Situating Mathematical Communication: the Settings, Interactions and Material Practices of Contemporary Mathematical Research

Katherine McCallum

A thesis submitted in partial fulfilment of the requirements of the University of Brighton for the degree of Doctor of Philosophy

January 2020

Abstract

This thesis is an artist's communicative ethnography of mathematics, using multiple methods and outputs to put forward a rich picture of mathematics as inextricably entwined with the minds and materials involved in its development. In this research I take an ethnographic position and approach mathematics as an artist explorer, taking excerpts from recordings of everyday mathematical work and engaging in a process of reflective sense-making with the material, carried out with person, medium and place in mind. The discussion proceeds through a blend of written analysis informed by linguistic pragmatics and the situated cognition paradigm, and creative practice experiments that enact and test the core ideas proposed. A particular goal of the research is to examine the role played by mathematical writing in the variety of different material and social situations that make up mathematical work.

The key claim made is that while mathematics is known for abstraction and pure ideas, and often seems to exist somewhat apart from messy human reality, its practitioners reach these sophisticated cognitive heights through a variety of heavily situated, interactive practices that are not so separate from the improvised, social world of everyday communication. A key perspective on communication is relevance theory, as put forward by Dan Sperber and Deirdre Wilson, which emphasises the inferences that communicators make about one another's minds; just such inferences are seen coming into play in essential ways throughout mathematical communication, which demonstrates that interpersonal interaction has an important role to play in the achievement of on-the-ground mathematical understanding. In addition the situated position on cognition taken by researchers such as Andy Clark and Edwin Hutchins, recognising the part that external resources play in cognition, allows us to understand mathematical writing as an important component in a collective cognitive system. Heavy use is made in mathematics of mark-making practices that are richly embedded in discourse in such a way as to extend the cognitive possibilities open to practitioners, and it is this that makes such incredibly complex ideas tractable. As such, this highly refined interaction between person and representation defines mathematics in an important way.

This thesis is a portrayal of mathematics as built up from interactions between persons and stabilised representations, and so situated and interactive in an essential way. It is also an application of relevance theory in such a way as to test the boundaries both of the theory and of what should be called communication. In its design, it proposes a method of doing artistic research that blends readily with other disciplines and is truly ethnographic while staying true to properly artistic aims.

Contents

Practice Contents

Multimedia materials

3.3 Whiteboard animation of lines 10 and 11. Gif.

3.3 Whiteboard animation of lines 20 – 22. Gif.

3.3 Whiteboard animation of lines 26 – 29. Gif.

5.3.2. Notes. Jpg.

5.3.2. Notes. Part 1. Pdf.

5.3.2. Notes. Part 2. Pdf.

8.2.2 Three operations website. Website folder (open file 0Intro3.html to view the website).

Pdf of thesis

Situating Mathematical Communication. Pdf.

List of figures

Abbreviations and conventions used

Transcription conventions taken from "Transcription Conventions in Conversation Analysis" from the *Handbook of Classroom Discourse and Interaction* (Numarkee, 2015).

D: pseudonym of an identified participant

SIMULTANEOUS UTTERANCES

Dan: [yes
He Hua: [yeh — simultaneous, overlapping talk by two speakers

Dan: [huh? [oh] I see]
He Hua: [what]
Feng Gang: [I don't get it] — simultaneous, overlapping talk by three (or more) speakers

CHARACTERISTICS OF SPEECH DELIVERY

?	rising intonation, not necessarily a question
!	strong emphasis, with falling intonation
yes.	a period indicates falling (final) intonation
so,	a comma indicates low-rising intonation suggesting continuation
descr↑iption↓	an upward arrow denotes marked rising shift in intonation, while a downward arrow denotes a marked falling shift in intonation
go:::d	one or more colons indicate lengthening of the preceding sound; each additional colon represents a lengthening of one beat
no-	a hyphen indicates an abrupt cut-off, with level pitch
because	underlined letters indicates marked stress
SYLVIA	large capitals indicate loud volume
SYLVIA	small capitals indicate intermediate volume
sylvia	lower case indicates normal conversational volume
°sylvia°	degree sign indicates decreased volume, often a whisper
.hhh	in-drawn breaths
hhh	laughter tokens
> the next thing<	>…< indicates speeded up delivery relative to the surrounding talk

< the next thing> <…> indicates slowed down delivery relative to the surrounding talk

COMMENTARY IN THE TRANSCRIPT

((coughs))	verbal description of actions noted in the transcript, including non-verbal actions
((unintelligible))	indicates a stretch of talk that is unintelligible to the analyst
… (radio)	single parentheses indicate unclear or probable item

OTHER TRANSCRIPTION SYMBOLS

co/l/al	slashes indicate phonetic transcription
→	an arrow in transcript draws attention to a particular phenomenon the analyst wishes to discuss
…	ellipsis
]	points of overlapped speech across two turns
::	lengthening of syllable
(( ))	researcher comments or translation
italics	non-English speech
-	short untimed pause
(x)	unclear word
word-	false-start or self-correction

Declaration

I declare that the research contained in this thesis, unless otherwise formally indicated within the text, is the original work of the author. The thesis has not been previously submitted to this or any other university for a degree, and does not incorporate any material already submitted for a degree.

Signed

Kate McCallum

Dated 31/01/2020

Acknowledgements

My heartfelt thanks go to my supervisors Dr. Ole Hagen and Dr. Andrew Fish, who supported and guided me through a somewhat unconventional PhD with openness and enthusiasm. I also want to express my deep gratitude to my lead supervisor Dr. Tim Wharton; it was his adventurousness, boundless enthusiasm and positivity that made such a project feel possible from the beginning, and he guided and supported me through a great many twists and turns, small and large, over the course of the PhD, providing help and encouragement in enormous, innumerable ways. Without such wonderful support it is hard to imagine how this thesis could have existed.

I would also like to thank a number of people who encouraged and helped me along the way: Tom Grimsey, the wonderfully inspiring artist without whom I may never have contemplated doing a PhD; Prof. Anne Boddington, then Dean of the University of Brighton, who gave me a start in postgraduate research; my wonderful cohort at Brighton, who provided support, fun and joy at throughout my PhD process; my loving family, who have always given me the support and motivation to be all that I can be; and my wonderful partner Dr. Ian O'Loughlin, who has endured first drafts, art projects, and half-baked ideas, given me discussions, apposite critiques, and inspiration, and done it all with patience and love.

0. Introduction

The fundamentals of mathematics give such an impression of existing as a timeless body of facts that there is almost a sense of incongruity with what goes on in the offices of real-world mathematicians working today: messy practices of discussing and scribbling, whiteboards and stacks of paper. In 'Mathematics has a Front and a Back,' Reuben Hersh describes certain 'myths' of 'unity, universality, objectivity, and certainty' (Hersh, 1991 p.131) commonly ascribed to mathematics, and puts forward an explanation of this occurrence by using a dramaturgical analogy to talk about the contrast between the polished, published 'front' of mathematics and the messier work done 'backstage':

> In this sense of the term, the 'front' mathematics is mathematics in 'finished' form, as it is presented to the public in classrooms, textbooks, and journals. The 'back' would be mathematics as it appears among working mathematicians, in informal settings, told to one another in an office behind closed doors.
>
> Compared to 'backstage' mathematics, 'front' mathematics is formal, precise, ordered and abstract. It is separated clearly into definitions, theorems, and remarks. To every question there is an answer, or at least, a conspicuous label: 'open question'. The goal is stated at the beginning of each chapter, and attained at the end.
>
> Compared to 'front' mathematics, mathematics 'in back' is fragmentary, informal, intuitive, tentative. We try this or that, we say 'maybe' or 'it looks like'. (Hersh, 1991 p.128)

Hersh's eventual claim is that '[i]f mathematics were presented in the same style in which it is created, few would believe in its universality, unity, certainty, or objectivity' (Hersh, 1991 pp.130–131); but the 'backstage' side of the story, he says, is obscured in the course of the concoction of the performance, the ideal version put together for publication. Indeed, while it is clear that new mathematics must be somehow being *done* by people somewhere for the field to be advancing, there seems to exist a certain temptation even by insiders[1] to view such practices as somehow 'just' symptomatic of our more imperfect engagements with mathematics, preserving the picture of the

---

[1] Hersh's essay claims only to be describing the way that *outsiders* are taken in by the 'front', stating that '[o]ne of the unwritten criteria separating the professional from the amateur, the insider from the outsider, is that the outsiders are taken in (deceived), the insiders are not taken in' (Hersh, 1991 p.129). Hersh's inclusion of journals alongside textbooks and classrooms in the 'front' (Hersh, 1991 p.128) immediately calls the coherence of these various divisions into question, since journal articles in mathematics are highly unlikely to be read, or indeed readable, by true outsiders, and all but the higher levels of students of mathematics—in which case, exactly how correct is it to call this group 'outsiders'? In addition, it is hardly uncommon for philosophising insiders to wax lyrical about qualities such as unity, universality, objectivity, and certainty when talking about mathematics in the abstract. A helpful amendment might be to say that a person's picture of mathematics is complicated by awareness of a 'backstage' depending on the extent to which that person is an insider; which side a person chooses to emphasise might depend on to whom they are speaking and on the purpose of the conversation. Where this and other dividing lines should be drawn are among the topics called into question in Christian Greiffenhagen and Wes Sharrock's ethnographic examination of Hersh's 'front' and 'back' (Greiffenhagen & Sharrock, 2011), to be discussed in the next chapter.

latter as a 'beautiful elsewhere' (Bourguignon & Casse, 2012) accessed, if at all, best by private meditation.

> 'Mathematicians do a lot of work in the shower or walking from a to b. They give a kind of 'not in this world' kind of impression, and I think it's true.' Prof. Guy Nason, interviewed for a video series on mathematical discovery (University of Bristol, n.d. 1'59")

> 'People can be thinking about mathematics all the time if they want to. They can carry it all around in their heads…' Prof. Elmer Rees, interviewed for a video series on mathematical discovery (University of Bristol, n.d. 2'34")

In general, it seems that the material practices of mathematics are not valued as a part of its essential nature, perhaps even habitually effaced from it by its practitioners, as one ethnography tentatively suggests: 'mathematical writing seems built to pass away, leaving something it is tempting to call "mathematics, itself" as its residue' (Barany, 2010 p.52). Traditional philosophy of mathematics has correspondingly exhibited a tendency to pay the greatest attention to foundational questions about the theoretical ideal of mathematics, a habit noted and challenged by a variety of researchers toward the end of the 20th Century (Lakatos, 1976; Kitcher, 1984; Asprey & Kitcher, 1988; Tymoczko, 1998).

An interesting question is what we might learn were we to reject the decision to downplay the importance of these everyday practices in the 'back', and instead chose to carefully observe them in order to understand how mathematical work is achieved. What is it that produces this sense of autonomy from human doings in the 'front', and where does this 'front' really begin? If an observer approaches on-the-ground mathematical work in its most everyday form, what can we learn about how mathematics is developed and what it consists of? What are we missing if we ignore the messy, contingent methods habitually used to understand and be understood in the course of bringing these ideas into being? Is mathematics really so different from other human activities?

An aim of this research is to cast mathematics in a new light, particularly one that will be meaningful to a lay reader. Paying attention to these messy human aspects of doing mathematics might be a means to lessen the perceived distance between mathematics and other aspects of human endeavour, to understand its methods and aims as somehow of a piece with those of our everyday lives and tasks.

In the last two decades, the school of philosophy of mathematical practice encourages serious attention to how mathematics is actually done. This research programme places focus on mathematical practices, such as proving, explaining and becoming convinced, even those that exist beyond the kind of rigour required of formal proofs. '[An alternative picture of mathematics] does not have at its core formal proofs, but rather the practices that are shared by a community of scholars who in their ordinary work consider informal proofs, such as proofs based on induction or visual tools, as sufficient for being justified in believing that some statement holds.' (Giardino, 2013 p.137). The methods tend to be theoretical, analysing some well-known practice or component of mathematical work, examined as case studies.

This research impetus is a substantial move in the direction of answering the questions above, and yet remains at a level or two of abstraction away from those paper-filled offices. I will approach the question of why that might be a problem by quoting the following from Bruno Latour, known by many for his promotion of ethnographic study of scientific practices, and here issuing a warning about the difficulty of pursuing a similar project in the case of mathematics:

> It is easy to study laboratory practices because they are so heavily equipped, so evidently collective, so obviously material, so clearly situated in specific times and spaces, so hesitant and costly. But the same is not true of mathematical practices: notions such as 'demonstration', 'modelling', 'proving', 'calculating', 'formalism', 'abstraction' resist being shifted from the role of indisputable resources to that of inspectable and accountable topics. (Latour, 2008: 444)

The range of ethnographic projects that have produced relevant results seems to belie the claim that such practices cannot be directly inspected at all (Livingston, 1989; Livingston, 1999; Livingston, 2006; Livingston, 2015; Merz & Cetina, 1997; Greiffenhagen, 2008; Greiffenhagen & Sharrock, 2011; Greiffenhagen, 2014; Barany, 2010; Barany & MacKenzie, 2014; Lane, 2016). It is easier to justify the claim that practices described in these sort of abstracted terms are defined in abstracted terms that already sign on to some kind of ideological commitment, in comparison with, for example, a physical action that need not be named in the traditional way to be examined and discussed. This for Latour was surely an important point since Latour and Woolgar's work in the laboratory emphasised 'seeing' practices according to alternative logics (Latour & Woolgar, 2013). In this way Latour seems, as many do, to overplay the private aspects of mathematical reasoning.

But must it be so? Were Latour a fly on the wall in a mathematician's office, watching two colleagues work through a problem, he might see practices of testing, reading-off, exemplification and adjustment playing out in real time in a way that were inspectable enough to satisfy him. The ethnographies listed above make quite different decisions when it comes to selecting the data to be used; Eric Livingston's auto-ethnography of reasoning, Lorenzo Lane and Michael Barany's interviews and observations of habits, and Christian Greiffenhagen's fine-grained analyses of blackboard use each allow a kind of observation of the way that mathematics progresses, whether from within or without. Each to some extent recommends the recognition of mathematical practices as an essential and undervalued area of study in understanding mathematics, and yet the actions of a particular mathematician going about daily business are most closely exhibited in studies like Greiffenhagen's video analysis of an undergraduate lecture, presenting the reader with direct evidence of organisation of thought by showing a set of material practices (Greiffenhagen, 2008 p.525). It is a challenge to provide this kind of direct evidence in studies of mathematical genesis, not least because, as Barany observes, 'mathematics is difficult' (Barany, 2010 p.7), a world of highly specialised subfields that is difficult to grasp and synthesise even for an insider. Perhaps this difficulty can be somewhat overcome if an ethnographic project makes cautious decisions about scope and scale.

If mathematics is to be observed in the wild, it is at its most observable when its practitioners are interacting either with one another or with some kind of medium of representation. When discussing emerging work at a research meeting, or scribbling notes on a sheet of paper alone at a desk, a mathematician's work is externalised. While this would seem to place particular restrictions on the scope of observational ethnography, there are those within the field making the case that both writing and interaction with colleagues sit very close to the heart of mathematics:

> Interviewer: what does it look like in your head, when you're doing mathematics?
> Terence Tao: um… [long pause] It's always a combination of working in your head, and speaking out loud, and working on the board… (Numberphile, n.d. 6'50")

> N: …I like to discuss maths with others, but at the level that is of motivation and general interests, questions. But when I want to think really deep, I need a pen… (Lane, 2016 p.179)

> [N:] So one important thing in writing is to help the mind to concentrate. Sometimes it's not important if it's right. I explained this to a collaborator once: that I just need to write. It's not strictly useful to write down, but it helps to think. (Lane, 2016 pp.220–1).

> G: I was mentioning these notebooks, they are with me essentially always, and sometimes I don't touch them for a week because nothing happens but they are with me anyhow, where ever I am. (Lane, 2016 p.218)

> Interviewer: When you shut your eyes, do you see something mathematical?
> No. […] I need not only open eyes, but paper and blackboard before I can think sensibly about anything mathematical. Don Zagier, interviewed in (Bourguignon & Casse, 2012 p.88)

> Knowledge in mathematics is embodied. Individuals act as repositories of knowledge, and many researchers prefer to ask a question directly to an expert, in order to find out more about a problem or idea, rather than consulting written material. (Lane, 2016 p.179)

> First of all, the proof of a theorem is a message. A proof is not a beautiful abstract object with an independent existence. No mathematician grasps a proof, sits back, and sighs happily at the knowledge that he can now be certain of the truth of his theorem. He runs out into the hall and looks for someone to listen to it. He bursts into a colleague's office and commandeers the blackboard. He throws aside his scheduled topic and regales a seminar with his new idea. He drags his graduate students away from their dissertations to listen. He gets onto the phone and tells his colleagues in Texas and Toronto. In its first incarnation, a proof is a spoken message, or at most a sketch on a chalkboard or a paper napkin. (De Millo et al., 1979 p.273)

In any case, the desirability of direct, in-the-wild observation suggests two things: that data on mathematicians' practices will come from situations of communication and of writing, and for this reason, that questions about how communication works and how minds work with representations will turn out to be pertinent to better understanding how mathematics works.

A second problem is that there has been a habit, even in ethnographies that intentionally foreground mathematical practices, to default eventually to essentially separating the messy world of the practices from some overarching reality of the mathematics (as in Greiffenhagen's description of the proof as reflected, not embodied, in its written form (Greiffenhagen, 2008 p.525), Barany's hesitant allusion to 'mathematics, itself' (Barany, 2010 p.52), Lane's attempt to separate the means of mathematics from 'the process of discovery, creativity, and assembly' (Lane, 2016 p.216), and so on). It is to be noted that each of them adopts a relatively traditional style of analysis that echoes this separation.

This thesis therefore faces a variety of challenges. How to stay close to the on-the-ground reality of mathematical practices? How, in spite of this, to make them comprehensible to the outsider observer? How to escape the ideology of mathematics, subjecting these practices to alternative logic? This thesis responds to those challenges by employing a variety of transgressive means of doing research to place inspectable data at the centre of the discussion, to take steps to unfold its content for a lay reader, and to do all it can to make unfamiliar the familiar. In this way, mathematics is illuminated from a different direction from usual, a new focused light cast on the subtle communicative and collaborative practices essential to its progress which are explored in a multifaceted blend of written analysis and practical experiment. In this way, a reader is presented with a multimedia means of coming to see mathematics in a new way that brings multiple ways of knowing, as well as theoretical perspectives from multiple disciplines, to bear in a subtle but important re-seeing of this intriguing corner of human endeavour.

The result is an artist's communicative ethnography of mathematics, using multiple methods and outputs to put forward a rich picture of mathematics as inextricably entwined with the minds and materials involved in its development. The key claim made is that while mathematics is known for abstraction and pure ideas, and often seems to exist somewhat apart from messy human reality, its practitioners reach these sophisticated cognitive heights through a variety of heavily situated, interactive practices that are not so separate from the improvised, social world of everyday communication. The mark-making practices seen in the mathematical world are richly embedded in discourse in such a way as to extend the cognitive possibilities open to practitioners, and it is this that makes such incredibly complex ideas tractable and perhaps even that makes mathematics unique.

As a motion toward the aforementioned inspectability of data, the chapters of this thesis will be interspersed with excerpts from the materials I gathered from mathematician participants, from interviews and elsewhere, that touch on the themes addressed in the main text. In this way, by showing these themes as addressed in the voices of working mathematicians, I hope to anchor the thesis as much as possible to the realities of their day-to-day work. Between each chapter you will find an Interlude, with materials from interviews on a particular topic. In the nine Interludes you will also find my photographs documenting the work sites of each of my nine participants, their offices, desks and institutional surroundings. In addition, the text of the thesis uses an intricate numbering structure that references that used in mathematical papers, exemplifying the kind of labelling and signposting used to organise thought in mathematical texts, albeit recontextualised.

The thesis begins with an outline on the key theoretical perspectives that inform its design, followed by an outline of the methods used in the course of the research. Then I present three chapters that analyse three excerpts from my collective data, and tell the story of the processes of performative sense-making that I went through to make sense of these excerpts from the perspective of a layperson. Next are three chapters of discussion that proceed through written analysis and practical experiment to make the key arguments of the thesis, followed by a conclusion.

Interlude 1: Writing and thinking

Interview with Subject G

00.44.41.536

G: The majority of the time I'm doing research I'm just and typing into LaTeX

OK so you just go straight into the…

G: Yeah. Straight in, just start typing […] then periodically if I'm stuck on some things I've got my pad- My pad! <holds up pad of paper with a few notes on> and pen! Then I'll scribble down a few things here, and once I've figured it out […] then I'll type up the calculation I did and keep going.

The majority of the time I'm writing and typing into LaTeX, and that's because, yeah, a lot of the figuring out arguments and what needs to be done in a paper I'll just do in my head, when I'm walking around or something. Or if I'm like at home, or I'm sleeping. Or I'll be sitting on the train in the morning to the office and I'll be thinking about mathematics. I'll be like, I'm stuck at this one point what do I do here and I'll be like, OK, that's how I'll work it out, when I get into the office I'll write it up.

00.54.18.112

So I'm interested in the role of writing and writing up, and you were talking about walking around through the world and thinking through stuff and coming in and writing up. Do you find that that process goes fairly straightforwardly?

G: No, I mean, it's never straightforward. Mathematics research isn't easy. […] But no, I mean, normally I'll have an idea and I'll start writing it up, but then I'll realise maybe later what I wrote didn't make any sense at all, it's just gibberish <laughs>

Especially when you're walking around thinking your thought processes can be quite naïve, and it's not 'til you sit down and start writing it that you properly work out what's going on. You can visualise it better if it's written down, I guess. Especially if you're doing some quite technical calculation, they're difficult. […] I can see basi- see my way through it but then I have to sit down and start writing it to work out the remaining details. And scribble a bit on my piece of paper <taps paper>

Work site 1

Work site 1 was the office of a lecturer at a technical university in continental Europe.

Figure 1. Communal spaces at the university, close to my subject's office. Original in colour.

Figure 2. Corridor space outside my subject's office. Original in colour.

Figure 3. Mathematical texts and geometric paper constructions on the windowsill of my subject's office.

Figure 4. Workspace in my subject's office, including desk with computer, whiteboard and chair for visitors. Original in colour.

1. Theoretical background

1.1. Study of an Ethnography

In the interests of situating the approach taken in this research, I will briefly discuss an existing ethnographic study that explicitly addresses Hersh's 'front' and 'back' of mathematics. This will help to clarify the problem space and the particular challenges that this thesis will aim to overcome. In 'Does mathematics look certain in the front, but fallible in the back?', Christian Greiffenhagen and Wes Sharrock carefully examine the notion of a 'front' and 'back' and test it by closely examining ethnographic data from recordings of practising mathematicians. One thing this study does is to complicate Hersh's picture in much-needed ways. Their conclusions can be summarised as follows: firstly, that the divide is not so strong as all that, the 'front' actually does reference the existence of the 'back', and the work in the 'back' holds the 'front' as a kind of aim; secondly, that the existence of some kind of divide really isn't so strange, since a proof is a product, not a description, of a process of discovery.

In this study, the researchers stick close to ethnographic data to exhibit evidence that has a bearing upon quite broad questions about mathematics. The mathematical content is, as is often the case, elided in favour of broad-strokes summaries, but the reader is presented with images of the participants at work in context (see Figure 5), which does a lot to ground the written discussion in their real-world practices.

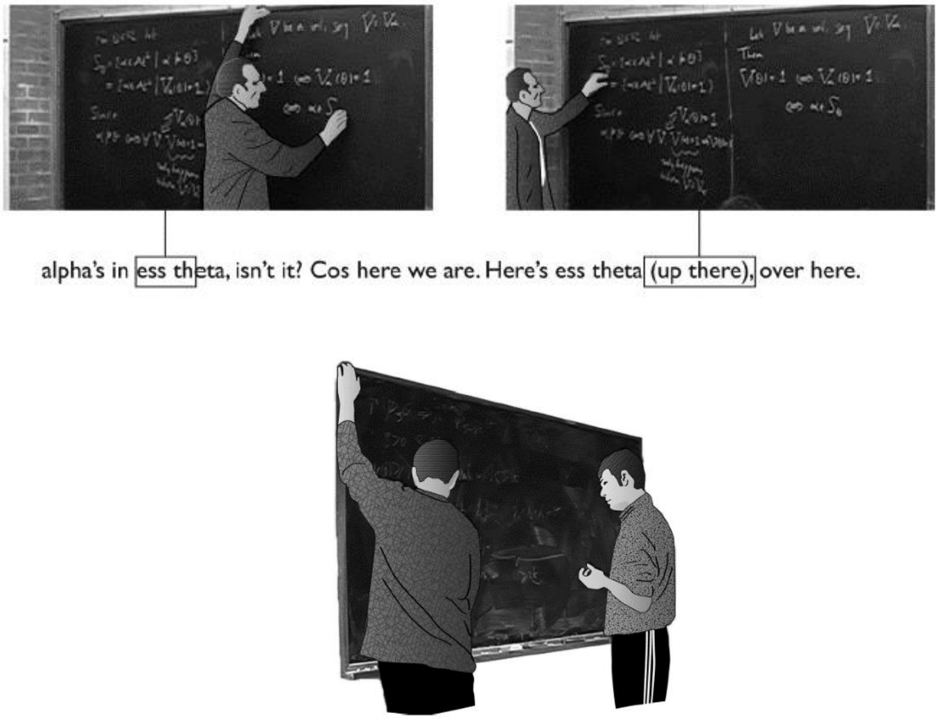


Figure 5. Figures 4 and 6 from Greiffenhagen and Sharrock (2011), showing participants' positioning and use of board notes

Greiffenhagen and Sharrock proceed by analysing a series of lectures, as satisfying Hersh's description of the 'front', and supervisory meetings of doctoral students for the 'back', and examining how comfortably each sits within its category, essentially concluding that rather than there being a firm divide in fact each is acknowledged and recognised in the other. For example, they describe how a lecturer guides students through a proof with commentary on its trickiness or unusualness, saying:

> When we look in detail at the way the lecturer presents the proof, we see that, although the written proof on the page or blackboard may have a 'formal, precise, ordered and abstract' character, the lecturer does not argue that understanding the proof is 'mechanical' or 'automatic'. […] although the lecturer presents a formal, deductive, abstract proof, he is not portraying mathematics (or mathematical reasoning) as monotonously formal, deductive, or abstract. (Greiffenhagen & Sharrock, 2011 p.850).

The authors thus suggest that the lecturer presents an ordered proof but recognises that students' understandings of that proof will be subjective, tentative, and so on. This finding already enriches our picture of the way mathematics is shared and yet leaves open the possibility that the battle lines simply ought to be redrawn a little further out, with the written proof perhaps becoming our new 'front', and the lecture acting as a kind of intermediary. In many ways finding reference to subjective experience in the face-to-face, intentionally pedagogical form of the lecture makes a kind of sense, where it might be even more surprising to find these kinds of traces in the rigorous form of a proof itself.

The big challenge that this paper levels at Hersh is to say that there is nothing strange about the difference between front and back, since '[t]he idea that the form of a finalised proof provides a description of the mathematical reasoning that produced it is about as wise as thinking that Hollywood blockbusters ought to be viewed as documentaries about their own production' (Greiffenhagen & Sharrock, 2011 p.858). The proof, they argue, is a polished product, and indeed should be different from the means used to reach it. This seems a very reasonable argument, and yet there is an assumption underpinning it that by seeking to recognise the 'fragmentary, informal, intuitive, tentative' in mathematics Hersh can only be referring to the history of mis-steps and guesses from its fallible practitioners in the course of producing the final correct (infallible, perhaps?) product. For an example, take this passage:

> The idea that the public mathematical record hides the fallibility of mathematics is based upon too narrow a consideration, concentrating only on the format of the single mathematical paper and focussing, for example, only on the fact that the individual paper setting out a theorem and proof usually does not also describe the difficulties and failings that may have featured in their creation. Such a 'failure' to describe should not be taken to suggest that there were no difficulties. (Greiffenhagen & Sharrock, 2011 p.860)

This is probably a reasonable reading of Hersh, who makes some gestures in this kind of direction and otherwise does not elaborate much on where exactly this tentativeness is to be found. But another reading is possible. Rather than recognising those qualities by looking at the history of a process of discovery and including the mis-steps, what if these qualities are still there when everything is going right? Hersh's call could be seen as something subtler: to recognise the basic impulses of mathematics, the reasoning processes at its heart, as themselves tentative, contingent, fragmentary and so on (not just the practitioners in their fallibility), in a way that is not *external to* good mathematical reasoning, and should surely continue to characterise it in some way even as its level of refinement increases. Otherwise, we fall once again into the trap of assuming that the uncertainty happens in the space before the 'real mathematics', the correct set of answers, is found. Ultimately the analysis of the research meeting is focused on broad-strokes characterisations of stages in problem-solving, and as such may be limited to this level of answer. If we are to find these qualities right at the heart of good mathematical reasoning then perhaps we need to look even more closely at the nitty-gritty of mathematical reasoning in progress, paying close attention to the cognitive side of what is going on; we also need to do yet more to break out of those tenacious assumptions about where the true substance of mathematics lies.

In any case, the other thing that we can note from this discussion is that if we are to see the mathematical proof as manifested in published papers as somehow a shift in register compared with the process of discovery, a product that is in some way characterised by and yet somehow set apart from on-the-ground mathematical reasoning, then understanding the nature of that shift seems an important aim. If the graduate student expects to work toward this kind of written 'front' presentation, and it is *not* simply a misrepresentation of a real process, then what is it? It makes sense to ask what it is that this earlier, exploratory work is working toward, and why. Returning to Figure 5, I note that the archetypal image that mathematics brings to mind is a person or persons writing, often working away at some notation on a blackboard. Mathematics is particularly known for its writing, particularly notations; in the Latourian tradition it is common to term the writing practices of a discipline 'inscriptions', and examine their forms and roles. When it comes to the highly specialised notations of mathematical writing, an interesting question is how this relates to and diverges from other kinds of writing (which we often expect to be communicative in nature) as it moves through different levels of formality.

From examining this study, we can conclude that grounded, real-world observation can flesh out answers to really broad questions about the nature of mathematics in unexpected directions. We also saw that more can be done to adhere closely to that data, to make available rather than summarise the mathematical content, to bring multiple media in to play, to choose moments of fine-grained idea development and dig deeply into exactly how mathematicians are reasoning and communicating to advance their work. It is also clear that we need to break out of established ways of interpreting what we see, and that this task will not be easy; indeed our best hope may be to adopt a truly experimental means of doing research. We can add these aims to those listed in the introduction:

- To carry out direct observation of mathematical practices in the wild, approaching the data in a way that embraces its multimodality
- To present an analysis of material from this difficult field in such a way as to welcome a lay reader in to understand the mathematical content, in spite of its challenging nature, while also making observations that are relevant to the contemporary work of mathematicians
- To subject these practices to interpretation according to alternative logics and ideologies, making unfamiliar the familiar

At this stage it makes sense to state some research questions:

1.1.1. Research questions

1. What role does mathematical writing play across different situations of communication and reasoning, and what are the cognitive and communicative forces at work in shaping that writing?
2. What can ethnographic data show us about how mathematics advances through private and collaborative reasoning, and the resources used in the course of that reasoning?

3. How do mathematicians reach consensus, and what is involved in gaining mathematical expertise?
4. What kind of research approach and presentation will give a reader access to the observed data, including the mathematical content, examine it in a truly multimodal way, and prioritise the aim of breaking out of tenacious ideological assumptions and subjecting the data to alternative ways of seeing?

A few words about research design: it is becoming clear that if we want to know more about how mathematics is *done*, and want to get answers from ethnographic data, then the question of how mathematicians in the real world communicate and come to understand one another (and indeed where the boundaries of communication lie, as we consider different forms of writing) is going to be central. In Chapter 6, I make the case for taking a particular theoretical perspective on communication in order to adequately explain the kinds of phenomena that are visible in the observed data. As a precursor to this discussion, in Lemma 1.2. Perspectives on communication I will outline some of the debates in the naturalistic study of communication, in order to situate the reader's consideration of the material in Chapters 3-5. The particular question of the status of mathematical writing, as a material representation, will be helpfully informed by a research programme that recognises the part that external resources play in cognition, and I will briefly introduce this in Lemma 1.3. The final methodological question is one that I respond to by adopting an experimental framework with a basis in artistic research. In the following sections, I set out some of the contexts that inform this research design and support the methodology that I have adopted.

Lemma 1.2. Perspectives on communication

My research questions revolve around themes of mutual understanding, consensus and collaboration, and it is clear that the question of how communication works in mathematics will be key. As such it is worth summarising some key perspectives on communication that will be useful to us and that will be considered in the subsequent discussion.

The code model of communication is an idea that dates back to Aristotle: that communication is essentially a case of taking a thought and encoding it in some transmissible format, which can then be decoded by a receiver. Today it is commonly described in terms that were originally delineated for information transmission by Shannon and Weaver (1949), summarised in Figure 6.

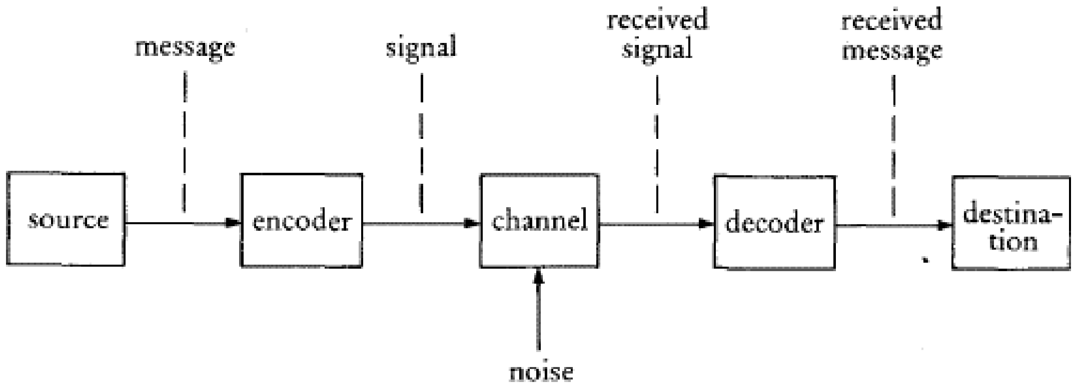


Figure 6. A widely quoted diagram of Shannon and Weaver (1949), shown in the slightly adapted form used in Sperber and Wilson (1986/1995)

The code model, then, describes communication as a linear process in which a thought is encoded, transmitted, received, and straightforwardly decoded using certain rules shared by speaker and hearer[2] to retrieve the original thought. The difficulty with this model is that often in practice, simple decoding is insufficient to explain how a 'signal' is successfully interpreted; real-world 'signals' tend to be packed with basic ambiguities that, in context, pose no problem for any human communicator to resolve, but the code model struggles when it comes to explaining how this resolution is found. For example, consider the following line:

(1) Dehn solved one of Hilbert's problems

From this, it might be possible to reach multiple valid interpretations:

(2) Dehn solved a problem in Hilbert's personal or professional life

(3) Dehn solved one of Hilbert's famously published list of 23 mathematical problems

Which of (2) and (3) would be selected as an interpretation by the hearer would be highly dependent on context, and could not be derived from the code alone (Sperber & Wilson, 1986/1995 p.13); however, in context, reaching the correct interpretation would be near-effortless. This is a key criticism of the code model: that it has no way of explaining the gap between what is said and what is meant (the interpretation that we would all quickly and effortlessly reach), other than proposing yet more codes. In Chapter 6, I will discuss this kind of problem in more detail using examples from Chapters 3-5.

An alternative perspective to the code model is some kind of *inferential* model, of the sort proposed by Paul Grice, which holds that communicators provide one another with not a fully encoded

---

[2] Throughout, reference will be made to speakers and hearers. This is simply to refer to the person communicating and the person being communicated to in a particular situation, but as we shall see, these may as well be writer and reader, diagram-drawer and audience, and so on.

thought but instead a kind of evidence, from which they are expected to make certain inferences about what the other was intending to convey. Grice's proposal here is that to *mean* something, in communicative terms, might simply *be* to have an intention to produce a certain effect in another person, which works because that intention is recognised: "[S] meant something by x" is (roughly) equivalent to "[S] intended the utterance of x to produce some effect in an audience by means of the recognition of this intention" (Grice, 1957/1971 p.58), where an 'utterance' is some kind of stimulus intended to be recognised by an interlocutor. This portrays the process by which we understand one another's communications as a rather more active, intelligent process than simple decoding, guided by our ability to recognise one another as thinking, intending beings.

The question of how these processes of inference might happen, and what kinds of principles might guide them, is crucial if we are to believe in this kind of explanation. Through subsequent publications and a famous series of lectures, Grice proposed that communicators adhere to a kind of principle of co-operation, which shapes the way that communicative stimuli are produced and interpreted (Grice, 1967; Grice, 1975). He developed this basic idea into nine *maxims,* which he classified in four categories:

> **Maxims of quantity**
> 1. Make your contribution as informative as is required (for the current purposes of the exchange).
> 2. Do not make your contribution more informative than is required.
>
> **Maxims of quality**
> *Supermaxim: Try to make your contribution one that is true.*
> 1. Do not say what you believe to be false.
> 2. Do not say that for which you lack adequate evidence.
>
> **Maxim of relation**
> Be relevant.
>
> **Maxims of manner**
> *Supermaxim: Be perspicuous.*
> 1 Avoid obscurity of expression.
> 2 Avoid ambiguity.
> 3 Be brief (avoid unnecessary prolixity).
> 4 Be orderly.

These maxims, Grice proposed, were a description of the basic principles that communicators follow, ones that might guide a hearer in interpreting an utterance in the way intended by the speaker.

Dan Sperber and Deirdre Wilson took these maxims as a point of departure and developed a somewhat more streamlined theory of communication on the basis of just one of them, from which, they argue, the others can be derived. This was the maxim of relation: to be *relevant*, a principle that they argue shapes all of communication, in a broad theory of naturalistic communication known as relevance theory. Relevance theory is based on two main principles: the *cognitive principle of relevance,* which states that human cognition is basically geared toward looking for relevant information, and the *communicative principle of relevance,* which proposes that communicators exploit this fact, shaping their communications to be relevant and interpreting them accordingly. Relevance, as used here, has a particular technical definition, given as a relation of cognitive effort and cognitive effects: an utterance produced by a speaker is expected by a hearer to be *optimally relevant,* in the sense that there is an expectation of a certain degree of useful cognitive effects, as balanced against the cognitive effort needed to interpret the utterance. This principle can guide a hearer's expectations about how much effort to invest, and what kind of effects to expect. For example, prolixity takes greater effort to interpret than more succinct communication, as might unnecessary obscurity or ambiguity, and so if extra cognitive effects are not to be found this would constitute a failure to be optimally relevant; what makes this useful is that prolixity, obscurity and ambiguity can be indicators that additional interpretive effort might be required, if different or more sophisticated effects are to be found. The principle is summarised as follows:

> (1) Relevance of an input to an individual
> a. Other things being equal, the greater the positive cognitive effects achieved by processing an input, the greater the relevance of the input to the individual at that time.
> b. Other things being equal, the greater the processing effort expended, the lower the relevance of the input to the individual at that time. (Wilson & Sperber, 2005 p.252)

This is a principle in a descriptive rather than hortative sense; relevance theory sets out not to offer a guide for behaviour but rather to describe how communicators already do behave. Exactly how to evaluate cognitive effort and effects from the outside is something that Sperber and Wilson own is difficult to pin down, a challenge faced in many areas of psychology; they point out that relevance theory need only 'describe how the mind assesses its own achievements and efforts from the inside' (Sperber & Wilson, 1986/1995 p.130). Broadly speaking, cognitive effects are evaluated in terms of the assumptions that a hearer is caused to strengthen, weaken or adopt. Some assumptions are more accessible than others; an assumption might be more accessible the more frequently accessed or salient it is. In *Relevance Theory*, an introduction to the theory with certain useful additions and interpretations, Billy Clark proposes a broader list of factors that processing effort might depend on:

1. Recency of use
2. Frequency of use
3. Perceptual salience
4. Ease of retrieval from memory

5. Linguistic/logical complexity
6. Size of the context (Clark, 2013 p.104)

A reader might notice a certain focus on cognition in the language used to describe this principle, and indeed relevance theory is considered a cognitivist theory of communication, described in terms of certain ideas about the way our minds work. It is to be noted that Sperber and Wilson consider their theory a felicitous amalgamation of code and inferential perspectives, since they recognise an element of decoding in a hearer's interpretation of the explicit content of an utterance. This theory was introduced as an *ostensive-inferential* explanation: one stated in terms of *ostension*, which is an act of deliberately and openly pointing out or exhibiting something, making clear that it is intentional, and *inference*, reaching conclusions on the basis of evidence and reasoning.

The strength of theories of communication that include inference is that their ability to make sense of the way communication works when it is impressionistic, ambiguous and subtle. Examination of ethnographic data reveals that this is very often the case, but it would be reasonable to suppose that in the well-defined, precise world of mathematical communication, a code model might be quite sufficient to explain how practitioners come to reach shared understandings of mathematical ideas. Having a sense of how mutual understanding is achieved is crucial to answering the stated research questions about how consensus is reached, how collaborative reasoning works and what role is played by writing in different communicative situations. As such this is a question to keep in mind through the data analysis, and will be revisited in Chapter 6.

Lemma 1.3. The situated cognition paradigm

I will here briefly introduce a second debate that will be pertinent in considering the particular question of mathematical writing, and the way that it relates to thought.

The idea that mind and body are essentially separate has been around a long time—again, dating back to Aristotle and beyond, and often associated with the notion of mortal body and immortal soul—but is now strongly associated with René Descartes, who famously expounded on the mind as a non-physical substance. In his *Meditations*, Descartes defined the mind/soul as *thinking*, and the body as matter, and *unthinking* (Descartes, 1641/1984). This is the view that mental phenomena are non-physical, and thus that the mind and body are of quite different kinds and so separable. This view is particularly seductive because it is quite intuitive to say that intra-mental states are felt differently from sensory phenomena, and so mental and physical phenomena are experienced in quite different ways, and it has deeply shaped the way that minds and bodies have been treated for centuries. In particular, it casts cognition as a purely internal phenomenon.

A challenge to this assumption has arisen and gained momentum in the situated cognition paradigm, a research programme of the last few decades that views cognition as importantly comprised of a whole person's actions in and interactions with the world. Proponents include Andy Clark, Edwin Hutchins and Alva Noë, the latter of whom summarised his position in particularly succinct form: 'to understand consciousness in human and animals, we must look not inward, into the recesses of our insides; rather, we need to look to the ways in which each of us, as a whole animal, carries on the processes of living in and with and in response to the world around us' (Noë, 2009 p.7). The research programme is known as dealing with the mind as embodied, embedded, enactive and extended, from which come the terms '4E' or 'E-turn', and the terms 'situated' or 'grounded' are often used. I will default to the simple term 'situated cognition' to refer to the broad movement.

This is an important perspective to take into account when it comes to considering the relationship between writing and thinking. In *The Extended Mind* (1998) Andy Clark and David Chalmers take the example of Inga and Otto, the latter a patient of Alzheimer's disease, both of whom want to visit a museum and need to know the location. Inga remembers the address of the museum, whereas Otto keeps notebooks with him at all times which serve as a kind of extended memory. The kind of move made by situated cognition thinkers is this: to see these external strategies not as essentially separate from the work of the mind, but as importantly bound up with and shaping of cognition as we know it (since human cognition has, after all, developed in a shapeable and interactive environment). Writing has revolutionised what humans have been able to achieve cognitively since its invention, and thus it seems reasonable to take this kind of technology seriously in its role as a component or constituent of cognition. With this in mind it is possible to look at something like mathematical writing and understand it as genuinely constitutive of the shaping and development of mathematical thought, and see practitioners' interactions with writing as part and parcel of thinking itself.

Person-environment interactions are thus seen as nuanced and intelligent negotiations, and we can begin to see even phenomena like perception as importantly characterised by action in an environment rather than simple absorption of information. An internalist, representationalist view might see cognition as a process of taking in perceptual information, encoding it in some internal representation, reasoning about it, and then translating this into action, whereas situated cognition thinkers talk about active, responsive perception that is informed by and feeds in to expected action. To see what this shift looks like in practice, it is worth quoting Alva Noë at length.

> Suppose you hold a bottle in your hands with your eyes shut. You feel it. You have the feeling of the presence of the whole bottle even though you only make finger-to-bottle contact at a few points. The standard account of this phenomenon proposes that the brain takes the little information it receives (at the isolated points of contact) and uses it to build up an internal model of the bottle (one capable of supporting the experience).

> But consider: this positing of a process of construction of an internal representation may be an unnecessary shuffle. For the bottle is right there, in your hands, to be probed as occasion arises. Why should the brain build models of the environment if the environment is present and so can serve as its own model, as an external but accessible repository for information (as has been argued by Brooks, 1991; O'Regan, 1992)?
>
> …
>
> The upshot of this discussion is that perceptual experience, in whatever sensory modality, is a temporally extended process of exploration of the environment on the part of an embodied animal. (Noë, 2000 pp.127–8)

As in relevance theory, where simple decoding is replaced by an interpretive process that is active, contextual and based on speaker and hearer's mutual awareness, interface with a material resource here becomes a case not of simply taking in information but of interacting, responding, and coming to a nuanced awareness of which an external resource is a part.

Situated understandings of cognition range from the view that the external world is a part of the thinking mind itself, as in Clark and Chalmers' characterisation, to more moderate framings such as that of Alvin Goldman (2012); for the latter, cognition is embodied in the sense that sensorimotor experience provides cognitive processes that can then be redeployed in more abstract reasoning. In the course of this thesis I will at many points describe how a person's process of thinking about an abstract topic is bound up with the manipulation of some external resource; in these situations embodied experience is not only source material for cognition but a method for reasoning, for specifying and testing hypotheses, analysing situations and so on, through manipulation and feedback. This kind of online back-and-forth goes beyond Goldman's picture of embodied cognition, but it is another question whether to adopt Clark and Chalmers' position that a resource used in the course of thinking, that is in some way constitutive of the process of thinking, should therefore be considered a part of the mind itself. To call it a *part of the mind* is a bold move that provokes fruitful reflection on where those boundaries ought be drawn, but in straying into questions about personhood and individuality, may go beyond what is helpful for the present research. Instead I will stay with a claim that can be supported by the material exhibited: that such resources play a constitutive part in certain acts of thinking.

Situated cognition is not only about interaction with material things; a useful ethnographic reference point for this project is an insightful ethnography of a navigation team written by Edwin Hutchins, in which a portrait is built up of a group of people as making up an extended cognitive system. In *Cognition in the Wild* (1995), Hutchins presents a fascinating ethnography of collaborative cognitive systems in the navigation of large ships, powerfully making the case that the team's interactions with one another and with certain tools constitute an external cognitive system and, indeed, that our understanding of cognition to date is far more rooted in external practices than has previously been seen. Known as the founder of modern cognitive ethnography, Hutchins is well qualified to make such a claim, and does so convincingly, detailing how many of the processes thought of as internal cognitive processes can be evidenced in these external interactions and co-ordinations.

This move places communication as an integral part of distributed cognition and, indeed, knowledge production, a point with clear implications in this discussion of communication and mathematical progress. We need not believe that mathematics is never done inside a single meditating thinker's head to see the importance of these insights for the discipline. Undoubtedly some mathematical work does happen privately but, as Hutchins notes in another paper, 'private disembodied thinking […] is relatively rare in the global cognitive ecology [… and] is a deeply cultural practice that draws on and is enacted in coordination with rich cultural resources' (Hutchins, 2010). These questions about the boundaries of cognition will be useful to keep in mind when considering the data presented, and these ideas will be revisited in the discussion.

Proposition 1.4. Art as research

My final research question pertains to my aim of coming to see mathematics somehow in a new light, and here I will make the case that a valuable ally in this will be the adoption of an artistic outlook in my research. Debates have raged about the validity of undertaking research in the arts, and its relationship to traditional academic research, particularly as art-practice-as-research has come to be recognised at the doctoral level in the last few decades (Borgdorff, 2006). A question is posed: whether art shares sufficient of the aims, methods and assumptions of other research practices to be understood as being of a kind with them, and whether treating it as such will further its goals, or will betray them in pursuit of an easier positioning within institutional and funding structures. Some consider art to be inherently research-oriented (O'Riley, 2011); it is for a position of this kind that I will argue.

The artistic process is not all about inspiration and bold gestures. It involves careful, exacting work, processes of reflection and adjustment that are honed and refined over years of study. To illustrate how this refinement is achieved, we might look to one of the most common teaching methods employed in arts education: the critique. In these critiques, students produce a piece of work and are exposed to their peers' reactions to that work, as the piece is discussed and evaluated. This gives students access to two important things: one, direct data about the responses that the art object produces in others, whether the object is somehow reminiscent of a place or state, produces an emotional response, causes the viewers to ask certain questions and so on; and two, an experience of developing those responses themselves, dialogically, as partial and tentative contributions to a field of possible responses. The success and prevalence of this method would seem to suggest that artistic work is heavily concerned with assessing the responses that a multiplicity of human minds might have, responses that are *expected* to vary, when contemplating a particular art object, and carefully adjusting those possibilities for response by adjusting the object. These responses are intended to be multiple, open-ended and sometimes unforeseen, so the process is one of carefully developing a prompt for an undefinable, and yet very particular, set of responses.

