

2. Formulation of equations for solving the octic equation:

Let the octic equation for which the solution is sought be:

$$x^8 + a_7x^7 + a_6x^6 + a_5x^5 + a_4x^4 + a_3x^3 + a_2x^2 + a_1x + a_0 = 0 \qquad (1)$$

where a_0, a_1, a_2, a_3, a_4, a_5, a_6, and a_7 are the real coefficients. In the method proposed here, we attempt to represent the octic (1), in the form as shown below.

$$[(x^4 + b_3x^3 + b_2x^2 + b_1x + b_0)^2 - p^2(x^4 + c_3x^3 + c_2x^2 + c_1x + c_0)^2\,]/(1 - p^2) = 0 \qquad (2)$$

where b_0, b_1, b_2, b_3, and c_0, c_1, c_2, c_3 are the unknown coefficients of the respective fourth-degree polynomials in the above equation. The parameter, p, is also an unknown to be determined. The merit of representing the octic (1) in the form of (2) is obvious: notice that (2) can be easily factorized as:

$$\{[(x^4 + b_3x^3 + b_2x^2 + b_1x + b_0) - p(x^4 + c_3x^3 + c_2x^2 + c_1x + c_0)]/(1 - p)\}$$

$$\{[(x^4 + b_3x^3 + b_2x^2 + b_1x + b_0) + p(x^4 + c_3x^3 + c_2x^2 + c_1x + c_0)]/(1 + p)\} = 0 \qquad (3)$$

When each factorial term is equated to zero, we obtain the following two fourth-degree polynomial equations.

$$[(x^4 + b_3x^3 + b_2x^2 + b_1x + b_0) - p(x^4 + c_3x^3 + c_2x^2 + c_1x + c_0)]/(1 - p) = 0 \qquad (4)$$

$$[(x^4 + b_3x^3 + b_2x^2 + b_1x + b_0) + p(x^4 + c_3x^3 + c_2x^2 + c_1x + c_0)]/(1 + p) = 0 \qquad (5)$$

The solution to the above quartic equations can be obtained easily by the Ferrari's method. The roots of the quartic equations, (4) and (5), will be the roots of the given octic (1), when it gets represented in the form of octic (2). To achieve this the coefficients of octic (1) should be same as that of octic (2). However notice that the coefficients of octic (2) are not explicitly written. Therefore we expand and rearrange (2) in descending powers of x as shown below in order to expose the coefficients for comparison.

$$x^8 + [2(b_3 - c_3p^2)/(1 - p^2)]x^7 + \{[(b_3^2 + 2b_2) - (c_3^2 + 2c_2)p^2]/(1 - p^2)\}x^6$$
$$+ \{2[(b_1 + b_2b_3) - (c_1 + c_2c_3)p^2]/(1 - p^2)\}x^5$$
$$+ \{[(b_2^2 + 2b_0 + 2b_1b_3) - (c_2^2 + 2c_0 + 2c_1c_3)p^2]/(1 - p^2)\}x^4$$
$$+ \{2[(b_0b_3 + b_1b_2) - (c_0c_3 + c_1c_2)p^2]/(1 - p^2)\}x^3 + \{[(b_1^2 + 2b_0b_2) - (c_1^2 + 2c_0c_2)p^2]/(1 - p^2)\}x^2$$
$$+ [2(b_0b_1 - c_0c_1p^2)/(1 - p^2)]x + [(b_0^2 - c_0^2p^2)/(1 - p^2)] = 0 \quad (6)$$

Equating the coefficients of (1) and (6), we obtain the following eight equations as given below.

$$[2(b_3 - c_3p^2)/(1 - p^2)] = a_7 \quad (7)$$

$$\{[(b_3^2 + 2b_2) - (c_3^2 + 2c_2)p^2]/(1 - p^2)\} = a_6 \quad (8)$$

$$\{2[(b_1 + b_2b_3) - (c_1 + c_2c_3)p^2]/(1 - p^2)\} = a_5 \quad (9)$$

$$\{[(b_2^2 + 2b_0 + 2b_1b_3) - (c_2^2 + 2c_0 + 2c_1c_3)p^2]/(1 - p^2)\} = a_4 \quad (10)$$

$$\{2[(b_0b_3 + b_1b_2) - (c_0c_3 + c_1c_2)p^2]/(1 - p^2)\} = a_3 \quad (11)$$

$$\{[(b_1^2 + 2b_0b_2) - (c_1^2 + 2c_0c_2)p^2]/(1 - p^2)\} = a_2 \quad (12)$$

$$[2(b_0b_1 - c_0c_1p^2)/(1 - p^2)] = a_1 \quad (13)$$

$$[(b_0^2 - c_0^2p^2)/(1 - p^2)] = a_0 \quad (14)$$

There are nine unknowns (b_0, b_1, b_2, b_3, c_0, c_1, c_2, c_3, and p), but only eight equations [(7) to (14)] to solve. Therefore one more equation has to be introduced so that all the nine unknowns can be determined. However introducing an equation (involving the unknowns) imposes certain condition on the roots (and hence the coefficients) of the octic equation (1). The equation chosen will dictate the type of solvable octic. The equation, we are introducing here is as follows:

$$b_3 = a_7/2 \quad (15)$$

As we notice from our further analysis given in this paper, the above equation leads to a solvable octic in which the sum of its four roots is equal to the sum of its remaining four roots. Observe that the equation as given by (15) is not the only equation (or condition) to be introduced to solve the octic by this method. There are several ways of choosing the equation, like $b_2 = c_2$, or $b_1 = c_1$, or, $b_0 = c_0$, etc. However the equation introduced will decide the type of solvable octic. Substituting the value of b_3 [using (15)] in equation (7), c_3 is evaluated as:

$$c_3 = a_7/2 \quad (16)$$

The values of b_3 and c_3 are substituted in equations, (8) to (11), resulting in expressions, (8A) to (11A), respectively, as given below.

$$b_2 = c_2p^2 + F_2(1 - p^2) \qquad (8A)$$

where F_2 is given by:

$$F_2 = (4a_6 - a_7^2)/8$$

$$b_1 = c_1p^2 + (a_5/2)(1 - p^2) - (a_7/2)(b_2 - c_2p^2) \qquad (9A)$$

$$b_0 = c_0p^2 + (a_4/2)(1 - p^2) - [(b_2^2 - c_2^2p^2)/2] - (a_7/2)(b_1 - c_1p^2) \qquad (10A)$$

$$b_0 = c_0p^2 + (a_3/a_7)(1 - p^2) - (2/a_7)(b_1b_2 - c_1c_2p^2) \qquad (11A)$$

Now we have seven equations, (8A), (9A), (10A), (11A), (12), (13), and (14), with seven unknowns, b_0, b_1, b_2, c_0, c_1, c_2 and p, to be determined from them. We attempt to determine these unknowns through the process of elimination. Using the expression (8A), we eliminate b_2 from the equations, (9A), (10A), (11A), and (12), resulting in the following expressions, (9B), (10B), (11B), and (12A), respectively.

$$b_1 = c_1p^2 + F_1(1 - p^2) \qquad (9B)$$

where F_1 is:

$$F_1 = (a_5 - a_7F_2)/2$$

$$b_0 = c_0p^2 - (a_7/2)(b_1 - c_1p^2) + [a_4 - F_2^2 + (c_2 - F_2)^2p^2][(1 - p^2)/2] \qquad (10B)$$

$$b_0 = c_0p^2 + (a_3/a_7)(1 - p^2) - (2/a_7)\{[c_2p^2 + F_2(1 - p^2)]b_1 - c_1c_2p^2\} \qquad (11B)$$

$$b_1^2 + 2b_0[c_2p^2 + F_2(1 - p^2)] - (c_1^2 + 2c_0c_2)p^2 = a_2(1 - p^2) \qquad (12A)$$

Using (9B) we now eliminate b_1 from equations (10B), (11B), (12A), and (13), resulting in expressions (10C), (11C), (12B), and (13A) respectively as shown below.

$$b_0 = c_0p^2 + F_0(1 - p^2) + [(c_2 - F_2)^2\,(1 - p^2)\,p^2/2] \qquad (10C)$$

where F_0 is given by:

$F_0 = (a_4 - a_7F_1 - F_2^2)/2$, and

$$b_0 = c_0p^2 + [(a_3 - 2F_1F_2)/a_7](1 - p^2) + (2/a_7)(1 - p^2)(c_1 - F_1)(c_2 - F_2)p^2 \quad (11C)$$

$$2b_0(c_2 - F_2)p^2 + 2F_2b_0 - 2c_0c_2p^2 = [a_2 - F_1^2 + (c_1 - F_1)^2p^2](1 - p^2) \quad (12B)$$

$$b_0(c_1 - F_1)p^2 + F_1b_0 - c_0c_1p^2 = (a_1/2)(1 - p^2) \quad (13A)$$

We are left with five equations, (10C), (11C), (12B), (13A), and (14), involving five unknowns namely b_0, c_0, c_1, c_2, and p. Using (10C) we eliminate b_0 from equations (11C), (12B), (13A), and (14), resulting in the expressions, (11D), (12C), (13B), and (14A) as shown below.

$$\{p^2[(a_7/4)(c_2 - F_2)^2 - (c_2 - F_2)(c_1 - F_1)] - F_3\}(1 - p^2) = 0 \quad (11D)$$

where F_3 is given by:

$F_3 = (a_3 - a_7F_0 - 2F_1F_2)/2.$

$$\{p^2[p^2(c_2 - F_2)^3 + F_2(c_2 - F_2)^2 - 2(c_2 - F_2)(c_0 - F_0) - (c_1 - F_1)^2] - F_4\}(1 - p^2) = 0 \quad (12C)$$

where F_4 is given by:

$F_4 = (a_2 - F_1^2 - 2F_0F_2)$

$$\{(c_1 - F_1)[p^4(c_2 - F_2)^2 - 2p^2(c_0 - F_0)] + p^2F_1(c_2 - F_2)^2 - 2F_5\}(1 - p^2) = 0 \quad (13B)$$

Where F_5 is defined as:

$F_5 = (a_1 - 2F_0F_1)/2$

$$[p^2(c_0 - F_0)^2 - p^4(c_0 - F_0)(c_2 - F_2)^2 - p^2F_0(c_2 - F_2)^2 - (p^4/4)(1 - p^2)(c_2 - F_2)^4 + F_6](1 - p^2) = 0$$

(14A)

where F_6 is given by:

$F_6 = a_0 - F_0^2$

We observe that the term $(1 - p^2)$ emerges as a factor in the above equations [(11D), (12C), (13B), and (14A)]. However we cannot equate $(1 - p^2)$ to zero, since this term appears in the denominator in equation (2), and subsequently in many equations also. Therefore factoring out this term from the equations, (11D), (12C), (13B), and (14A), we obtain following expressions respectively.

$$p^2[(a_7/4)(c_2 - F_2)^2 - (c_2 - F_2)(c_1 - F_1)] - F_3 = 0 \quad (15)$$

$$p^2[p^2(c_2 - F_2)^3 + F_2(c_2 - F_2)^2 - 2(c_2 - F_2)(c_0 - F_0) - (c_1 - F_1)^2] - F_4 = 0 \quad (16)$$

$$(c_1 - F_1)[p^4(c_2 - F_2)^2 - 2p^2(c_0 - F_0)] + p^2F_1(c_2 - F_2)^2 - 2F_5 = 0 \quad (17)$$

$$p^2(c_0 - F_0)^2 - p^4(c_0 - F_0)(c_2 - F_2)^2 - p^2F_0(c_2 - F_2)^2 - (p^4/4)(1 - p^2)(c_2 - F_2)^4 + F_6 = 0 \quad (18)$$

At this stage we have four equations [(15), (16), (17), and (18)] involving four unknowns, c_0, c_1, c_2, and p. Continuing the process of elimination, we now eliminate $(c_1 - F_1)$ from equations, (16) and (17), using equation (15). The equations, (16) and (17), get transformed into equations, (16A) and (17A), respectively as shown below.

$$2p^4(c_0 - F_0)(c_2 - F_2)^3 = p^6(c_2 - F_2)^5 + p^4F_7(c_2 - F_2)^4 + p^2F_8(c_2 - F_2)^2 - F_3^2 \quad (16A)$$

where F_7 and F_8 are given by:

$F_7 = (16F_2 - a_7^2)/16,$

$F_8 = (a_7F_3 - 2F_4)/2.$

$$[8F_3 - 2a_7p^2(c_2 - F_2)^2](c_0 - F_0)$$
$$= 8F_5(c_2 - F_2) + 4p^2F_3(c_2 - F_2)^2 - 4p^2F_1(c_2 - F_2)^3 - p^4a_7(c_2 - F_2)^4 \quad (17A)$$

Now there are three equations [(16A), (17A), and (18)] containing three unknowns, c_0, c_2, and p. Using equation (17A), we eliminate $(c_0 - F_0)$ from equations, (16A) and (18), to obtain the equations, (16B) and (18A), respectively as given below.

$$p^6(c_2 - F_2)^6 + F_9p^4(c_2 - F_2)^4 + F_{10}p^2(c_2 - F_2)^2 + F_{11} = 0 \quad (16B)$$

Where F_9, F_{10}, and F_{11} are given by:

$F_9 = (a_7F_8 + 8F_5 - 4F_3F_7)/(a_7F_7 - 4F_1)$

$F_{10} = -[(a_7F_3^2 + 4F_3F_8)/(a_7F_7 - 4F_1)]$

$F_{11} = 4F_3^3/(a_7F_7 - 4F_1)$

$$p^8(c_2 - F_2)^8 + F_{12}\,p^6(c_2 - F_2)^6 + F_{13}\,p^4(c_2 - F_2)^4 + F_{14}\,p^2(c_2 - F_2)^2 + F_{15} = 0 \quad (18A)$$

where F_{12}, F_{13}, F_{14}, and F_{15} are given by:

$F_{12} = (4a_7^2F_0 – 16F_1^2 – 8a_7F_3)/a_7^2$

$F_{13} = (16F_3^2 + 64F_1F_5 – 4a_7^2F_6 – 32a_7F_0F_3)/a_7^2$

$F_{14} = (32a_7F_3F_6 + 64F_0F_3^2 – 64F_5^2)/a_7^2$

$F_{15} = – 64F_3^2F_6/a_7^2$

We are left with two equations, (16B) and (18A), involving two unknowns, c_2 and p. However since these unknowns (c_2 and p) occur only as a inseparable product term $[p(c_2 – F_2)]$ in both the equations [(16B) and (18A)], it will not be possible to determine them separately. This situation can be illustrated more clearly by observing the following simple example.

Let us attempt to determine the two variables, u and v, from the following two equations.

$uv + k = 0,$
$u^2v^2 + muv + n = 0$

where k, m, and n are coefficients in the above equations. Notice that the variables, u and v, occur as an inseparable product, uv, in these equations. By substituting the value of uv from the first equation in the second equation, we obtain following expression.

$k^2 – mk + n = 0$

Thus instead of getting values for u and v from the above two equations, what resulted is a relation among the coefficients. In a later section of the paper we use this technique to derive the condition to be satisfied by the coefficients of the given octic equation (1), so that the octic is solvable through the proposed method.

The above example illustrates that, we cannot determine c_2 and p from equations, (16B) and (18A), instead we can get an expression relating the coefficients (of the given octic) from these equations. Since we are left with no further equation, determining c_2 and p (separately) appears to be an impossible task. In the next section, we describe a technique by which the unknowns, c_2 and p, are successfully evaluated.

At present let us observe the equations (16B) and (18A) more closely. The equation (16B) is a sixth-degree polynomial equation in $p(c_2 – F_2)$, while equation (18A) is a eighth-degree polynomial equation in $p(c_2 – F_2)$. However since both of these equations contain only even powers of $p(c_2 – F_2)$, the degrees of these equations can be reduced to half by the following variable transformation.

$$g = p^2(c_2 – F_2)^2 \qquad (19)$$

Thus the equations (16B) and (18A) are transformed into cubic and quartic equations respectively as shown below.

$$g^3 + F_9 g^2 + F_{10} g + F_{11} = 0 \qquad (20)$$

$$g^4 + F_{12} g^3 + F_{13} g^2 + F_{14} g + F_{15} = 0 \qquad (21)$$

The cubic equation (20) yields three roots of g, while the quartic equation (21) provides four roots of g. The root, which is common to both the equations, (20) and (21), is the desired value of g, we are looking for. In the next paragraph, we describe a method to extract this common root.

The equations, (20) and (21), are rewritten as follows for our convenience.

$$g^3 = -(F_9 g^2 + F_{10} g + F_{11}) \qquad (20A)$$

$$(g + F_{12}) g^3 = -(F_{13} g^2 + F_{14} g + F_{15}) \qquad (21A)$$

Using (20A), we substitute the value of g^3 in (21A) to obtain following expression.

$$F_9 g^3 + (F_{10} + F_9 F_{12} - F_{13}) g^2 + (F_{11} + F_{10} F_{12} - F_{14}) g + F_{11} F_{12} - F_{15} = 0 \qquad (21B)$$

Again using (20A), g^3 is eliminated from (21B), resulting in the following quadratic equation.

$$(F_{10} + F_9 F_{12} - F_{13} - F_9^2) g^2 + (F_{11} + F_{10} F_{12} - F_{14} - F_9 F_{10}) g + F_{11} F_{12} - F_{15} - F_9 F_{11} = 0 \qquad (21C)$$

For the sake of convenience, the above quadratic is rearranged as follows.

$$g^2 + h_1 g + h_0 = 0 \qquad (21D)$$

where h_0 and h_1 are given by:
$h_0 = (F_{11} F_{12} - F_{15} - F_9 F_{11})/(F_{10} + F_9 F_{12} - F_{13} - F_9^2)$
$h_1 = (F_{11} + F_{10} F_{12} - F_{14} - F_9 F_{10})/(F_{10} + F_9 F_{12} - F_{13} - F_9^2)$

Using (21D), the value of g^2 is substituted in the cubic equation (20A), to obtain the following quadratic equation.

$$h_1 g^2 + (h_0 + F_9 h_1 - F_{10}) g + F_9 h_0 - F_{11} = 0 \qquad (20B)$$

Again using (21D), we eliminate g^2 from (20B) to obtain a linear equation in g as follows.
$(h_0 + F_9h_1 - F_{10} - h_1^2)g + F_9h_0 - F_{11} - h_0h_1 = 0$

From the above linear equation the common root of g is found out as:
$g = h_2/h_3$ (20C)

where h_2 and h_3 are given by:
$h_2 = h_0h_1 + F_{11} - F_9h_0$
$h_3 = h_0 + F_9h_1 - F_{10} - h_1^2$

Once g is determined, the product, $p(c_2 - F_2)$, can be evaluated from (19). However to determine c_2 and p separately in the absence of any further equation requires some novel technique, which will be presented in the next section.

3. A discussion on the value of p^2

From equations (11D), (12C), (13B), and (14A), we notice that the term $(1 - p^2)$ emerges as a factor in these equations. However we were constrained not to equate p^2 to unity, as it amounts to division by zero in equation (2). Instead let us examine the consequences, when p^2 approaches unity, but will not attain unity value. In other words we are applying the limiting process and evaluating the expressions (that contain p), when p^2 tends to unity. Thus as a first step, let us evaluate c_2 in the limit as p^2 tends to unity, by rearranging the expression (19) and applying the limit as shown below.

$$c_2 = \lim_{p^2 \to 1} [F_2 \pm (g/p^2)^{1/2}] \qquad (22)$$

Simplifying (22) results in two values of c_2 as:

$c_{21} = F_2 + (g)^{1/2}$

$c_{22} = F_2 - (g)^{1/2}$ (23)

Consider the expression (15). After rearranging (15) and substituting for $(c_2 - F_2)$ by utilizing (19), we apply the limit as p^2 tends to unity to determine c_1, as shown below:

$$c_1 = \lim_{p^2 \to 1} \{F_1 + (a_7/4p)[\pm (g)^{1/2}] - (F_3/p)[\pm 1/(g)^{1/2}]\}$$

After simplifying, the above expression yields two values of c_1 (corresponding to two values of c_2 respectively) as:

$c_{11} = F_1 + \{(a_7/4)(g)^{1/2} - [F_3/(g)^{1/2}]\}$

$$c_{12} = F_1 - \{(a_7/4)(g)^{1/2} - [F_3/(g)^{1/2}]\} \tag{24}$$

In the same manner, the expression (17A) is rearranged and the term, $(c_2 - F_2)$, is eliminated using (19), and then limit as p^2 approaches unity, is applied as shown below, to facilitate evaluation of c_0.

$$c_0 = \lim_{p^2 \to 1} \{F_0 + [(4F_3g - a_7g^2)/(8F_3 - 2a_7g)] + [(\pm)(8F_5 - 4F_1g)(g)^{1/2}]/[p(8F_3 - 2a_7g)]\}$$

Simplifying the above expression, we obtain two values for c_0 (corresponding to two values of c_2 respectively) as:

$$c_{01} = F_0 + (g/2) + [(8F_5 - 4F_1g)/(8F_3 - 2a_7g)](g)^{1/2}$$

$$c_{02} = F_0 + (g/2) - [(8F_5 - 4F_1g)/(8F_3 - 2a_7g)](g)^{1/2} \tag{25}$$

In the same fashion, we determine b_0, b_1, and b_2, by applying the limit to the expressions (10C), (9B), and (8A) respectively. The values of b_0, b_1, and b_2 are obtained as shown below.

$$b_0 = c_0$$
$$b_1 = c_1$$
$$b_2 = c_2 \tag{26}$$

We have determined all the unknowns in the octic (2), which means we are able to successfully represent the given octic (1) in the form of (2). Using the results of (26), along with the earlier determined values of b_3 and c_3, the octic (2) gets converted into an interesting octic as shown below.

$$[x^4 + (a_7/2)x^3 + c_2x^2 + c_1x + c_0]^2 = 0 \tag{27}$$

Looking at the above equation (27) emerged after evaluation of all unknowns, one may feel little disappointed thinking that, what resulted after all the exhaustive mathematics, is a tame octic equation with repeated roots. However it is not the complete story, as the next section reveals that, the octics with distinct roots also can be solved using this approach.

4. Decomposition of the octic equation:

Since c_0, c_1, and c_2 have two values each, equation (27) yields two distinct quartic polynomials as factors of octic equation (2) as shown below.

$$[x^4 + (a_7/2)x^3 + c_{21}x^2 + c_{11}x + c_{01}][x^4 + (a_7/2)x^3 + c_{22}x^2 + c_{12}x + c_{02}] = 0 \quad (28)$$

Equation (28) proves that, we have arrived at the solution to the given octic equation (1), by decomposing it into a pair of quartic polynomials as its factors. When each of the quartic factorial term in the above equation is equated to zero, we obtain two quartic equations as shown below.

$$x^4 + (a_7/2)x^3 + c_{21}x^2 + c_{11}x + c_{01} = 0$$

$$x^4 + (a_7/2)x^3 + c_{22}x^2 + c_{12}x + c_{02} = 0 \quad (29)$$

These quartics can be solved by the well-known Ferrari's method to obtain four roots each, and eight roots in total, which are the required roots of the given octic equation (1). In the coming sections we shall study the behavior of the roots, and the condition to be satisfied by the coefficients.

5. Behavior of the roots

Let x_1, x_2, x_3, and x_4 be the roots of first quartic equation, and x_5, x_6, x_7, and x_8 be roots of second quartic equation in the equation set (29). Notice that the coefficients of x^3 in these quartics are equal, which means that the sum of the roots of first quartic is equal to that of second quartic, as shown below.

$$x_1 + x_2 + x_3 + x_4 = x_5 + x_6 + x_7 + x_8 \quad (30)$$

Individually each sum (of four roots) is given by:

$$x_1 + x_2 + x_3 + x_4 = -(a_7/2)$$

$$x_5 + x_6 + x_7 + x_8 = -(a_7/2) \quad (31)$$

From (30) we note that, the given octic equation is solvable by this method if the sum of one group of its four roots is equal to the sum of its remaining four roots. From (31) we observe that the sum of the four roots in each group (and hence in each quartic equation) is a real number. Since equation (30) relates the roots of the given octic equation (1), we note that one of the roots can be expressed in terms of remaining seven roots, and hence this root is not independent. If we denote x_8 as a dependent root, then it is expressed in terms of other independent roots as:

$x_8 = x_1 + x_2 + x_3 + x_4 – (x_5 + x_6 + x_7)$

Since the roots of the given octic are related, the coefficients also have to be related, and in the next section we shall derive the condition to be satisfied by the coefficients.

6. Condition for the coefficients

From the expression (20C) we note that the parameter, g, is evaluated in terms of parameters, F_9, F_{10}, F_{11}, h_0, and h_1. Again note that h_0 is a function of F_9, F_{10}, F_{11}, F_{12}, F_{13}, and F_{15}, whereas h_1 is a function of F_9, F_{10}, F_{11}, F_{12}, F_{13}, and F_{14}. Thus eventually g will be a function of F_9, F_{10}, F_{11}, F_{12}, F_{13}, F_{14}, and F_{15}, which are functions of the coefficients of octic (1). Hence g is ultimately a function of coefficients, a_0, a_1, a_2, a_3, a_4, a_5, a_6, and a_7. Using equation (20C) we substitute the value of g in the quadratic equation (21D) to obtain the following expression.

$$(F_9h_0 – F_{11} – h_0h_1)^2 + h_1(F_9h_0 – F_{11} – h_0h_1)(h_1^2 – F_9h_1 – h_0 + F_{10}) + h_0(h_1^2 – F_9h_1 – h_0 + F_{10})^2 = 0 \qquad (32)$$

The expression (32) provides a relation among the coefficients, a_0, a_1, a_2, a_3, a_4, a_5, a_6, and a_7. These coefficients have to satisfy the condition (32) in order that the given octic is solvable in this fashion. In the following numerical examples we find out the roots of the solvable octic, whose coefficients satisfy the condition (32).

7. Numerical examples

Following numerical examples enhance our understanding of the proposed method. Consider the octic equation shown below.

$$x^8 + 2x^7 – 25.1475x^6 – 62.86875x^5 + 51.94875x^4 + 95.47125x^3 – 78.72625x^2 + 17.3725x – 1.05 = 0$$

The coefficients of above octic have been obtained from the roots, which satisfy the relation (30). Using the expressions (15) and (16), the values of b_3 and c_3 obtained are: $b_3 = 1$, $c_3 = 1$. The parameters, F_0 to F_{15}, are determined as shown below, from the corresponding expressions derived in section 2.

$F_0 = – 41.12646255474101$, $F_1 = – 18.36062622070313$, $F_2 = – 13.07374954223633$,

$F_3 = – 151.1801398726457$, $F_4 = – 1491.192986908885$, $F_5 = – 746.4213570143718$,

$F_6 = – 1692.435922218831$, $F_7 = – 13.32374954223633$, $F_8 = 1340.012847036239$,

$F_9 = -242.5149916348265$, $F_{10} = 16339.83020529136$, $F_{11} = -295355.2632059779$,

$F_{12} = -908.2356715938659$, $F_{13} = 217987.6293012944$, $F_{14} = -19859925.14863872$,

$F_{15} = 618901739.0483127$.

Next, the parameters, h_0, h_1, h_2, and h_3, are determined from the corresponding expressions given in section 2 as shown below.

$h_0 = 10504.27396520806$, $h_1 = -216.0867489139377$

$h_2 = -17745.7612467463$, $h_3 = -124.7631910180035$

The parameter, g, is found out from (20C) as:

$g = 142.2355512226805$

From this value of g the coefficients of constituent quartic equations given in (29) are determined as shown below.

$c_{21} = -1.147495$, and, $c_{22} = -25$,

$c_{11} = 0.2787476$, and, $c_{12} = -37$,

$c_{01} = -0.01749611$, and, $c_{02} = 60$.

The quartic equations formed with these coefficients are:

$$x^4 + x^3 - 1.147495x^2 + 0.2787476x - 0.01749611 = 0$$

$$x^4 + x^3 - 25x^2 - 37x + 60 = 0$$

Solving the above quartic equations by Ferrari's method, all the eight roots of the octic equation are found out as: 0.1, 0.25, 0.4, – 1.75 (for the first quartic), and, 1, – 3, – 4, 5 (for the second quartic).

Let us solve another octic equation shown below.

$$x^8 + 2x^7 - 11.875x^6 - 12.5625x^5 + 45.66016x^4 + 14.58203x^3 - 53.34473x^2 + 1.620117x + 12.91992 = 0$$

As in the previous example, the coefficients of above octic have been obtained from the roots, which satisfy the relation (30). Using the expressions (15) and (16), the values of b_3 and c_3 obtained are: $b_3 = 1$, $c_3 = 1$. The parameters, F_0 to F_{15}, are determined as shown below, from the corresponding expressions derived in section 2.

$F_0 = 1.953125$, $F_1 = 0.15625$, $F_2 = -6.4375$, $F_3 = 6.34375$, $F_4 = -28.22265625$,

$F_5 = 0.5048828125$, $F_6 = 9.105224609375$, $F_7 = -6.6875$, $F_8 = 34.56640625$,

$F_9 = -17.34765625$, $F_{10} = 68.400634765625$, $F_{11} = -72.94073486328125$,

$F_{12} = -17.66015625$, $F_{13} = -72.42822265625$, $F_{14} = 2177.700668334961$,

$F_{15} = -5862.76876449585$.

Next, the parameters, h_0, h_1, h_2, and h_3, are determined from the corresponding expressions as shown below.

$h_0 = 40.2431640625$, $h_1 = -15.53515625$, $h_2 = 0$, and $h_3 = 0$.

Notice from (20C) that g becomes indeterminate since $h_2 = 0$, and $h_3 = 0$. This indicates there are more (than one) common roots between the cubic equation (20) and the quartic equation (21). Therefore we evaluate g from the quadratic equation (21D). The roots (g_1, g_2) of quadratic equation, (21D), are the common roots between (20) and (21), and are given by:

$g_1 = 3.28515618609$, $g_2 = 12.25$.

More common roots indicate that there are as many ways to group the eight roots, which satisfy the condition (30). Let us choose one of the two common roots, $g = 12.25$, and then proceed to determine the coefficients, c_{01}, c_{11}, c_{21}, and c_{02}, c_{12}, c_{22}, of quartic equations given in (29). These coefficients are found out as:

$c_{21} = -2.9375$, and, $c_{22} = -9.9375$.

$c_{11} = 0.09375$, and, $c_{12} = 0.21875$,

$c_{01} = 0.84375$, and, $c_{02} = 15.3125$.

The two quartic equations formed with these coefficients are as shown below.

$x^4 + x^3 - 2.9375x^2 + 0.09375x + 0.84375 = 0$

$x^4 + x^3 - 9.9375x^2 + 0.21875x + 15.3125 = 0$

Solving the above quartics by Ferrari's method, we obtain their roots as: 0.75, 1, – 0.5, – 2.25 for the first quartic; and 1.75, 2, – 1.25, – 3.5 for the second quartic.

If we choose the other common root, $g = 3.28515618609$, then the two quartic equations formed are as shown below.

$x^4 + x^3 - 8.25x^2 + 2.75x + 3.5 = 0$

$x^4 + x^3 - 4.625x^2 - 2.4375x + 3.691405 = 0$

The roots of the first quartic are: 1, 2, – 0.5, and, – 3.5, and the roots of second quartic are: 0.75, 1.75, – 1.25, and, – 2.25. Note that these roots are same as that obtained earlier (with $g = 12.25$).

Let us solve one more octic equation, whose roots are complex. Consider the following octic:

$x^8 - 10x^7 + 53x^6 - 166x^5 + 389x^4 - 790x^3 + 1787x^2 - 2314x + 1690 = 0$

Again the coefficients in the above equation are determined from the roots, which satisfy the condition (30). Using the expressions (15) and (16), b_3 and c_3 are obtained as: $b_3 = -5$, $c_3 = -5$. The parameters, F_0 to F_{15}, are determined as shown below, from the corresponding expressions derived in section 2.

$F_0 = 31.5$, $F_1 = -13$, $F_2 = 14$, $F_3 = -55.5$, $F_4 = 736$, $F_5 = -747.5$, $F_6 = 697.75$, $F_7 = 7.75$,

$F_8 = -458.5$, $F_9 = -12.76470588235294$, $F_{10} = 2783.705882352941$,

$F_{11} = 26816.29411764706$, $F_{12} = 54.56$, $F_{13} = -1673.36$, $F_{14} = -171585.76$,

$F_{15} = -1375516.44$.

Next, the parameters, h_0, h_1, h_2, and h_3, are determined from the corresponding expressions as shown below.

$h_0 = 884.1560157086755$, $h_1 = 107.2395573009639$

$h_2 = 132918.7853218638$, $h_3 = -14768.75392465153$.

Using (20C) the value of g is found out as: $g = -9$. Since the value of g is negative, the coefficients of the quartic equations shown in (29) are complex; as can be seen from the expressions (23), (24), and (25). Thus the coefficients, $[c_{21}, c_{22}]$, $[c_{11}, c_{12}]$, and, $[c_{01}, c_{02}]$, are evaluated from (23), (24), and, (25), as given below.

$c_{21} = 14 + 3i$, and $c_{22} = 14 - 3i$,

$c_{11} = -13 - 26i$, and, $c_{12} = -13 + 26i$,

$c_{01} = 27 + 31i$, and $c_{02} = 27 - 31i$.

The quartic equations formed with the above coefficients are:

$$x^4 - 5x^3 + (14 + 3i)x^2 - (13 + 26i)x + 27 + 31i = 0$$

$$x^4 - 5x^3 + (14 - 3i)x^2 - (13 - 26i)x + 27 - 31i = 0$$

Above quartic equations are solved to obtain the eight roots of the given octic equation. The roots are found out as:

$(1 - 1i)$, $(2 - 3i)$, $(3 + 2i)$, $(-1 + 2i)$, for the first quartic; $(1 + 1i)$, $(2 + 3i)$, $(3 - 2i)$, $(-1 - 2i)$, for the second quartic.

The numerical calculations in these examples are performed using BASICA software in double precision mode.

8. Conclusions

A novel polynomial decomposition technique is presented, to solve certain solvable octic equations. The criteria for the roots and the coefficients to satisfy, in order that the octic is solvable by the method given, are derived. Some numerical examples are solved using the proposed method.

Acknowledgments

The author thanks the management of Bharat Electronics for supporting this work.

References

[1]. G. Birkhoff and S. MacLane (1996). *A survey of modern algebra (*5th edition). Macmillan, New York.
[2]. B. R. King (1996). *Beyond the quartic equation*. Birkhauser, Boston.
[3]. M. I. Rosen (1995). Niels Hendrik Abel and equations of fifth degree. *American Mathematical Monthly*, 102, 495-505.
[4]. R. G. Kulkarni (2006). A versatile technique for solving quintic equations, *Mathematics and Computer Education*, 205 – 215.
[5]. Eric W. Weisstein. *Quintic Equation*. From MathWorld—A Wolfram Web Resource, http://mathworld.wolfram.com/QuinticEquation.html.
[6]. R. J. Drociuk. *On the complete solution to the most general fifth degree polynomial*. ,
3-May 2000, http://arxiv.org/abs/math.GM/0005026/.
[7]. Eric W. Weisstein. *Sextic Equation*. From MathWorld—A Wolfram Web Resource, http://mathworld.wolfram.com/SexticEquation.html.

Mathematics Education and Neurosciences: Relating Spatial Structures to the Development of Spatial Sense and Number Sense

Fenna van Nes[1] & Jan de Lange
Freudenthal Institute for Science and Mathematics Education, The Netherlands

Abstract

The Mathematics Education and Neurosciences (MENS) project is aimed at exploring the development of the mathematical abilities of young (four- to six-year old) children. It is initiated to integrate research from mathematics education with research from educational neuroscience in order to come to a better understanding of how the early skills of young children can best be fostered for supporting the development of mathematical abilities in an educational setting. This paper is specifically focused on the design research that is being conducted from the perspective of mathematics education in which we are investigating the relationship between young children's insight into spatial structures and the development of spatial and number sense. This should result in a series of classroom activities that may stimulate children's development of spatial and number skills.

Keywords: young children, spatial thinking, design research

1. Introduction: The Project in Context

It may come as no surprise that several publications support the point that we, the educational researchers, have been failing to properly value the cognitive capacities of young (three- to six-year old) children. A report from the National Research Council (NRC, 2005) concluded that

> early childhood education, in both formal and informal settings, may not be helping all children maximize their cognitive capacities.

It is also clear that there is an increasingly critical attitude towards some of Piaget's work. The aforementioned report concludes that 'modern research describes unexpected competencies in young children and calls into action models of development based on Piaget, which suggested

[1] Fenna van Nes, MSc, PhD student
Freudenthal Institute for Science and Mathematics Education, Utrecht University
Aidadreef 12 3561 GE Utrecht; PO Box 9432 3506 GK Utrecht, The Netherlands
phone +31 (0)30 2635570 (office direct); phone +31 (0)30 2635555 (office general)
fax +31 (0)30 2660430 (office)
www.fi.uu.nl

that children were unable to carry out sophisticated complex tasks, such as perspective taking' (NRC, 2005). Remarkably, scientists from a different discipline, the education neurosciences, have come to similar conclusions in a report on Numeracy and Literacy of young children (OECD, 2003).

The learning of young children is so intriguing, that it has engaged many different scientific disciplines. What is surprising, then, is that the brain scientists see no references in the educational research literature about the developments they see to be relevant and vice versa. Yet, the tide is changing. The OECD report 'Understanding the Brain. Towards a New Learning Science' (OECD, 2002) suggests trans-disciplinary research to be the way forward. This must bridge the brain sciences (called the 'hard sciences' by brain scientists) and ('more practical') educational research (Jolles et al., 2006).

Reflecting on these issues, a program called *TalentPower* (TalentenKracht) was initiated by van Benthem, Dijkgraaf and de Lange. Several universities and institutions in The Netherlands collaborate in this program to gain a better understanding of what talents, possibilities and qualities young children exhibit as they are engaged in scientific activities, how these talents and qualities may be enhanced, how they may be intertwined, and in what ways they may be connected to language development. Hence the goal of the project is to bring together scientists from various research perspectives, as well as parents and teachers in order to chart the talents of young children and to scientifically fundament how these talents may be used and developed in an optimal way (van Benthem, Dijkgraaf & de Lange, 2005).

As such, apart from the fundamental goal to investigate the possibilities of fostering young children's natural curiosity, an important goal in the methodology of *TalentPower* is the 'trans-disciplinary' approach. Given the abundance of research in the field of mathematics education, the project was designed to try to bridge the gap between the sciences of 'mathematics education' and 'educational neurosciences'. This is how the Mathematics Education and Neurosciences (MENS) project came into being.

The significance of the collaboration between the sciences lies in the grounding of research from the field of mathematics education in cognitive and neuroscientific theory while at the same time providing the research from the field of cognitive psychology and neurosciences with a strong practical basis from which testable predictions can be made. Many recent publications have emphasized how scientists from the disciplines of mathematics education, cognitive psychology and neuropsychology can and should contribute to each others research (Berninger & Corina, 1998; Byrnes & Fox, 1998; Davis, 2004; Griffin & Case, 1997; Jolles et al., 2006; Lester, 2007; Siegler, 2003; Spelke, 2002). As Cobb (2007) points out, comparing and contrasting research from various perspectives has the added benefit of deepening our understanding of the phenomena being studied and of broadening the practicality of the results.

Within the context of the development of mathematical abilities of young children, the authors of this paper are mainly concerned with the mathematics educational perspective. De Haan and Gebuis at Utrecht University are constructing and performing the educational neuroscientific experiments. In time, the results of research from these research perspectives will be compared, contrasted and combined in an effort to contribute to mathematics educational practices that can

foster the early talents of young children. Ultimately our findings may help stimulate those children who may be prone to experiencing problems in the development of mathematical thinking.

In the present paper we spiral into the work that is being performed from the perspective of mathematics education. In the first part of the paper, we lay out the theoretical framework of our research. This starts with a rationale for our focus on young children, on the constructs of spatial sense and number sense and on the role of spatial structures in the development of mathematical abilities. Next, we introduce preliminary experimental support that has contributed to the refinement of the research questions, and finally we explain our research methodology. We begin with the primary interest of all mathematics research in *TalentPower*: the importance of attending to how young children develop in their mathematical thinking.

2. Young Children Doing Mathematics

The overwhelming scientific attention to the mathematics education of young children can be attributed to seven factors that Clements and Sarama (2007) articulate: that a growing number of children attend early care and education programs, that the importance of mathematics is increasingly being recognized, that differences in performance between nations as well as between socioeconomic groups exist, that researchers are shifting to a perspective that recognizes innate mathematical competencies, that mathematics achievement is strongly predicted by specific quantitative and numerical knowledge, and finally, that knowledge gaps often appear because of poor bridging between informal knowledge and school mathematics.

What repeatedly stands out from studies on development in early childhood is how young children may be characterized by their natural drive to go out and explore the world. This is particularly illustrated in research stemming from Piaget's work. As mentioned in the introduction above, however, Piaget's methodology has strongly been criticized by researchers such as Freudenthal for depending too much on expert-use and interpretation of underlying concepts and on the child's language skills (Freudenthal, 1984, 1991). Freudenthal was greatly concerned about the intertwinement of children's cognitive competencies with their language skills, where relatively underdeveloped language skills could potentially suppress how children may express their understanding. Research methodologies that relied on children's ability to communicate their thinking could, in his view, only assess this language component and nothing more. Yet, Freudenthal's experiences with young children convinced him that children typically do possess remarkable cognitive competencies that develop through early learning processes.

Children's early competencies have been compared to the behavior of scientists in the Theory Theory (Gopnik, 2004; Gopnik, Meltzoff, & Kuhl, 1999). She suggests that children are born with certain theories about the world that they continuously test and amend as they gain new insights from daily experiences. Certain parallels are also drawn between children, scientists and poets who resemble one another in their sense of wonder and in the intense way in which they experience the world (Gopnik et al., 1999). As Dijkgraaf (2007) observes: 'It is often said that young children are ideal scientists. They are curious about the world around them. They ask questions, make up theories, and carry out experiments.' This is what is said to give both scientists as well as children their drive to learn (Gopnik, 2004).