If these responses are allowed to be so open-ended, then exactly what set of responses is it that is being sought? Even in the staggeringly diverse literature of the theory of art, there are certain common threads to be found. One is an understanding of art as putting forward a new way of being or seeing, as exemplifying an alternative world or worldview that allows us to better see our own (Murdoch, 1959; Rancière, 2013; Bourriaud, 1998). If the effect being sought is for an audience to somehow see the world differently, then perhaps this would account for the simultaneous exactitude and difficulty of definition noted above.

Jacques Rancière is a prominent theorist of art, and is well known for writing about how art can come to be truly political as well as truly art. His vision for a political art is based in making visible the system of what is seen and unseen in a particular community, which he calls the *partage du sensible*.[3] Rancière's basic proposal is that any community has a way of organising and prioritising that which it senses, and this is what defines what is seen by that community, what can be spoken about, and what therefore can be done.[4] Rancière's vision for a political art is that a work of art can propose something that exists outside of that system, and by so doing, can work to make that system visible (because from the inside, it is difficult to see what isn't being seen) and even to challenge it in the form of an alternative system—creating the staging of a conflict between two systems of sense, which Rancière terms *dissensus*. 'The distribution of the sensible reveals who can have a share in what is common to the community based on what they do and on the time and space in which this activity is performed […] it defines what is visible or not in a common space, endowed with a common language, etc. […] Politics revolves around what is seen and what can be said about it, around who has the ability to see and the talent to speak […]' (Rancière, 2013 pp.12–13). For Rancière, then, a work of art can intervene in the world beyond it by proposing a different way of seeing, a different system of sensing, with radical consequences.

[3] This term is a little difficult to properly translate. It is often rendered the 'distribution of the sensible', but that doesn't quite capture the original *partage*, which has a sense of the division, organisation or classification of the sensible.

4 This need not make any particular ontological claims—it says nothing about what actually is the case, rather what I given attention by the community.

Thinking in terms of this notion of re-seeing might explain some of the usefulness of ethnography. To use the mathematical world as an example, it is easier (or more accessible) for mathematicians to see and talk about certain aspects of what they do than others: the mathematical strategy upon which their minds are concentrated, say, as above the discursive practices that they are using. This might mean that it is easier for them to reflect on and change their approach with regard to a mathematical method than to reflect on and change their method of conversing. Bringing a sociologist's system of seeing into conflict with that used by the community might therefore make it possible to perceive and act upon totally different aspects of their practice than would be accessible if the practitioners were reflecting alone.

This way of speaking is reminiscent of the principles of Relational Aesthetics, a much-hyped and much-criticised mini-movement in art around the turn of the century. Critic and curator Nicolas Bourriaud gave the movement its name. Bourriaud talked about the aspects of art that were being picked up and developed by this movement, saying that '…what really good artists do is to create a model for a possible world, and possible bits of worlds […] any artwork is a relation to the world made visible' (Stretcher | Features | Nicolas Bourriaud and Karen Moss, n.d.). Once again, we have a sense that what art is supposed to be doing here is putting forward a model of an alternative way of being, and it seems important that these models should be consistent, functional, plausible: '…the role of artworks is no longer to form imaginary and utopian realities, but to actually be ways of living and models of action within the existing real, whatever the scale chosen by the artist' (Bourriaud, 1998 p.13).

There are echoes of this in the writings on art of Iris Murdoch, a philosopher and novelist who succeeded in being genuinely good at both. She references Kant's account of aesthetic appreciation, which involves the recognition of a kind of purposiveness but without an identifiable purpose, so that the work is deliberate without fulfilling some kind of objective end. Murdoch's neo-Kantian account stresses the importance of recognising that something can be utterly other and strange to oneself whilst also structured, existing as she puts it "…in accordance with a rule we cannot formulate" (Murdoch, 1959 p.43). Murdoch emphasises that true appreciation of an art object entails that it cannot simply be subsumed into the old way of seeing, but offers a new one with a rich integrity in itself that operates in some way that is incommensurable with the old. She illustrates this with a comparison to language-learning: "If I am learning, for instance, Russian, I am confronted by an authoritative structure which commands my respect. The task is difficult and the goal is distant and perhaps never entirely attainable. My work is a progressive revelation of something which exists independently of me" (Murdoch, 2001 p.89).

Derived from the above sources, here is an idea of what might properly be considered an aim of art: To put forward other ways of being that are self-consistent and that give the sense that they might be extended but are nonetheless strange and unexpected, to have a consistent kind of internal logic in order to be convincing, but to operate in some way that is markedly different to the world we know.

It is to be noted that proposing whole new ways of seeing is something that some would say is important for science to be doing—which is certainly something we understand as research. Thomas Kuhn described paradigm shifts in scientific research (Kuhn, 1962), and those are supposed to be respectively incommensurable ways of delivering intelligibility and fruitfulness. Paul Feyerabend takes up this idea, claiming that the incommensurability of competing theories is not only acceptable but crucial to the functioning of science, since '[w]e need an external standard of criticism, we need a set of alternative assumptions or, as these assumptions will be quite general, constituting, as it were, an entire alternative world, we need a dream-world in order to discover the features of the real world we think we inhabit' (Feyerabend 1975, 22). Indeed philosopher Catherine Elgin makes the case that science and art both have the aim of furthering understanding through thought experiments that exemplify ways of understanding the world and serve to reorient our thinking (Elgin, 2017). Quoting a description of hers of scientific work might help to clarify exactly the parallel that is being drawn: Elgin draws our attention to the way that scientists seek models, theories, and experiments that operate by concocting fictional worlds, writing that 'experiments… distance, isolate, and purify. They set up circumstances, sometimes quite unrealistic circumstances, and see how things play out. They devise contexts, including and omitting as necessary to bring out the features they seek to highlight. They can be vehicles for discovery' (Elgin, 2017 p.14). Stating 'understanding' as an aim of art may obscure some of the transgressive aspects of art more than Rancière's or Bourriaud's pictures do, but certain principles remain the same. If an art object can put forward some way of seeing a subject that is incommensurable with the accustomed way, then it might be serving to further something like our understanding of it in just the way that a new theory might.

A parallel has also been drawn between art and philosophy, a quite different but equally longstanding site of research, by philosopher Alva Noë. He writes of art as a way of pinning down, and so examining and developing, the patterned interactions we have with the world (which for Noë importantly includes perception). According to Noë, our interactions with the world involve a certain sort of organisation, back-and-forth exchanges that shape our actions. Noë argues that art is there to reflect on and reform our organised activity, itself a reorganising practice. This is illustrated by the case of choreography and dance. Dance emerges naturally in the lives of humans as an organised activity, movement that is shaped and developed in dialogue with external things, rhythm, environment and sound, and it is creatively and reflectively explored by—and subsequently performed with reference to—choreographed performance, which is done 'to fashion for us a representation of ourselves as dancers' (Noë, 2015 p.17). These representations are what allow us to test and change the boundaries, to see ourselves and our relationship to the world anew through experiment, to stabilise and examine our organised practices and so to greatly change and develop them. Noë's claim is that '[p]hilosophy is the choreography of ideas and concepts and beliefs' (Noë, 2015 p.17).

Art would appear to expect more in the way of unresolvability, of an absence of any one correct understanding of a particular work. This may be by virtue of its recognition of and engagement with subjective experience, its decision to engage with that which is most individual in our minds. Might this preclude art from producing knowledge? This is where the parallels with science and philosophy are useful to consider: in a highly-cited paper in the philosophy of science, Larry Laudan described even science as 'neither exclusively nor principally epistemic' (Laudan, 2004 p.15), since a model or description of the world has far more to offer than a group of data points. The map is not the territory; a model is not intended to be in complete correspondence with the world itself, and is instead most useful by virtue of its suggestive relationship to it, by the way in which it reorients thinking and generates new perspectives. By this description, art is not so much in a different realm to these endeavours but on a different place on the scale, dealing perhaps with more elusive aspects of worldview. Art, then, would seem to have a secure role to play in upending our assumptions and allowing us to see the world anew.

Corollary 1.5. Artist as ethnographer

This conception of the work of the artist gives a clear purpose to the artist-in-residence, on placement, or as ethnographer or cultural explorer. What, then, recommends itself as a subject for which artistic enquiry might have something to offer? An answer may be any that asks to be seen with fresh eyes, and in particular topics that pose challenges for other types of research. We might also think of areas of research in which systems of belief or understandings of the world are identified and examined.

A group that famously sought to bring artistic ways of seeing to bear in engagements with organised groups was the Artist Placement Group (APG), established in 1966 by Barbara Steveni and John Latham, and others. The APG worked to place artists in many industrial settings and large organisations, with the intention of positioning the work of the artist within a wider social context. It is interesting to note that in this initiative, artists were not required to produce tangible results in the course of their placements; their presence, the engagement itself, was considered to be the chief goal of the encounter, as the artists brought alternative perspectives and worldviews to bear in their dialogues with the organisations they worked with (The Individual and the Organisation: Artist Placement Group - Announcements - e-flux, n.d.).

Hal Foster, in the classic paper *The Artist as Ethnographer?*, considered the emerging tendency by artists to position themselves as ethnographers, recognising the ability of art 'for instance, to recover suppressed histories that are situated in particular ways, that are accessed by some more effectively than others' (Foster, 1996 p.306). Indeed, outsidership has been rendered an advantage in certain observational studies, since a group's portrayal of itself is bound to be formed by (to use a Rancièrian term) the particular regime of sense common to the group. In the field of Science and Technology Studies (STS), Latour and Woolgar's *Laboratory Life* used a strategy of 'making strange' in their observation of scientific practices to attempt to escape received interpretations.

> For us, the dangers of 'going native' outweigh the possible advantages of ease of access and rapid establishment of rapport with participants. Scientists in our laboratory constitute a tribe whose daily manipulation and production of objects is in danger of being misunderstood, if accorded the high status with which its outputs are sometimes greeted by the outside world. There are, as far as we know, no a priori reasons for supposing that scientists' practice is any more rational than that of outsiders. We shall therefore attempt to make the activities of the laboratory seem as strange as possible in order not to take too much for granted. (Latour & Woolgar, 2013 pp.29–30)

Scientific study produces a kind of knowledge that is accorded high status and universal validity, and cross-disciplinary study of the sites in academia in which this work is done seem bound to produce interesting perspectives on what it is to do research (notable examples include Traweek, 1988; Galison, 1997). Scrutinising knowledge-making procedures has the potential to reflect on subject and on researcher, and bringing the ways of seeing of two different disciplines into contact with one another will make visible their respective assumptions and values. What's more, art practice research is, by virtue of its diverse means of expression and consciousness of human audiences, perhaps uniquely placed to investigate the material and subjective aspects that might be 'suppressed histories' in these realms, and may similarly slip out of view in a written analysis.

Latour and Woolgar portray scientific work as an activity reliant and focused upon the distribution and use of increasingly stable representations (inscriptions), moving from lab readouts to publications, but the tendency sometimes observed in mathematicians to downplay the importance of its written manifestations promises a more complicated picture. Mathematics is a field that deals with complex abstractions whose ontology is mysterious, and presents great barriers to outsider engagement in the usual course of things, and so is a subject that almost demands an inventive kind of study in which a transgressive way of seeing is employed. A form of research that can be developed not only in the form of a text but also through material experimentation and participatory experiments may be able to go further than the ethnographies listen in the introduction in representing the material and embodied aspects of the field.

An artist encounter might be expected to be more transgressive, playful and exploratory than these, but this does not mean that it cannot engage in a theoretical dimension. Its insights can also proceed from a direct experimentation from the kinds of practices being observed. Gemma Anderson's residency at Imperial College, London, began with a drawing-based classification of various forms that have emerged from their work on Fano varieties,[5] and from examining the uses of drawing in mathematical work and as deployed by the artist observer, she and the mathematicians she was observing developed an account of the epistemological role of drawing in mathematics in which the drawing produced is a kind of record of an encounter between a logical Thinker and a mark-making Drawer (Anderson *et al.*, 2002; Anderson *et al.*, 2015). This image falls in perhaps too readily with a traditional Cartesian division of mind and body, but the characterisation of the drawing as their interaction is interesting nonetheless, and an indication of the potential theoretical contributions to be found through an artistic methodology.

Exploring such notions through a research programme of *drawing to understand* is a move that seems helpful to my aim of staying close to multimodal data and taking it on its own terms; it echoes the current impetus in post-qualitative research to explore the different insights that can be reached through non-traditional writing in research, as exemplified in the 'transgressive writing' adopted to expose different aspects of women's experiences by Gustafson *et al.*, (2019), in Anna Tsing and Laura Ogden's shaping of their writing after the forms of the mushrooms and rhizomes that are their subjects (Tsing, 2015; Ogden, 2011), the 'creative analytical practice ethnography' and story-writing of Laurel Richardson and Elizabeth St. Pierre (Richardson & St Pierre, 2008), and the variety of authors blending ethnography with poetry, fiction, memoir, and so on in *Crumpled Paper Boat: Experiments in Ethnographic Writing* (Pandian & McLean, 2017). Very broadly, each of these argues that including non-traditional academic writing in a research project can contribute a different kind of insight into a topic, particularly where form is carefully matched to research subject. In the methods section, I will describe the approach that I have adopted to take advantage of such a technique, one in which the practices observed become the medium for experimentation and through which insight is produced, enacting Alva Noë's conception of art as a practice through which existing practices can be examined and challenged. In this way I meet my aim of producing research that is truly multimodal and to subject ethnographic data to interpretation according to transgressive logics and ideologies.

Such projects successfully blend traditional and non-traditional approaches, and indeed there is a clear advantage to be gained from also taking advantage of the insights of more traditional ethnography. There is a danger of a certain arrogance to projects that seek to bring artists into contact with particular groups with an expectation of their inherent insightfulness, since it runs the

[5] A set of mathematical geometries thought of as the atomic pieces that make up other mathematical shapes

risk of duplicating the dubious power dynamics of early anthropology. It is important for such artist-explorers to take heed of the lessons learned in those disciplines; in *The Artist as Ethnographer?*, Foster warned against artists falling into the pitfalls of a naïvely conceived ethnographic position. Recognition and examination of the relationship of the artist to subject is one of the habits of artistic work that Foster suggests recommends it for ethnographic work; a researcher can also take steps to draw upon the traditional research frameworks that might be brought to bear on a project, in the hope of combining the better qualities of each.

Lemma 1.6. Ethnography of Communication

As Latour and Woolgar observed, the observable aspects of mathematics are few; the practices most central to it are generally taken to be cognitive in nature and the question of the ontology of its objects of study has raged for centuries.[6] As we have seen, the moments at which it is most observable appear to be at moments of exchange between mathematician and some kind of 'other': mathematician and collaborator, mathematician and pencil and paper, mathematician and the community. The question of how mathematical writing plays its part in the doing of mathematical work is to be central to this thesis, and even for an artist explorer the research design must take note of the insights gained by years of research of this nature in other fields. As such, the field of Ethnography of Communication (EOC)[7] provides an excellent stepping-off point for the design of the research. EOC was developed by Dell Hymes and enacts language-is-language-use perspective (in contrast to a Chomskyan view), taking the position that the analysis of a language (and communication more broadly) has to take into account the patterns and features of its social and cultural context (Hymes, 1964; Saville-Troike, 2008). Ethnographers of communication study topics such as the patterns and functions of language in a speech community, language ideology, how members acquire communicative competence, the relationship of language/communicative means to worldview, and the universals and inequalities to be found in a language community (Saville-Troike, 2008). EOCs have been carried out in academic communities, considering questions such as socialization into academic discourse (Duff, 2010) and the sticking points in communication between scientists and policy makers (Bernard, 1974).

---

[6] It is worth noting the impact that these debates have had through history, generating really important work by thinkers from Pythagoras and Plato to Descartes to Frege, with implications through philosophy, science, modern computer science and beyond.

[7] First termed 'Ethnography of Speaking' by Hymes, but later broadened to include the spectrum of communicative means

A question appropriate to the aims of this research seems to be that of how mathematicians come to understand one another, to come to a shared understanding of a topic, across the range of communicative means.[8] Over the course of my research I came across mathematical texts which comprise prose, notation and diagrams, speech events that involve speech, gesture, and public writing, and other ad hoc forms of communication such as email exchanges and mutual edits in a LaTeX document. One thing that an observer can do is observe the commonalities and differences between different communicative situations and the resources that are used; further to that, an outsider has certain advantages when it comes to carefully observing how research subjects come to understand one another, since having limited contextual knowledge relative to those participants can allow an ethnographer to examine the sticking points (and non-sticking points) in her own understanding in order to become aware of exactly what kind of context is employed. It is worth noting that this limited understanding is not such an unnatural thing to introduce into situations of mathematical communication; Michael Barany observed in the course of his own ethnography the 'embarrassing open secret that mathematicians tend to have comparatively little idea of what each other does' (Barany, 2010 p.32), a fact that does not prevent them from attending one another's talks, discussing one another's work, and gaining genuine insight from so doing. How and why this in fact seems to work is a worthy mystery to study. To emphasise the *how* of understanding in this way makes it likely that some kind of model of communication itself will need to be adopted; but more on that later.

As for particular research methods, there is one research approach (sometimes considered something of an *enfant terrible* of the sociological world) whose methods emphasise seeing anew in a way that could well be of use to this project. It is known to EOC for its association with discourse analysis, and could provide some of the methods needed to structure an attempt to see data in new ways.

Definition 1.6.1. Ethnomethodology

The questions being investigated here are about how people go about doing mathematics, reach consensus and advance their work. There is a branch of ethnographic study that is dedicated to examining the methods that people use to establish order in everyday life, how sense is made of the world and one another, and some of its techniques will prove useful in this research. This branch,

[8] The act of attempting a communicative ethnography of mathematics should not be mistaken for an argument for the view that mathematics *consists in* a language, or just *is* its formal manifestations, a view held by formalists and debated from realist and anti-realist positions (over whether mathematical objects exist independent of the language and people involved in studying it, or it is a system of doings and community activities, respectively) (Mancosu, 2008; Shapiro, 2001). It is not necessary to assert that nothing exists beyond a language in order to study that aspect of it. The position that this research does take is that the external manifestations of mathematics and how these figure in mathematics as a mental activity is worthy of study.

ethnomethodology, was borne out of a radical rejection of the coded structures hypothesised by traditional sociology; as proposed by its creator Howard Garfinkel it suggests adopting an agnosticism toward the concepts developed and sought in traditional sociological analysis, whether imposed from above from a societal perspective or below from a 'mentalistic' perspective, encouraging researchers to focus instead on the ways that the observed actions themselves bring order and meaning into being in the situation observed. Since then its techniques have been practiced in more moderate form, for their usefulness in paying serious attention to the way a group produces order from day to day. In ethnomethodology, the focus is on the way in which meaningful action is reflexively produced, holding that it is these everyday interactions that find and reinforce this meaning (Garfinkel, 1967; Garfinkel, 2002; Rawls, 2008; Saville-Troike, 2008; Francis & Hester, 2004).

Some of the principles and methods of ethnomethodology are as follows. The ethnomethodologist should look for the methods used by members of a group to construct the meaningful orderliness of their encounters. This has been deployed by researchers entering a situation in which they have a basic, vulgar competence, and observing what that competence consists of, observing how they and others make sense of those situations, not just in a social sense, but also in terms of a group's organisation of events under concepts (Livingston, 2015; Livingston, 1989; Cicourel, 1987). One principle is to attempt to describe these events as though they are a 'first time through', to become aware of how the observer organises the observed material under a description and how the actions observed shape and are shaped by such descriptions. Efforts are often made to make the data as available as possible for inspectability by the reader. When studying the record produced one night by John Cocke and Michael Disney of a pulsar, Garfinkel *et al* attempt to characterise the shareable knowledge of a pulsar, the data and the concept of its existence, not simply as *caused by* the pre-existing pulsar out in space (as it is characterised in the discovery announcement) but as a product of Cocke and Disney's efforts to construct something intelligible, exhibitable and analysable in the world of astronomy (Garfinkel *et al.*, 1981). Efforts to describe reflexive meaning-making can make ethnomethodological texts rather dense. A rather more spectacular method sometimes used in ethnomethodology is the breaching experiment, in which a rule is flouted to see what methods are used to reassert and reassemble order; an example might be driving the wrong way down a one-way street, and observing the results (Garfinkel, 2002).

An ethnomethodological study of mathematics was carried out by Eric Livingston in 1986, in which he laid out certain proofs and guided a reader through them, attempting to make clear exactly what it was in the course of these proofs that rendered them logically convincing, which he explains as depending on their 'natural accountability'. Livingston's conclusions are a kind of gloss on the activity of proving that is rather difficult to summarise, and indeed David Bloor levelled the charge that, far from explaining anything, his dense, terminology-laden description of how conviction occurs ends up being rather circular, giving answers that refer back to the questions and saying little beyond the existing common-sense picture of what is logically evident. Bloor's opinion is that the culprit here is ethnomethodology's refusal to theorise, that the research tradition is so determined to stay close to the processes that it is describing that it is unable to reflect and attain explanatory value (Bloor, 1987). An ethnomethodological approach, then, is not without its dangers, although this and other projects of Livingston's have the great advantage of having closely scrutinised, rather than taken for granted, mathematical reasoning, and examined the everyday practices that bring them forth in mathematical work (Livingston, 1999; Livingston, 2006; Livingston, 2015). In using an ethnomethodological approach, then, it might be worthwhile to consider which aspects to adopt, to relax certain constraints, and perhaps be a little choosy.

My fourth research question was a methodological one: how to give a reader access to the observed data, and subject this data to alternative ways of seeing. Ethnomethodology is an approach designed to do just this: to bring the data to the reader and to endeavour to really see with new eyes the organisation that is happening on the ground. With some modification, then, I use certain techniques adopted from and inspired by ethnomethodology in this research, recognising the value of this approach for meeting the aims of seeing a practice anew and bringing the data itself to any reader for inspection.

There were a number of requirements laid down for this kind of study by its creator Howard Garfinkel (Garfinkel, 1967). One was the *unique adequacy requirement* of a vulgar competency in the setting observed, i.e. observing as a competent participant. In this respect I intentionally push on certain boundaries and adopt an adaptation of that requirement, positioning myself as an interested outsider engaged in a process of sense-making and competent in my constructed role as a non-mathematician engaging with mathematics, a deliberate attempt to demystify mathematics and make the process of understanding it available to a broader audience. This deliberate adaptation of the ethnomethodological approach brings questions of insider- and outsidership into the scope of the analysis, asks which aspects of sense-making depend on expertise and which do not, and makes plain exactly how external resources come into play in that sense-making.

Another principle of ethnomethodology is that explanations should not be given in terms of other entities, including minds or social structures. This aim does seem to inhibit the explanatory power of ethnomethodological description, as Bloor observed;[9] indeed on the particular question of how communicating individuals understand one another it seems to produce a curious myopia in ethnomethodologists' writings. For example, when considering the question of sense-making in communication, David Francis and Stephen Hester strive to argue that any kind of 'mentalistic'

[9] More broadly, ethnomethodology has been described as problematically descriptive, and it has been widely argued that observation cannot help but be theory-laden (Kuhn, 1962; Feyerabend, 1975), and completely eliminating theoretical influences seems a faint hope.

explanation is unnecessary, claiming that sense-making is in general 'massively routine and unproblematic', that in general we find one another's actions 'transparent' and communication accordingly so, and that instances of underdetermined meaning can be easily decided by reference to 'context' (Francis & Hester, 2004 pp.5–9). This characterisation understates the scope of the problem of underdetermination of meaning that has been grappled with by philosophers of language and linguistic pragmatists (theorists of communication in context) for years (Grice, 1989; Sperber & Wilson, 1986/1995; Levinson & Levinson, 1983; Carston, 2009), and though its reference to 'context' provides part of a possible solution, what is omitted by such an exclusively externalist account is an account of *how* a hearer might successfully integrate the content of an utterance with the appropriate features of the appropriate context. To engage with these kinds of questions may indeed require some engagement with theoretical contributions from linguistics and the philosophy of mind, and this is where I will find myself diverging from the ethnomethodological philosophy.

What's more, though the subjects of investigation are the patterns and methods of day-to-day interactions, when dealing with something so seemingly cognitive as sense-making it may not be possible to properly make sense of the methods that are being used without some reference to theories that offer accounts of ways in which minds interact with their environments. It will be useful, then, to use some of ethnomethodology's means of setting aside assumptions and making data available to the reader, without subscribing to the whole radical approach. This, in fact, is not so uncommon; ethnomethodology, descriptive as it is, is treated by some simply as a data collection procedure (Saville-Troike, 2008 p.94), and while its usefulness to this project seems to extend a little beyond that, there is plenty of precedent for integrating it with more theory-inclined research. Instead the alterations of the ethnomethodological approach proposed above recall Garfinkel's use of *misreadings* of texts in the course of doing theory, putting forward alternative readings of texts or fragments of texts that suggest new ways of situating or deploying particular insights (Garfinkel, 2002). We might consider this research design to be a misreading of the ethnomethodological method itself, taking its methods and concerns and deploying them in a new way, as an aid in the attempt to set aside assumptions about what is important in the doing of mathematics and perceive new aspects of the on-the-ground, everyday work of its practitioners.

What ethnomethodology offers is a set of methods for research that suit the kind of exploratory, artistic research proposed, and give it a clear way to proceed. The ethnomethodological principle of inspectability of data is appropriate to research that holds as an aim the granting to its audience of direct access to a world that is often seen as inaccessible. Research that takes observational data from the mathematical world, reproduces excerpts for inspection by any reader, and takes that reader through a process of coming to make sense of it and recognising what is involved in that process, will produce not only conclusions of its own but also the means to access and problematise sense-making from other perspectives. The 'first time through' analysis with documentation of the researcher's own process of making sense of a situation offer a clear method for analysis, one in which the researcher's own position and routine self-reflection are documented and discussed. Ethnomethodology's more creative and transgressive method, the breaching experiment, mirrors the kind of proposal and testing of alternative ways of being that many see as centrally the goal of artistic work and demonstrates the true research insight that such experiments might produce.

With all of this in mind, in the next chapter I will outline the particulars of the methods used in this research.

Interlude 2. Embodied knowledge

Interview with subject J

00.20.00.000

J: A paper is not the best possible vehicle for a mathematical result. […]

What is the best vehicle?

J: I don't know. I can give you an example. So several months ago perhaps 5 months ago I asked a couple of people of department if they can prove a certain result for me – not actually prove it, you don't know what the result will be, so I asked them to investigate a certain object, and one of them came back within a couple of weeks saying yes, I – I proved such-and-such result. So the question then was whether this result is new or not. So I started searching just on Google, who can know something about that. I found one mathematician in the United States who could know something about that so I wrote to her. She said it's probably that. And I found one of her papers, and she thought yes, she thought that paper of hers might be relevant. So I started searching in the literature for that particular result again, and I couldn't find it. So then I asked some other people and another mathematician from Spain and he replied no actually it's not *that*, it's *that*, so I'm very grateful to him because he pointed me in the right direction. So now we know that that object that I was thinking about four months ago was this object which was known to knot theorists. So- uhh- after that I started looking in books and in that book I found that result which refers to a certain paper of 1987 written by two people, so then I found that paper and actually I have read that result and the proof of the result in that paper, I'm not interested in that paper, but I'm interested in that particular result. And actually we have an alternative proof now which is in a completely different style, so I had to read their proof to make sure what- that the way we proved their result is completely different.

And so you found yourself moving between papers from now and from a long time ago and books to find…

J: And live people.

Work site 2

Work site 2 was the office of a PhD student at the same technical university in continental Europe.

Figure 7. Office space shared by PhD students, with desks and shared whiteboard. Original in colour.

Figure 8. Work space of my participant, with work books and relevant posters pinned to the wall. Original in colour.

2. Methods

The overarching approach adopted in this research is as follows: to carry out ethnographic observations and conduct analysis in experimental ways from the perspective of an artist outsider, with the aim of building up an *artist's communicative ethnography* of mathematical research.

2.1. Data Gathering

The ambition of fine-grained analysis of particular situations of mathematical work recommended a case study approach to data, focusing in closely on certain short extracts from observations. What I most hoped to observe was mathematical work in progress: mathematicians working on and refining ideas, in dialogue with colleagues and with inscriptions. The intended outcomes then were not quantitative data of any kind but rather close discussion of examples that demonstrated some particular way of looking at them, of seeing what might not be noticed, that would have broader significance for perceptions of mathematics as a whole.

My aim in data gathering was to acquire a record of at least one piece of work as it moved from work-in-progress to publication of a formal mathematical text. The unpredictable nature of research meant that I judged that this aim was most likely to be fulfilled if I were to conduct observations of a number of mathematicians, and then to keep in contact with these subjects to see which projects had resulted in a publication. The advantage of this approach was that the additional observed material provided valuable contextualisation for the examples that were eventually chosen for analysis. I visited participants at their home institutions, and discussed with each of them how I might take some kind of record some aspect of their current work in progress. The data that I recorded tended to come from communications between collaborators, whether research meetings or email exchanges, and participants' own notes and note-taking practices. In the course of each visit, I also took the opportunity to conduct semi-structured interviews to gauge the mathematicians' own attitudes to mathematical writing and other relevant topics, as well as photographing several of the mathematicians' work spaces, excerpts from each of which you have been seeing in the interludes of this thesis.

I visited a total of nine mathematicians as my primary participants, whom I interviewed. Though I was not attempting a rigorously representative sample, I chose these participants with a consciousness of certain distributions. I visited five women and four men. My subjects are all members of the academic mathematical community, but it is a mistake to think of this community as without internal cultural variation; though all of my subjects work and publish in English, I made an effort to engage with academics at universities outside of my native UK, visiting four at institutions in the UK, three in the USA and two in Greece. My interest in different types of external representations and communication styles meant that I wanted a mixture of diagrammatic and non-diagrammatic fields in my observation, and I visited five mathematicians who for the most part work in fields like topology or applied mathematics that tend to involve diagrams, one who works with computer representations, and three who work in fields like analysis and number theory that, again for the most part, use fewer visual representations.

The qualifications that I have added to this last sentence reflect the discovery I made that mathematicians' specialisms give only an indication of what they might be working on at a given time; a project with a colleague might take an analyst into graph theory, for example. In fact, I found myself with a data set with more of a majority of diagrammatic projects than I had intended. After the fact, I reflected that a slight self-selection bias might have been at play, since a mathematician working on a project with interesting diagrams might be more excited to share it with an outsider than otherwise. In any case this fact contributed to the makeup of the excerpts that I ended up analysing.

The observations I carried out of their work in progress sometimes included colleagues, bringing the total number of mathematicians involved in the project to 17. I am immensely grateful to all of them. With enormous thanks to my nine interviewee mathematicians, and their participating colleagues:

Tim Browning

John Caughman

Charles Dunn

Christine Guenther

Neslihan Gügümcü

Mark Holland

James Isenberg

Vaughan Jones

Josh Laison

Sofia Lambropoulou

Alexei Lisitsa

Daniel Loughran

Nancy Ann Neudauer

Colin Starr

Jenny Venton

Alexei Vernitski

Of my dataset, I found that I had data from three subjects that documented some kind of progression toward formal mathematical writing. Though the manifestations are quite different from one another, in each case I was able broadly speaking to identify excerpts for analysis that included an idea being worked on in some preliminary stage, and also a later, the more formal presentation that this work led to.

2.2. Excerpts chosen

2.2.1. Excerpt from a research meeting, leading to a formal description

A group of collaborators are working on a lemma for a paper they are writing together, and are trying to write a line that will delineate when a particular algorithm will and will not work. The paper remains, at the time of writing, unpublished, but the group did by the end of the meeting formalise the description into a line that could be recalled at the beginning of the next meeting, and this provides the more formal presentation for comparison.

2.2.2. Excerpt from an email exchange, leading to edits to a paper

A pair of collaborators are editing together a paper for eventual publication, and run into a couple of different problems that they need to resolve: one a mathematical error, the other a clarification. They begin the discussion in notes made on a shared paper, and switch to email before resolving both issues and editing the paper accordingly.

2.2.3. Excerpt from a participant's notes, leading to a section of a paper

A mathematician is working on a tablet, laying out a proof intended to be used in a paper. Though the mathematician reported that the really early ideas stage tended to happen on sheets of paper, and that the tablet notes tended to come at a stage when the ideas were more clearly taking shape, there is still a great variation in the roles being played by different types of writing in the tablet notes as compared with the eventual presentation of the proof in a paper.

2.3 Thesis structure

2.3.1. Data Analysis

Each of these examples take place across multiple media, and are importantly bound up with the particular medium in which the mathematical work was taking place. In analysing them I chose to reflect some aspect of the original medium in my writing about the excerpts, whether adopting the form of a conversation, a correspondence, or sketched diagrams and notes. In this way I was able to learn more about those original situations by imitating them, by discovering that I needed to point at the parts of a diagram to get my point across, or annotate a shared document, or that it was by drawing a picture that I could come to understand a particular point.

In the course of my analysis I also went through a process of discovery when it came to understanding what I could of the mathematical content, and this seemed best represented if I attempted some kind of diachronic record of the process of coming to understand. To this end, each of the data analysis chapters is presented as a constructed narrative, a fictionalisation of that process as undergone by a fictional character or characters; the first a dialogue between two mathematical novices, speaking at a blackboard; the second a correspondence between those same interlocutors, discussing a document that is being annotated in a shared place; and the third a set of diagrams and notes made by one of the characters on and around a page of notes, extending rightward in time as the character digs in to the material. This is a kind of transgressive writing (Gustafson *et al.*, 2019), aiming to manifest its conclusions both by what it *says* and by what it *shows*.

As well as making available to the reader direct evidence of both the way that each medium can function and of the sticking points that I encountered in coming to understand, these fictionalised processes of discovery are also intended to guide a reader through the mathematical material, provide a humanised way in to even complex ideas. What I want to avoid is the situation in which a reader finishes a passage of my discussion and thinks, 'well, I can't understand the mathematics, but that sounds good.' In this case I want to give every reader the opportunity to understand as much as I have (limited though that is), whether or not any particular reader chooses to take it.

2.3.1.1. First time through

The above analyses were conducted with the intention of setting aside assumptions wherever possible, in a way informed by the ethnomethodological technique of describing a situation as though it were the *first time through* and such a thing had never been done before. Taking this perspective was an aid in remembering to examine how order and sense are made of material in each excerpt, and though quite challenging to maintain, this perspective is a useful aid in setting aside assumptions. In

Garfinkel et al. (1981) this work is oriented toward understanding just what the group does to extract the *scientific work* from their various doings that night, metaphorically described as bringing the 'animal from the foliage'. In this case, the aim was to understand what was done to bring 'the mathematics' from the conversations or scribblings, to transform that work into a useful, shareable piece of mathematics.

In addition, the fictionalised process of coming to understand serves a similar purpose; the characters in each fictionalisation access a range of different resources to help them to make as much sense as they can of each excerpt, and carefully document the resources (human and non-human) that they made use of. The characters thus experimentally learn what kinds of resources were being used by my subjects, including ones that may, for them, be so obvious as to lie forgotten. As well as the content of the excerpt, an aspect of the *first time through* reading was to pay attention to what exactly was being produced in the course of the work and how this was achieved. In all cases, the data is made as available as possible to the reader, and the outsider's process of making sense of the excerpt is also considered a central source of data.

2.3.1.3. Breaching experiments

Each data analysis chapter is accompanied by a *breaching experiment*, one of the experiments that I carried out in order to see what happens when some of the assumptions of mathematics are violated.

2.3.2. Discussion

In Corollary 1.3, I made the case that an artist ethnographer should be aware of the research traditions that intersect her work, and throughout this thesis traditional and non-traditional means of research rub up against one another. So it should be in the written discussion, the thesis making a genuine effort to contribute to different of the fields that it draws upon, and producing its research output, as well as methods, according to the logics of different types of research. As such, the written discussion takes place across two, interlinked modes:

2.3.2.1. Written discussion

This section is primarily, though not limited to, a contribution to the communicative theory that has informed my work, establishing the outcome of an application of relevance theory in a novel and borderline area of real-world communication. It takes the form of relatively traditional written research, and references observations from the Data Analysis to make the case for a situated, relevance theoretic account of mathematical communication as the best representation.

2.3.2.1. Practice discussion

This analysis builds on the observations and breaching experiments of the Data Analysis to put forward an alternative picture of how mathematical communication might be. The reader is invited to take part in a series of exercises, becoming the human actor working with external resources to solve, shape, or become entangled in, problems. By inhabiting the position of building and testing possible forms of mathematical writing, an audience is invited to inhabit the position of the working, inventing mathematician.

Interlude 3. Writing to understand

Subject L

0.23.53.00

L: I read books a lot, and I write, write even though I work with books, I just- I write everything that is written in the book. To a notebook, or to A4 paper, it doesn't matter.

What do you think that *does*, the writing whilst you're reading?

L: I think it… records… I mean it slows down the process of learning, and it's like- I repeat, you know, it's like if you're talking and you give me instructions, you say 'you need to go right' and I say, OK, I need to go right.

Subject M

0.05.49.000

M: …so then I worked on this project for a couple of hours, and that means me just sitting at my desk by myself, you know, with all of these books, just looking up results as I need them, and trying to work through papers. […] Then if I don't see why something is true, then I try to show it myself…

Work site 3

Work site 3 was the temporary office space used by my subject as a visiting researcher at a UK research institute.

Figure 9. View of outdoor spaces from a communal area commonly used by my subject. Original in colour.

Figure 10. Temporary office space used by my subject while visiting at this institution, with tablet and chalkboard. Original in colour.

3. Analysis of an excerpt from a research meeting, leading to a formal description

The first excerpt is taken from a research meeting.

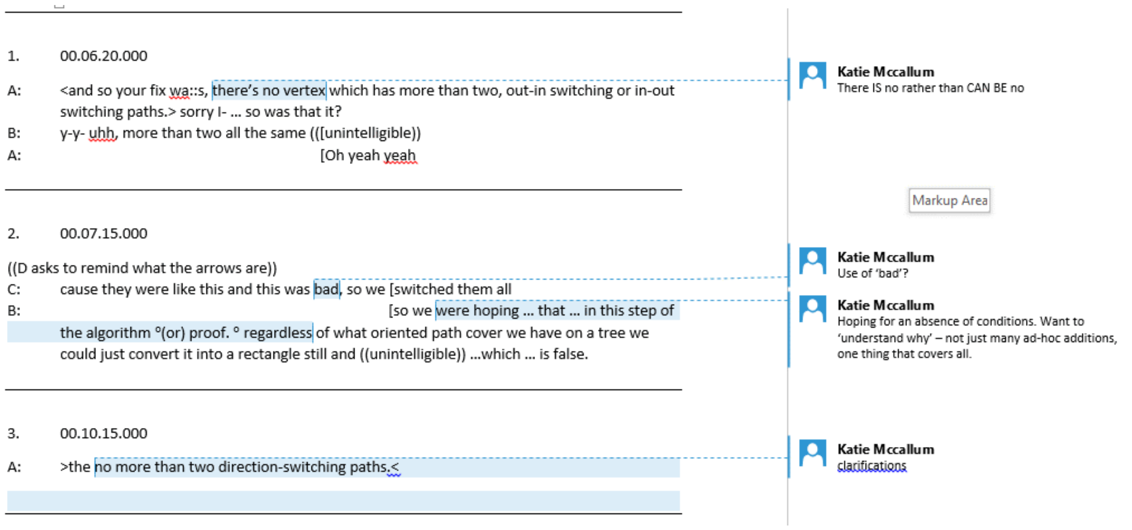
The main excerpt analysed (see 3.3.1. Transcripts for the full transcript) is 32 lines long and lasts roughly a minute. The section selected is taken from a meeting of a research group. It took place in a meeting room at a university in the early evening, and lasted three hours, from which I made recordings totalling 90 minutes. The excerpt was part of the video recorded, about one hour into the meeting, and half an hour into the recorded material. The excerpt was chosen in part because it appears to be an interesting turning-point in the discussion, and it exhibits multiple interesting communicative strategies. More importantly, even a non-expert reading of the excerpt would seem to indicate that what happens is that an individual proposes an idea that is queried by a collaborator who doesn't immediately understand, challenged by another, and then eventually accepted by the group and carefully restated. This progression must involve difficult communication and comprehension, as well as a moment in the developing, refining work of producing a mathematical text. The six participants are known simply as A, B, C, D, E and F.

3.1 Record of a first time through

3.1.1. Carrying out the analysis

This situation was made up of a configuration of multiple people around a whiteboard, the group producing brief, impressionistic utterances in the course of focused interaction with a set of shared inscriptions at the front of the room, which they adjusted and added to throughout the course of the meeting. The participants were engaging in a collective process of clarification and development that seemed to culminate in a short phrase; the atmosphere was one of a shared thought process, collective and exploratory.

First I watched through all of my recordings a couple of times, and then dug in to the recording and transcript of the excerpt line by line, often skipping back through the meeting to see where a term or idea had first arisen to see if I could get a grip on what it meant. The work was slow and confusing; often I would excitedly type up some notes on some new understanding of a particular section and read back through my transcript to see whether it fit, only to find that I just couldn't made sense of one participant's reaction according to that interpretation, or run up against a particular term that I hadn't made much progress on understanding yet. Often my biggest ally was being able to read the emotional content of the chatty, friendly exchanges of the group, noticing when a participant was surprised, or confused, or experiencing a sudden revelation. These clues let me know when to pay attention, which moments were most pivotal and which addition to the sketches on the whiteboard accompanied a shift in understanding. As I worked, I transcribed the sections of the meeting that I'd been making use of, trying to keep track of what I had seen that helped me to understand what was going on (see the annotated transcripts in Figure 11).



1. 00.06.20.000

A: <and so your fix wa::s, there's no vertex which has more than two, out-in switching or in-out switching paths.> sorry I- ... so was that it?

B: y-y- uhh, more than two all the same (([unintelligible))

A: [Oh yeah yeah

Katie Mccallum: There IS no rather than CAN BE no

Markup Area

2. 00.07.15.000

((D asks to remind what the arrows are))

C: cause they were like this and this was bad, so we [switched them all

B: [so we were hoping ... that ... in this step of the algorithm °(or) proof.° regardless of what oriented path cover we have on a tree we could just convert it into a rectangle still and ((unintelligible)) ...which ... is false.

Katie Mccallum: Use of 'bad'?

Katie Mccallum: Hoping for an absence of conditions. Want to 'understand why' – not just many ad-hoc additions, one thing that covers all.

3. 00.10.15.000

A: >the no more than two direction-switching paths.<

Katie Mccallum: clarifications

Figure 11. Notes on sections of the meeting. Original in colour.

I also looked carefully at my photographs from the meeting (see Figure 12) and cross-referenced with the video to make a replica of the board notes the group had made on a whiteboard (Figure 13) so that I could stare at it and pinpoint exactly what was being pointed to at what moment, and what the changes were when they happened. The group's work, I found, was so intensely oriented toward what was happening on the whiteboard that it made sense for me to imitate them, to stare closely at the diagram as I puzzled through the dialogue, looking for the all the joys and frustrations of the group in a few felt-tipped pen marks on a whiteboard.

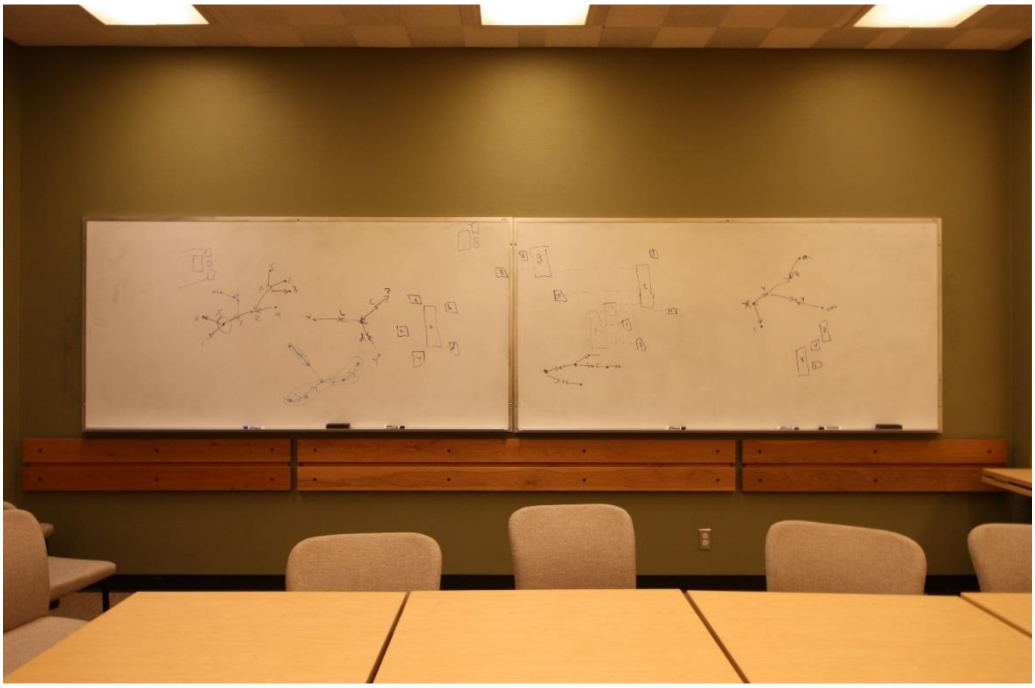


Figure 12. My photograph of the sum of the board notes at the close of the meeting. Original in colour.

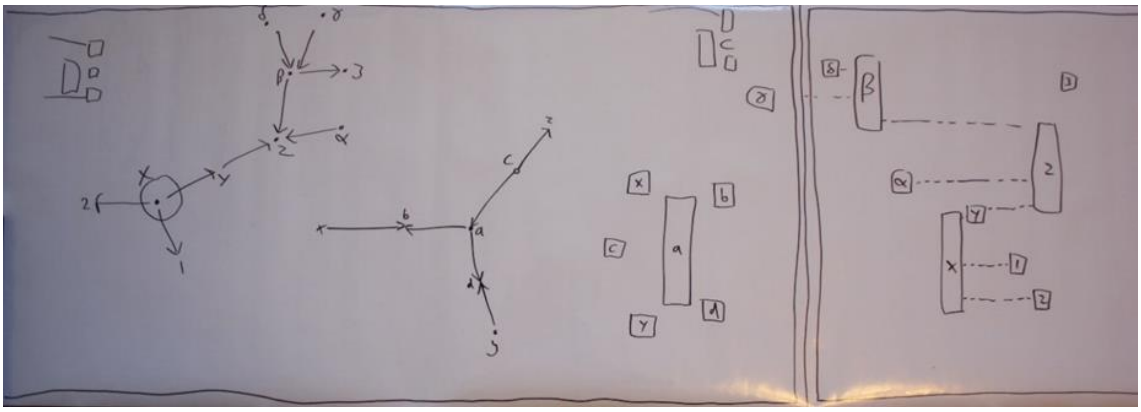
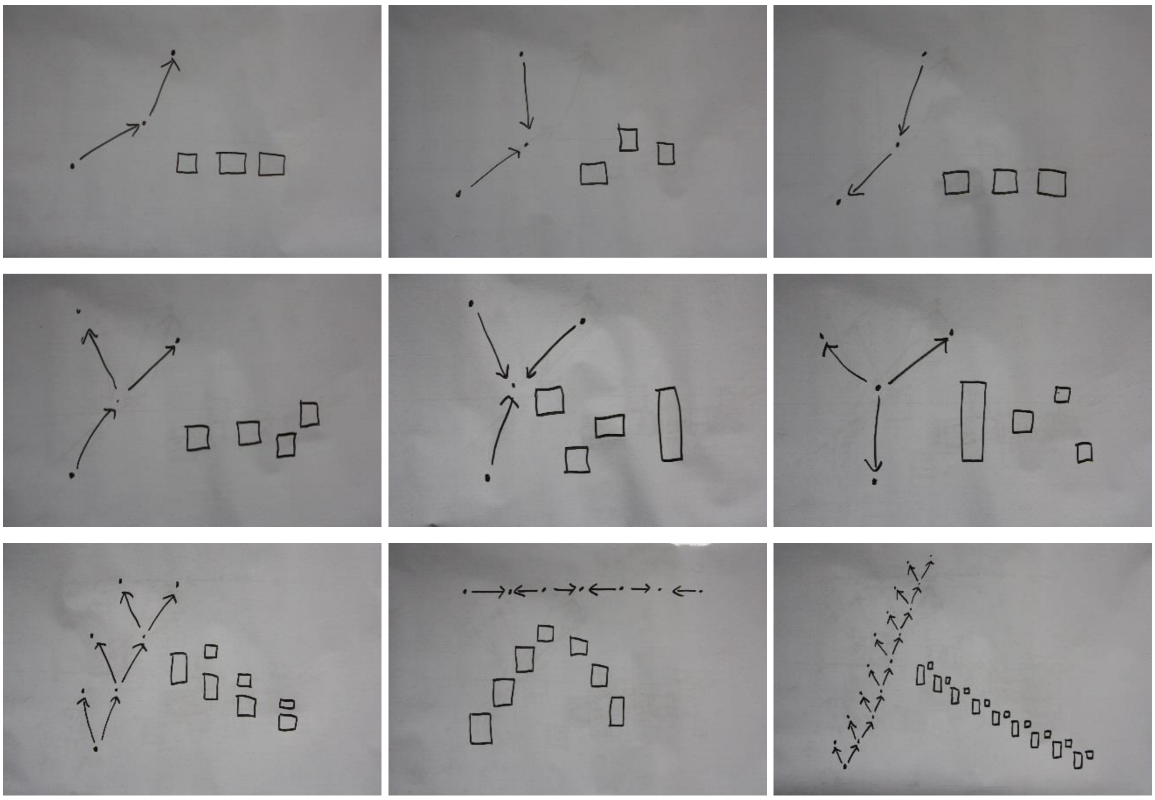


Figure 13. My recreation of the board notes at the time of the excerpt. Original in colour.