In summary, de Lange emphasizes the 'curious minds of young children' (de Lange, cited in Ros, 2006, p. 9) which 'have to be stimulated'. In this sense it is disconcerting to note that many early elementary mathematics curricula focus mainly on developing curricula that teach number sense (Casey, 2004; Clements & Battista, 1992). Indeed several researchers warn about the gap that has been observed at the start of formal schooling between children's informal, intuitive knowledge and interests, and the formal learning opportunities in school (cf. Griffin & Case, 1997; Hughes, 1986; Murphy, 2006). The key point that we are making, then, is that mathematics education for young children should intertwine with and originate from the natural experiences, the enthusiasm, and the interests of young children as they explore of the world.

Gopnik (2004) put the issue for science in general into the following words:

> If we could put children in touch with their inner scientists, we might be able to bridge the divide between everyday knowledge and the apparently intimidating and elite apparatus of formal science. We might be able to convince them that there is a deep link between the realism of everyday life and scientific realism (p 28).

Through acknowledging the early competencies of young children (concentrating on what the children can already do versus what they cannot yet do; see also Gelman & Gallistel, 1978), we should on the one hand be able to come to a greater understanding about what factors influence the development of mathematical thinking and learning, while, on the other hand, stimulating the child's innate curiosity and eagerness to learn mathematics. We focus our research on spatial sense and number sense, the core of mathematics in the early years (NCTM, 2000), and study whether and, if so, how the development of early spatial sense and emerging number sense may be related. For purposes of our argument, we now clarify what we understand to be number sense and spatial sense.

3. Emerging Number Sense

The concept of number sense can broadly be defined as the ease and flexibility with which children operate with numbers (Gersten & Chard, 1999). Berch (1999) compiled an extensive list of components that have been related to the construct of number sense from the literature of mathematical cognition, cognitive development, and mathematics education. As such, he states that

> possessing number sense ostensibly permits one to achieve everything from understanding the meaning of numbers to developing strategies for solving complex math problems; from making simple magnitude comparisons to inventing procedures for conducting numerical operations; and from recognizing gross numerical errors to using quantitative methods for communicating, processing, and interpreting information. (p. 334)

As children progress in their ability to count, they discover easier ways of operating with numbers and they come to understand that numbers can have different representations and can act as different points of reference (Berch, 1999; Griffin & Case, 1997; Van den Heuvel-

Panhuizen, 2001). Given the diversity of the definitions of number sense, we focus our research on the development of awareness of quantities, on learning to give meaning to quantities and on being able to relate the different meanings of numbers to each other. This knowledge can then be applied to determining a quantity, to comparing quantities and to preliminary adding and subtracting. Hence, a well-founded number sense is fundamental to the ease and level of understanding with which children progress to higher order mathematical skills and concepts.

Our focus on young children's ability to determine a quantity and to compare quantities is supported by the Central Conceptual Theory described by Griffin and Case (1997; Griffin, 2004b). This theory is grounded in cognitive research with findings on how children by the age of four can make global quantity comparisons and can count. As Gelman and Gallistel (1978) have shown, children by the age of four can count a set of objects and understand that the last named number word represents the quantity of the set. Much recent cognitive research has supported this finding and has extended it to mathematics operations. Berger, Tzur and Posner (2006), for instance, found that six-month old infants can recognize simple addition errors and that the corresponding brain activity can be compared to that of adults detecting an arithmetic error.

Apart from children's ability to count, research by Starkey (1992), for example, has shown that four-year olds possess numerical knowledge that is not yet numerical, but that allows them to make quantity comparisons. Indeed, more recent cognitive psychological research on children's numerical abilities has provided evidence on how infants as young as six months can differentiate between amounts of objects that differ by a 2.0 ratio (i.e. eight versus sixteen objects; Lipton & Spelke, 2003; Xu & Spelke, 2000). This ability has been seen to improve within months as nine-month old infants can already differentiate sets that differ in number at a 1.5 ratio (i.e. nine versus six objects).

Griffin and Case (1997) describe the ability to compare quantities and the ability to count initially as two separate schemas. At the age of four, children have difficulty integrating these competencies, as if 'the two sets of knowledge were stored in different "files" on a computer, which cannot yet be "merged"' (p. 8). A revolutionary developmental step is said to occur by the age of five or six, in which these two schemas merge into 'a single, super-ordinate conceptual structure for number' (Griffin, 2004a, p. 40) in a manner that is described in the Central Conceptual Structure Theory (Griffin, 2004b; Griffin & Case, 1997). Such a conceptual structure covers 'the intuitive knowledge that appears to underlie successful learning of arithmetic in the early years of formal schooling' (1997, p. 8). It connects an understanding of quantity with number and enables children to use numbers without having to rely on objects that are physically present. Hence, this new conceptual structure provides children with the conceptual foundation for number sense which is believed to fundament all higher-level mathematics (Griffin, 2004a).

The learning of number and operations in early childhood may be the best-developed area in mathematics education research (Baroody, 2004; Clements, 2004; Fuson, 2004; Steffe, 2004). Yet, other research has shown that spatial thinking skills and mathematics achievement of relatively older children are related (Bishop, 1980; Clements, 2004; Guay & McDaniel, 1977; Smith, 1964; Tartre, 1990a, 1990b). For this reason, the NCTM standards (1989, 2000) strongly recommend increasing the emphasis on the development of spatial thinking skills through the

teaching of geometry (the mathematics of space; Bishop, 1983) and spatial sense. In the next section we discuss three components of spatial sense that we consider to play an essential role in the development of young children's mathematical abilities.

4. Early Spatial Sense

Spatial sense can be defined as the ability to 'grasp the external world' (Freudenthal, in National Council of Teachers of Mathematics [NCTM], 1989, p. 48). In our view, this spatial sense consists of three main components that are most essential for enabling young children to 'grasp the world' and to develop mathematical thinking: spatial visualization, geometry ('shapes' in short), and spatial orientation ('space' in short). These components can be recognized in the foundations of comprehensive mathematics curricula for the middle grades such as Mathematics in Context (1998).

Spatial visualization involves the ability to imagine the movements of objects and spatial forms. In spatial visualization tasks, all or part of a representation may be mentally moved or altered (Bishop, 1980; Clements, 2004; Tartre, 1990a). This has been conceptualized as the ability to make object-based transformations where only the positions of the objects are moved with respect to the environmental frame of reference whereas the frame of reference of the observer stays constant (Zacks, Mires, Tversky & Hazeltine, 2000).

An example of a daily activity in which, already, young children have to apply spatial visualization skills, is when they imagine where in the kitchen it is that they can find their snack before they walk into the kitchen to get it. Recent cognitive research on children's spatial skills has shown how 16-24 month old infants can use the concept of distance to localize objects in a sandbox (Huttenlocher, Newcombe, & Sandberg, 1994). This has suggested an early competence to judge distances that is manifested regardless of the presence of any references in the direct surroundings of the child. Such an ability requires spatial visualization skills for creating a mental picture of the location of the object.

Geometry lessons in school should teach young children about shapes and figures and help them learn to refer to familiar structures such as their own body, to geometrical structures such as mosaics, and to geometrical patterns such as dot configurations on dominoes (cf. Clements & Sarama, 2007). This type of communication may help increase their vocabulary and enrich their imagination (Casey, 2004; Newcombe & Huttenlocher, 2000). Hence, geometric activities can stimulate the children's ability to sharpen and talk about their perceptions, which in turn helps develop children's spatial sense and reasoning skills (Van den Heuvel-Panhuizen & Buys, 2005). Indeed NCTM (1989, p. 48) has described spatial understandings as necessary for interpreting, understanding, and appreciating our inherently geometric world.

The third component that we name in the context of how children may 'grasp the world' is spatial orientation. This is the term that Clements (2004, p. 284) uses to describe how we 'make our way' in space. As children discover their surroundings, they gain experiences that help them to understand the relative positioning and sizes of shapes and figures (Van den Heuvel-

Panhuizen & Buys, 2005). As such, children learn to orientate themselves, to take different perspectives, to describe routes and to understand shapes, figures, proportions and relationships between objects.

Many of the activities in spatial orientation are examples of competencies that are typically manifested even before these children begin their formal schooling. A cognitive study with four and five-year olds, for example, provided evidence that at this age children can already compare proportions and figures (Sophian, 2000). The children in this study were able to match the correctly shrunken picture to the original picture without being distracted by pictures that not only were smaller, but also disproportional to the original picture. Studies such as this one exemplify the remarkably developed spatial sense that many children possess prior to the start of formal schooling.

Now that we have illustrated what we mean by emerging number sense and early spatial sense, we turn to why and how in our research we suspect a relationship to exist between these two constructs.

5. Relating Early Spatial Sense to Emerging Number Sense: Spatial Structures

To analyze the development of number and spatial sense of young children, we must first take a step back and find inspiration in how young children learn and think in general. In the process of learning and understanding, young children continuously try to organize new concepts and information about the world (de Lange, 1987; Gopnik, 2004; van den Heuvel-Panhuizen, 2001). Structuring is one fundamental method for children to organize the world (Freudenthal, 1987). In effect, this method of organization contributes to gaining insight into important mathematical concepts such as patterning, algebra, and the recognition of basic shapes and figures (Mulligan, Mitchelmore, & Prescott, 2006; Waters, 2004). Freudenthal even believed that there is no other science in which organization plays such a crucial role as in mathematics (1991). He described mathematics as

> an activity of solving problems, of looking for problems, but it is also an activity of organizing subject matter. This can be matter from reality which has to be organized according to mathematical patterns if problems from reality have to be solved. (1971, p. 413-414)

As children develop through experience, they improve their ability to organize incoming information and they learn to amend their organization schemes accordingly. Piaget regarded knowledge as structures that become increasingly complex through the processes of accommodation, assimilation, and equilibration. When a child with a certain method of thinking experiences something that no longer fits with this method of thinking (cannot assimilate), then it is put off balance until the method of thinking is adjusted (accommodated) and the system is balanced again (equilibrated). In this way, children are believed to reach more sophisticated means of thinking.

Van den Heuvel-Panhuizen (2001) gives an example of a practical mathematical situation in which the learning process illustrated above can be recognized. In this example, four-year old Anita is trying to connect meaning and purpose to the numbers that she is hearing:

> Anita is in a pancake restaurant with her father. They have just chosen a pancake from the menu. "I want pancake twelve," says her father to the waitress. "And pancake seven for this young lady." Anita cries: "But I can't eat *that* many pancakes...!" (p. 29)

Experiences such as these can set young children's thinking off balance and force them to adjust their definitions and frames of reference. Children learn from this, adjust the structure of their mode of thinking and, in doing so, reach a higher level of understanding.

The type of structure discussed thus far is mostly conceptual in nature in the way that it contributes to learning and understanding. Much research has concentrated on such a type of structure in thinking (cf. Dienes, 1960; Sriraman, 2004; Van Hiele, 1997). The particular type of structure that our study is concerned with is analogous to this conceptual structure, and yet it is more concrete. It is structure that fits with children's experiences and current levels of spatial reasoning and it is structure which they may impose on manipulatives to support their mathematical learning and understanding.

To illustrate what we define as structure, we make use of the definition that Battista (1999) gave to describe the act of spatial structuring. In his view, spatial structuring is

> the mental operation of constructing an organization or form for an object or set of objects. It determines the object's nature, shape, or composition by identifying its spatial components, relating and combining these components, and establishing interrelationships between components and the new object. (p. 418)

A spatial structure, then, is a product of this act of organizing space. Such a structure is an important element of a pattern. In line with Papic and Mulligan (2005), we may define a spatial structure in terms of a pattern. A pattern is a numerical or spatial regularity and the relationship between the elements of a pattern, then, is its structure. In particular, we refer to a spatial structure as a configuration of objects in space. This relates to the component 'spatial regularity' in the given definition of a pattern. The component 'numerical regularity' refers to numerical sequences that are not relevant to the mathematical abilities of four- to six-year old children. Examples of spatial structures that children of this age are typically familiar with are dot configurations on dice, finger counting images, rows of five and ten, bead patterns, and block constructions (illustrated in Figure 1).

In reference to the three components of early spatial sense that we elaborated on earlier, we suggest that spatial structures may play a supportive role in the development of number sense. Specifically, the intertwinement of the three components may contribute to children's understanding of quantities and relationships between numbers. We propose that once children can imagine (i.e. spatially visualize) a spatial structure of a certain number of objects (i.e. configuration of objects that makes up a shape) that are to be manipulated (in a space), then learning to understand quantities as well as the process of counting (i.e. emerging number sense)

should greatly be simplified. This hypothesized relationship between early spatial sense and emerging number sense is depicted in the figure below.

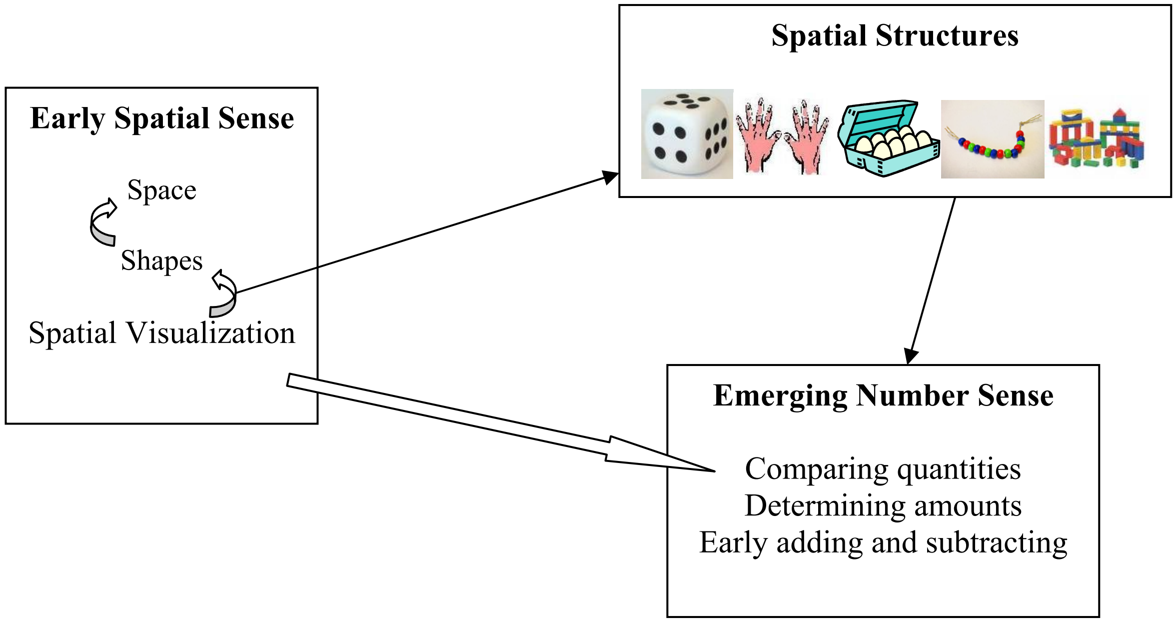


Figure 1. Spatial structures as a key factor in how early spatial sense may support the development of emerging number sense

After setting out why we suspect spatial structures to relate early spatial sense to emerging number sense, we continue our argument with illustrations of how spatial structures may play a supporting role in the development of mathematical abilities.

6. Spatial Structures in Early Numeracy Problems

To illustrate and support our concern with the role of spatial structures in the development of emerging number sense, we refer to Arcavi (2003) as one researcher who set out to define visualization and to analyze the various different roles that it may play in the learning and teaching of mathematics. Visualization, in his context, requires spatial visualization since it involves the interpretation and reflection upon pictures and images. Arcavi considers visualization to be at the service of problem solving because it may inspire the solution to a problem. In determining how many matches were needed to build an exemplar nxn square, for instance, most students used visual means to solve the problem. These visual means took different forms, one of which was the decomposition into what the students perceived to be easily countable units. This was a first step into changing the 'gestalt' (roughly the whole or the form) of the configuration.

It is the use of the term 'gestalt' in this context that supports our argument and indicates how students can simplify the mathematical problem by spatially visualizing objects into particular shapes in a space. For Arcavi's students, the 'gestalt' could involve 'breaking and rearranging the original whole' or 'imposing an "auxiliary construction" whose role consisted of providing

visual "crutches", which in themselves were not counted, but which supported and facilitated the visualization of a pattern that suggested a counting strategy' (Arcavi, 2003, p. 229).

Several studies have related the Gestalt laws to early development. Spelke and colleagues (1993), for example, found that while the perceptions of adults were strongly influenced by the Gestalt relations of color, texture similarity, good continuation, and good form, the perceptions of 5- and 9-month olds were only weakly affected, and the perceptions of 3-month olds were not at all affected. This suggests a developmental course of these particular Gestalt relations (cf. Quinn et al., 1993; 2002). Taken together, these studies highlight how even infants as young as three months are capable of distinguishing particular elements of and establishing crude perceptual coherence.

An anecdote of Richardson (2004) about the children in her preschool classroom illustrates how the extraction of spatial structures may occur in practice. Richardson had her children work with dot cards (showing configurations such as those on dice) so that they could learn to recognize amounts in such arrangements. When, one day, she asked the children to count out a certain number of counters, she was astonished to find that instead of correctly counting out the counters, the children made an 'X' shape to match what the children recognized to be the shape of five dots on a card, and they made a square shape to match what they recognized to be the arrangement of nine dots. Apparently, then, these children extracted a shape from the individual dots on cards and taught themselves that this shape should resemble a particular number.

Richardson (2004) concludes from this experience that teachers must always interact with the children to check whether what they are doing makes sense to them, because performing without understanding interferes with the development of their mathematical abilities. More than that, it is a practical example of how children extract a general shape from individual elements and it adds on to the finding that infants can deploy Gestalt principles to make sense of the real-world and to establish perceptual coherence. The ability to process the gestalt, the whole, is an important requirement for mathematical skill as it is one ability that should help simplify and shorten the children's process of learning to determine quantities (Van Eerde, 1996; Van Parreren, 1988). Such supporting evidence for children's tendencies to organize the world through the use of spatial structures, should encourage mathematics educators to take care to weave spatial abilities into early mathematics curricula.

Children typically begin to formalize their understanding of quantities by connecting a certain quantity with spatial structures such as a number of fingers that are being held up on a hand or dot configurations on a pair of dice. As Smith (1964, as cited in Tartre, 1990a) put it,

> the process of perceiving and assimilating a gestalt...[is] a process of abstraction (abstracting form or structure)... It is possible that any process of abstraction may involve in some degree the perception, retention in memory, recognition and perhaps reproduction of a pattern or structure" (p. 213-214).

These spatial structures require a child to use its spatial visualization skills for organizing and making sense out of visual information. The mental extraction of structures from spatial configurations is also what Arcavi (2003) found to aid the counting process of his students.

Although the students in Arcavi's study were older than the age group in our project, one can imagine how young children can also use 'gestalts' to rearrange objects that are to be counted, for example. The spatial structure that subsequently arises can help the child to oversee the quantity (Van Eerde, 1996; Van Parreren, 1988).

As illustrated in Figure 1, we propose that the spatial visualization abilities help the child to perceive the 'gestalt' or spatial structure, in order to either mentally or physically be able to rearrange the objects in a space. The spatial structure that subsequently arises can simplify early numerical procedures. When young children are asked to determine the quantity of a randomly arranged set of objects, they initially tend to count each object. As the set of objects grows, this procedure eventually confronts them with the difficulties of keeping track of which objects have already been counted and with the time-consuming process that accompanies the counting of larger sets.

The benefit of applying spatial structure to mathematical problems is evident, for instance, when reading off a quantity (i.e. seeing the quantity of six as being three and three), when comparing a number of objects (i.e. one dot in each of four corners is less than the same configuration with a dot in the center), when continuing a pattern (i.e. generalizing the structure) and when building a construction of blocks (i.e. relating the characteristics and orientations of the constituent shapes and figures). Here too, then, children's ability to grasp spatial structure appears essential for developing mathematical abilities such as ordering, comparing, generalizing and classifying (NCTM, 2000; Papic & Mulligan, 2005; Waters, 2004).

More formal mathematical skills require even further insight into and use of spatial structure. This is particularly the case for addition, multiplication and division (i.e. 8 + 6 = 14 because 5 + 5 = 10 and 3 + 1 = 4 so 10 + 4 = 14; Van Eerde, 1996), for using variables in algebra, for proving, predicting and generalizing, and for determining the structure of a shape in order to subsequently mentally rotate or manipulate it (Kieran, 2004). Various studies have shown that children with serious mathematical problems tend not to use any form of structure and continue to count objects one by one (Mulligan, Mitchelmore, & Prescott, 2005; Van Eerde, 1996). This accentuates the need for children to be familiar with various spatial structures in order to simplify the progression to more formal mathematical concepts and procedures.

7. Preliminary Experimental Support

Thus far, we have set out much of the theoretical support for why and how we propose that early spatial sense and emerging number sense may be related. Alongside this are some preliminary outcomes of a previously conducted explorative study (van Nes & de Lange, in press; van Nes & Doorman, 2006) in which we set out to investigate the strategies that four- to six-year old children use to solve various number sense and spatial thinking problems.