I typed up summaries of the mathematical problem and returned to them, examining mis-steps and adjusting my ideas. The rhythm of stating and then examining, proposing a solution and coming back to it with fresh eyes newly able to see the flaws, itself felt like a conversation, one rather analogous to the process that my participants were going through in the course of addressing their mathematical problem. I also paid close attention to the moments of confusion within the conversation I was analysing, and even of potential confusion that in the end troubled no one, treating these as key moments for me to make sense of: why did communication work, or not work, at this moment? How was mutual understanding reached? How did the group come to agree?

The group's aim, I came to understand, was to relate two types of diagram to one another; the tree-like kind on the left, and the groups of rectangles on the right. The diagrams made up of sets of rectangles are 'rectangle visibility graphs' (RVGs), the key question being which rectangles have a clear line of sight to one another in a particular direction. The group were attempting to 'build' an RVG representation equivalent to each tree diagram, respecting the direction of the arrows in each case, and had found that it wasn't always possible to do so; their aim in this meeting was to describe under what circumstances it was and wasn't possible.

As well as watching the participants, I would copy their actions, drawing pictures and experimenting to attempt to come to my own understanding by drawing. My own pictures were often playful, rhythmic, repetitive for the sake of generating interesting shapes. Nonetheless, my experiments helped me through my understanding, the final two examples in Figure 14 helping me to understand the distinction the group were making in describing three switching *vertices* downstream from one, rather than counting switching paths. The penultimate diagram causes no problem in terms of representation, despite its many forks. The final diagram has three separate vertices where switching 'begins' relative to the bottommost vertex, and is unrepresentable as an RVG. I drew this diagram expecting it to be representable as an RVG, and was forced to return to my transcripts to realise that it had precisely the property that the group had been describing: three switching vertices 'downstream' from a particular vertex.



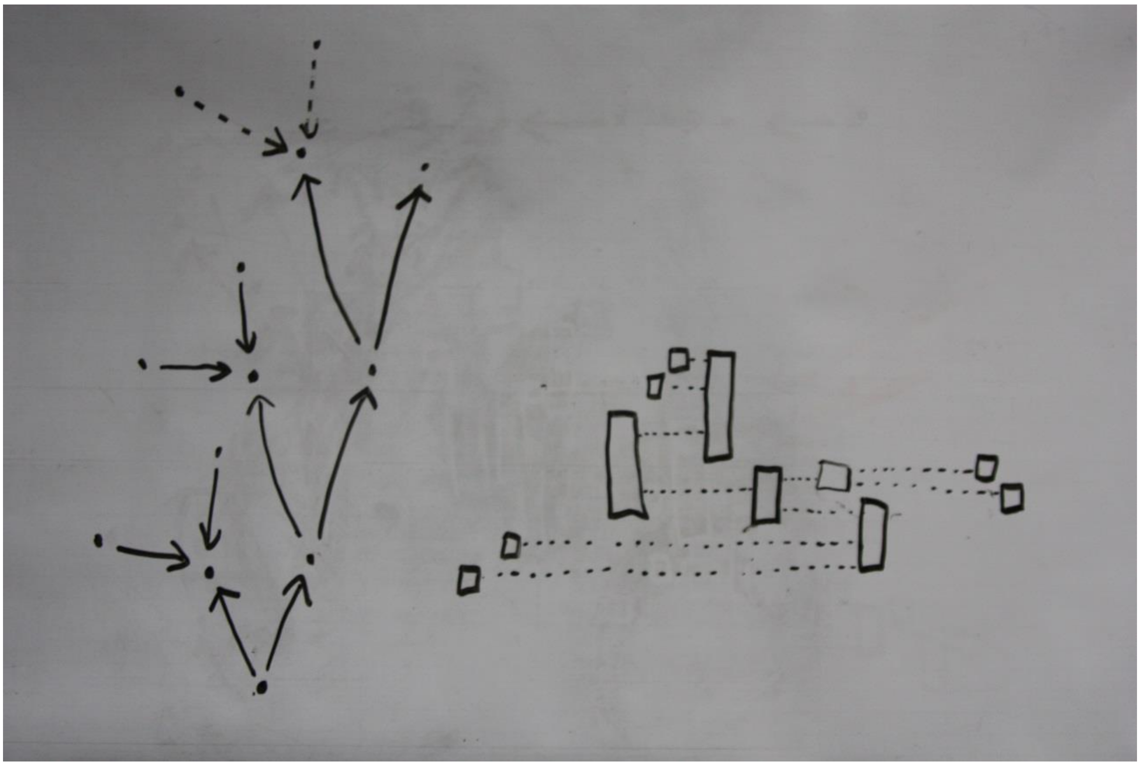
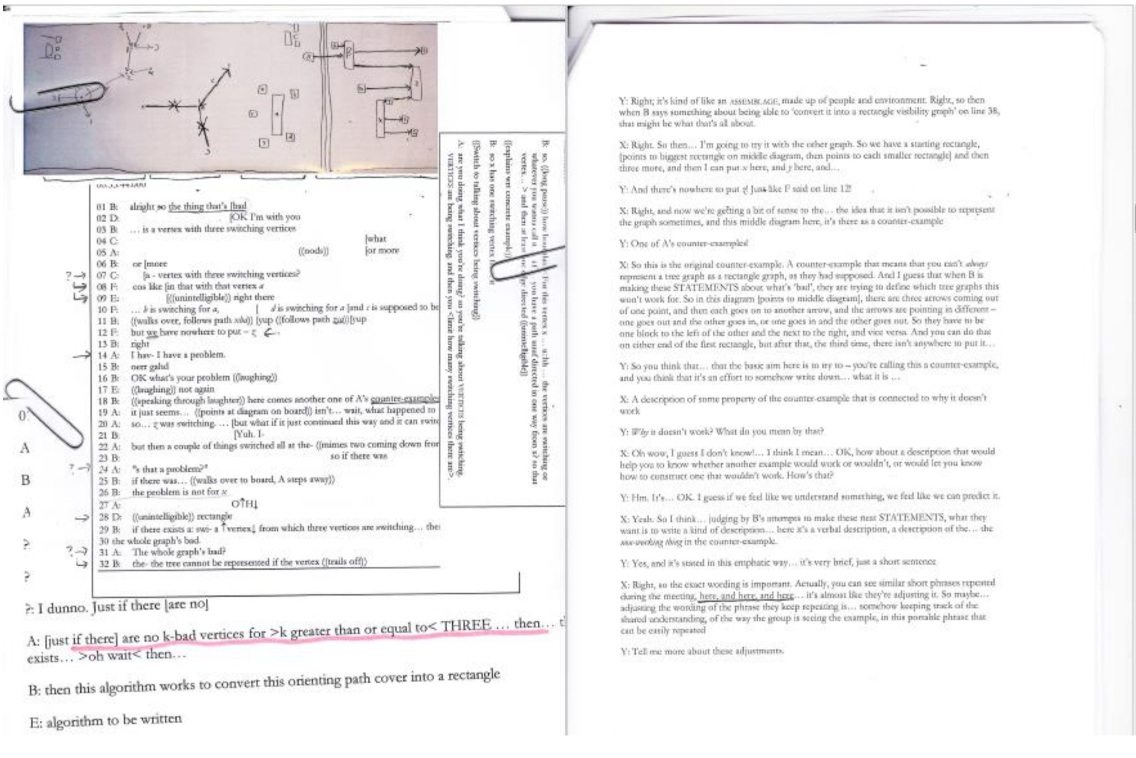


Figure 14. Drawing experiments

Once I felt I had come to the end of what I could come to know by re-watching a single meeting and looking up the words I heard on Google, I met up with one of the original meeting participants to discuss where I'd got to in my understanding. I found that I had been able to get a good understanding of the particular problem they were working on, but not of the overarching goals, the kinds of questions that are asked and answered in the subfield and so that the group were hoping to answer, or the context that had shaped the subfield, the history of constructions and extensions that had led to these questions being asked. My participant was able to give me a sense of the age of the field, of some of the key texts and developments, in short, to give me a map of the story that an insider was able to tell about the community's work in order to explain why the group wanted to achieve the things it wanted to achieve.

The most notable observation that I came back with from the process of analysis was that often what my participants actually *said* was hardly enough for them to understand one another, and yet communication was a success. A participant might walk up to the board, wave at a diagram and say one or two words, and be understood by the other collaborators as proposing a complex adjustment to the ideas being shared. How could this be? The whiteboard behaved almost as a member of the group, or more, a kind of enabling medium that presided over the meeting. As well as this there was the simple fact of the group's full awareness of one another, of their shared aims and history and of the significance of a surprised tone of voice or a groan.

The next step was for me to develop my emerging analysis by entering into a dialogue myself, to step further into the rhythm of proposal and response at a finer grain, working at and pointing to a whiteboard, uttering quick corrections and queries. I wrote a dialogue for myself to perform (see the spread from the script in Figure 15; these pages will be reproduced fully later on), playing two parts and going through the process of explanation and exploration in the space between the speakers. Character X was the one doing the bulk of the explaining, the one who had attended the meeting and had initial ideas about what was going on (although both were working together to find their way through the material). Character Y took on the role of asking questions and coming up with challenges, asking the questions that kept both of them attending to the way they were structuring their understanding of the material, according to the *first time through* approach. In writing the dialogue, I struggled to include anything quite like the vagueness and half-sentences that I was seeing in my observed material, still having trouble believing that real conversation, coherent though it always feels, on recording and examination is a collection of abandoned words and non-answers.


01 B: alright so the thing that's [bad
02 D: [OK I'm with you
03 B: … is a vertex with three switching vertices
04 C: [what
05 A: ((nods)) [or more
06 B: or [more
07 C: [a - vertex with three switching vertices?
08 F: cos like [in that with that vertex a
09 E: [((unintelligible)) right there
10 F: … b is switching for a, [d is switching for a [and c is supposed to be
11 B: ((walks over, follows path xba)) [yup ((follows path zya))[yup
12 F: but we have nowhere to put – z
13 B: right
14 A: I hav- I have a problem.
15 B: oerr galul
16 B: OK what's your problem ((laughing))
17 E: ((laughing)) not again
18 B: ((speaking through laughter)) here comes another one of A's counter-examples
19 A: it just seems… ((points at diagram on board)) isn't… wait, what happened to
20 A: so… z was switching. … [but what if it just continued this way and it can switc
21 B: [Yuh. I-
22 A: but then a couple of things switched all at the- ((mimes two coming down from
23 B: so if there was
24 A: °s that a problem?°
25 B: if there was… ((walks over to board, A steps away))
26 B: the problem is not for x
27 A: O↑H↓
28 D: ((unintelligible)) rectangle
29 B: if there exists a: swi- a ↑vertex↓ from which three vertices are switching… the
30 the whole graph's bad
31 A: The whole graph's bad?
32 B: the- the tree cannot be represented if the vertex ((trails off))

?: I dunno. Just if there [are no]

A: [just if there] are no k-bad vertices for >k greater than or equal to< THREE … then… t
exists… >oh wait< then…

B: then this algorithm works to convert this orienting path cover into a rectangle

E: algorithm to be written

Y: Right; it's kind of like an ASSEMBLAGE, made up of people and environment. Right, so then when B says something about being able to 'convert it into a rectangle visibility graph' on line 38, that might be what that's all about.

X: Right. So then… I'm going to try it with the other graph. So we have a starting rectangle, [points to biggest rectangle on middle diagram, then points to each smaller rectangle] and then three more, and then I can put x here, and y here, and…

Y: And there's nowhere to put z! Just like F said on line 12!

X: Right, and now we're getting a bit of sense to the… the idea that it isn't possible to represent the graph sometimes, and this middle diagram here, it's there as a counter-example

Y: One of A's counter-examples!

X: So this is the original counter-example. A counter-example that means that you can't *always* represent a tree graph as a rectangle graph, as they had supposed. And I guess that when B is making these STATEMENTS about what's 'bad', they are trying to define which tree graphs this won't work for. So in this diagram [points to middle diagram], there are three arrows coming out of one point, and then each goes on to another arrow, and the arrows are pointing in different – one goes out and the other goes in, or one goes in and the other goes out. So they have to be one block to the left of the other and the next to the right, and vice versa. And you can do that on either end of the first rectangle, but after that, the third time, there isn't anywhere to put it…

Y: So you think that… that the basic aim here is to try to – you're calling this a counter-example, and you think that it's an effort to somehow write down… what it is …

X: A description of some property of the counter-example that is connected to why it doesn't work

Y: *Why* it doesn't work? What do you mean by that?

X: Oh wow, I guess I don't know!… I think I mean… OK, how about a description that would help you to know whether another example would work or wouldn't, or would let you know how to construct one that wouldn't work. How's that?

Y: Hm. It's… OK. I guess if we feel like we understand something, we feel like we can predict it.

X: Yeah. So I think… judging by B's attempts to make these neat STATEMENTS, what they want is to write a kind of description… here it's a verbal description, a description of the… the *non-working thing* in the counter-example.

Y: Yes, and it's stated in this emphatic way… it's very brief, just a short sentence

X: Right, so the exact wording is important. Actually, you can see similar short phrases repeated during the meeting, here, and here, and here… it's almost like they're adjusting it. So maybe… adjusting the wording of the phrase they keep repeating is… somehow keeping track of the shared understanding, of the way the group is seeing the example, in this portable phrase that can be easily repeated

Y: Tell me more about these adjustments.

Figure 15. A spread from the script. Original in colour.

At the point of performance, I found that my conversation produced a table full of scraps, of the papers and diagrams and books that I wanted to point at and flip through to show my imaginary conversation partner what I meant, to allow her to see for herself and be convinced. The process of explaining and discussing, I found, was one of pointing, sketching, annotating, physical engagement with the diagrams at the centre of the discussion giving shape to the conversation. A simple written summary is barely enough for an expert to get to grips with a piece of mathematical work, as we can see in the fact that mathematicians go to one another in person to seek mathematical knowledge (see

Interlude 2). As such, the reader of this thesis is offered a multi-modal introduction to the mathematics.

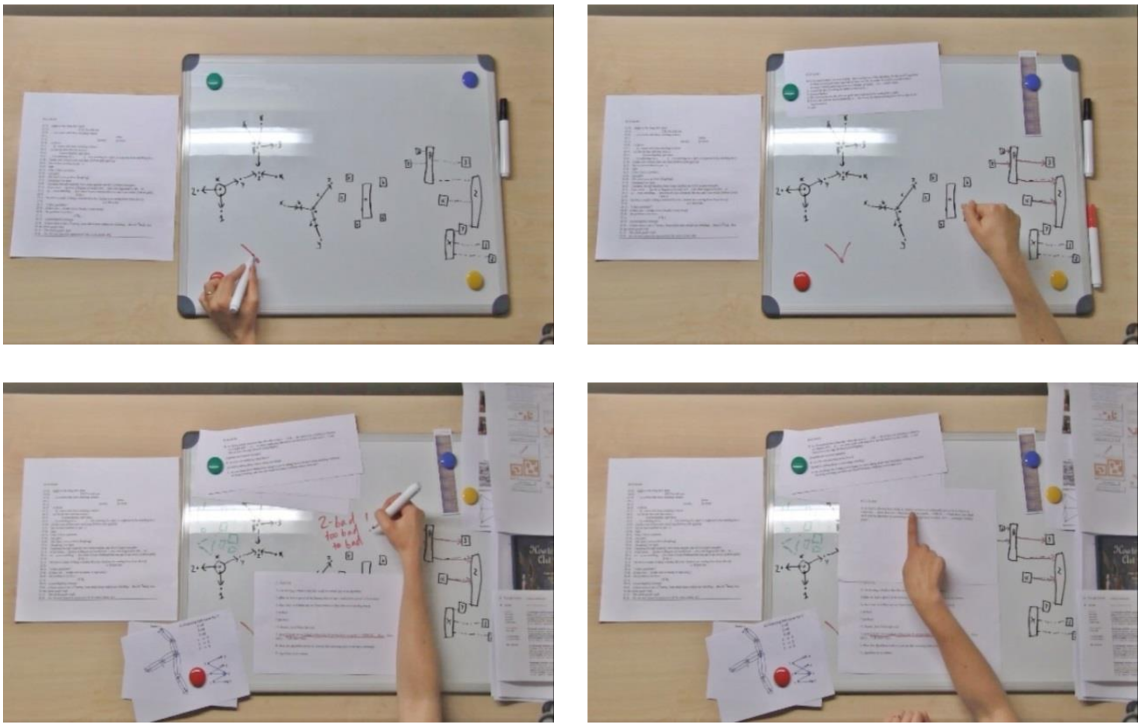


Figure 16. Pointing and writing in the course of my recreated conversation. Original in colour.

This performed conversation is documented by the script in 3.3.2. Script for a conversation. A reader of this thesis can follow this conversation as a beginner's guide through the mathematics, explaining in detail the problem that the group are working on and how I as an observer came to understand it; this text will also serve as an introduction to the observations that will be discussed in detail in the Written Discussion, situating each of these in the conversation and in the group's work.

3.2. Breaching experiments: bad diagramming

The *breaching experiments* associated with this piece of work concern the norms of diagramming.

It is usual to label a diagram with visually distinct shapes with known verbal equivalents, like the letters *a*, *b*, *c*, or numbers 1, 2, 3, so that each has a distinct name that can be written or spoken, and can operate across different media and senses. A transgressive or 'bad' diagram can be constructed by labelling in a way that does not meet these desiderata: failing to distinguish, failing to be nameable.

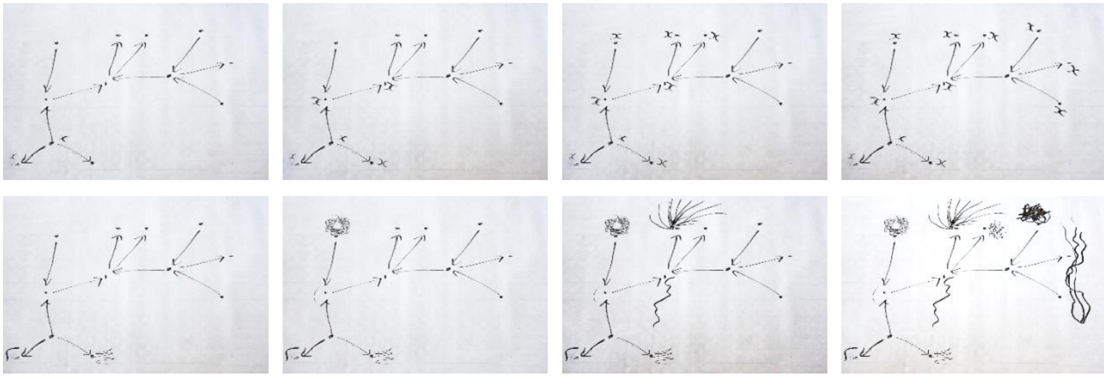


Figure 17. 'Bad' labelling

An experiment in media was to 'name' a vertex with a type of movement, one that could be imitated and replicated and that can even leave a visible trace. The marks are made by rolling, jiggling, stabbing and so on. They leave what is almost a distinct visible trace.

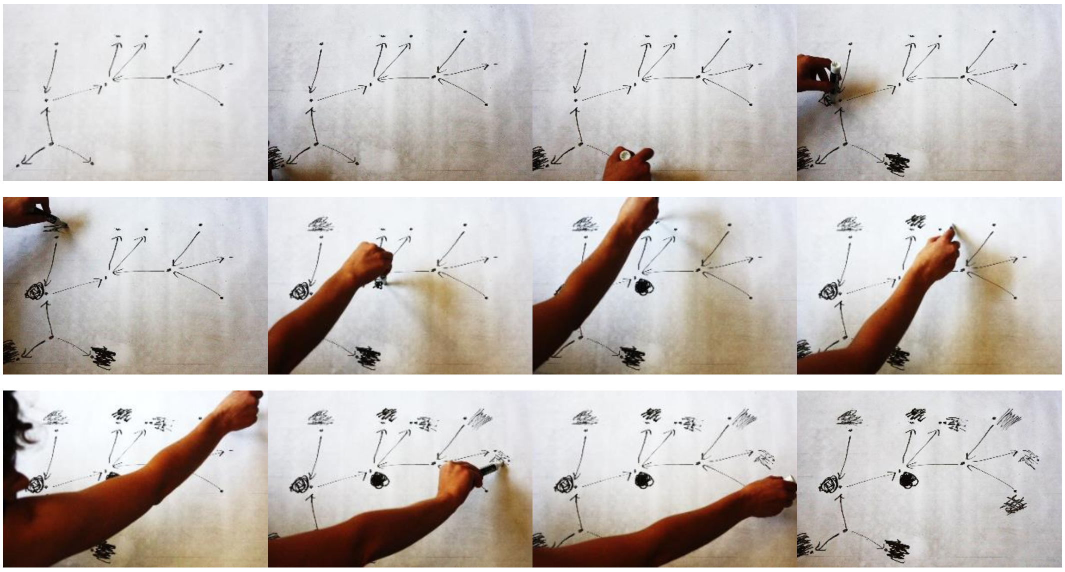


Figure 18. Labelling by movements. Original in colour.

A more interesting question is how to produce a transgressive diagram that uses the same means as the originals. If I adhere to the forms—dots, arrows— and basic rules of the original then the challenge is to see what additional norms exist that can be broken.

An unspoken norm is that diagrams are kept to what is digestible and necessary, a clear, minimal representation with the necessary features and no more. What happens when a tree expands beyond what the board can hold?

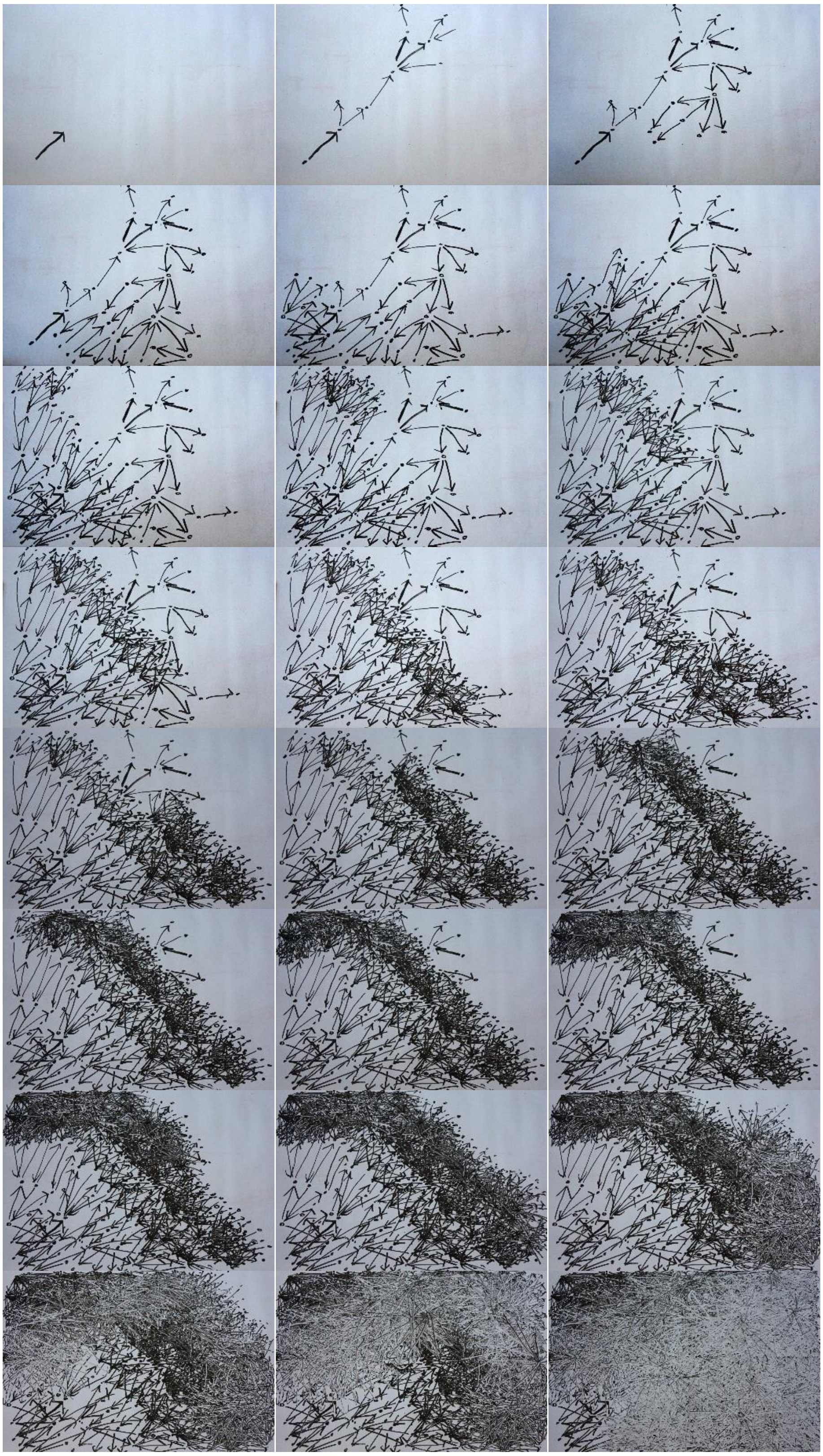


Figure 19. Excessive trees. Original in colour.

3.2.1. Seagull visibility graph

The perspective-taking involved in construction of an RVG has mathematicians imaginatively inhabiting a set of forms to evaluate what is 'seen' by each one. In this way even the utterly inanimate form of a rectangle is imbued with agency for a glittering moment, an example of the wild anthropomorphosis enacted on abstractions in the speech of mathematicians.

This drawing applies the same process to living beings, a freeze-frame that nonetheless changes the feeling of possibility of perspective-taking, among a flock of mutually aware beings.

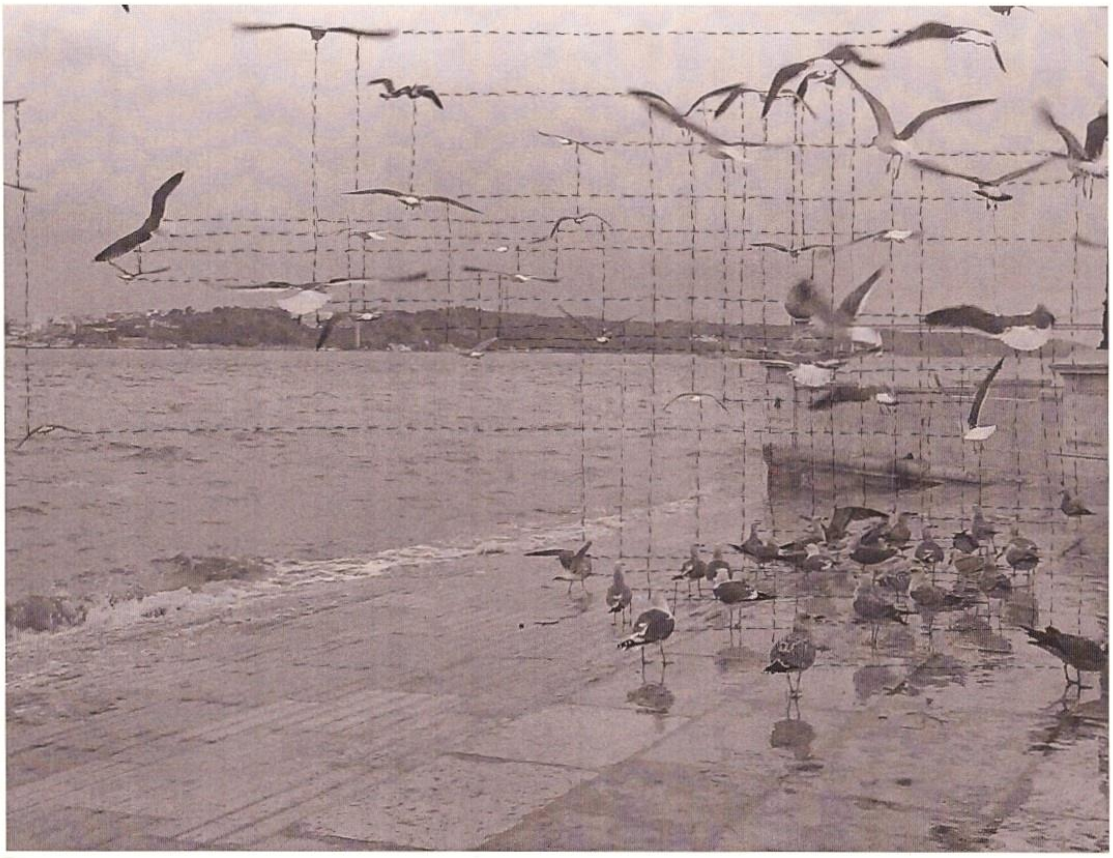


Figure 20. Seagull visibility graph. Original in colour.

Turning this RVG back into a tree graph creates a chaotic mess, one shaped not by the requirements of thought experiment but the contingent arrangement of a flock of seagulls at a moment in time one day in Lisbon, Portugal.

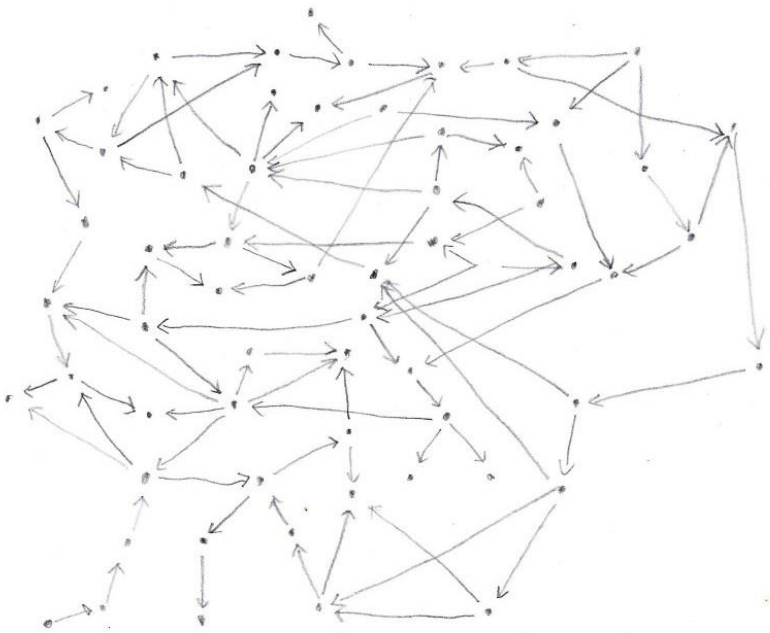


Figure 21. Seagull tree

This is a diagram whose logical structure is governed by the same rules as those used by the group, but whose purpose is quite different as, I would say, is the aesthetic of its organisation. There is reason to think that the broader, more subtle account of intention recognition put forward by relevance theory renders it more realistic to suppose that such recognition is involved in art appreciation, and that we readily read details of presentation to recognise intentions that point to subtle experiences rather than practical conclusions (Pignocchi, 2018; McCallum *et al.*, 2019). I return to the view on art espoused by Rancière, Noë and so on that was summarised in Chapter 1; if a proper goal of artistic practice is to test and transgress the rules of everyday interaction then perhaps simple recognition of a kind of deliberate weirdness, a kind of intentional uselessness, is what allows us to recognise that such boundary-testing is at play. But perhaps this diagram is not such a divergence after all; Catherine Elgin places art very much alongside science and mathematics for its experimental exemplification, its urge to exhibit fictions to round out the possibilities of our world. The group, after all, drew and redrew diagrams with boundary-testing as an explicit aim, and I only opted to step in a different direction. The rules of diagramming are already examined and contested all the time in mathematics; it is the rules of communication that are difficult to perceive and alter.

3.2.2. Dialogue without shared aims

My final experiment was with dialogue itself; an attempt to write a dialogue in which the usually assured premise of mutual aims is absent. The fictional interlocutors discuss a project and seem to be perpetually pulling in opposite directions, holding exactly opposite aims.

A: …so we want to get as close as we can to the point where these two elements are equal
B: well, ok, it would be really good if we got them as far apart as possible
A. so then they'll diverge
B: which is exactly what we want
A: yeah. <pause>
B: <pause> so if we take element one and… <pause>
A: …expand it as much as possible
B: can we do th-
A: -I mean blow it up with the thing we used before
B: oh yeah yeah great
A: then we'll want -to
B: - to really chop it right back and get it to <writing> do something like this to get the smallest value
A: ok so that looks quite good, that's looking nice. But how about element two, that's more of a problem
B: Yeah it's pretty similar right now, we need more divergence. ok
A: OK so. Let's say we take element two and … so maybe we can use one of those folding techniques
B: <excited> yeah! Yeah, that should help!
A: ok and that gives us… <writing>
B: Oh, ok. That's not really very good, that's way too small
A: no, ok, hang on, what if I raise it to the power of… that should get it a bit smaller…
B: yeah! That's better
A: … get us pretty close. ok so when we've applied that... then we... get something that looks like... <writing> and then, ok, yes. We're within, like, maybe .5 of element 1. Not at all bad
B: yep so then we apply it again
A: and then we're at maybe a... what's that, <writing> bigger by a scale of 7 or something
B: 72, yeah. OK great! This isn't bad, we're getting somewhere.
A: we still haven't got quite to equality, but these values aren't bad
B: -but we're getting somewhere, huh. So I suggest we think about what we can do to element 2
A: yeah element 2 is where we have the most play
B: and see how much divergence we can get

Figure 22. Script for a dialogue without shared aims

The utter impossibility of the speakers failing to recognise this as a problem in their conversation made the dialogue incredibly difficult to write. It gives the impression that the two must be in different rooms, on different planets, failing to listen properly to one another. Otherwise how could they believe themselves to be agreeing?

3.3. Evidence

3.3.1. Transcripts

In the transcription, I have adopted conventions taken from "Transcription Conventions in Conversation Analysis" from the *Handbook of Classroom Discourse and Interaction* (Numarkee, 2015) (see Abbreviations and conventions used).

```
00.33.44.000

01  B:    alright so the thing that's [bad
02  D:                                [OK I'm with you
03  B:    … is a vertex with three switching vertices
04  C:                                                            [what
05  A:                                              ((nods))  [or more
06  B:     or [more
07  C:        [a - vertex with three switching vertices?
08  F:     cos like [in that with that vertex a
09  E:              [((unintelligible)) right there
10  F:    … b is switching for a,          [    d is switching for a [and c is supposed to be switching for a
11  B:    ((walks over, follows path xba)) [yup ((follows path zca))[yup
12  F:    but we have nowhere to put – z
13  B:    right
14  A:    I hav- I have a problem.
15  B:    oerr gahd
16  B:    OK what's your problem ((laughing))
17  E:    ((laughing)) not again
18  B:    ((speaking through laughter)) here comes another one of A's counter-examples
19  A:    it just seems…  ((points at diagram on board)) isn't… wait, what happened to this… so
20  A:    so… z was switching. … [but what if it just continued this way and it can switch ((follows path))
21  B:                            [Yuh. I-
22  A:    but then a couple of things switched all at the- ((mimes two coming down from above))
23  B:                                                          so if there was
24  A:    °s that a problem?°
25  B:    if there was… ((walks over to board, A steps away))
26  B:    the problem is not for x
27  A:                               O↑H↓
28  D:    ((unintelligible)) rectangle
29  B:    if there exists a: swi- a ↑vertex↓ from which three vertices are switching… then it's ↑bad↓. then
30  the whole graph's bad.
31  A:    The whole graph's bad?
32  B:    the- the tree cannot be represented if the vertex ((trails off))
```

Figure 23. Transcript of the main excerpt

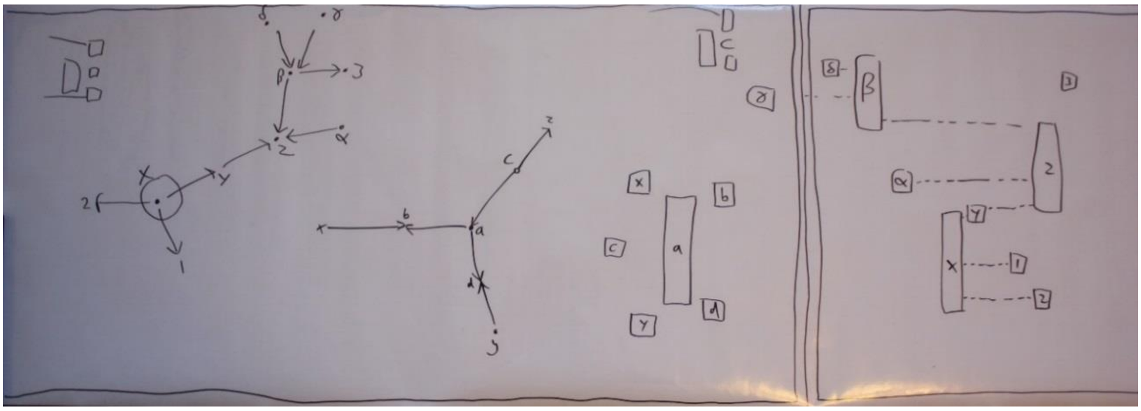


Figure 24. Board notes as they were for the duration of the excerpt. Original in colour.

Though the work done in the meeting has not yet been published, at the close of the meeting there was a sense of resolution at the moment when A gave a one-line summary of the conclusion they had reached:

01.13.05.750

33 A: I'm starting to believe that this could be written [up as an algorithm
34 B: [that we have a proof of the lemma, that we have outlined the proof of the lemma
35 A: that- that- so if there are no 3-bad vertices is that what we're [calling them]
36 ?: or four
37 ?: [or four]
38 ?: I dunno. Just if there [are no]
39 A: [just if there] are no k-bad vertices for >k greater than or equal to< three … then… there exists… >oh wait< then…
40 B: then this algorithm works to convert this orienting path cover into a rectangle
41 E: algorithm to be written

Figure 25. Excerpt at the close of the meeting

This one-line summary is introduced as a means to write up their work as an algorithm, or even the proof of a lemma. At this moment, then, what the group has in mind is the prospect of publication, of moving toward the 'front'. The summary could have another purpose, as well; at the beginning of this meeting, the group looked back over various materials from their shared Dropbox to reacquaint themselves with the topic they were studying (Figure 26). In a similar way, this one-line summary could be preserved in the participants' notes, or even in a similar Dropbox document, as a record of the work completed so far, a succinct summary. In Figure 26, a note has been added that it 'seems hard to add these paths as "rows" in an RVG; it is precisely this problem that the group goes on to addres in the research meeting.

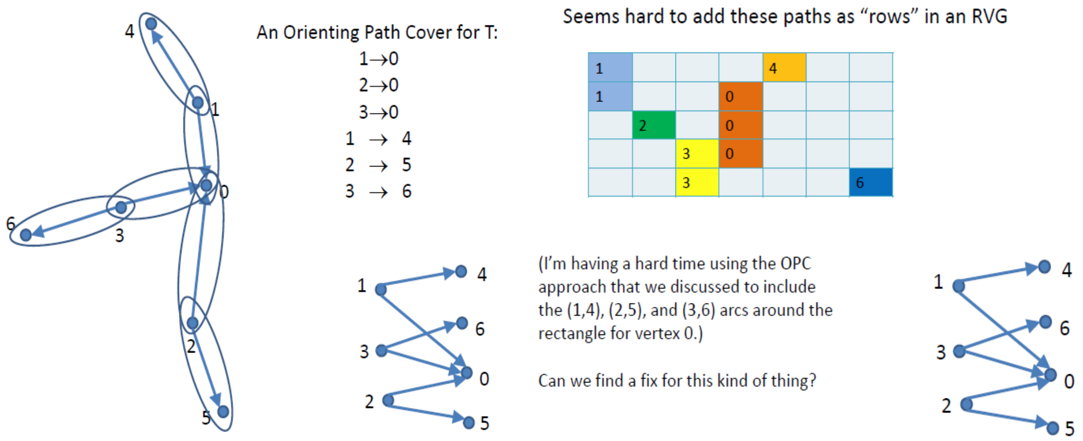


Figure 26. Materials from the Dropbox used by the meeting collaborators to reacquaint themselves with the topic they were studying. Original in colour.

The final line, then, can be summarised as follows:

> *if there are no k-bad vertices for k greater than or equal to three then this algorithm works to convert this orienting path cover into a rectangle*

The term k-bad needs a little explanation. In the conversation immediately preceding this summary, '1-bad' or '2-bad' are used to mean 'a vertex with one switching vertex following it' and 'a vertex with two switching vertices following it', as in the following excerpt:

01.10.30.000

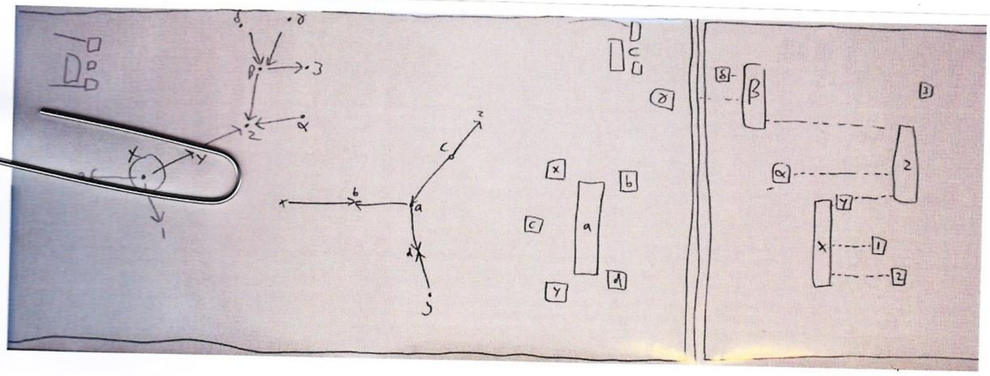
42 B: yeah so F's… POINT … o:f if you started a- if you start with a 2-bad vertex - you
43 ((unintelligible)) out from there, the vertices <can have only 1-bad vertices>((unintelligible))
44 cause if they have- they can have only one switching vertex a:fter that so if you have two
45 switching vertices after that and you have one switching vertex before…

Figure 27. Excerpt near to the close of the meeting

So we can add the following helpful definition to the closing line.

> *if there are no k-bad vertices for k greater than or equal to three then this algorithm works to convert this orienting path cover into a rectangle*
>
> *k-bad meaning a vertex followed by k switching vertices*

3.3.2. Script for a conversation



00.33.44.000

01 B: alright so the thing that's [bad
02 D: [OK I'm with you
03 B: … is a vertex with three switching vertices
04 C: [what
05 A: ((nods)) [or more
06 B: or [more
07 C: [a - vertex with three switching vertices?
08 F: cos like [in that with that vertex *a*
09 E: [((unintelligible)) right there
10 F: … *b* is switching for *a*, [*d* is switching for *a* [and *c* is supposed to be switching for *a*
11 B: ((walks over, follows path *xba*)) [yup ((follows path *zca*))[yup
12 F: but we have nowhere to put – *z*
13 B: right
14 A: I hav- I have a problem.
15 B: oerr gahd
16 B: OK what's your problem ((laughing))
17 E: ((laughing)) not again
18 B: ((speaking through laughter)) here comes another one of A's counter-examples
19 A: it just seems… ((points at diagram on board)) isn't… wait, what happened to this… so
20 A: so… *z* was switching. … [but what if it just continued this way and it can switch ((follows path))
21 B: [Yuh. I-
22 A: but then a couple of things switched all at the- ((mimes two coming down from above))
23 B: so if there was
24 A: °s that a problem?°
25 B: if there was… ((walks over to board, A steps away))
26 B: the problem is not for *x*
27 A: O↑H↓
28 D: ((unintelligible)) rectangle
29 B: if there exists a: swi- a ↑vertex↓ from which three vertices are switching… then it's ↑bad↓. then
30 the whole graph's bad.
31 A: The whole graph's bad?
32 B: the- the tree cannot be represented if the vertex ((trails off))

Original in colour

3.1 First time through: Script for a Conversation (after Imre Lakatos and Douglas Hofstadter)

This conversation takes place in an imaginary classroom, with a variety of relevant printouts scattered across the tables, as well as a laptop. The speakers, X and Y, are non-mathematicians, and are students of linguistics and communication. X has made a recording sitting in on a meeting held between a group of collaborating mathematicians. X and Y are here to discuss a certain excerpt from that meeting, attempting to describe it as a *first time through*[1], which means, rather than thinking of the mathematical objects as the driving force behind what's going on, to shift the focus to what the subjects do to shape themselves into *mathematicians* and *collaborators*, and to produce the work of that evening—the construction of a STATEMENT—as stateable, shareable *work on mathematics*. The speakers also decide to note which resources are needed, how, and when during their process of coming to understand the excerpt.

In their conversation the two are trying to make sense of the excerpt they have from the ground up; the audience has the sense that it is a genuine process of collaborative discovery. The two speakers are at ease, are not afraid to share their uncertainties, and it's clear that although they find the *first time through* approach tricky, they are making an effort to stay on track. They have before them a transcript of the excerpt as well as a reproduction of the board notes as they were at the time of the excerpt. They also have transcripts of other relevant sections of the meeting, printed papers, reproductions of the subjects' movements, and some other printed resources, as well as access to Google, and the points at which these are referenced are noted as clearly as possible in the text. At one point the presenters are joined by a guest, who was one of the participants in the original meeting.

A Layperson's Overview

X: So OK, we're talking about a meeting today. And to set the scene, this was all happening in a … in a room in a university, and it's early evening, and they've got snacks and drinks and, uh, everybody's kind of… talking and laughing a lot. And everybody's more or less oriented toward-- there's a whiteboard on one side of the room, and people stand up and write on it and talk about the writing. It seems like… there are two of them who are doing a lot of the talking, and tend to be next to the board so everybody's facing them-

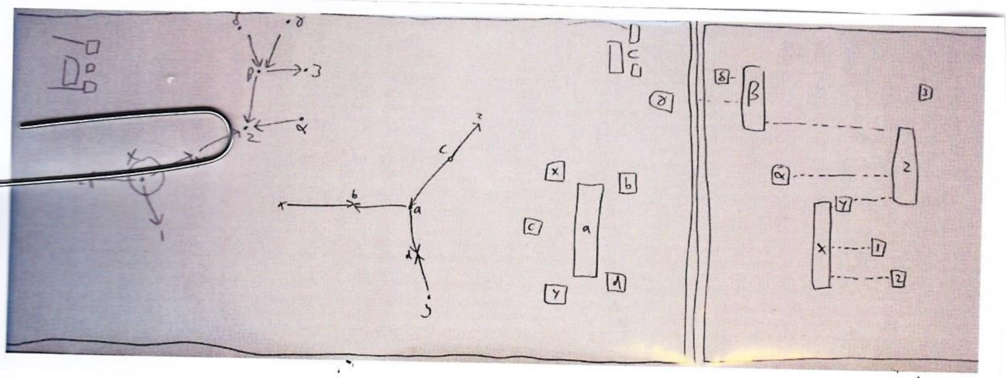
Y: So does it seem like they're… in charge?

X: I mean, not really! People interrupt them and talk over them so it doesn't seem like they're afforded more… power or anything. And it's not all them giving the explanation, you can see F explaining something here (line 08), and I think that's also what's happening with D there (28), and E there (09). And you get C asking a question here (07), and A kind of asking questions to the room here (24) and to B here (31), and they get answered by B and by F, so it's pretty… sort of mixed. I guess not quite equal though. I just get the feeling that… maybe these two people, A and B, they have more to say this time. But everyone's chipping in, so

Y: Right, so there's a kind of… there isn't a clear agreement that a particular person should be listened to most, but instead it's something like, if you feel you can explain something or have a query, you feel empowered by that contribution to speak, something like that? This is tricky. So we're supposed to be describing this as the *first time through*, right?

[1] A technique from ethnomethodology, in which an attempt is made to describe what is going on as though it is the first time it has ever happened (Garfinkel *et al.*, 1981)

Original in colour



00.33.44.000

01 B: alright so the thing that's [bad
02 D: [OK I'm with you
03 B: … is a vertex with three switching vertices
04 C: [what
05 A: ((nods))
06 B: or [more [or more
07 C: [a - vertex with three switching vertices?
08 F: cos like [in that with that vertex a
09 E: [((unintelligible)) right there
10 F: … b is switching for a, [d is switching for a [and c is supposed to be switching for a
11 B: ((walks over, follows path xba)) [yup ((follows path zca)) [yup
12 F: but we have nowhere to put – z
13 B: right
14 A: I hav- I have a problem.
15 B: oerr gahd
16 B: OK what's your problem ((laughing))
17 E: ((laughing)) not again
18 B: ((speaking through laughter)) here comes another one of A's counter-examples
19 A: it just seems… ((points at diagram on board)) isn't… wait, what happened to this… so
20 A: so… z was switching. … [but what if it just continued this way and it can switch ((follows path))
21 B: [Yuh. I-
22 A: but then a couple of things switched all at the- ((mimes two coming down from above))
23 B: so if there was
24 A: °s that a problem?°
25 B: if there was… ((walks over to board, A steps away))
26 B: the problem is not for x
27 A: O↑H↓
28 D: ((unintelligible)) rectangle
29 B: if there exists a: swi- a ↑vertex↓ from which three vertices are switching… then it's ↑bad↓. then
30 the whole graph's bad.
31 A: The whole graph's bad?
32 B: the- the tree cannot be represented if the vertex ((trails off))

Original in colour

X: Yes. Which is hard! We're supposed to try to look at what's actually being done by the people in this room as though it's the first time anyone's ever done this, and try to set aside our preconceptions about what's happening wherever possible.