One outcome from the explorative study was that four- to six-year old children with relatively stronger mathematical skills seemed to make more use of spatial structures than other children did. These children recognized the spatial structures that were presented and knew to implement these spatial structures for simplifying and speeding up counting procedures. Interestingly, however, there were several low achieving five- and six-year old children who seemed to

recognize the spatial structures, and yet who did not proceed to applying the structures to solve the problems. These particular cases triggered our interest into what role insight into spatial structures may play in the development of emerging number sense and, ultimately, in the child's level of mathematical achievement.

The findings from our explorative study complement research of Mulligan, Prescott and Mitchelmore (2004) in which they conducted an analysis of structure present in 103 first graders' representations for various tasks across a range of mathematical domains. They coded the individual profiles as one of four stages of structural development and found that mathematical structure in children's representations generalizes across various mathematical domains. Recently, Mulligan, Mitchelmore and Prescott (2005; 2006) developed a Pattern and Structure Assessment (PASA) interview and a Pattern and Structure Mathematics Awareness Program (PASMAP) to study whether the mathematics of low achieving students can be improved through explicit instruction about structures and patterns in mathematical domains. The preliminary results showed improved mathematical achievement, suggesting that explicit instruction of mathematical pattern and structure can stimulate student's learning and understanding of mathematical concepts and procedures.

Taking the theoretical background and the preliminary findings together, we summarize the research questions of the present study from the perspective of mathematics education as:

1. How are early spatial sense and emerging number sense related and what role may spatial structures play in this development?
2. How can spatial visualization be implemented in educational practices to support the development of number sense?

In order to answer these two research questions we concentrate on designing a teaching experiment in which we may study how the development of spatial sense and number sense may be stimulated in an educational setting. This last issue will be investigated in terms of a design research methodology.

8. An Instruction Experiment

In gaining an understanding of how children recognize and apply spatial structures to numerical problems, it is important to decide on a methodology that is appropriate for highlighting the processes that occur in the mind of the child from the perspective of the child. The methodology that appears to be most in line with the principles of TalentPower, is inspired by the main theoretical insights of researchers in mathematics education such as Freudenthal (1984, 1991), Dienes (1960) and Van Parreren (1988). This generally concerns a methodology that is focused on a child's learning processes, that applauds dialogue and interaction, that emphasizes the stimulation of the own actions of the child, and that rejects mechanistic mathematics education (Van Eerde, 1996).

The activities for the instruction experiment stem from the tasks that we developed, tried out and improved in the previous exploratory studies (van Nes & de Lange, in press; van Nes & Doorman, 2006). Next to being based on the abovementioned theoretical insights, these tasks were originally inspired by experimental outcomes and practical experiences as described in related literature (van den Heuvel-Panhuizen, 2001, for example) and developed with input from experts. We also assessed the appropriateness of the tasks in terms of their coherence with the outcomes of the Utrecht Numeracy Test (UNT, van Luit et al., 1994). This is a normed test for assessing the number sense of 4.5- to 7-year old children. We compared the children's scores on this test with their accuracy scores as well as with the level and types of strategies that they used on the tasks. As we were easily able to come to a consensus about the scoring of the tasks, the strategy classifications and their agreement with the UNT scores, we decided that the tasks would be suitable to work out into a series of activities for use in the instruction experiment.

As the methodology is based on the guidelines of 'design research' (Freudenthal, 1978; Gravemeijer, 1994, 2004; Gravemeijer, Bowers, & Stephan, 2003; Streefland, 1988), our theory will cohere with direct experiences from an educational setting. This should keep the findings both theoretical and practical. It will involve an iterative procedure of theory-driven adjustments to the intervention and amendments to the hypotheses that lead to an improved and evidence-based theory (Freudenthal, 1978; Gravemeijer, 1994; Streefland, 1988). Freudenthal (1991) referred to such a research design as an instruction experiment because the activities are meant to broaden the children's insight into spatial visualization, into the perception and application of spatial structures, and, ultimately, into the characteristics of quantities and numbers while, at the same time, providing the researchers with a greater understanding of the children's learning processes. The aim, then, is not necessarily to conclude *that* the series of activities teach the children about spatial structures, but more to come to an analysis about *why* the series of activities may have stimulated the children's thinking (Gravemeijer et al., 2003).

In order to study the children's thinking processes, the series of activities should guide the children along a so-termed conjectured local instruction theory (Gravemeijer, 1994; Simon, 1995). The conjectured local instruction theory is a learning trajectory based on mathematical, psychological, and didactical insights about how we expect that the children will progress from their original way of thinking to our aspired way of thinking. To ensure the practicality of our findings, we must take into account both the cognitive development of the individual students, as well as the social context (i.e. people, setting and type of instruction) in which the instruction experiment is to take place (Cobb & Yackel, 1996).

The cyclical process that characterizes design research is illustrated in the diagram below. In practice this means that we will implement the series of activities in an instruction experiment, perform retrospective analyses on the transcripts from these lessons, adjust our hypotheses accordingly in a thought experiment and improve the activities in line with the amended conjectured local instruction theory. Then we repeat the procedure by implementing the new set of activities in a subsequent cycle, and learning from the class-experiences for, once again, fuelling the next thought experiment. This process will contribute to establishing and refining our conjecture local instruction theory.

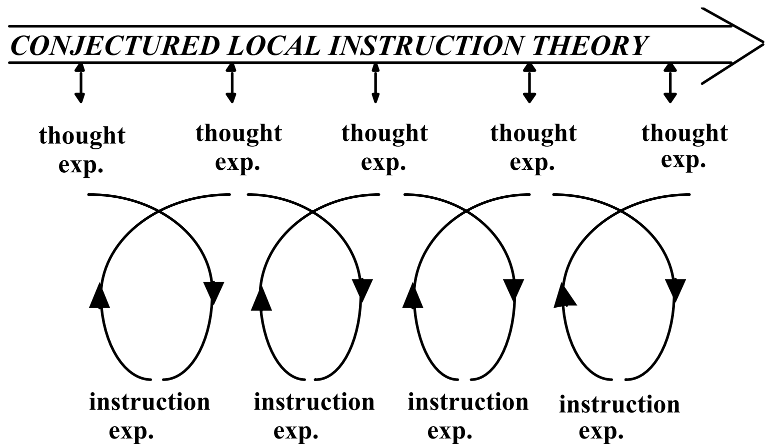


Figure 2. The cyclical procedure of design research (Gravemeijer, 2004)

9. Summary and Conclusion

After providing a broad overview of the theoretical framework that is propelling the MENS research, it is clear that young children possess spatial and numerical skills that should be cultivated in educational practice. As such, the aim of this research is to bring the spatial sense of young children to the fore and illustrate how spatial skills could function to stimulate the development of more formal mathematical skills that require number sense.

Supported by various fields of research, we consider spatial visualization, insight into shapes and an understanding of space to be three main components that make up young children's early spatial sense. As such, we suggest that children's spatial visualization skills contribute to their ability to organize representations of objects into spatial structures (such as dice configurations and finger images). These spatial structures relate to the children's conceptions of shapes with which they become familiar through exploring their surrounding space. Children's concepts of quantities and number, then, may greatly be stimulated when children are made aware of the simplifying effects of structuring manipulatives.

As soon as we have cycled through enough instruction and thought experiments to fundament our conjectured local instruction theory, we will turn to our colleagues for comparing and contrasting the results of the research perspectives of mathematics education and educational neurosciences. The neuroscientific perspectives may supplement our research with results from studies on brain behavior and neural correlates with respect to early spatial and numerical thinking. Ultimately, in line with the principles of *TalentPower*, the collaboration of these research perspectives should provide a more all-round and in-depth understanding of how education can foster the talents of young children and possibly stimulate those children who may be prone to experiencing problems in the development of mathematical skills.

As Tartre (1990a) stated in a discussion on spatial orientation,

> attempting to understand and discuss something like spatial orientation skill, which is by definition intuitive and nonverbal, is like trying to grab smoke: the very act of reaching out to take hold of it disperses it (p. 228).

She notes that any attempt to verbalize spatial thinking no longer is spatial thinking since spatial thinking is only a mental activity. We recognize that research into spatial sense is always an indirect attempt at trying to understand what is happening in the mind. Nevertheless, by taking into account the three components that we associate with spatial sense, and by relating them to each other in the way that we are, we aim to gain an understanding of how young children's early spatial skills may help them progress in their mathematical development. This is how we intend to better appreciate and more effectively cultivate young children's cognitive capacities that too often are underestimated or even neglected.

Acknowledgements
The MENS project is supported by the Research Council for Earth and Life Sciences (ALW) with financial aid from the Netherlands Organization for Scientific Research (NWO), project number 051.04.050.

References

Arcavi, A. (2003). The role of visual representations in the learning of mathematics. *Educational Studies in Mathematics, 52,* 215-241.

Battista, M.T. (1999). Fifth graders' enumeration of cubes in 3D arrays: conceptual progress in an inquiry-based classroom. *Journal for Research in Mathematics Education, 30*, 417-449.

Baltussen, M., Klep, J., & Leenders, Y. (1997). *Wiskunde-avonturen met Jonge Kinderen* [Mathematics Adventures with Young Children]. Amersfoort: CPS.

Baroody, A.J. (2004). The developmental bases for early childhood number and operations standards. In D.H. Clements & J. Sarama (Eds.), *Engaging Young Children in Mathematics Standards for Early Childhood Mathematics* (pp. 173-219). Mahwah: Lawrence Erlbaum Associates, Inc.

Benthem, J. van, Dijkgraaf, R., & de Lange, J. (1995). *TalentPower. A Summary.* Retrieved March 9, 2007, from http://www.TalentenKracht.nl.

Berch, D.B. (2005). Making sense of number sense: implications for children with mathematical disabilities. *Journal of Learning Disabilities,* 38 (4), 333-339.

Berger, A., Tzur, G., & Posner, M.I. (2006). Infant brains detect arithmetic errors. *Proceedings of the National Academy of Sciences, 103 (33),* 12649-12653.

Berninger, V.W., & Corina, D. (1998). Making cognitive neuroscience educationally relevant: creating bidirectional collaborations between educational psychology and cognitive neuroscience. *Educational Psychology Review, 10 (3),* 343-352.

Bishop, A.J. (1980). Spatial abilities and mathematics achievement—a review. *Educational Studies in Mathematics, 11,* 257-269.

Bishop, A.J. (1983). Space and geometry. In R. Lesh and M. Landau (Eds.), *Acquisition of Mathematics Concepts and Processes* (pp. 175-203). New York: Academic Press.

Byrnes, J.P., & Fox, N.A. (1998). The educational relevance of research in cognitive neuroscience. *Educational Psychology Review, 10 (3),* 297-342.

Casey, B. (2004). Mathematics problem-solving adventures: A language-arts-based supplementary series for early childhood that focuses on spatial sense. In D.H. Clements & J. Sarama (Eds.), *Engaging Young Children in Mathematics Standards for Early Childhood Mathematics* (pp. 377-389). Mahwah: Lawrence Erlbaum Associates, Inc.

Clements, D.H. (2004). Major themes and recommendations. In D.H. Clements & J. Sarama (Eds.), *Engaging Young Children in Mathematics Standards for Early Childhood Mathematics* (pp. 7-75). Mahwah: Lawrence Erlbaum Associates, Inc.

Clements, D.H. & Battista, M.T. (1992). Geometry and spatial reasoning. In D.A. Grouws (Ed.), *Handbook of Research on Mathematics Teaching and Learning* (pp. 420-464). New York: Macmillan.

Clements, D.H., & Sarama, J. (2007). Early childhood mathematics learning. In F. Lester (Ed.), *Handbook of Research on Teaching and Learning Mathematics* (2nd ed.). Greenwich, CT: Information Age Publishing.

Cobb, P. (2007). Putting philosophy to work: coping with multiple theoretical perspectives. In F. Lester (Ed.), *Handbook of Research on Teaching and Learning Mathematics* (2nd ed.). Greenwich, CT: Information Age Publishing.

Cobb, P., & Yackel, E. (1996). Constructivist, emergent and sociocultural perspectives in the context of developmental research. *Educational Psychologist, 31,* 175-190.

Davis, A. (2004). The credentials of brain-based learning. *Journal of Philosophy of Education, 38 (1),* 21-35.

De Lange, J. (1987). *Mathematics, Insight, and Meaning*. Utrecht: OW and OC.

Dienes, Z. (1960). *Building Up Mathematics* (4th ed.). London: Hutchinson Educational Ltd.

Dijkgraaf, R. (2007). *Nieuwsgierigheid en Verwondering* [Curiosity and Astonishment]. In TalentenKracht Program. The Hague: VTB.

Freudenthal, H. (1978). *Weeding and Sowing. Preface to a Science of Mathematical Education*. Dordrecht: D. Reidel Publishing Company.

Freudenthal, H. (1984). *Appels en Peren/Wiskunde en Psychologie: Gebundelde Opstellen. [Apples and Pears/Mathematics and Psychology: Collected Essays]*. Apeldoorn: van Walveren B.V.

Freudenthal, H. (1991). *Revisiting Mathematics Education: China Lectures.* Dordrecht: Kluwer Academic Publishers.

Fuson, K.C. (2004). Pre-K to grade 2 goals and standards: Achieving 21st century mastery for all. In D.H. Clements & J. Sarama (Eds.), *Engaging Young Children in Mathematics Standards for Early Childhood Mathematics* (pp. 105-148). Mahwah: Lawrence Erlbaum Associates, Inc.

Gelman, R., & Gallistel, C.R. (1978). *The Child's Understanding of Number*. London: Harvard University Press.

Gersten, R., & Chard, D. (1999). Number sense: rethinking arithmetic instruction for students with mathematical disabilities. *The Journal of Special Education, 44,* 18-28.

Gopnik, A. (2004). Finding our inner scientist. *Daedalus, 133 (1),* 21-28.

Gopnik, A., Meltzoff, A., & Kuhl, P. (1999). *The Scientist in the Crib: Minds, Brains, and How Children Learn.* Farfield, NJ: William Morrow and Company.

Gravemeijer, K. (1994). *Developing Realistic Mathematics Education.* Utrecht: CD-β Press.

Gravemeijer, K. (2004). Local instruction theories as means of support for teachers in reform mathematics education. *Mathematical Thinking and Learning, 6 (2),* 105-128.

Gravemeijer, K., Bowers, J., & Stephan, M. (2003). A hypothetical learning trajectory on measurement and flexible arithmetic. In M. Stephan, J. Bowers, P. Cobb, & K. Gravemeijer (Eds.). *Supporting Students' Development of Measuring Conceptions: Analyzing Students' Learning in Social Context* (pp. 51-66). Reston: NCTM.

Griffin, S. (2004a). Teaching number sense. *Educational Leadership, 61 (5),* 39-42.

Griffin, S. (2004b). Contributions of central conceptual structure theory to education. In A. Demetriou & A. Raftopoulos (Eds.). *Cognitive Developmental Change. Theories, Models and Measurement* (pp.264-295). Cambridge University Press.

Griffin, S., & Case, R. (1997). Re-thinking the primary school math curriculum: An approach based on cognitive science. *Issues in Education, 3 (1),* 1-49.

Guay, R.B., & McDaniel, E.D. (1977). The relationship between mathematics achievement and spatial abilities among elementary school children. *Journal for Research in Mathematics Education, 8 (3),* 211-215.

Hughes, M. (1987). *Children and Number. Difficulties in Learning Mathematics.* Oxford: Blackwell Publishers Ltd.

Huttenlocher, J., Newcombe, N., & Sandberg, E. (1994). The coding of spatial location in young children. *Cognitive Psychology, 27,* 115-147.

Jolles, J., de Groot, R., van Benthem, J., Dekkers, H., de Glopper, C., Uijlings, H., et al. (2006). *Brain Lessons*. Maastricht: Neuropsych Publishers.

Kieran, C. (2004, July). *The development of algebraic thinking and symbolization.* Paper presented to the PME Research Session, 10th International Congress on Mathematical Education. Copenhagen: Danish Technical University.

Lester, F. K. (2005). On the theoretical, conceptual and philosophical foundations for research in mathematics education. *Zentralblatt fur Didaktik der Mathematik, 37 (6),* 457-467.

Lipton, J., & Spelke, E.S. (2003). Origins of number sense: large-number discrimination in human infants. *Psychological Science, 14,* 396-401.

Mathematics in Context (1998). Austin: Encyclopaedia Britannica Educational Corporation.

Mulligan, J.T., Mitchelmore, M.C., & Prescott, A. (2005). Case studies of children's development of structure in early mathematics: a two-year longitudinal study. In Chick, H. & Vincent, J. (Eds.). *Proceedings of the 29th annual conference of the International Group for the Psychology of Mathematics Education: Vol. 4.* (pp. 1-8), Melbourne, Australia: PME.

Mulligan, J.T., Mitchelmore, M.C. & Prescott, A. (2006). Integrating concepts and processes in early mathematics: the Australian Pattern and Structure Mathematics Awareness Project (PASMAP). In J Novotná, H. Moraová, M. Krátká, & N. Stehlíková (Eds.). *Proceedings of the 30th annual conference of the International Group for the Psychology of Mathematics Education: Vol.4.* (pp. 209-216). Prague: PME.

Mulligan, J.T., Prescott, A., & Mitchelmore, M.C. (2004). Children's development of structure in early mathematics. In M. Høines & A. Fuglestad (Eds.). *Proceedings of the 28th annual conference of the International Group for the Psychology of Mathematics Education: Vol. 3* (pp. 393-401). Bergen, Norway: Bergen University College.

Murphy, C. (2006). Embodiment and reasoning in children's invented calculation strategies. In J. Novotná, H. Moraová, M. Krátká, M. & N. Stehlíková (Eds.). *Proceedings 30th Conference of the International Group for the Psychology of Mathematics Education: Vol. 4.* (pp. 217-224). Prague: PME.

National Council of Teachers of Mathematics (1989). *Curriculum and Evaluation Standards for School Mathematics.* Reston, VA: NCTM.

National Council of Teachers of Mathematics (2000). *Curriculum and Evaluation Standards for School Mathematics.* Reston, VA: NCTM.

National Research Council (2005). *Mathematical and Scientific Development in Early Childhood.* Washington D.C.: National Academic Press.

Organization for Economic Co-operation and Development (2002). *Understanding the Brain. Towards a New Learning Science.* Paris: OECD.

Papic, M. & Mulligan, J. (2005). Preschoolers' mathematical patterning. *The Proceedings of the 28th Mathematical Education Research Group of Australasia Conference* (pp. 609-616). Melbourne, Australia.

Quinn, P.C., Burke, S., & Rush, A. (1993). Part-whole perception in early infancy: evidence for perceptual grouping produced by lightness similarity. *Infant Behaviour & Development, 16,* 19-42.

Quinn, P.C., Bhatt, R.S., Brush, D., Grimes, A., & Sharpnack, H. (2002). Development of form similarity as a gestalt grouping principle in infancy. *Psychological Science, 13 (4),* 320-328.

Richardson, K. (2004). Making sense. In D.H. Clements & J. Sarama (Eds.). *Engaging Young Children in Mathematics: Standards for Early Childhood Mathematics* (pp. 321-324). Mahwah: Lawrence Erlbaum Associates Inc.

Ros, B. (2006). De sprankelcoëfficiënt van jonge kinderen [The sparkle coefficient of young children]. *Didactief, 4,* 8-9.

Siegler, R.S. (2003). Implications of cognitive science research for mathematics education. In J. Kilpatrick, W.B. Martin, & D.E. Schifter (Eds.). *A Research Companion to Principles and Standards for School Mathematics* (pp. 219-233). Reston, VA: National Council of Teachers of Mathematics.

Simon, M. (1995). Reconstructing mathematics pedagogy from a constructivist perspective. *Journal for Research in Mathematics Education, 26 (2),* 114-145.

Sophian, C. (2000) Perceptions of proportionality in young children: matching spatial ratios. *Cognition, 75,* 145-170.

Spelke, E.S. (2002). Developmental neuroimaging: a developmental psychologist looks ahead. *Developmental Science, 5 (3),* 392-396.

Spelke, E.S., Breinlinger, K., Jacobson, K., & Phillips, A. (1993). Gestalt relations and object perception: a developmental study. *Perception, 22,* 1483-1501.

Sriraman, B. (2004). Reflective abstraction, uniframes and the formulation of generalizations. *Journal of Mathematical Behaviour, 23,* 205-222.

Starkey, P. (1992). The early development of numerical reasoning. *Cognition, 43,* 83-126.

Steffe, L.P. (2004). PSSM From a constructivist perspective. In D.H. Clements & J. Sarama (Eds.). *Engaging Young Children in Mathematics Standards for Early Childhood Mathematics* (pp. 221-251). Mahwah: Lawrence Erlbaum Associates, Inc.

Streefland, L. (1988). *Realistisch Breukenonderwijs. Onderzoek en Ontwikkeling van een Nieuwe Leergang* [Realistic Instruction of Fractions. Research and Development of a New Program]. Utrecht: Vakgroep Onderzoek Wiskundeonderwijs & Onderwijscomputercentrum.

Tartre, L.A. (1990a). Spatial orientation skill and mathematical problem solving. *Journal for Research in Mathematics Education, 21 (3),* 216-229.

Tartre, L.A. (1990b). Spatial skills, gender, and mathematics. In E. Fennema & G.C. Leder (Eds.). *Mathematics and Gender* (pp. 27-59). New York: Teachers College Press.

Van den Heuvel-Panhuizen, M. (2001). *Children Learn Mathematics. A Learning-Teaching Trajectory with Intermediate Attainment Targets for Calculation with Whole Numbers in Primary School.* Utrecht: Freudenthal Institute, Utrecht University.

Van den Heuvel-Panhuizen, M., & Buys, K. (2005). *Young Children Learn Measurement and Geometry. A Learning-Teaching Trajectory with Intermediate Attainment Targets for the Lower Grades in Primary School.* Utrecht: Freudenthal Institute, Utrecht University.

Van Eerde, H.A.A. (1996). *Kwantiwijzer: Diagnostiek in Reken-wiskundeonderwijs* [Kwantiwijzer: Diagnostics in Mathematics Education]. Uitgeverij Zwijsen bv, Tilburg.

Van Hiele, P.M. (1997). *Structuur* [Structure]. Uitgeverij Thieme, Zutphen.

Luit, J.E.H. van, Rijt, B.A.M. van de, & Pennings, A.H. (1994). *De Utrechtse Getalbegrip Toets* [The Utrecht Numeracy Test]. Doetinchem: Graviant.

Van Nes, F.T. & de Lange, J. (in press). The role of spatial configurations in early numeracy problems. *Proceedings for the Fifth Congress of European Researchers for Mathematics Education. Working group 3.* Larnaca, Cyprus.

Van Nes, F.T. & Gebuis, T. (2006). Mathematics Education and Neurosciences (MENS). In J. Novotná, H. Moraová, M. Krátká, & N. Stehlíková (Eds.). *Proceedings 30th Conference of the International Group for the Psychology of Mathematics Education: Vol. 1.* (p. 352). Prague: PME.

Van Parreren, C.F. (1988). *Ontwikkelend Onderwijs*. Apeldoorn/Leuven: Acco.

Waters, J. (2004). Mathematical patterning in early childhood settings. *The Proceedings of the 27th Mathematical Education Research Group of Australasia Conference* (pp. 565-572). Townsville, Australia.

Xu, F. & Spelke, E.S. (2000). Large number discrimination in 6-month-old infants. *Cognition, 74,* B1-B11.

Zacks, J.M., Mires, J., Tversky, B., & Hazeltine, E. (2002). Mental spatial transformations of objects and perspective. *Spatial Cognition and Computation, 2,* 315-332.

Introduction of a new construct: The conceptual tool "Flexibility"

Mette Andresen[1]
Danish University of Education

Abstract

This paper presents a new construct: the conceptual tool 'Flexibility'. The construct was a result of an attempt to extract experiences of teaching and learning with the use of laptops. It was further developed and refined on the basis of four small-scale teaching experiments. The teaching experiments, being part of a development project in upper secondary school mathematics, investigated the use of laptops for teaching differential equations from a modeling point of view. The research was double-aimed: one objective was to conclude the project with some recommendations for the design of teaching, in the form of guidelines suitable for a wider dissemination amongst upper secondary mathematics teachers. The other main aim was to draw on the project's experiences for theory development within a framework based on Realistic Mathematics Education (RME) and related ideas. The construction of 'Flexibility' served both these aims.

Keywords: computer environments; flexibility; instrumental genesis; technology; teaching and learning; laptops; differential equations; Realistic mathematics education (RME)

1. Background

In Denmark, a recent reform of the structure and the curriculum in upper secondary school encompassed introducing the use of CAS in mathematics, chemistry and physics. Teaching with CAS was not required to follow authorized plans or materials: the design of teaching sequences, planning and preparation of extra teaching materials etc. were still the individual teacher's responsibilities. Thereby, the reform made new demands on the teachers' professional development: it was far from obvious how the use of CAS should be integrated in the individual teacher's repertoire of teaching instruments[2]. A metropolitan area six-year school development project, titled 'World Class Math and Science'[3], partly served as a precursor for the reform,

---

[1] Cand. Scient. Ph.d. Mette Andresen
Assistant professor of didactics of mathematics
Institute for curriculum research,
Danish University of Education
Tuborgvej 164
2400 København NV/
Tel : (+45) 8888 9688
e-mail mea@dpu.dk

Home page: http://www.dpu.dk/om/mea

[2] For a discussion of the generation of CAS use as the teacher's teaching instrument, see (Andresen 2006 p 265-275)
[3] www.matnatverdensklasse.dk

although the government did not sponsor the project. Part A of the development project[4] encompassed experiments with laptop classes in upper secondary school mathematics.

I was employed at the Danish University of Education to do the research in part A of the development project for my Ph.D. in mathematics education. Following the standpoint described in Wittmann (1998) and Lesh & Sriraman (2005), that mathematics education is a design science, I believe that research and development has to be linked to teaching and learning practice at crucial points. Improvement of teaching practice must be merged with the progress of the whole field of educational research and reversely, research progress linked to the development of teaching practice. So, the research for the Ph.D. project became double-aimed: one objective was to conclude the World Class project with some recommendations for the design of teaching, in the form of guidelines suitable for a wider dissemination amongst upper secondary mathematics teachers. The other main aim was to draw on the project's experiences for theory development within the theoretical framework based on Realistic Mathematics Education (RME) and related ideas. My construction of 'Flexibility' served both these aims.

1.1 Objectives of the construction

With this background, the initial goal of the research was to extract knowledge from the project's teaching experiments in a way that made this knowledge useful. More precisely, the aim was to identify, articulate and conceptualise the participants' shared experiences of improved learning and, subsequently, to turn these experiences into a form that would:

- Allow teachers to take them into account for improvement of their own teaching
- Serve as a contribution to math education theory

In yearly evaluation interviews, participants from the World Class A project repeatedly had expressed a shared experience of improved learning: in the interviews, a number of students and teachers described their feelings of getting "a better overview" and "a deeper understanding", apparently as a result of the use of laptops. These remarks were developed further with explanations like "because you can easily get series of graphs", "you do not get stuck in technical details", "it is easy to see examples" or "in stead of remembering a lot of techniques you are allowed to concentrate on the ideas". Such answers are in accordance with some expectations to CAS (and computer) use, widespread amongst math teachers. Teachers, like many researchers[5], believe that besides to favour visualisations, the CAS routines can be incorporated into the design of teaching as shortcuts, to facilitate students' focusing on ideas, structures and conceptions.

Hence, in the research project the challenge was to critically sort out examples of fruitful support of students' learning from the general feeling of 'flow in the classroom'. So the objective turned out to be pointing out important elements of learning activities, for teachers to aim at in their future design and preparation of teaching sequences and single lessons. These elements of learning activities are important, according to two criteria:

1. The activity should promote the student's actual work with mathematics in an observable way

[4] See (Andresen 2006 chapter 3, p 21-39) for a detailed description
[5] See for example (Drijvers, 2003, p 92). Hypothetically, the work with CAS may offer the students a shortcut to reification or a shortcut to working with "objects" as if the processes were reified.

2. Arguments based on the theoretical framework should support the claim that the activity could promote the student's learning

The mental actions 'change of perspective' and 'change of representation' turned out to be a common denominator for the important elements in focus of my interest. They gave inspiration to my choice of the term 'flexibility' to denote the new construct in the following definition.

1.2 Definition: Flexibility of mathematical conceptions

Definition

> *The flexibility of a mathematical conception constructed by a person is the designation of all the changes of perspective and all the changes between different representations the person can manage within this conception.* (Andresen, 2006, p. 136):

In this definition of the conceptual tool flexibility, the term 'change of perspective' means change between different facets of the mathematical conception in question, regarded as the student's construct[6]. My selection of a number of complementary pairs of perspectives intended to make flexibility an operational tool. Two main considerations determined my choice of pairs, in accordance with (1.) and (2.) above:

- Changes within the pairs of perspective should be recognisable for the teacher (or any observer)
- The changes should be pivots for the students' learning process.

Within the notion of flexibility, the term 'representation' is used in the sense of the media of expression or communication. The objective of the construction was not to categorise all mental changes. The overall aim was to offer teachers a few building blocks in the form of simple design heuristics. My design of these building blocks should be appropriately based on the research to ensure that the use of them was likely to support the students' learning.

Neither did the construction aim to establish any one-to-one correspondence between every mathematical conception and its exposure within each perspective. Supposedly, it is clear that any mathematical conception can be expressed in each perspective in more than one way, and it may be exposed in more than one way in each representation. The perspectives and representations, referred to in the definition, are presented in a later paragraph in this paper. Before that, the next paragraph tells about the theoretical basis for the concept of flexibility.

[6] According to L. P. Steffe and P. W. Thompson (Steffe & Thompson 2000 p 268-269), the experience of students' learning allows the researcher to inquire the students' mathematical realities. These realities are called *students' mathematics* and by (partly) knowing them, it is possible for the researchers to construct a model of students' mathematics, called *mathematics of the students*. Students' mathematics, which the students have constructed as a result of their interactions in their physical and sociocultural milieu, is indicated by what the students say and do when they engage in mathematical activity. In contrast, mathematics of the students is part of the shared knowledge in the classroom, compatible with the educational goals.

2. Basic ideas beyond the concept of 'flexibility'

Literature studies played an important role for the development and refinement of my earliest idea of 'flexibility'. This paper only presents three main ideas from the framework that forms the basis for my definition of 'flexibility'[7]: i) a specific dynamical approach to concept formation that combines main heuristics of Seymour Papert and Jean Piaget, introduced by Edith Ackermann, ii) vertical and horizontal mathematising in the RME sense realised in Koeno Gravemeijer's four level model and iii) the French theory of Instrumental Genesis.

2.1. A dynamical approach to concept formation

To facilitate cognitive growth, Edith Ackermann presents the idea of a bi-directional interplay between "diving in" and "stepping back" in (Ackermann 1990 p 6). Ackermann refers to both Jean Piaget and Seymour Papert[8] as constructivists who see children as the builders of their own cognitive tools, as well as builders of their external realities. Both Piaget and Papert consider knowledge as a personal experience to be constructed. Further, they both acknowledge adaptation as

> *.. the ability to maintain a balance between stability and change, closure and openness, continuity and diversity, or, in Piaget's words, between assimilation and accommodation. (...) The main difference is that Piaget's interest was mainly in the construction of internal stability whereas Papert is more interested in the dynamics of change. (Ackermann, 1990, p.4)*

The main point in Ackermann's description of Papert's view is: *diving into* situations rather than looking at them from a distance, *connectedness* rather than separation are powerful means of gaining understanding: becoming one with the phenomenon under study is a key to learning.
Ackermann's description of Piaget's view can be summarised like this: the way children progressively become *detached* from the world of concrete objects and local contingencies is closely related to their gradually becoming able to mentally manipulate symbolic objects within the realm of hypothetical worlds, so that rules and invariants are means of interpreting and organizing the world, and *abstract and formal thinking* are the most powerful way to handle complex environments.

Ackermann states that her own perspective is

> *..an integration of the above views. Along with Piaget I view separateness through progressive decentration as a necessary step toward reaching a deeper understanding. I see constructing invariants as the flipside of generating variation. (..) I share Papert's idea that diving into unknown situations, at the cost of experiencing a momentary sense of loss, is a crucial part of learning. (..) My claim is, that both "diving in" and "stepping back" are equally important in getting such a cognitive dance going. (Ackermann 1990 p 6)*

---

[7] The complete theoretical framework is presented and discussed in (Andresen 2006, chapter 5)

[8] In a footnote, Ackermann says: "*Describing the difference between Piaget and Papert has been useful for me, and might be of general interest for the reasons mentioned in the text. It is through working directly with both thinkers (first, at the Piaget Institute, and currently at MIT) that I became progressively convinced of the need for integrating structural and differential approaches in describing human development.*" (Ackermann 1990 p 27)

Ackermann's idea of a 'cognitive dance' gave me the inspiration to construct 'flexibility' as a means to capture the dynamics of concept formation rather than static descriptions of cognitive structures. My method was to identify and group changes of perspective and changes between representations.

2.2. Vertical and horizontal mathematising in RME

Within the paradigm of RME, guided reinvention (or progressive mathematising), didactical phenomenology and emergent models are the three key heuristics for the design of teaching. In contrast to the cognitive theories of concept formation, 'modelling' was acknowledged to be an issue of interest in the World Class project's teaching culture. Moreover, a 'guided reinvention' design is to a great extend in accordance with the prevailing norms for good teaching. So it seemed reasonable to include RME into the framework in the following way. The four-level-model of activity intends to capture the way students' thinking evolves (Gravemeijer & Stephan 2002 pp 159-160):

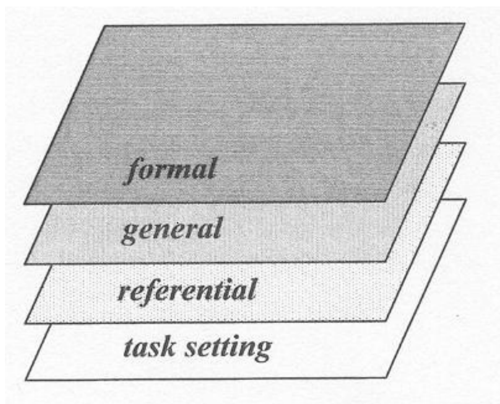


Fig 1. Levels of activity (Gravemeijer and Stephan, 2002, p. 159)

In the model, the activities at the level of task setting take place in a situation that is experientially real for the students. This enables the students to reason with a model of the problem but still think about acting in the situation. For experienced students an experientially real situation may also be mathematical. At the next, referential, level the students act with a model that is meaningful because it signifies an experientially real activity for them. At the general level the type of activity has changed since the students' attention has shifted from the contextual meaning to the mathematical relations involved. The students' activities are no longer dependent on situation-specific imagery. Finally, the students will no longer need the support of a model for their more formal mathematical reasoning, which constitutes the fourth level of activity.

So, progressive mathematising implies the students' raising up through the four levels. The level-raising does not happen as a total, one-directional movement: the student may in one context of content be at one level and in another context at another level. Besides, the raising from one level to another happens over time, where the student switches between two levels in both directions several times. The levels, further, have different character. It follows that in terms of RME,

concept formation by modelling can be expressed as changes between levels: horisontal mathematising happens as changes from *situational, task setting* level to *referential* level when models emerge, and vertical mathematising happens as changes from the level of *model of* (referential) – to the level of *model for* (general) and to further up to the *formal* level. In the description below, the first three levels in the four-level-model are considered as perspectives on the mathematical conceptions in question. Level-raising is considered as the result of bidirectional changes between these perspectives, which are included in 'flexibility'.

2.3. The French theory of Instrumental Genesis

The French theory of instrumental genesis is based on the idea that an artefact, for example a CAS calculator, does not in itself serve as a tool. It becomes useful, and then denoted an 'instrument', only after the user's formation of (one or more) mental utilisation scheme(s). Such utilisation schemes connect the artefact with conceptual knowledge and understanding of the way it may be used to solve a given task. The utilisation schemes contribute to the formation of instrumented action schemes. So, an instrument consists of the tool, for example a laptop with the CAS software Derive, the student's mental utilisation schemes and the task or problem to be solved. (Drijvers, 2003, pp. 96-97).

The process in which the artefact becomes an instrument is called 'instrumental genesis'. The process proceeds through activities in

> *The two-sided relationship between tool and learner as a process in which the tool in a manner of speaking shapes the thinking of the learner, but also is shaped by his thinking.* (Drijvers & Gravemeijer, 2005, p. 190).

The two directions of the process can be linked to the construction of epistemic and pragmatic knowledge, respectively. The distinction between construction of epistemic and pragmatic knowledge is reflected in the definition of flexibility, which discerns between pairs of perspectives with relation to the construction of epistemic knowledge and, partly corresponding, pairs with relation to pragmatic knowledge.

This outline of the theory of instrumental genesis reveals the underlying French framework: the scheme concept, encompassing utilisation schemes and instrumented action schemes, was introduced by Vergnaud (Trouche, 2005, p.149). Since the mental utilisation schemes are not directly accessible for study and analysis, the concept of 'instrumented techniques' is of special interest: instrumented techniques are the external, visible and manifest parts of the instrumented action scheme. Still, an instrumented technique involves conceptual elements, since the technique reflects the schemes. This leads me to the following two crucial conclusions:

- The study of a student's development and use of instrumented techniques is useful to enlighten the student's development of those instrumented action schemes, to which the techniques relate
- Development of mathematical conceptions cannot be studied if use of technology is considered separate from the student's other activities

The first point stresses the importance of empirical studies of students' work. The second point opposes my research to the standpoint, that teaching may be performed independently of what tools the students have at their disposal. It is also in contrast to the view, that the influence of for example computer use can be overlooked as if it were just a matter of 'digitising the pencil case'. This is in line with what Jean-Baptiste Lagrange stresses in (Lagrange, 2005, pp. 131-132):

> *The traditional opposition of concepts and skills should be tempered by recognising a technical dimension in mathematical activity, which is not reducible to skills. A cause of misunderstanding is that, at certain moments, a technique can take the form of a skill.*

Besides the very grouping of the pairs of perspectives, the construction of the 'tool – object' pair was another result of the impact of the theory of instrumental genesis. The pair composed by a 'tool' perspective on a mathematical conception and an 'object' perspective on the same conception is realised for example in problem-solving settings. The tool – object pair corresponds to Anna Sfard's process – object duality (Sfard 1991). The tool perspective on mathematical conceptions is opposed to the 'pure skill' viewpoint and it implies the technical dimension that Lagrange refers to in the above quotation. So, within the notion of perspectives, the term tool denotes mathematical processes, carried out by the student with his instrument to serve a concrete purpose. It should be remarked that the purpose of the activity makes the difference between the tool perspective and Sfard's process perspective, not the instrument. Activities that aim at construction of pragmatic knowledge, involves the tool perspective on the mathematical conception in question. If the purpose is construction of epistemic knowledge, the process perspective is involved. This tool perspective on mathematical conceptions is in accordance with Régine Douady's definition of concept as a tool:

> *We say that a concept is a tool when the interest is focused on its use for solving a problem. A tool is involved in a specific context, by somebody, at a given time. A given tool may be adapted to several problems, several tools may be adapted to a given problem.* (Douady, 1991, p. 115)

Mathematical activities, then, must be considered from a tool perspective when they are considered as parts or elements of a technique in the above mentioned sense. To take the corresponding object perspective, genesis of the instrument is requested. Therefore, the instrumental genesis is in a crucial way linked to and maybe a prerequisite for the change to object perspective.

2.4. Perspectives and representations, referred to in the definition of flexibility

The pairs of perspectives form three groups: perspectives intrinsic to mathematics, and perspectives relevant for construction of epistemic- and pragmatic knowledge respectively. As it was mentioned above, the distinction between epistemic and pragmatic knowledge intended to take into account that the conceptual tool flexibility should resonate with the theory of instrumental genesis. All the changes between perspectives and between representations are bidirectional. Each pair of perspectives can be developed reflexively during the bi-directional changes.

In the notion of flexibility the term representation means representation system or communication media in a functional sense. So, the representations function partly as the media for externalisations of internal conceptual systems (Lesh and Doerr, 2000, p. 364; Mousoulides, Sriraman & Christou, 2007) and resembling the use of the term in the KOM-report (Niss, 2002). There is no sharp distinction between the four representations. Especially, the technical representation in many cases widely overlaps with the others.

2.4.1. Perspectives intrinsic to mathematics

Local and global position

Local position and global position are intrinsic mathematical perspectives. For example, changes between them occur when a single member of a group of objects is picked out for closer examination or if a single object is incorporated into a collection or family of objects. Using CAS, this change is easily realised for example by the use of facilities as substitution and copy – paste.

General - specific

General status and specific status are intrinsic mathematical perspectives, which concern the *domain of validity*. Inductive reasoning is linked to changes from specific to general perspective, whereas deductive reasoning is linked to changes from general to specific perspective. The use of CAS allows a quick and easy generating of specific objects from general expressions or formulas.

Analytic- constructive

The analytic and constructive perspectives are well-known phases in working with geometry. Examples are *measuring a given thing* and *construction* of a figure with a given measure, respectively. Especially, changes between these two perspectives are of relevance in sequences of modelling at a functional level.

2.4.2. Perspectives linked to construction of epistemic knowledge

The process - object duality.

The process perspective is operational and the object perspective is structural. The distinction and the connection between these two perspectives are encompassed by the reification theory. In contrast with the viewpoint of reification theory, the students' mental activities captured in the notion of flexibility, though, may go in both directions and thereby support the development of both perspectives.

Situated - decontextualised

The situated perspective is used to capture concrete aspects of problems, and of handling actual problems and challenges. The decontextualised perspective is the result of extracting rules, aiming at internal stability, which can occur through abstraction. Changes between situated and decontextualised perspective are linked to reflections, but not necessarily to modelling at the functional level. Therefore, this pair of perspectives relates to formation of epistemic rather than pragmatic knowledge. Changes from situated to decontextualised perspective correspond to the raising from referential to general level in the four-level model described above.

2.4.3. Perspectives linked to construction of pragmatic knowledge

The tool - object duality
The tool perspective[9] focuses on the use of a mathematical conception for solving a problem in a specific context, with some specific aids. The corresponding object perspective means the distant overview-perspective on the tool. Flexibility in this case encompasses the student's distant overview over a collection of mathematical tools, besides his capability to change from the distant over-view perspective into a tool perspective in the actual context.