Y: Yes, that's perhaps even harder… let's see how it goes. It would be nice to try to make a note of everything we find ourselves using to make sense of what's going on, like-- every time we notice ourselves using a new resource, like other papers or things we remember about mathematics.

X: Yes, great. And the group uses different things during the meeting too, like papers and diagrams in their shared Dropbox, and what's on the board, and so on. So we'll start out with this text transcript and our recreation of what was on the whiteboard. So they've drawn some things on the board when this clip starts, there are numbers and letters and arrows; and boxes. They don't actually write anything during this excerpt, so it's just like this. So it starts off with B making a STATEMENT, 'the thing that's bad, is…'

Y: So what's B talking about? 'The thing that's bad…' Is this some kind of Sacksian 'IT'? I'm thinking of the kind described by Garfinkel et al in the paper about pulsar discovery, and attributed to an unpublished lecture by Oliver Sacks before he died – where an 'it' is mentioned, recognised and understood in the course of a conversation while it's still very much an unknown thing, and known to be an unknown thing, and so the word is sort of vague and undefined but is still this very essential tool for the speakers to be able to talk about something and thus come to define it

X: Huh, that's interesting. They do seem to be finding a way to describe the 'thing that's bad', working toward a STATEMENT of a definition maybe. Well, so B says this in the first line, 'the thing that's bad is a vertex with three switching vertices,' and then here (line 07) C repeats it wanting clarification…

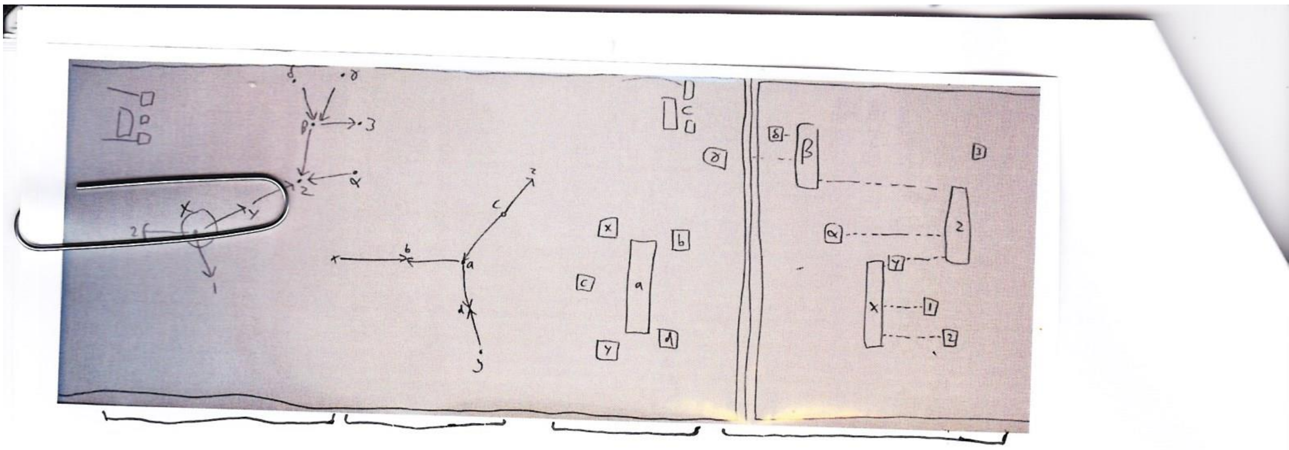
Y: Yeah! C just repeats it in a questioning tone, just reframes the same thing as a question, but we can guess that C's asking B to expand, it's like saying, 'tell me more about this'. So then F starts explaining, and – here (line 12) F is using the word 'we', saying 'we have nowhere…'

X: Right, so we get this kind of collaborative feel to what they're doing here. And in this animation of the gestures we can see B go up to this picture, this diagram here and follow along. And F is mentioning letters, and B is following the lines that are labelled with those letters. So it's connected with the- with this diagram here in some way.

Y: Oh, and so – F is explaining, like describing something, talking through an example maybe to explain what B means- to exhibit 'what's bad', maybe, to show the badness

X: So it's kind of this team effort, between F and B there's this shared understanding, and they can – they want to explain it in multiple ways at once, with words and waving at the picture, and it's important to them to share their understanding with C. And this understanding, it has to do with this – they want to convey it by talking about this picture, it happens between them and the picture. And each individual needs to have a part in this shared understanding. OK, so we have this sense that in this first section B has proposed some kind of STATEMENT of a definition for 'what's bad', and C asked for clarification, and F and B worked together to sort of… *exhibit* that diagram in the centre, as providing support for that definition. I don't think we're quite able to see how that works, yet.

Y: Agreed. Well, let's look at what happens next.



00.33.44.000

01 B: alright so the thing that's [bad
02 D: [OK I'm with you
03 B: … is a vertex with three switching vertices
04 C: [what
05 A: ((nods)) [or more
06 B: or [more
07 C: [a - vertex with three switching vertices?
08 F: cos like [in that with that vertex *a*
09 E: [((unintelligible)) right there
10 F: … *b* is switching for *a*, [*d* is switching for *a* [and *c* is supposed to be switching for *a*
11 B: ((walks over, follows path *xba*)) [yup ((follows path *zca*))[yup
12 F: but we have nowhere to put – *z*
13 B: right
14 A: I hav- I have a problem.
15 B: oerr gahd
16 B: OK what's your problem ((laughing))
17 E: ((laughing)) not again
18 B: ((speaking through laughter)) here comes another one of A's counter-examples
19 A: it just seems… ((points at diagram on board)) isn't… wait, what happened to this… so
20 A: so… *z* was switching. … [but what if it just continued this way and it can switch ((follows path))
21 B: [Yuh. I-
22 A: but then a couple of things switched all at the- ((mimes two coming down from above))
23 B: so if there was
24 A: °s that a problem?°
25 B: if there was… ((walks over to board, A steps away))
26 B: the problem is not for *x*
27 A: O↑H↓
28 D: ((unintelligible)) rectangle
29 B: if there exists a: swi- a ↑vertex↓ from which three vertices are switching… then it's ↑bad↓. then
30 the whole graph's bad.
31 A: The whole graph's bad?
32 B: the- the tree cannot be represented if the vertex ((trails off))

Original in colour

X: Then things get more complicated, because then we have – so on line 14 here A says, 'I have a problem.' And everything gets kind of chaotic, there's laughing, it's obviously quite informal. And then A walks up to the board and starts pointing at this other diagram [points to diagram on the left]-

Y: Wait, why are you saying 'diagram'? And how can you tell that they're separated out like that?

X: Yeah good question. Uhh… so they just… *look* like diagrams. What does that mean… They're simple, and there are arrows and relationships between elements. I've seen things like this before, called diagrams. And there are labels, you get points, dots, and those have a recognisable character next to them so you can easily refer to them, like they do in these two lines, using the labels. So there are – it's like a network. And there are kind of… four collections of connected things here.

Y: So what do you mean by 'diagram'? What is it you think these images are *doing* here?

X: Umm, I guess they seem kind of like tools for reasoning, like an image that is useful for organising thought, rather than… you know, having an aesthetic purpose. It's something that people can *use*, so you interact with it *as* you're – you're thinking, or you're talking

Y: Or if it were an engine diagram, as you were working on it

X: Right! So they have this functional aspect, and the way they are drawn is supposed to make something easier, somehow.

Y: OK, good. So… so A walks up to the board…

X: Right, and starts pointing at the *other* diagram, waving arms around it, it seems as though A is talking about it, directing action at this diagram. And it sounds like a continuation, not a change of topic.

Y: OK, so B makes this one clear STATEMENT at the beginning. And then A says 'I have a problem'. And this is announced with the same kind of gravity, emphasis, loud voice, waiting for people to stop talking, that kind of thing, so it seems kind of like a statement put on the same level as B's first one, like it's a response maybe. But then A is not pointing at the same diagram.

X: Hm. So A *could* be exhibiting 'badness' in the other diagram, but that doesn't seem right, somehow. It's more like a contradiction.

Y: Right. So … so they're talking about something that somehow exists between the diagrams, perhaps the *bad thing* is *in* one and not in the other, and that's what they want to talk about

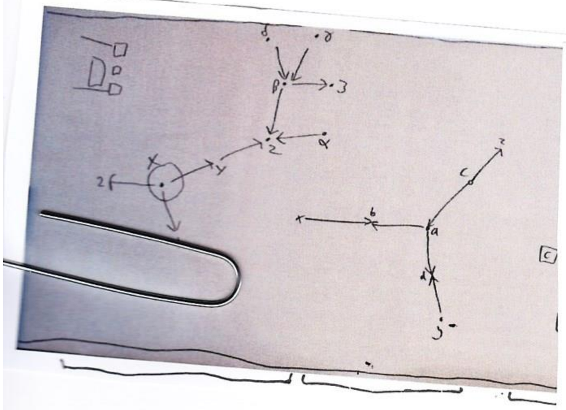
X. Yeah.

Y: OK, and this announcement has quite an impact, right? B says 'oh god', E says 'not again'…

X: Yeah they're acting like it's a disaster! And then they say, what was it… 'here comes another one of A's counter-examples.' But they're not serious about that, they're kind of doing this cartoon 'oh god' response, like they pretend they hate A for it but really this kind of setback is just part of the usual course of things

Y: Setback? Why that?

Original in colour



ique document, on file at the Center for History and Philosophy of Physics at the American Institute of Physics, was made available for our examination.[4] The tape was transcribed by us using the conventions of conversational analysis.

Our question was: 'What does the optically discovered pulsar consist of as Cocke and Disney's night's work?' The tape and transcript permitted us to treat some relevancies that are not otherwise available in science studies:

(1) That the discovery as their night's work had the property of 'first time through'.

(2) The local historicity of the night's collection of observations.

(3) The quiddity of their night's work.[5]

We tried to respect these properties of their night's work in asking what their discovery could be and so what we entertained their discovery to consist of may seem strange. With the hope of making it plain we begin by characterizing their discovery with the metaphoric use of a 'gestalt theme'. Their discovery and their science consists of astronomically 'extracting an animal from the foliage'. The 'foliage' is the local historicity of their embodied shop practices. The 'animal' is that local historicity done, recognized, and understood as a competent methodic procedure. The 'animal' formulates their embodiedly witnessable astronomical competent practices as the transcendental properties of the independent Galilean pulsar.[6] Their science consists of the optically discovered pulsar as the produced practical observability of their ordinary night's work.

00.33.44.000

01 B: alright so the thing that's [bad
02 D: [OK I'm with you
03 B: ... is a vertex with three switching vertices
04 C: [what
05 A: ((nods)) [or more
06 B: or [more
07 C: [a - vertex with three switching vertices?
08 F: cos like [in that with that vertex *a*
09 E: [((unintelligible)) right there
10 F: ... *b* is switching for *a*, [*d* is switching for *a* [and *c* is supposed to be switching for *a*
11 B: ((walks over, follows path *xba*)) [yup ((follows path *zca*)) [yup
12 F: but we have nowhere to put – *z*
13 B: right
14 A: I hav- I have a problem.
15 B: oerr gahd
16 B: OK what's your problem ((laughing))
17 E: ((laughing)) not again
18 B: ((speaking through laughter)) here comes another one of A's counter-examples
19 A: it just seems... ((points at diagram on board)) isn't... wait, what happened to this... so
20 A: so... *z* was switching. ... [but what if it just continued this way and it can switch ((follows path))
21 B: [Yuh. I-
22 A: but then a couple of things switched all at the- ((mimes two coming down from above))
23 B: so if there was
24 A: °s that a problem?°
25 B: if there was... ((walks over to board, A steps away))
26 B: the problem is not for *x*
27 A: O↑H↓
28 D: ((unintelligible)) rectangle
29 B: if there exists a: swi- a ↑vertex↓ from which three vertices are switching... then it's ↑bad↓. then
30 the whole graph's bad.
31 A: The whole graph's bad?
32 B: the- the tree cannot be represented if the vertex ((trails off))

Original in colour

X: Umm, I guess I get the sense that this is B suggesting something, and then A coming back with something that contradicts it. So it- derails what the first person said, it messes up that suggestion, but it advances the group's work in the process because it shows that there was something they weren't seeing.

Y: So we think that – it isn't just a counter-argument, it might really disrupt things

X: Right, it isn't like a difference of opinion.

Y: OK, so then what. Then things go kind of quiet, and A has trailed off, just kind of pointed and trailed off, and there's a silence. And then on line 26 here B says, 'the problem is not for *x*', and A says OH! Like, it seems like there's some revelation, like B has talked A into something.

X: Yes! Yes, that seems right, like that… saying 'not for *x*' is enough to completely change A's understanding somehow. But I don't yet see how! OK, well then what. Then B repeats the statement-

Y: Why does it seem like a statement?

X: Mm, I dunno, the phrasing has a kind of formality to it, starting with 'if there exists' like that. It sounds like something you'd read written down in a paper, whereas everything else is so much more informal. It's kind of a switch in tone.

Y: Oh, and you think the same is true of the first line?

X: Yeah, that seems less formal, I guess. There's a little of that in line 3, that it's so… specific, so technical. But also in line 1 when it's introduced B starts off with 'alright so', it's kind of like – fanfare

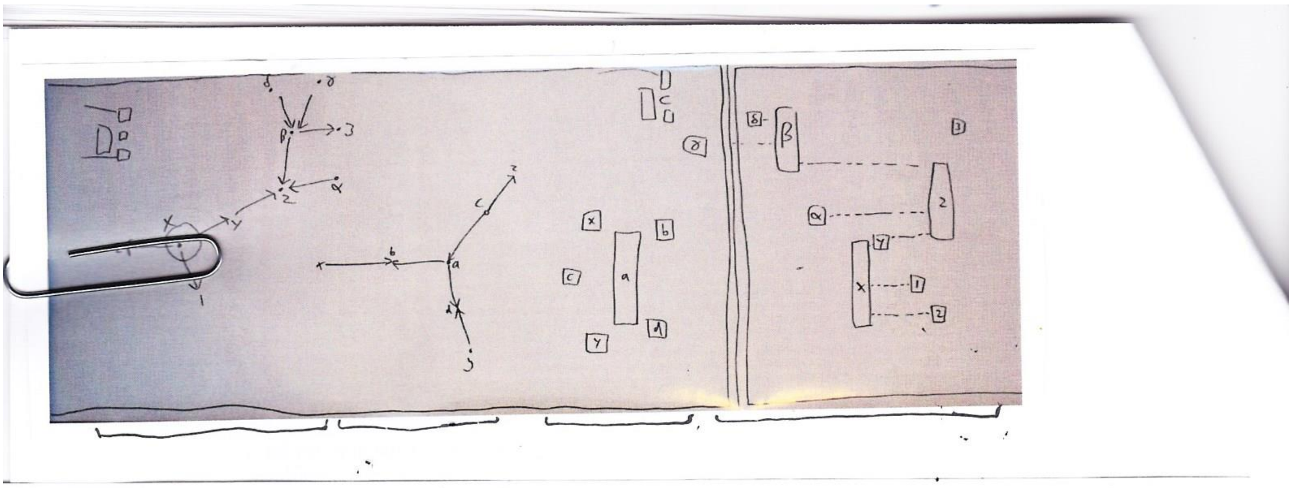
Y: Sure. And this is perhaps… in that Garfinkel *et al.* paper (1981) they talk about identifying what the group does to extract the animal from the foliage, which roughly speaking is extracting 'the *scientific work*' from their various doings that night. So in that paper, they argue that atemporal properties of the pulsar were still established as an exhibitable thing by people doing things with equipment one night. So here, we want to see what they do to extract the *mathematics* from their exchanges, sketches and so on, the *thing with wider use to the community*, that's produced according to the requirements and aims of the mathematical world. This fanfare seems like it's signalling something.

X: OK, so- maybe we should see what they work toward here, and what principles are employed. OK, so some general thoughts. They're talking about how to define what's 'bad', I think, especially B's lines at the beginning and the end. And then right at the end, A repeats what B has said, asking for an explanation, and B expands that 'bad', B replaces it with 'the tree cannot be represented'. So I think maybe when they talk about what's 'bad', they're talking about cannot be represented, that there's a representation that you can't build for those cases.

Y: Yeah, that makes sense. So one part of what they're doing is… the group wants to be able to represent the tree, and not being able to do so is 'bad'.

X: Right. And in another sense their work is to find a STATEMENT of a definition for what's 'bad'.

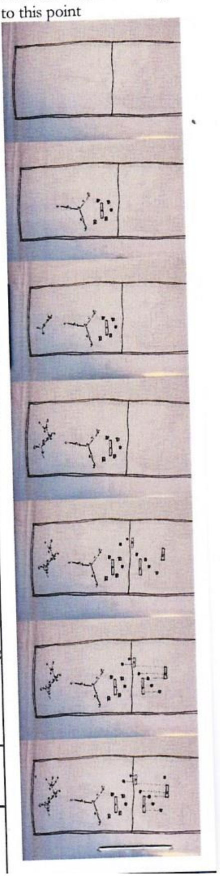
Y: OK. So what are they… talking about here?



00.33.44.000

01 B: alright so the thing that's [bad
02 D: [OK I'm with you
03 B: … is a vertex with three switching vertices
04 C: [what
05 A: ((nods)) [or more
06 B: or [more
07 C: [a - vertex with three switching vertices?
08 F: cos like [in that with that vertex *a*
09 E: [((unintelligible)) right there
10 F: … *b* is switching for *a*, [*d* is switching for *a* [and *c* is supposed to be switching fo
11 B: ((walks over, follows path *xba*)) [yup ((follows path *zca*))[yup
12 F: but we have nowhere to put – *z*
13 D: right
14 A: I hav- I have a problem.
15 B: oerr gahd
16 B: OK what's your problem ((laughing))
17 E: ((laughing)) not again
18 B: ((speaking through laughter)) here comes another one of A's counter-examples
19 A: it just seems… ((points at diagram on board)) isn't… wait, what happened to this… so
20 A: so… *z* was switching. … [but what if it just continued this way and it can switch ((follows pa
21 B: [Yuh. I-
22 A: but then a couple of things switched all at the- ((mimes two coming down from above))
23 B: so if there was
24 A: °s that a problem?°
25 B: if there was… ((walks over to board, A steps away))
26 B: the problem is not for *x*
27 A: O↑H↓
28 D: ((unintelligible)) rectangle
29 B: if there exists a: swi- a ↑vertex↓ from which three vertices are switching… then it's ↑bad↓. th
30 the whole graph's bad.
31 A: The whole graph's bad?
32 B: the- the tree cannot be represented if the vertex ((trails off))

Chronological record of what was on the board up to this point



Original in colour

X: Mm, I have some guesses, but we should mark this as bringing in a bit of general knowledge about the mathematical world, because some of this has to do with the terms they're using.

Basic Mathematical Knowledge

Y: Alright, well we both know a bit from experience about mathematics, from school and so on. Can we bring any of that to bear?

X: Yep OK. They're talking a lot about vertices and graphs and paths. A vertex is kind of like a corner, right?

Y: Right, an endpoint [draws dot], or like where two lines meet [draws corner]. And the mention of graphs, I don't know much about graphs but I guess I could imagine these two diagrams being graphs. And I know there's a field of study called 'graph theory'. So if there are people out there studying graphs and how they work, maybe that's what these people are doing.

X: Yes! Oh and look at the other types of drawing over on the other side, there are two basic types of drawing here, and maybe they're studying how they relate to each other.

Y: Right, but I think we'll have to look a bit at the rest of the meeting to understand that. What else do we have for mathematical terminology?

X: Oh, 'counter-example'! On line 18. A counter-example is normally something that you… that stands against something you're arguing for, that contradicts what you're saying by the fact that it exists. So when B says that 'here comes a… counter-example'… maybe that means B's imagining that A is about to come out with an example that contradicts B's original STATEMENT. And it … it kind of sounds like A has played that role before, they say 'another one of A's counter-examples'.

Y: Ooh, that's helpful. And thinking about examples, and counter-examples, might help us to make sense of what the diagrams are there to do, if they're both being discussed in the course of one piece of work.

The Rest of the Meeting

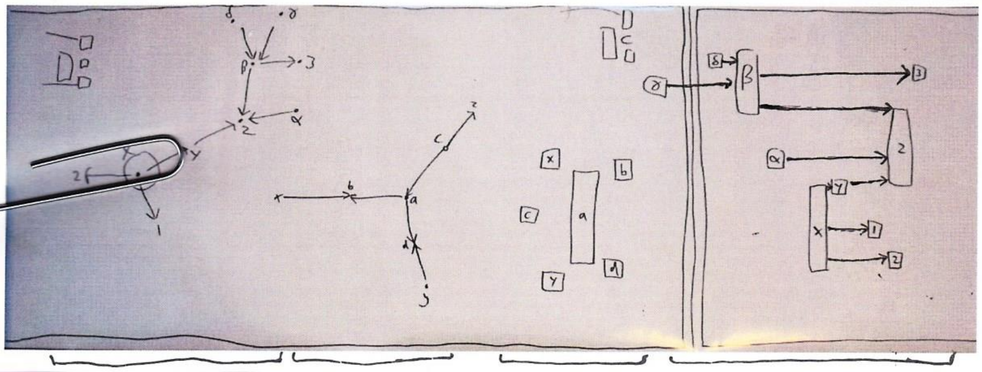
X: Alright, so let's think about how much more we can work out on the basis of the rest of the meeting. So really just the conversations that were had on that same day, in the same room. The first thing is these diagrams, some of them showed up at the same time as each other. You can see that in this chronological record of what was written on the board when up to this point.

Y: OK, yeah. So we can see that… There are two types of drawing, right, one with arrows and one with rectangles. And these two show up at the same time next to each other, and the two on the outside show up at the same time, on either side, each one on the side – arrows on the left, rectangles on the right. As well as the timing, we can see that the letter labels we get here, the *a, b, c, x, y*, they all appear in both of the inner diagrams. Oh, but no *z*?

X: Aah, you remember, F said 'we have nowhere to put *z*', on line 12.

Y: Ohhh, right. Ooh. What does that mean, I wonder. And the same for the other two, we have *x, y, z, 1, 2, 3, α, β, γ, δ*… wait, what? what a mixture…

X: Yes, I think I remember them laughing about this! It's an unspoken rule, normally you'd choose one alphabet or numbering system and stick to it, like Greek letters or Arabic numerals



00.07.16.000

B: So we were hoping – we were hoping – that – in this step of the algorithm, °nv the proof°, regardless of what oriented path cover ((gestures to 'tree' on left)) we could- we could js convert it into a rectangle visibility graph ((gestures to rectangles on right)).… um… which is false.

A: At least the idea of making the paths be rows in the …

B: [((unintelligible))

A: [the whole point was the coherent paths were supposed to be reading left to right

B: if you start with the fundamental idea of … the arrows are always pointing from left to right in the representation

A: yeah

00.33.44.000

01 B: alright so the thing that's [bad

02 D: [OK I'm with you

03 B: … is a vertex with three switching vertices

04 C: [what

05 A: ((nods)) [or more

06 B: or [more

07 C: [a - vertex with three switching vertices?

08 F: cos like [in that with that vertex *a*

09 E: [((unintelligible)) right there

10 F: … *b* is switching for *a*, [*d* is switching for *a* [and *c* is supposed to be switching for *a*

11 B: ((walks over, follows path *xba*)) [yup ((follows path *zca*))[yup

12 F: but we have nowhere to put – *z*

13 B: right

14 A: I hav- I have a problem.

15 B: oerr gahd

16 B: OK what's your problem ((laughing))

17 E: ((laughing)) not again

18 B: ((speaking through laughter)) here comes another one of A's counter-examples

19 A: it just seems… ((points at diagram on board)) isn't… wait, what happened to this… so

20 A: so… *z* was switching. … [but what if it just continued this way and it can switch ((follows path))

21 B: [Yuh. I-

22 A: but then a couple of things switched all at the- ((mimes two coming down from above))

23 B: so if there was

24 A: °s that a problem?°

25 B: if there was… ((walks over to board, A steps away))

26 B: the problem is not for *x*

27 A: O↑H↓

28 D: ((unintelligible)) rectangle

29 B: if there exists a: swi- a ↑vertex↓ from which three vertices are switching… then it's ↑bad↓. then

30 the whole graph's bad.

31 A: The whole graph's bad?

32 B: the- the tree cannot be represented if the vertex ((trails off))

Original in colour

or whatever. You use these systems, so that everything gets a different label but they're all taken from the same set so that they're all kind of… equal, you show that all of the vertices or whatever are of a kind. But they kept mixing them up. And it really cracked everybody up, because it was so against convention.

Y: That's funny. They're supposed to be all of a kind, and just messing with that a little makes them so much more visible as markers.

X: OK so I'm getting the impression that for each arrow-graph there's a rectangle-graph, they come in pairs like that.

Y: Right. So what's the relationship?

X: Uhh, I think this comes up, early on in the meeting. Yeah, here, about 7 minutes in. They talk about… they were *hoping* they could take an 'oriented path cover', and that was something B said while pointing at one of the trees… and thinking they could 'just convert it into a rectangle visibility graph'. So there's some kind of conversion going on here.

Y: Right, great. So you're thinking that they're saying that… if you had a tree diagram, an 'oriented path cover', then it might be possible to draw a rectangle visibility graph that was somehow equivalent to it. An equivalence that has something to do with those labels that we were noticing were the same. And looking at where those labels are, it looks like maybe each character [points to x, y, z in outer two diagrams] is assigned to a point, a vertex, in the oriented path cover, and to a rectangle in the rectangle visibility graph. And in the tree diagram, they're connected by arrows, and in the rectangle one, they're connected by dotted lines…?

X: Yes and I think there's something about that here too. Further on in the same excerpt, here, they keep talking about *going from left to right*, and they say 'arrows are always pointing from left to right in the representation'. So while one type of diagram [points to tree diagram] has… directionality, the connections between the elements are directed, the other type [points to rectangle diagram] maybe it's kind of built in to their relative positions. Let me show you by drawing over the dotted lines [draws arrows from left to right on rightmost diagram], I think it's about which is lined up with which, so that… it says 'visibility', it's about which rectangle can 'see' which other rectangle if they're all facing to the right…

Y: Except that rectangles don't have faces

X: Except that rectangles don't have eyes, right. But it still makes sense, right? If you figure out who can see whom, and who's to the left, then it matches up here with the arrows.

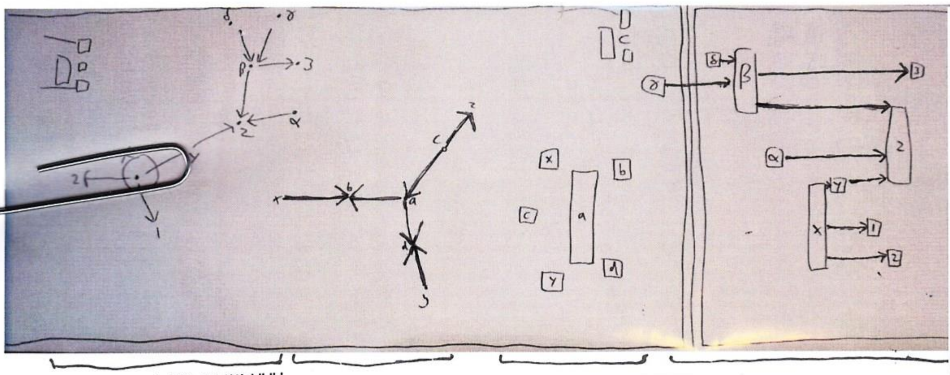
Y: It's funny, I get what you mean about them 'seeing' each other but in the actual diagram they've drawn, they barely line up, or *do* line up when they aren't meant to

X: Yes, that doesn't seem to matter. The representation, it doesn't have to actually have the property.

Y: Why not? Surely we want it to?

X: Well, I guess the point is that it's embedded in a discussion, everybody's talking about it and… the diagram doesn't stand alone, it's something that happens in combination with a bunch of acting people

Original in colour



01 B: alright so the thing that's [bad
02 D: [OK I'm with you
03 B: … is a vertex with three switching vertices
04 C: [what
05 A: ((nods)) [or more
06 B: or [more
07 C: [a - vertex with three switching vertices?
08 F: cos like [in that with that vertex *a*
09 E: [((unintelligible)) right there
10 F: … *b* is switching for *a*, [*d* is switching for *a* [and *c* is supposed to be
11 B: ((walks over, follows path *xba*)) [yup ((follows path *zca*))[yup
12 F: but we have nowhere to put – *z*
13 B: right
14 A: I hav- I have a problem.
15 B: oerr gahd
16 B: OK what's your problem ((laughing))
17 E: ((laughing)) not again
18 B: ((speaking through laughter)) here comes another one of A's counter-examples
19 A: it just seems… ((points at diagram on board)) isn't… wait, what happened to
20 A: so… *z* was switching. … [but what if it just continued this way and it can switc
21 B: [Yuh. I-
22 A: but then a couple of things switched all at the- ((mimes two coming down fror
23 B: so if there was
24 A: °s that a problem?°
25 B: if there was… ((walks over to board, A steps away))
26 B: the problem is not for *x*
27 A: O↑H↓
28 D: ((unintelligible)) rectangle
29 B: if there exists a: swi- a ↑vertex↓ from which three vertices are switching… the
30 the whole graph's bad.
31 A: The whole graph's bad?
32 B: the- the tree cannot be represented if the vertex ((trails off))

?: I dunno. Just if there [are no]

A: [just if there] are no k-bad vertices for >k greater than or equal to< THREE … then… t exists… >oh wait< then…

B: then this algorithm works to convert this orienting path cover into a rectangle

E: algorithm to be written

B: so. ((long pause)) how bou this. <For this vertex x … u::hh … the vertices are switching or whatever you wanna call it … if … you have a path uuuf directed in one way from x? so that vertex… > and then at least one edge directed ((unintelligible))
((explains wrt concrete example))
B: so x has one switching vertex from it
((Switch to talking about vertices being switching))
A: are you doing what I think you're doing? so you're talking about VERTICES being switching. VERTICES are being switching. and then you <limit how many switching vertices there are>.

Original in colour

Y: Right; it's kind of like an ASSEMBLAGE, made up of people and environment. Right, so then when B says something about being able to 'convert it into a rectangle visibility graph' on line 38, that might be what that's all about.

X: Right. So then… I'm going to try it with the other graph. So we have a starting rectangle, [points to biggest rectangle on middle diagram, then points to each smaller rectangle] and then three more, and then I can put *x* here, and *y* here, and…

Y: And there's nowhere to put *z*! Just like F said on line 12!

X: Right, and now we're getting a bit of sense to the… the idea that it isn't possible to represent the graph sometimes, and this middle diagram here, it's there as a counter-example

Y: One of A's counter-examples!

X: So this is the original counter-example. A counter-example that means that you can't *always* represent a tree graph as a rectangle graph, as they had supposed. And I guess that when B is making these STATEMENTS about what's 'bad', they are trying to define which tree graphs this won't work for. So in this diagram [points to middle diagram], there are three arrows coming out of one point, and then each goes on to another arrow, and the arrows are pointing in different – one goes out and the other goes in, or one goes in and the other goes out. So they have to be one block to the left of the other and the next to the right, and vice versa. And you can do that on either end of the first rectangle, but after that, the third time, there isn't anywhere to put it…

Y: So you think that… that the basic aim here is to try to – you're calling this a counter-example, and you think that it's an effort to somehow write down… what it is …

X: A description of some property of the counter-example that is connected to why it doesn't work

Y: *Why* it doesn't work? What do you mean by that?

X: Oh wow, I guess I don't know!… I think I mean… OK, how about a description that would help you to know whether another example would work or wouldn't, or would let you know how to construct one that wouldn't work. How's that?

Y: Hm. It's… OK. I guess if we feel like we understand something, we feel like we can predict it.

X: Yeah. So I think… judging by B's attempts to make these neat STATEMENTS, what they want is to write a kind of description… here it's a verbal description, a description of the… the *non-working thing* in the counter-example.

Y: Yes, and it's stated in this emphatic way… it's very brief, just a short sentence

X: Right, so the exact wording is important. Actually, you can see similar short phrases repeated during the meeting, here, and here, and here… it's almost like they're adjusting it. So maybe… adjusting the wording of the phrase they keep repeating is… somehow keeping track of the shared understanding, of the way the group is seeing the example, in this portable phrase that can be easily repeated

Y: Tell me more about these adjustments.

Original in colour

00.30.48.000

B: so. ((long pause)) how bout this. <For this vertex x … u::hh … the vertices are switching or whatever you wanna call it … i:f … you have a path uuuf directed in one way from x? so that vertex… > and then at least one edge directed ((unintelligible))

((explains wrt concrete example))

B: so x has one switching vertex from it

((Switch to talking about vertices being switching))

A: are you doing what I think you're doing? so you're talking about VERTICES being switching. VERTICES are being switching. and then you <limit how many switching vertices there are>.

00.16.20.000

B: ((finishes drawing diagram)) So- does that count as one- so I'm looking at that vertex ((circles x)). Does that count as one switching path or two

00.18.04.000

F: Yeah so what do we mean by a switching path
E: Yeah so what's a switching path
A: WELL UH- TH- K- OK. I was hung up on this also. It doesn't mean <path of the path cover>. It's [just
C: [Oh
A: a path in the graph RELATIVE TO the path cover is switching.
A: So this is a path in the graph ((follows *x* to *δ*))

00.18.46.973

F: So whether we have to channel from whatever vertex we're looking at all the way to a leaf. … before we decide whether it's switching.

00.20.05.000

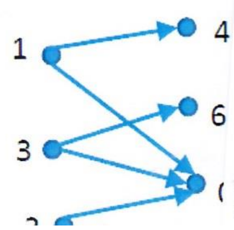
D: So am I right for that vertex we're saying that's- we got THREE switching paths?
C: Three- well we don't know that
A: ((following paths to ends. follows *x* to *δ*)) one… ((follows *x* to *a*)) two… ((follows *x* to *γ*)) three. Yes
B: I- uh- does it
C: Cause you go all the way to the leaf.
F: I mean if- I mean IF, we haven't- determined that
A: Or it has just one!

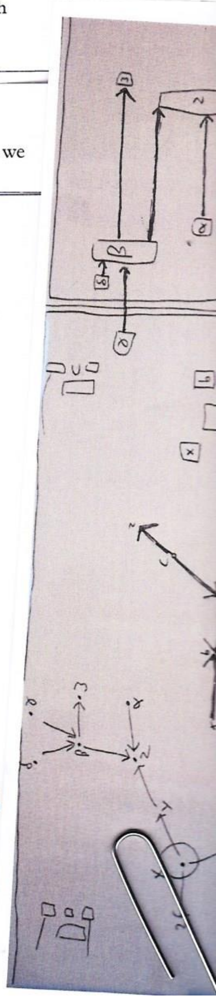
Seems hard to add these paths as "rows" in an RVG

1				4		
1			0			
	2		0			
		3	0			
		3				6

(I'm having a hard time using the OPC approach that we discussed to include the (1,4), (2,5), and (3,6) arcs around the rectangle for vertex 0.)

Can we find a fix for this kind of thing?





Original in colour

X: Alright, so right now B is saying that… 'the thing that's bad… is a *vertex* with three switching *vertices*'. Earlier on we see this line here that talks about switching vertices. There's a *change* here. Before this line the group had been talking about switching *paths*, so – I think by that they mean… if you follow a line of arrows, you follow one to a point, then the next, then the next, like this. So then the word 'switching' refers to the direction of the arrows, whether it changes direction as you go along.

Y: Wait, what do you mean by 'changes direction'? Relative to what? I understand what you mean by 'as you go along', of course, but why should it matter – I mean, it's changing direction relative to what?

X: Yeah, I'm not sure. I guess I'm just talking about—I'm talking about when two arrows point away from one another, or toward one another. Actually I think this is discussed earlier in the meeting, at around 16-18 minutes in [finds excerpt], they talk about what counts as a switching path. It's relative to a thing they call a path cover, which I *think* is what decides where those arrows are pointing, like on this [shows diagram from Dropbox]- this is something from their Dropbox, which I think is the first presentation of this middle diagram here, of the first thing they had trouble representing.

Y: Right, I think I see that, but what about where you… *count from*, if you see what I mean?

X: Well, I think you can start anywhere, actually. F says here, 'from whatever vertex we're looking at'

Y: OK. And then what counts as a path, a string like this [follows path along on leftmost diagram]? And what happens if it branches?

X: Right, you have to decide where it ends and how to count them.

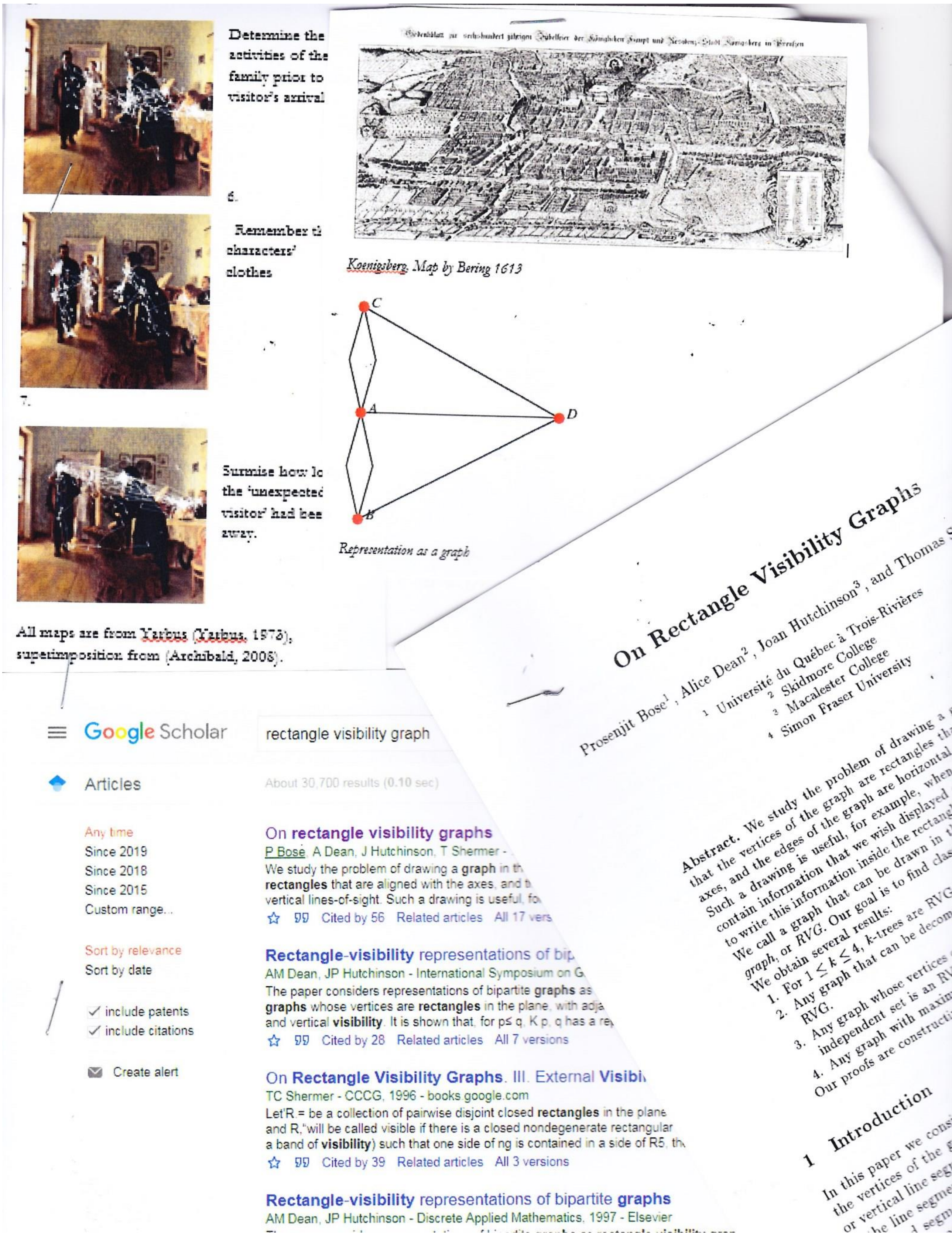
Y: Oh I see, so whether – whether a path has to end at a 'leaf', which I take it is one of these endpoints here. And how many times you count it if it later branches.

X: And then there's a shift, here, to talking instead about switching *vertices,* the first time that a path switches. Look, they emphasise it, it's a total shift in focus.

Y: Oh, wow! OK, so by changing that word a person is changing the focus from… rather than talking about a path, which might go from point a to point b or to point c, they're talking instead about the switching being sort of *decided at one point*, where the arrows point together or apart. [points to first switch in leftmost diagram] It's localised at the first moment that it happens, as you encounter it coming from a starting vertex.

X: Right, so the switch in focus, it subtly evades those questions. And then I think that the 'vertex with three switching vertices' that B talks about, or it's a little clearer at the end here when B says '↑*vertex*↓ from which three *vertices* are switching', that means… B's emphasising that first 'vertex', trying to say that it's a question of whether you can find *any vertex anywhere* with three of those first-time switching vertices, and we just have to look for the first time it happens down a path, so the rest of the path isn't important. It changes what you look for. And it's not just the content of the statement, it's where the emphasis is placed.

Y: and so then we can follow this, and say… *c* is switching for *a*, *d* is switching for *a* – just like F said!


Determine the activities of the family prior to visitor's arrival
6.
Remember the characters' clothes
7.
Surmise how lo the 'unexpected visitor' had bee away.
All maps are from Yarbus (Yarbus, 1973), superimposition from (Archibald, 2008).
Koenigsberg. Map by Bering 1613
C
A
D
B
Representation as a graph
Google Scholar
rectangle visibility graph
Articles
About 30,700 results (0.10 sec)
Any time
Since 2019
Since 2018
Since 2015
Custom range...
Sort by relevance
Sort by date
include patents
include citations
Create alert
On rectangle visibility graphs
P Bose, A Dean, J Hutchinson, T Shermer -
We study the problem of drawing a graph in th
rectangles that are aligned with the axes, and t
vertical lines-of-sight. Such a drawing is useful, fo
Cited by 56 Related articles All 17 vers
Rectangle-visibility representations of bip
AM Dean, JP Hutchinson - International Symposium on G
The paper considers representations of bipartite graphs as
graphs whose vertices are rectangles in the plane, with adja
and vertical visibility. It is shown that, for p≤ q, K p, q has a re
Cited by 28 Related articles All 7 versions
On Rectangle Visibility Graphs. III. External Visibi
TC Shermer - CCCG, 1996 - books.google.com
Let R = be a collection of pairwise disjoint closed rectangles in the plane
and R, will be called visible if there is a closed nondegenerate rectangular
a band of visibility) such that one side of ng is contained in a side of R5, th
Cited by 39 Related articles All 3 versions
Rectangle-visibility representations of bipartite graphs
AM Dean, JP Hutchinson - Discrete Applied Mathematics, 1997 - Elsevier
On Rectangle Visibility Graphs
Prosenjit Bose[1], Alice Dean[2], Joan Hutchinson[3], and Thomas S
1 Université du Québec à Trois-Rivières
2 Skidmore College
3 Macalester College
4 Simon Fraser University
Abstract. We study the problem of drawing a g
that the vertices of the graph are rectangles th
axes, and the edges of the graph are horizontal
Such a drawing is useful, for example, when
contain information that we wish displayed
to write this information inside the rectang
We call a graph that can be drawn in th
graph, or RVG. Our goal is to find clas
We obtain several results:
1. For 1 ≤ k ≤ 4, k-trees are RVG
2. Any graph that can be decom
RVG.
3. Any graph whose vertices
independent set is an RV
4. Any graph with maxim
Our proofs are constructi
1 Introduction
In this paper we cons
the vertices of the g
or vertical line seg

X: And if there are three switching vertices coming off of one vertex, then we have nowhere to put z!

Y: This use of 'vertex' to focus attention on the beginning, it reminds me of those eye tracking studies of Alfred Yarbus', showing how attention is directed when people are asked to answer certain questions about a picture. Those sort of… it's like a map of perception being active and guided by the mind, guided by what you're looking for even if you aren't that aware of it. So in a way these subtle shifts in wording and emphasis are sending the group's eyes and minds around the diagram, paying attention to different parts. They're an ATTENTION-DIRECTING TOOL, and the diagram can be a very different thing depending on how you look at it.

The Mathematical Community

X: OK, so we think this connects up with graph theory, maybe

Y: Yeah. So that's… have you heard of the Königsberg Bridge Problem?

X: That's the Russian city with an island, and a river, and seven bridges, right?

Y: Right, and the problem is trying to find out whether there's a route around the city that would cross each bridge only once. And nobody could find one, but nobody had been able to demonstrate that it wasn't possible at all until Euler found a way to study this and prove it in the 1730s. I think it was…

X: OK, so this is supposed to be a kind of… graph theory representation, right? With, uh, vertices and connecting lines. So all the landmasses are vertices, and each line is a bridge-route.

Y: Yeah, so then a city is represented as a graph, and then it's possible to… to say what can and can't be done with that graph. It's the study of these constructions.

X: Yep. And I think… they mention rectangle visibility graphs, somewhere…

Y: Oh, great! So if I search that phrase on Google scholar, then I get some results. Bose, Dean, Hutchinson. There's this one, Bose, Dean, Hutchinson, Shermer (Bose et al., 1997). They say 'Our goal is to find classes of graphs that are RVGs.'

X: OK, so it seems like people are concerned with working out what can and can't be represented. But it looks like these are just with lines, not arrows.

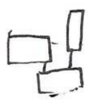
Y: OK, so there are a few things that look different and we don't know why? Shall we ask C?

X: Yes!

C: So… there was something called bar visibility graphs, and then interval graphs. Rectangle visibility graphs were an extension of bar visibility graphs. I can give you some papers … there's a paper here on rectangle visibility graphs from about '97, this is Dean & Hutchinson, 1997, which is probably roughly as old as the field is…

X: Oh, great! So we have here… a couple from the '80s on bar visibility graphs, which are Wismath, 1985; Kirkpatrick & Wismath, 1989

C: Yeah, so they're a little different - for the bars it's literally just bars like this [draws lines], and then you only get visibility in one direction. Rectangle visibility you have both, [draws rectangle]



Original in colour

On Rectangle Visibility Graphs

Prosenjit Bose[1], Alice Dean[2], Joan Hutchinson[3], and Thomas Shermer[4] *

[1] Université du Québec à Trois-Rivières
[2] Skidmore College
[3] Macalester College
[4] Simon Fraser University

Characterizing Bar Line-of-Sight Graphs

Stephen K. Wismath
University of Lethbridge

Weighted Visibility Graphs of Bars and Related Flow Problems (Extended Abstract)

David G. Kirkpatrick *Computer S...

ELSEVIER

Discrete Applied Mathematics 75 (1997) 9–25

DISCRETE APPLIED MATHEMATICS

Rectangle-visibility representations of bipartite graphs

Alice M. Dean[a,1], Joan P. Hutchinson[b,*]

Original in colour

interior points. In this situation, we will call two rectangles u and v *visible* if there is a *band of visibility* $B_{u,v}$ between them: $B_{u,v}$ is a rectangular region with two opposite sides that are subsets of u and v, and such that $B_{u,v}$ intersects no other rectangle of $\mathcal{R}$. The *visibility graph* of $\mathcal{R}$ is the graph of the visibility relation on the elements of $\mathcal{R}$. We call a graph G a *rectangle-visibility graph* or RVG if it is the visibility graph of some collection $\mathcal{R}$ of rectangles; in this situation, $\mathcal{R}$ is called a *layout* of G. The edges of a rectangle-visibility graph G can be partitioned into the two sets representing horizontal and vertical visibility; each of these two edge sets forms a BVG. Thus G, as a union of two planar graphs, is said to have *thickness-two*. Much less is known about thickness-two graphs than about planar ones, although their recognition is known to be NP-complete.

Wismath [14] has shown that every planar graph has a rectangle-visibility layout. Hutchinson, Shermer, and Vince [8] show that a rectangle-visibility graph with n vertices has at most $6n-20$ edges, in contrast with thickness-two graphs, which may have at most $6n-12$ edges; in both cases these bounds are attainable. Dean and Hutchinson [4] show that $K_{5,5}$ is *not* a rectangle-visibility graph though $K_{5,5}$ plus any edge is; see Figure 2. Thus rectangle-visibility graphs are not closed under the formation of subgraphs.

Each of the classes of BVGs and RVGs has two important subclasses: graphs with *noncollinear* layouts and those with *strong* layouts, as defined below.