Model - reality
The changes between reality and model perspective result in level-raising from the first to the second level in Gravemeijer's four-level model. In a number of cases, reality is expressed in natural language, and changes from reality to model then happens in connection with a change of representation from natural- to formal language.

Model of - model for
The changes between model of and model for perspective result in raising from the referential to the general level in Gravemeijer's four-level model. The change is exemplified in the case of calculus by K. Gravemeijer and M. Doorman (Gravemeijer & Doorman, 1999, p.111).

2.5. Representations

The term representation is used in a very broad sense. Consequently, changes between the four representations may occur in different contexts, and happen at a variety of levels. As part of the students' modelling and problem solving activities, the communication media may for example change at a functional level. In contrast, in other cases a single change to natural language may involve cognitive challenges like interpretation of a graph or ascribing meaning to a symbol.
The four representations do not intend to form any classification of media for expression of mathematical conceptions. In line with this, computer language is included as a representation on its own although it overlaps with the other representations, because changes to and from computer language is one part of the instrumental genesis.

Analytic representation (formal language)
Analytic representation includes formal expressions and formulas, algebraic expressions and symbols. Changes to analytic representation are related with symbolising whereas changes from analytic representation are connected with interpretation of symbols and the ascribing of meaning to symbols. The changes may cover more or less complex actions: For example, a shift to formal language may imply a routine translation from the graph of a linear function into the corresponding formal expression. Or it may describe the level raise from situated to formal level or from referential to general level. During modelling processes, the changes to formal language often happens stepwise, passing stages of partly formalised in-between-expressions.

---

[9] Here, the term tool is used in a broad meaning, not synonymous with artifact in contrast to instrument like in the theory of instrumental genesis

Graphical representation
Graphic representation includes graphs, curves, diagrams and tables, drawings etc. and explanatory gesticulation. Changes to and from graphic representation may be the result of for example the student's construction of a spatial conception of a curve that corresponds to the formal expression of a mathematical conception. Though, the curved line in the calculator's window, created by one press on the button is regarded as a graphic representation of the conception in question as well.

Natural language
The representation natural language encompasses spoken as well as written expressions and includes talk, explanations and negotiations, some texts written by the students or the by teacher, and some textbook texts.

The teacher's assessment has an impact on the communication in the classroom: depending on the social norms in the classroom, the students' can learn standard phrases of explanation by rote. The teacher's assessment of the student's understandings often relies on the student's capability to handle the content 'in his or her own words'. Standard phrases resembling own words, of course, can still be learned by root. Nevertheless, in the notion of flexibility such standard phrases are considered natural language. Natural language can overlap with formal language when technical terms are used, for example in group discussions.
The changes to natural language in many cases cover complex processes of interpretation and construction of meaning.

Technical representation (computer language)
Changes to technical representation encompass the translation from other representations into the version, adapted to the computer or calculator in question. Changes to computer language span from simple routines where students choose a well known instrument, to complex processes of instrumental genesis. Expressions in the other representations and changes between them are mediated in computer language: graphs, formal expressions and natural language can all be expressed in technical language too. In some cases, further, the software allows simultaneously use of formal language and graphic representation.

3. Methods and modes of the research

The initial construction of flexibility was based on teaching experiences, a preparatory classroom study and literature studies[10]. The teaching experiences were partly my personal ones; partly from the World Class A project referred to in the evaluation interviews, and partly second-hand ones, collected over years in informal talk and discussions with teacher-friends and colleagues.
Subsequently, I carried out an empirical study which aimed to

- Identify signs which could indicate flexibility in the students' mathematical conceptions
- Inquire how the teaching, the task, the teacher's questions etc. provoked the students to demonstrate flexibility
- Interpret the role of flexibility for the students' further working with mathematics
-

[10] The research design, the empirical studies and inquiries and their relations and roles in the design are presented and discussed in (Andresen, 2006, chapter 6)

The outcome of the study appeared at different levels. For example, one of the overall conclusions was that flexibility of the students' conceptions in general was prerequisite for acting with competence in the KOM sense (Niss, 2002). In several concrete cases I concluded that the use of certain software commands enhanced the flexibility of the students' mathematical conceptions. Finally, at the level of analysis I concluded that the conceptual tool flexibility was useful to throw light on the students' process of getting used to computer use.

Four small-scale teaching experiments served as cases for the empirical study. The participating teachers designed teaching sequences on differential equations for the experiments based on shared written materials in the form of a booklet, prepared by a group of teachers from the development project (Hjersing, N., Hammershøj, P. and Jørgensen, B. 2004). For the teachers, the aim of the teaching experiments was to develop good practices of teaching differential equations from a dynamical systems point of view. The booklet focused on modeling with a problem solving approach.

Data from the teaching experiments had the form of field notes and film recordings from classroom observations, students written reports, group interviews with teachers and students, and teaching materials. I observed fifty lessons spread over the four classes with 22, 12, 10 and 6 lessons during the winter 2003-2004. The lessons were chosen, restricted by practical circumstances. The four teachers who had voluntarily agreed were very helpful and obliging supportive. Thanks to them, the students were also helpful and open-minded.

I observed almost all lessons on the subject differential equations, taught by one of the booklet's teacher-authors. The goals were: i) to study the authors' overall intensions with the modelling approach to the subject, ii) to qualify my reactions on the booklet to the authors and iii) to inquire the use of laptops in class room teaching. In one other class, my observations focused on the modelling aspect. In the last two classes, I observed students' group work on project tasks. It was my hope that the dialogue and negotiations between the students in the groups would reveal signs of their learning process at a closer and more personal level, compared to the classroom observations.

Data were studied during interpretative analysis, taking social and psychological perspectives into account. The analysis followed the approach, designed by Paul Cobb et al. to meet the following three criteria (Cobb, Stephan, McClain & Gravemeijer 2001 p 116):

- Enable documentation of the collective mathematical development of the classroom community over the extended periods of time covered by instructional sequences
- Enable documentation of the developing mathematical reasoning of individual students as they participate in the practices of the classroom community
- Result in analyses that feed back to inform the improvement of the instructional designs.

4. Episode from a case study

The rest of this paper presents an excerpt from one of the Ph.D. project's nine cases. The objective of this part of the research was to inquire into the hypothesis: the construct 'flexibility' may capture important elements of learning activities, like they were described in the first parts of this paper. The excerpt is the larger part of one (the second) of the case's three episodes. It intends to illustrate the analyses of data following the three main aims, listed above. My analysis

of the episode mostly concentrates on the changes between model of perspective and model for perspective and on the inquiry of the model for perspective as it appears in two groups of students' work. Obviously, other perspectives and representations are involved and might have been chosen as the object of analysis as well. In the thesis, other cases concentrate on other perspectives and representations[11].

4.1. Case 8, chemical reactions

The theme of case 8 was changes between model of perspective and model for perspective. This case encompassed episodes 8.1 to 8.3. The main aim of the analyses of the case was to inquire the model for perspective as it appeared in some of the students work. Two groups of students (group 9 and group 11) in two different schools worked with the same project task. The project task concerned with exploration of differential equations' models of the rate of chemical reactions of order zero, one and two. The episodes in case 8 were based on the project task, the written reports from the two groups of students, a transcription of film recordings of one of the group's (group 9) work with the project task during one lesson, and my field notes from the same lesson. The project task took the model of chemical reactions as its starting point[12]. After a few introductory statements concerning the technical handling of the amounts and concentrations of the compounds, for example in the case of precipitation of a compound, the model of the reaction rate was introduced in the task:

> Theoretically, the rate of chemical reactions in general is expressed as:
>
> The rate of combination of two or more chemical compounds is proportional to the product of their concentrations
>
> The relation is expressed like this:
>
> $$Rate = \frac{dc}{dt} = \mp k[A]^x[B]^y[C]^z \qquad (1.1)$$
>
> That is, the rate of production/consumption of C is proportional to the concentration of the reactants ([A], [B], [C]) raised to the power of x, y and z respectively. The degree of the exponent denotes the order of the reaction.

Fig 2. Excerpt from page 51 in (Hjersing et al. 2004), (author's translation)

The task's design intended to encourage the student's explorations of this model, aiming to see special cases and recognise them as models of certain reactions of order zero, one and two. The students in the two groups were taught these topics in chemistry in advance. In their study program, they all combined high-level chemistry and high-level mathematics. In the textbook the special cases of the model were deduced from the general expression and then treated mathematically. In that sense, the project task dealt with changes between model of perspective with reference to the chemical setting and model for perspective with reference to the more general model in the mathematical setting. In the following, only episode 8.2 is considered.

---

[11] See the overview over Cases and episodes p. 202-204 in (Andresen 2006)

[12] For an introduction to the rates of chemical reactions and equilibrium see H. F. Holtzclaw, W.R. Robinson and W.H. Nebergall (1984): *General chemistry*, D.C. Heath and Company, USA pp 407-449 or P.W. Atkins (1990): *Physical Chemistry.* Oxford University Press pp775-810

4.2. Episode 8.2: group 9 answering question 3

When the episode took place, two students from group 9 worked with the chemical reactions project whereas the third member of the group was absent. At this time the students had passed the first one and a half page of the task's text, where the reaction rate was modelled for reactions of the different orders. The two students worked with page 53 in the booklet:

Irreversible second order reactions

Consider the irreversible reaction $A + B \to X + Y$.
One molecule A and one molecule B combine to one molecule of each of the compounds X and Y. The rate of consumption of A and B equals the production rate of X and Y.

We have inquired second order reactions in the simple cases, where the initial concentrations of the reactants were equal. But what happens, if [A] does not equal [B] from the beginning, or if some X or Y is already produced?

This time, we will model the production of X. The differential equation, mentioned earlier, now changes into:

$$\frac{dx}{dt} = kab \qquad (1.6)$$

We now want an expression on the right side, which only depends on the immediate [X], and on some initial values of [A] and [B]. So the goal is to express a and b as functions of x.
The immediate concentration of a equals: The initial concentration of a (a_o) minus the 'new' x. The 'new' x equals the actual concentration of x minus the initial concentration of x (x_o). Similar to [B], so all in all gives (1.6) the equation:

$$\frac{dx}{dt} = k(a_0 - (x - x_0))(b_0 - (x - x_0)) \qquad (1.7)$$

$$\frac{dx}{dt} = k(a_0 + x_0 - x)(b_0 + x_0 - x) \qquad (1.8)$$

Example

2 mol silver nitrate $AgNO_3$ is mixed with 3 mol hydrochloric acid HCl. White silver chloride precipitates and the reaction runs completely. In this case there is ½ mol silver chloride when the reaction starts. The rate constant k is 1.

$$AgNO_3 + HCl \to AgCl + HNO_3$$

Based on the text we state the following:
$k = 1$
$a_o = 2$
$b_o = 3$
$x_o = ½$

Then, the differential equation is:

$$\frac{dx}{dt} = 1(2 + \frac{1}{2} - x)(3 + \frac{1}{2} - x) \qquad (1.9)$$

Task3

Find the equilibrium points, where the rate equals zero for the differential equation in (1.8) and explain, what this means in practice

Fig3. Page 53 from the booklet (Hjersing et al. 2004), author's translation

In terms of flexibility, (1.1) gives a model for perspective since the level of the equation is general, whereas the perspective of (1.9) is model of because the level is referential – it refers to the actual experiment in the task.
Therefore, (1.8) can give both perspectives on the equation, depending on which of the two others it relates to:
<u>(1.8) in model of perspective with regard to (1.1)</u>
The model (1.8) is a general application of (1.1) to the case of a second order irreversible reaction with two reactants.

In (1.8), $\left[X\right]=x$ equals the concentration of the product,
$\left[A\right]=a=(a_0+x_0-x)$ equals the concentration of one of the reactants, and
$\left[B\right]=b=(b_0+x_0-x)$ equals the concentration of the other reactant. a_0 is the initial concentration of the compound a, similar for b and x.

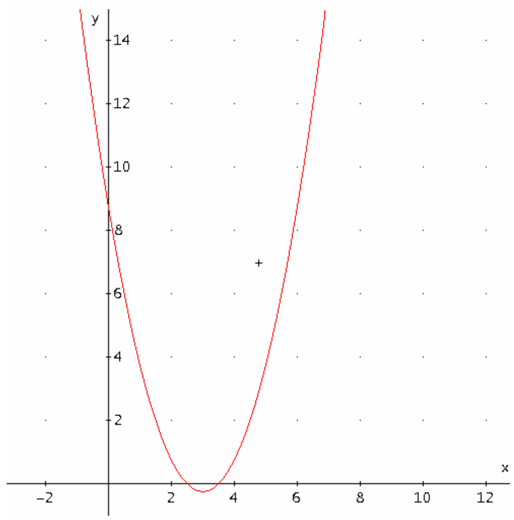
Fig 4. Explanation of (1.8)

So, with regard to (1.1) (1.8) serves as a model at referential level, referring to second order reactions. The perspective of (1.8), then, is model of whereas the perspective of (1.1) is model for.
<u>(1.8) in model for perspective with regard to (1.9)</u>
The model (1.9) is an application of (1.8) to the case of the reaction between Ag^+ and Cl^- where *a* stands for the concentration of Ag^+ (or $AgNO_3$) and *b* stands for the concentration of Cl^- (or HCl). *x* stands for the concentration of AgCl and dx/dt is the formation rate of AgCl. Since (1.9) is concerning with the actual formation of silver chloride, (1.9) is even more concrete than (1.8). So, (1.8) is a more general and decontextualised equation than (1.9). Consequently, in this case (1.8) serves to give the model for perspective and (1.9) serves to give the model of perspective on the equation.

The example and task 3 (in fig 3.) both aim to make the students change in both directions between the general mathematical model (1.8.) and the model of the actual reaction (1.9). Therefore, in terms of flexibility, they encourage to changes between model for and model of perspective. These changes happened stepwise in the actual case.
In the written report, the answer from group 9 to task 3 (fig 3.) was:

- Where does the rate for the differential equation equal 0? Where the rate equals zero are the equilibrium points.
To make the differential equation equal zero, one of the 3 factors in the differential equation has to equal zero.

$$\#35: \quad f(x) := 1\cdot\left(2 + \frac{1}{2} - x\right)\cdot\left(3 + \frac{1}{2} - x\right)$$




```
#36:   SOLVE(f(x), x, Real)

#37:                      x = 3.5 v x = 2.5
```

It is seen from the graph, that the equilibrium points are 2,5 and 3,5.
This means that the concentrations are unchanged in these points. From a mathematical point of view, both these points are useful, but in practice only 2,5 is of relevance because the reaction will stop here.

Fig 5. The students' answer

Apparently, the students linked the notion of equilibrium points with the roots of the polynomial. In their text, they interpreted the term equilibrium point in natural language in three (slightly) different ways, namely as

1) Points of the reaction where the reaction rate equals zero
2) Points of the reaction where the concentration is unchanged
3) A point where the reaction stops

In all three interpretations the equilibrium points were seen in a model of perspective. The students made a change to formal language when they turned to the equation #35 (fig 5.) and simultaneously to graphic representation, which was used in parallel (mixed in between to analytic expressions) with the formal language.

The students explicitly interpreted the actual values of the roots as concentrations. This interpretation revealed that they could handle the changes between equilibrium points in the model perspective (referential level) to the concentrations in the reality perspective. Later on, the students changed to a model for perspective on the equilibrium points, where they interpreted the equilibriums as the roots of the polynomial. They referred to the distinction between referential level (that is, model of perspective) and general level (that is, model for perspective) with the remark "*From a mathematical point of view*" in contrast to practice, which was understood to be at the referential level, that is, in a model of perspective.

By the students interpretation of the equilibrium point as a point where the reaction stops and where the rate equals zero, they linked the conception of equilibrium closely to the conception of rate. From the excerpt of text from the report (fig 5.), it is uncertain to say whether the students' chemical conception of equilibrium was a dynamic or a static one. The third point, (3), above and the following excerpts from the recordings of the students' work with task 3 may suggest that the students perceived the conception of chemical equilibrium in a static way even if the conception of rate was well described in their report's introduction (mentioned in the previous episode[13]). The students tended to confuse the rate and the rate constant:

S2: We look for the equilibrium point. Rate and rate constant is not the same, isn't it?
S1: The rate…
S1: Let us just…
S2: The rate, is it k or is that simply the rate constant
S1: It is the rate constant
S2: Then I guess we have to isolate k

In this way, then, the students were tempted to completely reduce the complexity of the problem. The dialogue revealed that they might doubt whether the equations (1.1) (fig 2.) , (1.8) and (1.9) (both fig 3.) modelled the reaction rate mathematically (general level) as well as chemically (referential level). According to my interpretation, hence, what confused them was the combination of mathematical and chemical models. Combining them is the very issue of changing between model for perspective and model of perspective in this case.
The dialog continued:

S1: No, because the rate is zero
S2: Then we simply have to write it…
S1: I do not understand what we are supposed to do. (reads) "The equilibrium points where the rate is zero"
S2: Where equal amounts are being produced. Where the amount left is the same as what is produced, isn't it?

Here, S2 referred to the chemical conception of dynamical equilibrium. In the next remark, S1 interpreted this involvement of a chemical conception as a change from (mathematical) model perspective to (chemical) reality perspective, but S2 maintained the model perspective by referring to *theory*:

[13] Episode 8.1 page 251-252 in (Andresen 2006)

S1: This is explaining what it means in real world, isn't it?
S2: But didn't we learn theory about this in chemistry?
S1: I do not understand what they... when the rate is zero, is it simply the rate constant?

The students called the teacher who referred to page 16-17 in the booklet, where the equilibrium solutions were considered:

$$\frac{dP}{dt} = k \cdot P \cdot \left(1 - \frac{P}{N}\right)$$

(...)

The right side of the equation is a polynomial of second degree in P with the roots

$P(t) = 0$, $P(t) = N$

It is positive between the roots and negative outside.

Fig 6. Excerpt from (Hjersing et al. 2004, p.16), author's translation

S1: (reads) the right hand side of the equation is a polynomial of second degree, find the roots,..
T: So what is this?
S1: I am reading... Equilibrium solutions – that is what we are looking for
T: Equilibrium solutions, what is it then?
S2: It was negative and positive second degree

The problem had now been transformed to the simpler one of finding the roots of a polynomial of second degree. This problem was expressed in graphic representation, natural and formal language simultaneously. The problem was seen in a model for perspective[14] as far as it had to be solved in a purely technical context with no connection with the chemical model.

When the dialogue continued, the teacher tried to establish the connection between the simpler problem and the question of equilibrium solutions to the differential equation. In terms of flexibility, the teacher insisted on keeping the changes between model for perspective and model of perspective in focus of interest. Shortly after, the episode was concluded with the teacher having accomplished the guidance of the students. The students went to the computer room for doing some graphs.

4.3. Conclusion of case 8

The changes between model of and model for perspective were in this case realised as changes between a mathematically formalised chemical model and a more general mathematical model. It was understood in the project task's text that the students knew the formalised chemical models in advance. So, the students were supposed in advance to manage changes between the chemical reality, in the form of proceedings of the actual reactions, and the formalised chemical models. In the episodes, the changes between the two perspectives went in both directions: the report from group 11 for the major part took the model for perspective (illustrated in episode 8.1), but the

[14] This situation illustrates a tool perspective on the solving of quadratic equation

task's questions provoked changes to and fro a model of perspective and further interchanges between model and reality perspective (it was demonstrated in episode 8.3). The report from group 9 divided into parts that took the model of perspective and other parts that took a model for perspective. The links between the two perspectives, though, were not always clearly revealed (it was shown in episode 8.3).

Episode 8.2 demonstrated how the teacher guided the students to change between model of and model for perspective. The episode pointed to the important role of the guidance from the teacher, as well as from the task's questions.

During the complete case, changes between reality and model perspectives and between model of and model for perspectives were realised on the relation, modelled by the differential equations (1.1), (1.8) and (1.9). The same relation was expressed in all the four different representations. According to the definition of flexibility, the individual student's flexibility of the conception of this mathematical relation (e.g. the differential equations model), in this context, was developed to the extent that he or she was able to manage these changes. Working with the tasks in the case seemed to support this flexibility since the individual student was more likely to manage the changes after the teaching sequence.

5. Conclusion

To conclude this paper three questions are discussed: what contribution does the construction of flexibility add to the field, are the promises, criteria and requests from the first part of the paper kept, and what questions does this paper not answer?

5.1. What is new?

The introduction of a new construct like flexibility with its definition and technical notion to the field of math education, which may seem almost overloaded with a diversity of notions and terms already, has to be justified. One important argument is that the introduction of flexibility summarises, connects and simplifies key elements of well established and acknowledged theory. The issue of visualisation, for example, is included in flexibility in terms of change to graphic representation. The relation between the levels in Gravemeijer's model is another example. The novel idea beyond the construction of flexibility is to focus on the dynamics as a common denominator, which leads to consider all these key elements in a new light and in a new combination.