2.3 Collinear and noncollinear layouts

A bar-visibility layout is called *noncollinear* if no two line segments have collinear endpoints; a rectangle-visibility layout is noncollinear if no two rectangles have collinear sides. The 4-cycle (see Figure 1a) does not have a noncollinear bar-visibility layout; Figure 3 shows a (collinear) rectangle-visbility layout of $K_{4,4}$ minus an edge; by a result in [3] it has no noncollinear layout, but a noncollinear layout of $K_{4,4}$ minus two edges is shown in Figure 10.

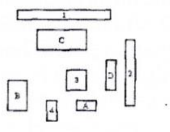


Fig. 3. A collinear (and strong) layout of $K_{4,4} - e$.

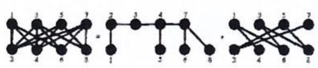
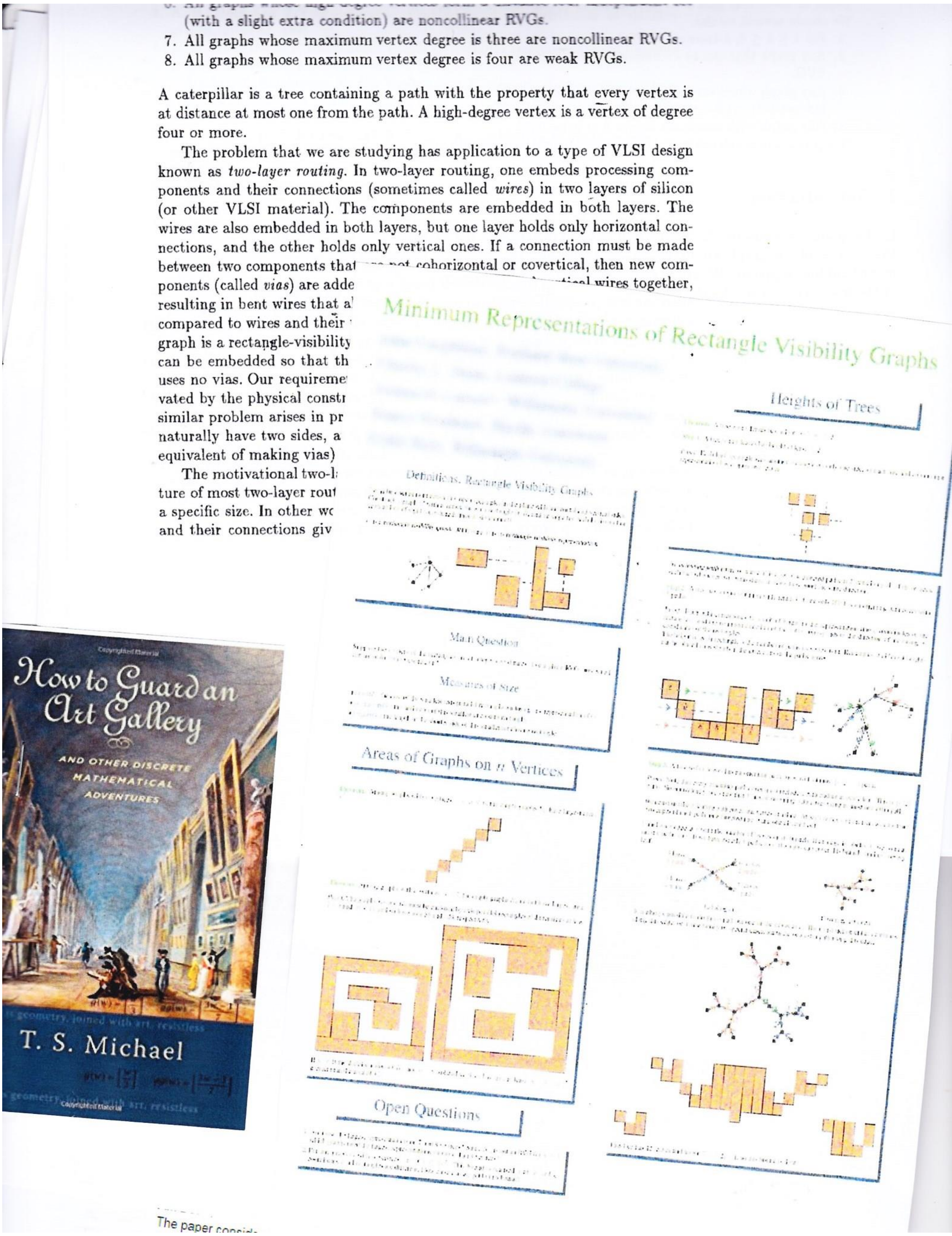


Fig. 4. $K_{4,4}$ minus two edges decomposed into two caterpillars.



(with a slight extra condition) are noncollinear RVGs.
7. All graphs whose maximum vertex degree is three are noncollinear RVGs.
8. All graphs whose maximum vertex degree is four are weak RVGs.

A caterpillar is a tree containing a path with the property that every vertex is at distance at most one from the path. A high-degree vertex is a vertex of degree four or more.

The problem that we are studying has application to a type of VLSI design known as *two-layer routing*. In two-layer routing, one embeds processing components and their connections (sometimes called *wires*) in two layers of silicon (or other VLSI material). The components are embedded in both layers. The wires are also embedded in both layers, but one layer holds only horizontal connections, and the other holds only vertical ones. If a connection must be made between two components that are not cohorizontal or covertical, then new components (called *vias*) are adde
resulting in bent wires that a
compared to wires and their
graph is a rectangle-visibility
can be embedded so that th
uses no vias. Our requireme
vated by the physical constr
similar problem arises in pr
naturally have two sides, a
equivalent of making vias)

The motivational two-l
ture of most two-layer rout
a specific size. In other wc
and their connections giv

The paper consid

Original in colour

they can see each other from up here and down here. They were an extension from bar visibility graphs, which have been studied for a lot longer. Then you can put restrictions on this, like studying unit bar visibility graphs where you can only use squares, and that restricts what you can do.

Y: OK. It looks like the rectangles in these diagrams weren't 'seeing' each other from above or below, though

C: Because we were doing trees, we don't have any cycles. So I think that's why we were trying to avoid using up-down visibility, you have to kind of limit it so you don't get a cycle. It's something you can use when you need three to see each other, like this.

X: Aah! Yes, I see. Right, so now we have this '97 paper. Here they say that their 'goal is to characterize those complete bipartite graphs that can be so represented…', So working out which of these vertices-and-lines kinds of graph can be represented as…

Y: Can be represented, like, we can build a rectangle graph that has the same number of vertices, related in ways that are parallel?

X: …right, can be represented as a rectangle visibility graph, uh, that's a – that's the kind of thing that people study.

C: Yes. And then we can say things about the properties of those rectangle visibility graphs, like how big they have to be to work.

Y: Also, I notice that these graphs have *lines*, rather than *arrows*.

C: Yes, that's right, so we're using an Oriented Path Cover like for example this poster that our research group made, that dealt with the area of the graphs.

Y: OK. Just out of interest, what do these… get used for? I mean applications?

C: A friend of mine who was also a student of my PhD adviser wrote a fantastic book called *How to Guard an Art Gallery* (Michael, 2009).[picks up book]. So you can imagine, a problem being that you have an art gallery with certain rooms, and you want the minimum number of guards, where do you position them so they can see all of the rooms…

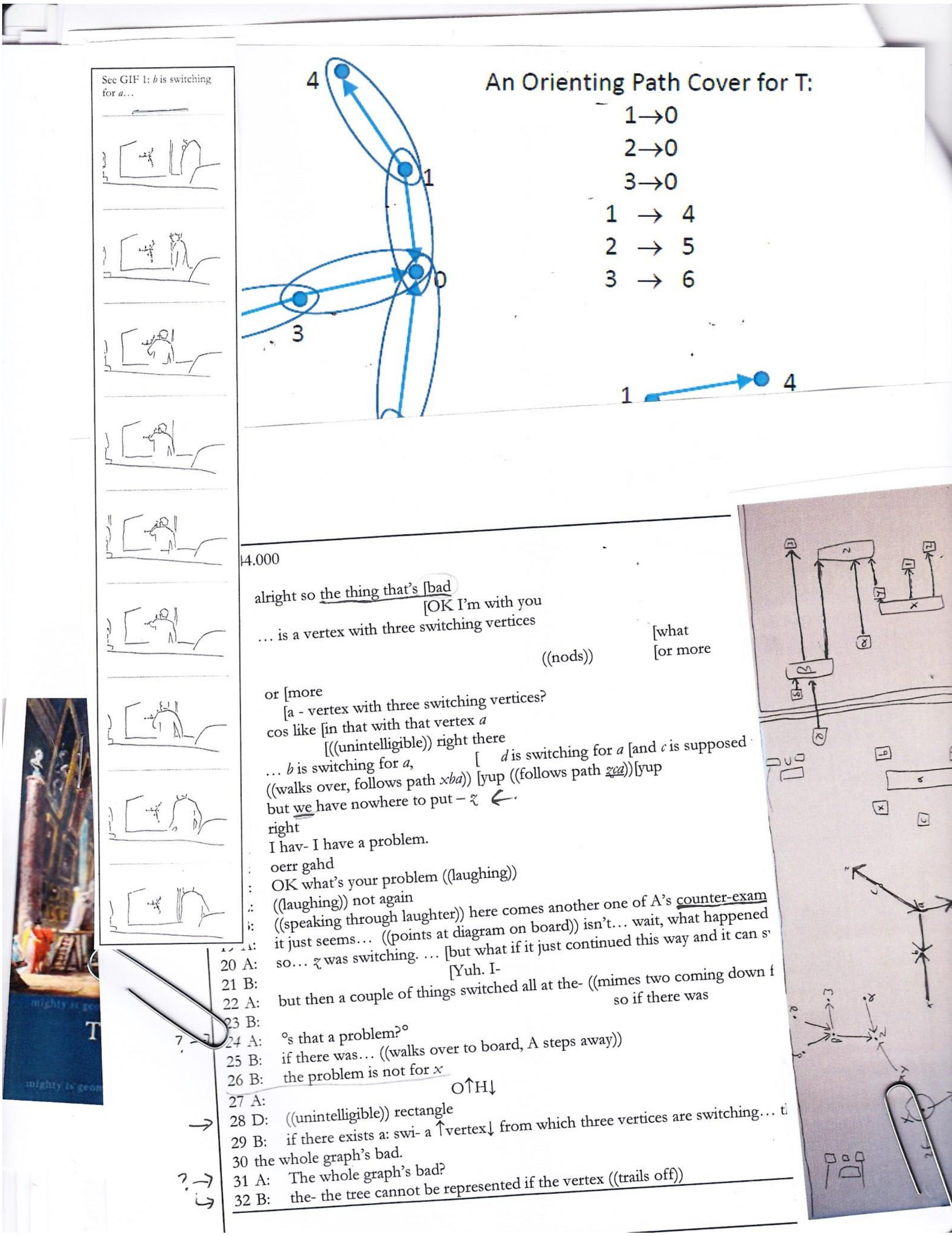
X: Yeah, I see! That's great. Oh and this chapter here starts with a Georgia O'Keefe quote, 'I found I could say things with colours and shapes that I couldn't say any other way' – I guess that's about reasoning with diagrams?

Y: Yes, though I note that diagrams are sometimes rightly treated with suspicion in mathematics. When you're looking at a particular diagram, the particular features in the example you choose might give you a misleading impression about the general claims that can be made.

X: I notice in this '97 paper that reference is made to… VLSI design, that's kind of circuit board design, right? I guess that makes sense, as a useful application. Hey, thank you, C, that's been really helpful.

C: No problem!

Y: Yes, so let's reflect a moment—what kind of thing is it we've learned from talking to C?



Original in colour

X: Hmm. I guess we know a lot more about the history, who else has done what and why… the *why* seems quite important, I've been interested to learn what kind of questions people want to answer, what constitutes a contribution to this field.

Y: Yep! And questions we have, like why there are arrows instead of lines, and why in these examples we're just using left-to-right visibility, how it relates to the other work, those are hard to answer, because we don't have the right words to look for them. A broader picture of an ASSEMBLAGE might include the network of books, papers, online resources, and human repositories of knowledge.

Guiding Perception

X: OK, so let's get in to some of this detail. On lines 01 and 03, B proposes a STATEMENT of a definition of the restriction, a way to define which examples will be unrepresentable. C doesn't immediately accept the definition of the restriction

Y: Yep, and then F jumps in and talks through the example

X: So… A thing I want to talk about is this context that is shared by the group, this mutually manifest landscape. So that includes the – what's been said in the conversation so far, their memory of previous meetings… I know they had some items in their shared Dropbox that we've seen already, diagrams they'd made themselves and papers that they'd shared… then there's also what's on the board in the room…

Y: Yeah, so they have things they've all been looking at and thinking about. And a lot of it, they all know that they've all seen it, if you see what I mean, they *know* what they're all seen.

X: Right, but they still don't all *know* the same things. They have access to all of this shared material but they're only actually thinking about parts of it, and *which* parts might differ. There will be differences even if they had all seen the same material.

Y: Like if one person had really thought hard about one part, and then had been eating snacks when another was mentioned, or had worked on something else recently that was related…

X: Right, and so here (07) we see that B and F, and C, aren't quite on the same page.

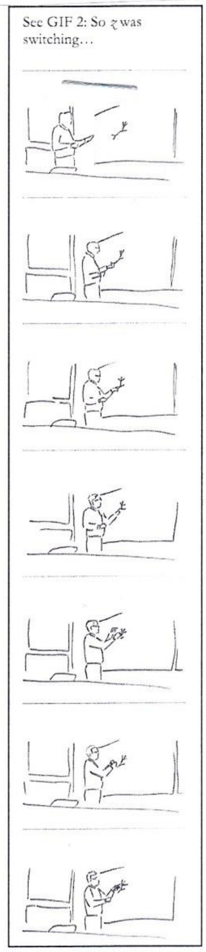
Y: OK, yeah. So then we can think a little about what F does to fix that. Let's look at this section… So F says that '*b* is switching for *a*, *d* is switching for *a* and *c* is supposed to be switching for *a*', and B comes over, and follows each line to show how it animates the diagram. So that's interesting, it's the *ordering* of F's explanation that clarifies that it is the *vertices*, not the paths, that are switching, because F says *b*, *c* and *d* first. Listing each of these letters locates the 'problem' in the points with these labels, as opposed to the paths. And it's – there's repetition, '*b* is switching for *a*, *d* is switching for *a* and *c* is supposed to be switching for *a*' with emphasis on b, d and c, so we see that it's the same each time, the emphasis is on the vertex. It's an ATTENTION-DIRECTING NARRATION.

X: Right! And then… so F simply lists each of the vertices, and B just gestures to the same paths. So… what they are doing is taking the static diagram, and animating it, making the first circle of vertices the focus and changing the role of the second circle to be 'things that need to be placed'.

Y: Yes, it's sort of… filling out the diagram, they are using their engagement with the diagram to shape C's understanding of it, and of the STATEMENT of a definition that B gave. So F doesn't try to explain B's condition. Instead F just tries to talk through the example to *exhibit* it as something that has certain features.

X: So… rather than further explicating the condition, F actually just identifies how example 1 meets it, the implication being that the example has three switching vertices and so cannot be represented. So F makes it clear that it's impossible in that case, and… stating it narratively like that shows that the

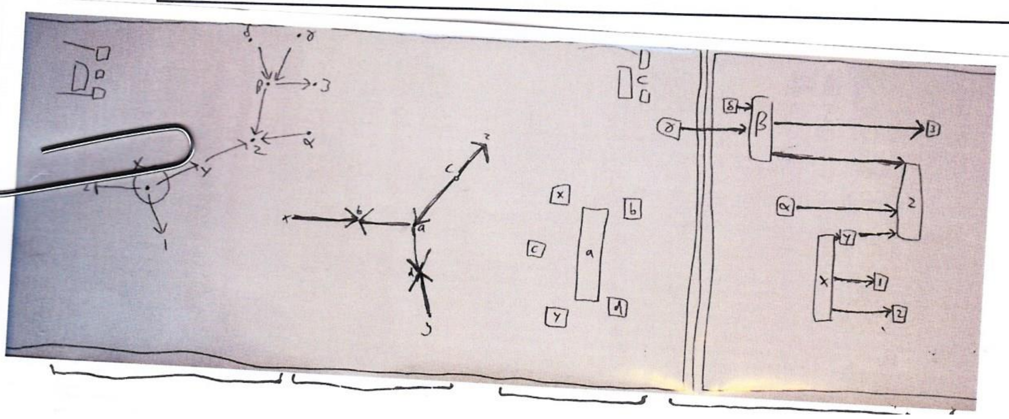
Original in colour



000

ılright so the thing that's [bad
[OK I'm with you
.. is a vertex with three switching vertices
[what
((nods)) [or more
ɔr [more
[a - vertex with three switching vertices?
cos like [in that with that vertex *a*
[((unintelligible)) right there
.. *b* is switching for *a*, [*d* is switching for *a* [and *c* is supposed to be switching for *a*
walks over, follows path *xba*)) [yup ((follows path *zca*))[yup
ut we have nowhere to put – *z*
ght
hav- I have a problem.
err gahd
K what's your problem ((laughing))
laughing)) not again
speaking through laughter)) here comes another one of A's counter-examples
just seems… ((points at diagram on board)) isn't… wait, what happened to this… so
… *z* was switching. … [but what if it just continued this way and it can switch ((follows path))
[Yuh. I-
ıt then a couple of things switched all at the- ((mimes two coming down from above))
so if there was
that a problem?°
there was… ((walks over to board, A steps away))
e problem is not for *x*

27	A:	O↑H↓
28	D:	((unintelligible)) rectangle
29	B:	if there exists a: swi- a ↑vertex↓ from which three vertices are switching… then it's ↑bad↓. then
30		the whole graph's bad.
31	A:	The whole graph's bad?
32	B:	the- the tree cannot be represented if the vertex ((trails off))



Original in colour

definition points to what it is that makes it impossible. You get to switching vertex 3, and it fails. And C can now see that.

Y: Yes, and the board notes are in this shared space, and they're almost *there for* people to interact with. We can tell who's holding the baton, so to speak, because they're standing next to the board.

X: Right, so you and I right now are making sense of these diagrams *by virtue of* the way they're used by the people, like that this one is exhibiting a property that needs defining, and this one *isn't* exhibiting that property, and they've been placed here so that people can interact with them, add bits and see what happens, find different ways to pay attention to different parts, physically interact with them—for their role in the ASSEMBLAGE. They're not just there to be decoded, they're external tools for reasoning.

Y: Yes, and it occurs to me that… Figuratively speaking, when STATEMENTS are proposed in lines 01 and 03, and 29, 30, 32, these are also placed in a shared space for the group to consider and critique, but are not set down on the board. The STATEMENTS can be discussed while they are fresh in the group's memory, and that might be why they're stated over and over again. Since the work of the group is to refine the STATEMENT, this is continuously helpful.

X: So this seems to involve a kind of sharing of mental states, a physical enacting of an understanding such that others can access the speaker's personal experience of what is going on and adopt it, or act upon it, change it. An *understanding* is a very difficult thing to directly share, and this acting-out-on-the-diagram allows something like *showing how you understand something* to occur through intentional highlighting and guidance. The speakers performatively think through the problem, and as they do it they talk through the process a little bit extra to make that thinking available to the watching participants.

Sharing Mental States

Y: OK, so then we have the 'problem is not for *x*', this part we're still not that sure about.

X: Yep. So in lines 20 and 22, we have this interesting action by A. [shows animation of lines 20-22]. A moves the focus back to the outer two diagrams by physically moving closer to this part. And A says, 'so… z was switching. … but what if it just continued this way and then switch' and follows the line up and then says 'but then a couple of things switched all at the' and seems to mime two coming down from above. So same here, we're seeing A sort of animate the diagram. A's talking about two coming down… I think the movements here are important, A follows the line up, so A's asking about the continuation of a switching path, what happens after it's switched once, because B has changed the focus to the *first* switch

Y: Hm, right. And then the 'two from above', A's actually talking about 'a couple of things switched all at the same time' so really I think that… A's asking a question about further switching downstream from an existing switching vertex. But A's kind of illustrating it by adding switching *paths*, not vertices. Because of this action, miming these two coming down from above, but that would only add one switching vertex.

X: Oh, I think I see what you mean. Yeah. But still, we can get the sense… I mean, A's worried about only counting this first one, about what happens afterward. It's still this question about *how to count* them.

Y: Ooh, yes. [slowly] *But* it's not – it's not as though A has *said* this, exactly, more like kind of … shared a way of looking at this diagram as having a path going up here and then things happening later on, and it's kind of vague, but still we can… kind of extrapolate from that what the miscommunication is

00.33.44.000

01 B: alright so the thing that's [bad

02 D: [OK I'm with you

03 B: ... is a vertex with three switching vertices

04 C: [what

05 A: ((nods)) [or more

06 B: or [more

07 C: [a - vertex with three switching vertices?

08 F: cos like [in that with that vertex *a*

09 E: [((unintelligible)) right there

10 F: ... *b* is switching for *a*, [*d* is switching for *a* [and *c* is supposed to be switching for *a*

11 B: ((walks over, follows path *xba*)) [yup ((follows path *zca*))[yup

12 F: but we have nowhere to put – *z*

13 B: right

14 A: I hav- I have a problem.

15 B: oerr gahd

16 B: OK what's your problem ((laughing))

17 E: ((laughing)) not again

18 B: ((speaking through laughter)) here comes another one of A's counter-examples

19 A: it just seems... ((points at diagram on board)) isn't... wait, what happened to this... so

20 A: so... *z* was switching. ... [but what if it just continued this way and it can switch ((follows path))

21 B: [Yuh. I-

22 A: but then a couple of things switched all at the- ((mimes two coming down from above))

23 B: so if there was

24 A: °s that a problem?°

25 B: if there was... ((walks over to board, A steps away))

26 B: the problem is not for *x*

27 A: O↑H↓

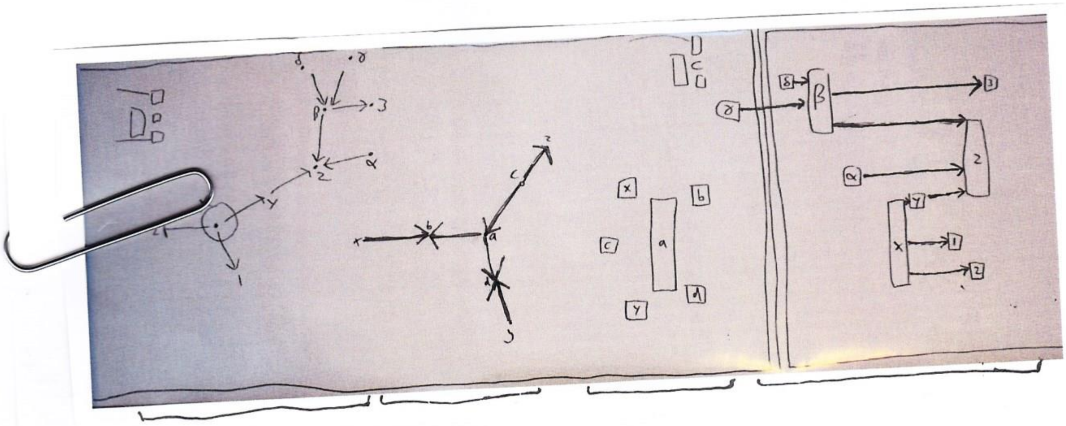
28 D: ((unintelligible)) rectangle

29 B: if there exists a: swi- a ↑vertex↓ from which three vertices are switching... then it's ↑bad↓. then

30 the whole graph's bad.

31 A: The whole graph's bad?

32 B: the- the tree cannot be represented if the vertex ((trails off))



Original in colour

X: Right. And then B, B kind of thinks about it for a while and just says: [with emphasis] 'the problem is not *for x*'. This is great. Earlier on they had been using *x* as the vertex they were tracing paths from. But now… B's pointing out that it matters where you count it from, according to this new 'vertices' description, that if you introduced a bunch more switching vertices up there then if you were looking from *x* then it wouldn't look like a problem, but it *would* from another vertex. It's yet another interesting question about counting, about how to count

Y: Oh yeah! and I'm finding it funny that you're inhabiting these vertices now, you're giving them perspective

X: Yep, and A was kind of caught in *x*'s perspective, and B was trying to make the point that it doesn't matter which vertex it doesn't work for, as long as there's one it looks that way for

Y: OK yeah, so it's this kind of subtle question about how to 'count' them, which is a little hard to get at in words. It's really important for them to decide what gets counted, and what doesn't. You can have a diagram with a whole load of arrows switching directions, and you can still build the rectangle version, and you can have a diagram with just a few switches of direction and you can't build the rectangle. And they're just trying to find the right way of counting them, the right way to go through the diagram – really literally, it seems, we can see how it's helping to direct their attention around this drawn picture – and say, yes this one has too many of the *bad sort,* no, this one doesn't and should be fine

X: Yeah. And then A wasn't seeing that it could be counted in this other way. And so B just – just says something really brief to contradict the assumption that A was making.

Y: Right, so … rather than saying something like, 'you can start at any vertex', it's sort of as though B has just *perceived* A's picture of the problem as still being kind of… hooked on following the lines from *x*, and so B just contradicts that one assumption, and then everything sort of… hangs together in a different way for A

X: Right. This is a really tricky one, huh. I think we've more or less understood it, though.

Y: What I like is that neither A nor B really *say* what – neither of them really states anything clearly, the actual *words they utter* don't explicitly say any of this. They make points in these really indirect ways, and yet it clearly works perfectly.

X: Right, and… and if A hadn't gone through this process of thinking out loud with the diagram, or just thinking through it but doing so in a way that's audible and visible, then B could not have got such a sense of where A's understanding was at, and been able to know what to say to change it. THINKING OUT LOUD GIVES ACCESS TO MENTAL STATES.

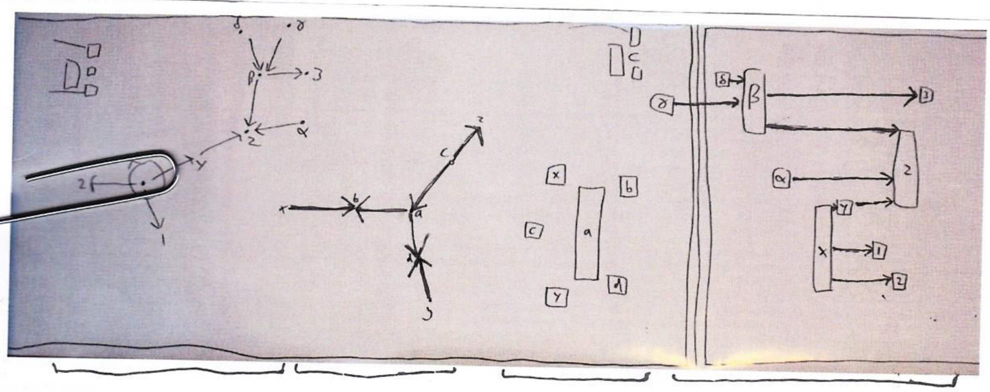
Y: Yeah. So it really was as though A's thought process just happened, for that moment, externally, so that B could see it and replicate it. And so know how to alter it.

X: So after this exchange, after helping A to see the condition the way B is, then B restates it. B says 'if there exists a ↑vertex↓ from which three vertices are switching… then it's ↑bad↓. then the whole graph's bad.' So now emphasis is on that vertex *just existing*, wherever it might be. So now that B is confident that A has seen it, then B tries to state it as a sentence, to find the right STATEMENT.

Original in colour

00.33.44.000

01 B: alright so the thing that's [bad
02 D: [OK I'm with you
03 B: … is a vertex with three switching vertices
04 C: [what
05 A: ((nods)) [or more
06 B: or [more
07 C: [a - vertex with three switching vertices?
08 F: cos like [in that with that vertex *a*
09 E: [((unintelligible)) right there
10 F: … *b* is switching for *a*, [*d* is switching for *a* [and *c* is supposed to be switching for *a*
11 B: ((walks over, follows path *xba*)) [yup ((follows path *zca*)) [yup
12 F: but we have nowhere to put – *z*
13 E: right
14 A: I hav- I have a problem.
15 B: oerr gahd
16 B: OK what's your problem ((laughing))
17 E: ((laughing)) not again
18 B: ((speaking through laughter)) here comes another one of A's counter-examples
19 A: it just seems… ((points at diagram on board)) isn't… wait, what happened to this… so
20 A: so… *z* was switching. … [but what if it just continued this way and it can switch ((follows path))
21 B: [Yuh. I-
22 A: but then a couple of things switched all at the- ((mimes two coming down from above))
23 B: so if there was
24 A: °s that a problem?°
25 B: if there was… ((walks over to board, A steps away))
26 B: the problem is not for *x*
27 A: O↑H↓
28 D: ((unintelligible)) rectangle
29 B: if there exists a: swi- a ↑vertex↓ from which three vertices are switching… then it's ↑bad↓. then
30 the whole graph's bad.
31 A: The whole graph's bad?
32 B: the- the tree cannot be represented if the vertex ((trails off))



Y: Right. And there's this… 'it's bad, the whole graph is bad'. We should talk a bit about this term 'bad', this very emotive word that sort of refers to the work of the group but also has a particular sense, to do with difficulty with building one of these rectangle graphs that is parallel to a tree graph.

X: Yeah, so in lines 01-03 'bad' referred to the local problem happening in a particular vertex. The badness isn't so much something that is happening in the whole graph, but it *is* something that *can't be done* for the whole graph. The group want to be able to represent an entire graph, not just part—so if a part does not work, this counts as the entire example not working, for them.

Y: Yeah. So in these two lines (29 and 30), B aligns badness of one vertex with badness of the whole thing, but of a different kind. And A isn't – just asks for clarification, again by repeating word for word, like C did.

X: Right. B is already thinking in terms of what can or can't be represented, and evaluating everything relative to that goal; this is what B means by 'the whole graph's bad', which is understood at least by F. And that clarification, I think it… it answers the question by finishing the sentence in a more stateable way

Y: What do you mean by that?

X: I guess it's something that is clearer, more formal, less dependent on the group's informal shared understanding, more… portable?

Y: Hmm, interesting. So these clarifications- this time the STATEMENT is restated in a way that sounds like it's closer to the language of a paper. But the condition is only stated as being 'from which three vertices are switching', which is only one particular case, and they've already said that *anything* more than two will do it. So it's still stated in a way that's closely tied to a particular example.

X: OK. I just want to talk about valenced language for a moment. Like 'bad', and… They use the word 'problem' on four separate lines, on 14, 16, 24 and 26, but I get the impression that it means quite different things

Y: OK, how so?

X: Well on line 14, A is expressing a concern about B's STATEMENT, and I think it's used in the same spirit on line 16 and 24, though 24 is less clear. But in 26, I *think* the 'problem' seems to be within the example, it's about what's unrepresentable so that the STATEMENT holds. And it seems on paper kind of incredible that these people successfully arrive at these different interpretations in such a short time.

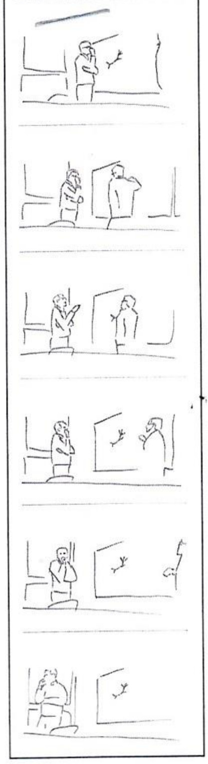
Y: Sure. I mean, it's a common word, with lots of meanings. And when we hear it, we have to find the right one. I guess through context?

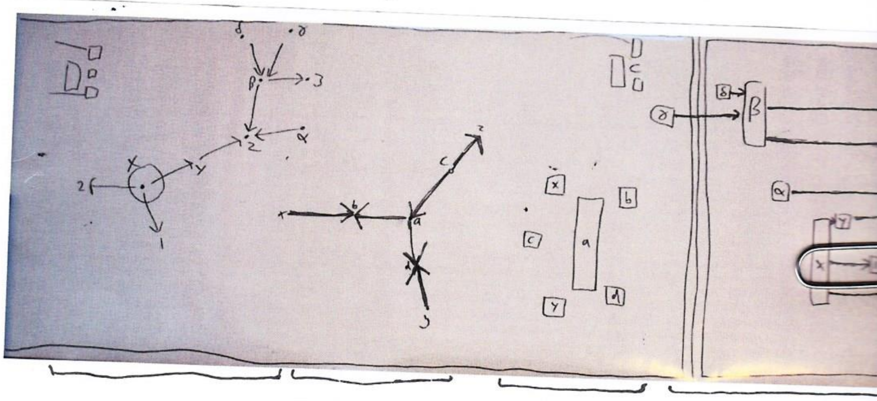
X: So that'll depend on things like how a word's been used recently, like a priming effect. But also in this context there are multiple different meanings in a short space of time, so there must be something else, something like what kind of answer you're expecting, what will answer the question, what makes sense in the context.

Y: Hm. OK, so let's look at the context of each a bit. So in each case, some kind of suggestion has just been tabled, so the- rather than 'hearer' let's say audience, since this stuff is kind of directed to the room, so the speaker knows that the audience expects effects related to that

Original in colour

See GIF 3: Problem is not for *x*





.33.44.000

B: alright so the thing that's [bad
D: [OK I'm with you
B: … is a vertex with three switching vertices
C:
A: [what
B: or [more ((nods)) [or more
C: [a - vertex with three switching vertices?
F: cos like [in that with that vertex *a*
E: [((unintelligible)) right there
F: … *b* is switching for *a*, [*d* is switching for *a* [and *c* is supposed to be switching for *a*
B: ((walks over, follows path *xba*)) [yup ((follows path *zca*))[yup
F: but we have nowhere to put – *z*
B: right
A: I hav- I have a problem.
B: oerr gahd
B: OK what's your problem ((laughing))
E: ((laughing)) not again
B: ((speaking through laughter)) here comes another one of A's counter-examples
A: it just seems… ((points at diagram on board)) isn't… wait, what happened to this… so
A: so… *z* was switching. … [but what if it just continued this way and it can switch ((follows path))
B: [Yuh. I-
A: but then a couple of things switched all at the- ((mimes two coming down from above))
B: so if there was
A: °s that a problem?°
B: if there was… ((walks over to board, A steps away))
B: the problem is not for *x*
A: O↑H↓
D: ((unintelligible)) rectangle
B: if there exists a: swi- a ↑vertex↓ from which three vertices are switching… then it's ↑bad↓. then
the whole graph's bad.
A: The whole graph's bad?
B: the- the tree cannot be represented if the vertex ((trails off))

Original in colour

context. In the first case, it's a problem with the STATEMENT that's just been presented to the room, and in the second it's a problem with representing the tree (which means that the STATEMENT holds).

X: Yeah. Their physical positions in the room seem to reinforce that. In 14, A is facing out into the room, and in 26 B has physically walked up to the diagram, B *physically moves* into the context of the diagram:

Y: So looking at the context the first two examples are talking about a flaw in *the definition of* unrepresentability, and the last one is talking about… a property that makes something unrepresentable. So if we try to pin down what's meant in the first, it's something like 'objection to the definition' or 'counter-example'—we can call that problem$_1$? And the final usage narrows down to 'unrepresentability', or 'specific locus of unrepresentability', and we can call it problem$_2$.

X: So they're really quite different, especially if you think about what's good for the group… the first usage, A's usage, really *is* a prob- going to cause difficulty for the work of the group, but B's usage is to show that because the diagram has a 'problem' then there *isn't* a new difficulty for the group's work! It seems like a heck of a perspective shift, to switch from what's a problem *for us* to what's a problem *for the diagram*

Y: Well, it's a bit more complicated than that because that second usage references the context of their previous work, wherein A proposed a counter-example that they described as a 'problem', that was a counter-example *because* it couldn't be represented. So the- there's a shared context in which the word 'problem' was recently used to refer to something that contradicted an expectation like problem$_1$, was unrepresentable like problem$_2$, *and* was a problem for the group's progress. And that old usage kind of encapsulates the two new narrowings, so they aren't *so* far apart.

X: Aah. So actually problem$_2$ is kind of an old problem, as far as the group's concerned. Ha. It's kind of… it's changed from a *problem* for them, frustration and all, into just a word to use to talk about a property of the diagram. It's like… they were in this world of trying to represent things as rectangle graphs, and in that world there was something that really was *bad* for their work, but now they've zoomed out into a world of trying to construct and evaluate STATEMENTS of definitions, and there's a new sense of 'problem' for the group, evaluated relative to the new aim, but the old usage still persists

Y: Like a sort of TELESCOPING, zooming out but LEAVING TRACES behind. Words like 'bad' or 'nice' pop up all the time in mathematics, with certain technical definitions. I wonder if this kind of thing is the reason.

X: Yes, and it strikes me that every time they are used, they still have this sense of referring to… people trying to do things, and that might even help people to connect with the content! To inhabit it, to get the sense of what people are trying to achieve, to get the sense of how something fits in with the aim of a piece of work. If a term hints at the role something played in that history of endeavour, then so much the better for the person reading and trying to make sense of it.

Y: Right. It seems relevant perhaps that there's evidence that really funny bits of polysemy can come into play in comprehension. There's been evidence that different but related meanings of words can 'prime' one another for faster comprehension (Williams, 1992), and that this effect is dependent on whether the meanings are related (MacGregor et al., 2015). And then there's research showing really surprising stuff that seems to relate interestingly to these ideas about

01.13.05.750

A: I'm starting to believe that this could be written [up as an algorithm

B: [that we have a proof of the lemma, that we have outlined the proof of the lemma

A: that- that- so if there are no 3-bad vertices is that what we're [calling them]

?: or four

?: [or four]

?: I dunno. Just if there [are no]

A: [just if there] are no k-bad vertices for >k greater than or equal to< THREE … then… th exists… >oh wait< then…

B: then this algorithm works to convert this orienting path cover into a rectangle

E: algorithm to be written

B: alright so the thing that's [bad
D: [OK I'm with you
B: … is a vertex with three switching vertices
C: [what
A: ((nods)) [or more
B: or [more
C: [a - vertex with three switching vertices?
F: cos like [in that with that vertex *a*
E: [((unintelligible)) right there
F: … *b* is switching for *a*, [*d* is switching for *a* [and *c* is supposed to be switching for *a*
B: ((walks over, follows path *xba*)) [yup ((follows path *zca*))[yup
F: but we have nowhere to put – *z*
[illegible]: right
A: I hav- I have a problem.
B: oerr gahd
B: OK what's your problem ((laughing))
E: ((laughing)) not again
B: ((speaking through laughter)) here comes another one of A's counter-examples
A: it just seems… ((points at diagram on board)) isn't… wait, what happened to this… so
A: so… *z* was switching. … [but what if it just continued this way and it can switch ((follows path))
B: [Yuh. I-
A: but then a couple of things switched all at the- ((mimes two coming down from above))
B: so if there was
A: °s that a problem?°
B: if there was… ((walks over to board, A steps away))
B: the problem is not for *x*
A: O↑H↓
D: ((unintelligible)) rectangle
B: if there exists a: swi- a ↑vertex↓ from which three vertices are switching… then it's ↑bad↓. then
the whole graph's bad.
A: The whole graph's bad?
B: the- the tree cannot be represented if the vertex ((trails off))

Original in colour

words with emotive content, like studies that show that positive-valenced words are understood more quickly when they're positioned high up, and negative when they're positioned low down (Meier & Robinson, 2004). That kind of thing can influence our interpretation in really subtle ways.

X: Yeah, and so… 'bad' started off being a value judgement, they were frustrated that the representation didn't work, and then right near the end of the meeting we see this kind of more formalised statement that they've settled on, which is still using the word 'bad'!

Y: Shall we take a look at that?

The Outcome

X: So they… as the meeting wears on, the group keeps stating and testing these sentences. And here they've come up with this, *the tree cannot be represented if there exists a vertex from which three vertices are switching.*

Y: OK, and do they refine it any more after that? It would be interesting to see what they are aiming for, if you see what I mean. The animal from the foliage.

X: At the close of the meeting, it's been refined into another line: I'm going to have to paraphrase a bit, but it goes like this: '*If there are no k-bad vertices for k greater than or equal to three then [a certain algorithm works to convert this orienting path cover into a rectangle]*'

Y: And we think that k-bad is vertices that lead on to k switching vertices, where the switching vertex is the first vertex in the path to switch

X: Yeah, and so it's when they're more than 2-bad that's a problem

Y: What was that? 2-bad?

X: Oh, yeah, 2-bad [writes it down]! You could hear it as this [writes too bad] or this [to bad]… I think the group also had difficulty with that, spoken it could be any of those! I guess that just goes to show that really the... the aim is to produce something in text form, something written.

Y: Oh, that's interesting. WRITING TRUMPS SPEAKING.

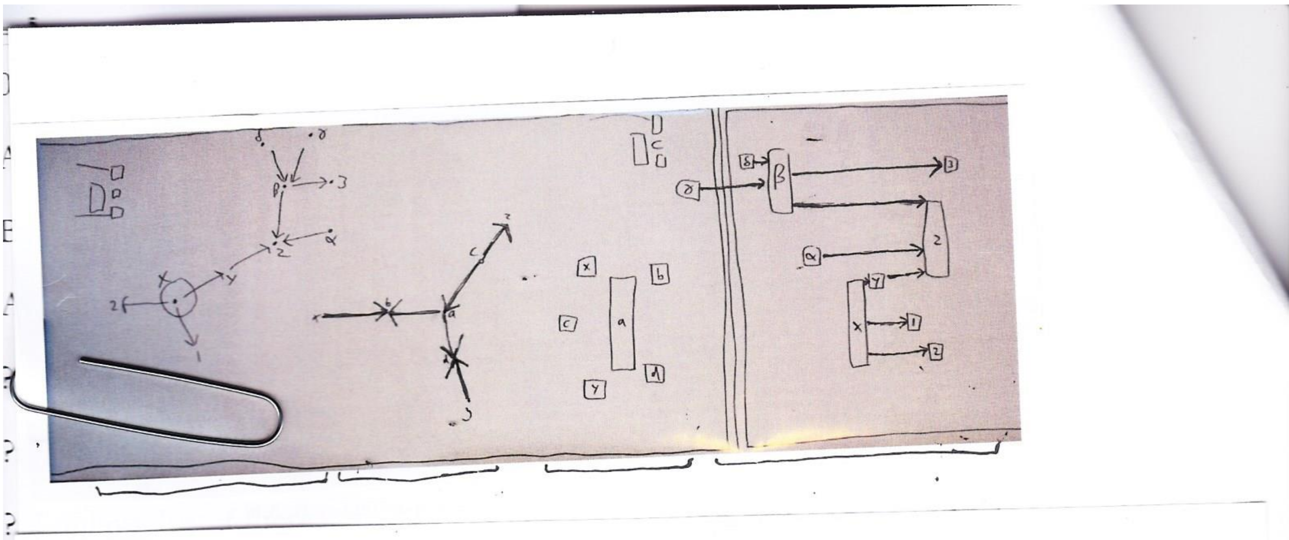
X: Yes, and there's also… they talk all the time about producing a lemma, that the outcome is a lemma, right from here 12 minutes in (00.12.52.000) to right at the end (01.13.05.750). They're aiming for a written form, what they have in mind throughout is how to produce it a solidified, shareable form.

Y: Yeah. This is also the mirror-image of the draft STATEMENT passed around during the meeting, stated in terms of what needs not to be the case for the graph to be represented, rather than what needs to be the case for the graph to be unrepresentable.

X: Oh, that's interesting. And before they wanted to *exemplify* what can't be represented, so they could stare at it, and that would be three or more. And now it's a STATEMENT about what *can* be done as long as it *isn't* three or more.

Y: Why do you think that is?

Original in colour



00.33.44.000

```
01 B:  alright so the thing that's [bad
02 D:                              [OK I'm with you
03 B:  ... is a vertex with three switching vertices
04 C:                                                  [what
05 A:                                   ((nods))       [or more
06 B:  or [more
07 C:     [a - vertex with three switching vertices?
08 F:  cos like [in that with that vertex a
09 E:           [((unintelligible)) right there
10 F:  ... b is switching for a,       [  d is switching for a [and c is supposed to be switching for a
11 B:  ((walks over, follows path xba)) [yup ((follows path zca)) [yup
12 F:  but we have nowhere to put – z
13 B:  right
14 A:  I hav- I have a problem.
15 B:  oerr gahd
16 B:  OK what's your problem ((laughing))
17 E:  ((laughing)) not again
18 B:  ((speaking through laughter)) here comes another one of A's counter-examples
19 A:  it just seems... ((points at diagram on board)) isn't... wait, what happened to this... so
20 A:  so... z was switching. ... [but what if it just continued this way and it can switch ((follows path))
21 B:                             [Yuh. I-
22 A:  but then a couple of things switched all at the- ((mimes two coming down from above))
23 B:                                                  so if there was
24 A:  °s that a problem?°
25 B:  if there was... ((walks over to board, A steps away))
26 B:  the problem is not for x
27 A:                       O↑H↓
28 D:  ((unintelligible)) rectangle
29 B:  if there exists a: swi- a ↑vertex↓ from which three vertices are switching... then it's ↑bad↓. then
30 the whole graph's bad.
31 A:  The whole graph's bad?
32 B:  the- the tree cannot be represented if the vertex ((trails off))
```

Original in colour

X: Well, because now the STATEMENT isn't *for* knowing what can't be represented, it's for knowing what *can*, because that's what they eventually want to do with the paper, that's what would interest, or could be used by, others, to show what *can* be represented.

Y: Oh, interesting. But now just this sentence is more kind of convoluted, with this negation at the beginning, the focus of the sentence being elsewhere. Lakatos talks about this in *Proofs and Refutations* (Lakatos, 1976), how definitions get adjusted and adjusted over time, they're the product of the efforts over time to find a proof. And this creates these intimidating, weird definitions where the thing you should focus on if you want to understand it is obscured. In this case, it's just that the STATEMENT no longer points *to* the examples they were discussing, it points *away from* them.

X: Yes, and I think that probably affects how easy it is to make sense of these STATEMENTS, and that can affect our reasoning. So for example: studies using the Wason selection task have shown that, all things being equal, participants will tend to choose the cards that have been mentioned, because they're nudged in that direction by an assumption that their interlocutor will be mentioning the most relevant parts (Girotto et al., 2001; Sperber et al., 1995; Mercier & Sperber, 2011; Allott, 2002). The trouble there is that the formulations of the task aren't aligned with our usual means of communicating, and their artificiality pushes participants to make mistakes. So in the case of mathematics, these adjusted framings will include and exclude the cases that they want, but they might also become more and more effortful to read, since they may not immediately suggest the most central case, and instead introduce a lot of convoluted conditions that point in many other directions. When STATEMENTS are adjusted in a distributed way, the ATTENTION DIRECTION is sometimes left as a TRACE.

Y: That might cast some light on what it means for language to become more technical, rather than conversational. Specialised formulations that have been subjected to these purposive adjustments are less intuitively framed than everyday speech, or just diverge more from ordinary communication the more it happens, and that might mean that successful interpretation becomes more reliant on conventional familiarity.

X: Purposive like, with a *different* purpose than conversational communication. Like the usefulness to the community that we discussed. On the other hand I quite like the way that their emotive descriptions have been baked in, the wording that refers to the history of the group's work, using 'bad' as a term.

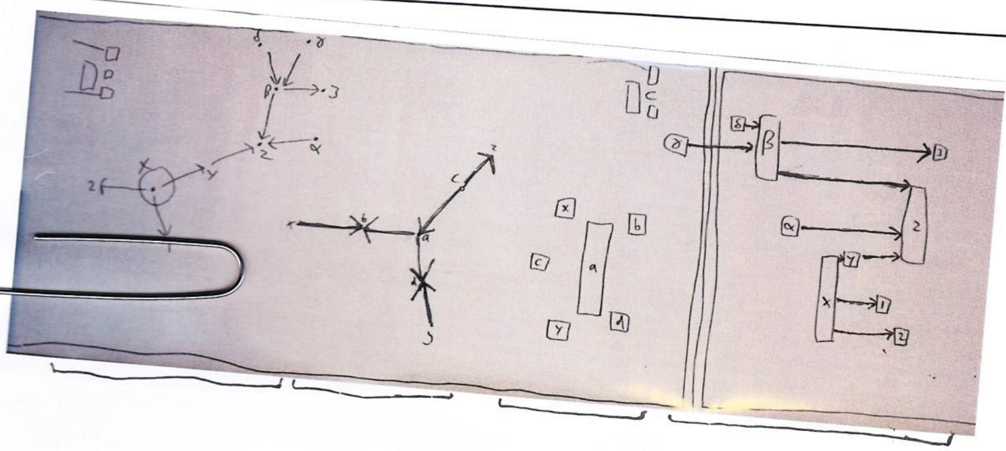
Y: Yeah. The description used is 'k-bad vertices for k greater than or equal to three', rather than simply 'greater than two'; this *does still* focus on the 'bad' cases in the same way that the meeting did, taking a particular case that was '3-bad' and examining it. It is interesting to look at the wording here, since there's nothing *necessary* about it- you could describe this problem in a multitude of different ways, like 'more than two folded vertices connected to one rectangle', 'maximum two double-ended vertices can be supported by one vertex'… If we used some inventive Greek-based terminology (from homo vs. hetero, and akri for tipped) we might say 'maximum two homoakrinous vertices can be held by one vertex'

X: At one point the group was calling them 'locations of incoherence'.