The case illustrates how flexibility intends to serve as a tool for clarification: The case's task presented a rich and complex structure of models at different levels, and links between them. The task was suitable for the students' exploration of the relation, modelled by the differential equations, exactly because it was sufficiently complex to offer possibilities of open ended inquiries. During the case, one of the students' main difficulties concerned the changes between the levels of the model, represented in the three equations. Another main difficulty was the transformation of the problem to the simpler one of finding the roots of a second degree polynomial, followed by interpretation of the result. These main difficulties are closely related to the main learning potentials of the task, so attempts to avoid them by omitting parts of them for simplification would be of little use. The analysis of the case demonstrates how these main issues of the task can be interpreted in terms of flexibility. My claim is that the teacher could strengthen

the guidance of the students and make it more explicit without giving the answers, if he or she had such an interpretation in mind.
The case also illustrates that the apparent conflict of having the equation (1.8) at two different levels is cancelled when focus, in terms of flexibility, is on the dynamics of changing between levels rather than on placing conceptions at the levels.

5.2. Does flexibility keep the promises?

According to the initial goal of the research mentioned in this paper, flexibility should serve as a tool for teachers to take the project's experiences into account for improvement of their own teaching. The case gives an example of flexibility's potentials for improved guidance, according to my interpretation. To meet the goal, though, more focused teaching experiments, based on an elaborated description of the single elements of flexibility would be the next step.
The second request, which concerns the contribution to math education theory, was already discussed in a previous paragraph. The objectives of the construction set two demands on flexibility: signs of flexibility should be observable, and claims of its relevance should be theoretically founded. The construction of flexibility intended to capture important learning activities which promote the student's actual work with mathematics in an observable way. The case demonstrates that the changes of perspective and changes between different representations are observable. The presentation in this paper of three basic ideas beyond the concept of flexibility serves to justify[15] the claim, that the changes promote the student's learning.

5.3. What is left?

It follows from the research design that the effect was not tested on teaching designs, which aim at development of flexibility of the mathematical conceptions in question. A number of guidelines are presented to conclude the Ph.D. thesis. The guidelines are meant for teachers who want to aim at flexibility in the design of their teaching. A large scale inquiry of the effect could include teaching materials and teaching designs based on these guidelines, pre- and after tests and qualitative evaluation. The guidelines are (Andresen, 2006, pp. 294-295):

Teaching that aims to support flexibility in the mathematical conceptions with and without the use of CAS should be based on the following principles:

- The design of tasks and problems ensures that changes of perspective go in both directions in all the pairs of perspectives that form "flexibility":
    1. Local - global
    2. General – specific
    3. Analytic- constructive
    4. Process - object
    5. Situated – decontextualised
    6. Tool – object
    7. Model - reality
    8. Model of - model for
- The design of tasks and problems ensures that changes between representations go in both directions between *graphic representation, analytic representation (or formal language), natural language* and *technical representation (or computer language)*

[15] Further justification builds on the discussions all through chapters 5 and 7 in (Andresen, 2006).

- Expressive work and explorative work with mathematical models are both important. The teaching is designed to ensure that the students over time are encouraged to model a number of key conceptions including all the four levels situated, referential, general and formal.
- A diversity of strategies is not only accepted, but appreciated in the classroom. The students are encouraged to try out ideas and techniques. Results, ideas and strategies are discussed and negotiated with open minds in the classroom.

The relations between flexibility and the shortcuts mentioned in the first part of this paper, and the role of flexibility in the instrumental genesis, apparently, are issues in focus of interest for the continued work with development and refinement of the construct flexibility. It should be remarked that although the project took the use of laptops as its starting point, flexibility is not restricted to mathematical conceptions within a computer environment.

References

Ackermann, E. (1990): *From Decontextualised to Situated Knowledge: Revisiting Piaget's Water-Level Experiment*. MIT Media Lab, E&L Memo N.5. September 90

Andresen, M. (2006). *Taking advantage of computer use for increased flexibility of mathematical conceptions.* Danish University of Education. Copenhagen. (Ph.D. thesis)

Cobb, P., Stephan, M., McClain, K. & Gravemeijer, K. (2001). Participating in Classroom Mathematical Practices. *The Journal of Learning Sciences 10*(1&2), 113-163

Cobb, P., Yackel, E. & McClain, K. (Eds) (2000). *Symbolising and communicating in mathematics classrooms. Perspectives on discourse, tools, and instructional design*. Mahwah: Lawrence Erlbaum Associates.

Douady, R. (1991). Tool, Object, Setting, Window: Elements for analysing and Constructing didactical situations in mathematics. In A. Bishop., S. Mellin-Olsen, & J. van Dormolen (Eds), *Mathematical Knowledge: It's Growth Through Teaching.* Kluwer Academic Publishers.

Drijvers, P. (2003). *Learning Algebra in a computer environment. Design research on the understanding of the concept of parameter*" dissertation, Utrecht, CD-Beta press

Drijvers, P. and Gravemeijer, K. (2005). Computer Algebra as an Instrument. In D. Guin., K. Ruthven and L. Trouche (Eds). *The Didactical Challenge of Symbolic Calculators* (pp. 163-196). Springer.

Gravemeijer, K. and Doorman, M. (1999). Context problems in realistic mathematics education: A calculus course as an example. *Educational Studies in Mathematics* 39, 111-129

Gravemeijer, K. and Stephan, M. (2002). Emergent models as an instructional design heuristic. In K. Gravemeijer, R. Lehrer et.al. (Eds). *Symbolizing, Modeling and Tool use in Mathematics* (pp.145-169). Dordrecht: Kluwer Academic Publishers.

Guin, D., Ruthven, K. and Trouche, L. (2005). *The Didactical Challenge of Symbolic Calculators*. Springer

Hjersing, N., Hammershøj, P. and Jørgensen, B. (2004). *Modeller i Derive. Differentialligninger og modelbygning. Højt niveau i matematik.* (Models in Derive. Differential equations and modelling. High level mathematics) Matematiklærerforeningen. Copenhagen

Lagrange, J.-B. (2005). Using symbolic calculators to study mathematics. In D. Guin., K. Ruthven and L. Trouche (Eds). *The Didactical Challenge of Symbolic Calculators* (pp. 113-136). Springer.

Lesh, R. and Doerr, H. (2000). Symbolising, communicating and mathematising: Key components of models and modelling. In P. Cobb., E. Yackel, E. & K. McClain (Eds) *Symbolising and communicating in mathematics classrooms. Perspectives on discourse, tools, and instructional design* (pp. 361 – 383). Mahwah: Lawrence Erlbaum Associates.

Lesh, R., & Sriraman, B. (2005). Mathematics education as a design science. *Zentralblatt für Didaktik der Mathematik*, 37(6), 490-505.

Mousoulides, N., Sriraman,B., & Christou, C. (2007). From problem solving to modeling: The emergence of models-modeling perspectives, *Nordic Studies in Mathematics Education* (NOMAD), 12(1), 23-48.

Niss, M. (2002). *Mathematical competencies and the learning of mathematics: The Danish KOM project,* IMFUFA, Roskilde University

Sfard, A. (1991). On the dual nature of mathematical conceptions: Reflections on processes and objects as different sides of the same coin. *Educational Studies in Mathematics*, 22, 1-36

Sierpinska, A. and Kilpatrick, J. (1998). *Mathematics Education as a Research Domain: A search for Identity.* Dodrecht: Kluwer Academic Publishers.

Steffe, L. and Thompson, P. (2000). Teaching experiment methodology: Underlying principles and essential elements. In A. Kelly and R. Lesh (Eds). *Handbook of research design in mathematics and science education* (pp. 267-306). Mahwah: Lawrence Erlbaum and Associates.

Trouche, L. (2005). An instrumental approach to mathematics learning in symbolic calculator environments. In D. Guin., K. Ruthven and L. Trouche (Eds). *The Didactical Challenge of Symbolic Calculators* (pp. 137-162). Springer

Wittmann, E. (1998). Mathematics education as a “Design science”. In A. Sierpinska and J. Kilpatrick (Eds). *Mathematics Education as a Research Domain: A search for Identity* (pp.87-103). Dodrecht: Kluwer Academic Publishers.

Non-linear functions in secondary school of lower qualification level (German Hauptschule)

Astrid Beckmann[1]
University of Education, Schwäbisch Gmünd, Germany

Abstract
This article reports of the effectiveness of introducing non-linear functions in a German Hauptschule. Some research based recommendations are provided for practitioners with the goal of improving student competencies in lieu of PISA..

Keywords: classroom experiments; non-linear functions; non-linear relations; student competencies; PISA competencies; German Hauptschule

1. The Situation in the German Hauptschule (secondary school of lower qualification level)

The results of international comparison studies like PISA recommended that German educators increased their focus on the mathematical competencies of students in secondary schools of lower qualification level (Hauptschule). According to the results of PISA 2003 not even one-fifth of the students achieved level 1 of competence. Tasks at level 1 "demand to gather information from a simple table given in standard form or from a simple graph and to execute simple calculations which refer to relations between two familiar variables" (PISA, 2004, p 56). A further striking break is to be seen in the field of level 3 to 4. While almost a quarter of junior high school students (German Realschule) are at level 4, this does not even apply to 5 % of the students of the Hauptschule - apart from higher levels which are hardly reached by one. Tasks of level 4 require additionally "to argue also in less familiar functional contexts and to

[1] Astrid Beckmann
Professor Dr. habil.
Department of Mathematics and Computer Science/ head
University of Education Schwäbisch Gmünd
Oberbettringer Straße 200
D - 73525 Schwäbisch Gmünd
GERMANY
Tel.: +497171983359
email: astrid.beckmann@ph-gmuend.de
Web: http://mathematik.ph-gmuend.de/lehrende/beckmann/

The Montana Mathematics Enthusiast, **ISSN 1551-3440**, **Vol. 4, no.2, pp. 251-257**

communicate these arguments as well as to deal with given linear models of real situations" (PISA, 2004, p 56). Figure 1 shows a linear task, (not necessarily) a familiar real situation.

The picture shows the movement of a car.
In which traffic situation is the car?

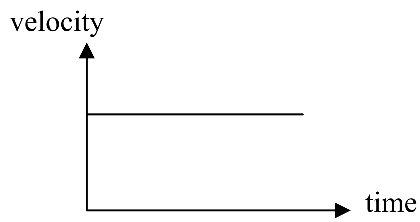


Figure 1

A characteristic answer from our own random sampling at the Hauptschule is the following: "The car is always driving straight on". This answer is in accordance with the observations from PISA and further studies according to which pupils interpret rather the optical picture than the functional interrelation (compare known racecourse task in PISA). Since at the same time they describe the graph in figure 2 as follows: "The car is rolling downhill."

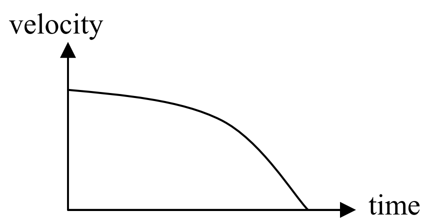


Figure 2

The question is how can Hauptschule students be supported in their learning in order to reach level 4 easier and be more successful?

2. About the importance of non-linear functions

Thesis:
For the achievement of level 4 the acquaintance with non-linear functions is beneficial.

A good reason for this thesis is that so far in the Hauptschule only linear functions are treated and take up a big space of time in the curriculum in comparison to other function types. These students experience functional contexts as mainly linear. This may lead to a restricted view and functional thinking becomes dispensable. For example the overemphasis of linear functions may lead to the perception that any functional coherence can be described by a straight line. Experiences from school projects show that students often think accordingly. If you ask them to demonstrate the relation between air volume and pressure graphically they mark a straight line with positive gradient even though the relation is inversely proportional. Furthermore the presumption can be supported that even two points are enough to specify the graph of a function. This has also been noticed in school projects when cubic relations have been rashly interpreted

as straight lines. Examples like these show that working with different types of functions requires more flexible thinking. It is not clear from the beginning how the graph is running, which output belongs to a certain x etc. Non-linear functions require thinking in changes and connections and lead to aspects of covariation (DeMorois & Tall 1996; Malle, 2000).

However non-linear functions can be complex and have typically been described algebraically. For example, intensity of lighting diminishes quadratically with the distance from light source. In the Hauptschule the functional term should not simply be the objective but the functional correlations with their dependencies, changes etc should be emphasized. However there is a certain difficulty included which is related to the graphical description. A better access to the functional connection can be produced by equalizing curves plotted instead of points connecting lines. According to our experience it has to be reflected before and during class in order to approach this result.

3 .The School Project

In November 2006 a first testing of the treatment of non-linear functions took place in a Hauptschule in Baden-Württemberg (Ostalbkreis, Germany). In different stations 9th grade students learned about non-linear functional relations in experimental activities. The students had already been used to experiments. At the end of 8th grade there had already been a course about linear functions only. Experiments had been chosen because of prior experiences in this area conducted in different school types with positive results from such experiments in connection with functions (Beckmann 2006, 2007). In the performance of these experiments the aspects of the contents of function like correspondence and covariation can be experienced by action. For example, in dipping a ball with a certain radius into a jug filled with water a certain water volume is edged out. The experience is made that any radius corresponds to a certain volume. With a running car the concurrent change of line and time, i.e. covariation, can be witnessed directly etc. (see Dubinsky& Harel, 1992; Vollrath 1978). In the school project the experiments were conducted together with worksheets. Worksheets proved to be very helpful in order to activate certain trains of thoughts, which could be deepened in the final presentation. Main aspects are summarized in table 1.

Aspects	Example
Everyday activities Forming the Thesis	Pumping a closed bicycle pump and then discuss on it
General Task	Describe the correlation between air volume and air pressure. Check: does the correlation confirm your answer to the above question? Describe the special characteristics of the correlation.
Operation of the Experiment	Get familiar with the parts of the experiment and the components of the measuring possibilities.
Correspondence	Adjust a certain value for the air volume and record the value in the box. Read the corresponding pressure. Record …

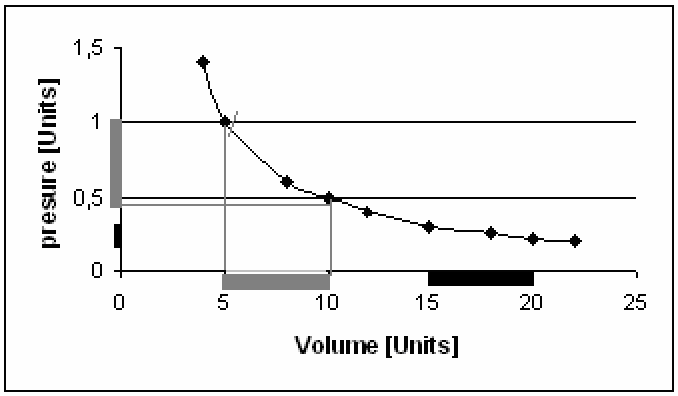
Change between different forms of presentation	Act in a systematic manner and measure. Record in the table…. Characteristics of the table… Enter the values from the table in the system of co-ordinates. Regard the graph. Describe it.
Correspondence	Which pressure corresponds to a volume of 5 units?
Covariation	Highlight in the graph by a thick line only the change from 5 to 10 (from 15 to 20) volume units at the x-axis and only the corresponding change at the y-axis. Compare the particular changes… 
Closure	Document the results of this station clearly and neatly on the prepared poster. Remind the general task.

Table 1 Structure of a worksheet

In the school project the following experiments were chosen out of a big collection (Beckmann 2006)[2]:

- Functional relation between air pressure and air volume
- Relation between intensity of light and distance of the light source
- Relation between radius and volume of cylinders of the same height
- Relation between height and time of a falling ball
- Relation between distance and time of an accelerated car
- Relation between force and lever of force

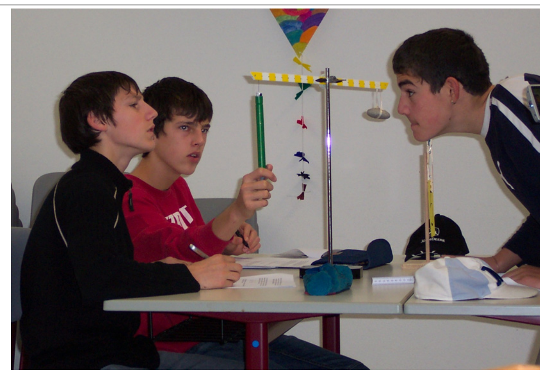


Figure 3
Experimenting Students of the Hauptschule

Implementing these experiments needed two double-lessons, in which the students experienced two or three relations. The division into student groups was decided by the teacher. Before starting the experiments a short introduction including questions of graphical description was given.

4. Observations and results

The lessons were video and audio taped, the worksheets were evaluated and single episodes were recorded. The evaluation shows that the experimental activities in non-linear contexts leaded to some aspects according functional thinking. The difference to the (expected) simple proportional relation and the missing possibility of predicting more values led to discussions. The students experienced and applied the idea of covariation. They recognized that only two values do not lead to an indisputable statement about the function.

Reflection about the table; extracts from worksheets:

> "The lower the distance [height of fall], the shorter the needed time [of the falling ball]."
> "The distance does not double." [comparing the distances of the accelerated car between 1 and 3 s with the distances between 2 and 4 s time of going].

Forming a thesis using the aspect of covariation:

> S1: Now we hang it at the end.
> S1 measures 25.5 cm
> S2 measures: 0.6 N. Actual this has to be at 36 cm.
> S3: It should be one less the more you hang it at the back.
> S2: Now it is 18.
> S1 measures: It is 1.3 N.
> S3: The next should be 0.1 N more.
> S1: At 13 cm it is 2.1 N.

Checking the thesis and gaining in through experiments; extract from a taking down:

[2] As agreed with the mathematics teacher Mr. Arthur Litz, Sebastian-von-Drey-Schule in Ellwangen-Röhlingen

The students discuss the relation between the intensity of light and the distance from the light source. In a first measurement they have obtained two values.
S1: Normally it is not proportional
S2: Eh, than let´s do it; then we will know.
S1 draws a system of co-ordinates. Together the students discuss the division of the axis and plot it. Finally the measurement is filled in.
S2: That is not proportional.
S1: Sure?
S3: Except that the line runs like this.
S2: That is not proportional, as I said before.
The students take more tubes and take many measurements, before they bear out their assumption.

Figure 4
Measuring the intensity of light at a tube

The following final report of a student shows that he tackled with correspondence and covariation:

> "The second project - we called ´light and tunnel´. The more the car drives into the tunnel, the more it becomes dark. The first tube had a length of 9.7 cm. When we held it at the window the intensity of light was very good 36 lux. When we held a tube with 30 cm length at the window, we could only measure 0.1 lux. In an extra drawn graph we could read exactly, how much intensity there is in the beginning of the tunnel and in how far it decreases."

5. Summary and Perspectives

Starting point is the thesis that working with non-linear functions supports flexible functional thinking, thinking in correspondences, dependencies and changes. This kind of thinking is seen to be a condition for competencies on level 4, which only few students of the Hauptschule reach. An access to non-linear functions is given by simple experiments. While doing experiments the students can experience the aspects of functional relations.

In this school project 9th grade students of a secondary school of lower qualification level (German Hauptschule) were confronted with non-linear relations in experiments. Obviously the non-linear functions led to discussions about dependencies and the course of the graph. Possibly the experiences of the first run of the project played a role, in which the students learned linear

relations only. Summing up the worksheets with detailed questions and the final classroom presentations with critical discussions were an important basis for a fruitful tackle with functional relations. A certain difficulty was the inclination of students to draw a straight line instead of equalizing curves. This could be cleared up in appropriate follow up discussions.

This single school project cannot lead to a conclusive answer to the question to what extent the non-linear experiments led to success regarding competence level 4. But the results show that non-linear functions are an appropriate theme for 9th grade students in a secondary school of lower qualification level. It gives great hope.

References

Beckmann, A. (2006): *Experimente zum Funktionsbegriffserwerb im Mathematikunterricht der Sekundarstufe I und Anfang der Sekundarstufe II*, in press.

Beckmann, A. (2007*). Funktionsbegriffserwerb durch Experimente*. In press in *Mathematik Lehren*.

DeMorois, PH.& Tall, D. (1996). Facets and Layers of the Function Concept. Puig, L., Gutierrez, A. (Eds.). *Proceedings of the 20th Conference of the International Group for Psychology of Mathematics Education* (PME 20), 2, 297-304

Dubinsky, E. & Harel, G. (Ed.). *The Concept of function. Aspects of epistomology and pedagogy*. Mathematical. Association of America.

Malle, G. (2000). Zwei *Aspekte von Funktionen: Zuordnung und Kovariation*. *Mathematik Lehren,* 103, 8-11

PISA-Konsortium (Hg.): *PISA 2003, Der Bildungsstand der jugendlichen in Deutschland – Ergebnisse des zweiten internationalen Vergleichs*, Münster (Waxmann)

Vollrath, H.-J. (1978). Schülerversuche zum Funktionsbegriff. *Der Mathematikunterricht* 4, 90-101

Looking back at the beginning: Critical thinking in solving unrealistic problems

Mark Applebaum[1]
Kaye Academic College of Education, MOFET Association, Israel
&
Roza Leikin
University of Haifa, Israel

Abstract

We believe that problem-solving skills engage critical thinking at every phase of problem solution. In this research a special attention is given to the fist phase - "understanding the problem". We consider this phase as a continuation of all the previous mathematical experience, in which understanding of new problems requires "looking back" at those solved in the past. Evaluation of the givens in the problem sometimes allows immediate solution whereas in other cases it shows that solution does not exist. We found that it is not easy for mathematics teachers to discover that a problem includes contradictory (i.e. unrealistic) conditions. We suggest that such problems should be included into teachers' professional development programs to develop teachers' awareness of the importance of mathematical accuracy and connectedness.