Y: Nice. But instead the language used throughout is that of switching, which- it's a term that worked really well with their earlier path-based description, it works really well for the perspective-taking method of counting that they talked about.

00.33.44.000

01 B: alright so the thing that's [bad
02 D: [OK I'm with you
03 B: … is a vertex with three switching vertices
04 C:
05 A: [what
06 B: or [more ((nods)) [or more
07 C: [a - vertex with three switching vertices?
08 F: cos like [in that with that vertex *a*
09 E: [((unintelligible)) right there
10 F: … *b* is switching for *a*, [*d* is switching for *a* [and *c* is supposed to be switching for *a*
11 B: ((walks over, follows path *xba*)) [yup ((follows path *zca*)) [yup
12 F: but we have nowhere to put – *z*
13 B: right
14 A: I hav- I have a problem.
15 B: oerr gahd
16 B: OK what's your problem ((laughing))
17 E: ((laughing)) not again
18 B: ((speaking through laughter)) here comes another one of A's counter-examples
19 A: it just seems… ((points at diagram on board)) isn't… wait, what happened to this… so
20 A: so… *z* was switching. … [but what if it just continued this way and it can switch ((follows path))
21 B: [Yuh. I-
22 A: but then a couple of things switched all at the- ((mimes two coming down from above))
23 B: so if there was
24 A: °s that a problem?°
25 B: if there was… ((walks over to board, A steps away))
26 B: the problem is not for *x*
27 A: O↑H↓
28 D: ((unintelligible)) rectangle
29 B: if there exists a: swi- a ↑vertex↓ from which three vertices are switching… then it's ↑bad↓. then
30 the whole graph's bad.
31 A: The whole graph's bad?
32 B: the- the tree cannot be represented if the vertex ((trails off))



Original in colour

X: Yeah, so it's tied to the ways that these very human actors have been moving around these diagrams and how they've been helping one another to 'see' them, there's a TRACE LEFT BEHIND. The sentence has in it this kind of record of the group's discovery and management of a counter-example, even in quite subtle ways.

Y: That's an interesting way of seeing it. I suppose it's- if I were to describe it to you using the word 'switching', then maybe that already gives you a sense of following paths along and changing direction, more than 'folded' or 'double-ended' or 'homoakrinous'.

X: Yes, so retaining the word 'switching' might well be just enough to nudge a reader of the eventual paper into looking at a graph rather like the way they did in the meeting, and so reaching an understanding similar to theirs. And that isn't – it isn't that the language here *isn't technical enough* somehow, it's just in the nature of technical language, that it's borne from this stepwise refinement of language from situated, impressionistic ways of communicating gradually toward something more general and portable that can become part of the community's work. It could be expressed otherwise. But the very small details of its expression might help or hinder the creation of shared understanding in subtle and generally unseen ways.

Y: Hm. Especially if- so we basically think that the outcome of this work is this STATEMENT, right? That they were working toward this portable, repeatable STATEMENT, which was this *way to direct attention around* a whole set of possible examples, where they- they worked with these two examples, but they're supposed to *stand in* for all kinds of other graphs. And what they want is this short collection of words that'll help somebody to see things – 'the problem' – in the right way.

X: Yeah, though actually the reader of a paper might never even care that much about which examples are excluded. The STATEMENT might just draw a line around what can be done. But it *can be found* again.

Y: On the other hand, we were talking about how convoluted the STATEMENT had got, and Lakatos' observation about the development of definitions as they're passed around. And that's kind of unavoidable because mathematics is this collaborative activity across a whole *big* group.

X: Yes, the process of argument and counter-argument is greatly extended in the mathematical community. You see clarifications and counter-examples being traded over the course of many years, and definitions are greatly refined and developed by this process.

Summary

Y: OK, so are we finished?

X: I think so, more or less. We've gone through our own process of making sense of things, and tried also to make sense of how our subjects understood one another

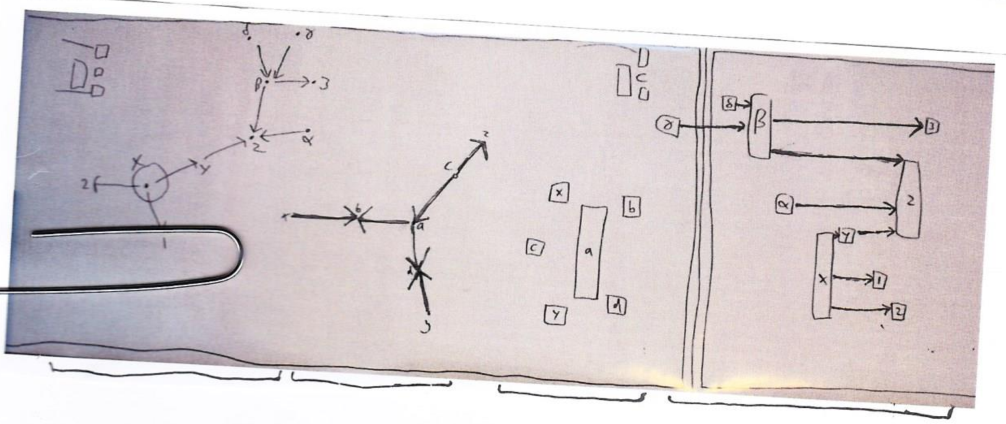
Y: Which I guess is sort of a shortcut to seeing more clearly how we understood what they were doing.

X: Right. And how we read them, or read their actions in the shared environment, as part of the PEOPLE-ENVIRONMENT ASSEMBLAGE. And we found that the subjects DIRECTED ONE ANOTHER'S ATTENTION around the shared materials in really subtle ways, even when what they were trying to convey was not something they could have paraphrased, it was more like SHARING A MENTAL STATE, a way of seeing. And encapsulating the way of seeing as a STATEMENT was part of the work of bringing the animal from the trees.

Original in colour

00.33.44.000

01 B: alright so the thing that's [bad
02 D: [OK I'm with you
03 B: … is a vertex with three switching vertices
04 C:
05 A: [what
06 B: or [more ((nods)) [or more
07 C: [a - vertex with three switching vertices?
08 F: cos like [in that with that vertex *a*
09 E: [((unintelligible)) right there
10 F: …, *b* is switching for *a*, [*d* is switching for *a* [and *c* is supposed to be switching for *a*
11 B: ((walks over, follows path *xba*)) [yup ((follows path *zca*)) [yup
12 F: but we have nowhere to put – *z*
13 B: right
14 A: I hav- I have a problem.
15 B: oerr gahd
16 B: OK what's your problem ((laughing))
17 E: ((laughing)) not again
18 B: ((speaking through laughter)) here comes another one of A's counter-examples
19 A: it just seems… ((points at diagram on board)) isn't… wait, what happened to this… so
20 A: so… *z* was switching. … [but what if it just continued this way and it can switch ((follows path))
21 B: [Yuh. I-
22 A: but then a couple of things switched all at the- ((mimes two coming down from above))
23 B: so if there was
24 A: °s that a problem?°
25 B: if there was… ((walks over to board, A steps away))
26 B: the problem is not for *x*
27 A: O↑H↓
28 D: ((unintelligible)) rectangle
29 B: if there exists a: swi- a ↑vertex↓ from which three vertices are switching… then it's ↑bad↓. then
30 the whole graph's bad.
31 A: The whole graph's bad?
32 B: the- the tree cannot be represented if the vertex ((trails off))



Original in colour

1.

Y: Right, but then as the group zoomed out from one field of enquiry to a broader one, the STATEMENT kind of… solidified. The group still understood it, but this was rather more because they've repeated it and repeated it, they kind of *knew* it rather than *understood* it

X: Right, and the STATEMENT became more of an artefact than the genuinely communicative, contextual utterance that they began with. And this might explain the sense that we had that WRITING TRUMPS SPEAKING- if they were aiming toward a solidified, portable version in a paper then the expectation that the written form is the defining one makes sense. And yet, still it bore the TRACES of this genuinely mental-state-sharing function it once had.

Y: Right. And while we had access to the marks they'd made on the board, and the words they'd said, it wasn't just understanding the definitions of all of those that helped us to understand what was going on. We spent most of this time talking about where the attention of the group was, what was being emphasised, and working out what would make sense as an answer given the group's particular aims and motivations.

X: Yes! So we've seen that what is being communicated goes far beyond the content of the diagrams or even of the sentences uttered—they're using these really subtle strategies to shift each other's *way of seeing* the problem.

Y: Great! So shall we wrap up?

X: Sure! Til the next time.

3.4. Summary

This analysis is written in the form of a script for a conversation (after Imre Lakatos (1976) and Douglas Hofstadter (1980)), in which two characters, X and Y, make sense of the meeting and reflect on how they have done so. The concept of a mathematical STATEMENT is introduced and the definition of this term is developed through the course of the meeting by the interlocutors. The fictional pair discuss the content of the excerpt first from a lay perspective, then with a context of basic mathematical knowledge, and then extend their scope to include the context of the rest of the meeting and the context of the mathematical world (including bringing in an expert to answer questions). Some of the key observations from the discussion are as follows.

The reasoning can be described as taking place in a people-environment assemblage, a cognitive system of the kind described in *Cognition in the Wild* (Hutchins, 1995). The collaborating group and its physical environment (in particular the whiteboard) are all taking part in a process of reasoning, testing and development. The group operates under a shared system of definitions, which are themselves under constant revision (take for example the meaning of 'bad'), and this allows them to coordinate their actions toward a common goal, sometimes playing the role of proposing an idea, sometimes challenging or pushing for clarification. Material resources are particularly notable when participants interact with diagrams on the whiteboard, the external representation clearly playing an active part in their reasoning and communication in an essential way.

Participants are observed producing *attention-directing narrations* that guide the perception of their colleagues in such a way as to share a way of seeing a diagram or problem. Subtle shifts in wording send the group's eyes and minds around a diagram, causing them to pay slightly more or less attention to different parts or aspects of the form. In this way, a speaker does something like 'thinking out loud' in interaction with the diagram, and doing this gives a hearer a kind of access to the speaker's subtle mental states, including confused or partial understandings that would be difficult to consciously summarise. In this way speakers seem able to open up a way of looking at the problem for correction or challenge from the group without ever explicitly describing it. This is the kind of phenomenon that is difficult to account for using a code model framework, and that may be better explained with an ostensive-inferential model.

Indeed often what my participants actually *said* was hardly enough for them to understand one another, and yet communication was a success. In the rich shared context of the group's aims and history and with the shared external representations as a resource, it was possible for vastly underdetermined stimuli, like a couple of words or a hand-wave, to convey rich meaning to the group.

Interestingly, the messy back-and-forth of discussion turns out to shape even the relatively formal language of a boiled-down STATEMENT, the particulars of the terms used containing vestigial remnants of the group's earlier ways of looking at the problem. As such it is possible to see even a quite formal-sounding sentence as containing a kind of trace of the group's discovery and management of a counter-example, of the twists and turns of a research process.

A final interesting observation is that the group clearly took their work to be working in the direction of a written output, even without a written paper being yet in progress. Throughout the meeting participants make reference to their work building toward a 'Lemma', a clear indication that an eventual written paper is intended. Even more remarkably, the group agree on a piece of terminology, '2-bad', that is ambiguous when spoken and even causes confusion within the meeting (since it could be understood as 'too bad', or 'to bad'), but is unambiguous and informative when written. This demonstrates that to some extent, their work is perceived as essentially geared toward producing a written output.

Interlude 4. Shared writing

Collaborators can meet over Skype, but it has its limitations. At times, the mathematicians really need to share something that they are able to write, but unable to say. To get around this problem, collaborators will sometimes scribble expressions or diagrams down on pieces of paper and hold them up to the camera, a creative way to approximate the missing shared writing space of the chalk- or whiteboard.

Subject L

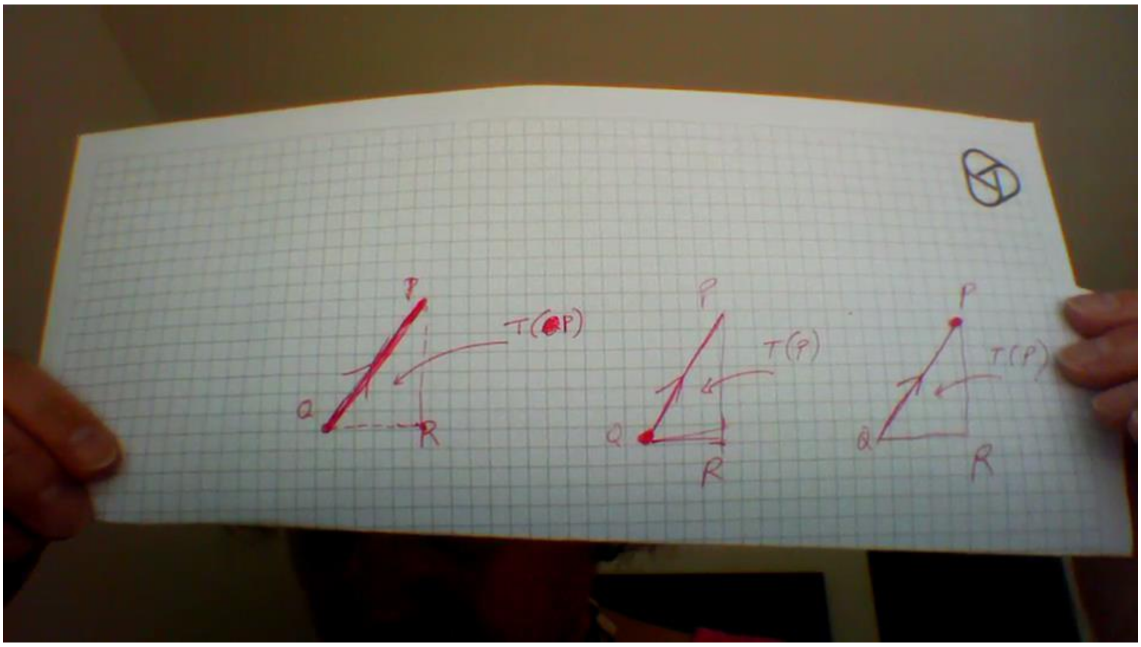


Figure 28. Holding up a diagram to share with a collaborator in the course of a Skype call. Original in colour.

Subject M

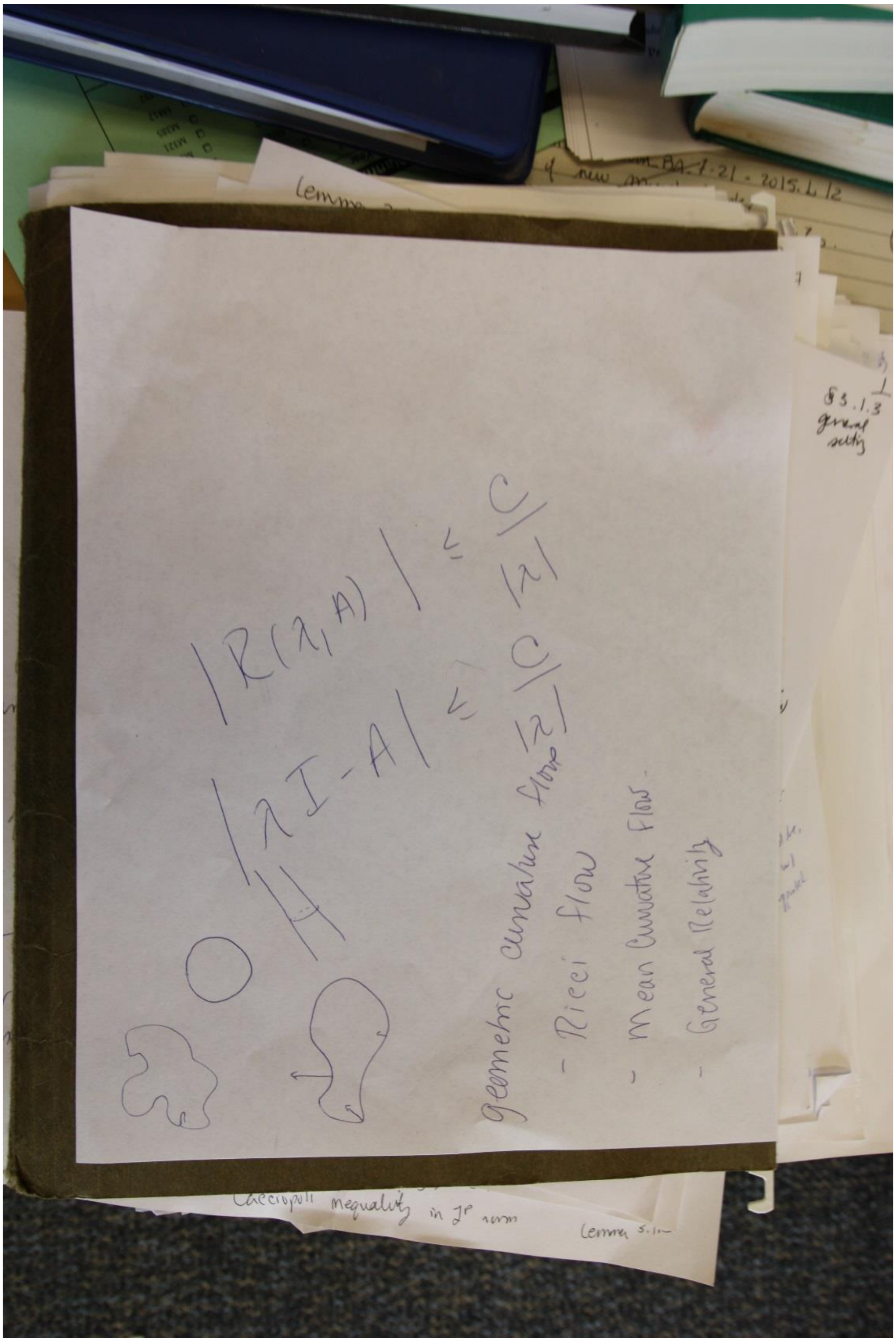


Figure 29. Notes scribbled on a scrap of paper and held up to the camera during a Skype meeting. Original in colour.

Napkin notes from subject J

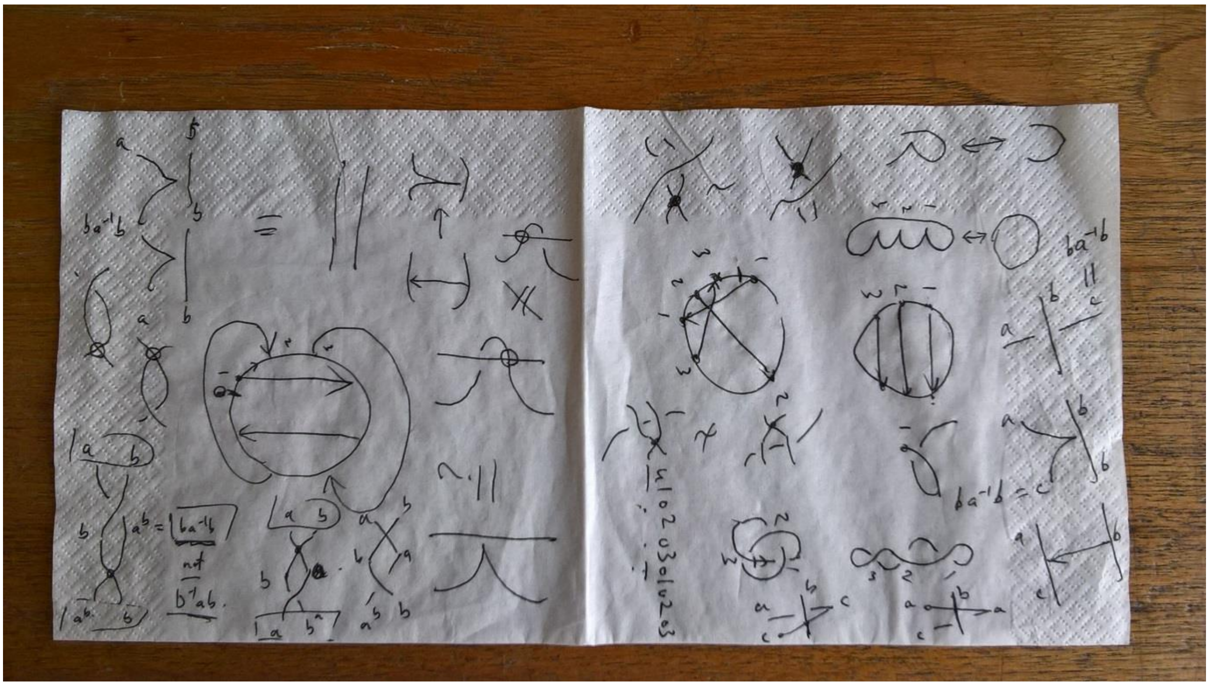


Figure 30. Notes scribbled on a napkin while discussing a project in a canteen. Original in colour.

Work site 4

Work site 4 was the office of a lecturer at a university in the UK.

Figure 31. Photocopiers in a corridor space at the university outside my subject's office, a common meeting place for colleagues. Original in colour.

Figure 32. Comfortable seating near the lifts, another space regularly used for casual discussion of research. Original in colour.

Figure 33. My participant's office space, including shelves of books and notes and a table and chairs to use with visitors. Original in colour.

4. Analysis of an excerpt from an email exchange, leading to edits to a paper

The second excerpt is taken from an email exchange and the Dropbox history of edits to a paper.

The excerpt analysed (see 4.3.1. Transcripts for the full transcript) is a record of an exchange between two collaborators as they brought a paper that they were working on to a state in which they felt happy to submit it. The edit history showed that the two of them had contributed sections to make up a paper and had been taking turns to refine the document. As well as making changes and marked commentary in the document, the collaborators also exchanged emails in which they discussed what they were doing in greater detail. Unlike some other mathematicians the participant whom I interviewed reported often working directly into LaTeX, the popular typesetting programme that the pair were using, and they are refining both the presentation and the content at this writing stage. The two participants are known as G and H.

4.1. Record of a first time through

4.1.1. Carrying out the analysis

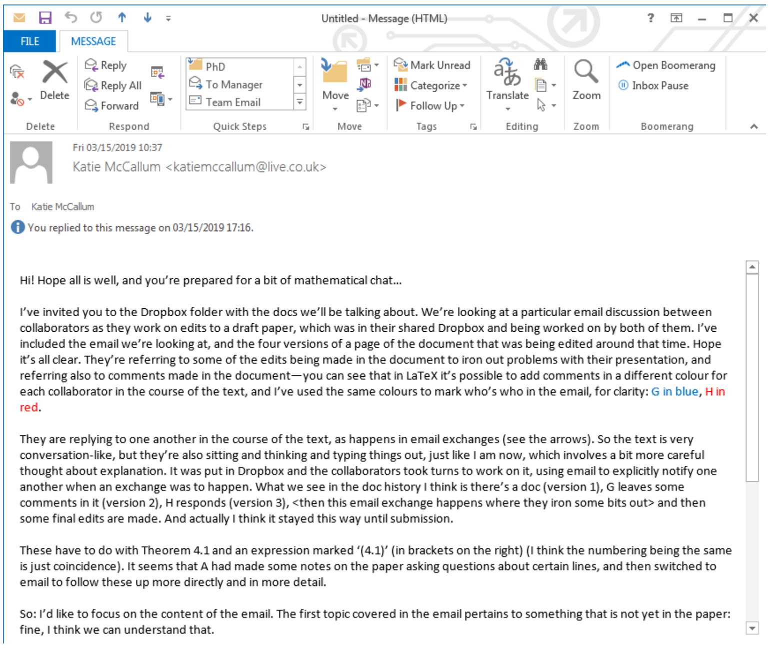
The situation analysed in this case had two mathematicians using multiple shareable inscriptions as a conduit for a collaborative process, whose aim was to refine a textual presentation of their work. The participants kept this latter component in a shared place, taking turns to make changes to it that the other was then able to see, while also producing textual interventions in the document that were explicitly directed to one another, set aside by colours and markers as commentary. In addition, the participants maintained a separate textual channel where needed for direct communication with one another about tricky questions, wherever this was needed.

The analysis of this excerpt was quite challenging. The mathematical work was at a much later stage than the previous excerpt, and the fact that this non-diagrammatic topic was much less visual played a part. My observation was also limited to text-based media. Each of these made understanding depend much more heavily on the text itself, and on the extent to which I knew the definitions of terms or notations.

Another limiting factor was the extent to which I was able to recognise which sections, or even characters, of an expression would prove to be key. I ran into a serious obstacle in understanding an explanation that one of my participants had sent to me because of not appreciating the significance of different types of brackets: ‘()’ vs. ‘[]’. It is easy to imagine that a face-to-face explanation could give a novice guidance through these kinds of details, by calling attention in myriad subtle ways to certain aspects of the written materials.

Again, I documented my path through the material, this time writing my understanding into correspondence between my two characters as I went along. My limitations at a given moment were reflected in the pauses and struggles of the characters. I began by working into email (see Figure 34), typing out my thoughts and sending them to myself, examining the ways in which this form limited my own explanation in just the ways I was finding my understanding of my participant’s explanation limited: the restrictive nature of typesetting does not leave much space for emphasis, crossings-out, or the organising marks we saw in the previous chapter. As an experiment I migrated my correspondence to paper (see Figure 35). I wanted to examine the contrast with the clean text I was analysing but also to recognise the continuity between this older form and the newer one; letters are the form that arguably inspired and still shapes our email behaviour, an older form whose advantages we are still trying to preserve even in newer forms with more versatility.

As I went, I was also annotating the excerpts being discussed, just as my participants had been annotating their document, but I was doing so by hand while cross-referencing with the tentative explanations I was writing out for myself. As I packed up a letter in an envelope with the annotated documents under discussion, imaginatively sending it on to find its reply, I reflected on the parallels and differences between a document shared in a Dropbox and a document physically sent back and forth: the latter can be spread across a desk, cross-referenced, will bear the history of the notes made upon it on its face; the former can be polished and adjusted and itself move toward a ‘front’ presentation, no final writing-up required, the perfected final document simply the aggregate of a long history of edits.



Untitled - Message (HTML)

Fri 03/15/2019 10:37
Katie McCallum <katiemccallum@live.co.uk>

To Katie McCallum

You replied to this message on 03/15/2019 17:16.

Hi! Hope all is well, and you're prepared for a bit of mathematical chat…

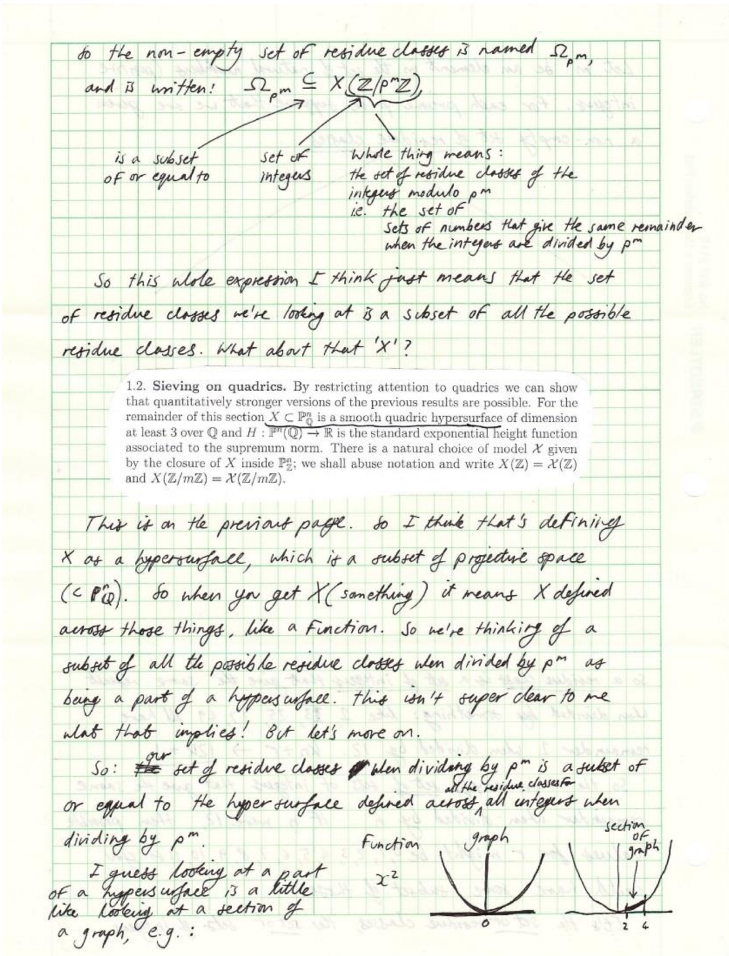
I've invited you to the Dropbox folder with the docs we'll be talking about. We're looking at a particular email discussion between collaborators as they work on edits to a draft paper, which was in their shared Dropbox and being worked on by both of them. I've included the email we're looking at, and the four versions of a page of the document that was being edited around that time. Hope it's all clear. They're referring to some of the edits being made in the document to iron out problems with their presentation, and referring also to comments made in the document—you can see that in LaTeX it's possible to add comments in a different colour for each collaborator in the course of the text, and I've used the same colours to mark who's who in the email, for clarity: G in blue, H in red.

They are replying to one another in the course of the text, as happens in email exchanges (see the arrows). So the text is very conversation-like, but they're also sitting and thinking and typing things out, just like I am now, which involves a bit more careful thought about explanation. It was put in Dropbox and the collaborators took turns to work on it, using email to explicitly notify one another when an exchange was to happen. What we see in the doc history I think is there's a doc (version 1), G leaves some comments in it (version 2), H responds (version 3), <then this email exchange happens where they iron some bits out> and then some final edits are made. And actually I think it stayed this way until submission.

These have to do with Theorem 4.1 and an expression marked '(4.1)' (in brackets on the right) (I think the numbering being the same is just coincidence). It seems that A had made some notes on the paper asking questions about certain lines, and then switched to email to follow these up more directly and in more detail.

So: I'd like to focus on the content of the email. The first topic covered in the email pertains to something that is not yet in the paper: fine, I think we can understand that.

Figure 34. Early email correspondence. Original in colour.



So the non-empty set of residue classes is named Ω_{p^m},
and is written: $\Omega_{p^m} \subseteq X(\mathbb{Z}/p^m\mathbb{Z})$

is a subset of or equal to

set of integers

Whole thing means: the set of residue classes of the integers modulo p^m
ie. the set of sets of numbers that give the same remainder when the integers are divided by p^m

So this whole expression I think just means that the set of residue classes we're looking at is a subset of all the possible residue classes. What about that 'X'?

1.2. **Sieving on quadrics.** By restricting attention to quadrics we can show that quantitatively stronger versions of the previous results are possible. For the remainder of this section $X \subset \mathbb{P}^n_{\mathbb{Q}}$ is a smooth quadric hypersurface of dimension at least 3 over $\mathbb{Q}$ and $H : \mathbb{P}^n(\mathbb{Q}) \to \mathbb{R}$ is the standard exponential height function associated to the supremum norm. There is a natural choice of model $\mathcal{X}$ given by the closure of X inside $\mathbb{P}^n_{\mathbb{Z}}$; we shall abuse notation and write $X(\mathbb{Z}) = \mathcal{X}(\mathbb{Z})$ and $X(\mathbb{Z}/m\mathbb{Z}) = \mathcal{X}(\mathbb{Z}/m\mathbb{Z})$.

This is on the previous page. So I think that's defining X as a hypersurface, which is a subset of projective space ($\subset \mathbb{P}^n_{\mathbb{Q}}$). So when you get X(something) it means X defined across those things, like a function. So we're thinking of a subset of all the possible residue classes when divided by p^m as being a part of a hypersurface. This isn't super clear to me what that implies! But let's move on.

So: ~~the~~ our set of residue classes when dividing by p^m is a subset of or equal to the hypersurface defined across all the residue classes for all integers when dividing by p^m.

I guess looking at a part of a hypersurface is a little like looking at a section of a graph, e.g.:

Function x^2 graph — section of graph

Figure 35. Page from a letter. Original in colour.

My letters were full of colours, crossings-out, ink splodges, and the complex arrangements of subscripts and superscripts that make up mathematical notation and remain outside of the grasp of the basic text form of emails. As in Interlude 4, I found that my participants had come up with a fascinating workaround to solve their pressing need for the capabilities of nothing more complex than a pen. In the interlude, the mathematicians, unable to write in front of one another on a blackboard or whiteboard, scribbled on pieces of paper and held them up to the camera; in this excerpt the mathematicians, needing to construct complex expressions and limited to plain text, simply used the textual codes used as inputs on the 'back end' of LaTeX, codes that would be very familiar and intuitive to the mathematicians who use LaTeX every day.

In the end, I simply did find that my eventual understanding of the material was more limited than in the previous example. I was, however, able to conclude that the difficulties being addressed by the participants were somehow of at least two different kinds, and identifying this difference was something I found difficult to achieve according to the *first time through* principle being employed. One of these difficulties was what my participant termed a 'mathematical error', an invalid move in the statement of a proof; another posed no problem in terms of 'correctness' but demanded explanation for a reader of the paper not to find it confusing, a quite different kind of problem. In pursuit of a *first time through* explanation I attempted to focus in my discussion on what the participants were attempting to produce in their document in each case, in functional terms. In the first case, the two placed great value on adherence to the structure of agreed rules known to and accepted by the discipline, deviating from which would mean that a text would be would be invalidated as a part of mathematics, left disconnected and unable to play a part in the mathematical world. In the second, it was a case of manifesting the organisation of the argument, caring not only about whether the moves made were orderly but also about how evident that orderliness would be to an individual reader.

The question being investigated was that of the density of primes: how many numbers in a particular interval are prime. The pair were using a technique known as 'sieving', a pleasantly figurative term for a technique aimed at estimating the density of primes by gradually taking out all of the numbers with divisors to approach the point at which only primes are left.

This performed conversation is documented by the fictional correspondence in

4.3.2. Correspondence. A reader of this thesis can follow this conversation as a beginner's guide through the mathematics, explaining in detail the problem that the group are working on and how I as an observer came to understand it; this text will also serve as an introduction to the observations that will be discussed in detail in the Written Discussion, situating each of these in the conversation and in the group's work.

4.2. Breaching experiments

4.2.1. Reimagining the mathematical paper

What follows is from a workshop that I ran for an audience with specialisms in mathematics and in creative practice, at Bridges Conference for Mathematics and the Arts 2018 in Stockholm.

This workshop consisted of exercises designed to encourage new and experimental engagements with mathematical papers and reconsider the possibilities in their presentation. Participants were given a number of prompts for transgressive, creative engagements with existing published papers, physically taking the texts apart and finding new ways to put them back together.

4.2.1.1 Poetry

These poems were produced by slicing up and rearranging a page of a mathematical text to find new and unexpected meanings.

to implement but we shall use the idea. This procedure is tricky

the action of $\mathrm{Diff}(S^1)$ is discontinuous. hopelessly

using structures suggested by planar algebra. It is this relation

This is just block spin renormalisation.

monotone

for any monotone sequence
any sequence of balls

for all

such sequences

For any $t > 0$,

for any

$t > 0$

for all $t > 0$:

consider

Hence
Therefore
It follows that

This completes the proof

proponents of extreme values,

proved that infinite-time counterparts

are almost everywhere

Our respect to infinitely many conditions.

The problem is that the limit of the bounds are not realistic

Let X be false, not rational

As discovered by taking over the

one

irreducible

component

accumulating fibres is not special

The large sieve was employed

to understand how many varieties in the family are not realistic

A natural question is whether

this end is *split*, with the following types of behaviour

counting

following

removing

i.e. contains a multiplicity

suppose that we discuss a range of maps and balls

From a point of view

This is satisfying

This procedure is tricky to implement

We have no answer to the discontinuity problem.

hopelessly discontinuous.

we have done nothing

this construction fails to be of any interest

Figure 36. Poems produced by participants from a cut up mathematical text

4.2.1.2 Deconstruction Task

Participants were asked to find different methods for taking apart and classifying a few pages of a published paper, finding different ways to reconceptualise the content.

Figure 37. Participants examining a paper. Original in colour.

Some of the methods they came up with were as follows:

- Ordering sections according to symbolic content versus prose content
- Noticing which sentences were in the imperative mood
- Discussing alternative readings of particular sentences

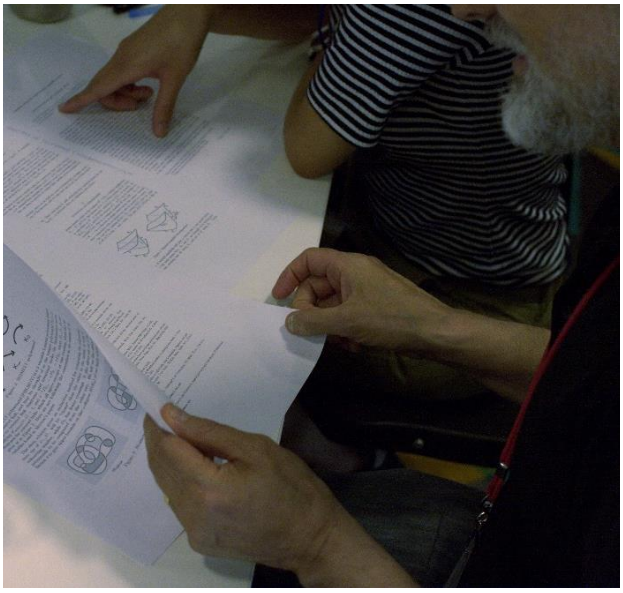
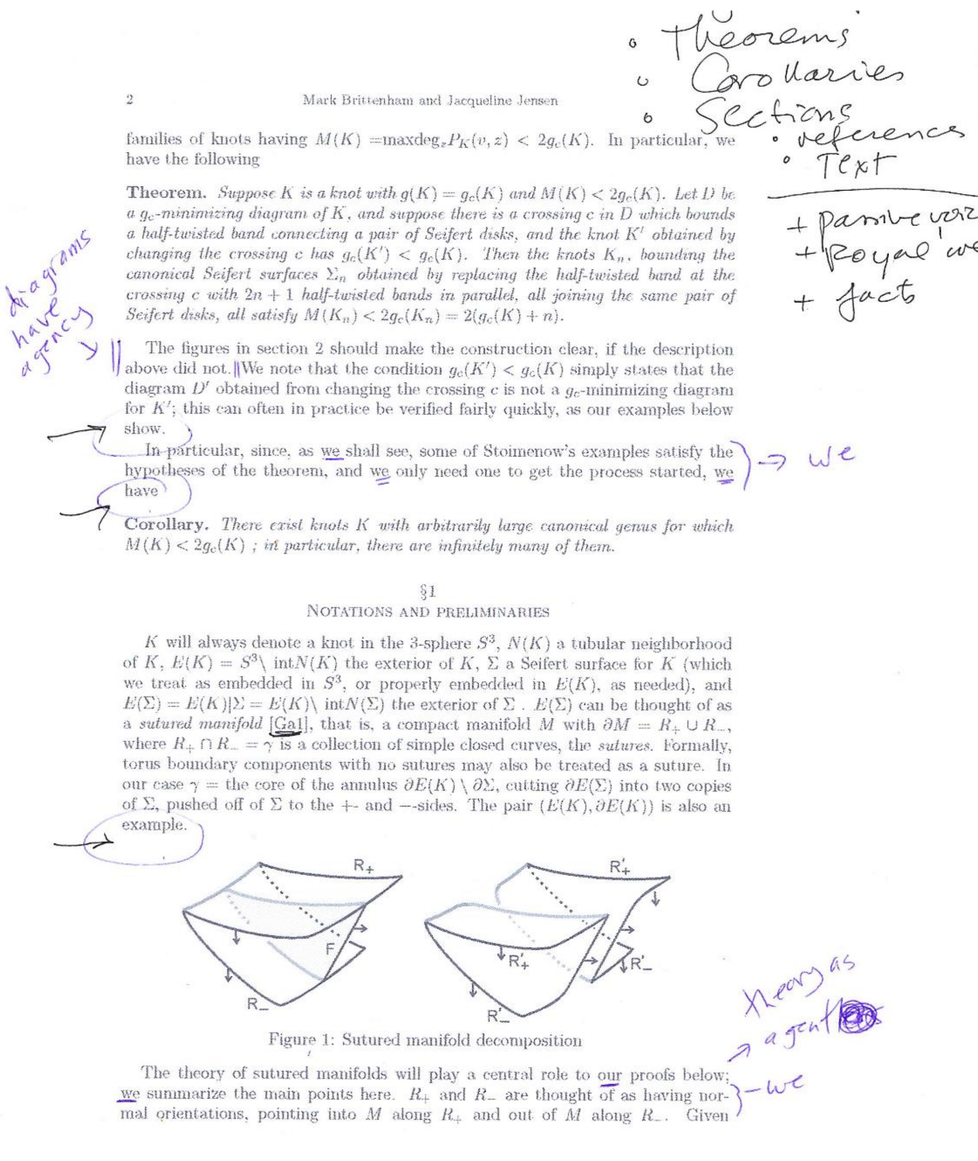


Figure 38. Participants examining a paper. Original in colour.

- Noticing the agency of the sentences: one group observed that often there was a rhythm wherein an agentless sentence was followed by a 'we'-formulation
- Considering different types of 'we': the exclusive we, the inclusive we, the royal we, and so on
- Finding five different ways to read a paper: backward, first sentence of each paragraph only, figure captions only, etc. A participant reported using this technique as part of the journal refereeing process, as a means to really 'see' the text in new ways.



2 Mark Brittenham and Jacqueline Jensen

families of knots having $M(K) = \text{maxdeg}_z P_K(v, z) < 2g_c(K)$. In particular, we have the following

Theorem. *Suppose K is a knot with $g(K) = g_c(K)$ and $M(K) < 2g_c(K)$. Let D be a g_c-minimizing diagram of K, and suppose there is a crossing c in D which bounds a half-twisted band connecting a pair of Seifert disks, and the knot K' obtained by changing the crossing c has $g_c(K') < g_c(K)$. Then the knots K_n, bounding the canonical Seifert surfaces Σ_n obtained by replacing the half-twisted band at the crossing c with $2n + 1$ half-twisted bands in parallel, all joining the same pair of Seifert disks, all satisfy $M(K_n) < 2g_c(K_n) = 2(g_c(K) + n)$.*

The figures in section 2 should make the construction clear, if the description above did not. We note that the condition $g_c(K') < g_c(K)$ simply states that the diagram D' obtained from changing the crossing c is not a g_c-minimizing diagram for K'; this can often in practice be verified fairly quickly, as our examples below show.

In particular, since, as we shall see, some of Stoimenow's examples satisfy the hypotheses of the theorem, and we only need one to get the process started, we have

Corollary. *There exist knots K with arbitrarily large canonical genus for which $M(K) < 2g_c(K)$; in particular, there are infinitely many of them.*

§1

NOTATIONS AND PRELIMINARIES

K will always denote a knot in the 3-sphere S^3, $N(K)$ a tubular neighborhood of K, $E(K) = S^3 \setminus \text{int}N(K)$ the exterior of K, Σ a Seifert surface for K (which we treat as embedded in S^3, or properly embedded in $E(K)$, as needed), and $E(\Sigma) = E(K)|\Sigma = E(K) \setminus \text{int}N(\Sigma)$ the exterior of Σ . $E(\Sigma)$ can be thought of as a *sutured manifold* [Ga1], that is, a compact manifold M with $\partial M = R_+ \cup R_-$, where $R_+ \cap R_- = \gamma$ is a collection of simple closed curves, the *sutures*. Formally, torus boundary components with no sutures may also be treated as a suture. In our case $\gamma =$ the core of the annulus $\partial E(K) \setminus \partial\Sigma$, cutting $\partial E(\Sigma)$ into two copies of Σ, pushed off of Σ to the +- and −-sides. The pair $(E(K), \partial E(K))$ is also an example.

R_+ F R_- R'_+ R'_+ R'_- R'_-

Figure 1: Sutured manifold decomposition

The theory of sutured manifolds will play a central role to our proofs below; we summarize the main points here. R_+ and R_- are thought of as having normal orientations, pointing into M along R_+ and out of M along R_-. Given

Figure 39. Participants' notes on a paper. Original in colour.

4.2.1.3 Reconstruction Task

The second task was for participants to put forward a reimagining of a particular six-page paper. They were invited to lay out and alter the paper in whatever way they saw fit, adding in colour and string connections to give articulation to the content.

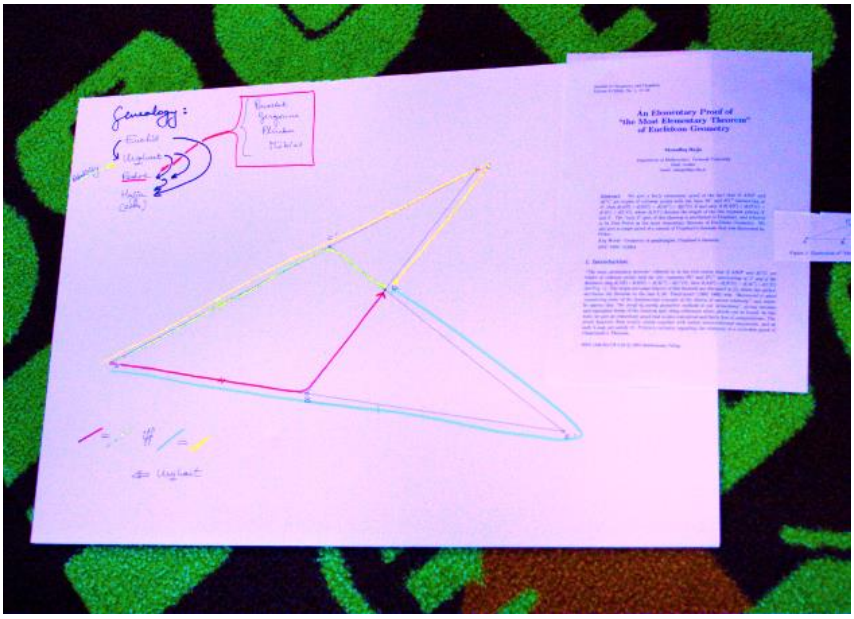


Figure 40. A colour-coding presentation of an argument, attempting to present the entire argument on one page. Original in colour.

Since all of the information is here presented simultaneously rather than sequentially, the coloured lines are used to direct the reader's attention to different sections in turn and so move through the argument.

Figure 41. Physical engagement with the parts of the text. Original in colour.

One group laid out the diagram and paper on a large sheet, such that reading the paper would cause a person to physically move around the diagram as the argument progresses.

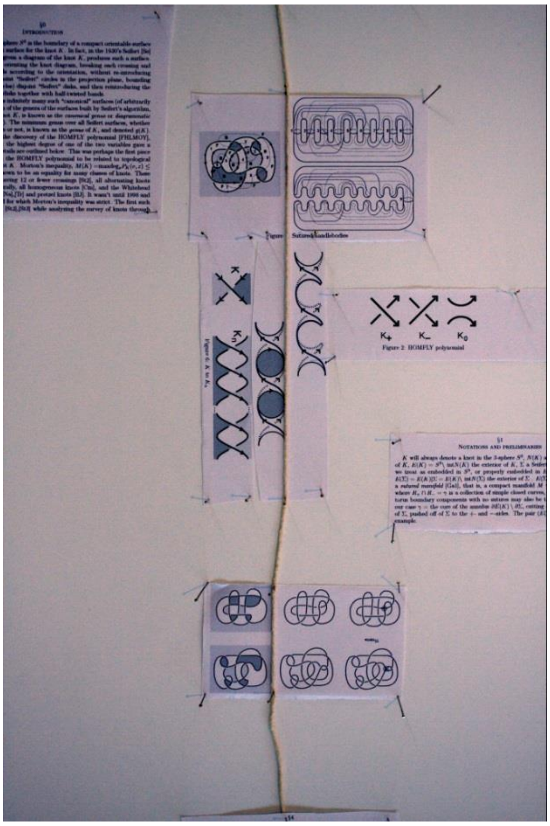


Figure 42. Classification system. Original in colour.

This group noted that the diagrams divided into those that were outlines and those that were filled in, and used this as a metaphorical classification system for the entire paper, dividing their workspace into two sections and arranging the text and images accordingly. Thus the section on notation is classified as 'outline', and the introduction 'fills in' the context for the paper.

Here the paper is presented as a comic strip, prioritising diagrams and adding a novel diagrammatic representation of the theorem, and bringing annotations into the diagram to obviate the need to switch between diagram and text.