Keywords: Algebraic and geometric tasks; Critical thinking; Problem solving; Polya style heuristics; Teachers professional development

1. The background

1.1. On the importance of unrealistic tasks

In his extensive study on students' mathematical abilities Krutetskii (1976) included *unrealistic problems*, i.e., those that include contradictory givens, as one of the types of problems that allow examination of understanding of mathematical material learned by a student, "which shows up in its processing and retention". An example of such a problem is the following:

---

[1] The Head of Mathematics' Department,
Kaye Academic College of Education,
Beer Sheva, Israel
E-mail: mark@kaye.ac.il

***The Montana Mathematics Enthusiast*, ISSN 1551-3440, Vol. 4, no.2, pp. 258-265**

Task 1: *What is the area of an isosceles right triangle with leg equal to 5a cm and hypotenuse to 12a cm?* (ibid. p. 133)

It is clear that such a triangle does not exist. The given measures of the sides of the triangle and its type (isosceles right triangle) are contradictory. Krutetskii assumed that solvers who solve those problems out of specific context would not be able to realize the unrealistic character of the situation described by the problem.

Our interest in teachers' solving unrealistic problems is based on our belief that identification of unrealistic problems is an integral part of teachers' mathematical knowledge: Understanding of the unrealistic conditions of a problem demonstrates connectedness and consistency of ones mathematical knowledge while identification of inconsistent data when solving a problem demonstrates person's critical thinking. From the pedagogical point of view the ability to identify unrealistic situations may prevent generation of ill-defined examples in the course of mathematics lessons as well as strengthen teachers' critical view of textbooks and other instruction materials.

This research was motivated by our observation that shows that pre-service teachers usually do not identify unrealistic conditions of a problem. It is supported by the analysis of teachers' performance on sorting conditional statements task (Zaslavsky and Leikin, 2004). Zaslavsky & Leikin demonstrated that teachers usually relied on external features of the equations and inequalities (e.g., types of functions, types of conditional statement), and they seldom considered the internal features – the domain and the range of the functions. For example, equation $\log(4-x)-\log(x-6)=1$ (task 2) does not have a solution (the solution is an empty set) since domains of the two functions are disjoint sets. Thus, there is no need to perform algebraic manipulations in order to find a solution; the result is immediate. Similarly no manipulative solution is needed for the inequality $\sqrt{x^2+px+1}<0$ (task 3): it does not have a solution since radical denotes "arithmetic square root" which is non-negative for any non-negative value of the function under the radical. To solve these tasks shortly, one has not only to think about algorithms of solutions of these problems but also to see the whole context, to connect all the pieces of knowledge related to the task.

There is a similarity in approaching the tasks used by Zaslavsky and Leikin (2004) and unrealistic tasks offered by Krutetskii (1976). Subjects with procedural understanding (Skemp, 1976) of the topic will approach the tasks algorithmically by applying formulas (e.g., task 1 – formula of the area of right triangle) and manipulations (e.g., for task 2 and task 3). Those with relational understanding will consider connections between the given mathematical objects, their properties and components and will conclude that solution does not exist.

1.2. On the cyclic nature of reflection on a solution and understanding a problem

Polya (1973) highlighted four main phases of problem-solving process: understanding the problem, devising a plan, carrying out the plan and looking back at the completed solution. He recommended checking the results and checking the argument in order to make sure that it is correct.

> By looking back at the completed solution, by reconsidering and reexamining the results and the path that led to it, they [students] could consolidate their knowledge and develop their ability to solve problems. (Polya, 1973, pp. 14-15)

We argue that in the case of unrealistic problems "looking back" should start at the planning stage of the solution. On the one hand careful analysis of the givens in the problem may outline a solution plan, on the other hand, it may reveal discrepancy of the givens and immediately show that the problem does not have a solution. As noted above we believe that critical thinking is one of the basic cognitive skills supporting and encouraging solution checking.

1.3. On critical thinking in mathematical problem solving

NCTM *Curriculum and Evaluation Standards* (1989) pointed out that "A climate should be established in the classroom that places *critical thinking* at the heart of instruction ... To give students access to mathematics as a powerful way of making sense of the world, it is essential that an emphasis on *reasoning* pervades all mathematical activity." (ibid. p. 25). Erroneously unrealistic problems may be seen as those that prepare students to real life through developing their critical thinking. The ability to think critically is essential if individuals are to live, work, and function effectively in our changing society.

Critical thinking includes the use of cognitive skills or strategies directed at desirable outcomes of human activities of different kinds: solving problems, formulating inferences, calculating likelihoods, and making decisions. It also includes using skills that are effective for the particular context and type of thinking task (Halpern, 1998).

> Critical thinking is a mental process of analyzing or evaluating information, particularly statements or propositions that are offered as true. It is a process of reflecting upon the meaning of statements, examining the offered evidence and reasoning, and forming judgments about the facts. Such information may be gathered from observation, experience, reasoning, or communication. Critical thinking has its basis in intellectual values that go beyond subject matter divisions and include: clarity, accuracy, precision, evidence, thoroughness and fairness (Wikipedia, 2005).

Ferrett (2002) included among other attributes of critical thinking the following: assessment of statements and arguments, admitting a lack of information, ability to clearly define a set of criteria for analyzing ideas, examining problems closely, being able to reject information that is incorrect or irrelevant. Critical thinking involves evaluation of the thinking process - the reasoning that went into the conclusion we arrived at.

2. The Investigation

2.1. The purpose

This investigation was aimed at exploring teachers' mathematical performance on unrealistic problems, critical reasoning associated with unrealistic tasks. In particular we examined whether teachers understand the contradictory nature of conditions given in the problem. We also analyzed teachers' views on such kind of tasks.

2.2. Population

We assumed that ET's may succeed better in unrealistic tasks both because of their teaching experience and of their educational background. Thus the population of our study included three groups of mathematics teachers as follows: Seventeen pre-service mathematics teachers (PT) and

48 experienced high school teachers (ET) from two groups participated in our study. PT's had BA in mathematics and were learning for teaching certificate. ETs' experience varied from 5 years to 28 years. Most of the ET's had MA in mathematics or mathematics education. These teachers participated in the study in two groups: ET1 included 27 teachers and ET2 included 21 teachers.

2.3. The instrument

The teacher were asked to complete a written questionaire. We assumed that conditional formulation of the tasks might evoke teachers' critical thinking. Thus the tasks were formulated in two versions. In one (*non-conditional*) version we asked the teachers to "solve problems". In other (*conditional*) version we asked the teachers to "solve problems if possible". PT's and ET's from one group (ET1) were presented with a non-conditional questionnaire while ET's from group ET2 were presented with a conditional questionnaire.

Our questionnaire included algebraic and geometric problems. Figure 2 shows the tasks presented to the subjects. It also describes correct, alternative and incorrect solutions to the tasks.

Algebraic task required from the teachers "to find sum of the squares of the real roots of the equations without calculating the roots" (A1: "find"; A2 "find if possible"). The teachers were given two equations that did not have real roots. As an integral part of the solution the teachers had to check whether the roots exist. There was no need to perform algebraic manipulations since the equation does not have real roots.

Task Ab had an additional control level: When missing the contradiction in the question at the beginning of the solution, one could find that the sum of the squares of the two numbers is negative ($\alpha^2 + \beta^2 = -\frac{3}{4}a^2 - 1$. and claim that the question does not have answer on the set of real numbers. Alternatively teachers could state that they found sum of the squares of complex roots of the equation.

Geometry task required form the teachers finding area of a right triangle according to its hypotenuse and altitude to the hypotenuse (G1: "find"; G2 "find if possible"). In these two tasks the length of the altitude was bigger than half of the side, thus these measures were inconsistent with the following property of a right triangle: altitude of a right triangle is not bigger than half of the side (see Figure 2). The solution was very simple both when noticing the contradiction and when missing it. Note that all the teachers in the sample group were familiar with the property.

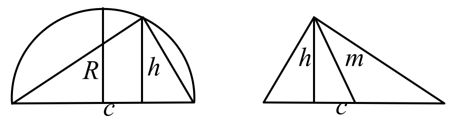
After performing the written assignment teachers checked and corrected their works. They did it in a different color so we could keep track of their initial solutions. Each session was concluded with a whole-group discussion. The discussion in group PT1 was video-recorded and transcribed, discussions with ET's were recorded in writing.

<table>
<tr><th>The task</th><th>Correct (content-connected) solution</th><th>Incorrect (algorithmic) solution</th><th>Alternative solution</th></tr>
<tr><td colspan="4">A1*: Without calculating roots of the equations (α and β) find* sum of the squares of the real roots ($\alpha^2+\beta^2$), for each one of the following equations:</td></tr>
<tr><td>Aa.
$x^2-5x+7=0$</td><td><u>No solution:</u>
$\Delta = 25-28<0$
$\Rightarrow$ no real α and β.</td><td>$\alpha^2+\beta^2=(\alpha+\beta)^2-2\alpha\cdot\beta=$
$=(-5)^2-2\cdot 7=11$</td><td>$\Delta<0 \Rightarrow \alpha$ and β are complex numbers
$\alpha^2+\beta^2=(-5)^2-2\cdot 7=$
$=11$</td></tr>
<tr><td>Ab.
$2x^2-ax+a^2+1=0$</td><td><u>No solution:</u>
$\Delta = a^2-8a^2-8<0$
$\Rightarrow$ no real α and β</td><td>$\alpha^2+\beta^2=(\alpha+\beta)^2-2\alpha\cdot\beta=$
$=\left(\frac{a}{2}\right)^2-2\cdot\frac{a^2+1}{2}=-\frac{3}{4}a^2-1$</td><td>$\alpha^2+\beta^2=-\frac{3}{4}a^2-1$
$\alpha^2+\beta^2<0$: impossible for real α and $\beta \Rightarrow$
1) no solution
2) α and β are complex numbers</td></tr>
<tr><td colspan="4">G1*: Find*area of the right angle triangles in each one of the following cases:</td></tr>
<tr><td>Ga. Hypotenuse is $\sqrt{11}+1$ cm, the altitude to the hypotenuse is $\sqrt{11}-1$cm.</td><td rowspan="2"><u>No solution</u>:
Altitude to the hypotenuse in right triangle is not longer than half of the hypotenuse</td><td>$S=\frac{ch}{2}$
$S=\frac{(\sqrt{11}+1)(\sqrt{11}-1)}{2}=5$</td><td>The condition that the triangle is right angled is surplus
If not right triangle: $S = 5$</td></tr>
<tr><td>Gb. Hypotenuse is $2a-1$ cm, the altitude to the hypotenuse is a cm.</td><td>$S=\frac{ch}{2}=\frac{a\cdot(2a-1)}{2}$</td><td>If not right triangle:
$S=\frac{a\cdot(2a-1)}{2}$</td></tr>
<tr><td colspan="4">* Questionnaires A1 and G1 included non-conditional requirement "find".
Questionnaires A2 and G2 included conditional requirement "find if possible"</td></tr>
</table>

Figure 1: The tasks in the questionnaires

3. Solving the tasks

For both versions of the tasks majority of the teachers produced algorithmic (wrong) solutions. Only one ET identified contradiction in the geometric task. All other teachers in both versions of Task G calculated area of the triangles using formula. Teachers succeeded better to some extent in solving the algebraic tasks (see Table 1).



Altitude to the hypotenuse in right triangle is not longer than half of the hypotenuse

Figure 2: Task G - explanation

On the non-conditional task, 4 of 17 PT's and 2 of 29 ET's examined delta both in tasks Aa and Ab and concluded that the tasks do does not have a solution. In task Ab, which included additional control level, 2 PT's and 6 ET's decided that the task cannot be solved since a sum of the squares of real numbers cannot be negative. When solving the conditional tasks 5 of 21 teachers examined delta at the beginning of the solution and did not perform algebraic manipulations.

Group of teachers Task		PT (N=17)	ET1(N=29)	ET2 (N=21)
Aa	Non-conditional	4(Δ<0) 23.5%	2(Δ<0) 7%	
	Conditional			5(Δ<0) 28%
Ab	Non-conditional	4(Δ<0)&2($\alpha^2+\beta^2$<0) 35%	2(Δ<0)&6($\alpha^2+\beta^2$<0) 27%	
	Conditional			5(Δ<0) 28%

Table 1: Teachers performance on Algebraic Tasks

As we expected, conditional task Aa was somewhat easier for the teachers than non-conditional one (Task Aa: 7% of teachers in group ET1 vs. 28% in ET2, Table 1). This tendency was also clear in task Ab at the planning phase of solution (Task Ab: 7% [2(Δ<0)] ET1 vs. 28% [5(Δ<0)] of ET2, Table 1). Then again additional control tool [$\alpha^2+\beta^2<0$] helped 6 ET's realize there was no solution for non-conditional Task Aa. Surprisingly, the conditional task, negative value of $\alpha^2+\beta^2$, did not help teachers who failed to realize that the equation did not have real roots at the beginning of the solution (by checking value of Δ). We assumed that conditional formulation of the problem could evoke critical reasoning to the same extent as the second level of control in Task Ab.

Opposite to expectations ET's did not appear to be more successful in performing the tasks then pre-service mathematics teachers. As a result of the experiment and subsequent group

discussions we have come to the conclusion that the fact that PT's are not less successful in solving this kind of tasks has a reasonable explanation that we outline in the next section of this paper.

4. Discussion of Tasks and Solutions

In the course of the whole group discussions all the teachers who had not realized contradictions in the problems conditions were upset. Some of them were annoyed by the fact that they had missed the inconsistencies, and some were even angry with us for presenting "such unfair tasks". However, majority of the teachers reported that they had enjoyed the experience. In each group there was a disagreement on the "unfairness" of the problems. Only a small number of teachers thought their experience was bad whereas most of them claimed that it was very positive.

There was clear distinction in the teachers' attitude to Task A and Task B. The teachers agreed that Task A was reasonable and "reminded them about the necessity of mathematical accuracy" since "any solution related to quadratic equation should start with examination of delta". One of the teachers said:

> How could I miss this? Finding delta when solving the tasks related to the quadratic equation is a part of the algorithm. It is a regular procedure. I always tell my students: "First check whether the equation has real roots, then find the roots, or do whatever the task requires". I just did not think [about delta]. How can you talk about something when you do not know whether it exists.

We agree with the teachers that Task A in our study was less provocative and more regular than Task G. Based on the teachers' reactions in the discussion, we think that performance on algebraic task reflects mathematical culture of their classrooms where under the pressure of time teachers sometimes do not require from their students precise and accurate mathematical performance, where algebraic manipulation and procedures are in the heart of the instructional processes. Similarly to Task 1 ($\log(4-x)-\log(x-6)=1$) accuracy as characteristic of critical reasoning and relational perspective, allows to shorten the procedure, even not perform it at all. Note that slightly better performance of PT's than of ET's on Task A1 we address to the classroom routines in which ET's are involved every day while PT's are still learning and more challenged by the courses in which they participate.

Task G presented to the teachers was found more "tricky". At the beginning of the discussion, the teachers felt they had never met such kind of tasks before. Though after discussing the task they made an analogy between Task B and other "tasks from the textbooks that include mistakes". Contrary to Task A, which teachers saw as "pretty regular for mathematics classes" and for which wrong solutions they considered "just the result of a mistake", Task G was considered by them inapplicable for the classroom situation. They saw this task as mistaken and claimed that teachers' duty is to avoid such tasks in classroom activities. In teachers' opinion, using textbooks is and should be "safe" and "the authors have to check many times the problems and the solutions that they include in the textbooks". The teachers were certain that problems of this kind confuse pupils. On the other hand, they agreed the task is good for teachers as it

requires thinking about mathematical connections, i.e., "other theorems related to the task, not only those you need for the solution of a specific problem".

During the discussion with PT's the distinction between solving proof tasks from the books and exploring conjectures raised in the course of an inquiry-based lesson was mentioned. Conjectures and hypotheses may be unrealistic and the inquiry procedure has to verify their realistic nature. Refuting a conjecture at the proof stage of inquiry is a natural procedure whereas when meeting a proof or computational task teachers presume that those who ask to prove or find something have already checked that this task is realistic.

5. Concluding remark

We finish this experiment with many open questions. Some of them are raised from our communications with the teachers and their replies while others raised from the mathematical analysis of the tasks we performed in the course of the study. Among other questions we ask: How can we formulate the tasks so that teachers' critical thinking will be evoked? We find Task 1 more transparent than Task G. How different will be teachers' solutions of Task 1 and Task G? How different or similar will be teachers reasoning associated with unrealistic problems and problems with consistent surplus conditions (See, for example, Krutetskii (1976)?

References

Ferrett, S. (2002) *Peak performance: Success in college and beyond*. McGraw-Hill Clooge.

Halpern, D. F. 1998. Teaching critical thinking for transfer across domains. *American Psychologist.* 53, 449–55.

Krutetskii, V. A. (1976). *The psychology of mathematical abilities in schoolchildren.* J. Kilpatrick & I. Wirszup: Eds., J. Teller: Translation from Russian. Chicago: The University of Chicago Press.

National Council of Teachers of Mathematics. (1989). *Curriculum and evaluation standards for school mathematics.* Reston, VA: Author.

Polya, G. (1973). *How to solve it.* Princeton, NJ: Princeton University Press.

Schoenfeld, A. H. (1986). On having and using geometric knowledge. In J. Hiebert (Ed.). *Conceptual and Procedural Knowledge: The Case of Mathematics*, (pp. 225-64). Hillsdale, N.J: Erlbaum.

Skemp, R. R. (1987). *The psychology of learning mathematics.* Hillsdale, NJ: Erlbaum.

Wikipedia (2005). The free encyclopedia. http://en.wikipedia.org/wiki/Critical_thinking. (downloaded: Dec. 2005)

Zaslavsky, O. & Leikin, R. (2004). Professional development of mathematics teacher-educators: Growth through practice. *Journal of Mathematics Teacher Education,* 7, 5-32.

New and Noteworthy

Handbook of Research on Psychology of Mathematics Education: *Past, Present and Future*
Angel Gutiérrez, *Universidad de Valencia* and **Paolo Boero,** *Università di Genova* **(eds.)**
This volume is a compilation of the research produced by the International Group for the Psychology of Mathematics Education (PME) since its creation, 30 years ago. It has been written to become an essential reference for Mathematics Education research in the coming years. This handbook shall be of interest to both experienced researchers and doctoral students needing detailed synthesis of the advances and future directions of research in Mathematics Education, and also to mathematics teacher trainers who need to have a comprehensive reference as background for their courses on Mathematics Education.
paperback: US$49.00/€45.00 ISBN:90-77874-19-4
hardback: US$147.00/€135.00 ISBN:90-77874-66-6
March 2006, 544 pp

Mathematics Classrooms In Twelve Countries: *The Insider's Perspective*
David Clarke, *Univ. of Melbourne,* **Christine Keitel,** *Freie Univ. Berlin* and **Yoshinori Shimizu,** *Univ. of Tsukuba, Japan* **(eds.)**
This book reports the accounts of researchers investigating the eighth grade mathematics classrooms of teachers in Australia, China, the Czech Republic, Germany, Israel, Japan, Korea, The Philippines, Singapore, South Africa, Sweden and the USA. This combination of countries gives good representation to different European and Asian educational traditions, affluent and less affluent school systems, and mono-cultural and multi-cultural societies. Researchers within each local group focused their analyses on those aspects of practice and meaning most closely aligned with the concerns of the local school system and the theoretical orientation of the researchers. The Learner's Perspective Study is guided by a belief that we need to learn from each other. The resulting chapters offer deeply situated insights into the practices of mathematics classrooms.
paperback: US$49.00/€45.00 ISBN 90-77874-95-X
hardback: US$147.00/€135.00 ISBN 90-77874-99-2
SERIES: THE LEARNER'S PERSPECTIVE STUDY 1
August 2006, 402 pp

Review forthcoming in ZDM-The International Journal on Mathematics Education, vol. 40, 2008

Making Connections: *Comparing Mathematics Classrooms Around The World*
David Clarke, *Univ. of Melbourne,* **Jonas Emanuelsson,** *Gotenborgs Universitet,*
Eva Jablonka, *Freie Universitat Berlin* and **Ida Ah Chee Mok,** *The University of Hong Kong* **(eds.)**
In this book, comparisons are made between the practices of classrooms in a variety of different school systems around the world. The abiding challenge for classroom research is the realization of structure in diversity. The structure in this case takes the form of patterns of participation: regularities in the social practices of mathematics classrooms. The expansion of our field of view to include international rather than just local classrooms increases the diversity and heightens the challenge of the search for structure, while increasing the significance of any structures, once found. In particular, this book reports on the use of 'lesson events' as an entry point for the analysis of lesson structure.
paperback: US$49.00/€45.00 ISBN 90-77874-79-8
hardback: US$147.00/€135.00 ISBN 90-77874-90-9
SERIES: THE LEARNER'S PERSPECTIVE STUDY 2
September 2006, 280 pp

Review forthcoming in ZDM-The International Journal on Mathematics Education, vol. 40, 2008