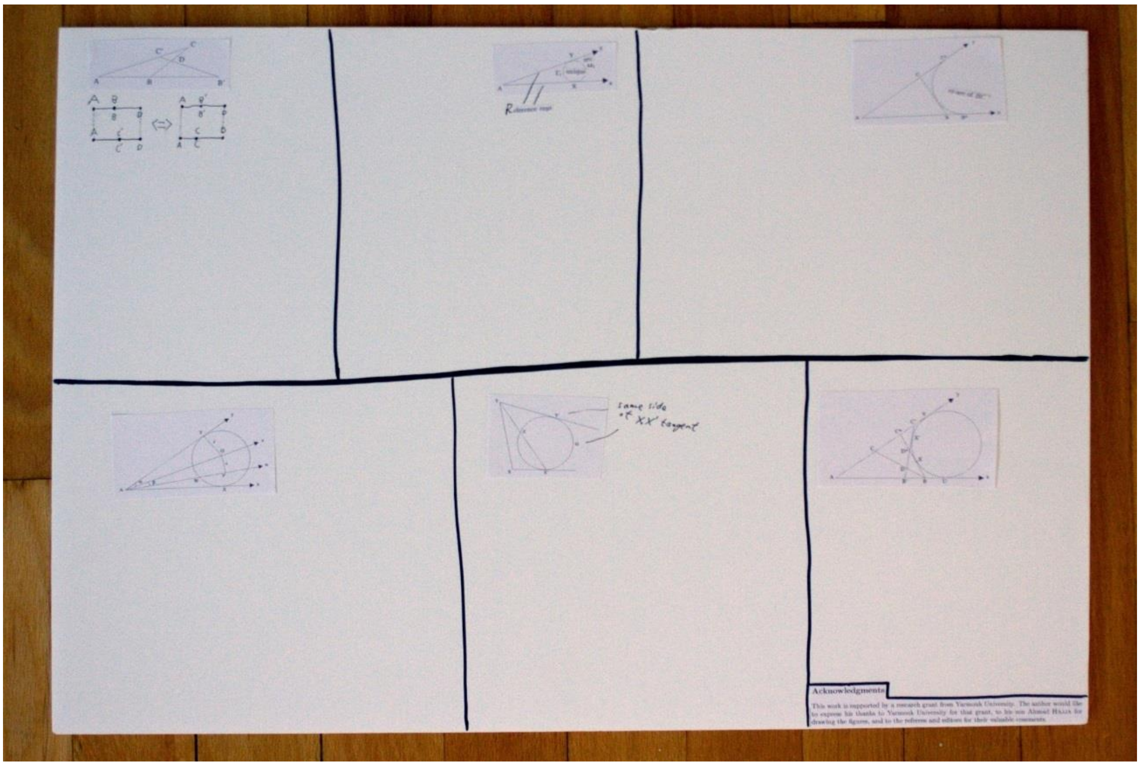


Figure 43. Comic strip. Original in colour.

This workshop was constructed with the aim of prompting a group to 'see' the content of a mathematical paper, a form with which many of them were intimately familiar, in a new and transgressive way. The groups' reconstructions of the paper ranged from re-presentations of the argument that arguably followed the aims of the original to quite inventive restructuring according to quite different classificatory logic. Perhaps the task that yielded the most surprises though was the poetry task; participants commented that they were amazed to find such emotional narratives popping out of a page from such an austere, impersonal-seeming text. This exercise really did seem to reveal an unseen undercurrent of emotion and endeavour running through mathematical writing, in the subtleties of turns of phrase and loaded terminology, obscured perhaps by the professional tone of the whole.

This experiment engaged largely with the content of the text, text itself being certainly an important medium to examine in the course of investigating mathematical publications. In the course of the analysis, though, I found that the back-end functioning of the particular typesetting medium of LaTeX even became visible in the texts I was analysing, which highlighted the central role that it plays in the practical production of mathematical texts. This medium, then, presented itself as a candidate for experimentation.

I present the results of this workshop with very great thanks to the participants:

Liora Butov

Carol Bier

Amenda Chow

Loe Feijis

Paul Gailiunas

Susan Gerofsky

Jana Kopfovà

Amy Selikoff

Donald Spector

Hamish Todd

Marco Torredimare

Martin Weissman

4.2.2. Pulling a paper apart

Since I had access to the LaTeX files used by the collaborators in this case study, I made a copy and began to experiment with them, pulling the paper apart to see what I could find or reshape within it.

TAKING PARTS APART

X AND Y

ABSTRACT. .
...
:::
::::::::

CONTENTS

1. INTRODUCTION

1.1. First.

1.1.1. *this.* We begin by
We work with the following.
A natural question is

We show that
If this is how
Then we can
We would like

Here is

Date: 1st May 2019.
2010 *Mathematics Subject Classification.* 14G05; 11N36, 11P55, 14D10,

Figure 44. Experiment with the beginning of a paper, removing content and leaving connecting phrases. Original in colour.

As in the poetry experiment above, my first move was to select parts of the text to remove, to rephrase and to keep, leaving short phrases that were strange and out of context and yet still aesthetically mathematical-seeming thanks to the typesetting efforts of the programme. The robustness of its attractive, symmetrical positioning even of expressions that I pushed into more and more nonsensical configurations was quite charming.

Proof.

$$\frac{\mathrm{d}}{\mathrm{d}x}(m(g\Bigg(x\bigg(m(n\Bigg(A\Bigg))\bigg)\Bigg))) \tag{2.9}$$

$$(m(g\Bigg(x\bigg(m(n\Bigg(A\Bigg))\bigg)\Bigg))) \tag{2.10}$$

$$(m(g\Bigg(x\bigg(m(n\Bigg(\Bigg))\bigg)\Bigg))) \tag{2.11}$$

$$(m(g\Bigg(x\bigg(m(\Bigg(\Bigg))\bigg)\Bigg))) \tag{2.12}$$

$$(m(g\Bigg(x\bigg((\Bigg(\Bigg))\bigg)\Bigg))) \tag{2.13}$$

$$(m(g\Bigg(\bigg((\Bigg(\Bigg))\bigg)\Bigg))) \tag{2.14}$$

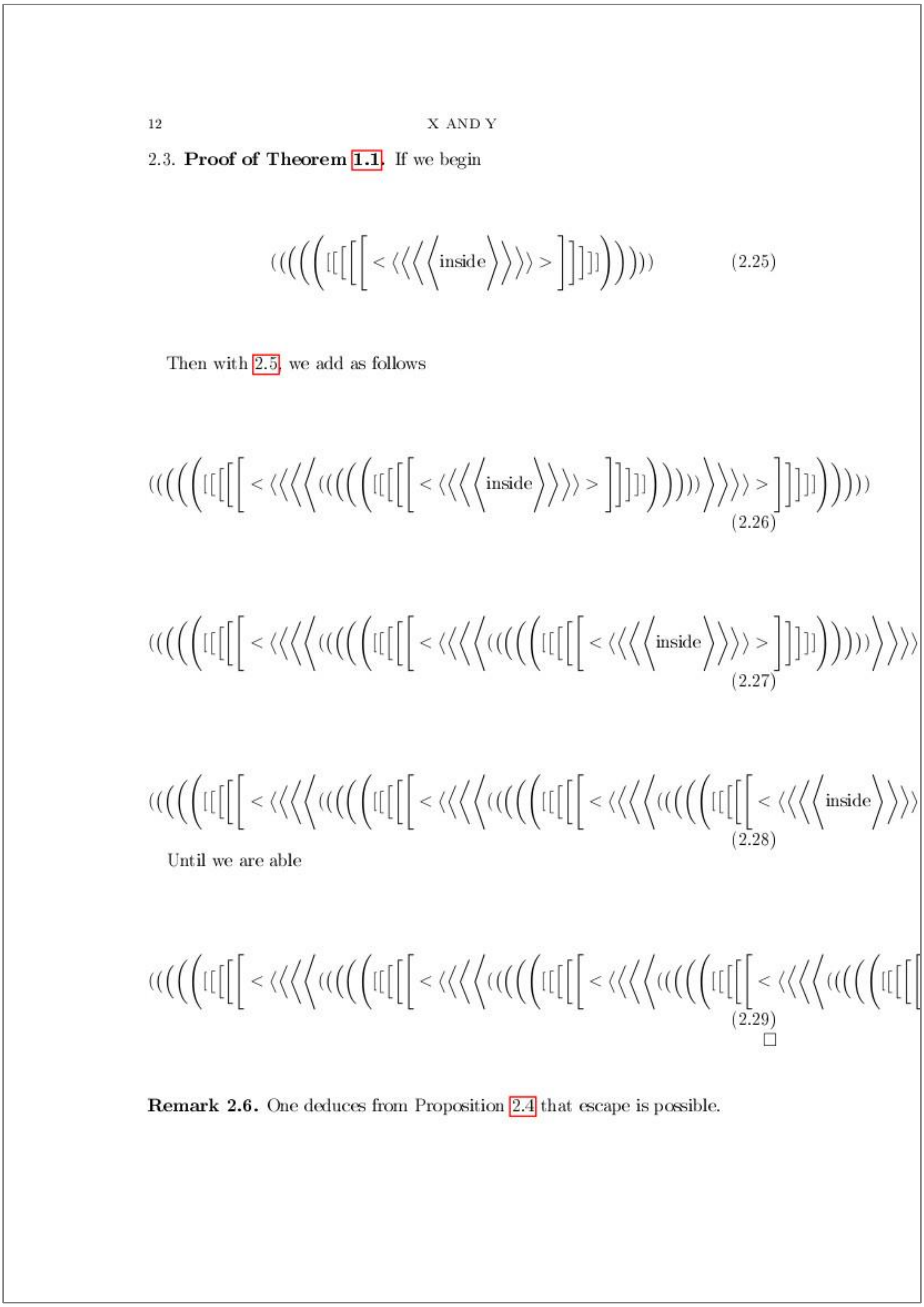
$$(m(\Bigg(\bigg((\Bigg(\Bigg))\bigg)\Bigg))) \tag{2.15}$$

$$((\Bigg(\bigg((\Bigg(\Bigg))\bigg)\Bigg))) \tag{2.16}$$

□

Figure 45. Symmetrical configurations of brackets, losing their content

It came to be a challenge, then, to push that system until it broke, until the absurd expressions began escaping the page.



2.3. **Proof of Theorem 1.1.** If we begin

$$((((\,[[[[\,<\langle\langle\langle\langle\,\text{inside}\,\rangle\rangle\rangle\rangle>\,]]]]\,)))) \tag{2.25}$$

Then with 2.5, we add as follows

$$((((\,[[[[\,<\langle\langle\langle\langle\,((((\,[[[[\,<\langle\langle\langle\langle\,\text{inside}\,\rangle\rangle\rangle\rangle>\,]]]]\,))))\,\rangle\rangle\rangle\rangle>\,]]]]\,)))) \tag{2.26}$$

$$((((\,[[[[\,<\langle\langle\langle\langle\,((((\,[[[[\,<\langle\langle\langle\langle\,((((\,[[[[\,<\langle\langle\langle\langle\,\text{inside}\,\rangle\rangle\rangle\rangle>\,]]]]\,))))\,\rangle\rangle\rangle\rangle \tag{2.27}$$

$$((((\,[[[[\,<\langle\langle\langle\langle\,((((\,[[[[\,<\langle\langle\langle\langle\,((((\,[[[[\,<\langle\langle\langle\langle\,((((\,[[[[\,<\langle\langle\langle\langle\,\text{inside}\,\rangle\rangle\rangle \tag{2.28}$$

Until we are able

$$((((\,[[[[\,<\langle\langle\langle\langle\,((((\,[[[[\,<\langle\langle\langle\langle\,((((\,[[[[\,<\langle\langle\langle\langle\,((((\,[[[[\,<\langle\langle\langle\langle\,((((\,[[[[\tag{2.29}$$

□

Remark 2.6. One deduces from Proposition 2.4 that escape is possible.

Figure 46. Pushing the software's layout capabilities to breaking. Original in colour.

Repeating,

$$x = a + \cfrac{1}{a + \cfrac{1}{a} + \cfrac{1}{a} + \cfrac{1}{a} + \cfrac{1}{a + \cfrac{1}{a + \cfrac{1}{a + \cfrac{1}{a}}}}} \tag{3.4}$$

$$x = a + \cfrac{1}{a + \cfrac{1}{a} + \cfrac{1}{a} + \cfrac{1}{a} + \cfrac{1}{a + \cfrac{1}{a} + \cfrac{1}{a} + \cfrac{1}{a} + \cfrac{1}{a} + \cfrac{1}{a} + \cfrac{1}{a + \cfrac{1}{a + \cfrac{1}{a}}}}} \tag{3.5}$$

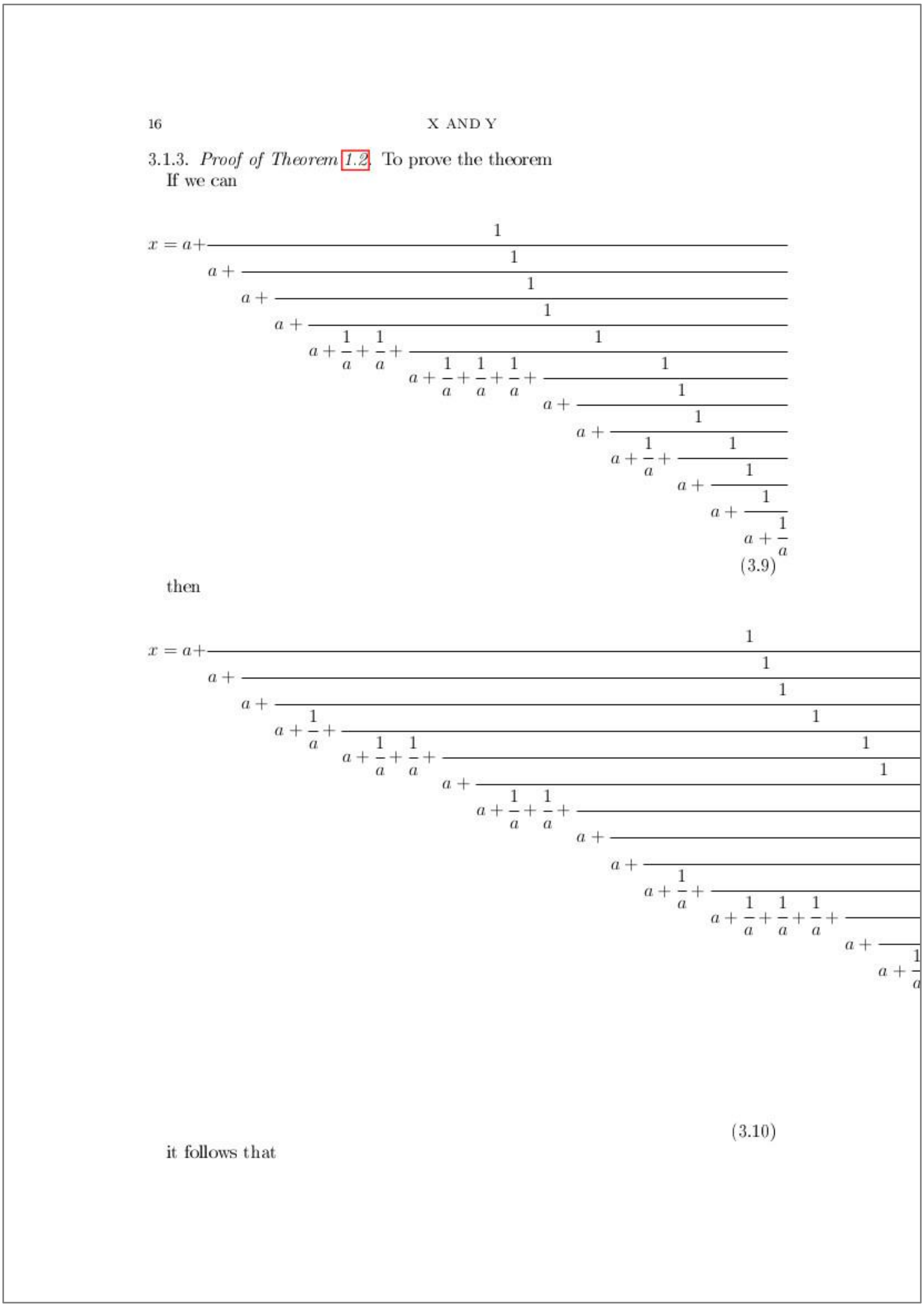
$$x = a + \cfrac{1}{a + \cfrac{1}{a + \cfrac{1}{a + \cfrac{1}{a + \cfrac{1}{a} + \cfrac{1}{a} + \cfrac{1}{a + \cfrac{1}{a} + \cfrac{1}{a} + \cfrac{1}{a} + \cfrac{1}{a + \cfrac{1}{a + \cfrac{1}{a + \cfrac{1}{a} + \cfrac{1}{a + \cfrac{1}{a + \cfrac{1}{a + \cfrac{1}{a}}}}}}}}}}} \tag{3.6}$$

□

Remark 3.3. We begin to see that we might overbalance, as in

$$x = a + \cfrac{1}{a + \cfrac{1}{a} + \cfrac{1}{a} + \cfrac{1}{a} + \cfrac{1}{a + \cfrac{1}{a} + \cfrac{1}{a} + \cfrac{1}{a} + \cfrac{1}{a} + \cfrac{1}{a} + \cfrac{1}{a + \cfrac{1}{a} + \cfrac{1}{a} + \cfrac{1}{a} + \cfrac{1}{a} + \cfrac{1}{a} + \cfrac{1}{a + \cfrac{1}{a + \cfrac{1}{a}}}}}} \tag{3.7}$$

Figure 47. Increasingly absurd fractions



16 X AND Y

3.1.3. *Proof of Theorem 1.2.* To prove the theorem
If we can

$$x = a + \frac{1}{a + \frac{1}{a + \frac{1}{a + \frac{1}{a + \frac{1}{a} + \frac{1}{a} + \frac{1}{a + \frac{1}{a} + \frac{1}{a} + \frac{1}{a} + \frac{1}{a + \frac{1}{a + \frac{1}{a + \frac{1}{a} + \frac{1}{a + \frac{1}{a + \frac{1}{a + \frac{1}{a}}}}}}}}}}} \qquad (3.9)$$

then

$$x = a + \frac{1}{a + \frac{1}{a + \frac{1}{a + \frac{1}{a} + \frac{1}{a + \frac{1}{a} + \frac{1}{a} + \frac{1}{a + \frac{1}{a + \frac{1}{a} + \frac{1}{a} + \frac{1}{a + \frac{1}{a + \frac{1}{a + \frac{1}{a} + \frac{1}{a + \frac{1}{a} + \frac{1}{a} + \frac{1}{a} + \frac{1}{a + \frac{1}{a + \frac{1}{a}}}}}}}}}}}} \qquad (3.10)$$

it follows that

Figure 48. Fractions escaping the page

Working through these strange exercises I came to a selective familiarity with the back end of LaTeX, working rhythmically and repetitively with components of expressions in my efforts to break the boundaries of the neat formatting.

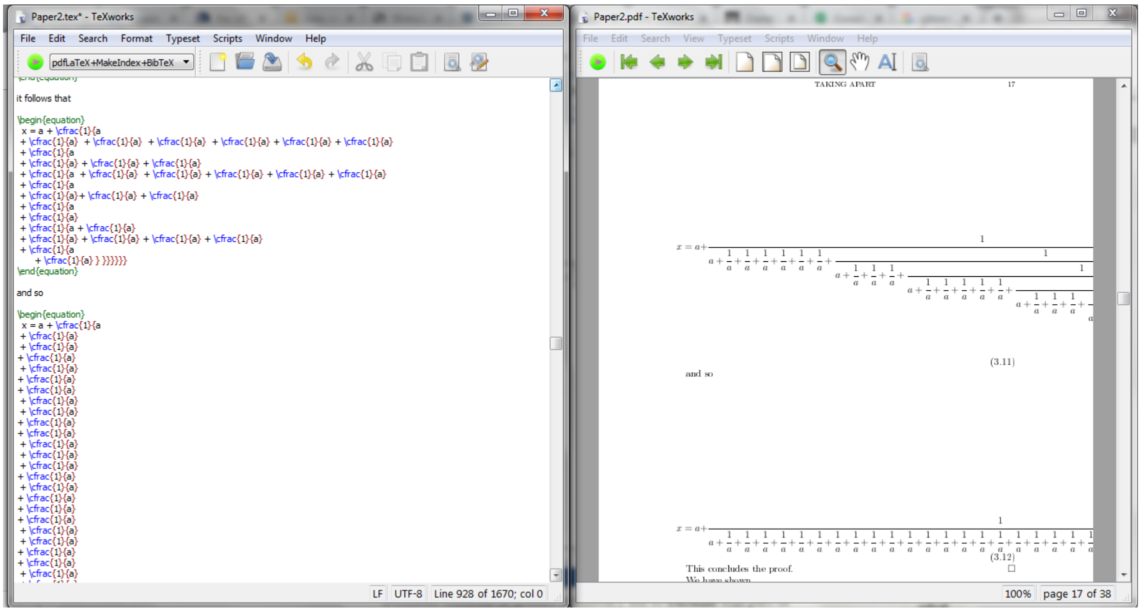


Figure 49. LaTeX and typeset result. Original in colour.

As a final experiment I pushed that rhythmic repetition into the textual output, having the authorial voice slip into weird, repetitive computer-speak.

5.2. **Proof of Theorem 1.6.** Let be , with dimension .
Let .
Let .
Let .
To prove Theorem 1.6, we may repeat , in the sense of §1.3.

We begin the more difficult case. According to Proposition 3.5 there exists
, such that for each
we have
and for each
we have
and for each
we have
and for each
we have
and for each
we have
and for each
we have
and for each
we have
and for each
we have .
Taking , for such we therefore have .

An application of Lemma 3.2 now yields
for any .

It follows that there exists such that for large enough .
For each
there exists
. For each
there exists.
For each
there exists
.
For each
there exists
.
For each
For each
For each
Let denote .

Figure 50. Rhythmic connecting phrases. Original in colour.

4.3. Evidence

4.3.1. Transcripts

Email 1

katiemccallum@live.co.uk [redacted]

From: G [redacted]

Sent: 20 April 2017 10:16

To: H [redacted]

Subject: Re: Sieves paper

I know I did not get round to writing the proof of Theorem 1.6. But as should be clear to you it will be very easy. I wanted to discuss a bit with you in Chicago how much of this stuff you wanted to include, as I'm not convinced how interesting it is.

Theorem 4.1: I think you are overlooking some subtleties here. I don't think there is a nice way to define the Omega_{p^m}. The reason is that the subset of a product of sets is not a product of subsets. So you cannot in general write Omega_M as a product of sets of the form Omega_{p^m}. I think this statement and the proof might need rewriting.

Also I find the condition in (4.1) that gcd(\x,M) = 1 a bit strange. Why is this condition present? For example, you don't allow Omega_M to consist of the zero vector. Why not? I guess this does not matter for the application, as we are dealing with primitive integer points?

[redacted] G

Figure 51. Email 1, from G to H. Original in colour.

Email 2

katiemccallum@live.co.uk

From: H [redacted]
Sent: 21 April 2017 05:56
To: G [redacted]
Subject: Re: Sieves paper

> I know I did not get round to writing the proof of Theorem 1.6. But as should be clear to you it will be very easy. I wanted to discuss a bit with you in Chicago how much of this stuff you wanted to include, as I'm not convinced how interesting it is.

me neither, but it doesn't take much space.

> Theorem 4.1: I think you are overlooking some subtleties here. I don't think there is a nice way to define the Omega_{p^m}. The reason is that the subset of a product of sets is not a product of subsets. So you cannot in general write Omega_M as a product of sets of the form Omega_{p^m}. I think this statement and the proof might need rewriting.

right; i always need to be a product. i guess it will be better to assume the existence of a subset for each prime in (4.1), and to let \Omega_M=prod_p \Omega_{p^m}.

> Also I find the condition in (4.1) that gcd(\x,M) = 1 a bit strange. Why is this condition present? For example, you don't allow Omega_M to consist of the zero vector. Why not? I guess this does not matter for the application, as we are dealing with primitive integer points?

the problem comes in using Hensel to get a nice simple form of the singular series for the primes p\mid M. At the top of page 24, one wants to prove N(k+1)=p^{n-1}N(k) for k\geq m. This fails if p\mid \grad F(x); ie. if p|x, or if p|Delta_F.
since we are ultimately working with projective points, I thought it was simplest to assume it in this section. one can derive a version without the primitivity condition (or the (M,2Delta)=1 condition), but the singular series will not be so explicit.

i could record a version with the singular series uncomputed (ie just a product \prod_p \sigma_p(\OMega)), valid without either of these 2 assumptions, and then a remark/lemma about what happens to it when p\nmid \x in \Omega and p\nmid 2 Delta_F?

Best, [redacted] H

1

Figure 52. Email 2, from H to G. Original in colour.

Email 3

p. 1

katiemccallum@live.co.uk

From:	H [redacted]
Sent:	28 April 2017 09:41
To:	[redacted] G
Subject:	Re: Sieves paper

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[redacted] H

On 23 April 2017 at 19:59, G [redacted] wrote:

Hi [redacted]

Thanks for the response. I've been going through what you have written and understanding it a bit better now.

On 21/04/17 13:59, H [redacted] wrote:

> ps. The nabla F(x)\gg 1 condition is used in analysing the singular integral (p24).

Yes I think I see this now. Could you please add a small comment saying something like this condition is something to do with choosing a "good" weight function which simplifies the analysis of the singular integral?

On 21 April 2017 at 13:56, H [redacted] wrote:

> I know I did not get round to writing the proof of Theorem 1.6. But as should be clear to you it will be very easy. I wanted to discuss a bit with you in Chicago how much of this stuff you wanted to include, as I'm not convinced how interesting it is.

me neither, but it doesn't take much space.

> Theorem 4.1: I think you are overlooking some subtleties here. I don't think there is a nice way to define the Omega_{p^m}. The reason is that the subset of a product of sets is

Email 3

p. 2

not a product of subsets. So you cannot in general write Omega_M as a product of sets of the form Omega_{p^m}. I think this statement and the proof might need rewriting.

right; i always need to be a product. i guess it will be better to assume the existence of a subset for each prime in (4.1), and to let \Omega_M=prod_p \Omega_{p^m}.

> Also I find the condition in (4.1) that gcd(\x,M) = 1 a bit strange. Why is this condition present? For example, you don't allow Omega_M to consist of the zero vector. Why not? I guess this does not matter for the application, as we are dealing with primitive integer points?

the problem comes in using Hensel to get a nice simple form of the singular series for the primes p\mid M. At the top of page 24, one wants to prove N(k+1)=p^{n-1}N(k) for k\geq m. This fails if p\mid \grad F(x); ie. if p|x, or if p|Delta_F.
since we are ultimately working with projective points, I thought it was simplest to assume it in this section. one can derive a version without the primitivity condition (or the (M,2Delta)=1 condition), but the singular series will not be so explicit.

i could record a version with the singular series uncomputed (ie just a product \prod_p \sigma_p(\OMega)), valid without either of these 2 assumptions, and then a remark/lemma about what happens to it when
p\nmid \x in \Omega and p\nmid 2 Delta_F?

Thanks for the explanation. I guess it's your call what you would like to do. I guess a small comment these conditions are to simplify the calculation of the singular series, would not go a miss.

Best,

[redacted] G

[redacted] H

Figure 53. Email 3, from G to H. Original in colour.

Edit 1

p. 1

We may therefore assume that the generic fibre of π is geometrically integral. Let $z > 0$. We clearly have

$$N(U,H,\pi,B) \leqslant \#\{x \in U(k) : H(x) \leqslant B,\ x \in \pi(Y(k_{\mathfrak{p}}))\forall\ \mathrm{N}\,\mathfrak{p} \leqslant z\}$$
$$\leqslant \#\{x \in U(k) : H(x) \leqslant B,\ x \bmod \mathfrak{p}^2 \in \pi(Y(k_{\mathfrak{p}})) \bmod \mathfrak{p}^2\,\forall\ \mathrm{N}\,\mathfrak{p} \leqslant z\}.$$

Thus using equidistribution, (3.2), and (3.6), we obtain

$$\lim_{B\to\infty} \frac{N(U,H,\pi,B)}{N(U,H,B)} \ll \prod_{\mathrm{N}\,\mathfrak{p}\leqslant z} \frac{\#(\pi(Y(\mathfrak{o}_{\mathfrak{p}})) \bmod \mathfrak{p}^2)}{\#X(\mathfrak{o}_k/\mathfrak{p}^2)} \ll \frac{1}{(\log z)^{\Delta(\pi)}},$$

where the implied constant is independent of z. Our assumption that there is a non-pseudo-split fibre over some codimension 1 point implies that $\Delta(\pi) > 0$. Taking $z \to \infty$ completes the proof of Theorem 1.4. □

4. Zeros of quadratic forms in fixed residue classes

Let $F \in \mathbb{Z}[x_1, \dots, x_n]$ be an isotropic quadratic form with non-zero discriminant $\Delta_F \in \mathbb{Z}$. For any positive integer M, suppose that we are given a non-empty subset

$$\Omega_M \subseteq \{\mathbf{x} \in (\mathbb{Z}/M\mathbb{Z})^n : (\mathbf{x}, M) = 1,\ F(\mathbf{x}) \equiv 0 \bmod M\}. \tag{4.1}$$

We shall write $[\mathbf{x}]_M \in \Omega_M$ to denote that the reduction of $\mathbf{x}$ modulo M lies in Ω_M. In this section we shall use the Hardy–Littlewood circle method to produce an asymptotic formula for the counting function

$$\widehat{N}(B, \Omega_M) = \sum_{\substack{\mathbf{x}\in\mathbb{Z}^n \\ F(\mathbf{x})=0 \\ [\mathbf{x}]_M\in\Omega_M}} w(\mathbf{x}/B), \tag{4.2}$$

where $w : \mathbb{R}^n \to \mathbb{R}_{\geqslant 0}$ is an infinitely differentiable function with compact support. In this section we shall allow all of our implied constants to depend on F, but we require complete uniformity in M.

Associated to F and w is the weighted real density $\sigma_\infty(w)$, as defined in [10, Thm. 3]. It satisfies $1 \ll_{F,w} \sigma_\infty(w) \ll_{F,w} 1$. Moreover, we have the associated p-adic density

$$\sigma_p = \lim_{k\to\infty} p^{-(n-1)k} \#\{\mathbf{x} \in (\mathbb{Z}/p^k\mathbb{Z})^n : F(\mathbf{x}) \equiv 0 \bmod p^k\}, \tag{4.3}$$

for each prime p. The goal of this section is to prove the following result.

Theorem 4.1. *Assume that $n \geqslant 5$ and that $\nabla F(\mathbf{x}) \gg 1$ for all $\mathbf{x} \in \mathrm{supp}(w)$. Assume that M is coprime to $2\Delta_F$. Then*

$$\widehat{N}(B,\Omega_M) = \sigma_\infty(w) B^{n-2} \prod_{p\nmid M} \sigma_p \prod_{p^m \| M} \frac{\#\Omega_{p^m}}{p^{m(n-1)}} + O_{\varepsilon,F,w}\left(B^{n/2+\varepsilon} M^{n/2+\varepsilon}\right),$$

for any $\varepsilon > 0$.

In this result and henceforth in this section, the implied constant is allowed to depend on the choice of ε, the form F and the weight function w, but not on the modulus M. To ease notation we shall suppress this dependence in what follows.

The Euler product $\prod_p \sigma_p$ is absolutely convergent for $n \geqslant 5$. Theorem 4.1 can be improved in several directions. Firstly, an inspection of the proof reveals that

Edit 1

p. 2

H and G

18 [redacted]

one does rather better in the B-aspect of the error term when n is odd. Secondly, it would not be hard to deal with quadratic forms in $n = 4$ variables. Finally, when M is square-free it is possible to improve the error term to $O(B^{n/2+\varepsilon}\#\Omega_M^{1/2})$. In order to simplify our exposition we have decided not to pursue any of these improvements in the present investigation. In our application Ω_M will be comparable in size to the set of $\mathbf{x} \in (\mathbb{Z}/M\mathbb{Z})^n$ for which $F(\mathbf{x}) \equiv 0 \bmod M$, and so we have relaxed the dependence on $\#\Omega_M$. In fact, although wasteful, we shall often employ the trivial inequality $\#\Omega_M \leqslant M^n$.

We begin the proof of Theorem 4.1 by invoking the version of the circle method developed by Heath-Brown [10, Thm. 1]. This implies that

$$\widehat{N}(B,\Omega_M) = \frac{c_Q}{Q^2}\sum_{q=1}^{\infty}\sideset{}{^*}\sum_{a \bmod q}\sum_{\substack{\mathbf{x}\in\mathbb{Z}^n\\ [\mathbf{x}]_M\in\Omega_M}} w(\mathbf{x}/B)e_q(aF(\mathbf{x}))h\left(\frac{q}{Q},\frac{F(\mathbf{x})}{Q^2}\right),$$

for any $Q > 1$. Here c_Q is a positive constant satisfying $c_Q = 1 + O_A(Q^{-A})$ for any $A > 0$ and, moreover, $h(x,y)$ is a smooth function defined on the set $(0,\infty)\times\mathbb{R}$ such that $h(x,y) \ll x^{-1}$ for all y, with $h(x,y)$ non-zero only for $x \leqslant \max\{1, 2|y|\}$. In particular, we are only interested in $q \ll Q$ in this sum.

We will henceforth take $Q = B$. It is natural to break the sum into residue classes modulo the least common multiple $[q,M]$ and then apply Poisson summation, as in the proof of [10, Thm. 2]. This leads to the expression

$$\widehat{N}(B,\Omega_M) = \frac{c_B}{B^2}\sum_{q\ll B}\sum_{\mathbf{c}\in\mathbb{Z}^n}[q,M]^{-n}S_{q,M}(\mathbf{c})J_{q,M}(\mathbf{c}),$$

where

$$S_{q,M}(\mathbf{c}) = \sideset{}{^*}\sum_{a \bmod q}\sum_{\substack{\mathbf{y} \bmod [q,M]\\ [\mathbf{y}]_M\in\Omega_M}} e_q(aF(\mathbf{y}))\,e_{[q,M]}(\mathbf{c}.\mathbf{y}) \tag{4.4}$$

and

$$\begin{aligned} J_{q,M}(\mathbf{c}) &= \int_{\mathbb{R}^n} w(\mathbf{x}/B)h\left(\frac{q}{B},\frac{F(\mathbf{x})}{B^2}\right)e_{[q,M]}(-\mathbf{c}.\mathbf{x})\mathrm{d}\mathbf{x}\\ &= B^n\int_{\mathbb{R}^n} w(\mathbf{x})h\left(\frac{q}{B},F(\mathbf{x})\right)e_{[q,m]}(-B\mathbf{c}.\mathbf{x})\mathrm{d}\mathbf{x}. \end{aligned}$$

It will be convenient to set

$$I_r^*(\mathbf{v}) = \int_{\mathbb{R}^n} w(\mathbf{x})h\left(r,F(\mathbf{x})\right)e_r(-\mathbf{v}.\mathbf{x})\mathrm{d}\mathbf{x}, \tag{4.5}$$

for any $r > 0$ and $\mathbf{v}\in\mathbb{R}^n$. In this notation, which coincides with that introduced in [10, §7], we may clearly write $J_{q,M}(\mathbf{c}) = B^nI_r^*(M'^{-1}\mathbf{c})$, where $r = q/B$ and $M' = [q,M]/q = M/(M,q)$. Thus

$$\widehat{N}(B,\Omega_M) = c_BB^{n-2}\sum_{q\ll B}\sum_{\mathbf{c}\in\mathbb{Z}^n}[q,M]^{-n}S_{q,M}(\mathbf{c})I_r^*(M'^{-1}\mathbf{c}). \tag{4.6}$$

Figure 54. Version 1 of edits to the paper. Original in colour.

Edit 2

p. 1

Let $z > 0$. Imposing the above local conditions for all $\mathfrak{p}$ with $\mathrm{N}\,\mathfrak{p} \leqslant z$, we may use equidistribution, Example 3.4, and (3.5), to obtain

$$\lim_{B\to\infty} \frac{N(U,H,\pi,B)}{N(U,H,B)} \ll \prod_{\mathrm{N}\,\mathfrak{p}\leqslant z} \frac{\#(\pi(\mathcal{Y}(\mathfrak{o}_\mathfrak{p})) \bmod \mathfrak{p}^2)}{\#\mathcal{X}(\mathfrak{o}_k/\mathfrak{p}^2)} \ll \frac{1}{(\log z)^{\Delta(\pi)}},$$

where the implied constant is independent of z. Our assumption that there is a non-pseudo-split fibre over some codimension 1 point implies that $\Delta(\pi) > 0$. Taking $z \to \infty$ completes the proof of Theorem 1.4. □

4. Zeros of quadratic forms in fixed residue classes

Let $F \in \mathbb{Z}[x_1, \dots, x_n]$ be an isotropic quadratic form with non-zero discriminant $\Delta_F \in \mathbb{Z}$. For any positive integer M, suppose that we are given a non-empty subset

G
$$\Omega_M \subseteq \{\mathbf{x} \in (\mathbb{Z}/M\mathbb{Z})^n : (\mathbf{x}, M) = 1,\ F(\mathbf{x}) \equiv 0 \bmod M\}. \tag{4.1}$$

♣♣♣ [redacted] [What does $(\mathbf{x}, M) = 1$ mean?]

For $\mathbf{x} \in \mathbb{Z}^n$, we write $[\mathbf{x}]_d$ for its reduction modulo M. ♣♣♣ [redacted] G [Rewrote this slightly] In this section we shall use the Hardy–Littlewood circle method to produce an asymptotic formula for the counting function

$$\widehat{N}(B, \Omega_M) = \sum_{\substack{\mathbf{x}\in\mathbb{Z}^n \\ F(\mathbf{x})=0 \\ [\mathbf{x}]_M \in \Omega_M}} w(\mathbf{x}/B), \tag{4.2}$$

where $w : \mathbb{R}^n \to \mathbb{R}_{\geqslant 0}$ is an infinitely differentiable function with compact support. In this section we shall allow all of our implied constants to depend on F, but we require complete uniformity in M.

Associated to F and w is the weighted real density $\sigma_\infty(w)$, as defined in [10, Thm. 3]. It satisfies $1 \ll_{F,w} \sigma_\infty(w) \ll_{F,w} 1$. Moreover, we have the associated p-adic density

$$\sigma_p = \lim_{k\to\infty} p^{-(n-1)k} \#\{\mathbf{x} \in (\mathbb{Z}/p^k\mathbb{Z})^n : F(\mathbf{x}) \equiv 0 \bmod p^k\}, \tag{4.3}$$

for each prime p. The goal of this section is to prove the following result.

Theorem 4.1. *Assume that $n \geqslant 5$ and that $\nabla F(\mathbf{x}) \gg 1$ for all $\mathbf{x} \in \mathrm{supp}(w)$* ♣♣♣ [redacted] G [Where does this last condition come from? Is it just saying that the quadratic form is non-singular? Also is this for $\mathbf{x} \in \mathbb{R}^n$ or $\mathbf{x} \in \mathbb{Z}^n$? And you probably want to ignore the zero vector for $\mathbf{x}$] *Assume that M is coprime to $2\Delta_F$.* ♣♣♣ G [redacted] [How important is it that $\gcd(M, 2\Delta_F) = 1$? I mean, these are finitely many primes which should not cause too many problems] *Let Ω_M be as in (4.1). Then*

$$\widehat{N}(B, \Omega_M) = \sigma_\infty(w) B^{n-2} \prod_{p\nmid M} \sigma_p \prod_{p^m \| M} \frac{\#\Omega_{p^m}}{p^{m(n-1)}} + O_{\varepsilon,F,w}\left(B^{n/2+\varepsilon} M^{n/2+\varepsilon}\right),$$

for any $\varepsilon > 0$. ♣♣♣ [redacted] [Ω_{p^m} has not been defined in this section, only Ω_M].

In this result and henceforth in this section, the implied constant is allowed to depend on the choice of ε, the form F and the weight function w, but not on the modulus M. To ease notation we shall suppress this dependence in what follows. ♣♣♣ [redacted] [You are just repeating here what you already said earlier.]

Edit 2

p.2

H and G

18 [redacted]

The Euler product $\prod_p \sigma_p$ is absolutely convergent for $n \geqslant 5$. Theorem 4.1 can be improved in several directions. Firstly, an inspection of the proof reveals that one does rather better in the B-aspect of the error term when n is odd. Secondly, it would not be hard to deal with quadratic forms in $n = 4$ variables ♣♣♣ [redacted]. [What about $n = 3$?]. Finally, when M is square-free it is possible to improve the error term to $O(B^{n/2+\varepsilon}\#\Omega_M^{1/2})$. In order to simplify our exposition we have decided not to pursue any of these improvements in the present investigation. In our application Ω_M will be comparable in size to the set of $\mathbf{x} \in (\mathbb{Z}/M\mathbb{Z})^n$ for which $F(\mathbf{x}) \equiv 0 \bmod M$, and so we have relaxed the dependence on $\#\Omega_M$. In fact, although wasteful, we shall often employ the trivial inequality $\#\Omega_M \leqslant M^n$.

We begin the proof of Theorem 4.1 by invoking the version of the circle method developed by Heath-Brown [10, Thm. 1]. This implies that

$$\widehat{N}(B,\Omega_M) = \frac{c_Q}{Q^2}\sum_{q=1}^{\infty}\sideset{}{^*}\sum_{a \bmod q}\sum_{\substack{\mathbf{x}\in\mathbb{Z}^n\\ [\mathbf{x}]_M\in\Omega_M}} w(\mathbf{x}/B)e_q(aF(\mathbf{x}))h\left(\frac{q}{Q},\frac{F(\mathbf{x})}{Q^2}\right),$$

for any $Q > 1$. Here c_Q is a positive constant satisfying $c_Q = 1 + O_A(Q^{-A})$ for any $A > 0$ and, moreover, $h(x,y)$ is a smooth function defined on the set $(0,\infty)\times\mathbb{R}$ such that $h(x,y) \ll x^{-1}$ for all y, with $h(x,y)$ non-zero only for $x \leqslant \max\{1, 2|y|\}$. In particular, we are only interested in $q \ll Q$ in this sum.

We will henceforth take $Q = B$. It is natural to break the sum into residue classes modulo the least common multiple $[q,M]$ and then apply Poisson summation, as in the proof of [10, Thm. 2]. This leads to the expression

$$\widehat{N}(B,\Omega_M) = \frac{c_B}{B^2}\sum_{q\ll B}\sum_{\mathbf{c}\in\mathbb{Z}^n}[q,M]^{-n}S_{q,M}(\mathbf{c})J_{q,M}(\mathbf{c}),$$

where

$$S_{q,M}(\mathbf{c}) = \sideset{}{^*}\sum_{a \bmod q}\sum_{\substack{\mathbf{y} \bmod [q,M]\\ [\mathbf{y}]_M\in\Omega_M}} e_q\left(aF(\mathbf{y})\right)e_{[q,M]}\left(\mathbf{c}.\mathbf{y}\right) \qquad (4.4)$$

and

$$\begin{aligned} J_{q,M}(\mathbf{c}) &= \int_{\mathbb{R}^n} w(\mathbf{x}/B)h\left(\frac{q}{B},\frac{F(\mathbf{x})}{B^2}\right)e_{[q,M]}(-\mathbf{c}.\mathbf{x})\mathrm{d}\mathbf{x}\\ &= B^n\int_{\mathbb{R}^n} w(\mathbf{x})h\left(\frac{q}{B},F(\mathbf{x})\right)e_{[q,m]}(-B\mathbf{c}.\mathbf{x})\mathrm{d}\mathbf{x}. \end{aligned}$$

It will be convenient to set

$$I_r^*(\mathbf{v}) = \int_{\mathbb{R}^n} w(\mathbf{x})h\left(r,F(\mathbf{x})\right)e_r(-\mathbf{v}.\mathbf{x})\mathrm{d}\mathbf{x}, \qquad (4.5)$$

for any $r > 0$ and $\mathbf{v}\in\mathbb{R}^n$. In this notation, which coincides with that introduced in [10, §7], we may clearly write $J_{q,M}(\mathbf{c}) = B^n I_r^*(M'^{-1}\mathbf{c})$, where $r = q/B$ and $M' = [q,M]/q = M/(M,q)$. Thus

$$\widehat{N}(B,\Omega_M) = c_B B^{n-2}\sum_{q\ll B}\sum_{\mathbf{c}\in\mathbb{Z}^n}[q,M]^{-n}S_{q,M}(\mathbf{c})I_r^*(M'^{-1}\mathbf{c}). \qquad (4.6)$$

Figure 55. Version 2 of edits to the paper. Original in colour.

Edit 3

p. 1

Let $z > 0$. Imposing the above local conditions for all $\mathfrak{p}$ with $\mathrm{N}\,\mathfrak{p} \leqslant z$, we may use equidistribution, Example 3.4, and (3.5), to obtain

$$\lim_{B\to\infty} \frac{N(U,H,\pi,B)}{N(U,H,B)} \ll \prod_{\mathrm{N}\,\mathfrak{p}\leqslant z} \frac{\#(\pi(\mathcal{Y}(\mathfrak{o}_\mathfrak{p})) \bmod \mathfrak{p}^2)}{\#\mathcal{X}(\mathfrak{o}_k/\mathfrak{p}^2)} \ll \frac{1}{(\log z)^{\Delta(\pi)}},$$

where the implied constant is independent of z. Our assumption that there is a non-pseudo-split fibre over some codimension 1 point implies that $\Delta(\pi) > 0$. Taking $z \to \infty$ completes the proof of Theorem 1.4. □

4. Zeros of quadratic forms in fixed residue classes

Let $F \in \mathbb{Z}[x_1, \dots, x_n]$ be an isotropic quadratic form with non-zero discriminant $\Delta_F \in \mathbb{Z}$. For any positive integer M, suppose that we are given a non-empty subset

$$\Omega_M \subseteq \{\mathbf{x} \in (\mathbb{Z}/M\mathbb{Z})^n : \gcd(\mathbf{x}, M) = 1,\ F(\mathbf{x}) \equiv 0 \bmod M\}. \tag{4.1}$$

For $\mathbf{x} \in \mathbb{Z}^n$, we write $[\mathbf{x}]_d$ for its reduction modulo M. In this section we shall use the Hardy–Littlewood circle method to produce an asymptotic formula for the counting function

$$\widehat{N}(B, \Omega_M) = \sum_{\substack{\mathbf{x}\in\mathbb{Z}^n \\ F(\mathbf{x})=0 \\ [\mathbf{x}]_M \in \Omega_M}} w(\mathbf{x}/B), \tag{4.2}$$

where $w : \mathbb{R}^n \to \mathbb{R}_{\geqslant 0}$ is an infinitely differentiable function with compact support.

Associated to F and w is the weighted real density $\sigma_\infty(w)$, as defined in [10, Thm. 3]. It satisfies $1 \ll_{F,w} \sigma_\infty(w) \ll_{F,w} 1$. Moreover, we have the associated p-adic density

$$\sigma_p = \lim_{k\to\infty} p^{-(n-1)k} \#\{\mathbf{x} \in (\mathbb{Z}/p^k\mathbb{Z})^n : F(\mathbf{x}) \equiv 0 \bmod p^k\}, \tag{4.3}$$

for each prime p. The goal of this section is to prove the following result.

Theorem 4.1. *Assume that $n \geqslant 5$ and that $\nabla F(\mathbf{x}) \gg 1$ for all $\mathbf{x} \in \mathrm{supp}(w)$* ♣♣♣ [redacted] G [Where does this last condition come from? Is it just saying that the quadratic form is non-singular? Also is this for $\mathbf{x} \in \mathbb{R}^n$ or $\mathbf{x} \in \mathbb{Z}^n$? And you probably want to ignore the zero vector for $\mathbf{x}$] *Assume that M is coprime to $2\Delta_F$* ♣♣♣ [redacted] G [How important is it that $\gcd(M, 2\Delta_F) = 1$? I mean, these are finitely many primes which should not cause too many problems] *and let Ω_M be as in* (4.1)*. Then*

$$\widehat{N}(B, \Omega_M) = \sigma_\infty(w) B^{n-2} \prod_{p \nmid M} \sigma_p \prod_{p^m \| M} \frac{\#\Omega_{p^m}}{p^{m(n-1)}} + O_{\varepsilon,F,w}\left(B^{n/2+\varepsilon} M^{n/2+\varepsilon}\right),$$

for any $\varepsilon > 0$. ♣♣♣ [redacted] G [Ω_{p^m} has not been defined in this section, only Ω_M]. ♠♠♠ H [redacted] [It is too obvious to repeat here. I'll answer the remaining concerns in Chicago..]

In this result and henceforth in this section, the implied constant is allowed to depend on the choice of ε, the form F and the weight function w, but not on the modulus M. To ease notation we shall suppress this dependence in what follows. ♣♣♣ [redacted] G [You are just repeating here what you already said earlier.] ♠♠♠ [redacted] H [Deleted earlier appearance.]

H and G

18 [redacted]

The Euler product $\prod_p \sigma_p$ is absolutely convergent for $n \geqslant 5$. Theorem 4.1 can be improved in several directions. Firstly, an inspection of the proof reveals that one does rather better in the B-aspect of the error term when n is odd. Secondly, it would not be hard to deal with quadratic forms in $n = 3$ or 4 variables. Finally, when M is square-free it is possible to improve the error term to $O(B^{n/2+\varepsilon}\#\Omega_M^{1/2})$. In order to simplify our exposition we have decided not to pursue any of these improvements in the present investigation. In our application Ω_M will be comparable in size to the set of $\mathbf{x} \in (\mathbb{Z}/M\mathbb{Z})^n$ for which $F(\mathbf{x}) \equiv 0 \bmod M$, and so we have relaxed the dependence on $\#\Omega_M$. In fact, although wasteful, we shall often employ the trivial inequality $\#\Omega_M \leqslant M^n$.

We begin the proof of Theorem 4.1 by invoking the version of the circle method developed by Heath-Brown [10, Thm. 1]. This implies that

$$\widehat{N}(B,\Omega_M) = \frac{c_Q}{Q^2}\sum_{q=1}^{\infty}\ \sideset{}{^*}\sum_{a \bmod q}\ \sum_{\substack{\mathbf{x}\in\mathbb{Z}^n\\ [\mathbf{x}]_M\in\Omega_M}} w(\mathbf{x}/B)e_q(aF(\mathbf{x}))h\left(\frac{q}{Q},\frac{F(\mathbf{x})}{Q^2}\right),$$

for any $Q > 1$. Here c_Q is a positive constant satisfying $c_Q = 1 + O_A(Q^{-A})$ for any $A > 0$ and, moreover, $h(x,y)$ is a smooth function defined on the set $(0,\infty)\times\mathbb{R}$ such that $h(x,y) \ll x^{-1}$ for all y, with $h(x,y)$ non-zero only for $x \leqslant \max\{1, 2|y|\}$. In particular, we are only interested in $q \ll Q$ in this sum.

We will henceforth take $Q = B$. It is natural to break the sum into residue classes modulo the least common multiple $[q,M]$ and then apply Poisson summation, as in the proof of [10, Thm. 2]. This leads to the expression

$$\widehat{N}(B,\Omega_M) = \frac{c_B}{B^2}\sum_{q\ll B}\sum_{\mathbf{c}\in\mathbb{Z}^n}[q,M]^{-n}S_{q,M}(\mathbf{c})J_{q,M}(\mathbf{c}),$$

where

$$S_{q,M}(\mathbf{c}) = \sideset{}{^*}\sum_{a \bmod q}\ \sum_{\substack{\mathbf{y} \bmod [q,M]\\ [\mathbf{y}]_M\in\Omega_M}} e_q\left(aF(\mathbf{y})\right)e_{[q,M]}\left(\mathbf{c}.\mathbf{y}\right) \tag{4.4}$$

and

$$\begin{aligned} J_{q,M}(\mathbf{c}) &= \int_{\mathbb{R}^n} w(\mathbf{x}/B)h\left(\frac{q}{B},\frac{F(\mathbf{x})}{B^2}\right)e_{[q,M]}(-\mathbf{c}.\mathbf{x})\mathrm{d}\mathbf{x}\\ &= B^n\int_{\mathbb{R}^n} w(\mathbf{x})h\left(\frac{q}{B},F(\mathbf{x})\right)e_{[q,m]}(-B\mathbf{c}.\mathbf{x})\mathrm{d}\mathbf{x}. \end{aligned}$$

It will be convenient to set

$$I_r^*(\mathbf{v}) = \int_{\mathbb{R}^n} w(\mathbf{x})h\left(r,F(\mathbf{x})\right)e_r(-\mathbf{v}.\mathbf{x})\mathrm{d}\mathbf{x}, \tag{4.5}$$

for any $r > 0$ and $\mathbf{v}\in\mathbb{R}^n$. In this notation, which coincides with that introduced in [10, §7], we may clearly write $J_{q,M}(\mathbf{c}) = B^nI_r^*(M'^{-1}\mathbf{c})$, where $r = q/B$ and $M' = [q,M]/q = M/(M,q)$. Thus

$$\widehat{N}(B,\Omega_M) = c_BB^{n-2}\sum_{q\ll B}\sum_{\mathbf{c}\in\mathbb{Z}^n}[q,M]^{-n}S_{q,M}(\mathbf{c})I_r^*(M'^{-1}\mathbf{c}). \tag{4.6}$$

Figure 56. Version 3 of edits to the paper. Original in colour.

Edit 4

p. 1

4. Zeros of quadratic forms in fixed residue classes

Let $F \in \mathbb{Z}[x_1, \dots, x_n]$ be an isotropic quadratic form with non-zero discriminant $\Delta_F \in \mathbb{Z}$. For any positive integer M and each prime power factor $p^m \| M$ suppose that we are given a non-empty subset

$$\Omega_{p^m} \subseteq \{\mathbf{x} \in (\mathbb{Z}/p^m\mathbb{Z})^n : p \nmid \mathbf{x},\ F(\mathbf{x}) \equiv 0 \bmod p^m\}. \tag{4.1}$$

Put $\Omega_M = \prod_{p^m \| M} \Omega_{p^m}$. For $\mathbf{x} \in \mathbb{Z}^n$, we write $[\mathbf{x}]_M$ for its reduction modulo M. In this section we shall use the Hardy–Littlewood circle method to produce an asymptotic formula for the counting function

$$\widehat{N}(B, \Omega_M) = \sum_{\substack{\mathbf{x} \in \mathbb{Z}^n \\ F(\mathbf{x})=0 \\ [\mathbf{x}]_M \in \Omega_M}} w(\mathbf{x}/B), \tag{4.2}$$

where $w : \mathbb{R}^n \to \mathbb{R}_{\geqslant 0}$ is an infinitely differentiable function with compact support.

Associated to F and w is the weighted real density $\sigma_\infty(w)$, as defined in [10, Thm. 3]. It satisfies $1 \ll_{F,w} \sigma_\infty(w) \ll_{F,w} 1$. Moreover, we have the associated p-adic density

$$\sigma_p = \lim_{k \to \infty} p^{-(n-1)k} \#\{\mathbf{x} \in (\mathbb{Z}/p^k\mathbb{Z})^n : F(\mathbf{x}) \equiv 0 \bmod p^k\}, \tag{4.3}$$

for each prime p. The goal of this section is to prove the following result.

Theorem 4.1. *Assume that $n \geqslant 5$ and that $\nabla F(\mathbf{x}) \gg 1$ for all $\mathbf{x} \in \operatorname{supp}(w)$ Assume that M is coprime to $2\Delta_F$ and let Ω_M be as in (4.1). Then*

$$\widehat{N}(B, \Omega_M) = \sigma_\infty(w) B^{n-2} \prod_{p \nmid M} \sigma_p \prod_{p^m \| M} \frac{\#\Omega_{p^m}}{p^{m(n-1)}} + O_{\varepsilon,F,w}\left(B^{n/2+\varepsilon} M^{n/2+\varepsilon}\right),$$

for any $\varepsilon > 0$.

In this result and henceforth in this section, the implied constant is allowed to depend on the choice of ε, the form F and the weight function w, but not on the modulus M. To ease notation we shall suppress this dependence in what follows.

Some comments are in order about the statement of this result. The condition that $\nabla F(\mathbf{x}) \gg 1$ for any $\mathbf{x}$ in the support of w is required to simplify the analysis of the oscillatory integrals that appear in the argument. The assumptions $(M, 2\Delta_F) = 1$ and $(\mathbf{x}, M) = 1$ for any $\mathbf{x} \in \Omega_M$ are made purely to simplify the expression for the leading constant in the asymptotic formula for $\widehat{N}(B, \Omega_M)$.

This second part is what's mentioned in the email!

Theorem 4.1 can be improved in several directions. Firstly, an inspection of the proof reveals that one does rather better in the B-aspect of the error term when n is odd. Secondly, it would not be hard to deal with quadratic forms in $n = 3$ or 4 variables. Finally, when M is square-free it is possible to improve the error term to $O(B^{n/2+\varepsilon} \#\Omega_M^{1/2})$. In order to simplify our exposition we have decided not to pursue any of these improvements in the present investigation. In our application Ω_M will be comparable in size to the set of $\mathbf{x} \in (\mathbb{Z}/M\mathbb{Z})^n$ for which $F(\mathbf{x}) \equiv 0 \bmod M$, and so we have relaxed the dependence on $\#\Omega_M$. In fact, although wasteful, we shall often employ the trivial inequality $\#\Omega_M \leqslant M^n$.

Edit 4

p. 2

H and G

18 [redacted]

We begin the proof of Theorem 4.1 by invoking the version of the circle method developed by Heath-Brown [10, Thm. 1]. This implies that

$$\widehat{N}(B,\Omega_M) = \frac{c_Q}{Q^2}\sum_{q=1}^{\infty}\sideset{}{^*}\sum_{a \bmod q}\sum_{\substack{\mathbf{x}\in\mathbb{Z}^n\\ [\mathbf{x}]_M\in\Omega_M}} w(\mathbf{x}/B)e_q(aF(\mathbf{x}))h\left(\frac{q}{Q},\frac{F(\mathbf{x})}{Q^2}\right),$$

for any $Q > 1$. Here c_Q is a positive constant satisfying $c_Q = 1 + O_A(Q^{-A})$ for any $A > 0$ and, moreover, $h(x,y)$ is a smooth function defined on the set $(0,\infty)\times\mathbb{R}$ such that $h(x,y) \ll x^{-1}$ for all y, with $h(x,y)$ non-zero only for $x \leqslant \max\{1, 2|y|\}$. In particular, we are only interested in $q \ll Q$ in this sum.

We will henceforth take $Q = B$. It is natural to break the sum into residue classes modulo the least common multiple $[q,M]$ and then apply Poisson summation, as in the proof of [10, Thm. 2]. This leads to the expression

$$\widehat{N}(B,\Omega_M) = \frac{c_B}{B^2}\sum_{q\ll B}\sum_{\mathbf{c}\in\mathbb{Z}^n}[q,M]^{-n}S_{q,M}(\mathbf{c})J_{q,M}(\mathbf{c}),$$

where

$$S_{q,M}(\mathbf{c}) = \sideset{}{^*}\sum_{a \bmod q}\sum_{\substack{\mathbf{y} \bmod [q,M]\\ [\mathbf{y}]_M\in\Omega_M}} e_q(aF(\mathbf{y}))\,e_{[q,M]}(\mathbf{c}.\mathbf{y}) \tag{4.4}$$

and

$$J_{q,M}(\mathbf{c}) = \int_{\mathbb{R}^n} w(\mathbf{x}/B)h\left(\frac{q}{B},\frac{F(\mathbf{x})}{B^2}\right)e_{[q,M]}(-\mathbf{c}.\mathbf{x})\mathrm{d}\mathbf{x}$$
$$= B^n\int_{\mathbb{R}^n} w(\mathbf{x})h\left(\frac{q}{B},F(\mathbf{x})\right)e_{[q,m]}(-B\mathbf{c}.\mathbf{x})\mathrm{d}\mathbf{x}.$$

It will be convenient to set

$$I_r^*(\mathbf{v}) = \int_{\mathbb{R}^n} w(\mathbf{x})h(r,F(\mathbf{x}))\,e_r(-\mathbf{v}.\mathbf{x})\mathrm{d}\mathbf{x}, \tag{4.5}$$

for any $r > 0$ and $\mathbf{v}\in\mathbb{R}^n$. In this notation, which coincides with that introduced in [10, §7], we may clearly write $J_{q,M}(\mathbf{c}) = B^nI_r^*(M'^{-1}\mathbf{c})$, where $r = q/B$ and $M' = [q,M]/q = M/(M,q)$. Thus

$$\widehat{N}(B,\Omega_M) = c_BB^{n-2}\sum_{q\ll B}\sum_{\mathbf{c}\in\mathbb{Z}^n}[q,M]^{-n}S_{q,M}(\mathbf{c})I_r^*(M'^{-1}\mathbf{c}). \tag{4.6}$$

4.1. **The exponential sum.** The purpose of this section is to analyse the sum $S_{q,M}(\mathbf{c})$ in (4.4) for $q, M\in\mathbb{N}$ such that M is coprime to $2\Delta_F$. We begin by establishing the following factorisation property.

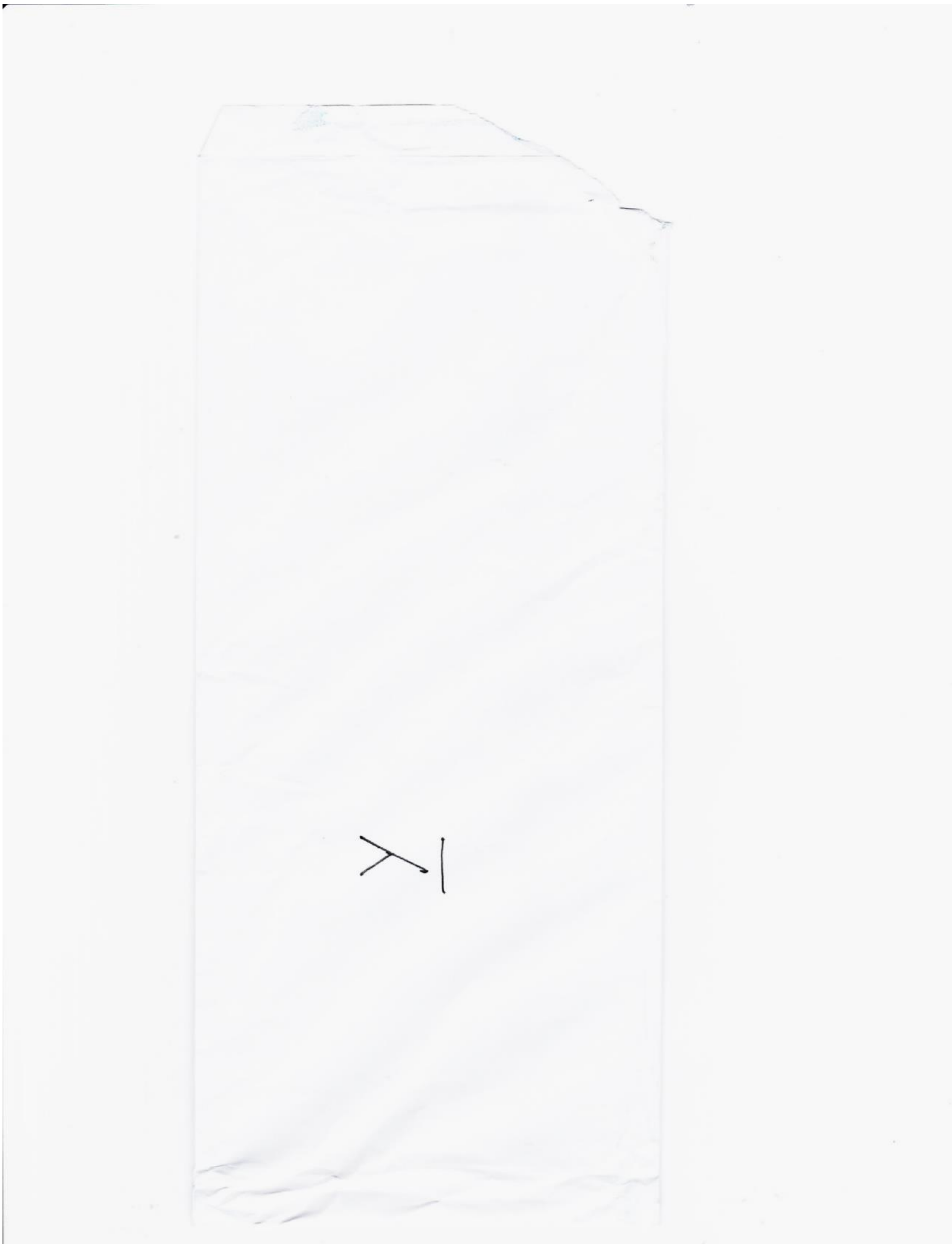
Lemma 4.2. *Let $M = M_1M_2$. Suppose that $(q_1M_1, q_2M_2) = 1$ and choose integers s, t such that $[q_1, M_1]s + [q_2, M_2]t = 1$. Then*

$$S_{q_1q_2,M}(\mathbf{c}) = S_{q_1,M_1}(t\mathbf{c})S_{q_2,M_2}(s\mathbf{c}).$$

Proof. Note that $[q_1q_2, M] = [q_1, M_1][q_2, M_2]$. As $\mathbf{y}_1$ runs modulo $[q_1, M_1]$ and $\mathbf{y}_2$ runs modulo $[q_2, M_2]$, so $\mathbf{y} = \mathbf{y}_1[q_2, M_2]t + \mathbf{y}_2[q_1, M_1]s$ runs over a full set of residue classes modulo $[q_1q_2, M]$. Now let $\bar{q}_1, \bar{q}_2\in\mathbb{Z}$ be such that $q_1\bar{q}_1 + q_2\bar{q}_2 = 1$. Then $a = a_1q_2\bar{q}_2 + a_2q_1\bar{q}_1$ runs over $(\mathbb{Z}/q_1q_2\mathbb{Z})^*$ as a_1 (resp. a_2) runs over $(\mathbb{Z}/q_1\mathbb{Z})^*$

Figure 57. Version 4 of edits to the paper. Original in colour.

4.3.2. Correspondence



Original in colour

Dear Y,

Hi there! I hope all is well, and you're prepared for a bit of mathematical chat...

I am enclosing the docs we're going to be talking about. I hope it's reasonably clear. We ~~have~~ are going to be looking at the email correspondence between two collaborators, G and H, as they work on edits to a draft paper – so we have three emails in sequence, and the four stages of edits made to the paper as those emails were going back and forth. We're just looking at a short section, which is marked on the printouts. The two of them discuss various edits they're making to the draft paper to improve it in some way, and in the emails you can see they're also responding to notes they've left one another in the text (there's a way to leave marked commentary in LaTeX, which is what they're using). Throughout, I've used a color for each of them to make it clear who's who – G is blue, H is red. What's nice here is that as well as looking at how G and H come to understand each other in the email exchange, the fact that they're working on a paper that's so close to publication means that we're also getting to see how they think about communicating their work with a broader audience.

You can see, the two of them are replying to one another in-text in the email, so the text is quite conversation-like. That said it isn't quite like the conversation we looked at before, where the speakers were using gestures and half-sentences to convey their meaning. Instead they do sit and think and even though it is still quite informal!

Original in colour

write out what they're going to say, much like I am now, which involves more careful thought about explanation, more full sentences, and a different pace to the exchange. The paper draft was put in Dropbox and the collaborators took turns to work on it, figuratively passing it back and forth by letting each other know by email when they would be finished with it. This email exchange I think all happens between edits 3 and 4; ~~at the point where~~ before that G and H had been conversing in those little notes in the document, then switched to email (because they needed to discuss in more depth, perhaps?), and then some final edits were made. And actually I think it then stayed that way until submission.

What's being discussed is edits to the statement of Theorem 4.1 and an expression marked '(4.1)' (in brackets to the right) (I actually think the numbering being the same is just coincidence). G had asked some questions in the text of the paper, but was not satisfied by H's response and so switched to email to follow up – I guess more directly, and with more space to explain.

So: first off, what's in the email? The first thing discussed is whether or not to include something in the paper. It's really interesting that 'it doesn't take much space' is clearly written as a response to 'I'm not convinced how interesting it is' – for this to make any sense as a response, we might guess that it's obvious to G and H that interestingness and space should be weighed against one another when deciding whether to include something in the text. Brevity is good, but so is saying as much as you can so you and others can use it later. THE CONTENT NEEDS to BE DEEMED INTERESTING ENOUGH TO JUSTIFY THE SPACE DEDICATED TO IT.

The second bit. This refers to Theorem 4.1. If you take a look at the edit history you can see G add a comment saying '$\Omega_{p\cap m}$ has not been defined in this section, only Ω_m', and then H replies in the document with something kind of brief. I think that this is what G is then referring to in the email, since I can recognise G making reference to 'Omega_{p∩M}' and 'Omega_M'. So I think that G can tell from H's response in the document that G's concern has not really been understood. G's comment pointed to a serious problem, but G's in-text comment was quite brief, and H took it to be something relatively minor – it seems to me that part of the original misunderstanding was that the two of them had misaligned impressions of the <u>size</u> of the problem, so to speak, and how much thought it needed. So then, G switches to email, beginning with 'I think you are overlooking some subtleties here' – this is nice since we can spot that this is G letting H know, OK, you need to pay closer attention here, there is something you're missing. It's interesting to note that pointing out what 'has not been defined' was something that G figured would do the job, a <u>very</u> inexplicit way to say that G didn't think it was <u>possible</u> to define it as things were, but it didn't quite work. G had to warn H to pay really careful attention, and then spell out

I guess borne out in the solution they eventually adopt, which is to include this comment in the text just to <u>explain</u> that the condition is there just to simplify a later section. (I've underlined that part in Edit 4 so you can see what I mean)

OK – take a look, see what you think and send it back to me. I guess one thing we'll be doing is understanding the nature of the edits being made, which in some cases means digging in to the mathematical content. We'll see how we do!

Best,

X

Edit 4

p.1

4. Zeros of quadratic forms in fixed residue classes

Let $F \in \mathbb{Z}[x_1, \dots, x_n]$ be an isotropic quadratic form with non-zero discriminant $\Delta_F \in \mathbb{Z}$. For any positive integer M and each prime power factor $p^m \| M$ suppose that we are given a non-empty subset

$$\Omega_{p^m} \subseteq \{\mathbf{x} \in (\mathbb{Z}/p^m\mathbb{Z})^n : p \nmid \mathbf{x},\ F(\mathbf{x}) \equiv 0 \bmod p^m\}. \tag{4.1}$$

Put $\Omega_M = \prod_{p^m \| M} \Omega_{p^m}$. For $\mathbf{x} \in \mathbb{Z}^n$, we write $[\mathbf{x}]_M$ for its reduction modulo M. In this section we shall use the Hardy–Littlewood circle method to produce an asymptotic formula for the counting function

$$\widehat{N}(B, \Omega_M) = \sum_{\substack{\mathbf{x} \in \mathbb{Z}^n \\ F(\mathbf{x})=0 \\ [\mathbf{x}]_M \in \Omega_M}} w(\mathbf{x}/B), \tag{4.2}$$

where $w : \mathbb{R}^n \to \mathbb{R}_{\geqslant 0}$ is an infinitely differentiable function with compact support.

Associated to F and w is the weighted real density $\sigma_\infty(w)$, as defined in [10, Thm. 3]. It satisfies $1 \ll_{F,w} \sigma_\infty(w) \ll_{F,w} 1$. Moreover, we have the associated p-adic density

$$\sigma_p = \lim_{k \to \infty} p^{-(n-1)k} \#\{\mathbf{x} \in (\mathbb{Z}/p^k\mathbb{Z})^n : F(\mathbf{x}) \equiv 0 \bmod p^k\}, \tag{4.3}$$

for each prime p. The goal of this section is to prove the following result.

Theorem 4.1. *Assume that $n \geqslant 5$ and that $\nabla F(\mathbf{x}) \gg 1$ for all $\mathbf{x} \in \operatorname{supp}(w)$. Assume that M is coprime to $2\Delta_F$ and let Ω_M be as in (4.1). Then*

$$\widehat{N}(B, \Omega_M) = \sigma_\infty(w) B^{n-2} \prod_{p \nmid M} \sigma_p \prod_{p^m \| M} \frac{\#\Omega_{p^m}}{p^{m(n-1)}} + O_{\varepsilon,F,w}\left(B^{n/2+\varepsilon} M^{n/2+\varepsilon}\right),$$

for any $\varepsilon > 0$.

In this result and henceforth in this section, the implied constant is allowed to depend on the choice of ε, the form F and the weight function w, but not on the modulus M. To ease notation we shall suppress this dependence in what follows.

Some comments are in order about the statement of this result. The condition that $\nabla F(\mathbf{x}) \gg 1$ for any $\mathbf{x}$ in the support of w is required to simplify the analysis of the oscillatory integrals that appear in the argument. The assumptions $(M, 2\Delta_F) = 1$ and $(\mathbf{x}, M) = 1$ for any $\mathbf{x} \in \Omega_M$ are made purely to simplify the expression for the leading constant in the asymptotic formula for $\widehat{N}(B, \Omega_M)$.

This second part is what's mentioned in the email!

Theorem 4.1 can be improved in several directions. Firstly, an inspection of the proof reveals that one does rather better in the B-aspect of the error term when n is odd. Secondly, it would not be hard to deal with quadratic forms in $n = 3$ or 4 variables. Finally, when M is square-free it is possible to improve the error term to $O(B^{n/2+\varepsilon} \#\Omega_M^{1/2})$. In order to simplify our exposition we have decided not to pursue any of these improvements in the present investigation. In our application Ω_M will be comparable in size to the set of $\mathbf{x} \in (\mathbb{Z}/M\mathbb{Z})^n$ for which $F(\mathbf{x}) \equiv 0 \bmod M$, and so we have relaxed the dependence on $\#\Omega_M$. In fact, although wasteful, we shall often employ the trivial inequality $\#\Omega_M \leqslant M^n$.

Original in colour

Hi there, X!

Yes, all makes sense so far!

Well, not quite – you talked about recognising Ω_{p^m} and Omega_{p^M} and that completely confused me at first until I figured out that the latter is the LaTeX for the former! (I found an online tool that lets you type in the LaTeX and it show you the expression: mathquill.com. That helped me with translation.) I realised that it was important for me to <u>RECOGNISE THE SHAPE OF THE EXPRESSION</u> to quickly see what was going on. It's funny – everything used to be handwritten like this, and mathematicians developed all these interesting spatial notations, like

$$\sum_{i=0}^{10}(i^3+1)$$

and so on, that use top and bottom and superscripts and subscripts and big symbols and so on. And then everything switched to computers and it's a huge pain. So people developed LaTeX to typeset these complex formulae, and all the mathematicians got familiar with it, but still when they want to talk to each other it's over email which is restricted to that linear plain text. So they just use what they type in on the <u>back end</u> of LaTeX, since they're both also familiar with what that looks like! But it loses those recognisable shapes. And that made things more confusing for me, at least. It's really interesting to see the way they're <u>adapting</u> email as a means of communication – making it conversation-like by replying inline,

Original in colour

turning it into LaTeX to overcome the typesetting limitations. Sort of adapting the medium to approximate aspects of the traditional shared blackboard scenario.

So, stepping back. These are two people trying to write something that others eventually will read. They want it to be consistent with certain rules in each step. By that I mean the 'rules' of mathematics, the rules agreed upon by the mathematical community. But they also seem to want to consider something that we might call clarity, or we might call 'rules' of communication. So there are adjustments being made that are of quite different characters.

Para 1. G is expecting to write something up, thinks it will be easy, wonders how interesting it is. There's an INTERESTINGNESS VS. SPACE BALANCE going on here (which it strikes me might be quite a crucial notion throughout mathematics—there's such a lot of talk of elegance). So as they write up even things they think they have figured out and that obey all of the mathematical rules, they are also concerned about what is going to be worth saying. Things being correct doesn't make them worth saying—writing this letter I could sit here and tell you that the fourth word of email 1 is 'not' and the second to last is 'Best'. But why would I? So there are some similar considerations happening here. Back when David Hilbert was posing his famous problems, there was an effort to find a set of clear and provably consistent axioms from which all mathematics could be derived. This was known as Hilbert's program, and the 'provably' part was famously proved impossible by Gödel's incompleteness theorems. I'm wandering into irrelevancy myself, here - point is, if you took a set of axioms like that and derived everything that they implied, you wouldn't get something that looks like mathematics as we know it. You would get tons and tons of very boring but true statements, with the ones that we think are interesting, that people care about one way or the other, interspersed among them. And here it sounds like even statements that human mathematicians have considered and written might not make the interestingness cut'. So what's that? Something that might mean something to somebody, might help answer something with broader implications, in short, might have a few implications. Sometimes it's kind of hard to predict what might be useful later. So they have to weigh all of that up.

Para 2. I've stared at this a lot and here I think this might be a question of correctness, of adhering to mathematical rules. Which is cool! Mathematics has these agreed systems, ways of deciding the relationships of various elements, and G is saying that H has unknowingly violated one of those. It's very important for them to stay in accordance with those agreed systems, or else what they're doing won't even be mathematics. So I guess it might be interesting to figure out what has been violated. It has to do with defining Ω_{pM}, and G's explanation is that 'the subset of a product of sets is not a product of subsets! So our next task might be to understand what that means, what kind of rule it might be, and how it applies.

Para 3. What I think is happening is that they are adding some kind of restriction to (4.1) to make it easier (or clearer) to write something that comes 7 pages later. G suggests adding a note to explain it and it looks like this is what they do. So this is the other kind of rule, not mathematical because there wasn't anything that was 'wrong', so what, then? When you're reading page 17, the reason for the condition is not clear, it would seem to be unnecessary or rather unmotivated. It might be a bit of unnecessary complexity, or a bit of a red herring, causing a reader to look for a reason for its inclusion that actually won't show up until page 24. So we're looking at some kind of consideration of a person reading the paper as a person who will LOOK FOR PURPOSE in the paper they're reading, looking for what a STATEMENT is to be used for (intentions perhaps?), for the structure of an argument, and it's a part of the authors' work as they see it to consider that. And they don't want red herrings, they want to GUIDE ATTENTION in all the right ways. And extra conditions aren't in themselves desirable, since they restrict the implications of a STATEMENT. So an explanation will be sought.

gcd (x, M) = 1

this means that the greatest common denominator of x and M is 1.

So what H said is:

the problem comes in using Hensel to get a nice simple form of the singular series for the primes p\mid M.
At the top of page 24, one wants to prove N(k+1)=p^{n-1}N(k) for k\geq m. This fails if p\mid \grad F(x);
ie. if p|x, or if p|Delta_F.
since we are ultimately working with projective points, I thought it was simplest to assume it in this section.
one can derive a version without the primitivity condition (or the (M,2Delta)=1 condition), but the singular series will not be so explicit.

i could record a version with the singular series uncomputed (ie just a product \prod_p \sigma_p(\OMega)), valid without either of these 2 assumptions, and then a remark/lemma about what happens to it when p\nmid \x in \Omega and p\nmid 2 Delta_F?

So whatever it is on page 24 'fails' if p divides x, or if p divides ΔF. I'm not sure about the latter but if p divides M, and x and M have a common denominator (other than 1), then p might divide x.

I think I'll just 'translate' those LaTeX phrases I was having trouble with. I'll get on better just 'seeing' i.e. quickly perceiving links between the paper and emails if I, as a LaTeX novice, can just RECOGNISE THE FAMILIAR SHAPES. I think some of it isn't quite straightforward but I'll muddle through Ok. Thank goodness we're writing the old-fashioned way, eh!

All best

Y

Original in colour

Email 3

p. 1

katiemccallum@live.co.uk

From:	H [redacted].ac.uk>
Sent:	28 April 2017 09:41
To:	[redacted] G
Subject:	Re: Sieves paper

[illegible]

[redacted]

On 23 April 2017 at 19:59, G [redacted] wrote:

Hi [redacted]

Thanks for the response. I've been going through what you have written and understanding it a bit better now.

On 21/04/17 13:59, H [redacted] wrote:

> ps. The nabla F(x)\gg 1 condition is used in analysing the singular integral (p24).

Yes I think I see this now. Could you please add a small comment saying something like this condition is something to do with choosing a "good" weight function which simplifies the analysis of the singular integral?

On 21 April 2017 at 13:56, H [redacted]ac.uk> wrote:

> I know I did not get round to writing the proof of Theorem 1.6. But as should be clear to you it will be very easy. I wanted to discuss a bit with you in Chicago how much of this stuff you wanted to include, as I'm not convinced how interesting it is.

me neither, but it doesn't take much space.

> Theorem 4.1: I think you are overlooking some subtleties here. I don't think there is a nice way to define the Omega_{p^m}. The reason is that the subset of a product of sets is

Ω_{p^m}

Email 3

p. 2

not a product of subsets. So you cannot in general write Omega_M as a product of sets of the form Omega_{p^m}. I think this statement and the proof might need rewriting.

right; i always need to be a product. i guess it will be better to assume the existence of a subset for each prime in (4.1), and to let \Omega_M=prod_p \Omega_{p^m}.

> Also I find the condition in (4.1) that gcd(\x,M) = 1 a bit strange. Why is this condition present? For example, you don't allow Omega_M to consist of the zero vector. Why not? I guess this does not matter for the application, as we are dealing with primitive integer points?

the problem comes in using Hensel to get a nice simple form of the singular series for the primes p\mid M. At the top of page 24, one wants to prove N(k+1)=p^{n-1}N(k) for k\geq m. This fails if p\mid \grad F(x); ie. if p|x, or if p|Delta_F. since we are ultimately working with projective points, I thought it was simplest to assume it in this section. one can derive a version without the primitivity condition (or the (M,2Delta)=1 condition), but the singular series will not be so explicit.

i could record a version with the singular series uncomputed (ie just a product \prod_p \sigma_p(\OMega)), valid without either of these 2 assumptions, and then a remark/lemma about what happens to it when p\nmid \x in \Omega and p\nmid 2 Delta_F?

Thanks for the explanation. I guess it's your call what you would like to do. I guess a small comment these conditions are to simplify the calculation of the singular series, would not go a miss.

Best,

[redacted] G

[redacted] H

Original in colour

Translation!

Email Translation

On 21 April 2017 at 13:56, [redacted] <[redacted]@[redacted].ac.uk> wrote:

>> I know I did not get round to writing the proof of Theorem 1.6. But as should be clear to
>>you it will be very easy. I wanted to discuss a bit with you in Chicago how much of this
>>stuff you wanted to include, as I'm not convinced how interesting it is.
>me neither, but it doesn't take much space.

>> Theorem 4.1: I think you are overlooking some subtleties here. I don't think there is a nice
>>way to define the Ω_{p^m}. The reason is that the subset of a product of sets is not a
>>product of subsets. So you cannot in general write Ω_M as a product of sets of the
>>form Ω_{p^m}. I think this statement and the proof might need rewriting.
>right; i always need to be a product. i guess it will be better to assume the existence of a
>subset for each prime in (4.1), and to let $\Omega_M = \Pi_p \Omega_p{}^m$.

>> Also I find the condition in (4.1) that gcd(x,M) = 1 a bit strange. Why is this condition
>>present? For example, you don't allow Ω_M to consist of the zero vector. Why not? I
>>guess this does not matter for the application, as we are dealing with primitive integer
>>points?
>the problem comes in using Hensel to get a nice simple form of the singular series for the
>primes p | M. At the top of page 24, one wants to prove $N(k+1)=p^{n-1}N(k)$ for k≥
>m. This fails if p | grad F(x); ie. if p|x, or if $p|\Delta_F$.
>since we are ultimately working with projective points, I thought it was simplest to assume
>it in this section. one can derive a version without the primitivity condition (or the
>(M,2Δ)=1 condition), but the singular series will not be so explicit.

>i could record a version with the singular series uncomputed (ie just a product Π_p
>$\sigma_p(\Omega)$), valid without either of these 2 assumptions, and then a remark/lemma
>about what happens to it when
> $p \nmid x$ in Ω and $p \nmid 2\Delta_F$?
Thanks for the explanation. I guess it's your call what you would like to do. I guess a small
comment these conditions are to simplify the calculation of the singular series, would not go
a miss.

Best,

[G]

>Best, [H]

Original in colour

Dear Y,

Thanks for the reply! A couple of comments:

I loved your comment about (4.1), that "in mathematics you generally want to express things as simply as possible and for them to have as great a scope as possible. I absolutely agree, and that echoes something we saw in the previous excerpt where the group were working up their STATEMENT and adjusting the wording. Specifics of wording become incredibly important in that kind of situation, since 1. you want your statement to be as clear and succinct as possible and 2. you need to be very careful about what it includes and excludes, you can't have a carelessly worded statement exclude something or include something it oughtn't. So there's this very intense focus on minute particulars of wording, and on getting the scope of each part of the STATEMENT precisely correct. And you really want to maximise scope wherever you can since that's likely to maximise interestingness, understood as some function on the possible implications of a STATEMENT. I absolutely think that the concept of elegance applies here— THE GREATER THE RATIO OF EFFECTS TO MEANS, THE GREATER THE ELEGANCE, or something along those lines.

And on the topic of relevance, there's been some great research on beauty in mathematics. It's a common habit to associate beauty with simplicity when talking about mathematical theorems, but there's good reason to question that association. In Inglis and Aberdein (2014) a team had mathematicians write down adjectives to describe different pieces of work, and did a factor analysis to see whether beauty-assessments and simplicity-assessments correlated. They totally didn't! What *did* correlate with beauty were things like elegance and insightfulness. So perhaps there's some assessment of effects-for-means at play when making aesthetic judgements. (I actually think this might be the case elsewhere)

On (4.1), I think your 'red herring' explanation is a nice one. The writers have something they're trying to achieve, and the reader is trying to follow them through that endeavour, so unexpected moves need some explanation. If I were watching a video of someone chopping up vegetables, putting oil into a pan, turning on the hob, etc. then I could follow along and make guesses about what that person was intending to do. If they then started setting aside every fifth carrot then it would be helpful if they would look at the camera and tell me why. I think what we're seeing here is similar to that. The following paragraph, where they added the comment about the condition, is more clearly directed to a reader, and is explicitly guiding a person through the text. But the STATEMENT of (4.1), that's not the place to make that comment, it's more like the authors are just laying something out, just *showing*, and the commentary happens separately. THE TEXT IS SOMETIMES EXHIBITED ACTION AND SOMETIMES COMMENTARY, maybe, or something like that. But the 'exhibited' part is important, they're still clearly thinking about how that STATEMENT will be received.

It's a bit like us, sharing those sheets we're looking at and scribbling on and also sharing our thoughts on them in these letters. Actually I could've used a little commentary on what you were up to starting to cut up the pages like that, then there were bits falling out of the envelope! But once I got to that page I understood what you were trying to do. I guess emails are a little easier for quoting, huh? Though I suppose doing it your way everything looks exactly as it does in the original!

Best,

T

Original in colour

Original in colour

Hey X!

Yeah, interesting! This EXHIBITED ACTION thing: I like the analogy with a cooking show, especially since mathematical papers are so often written in the imperative! On this page we have 'Let', 'suppose', 'Put', 'Assume'. It's as though some of it takes the form of instruction or action-speak, and then other parts are commentary. I've done some highlighting on Edit 4 to show what I mean, instruction/action in blue and commentary in red. Some parts overlap, where the authors seem to be explaining where things are going next. Others seem to be almost directed to the assessments that the authors are expecting the reader to be making of the work, or at least to the reader's experience, in a way that I think is really clearly addressed to a person. The overlapping parts sort of lay out what's happening and what's happening next, in a way that has to be helpful to a reader, but is also expressed as statement of fact in a way that is less clearly reader-oriented. The blue parts seem to be statements in the imperative interspersed with expressions in notation.

And yes, sorry! That must have confused you!

All best

Y

Original in colour

Edit 4

p. 1

4. Zeros of quadratic forms in fixed residue classes

Let $F \in \mathbb{Z}[x_1, \ldots, x_n]$ be an isotropic quadratic form with non-zero discriminant $\Delta_F \in \mathbb{Z}$. For any positive integer M and each prime power factor $p^m \| M$ suppose that we are given a non-empty subset

$$\Omega_{p^m} \subseteq \{\mathbf{x} \in (\mathbb{Z}/p^m\mathbb{Z})^n : p \nmid \mathbf{x},\ F(\mathbf{x}) \equiv 0 \bmod p^m\}. \tag{4.1}$$

Put $\Omega_M = \prod_{p^m \| M} \Omega_{p^m}$. For $\mathbf{x} \in \mathbb{Z}^n$, we write $[\mathbf{x}]_M$ for its reduction modulo M. In this section we shall use the Hardy–Littlewood circle method to produce an asymptotic formula for the counting function

$$\widehat{N}(B, \Omega_M) = \sum_{\substack{\mathbf{x} \in \mathbb{Z}^n \\ F(\mathbf{x})=0 \\ [\mathbf{x}]_M \in \Omega_M}} w(\mathbf{x}/B), \tag{4.2}$$

where $w : \mathbb{R}^n \to \mathbb{R}_{\geqslant 0}$ is an infinitely differentiable function with compact support.

Associated to F and w is the weighted real density $\sigma_\infty(w)$, as defined in [10, Thm. 3]. It satisfies $1 \ll_{F,w} \sigma_\infty(w) \ll_{F,w} 1$. Moreover, we have the associated p-adic density

$$\sigma_p = \lim_{k \to \infty} p^{-(n-1)k} \#\{\mathbf{x} \in (\mathbb{Z}/p^k\mathbb{Z})^n : F(\mathbf{x}) \equiv 0 \bmod p^k\}, \tag{4.3}$$

for each prime p. The goal of this section is to prove the following result.

Theorem 4.1. *Assume that $n \geqslant 5$ and that $\nabla F(\mathbf{x}) \gg 1$ for all $\mathbf{x} \in \operatorname{supp}(w)$ Assume that M is coprime to $2\Delta_F$ and let Ω_M be as in (4.1). Then*

$$\widehat{N}(B, \Omega_M) = \sigma_\infty(w) B^{n-2} \prod_{p \nmid M} \sigma_p \prod_{p^m \| M} \frac{\#\Omega_{p^m}}{p^{m(n-1)}} + O_{\varepsilon,F,w}\left(B^{n/2+\varepsilon} M^{n/2+\varepsilon}\right),$$

for any $\varepsilon > 0$.

Same

In this result and henceforth in this section, the implied constant is allowed to depend on the choice of ε, the form F and the weight function w, but not on the modulus M. To ease notation we shall suppress this dependence in what follows.

Some comments are in order about the statement of this result. The condition that $\nabla F(\mathbf{x}) \gg 1$ for any $\mathbf{x}$ in the support of w is required to simplify the analysis of the oscillatory integrals that appear in the argument. The assumptions $(M, 2\Delta_F) = 1$ and $(\mathbf{x}, M) = 1$ for any $\mathbf{x} \in \Omega_M$ are made purely to simplify the expression for the leading constant in the asymptotic formula for $N(B, \Omega_M)$.

This second part is what's mentioned in the email!

Theorem 4.1 can be improved in several directions. Firstly, an inspection of the proof reveals that one does rather better in the B-aspect of the error term when n is odd. Secondly, it would not be hard to deal with quadratic forms in $n = 3$ or 4 variables. Finally, when M is square-free it is possible to improve the error term to $O(B^{n/2+\varepsilon} \#\Omega_M^{1/2})$. In order to simplify our exposition we have decided not to pursue any of these improvements in the present investigation. In our application Ω_M will be comparable in size to the set of $\mathbf{x} \in (\mathbb{Z}/M\mathbb{Z})^n$ for which $F(\mathbf{x}) \equiv 0 \bmod M$, and so we have relaxed the dependence on $\#\Omega_M$. In fact, although wasteful, we shall often employ the trivial inequality $\#\Omega_M \leqslant M^n$.

Edit 4

p. 2

H and G

We begin the proof of Theorem 4.1 by invoking the version of the circle method developed by Heath-Brown [10, Thm. 1]. This implies that

$$\widehat{N}(B,\Omega_M)=\frac{c_Q}{Q^2}\sum_{q=1}^{\infty}\ \sideset{}{^*}\sum_{a \bmod q}\ \sum_{\substack{\mathbf{x}\in\mathbb{Z}^n\\ [\mathbf{x}]_M\in\Omega_M}} w(\mathbf{x}/B)e_q(aF(\mathbf{x}))h\left(\frac{q}{Q},\frac{F(\mathbf{x})}{Q^2}\right),$$

for any $Q>1$. Here c_Q is a positive constant satisfying $c_Q=1+O_A(Q^{-A})$ for any $A>0$ and, moreover, $h(x,y)$ is a smooth function defined on the set $(0,\infty)\times\mathbb{R}$ such that $h(x,y)\ll x^{-1}$ for all y, with $h(x,y)$ non-zero only for $x\leqslant\max\{1,2|y|\}$. In particular, we are only interested in $q\ll Q$ in this sum.

We will henceforth take $Q=B$. It is natural to break the sum into residue classes modulo the least common multiple $[q,M]$ and then apply Poisson summation, as in the proof of [10, Thm. 2]. This leads to the expression

$$\widehat{N}(B,\Omega_M)=\frac{c_B}{B^2}\sum_{q\ll B}\sum_{\mathbf{c}\in\mathbb{Z}^n}[q,M]^{-n}S_{q,M}(\mathbf{c})J_{q,M}(\mathbf{c}),$$

where

$$S_{q,M}(\mathbf{c})=\sideset{}{^*}\sum_{a \bmod q}\ \sum_{\substack{\mathbf{y} \bmod [q,M]\\ [\mathbf{y}]_M\in\Omega_M}} e_q\left(aF(\mathbf{y})\right)e_{[q,M]}\left(\mathbf{c}.\mathbf{y}\right) \tag{4.4}$$

and

$$\begin{aligned}J_{q,M}(\mathbf{c})&=\int_{\mathbb{R}^n}w(\mathbf{x}/B)h\left(\frac{q}{B},\frac{F(\mathbf{x})}{B^2}\right)e_{[q,M]}(-\mathbf{c}.\mathbf{x})d\mathbf{x}\\&=B^n\int_{\mathbb{R}^n}w(\mathbf{x})h\left(\frac{q}{B},F(\mathbf{x})\right)e_{[q,m]}(-B\mathbf{c}.\mathbf{x})d\mathbf{x}.\end{aligned}$$

It will be convenient to set

$$I_r^*(\mathbf{v})=\int_{\mathbb{R}^n}w(\mathbf{x})h\left(r,F(\mathbf{x})\right)e_r(-\mathbf{v}.\mathbf{x})d\mathbf{x}, \tag{4.5}$$

for any $r>0$ and $\mathbf{v}\in\mathbb{R}^n$. In this notation, which coincides with that introduced in [10, §7], we may clearly write $J_{q,M}(\mathbf{c})=B^nI_r^*(M'^{-1}\mathbf{c})$, where $r=q/B$ and $M'=[q,M]/q=M/(M,q)$. Thus

$$\widehat{N}(B,\Omega_M)=c_BB^{n-2}\sum_{q\ll B}\sum_{\mathbf{c}\in\mathbb{Z}^n}[q,M]^{-n}S_{q,M}(\mathbf{c})I_r^*(M'^{-1}\mathbf{c}). \tag{4.6}$$

4.1. **The exponential sum.** The purpose of this section is to analyse the sum $S_{q,M}(\mathbf{c})$ in (4.4) for $q,M\in\mathbb{N}$ such that M is coprime to $2\Delta_F$. We begin by establishing the following factorisation property.

Lemma 4.2. *Let $M=M_1M_2$. Suppose that $(q_1M_1,q_2M_2)=1$ and choose integers s,t such that $[q_1,M_1]s+[q_2,M_2]t=1$. Then*

$$S_{q_1q_2,M}(\mathbf{c})=S_{q_1,M_1}(t\mathbf{c})S_{q_2,M_2}(s\mathbf{c}).$$

Proof. Note that $[q_1q_2,M]=[q_1,M_1][q_2,M_2]$. As $\mathbf{y}_1$ runs modulo $[q_1,M_1]$ and $\mathbf{y}_2$ runs modulo $[q_2,M_2]$, so $\mathbf{y}=\mathbf{y}_1[q_2,M_2]t+\mathbf{y}_2[q_1,M_1]s$ runs over a full set of residue classes modulo $[q_1q_2,M]$. Now let $\bar{q}_1,\bar{q}_2\in\mathbb{Z}$ be such that $q_1\bar{q}_1+q_2\bar{q}_2=1$. Then $a=a_1q_2\bar{q}_2+a_2q_1\bar{q}_1$ runs over $(\mathbb{Z}/q_1q_2\mathbb{Z})^*$ as a_1 (resp. a_2) runs over $(\mathbb{Z}/q_1\mathbb{Z})^*$

Original in colour

Dear Y,

Yeah, so I feel like I've been making the case a lot that emails are kind of trying to approximate other forms like handwriting and conversation, so it's nice for us to see a way that the email medium really comes into its own. Actually ~~the~~ it seems as though maybe this medium, of correspondence, really suits the kind of refining, adjusting work they do, more than it does the quicker messier work of early-stage ideas. There's a lot less testing and discarding of possible ideas going on, and a lot more standing back and looking – and looking at a text (shared text) in the way a reader might, which is the whole objective. I wonder if that's a theme – mathematicians meeting in person to hash out ideas, then working on how to stabilise and communicate them remotely, by correspondence. If the work is to be communicated it has to become somehow more stable in its presentation, less dependent on circumstances and corrections. And that's the direction the paper is going.

Right, so how about we dig in a little to the second part? Theorem 4.1. I Googled G's explanation a little, stuff about sets, subsets and products. So G said: if you take the product of (kind of like multiplying) two sets and take ~~the~~ a subset of that, it isn't the same as if you take two subsets and take the product of them. I haven't got much further than remembering a few terms and what 'taking the product' entails.

I have done a bit of work on just understanding some of the notation, though. Capital pi, Π, is used to talk about the product of a series, and we see that in the theorem, along with mentions of Ω_M and Ω_{p^m}. I've tried finding the first time that p^m shows up in the paper since that's likely to give us a bit of an introduction.

SIEVING RATIONAL POINTS 5

1.2.1. *A version of the Selberg sieve.* Our fundamental tool will be a version of the Selberg sieve for rational points on quadrics. Let $m \in \mathbb{N}$ be fixed once and for all. For each prime p we suppose that we are given a non-empty set of residue classes $\Omega_{p^m} \subseteq X(\mathbb{Z}/p^m\mathbb{Z})$. Our goal is to measure the density of points $x \in X(\mathbb{Q})$ whose reduction modulo p^m lands in Ω_{p^m} for each prime p. Namely, we are interested in the behaviour of the counting function

$$N(X, H, \Omega, B) = \#\{x \in X(k) : H(x) \leqslant B,\ x \bmod p^m \in \Omega_{p^m} \text{ for all } p\}$$

as $B \to \infty$, where $\Omega = (\Omega_{p^m})_p$. This function has order of magnitude B^{n-1} when $\Omega_{p^m} = X(\mathbb{Z}/p^m\mathbb{Z})$ for all p, but we expect it to be significantly smaller when Ω_{p^m} is a proper subset of $X(\mathbb{Z}/p^m\mathbb{Z})$ for many primes p. We define the density function

$$\omega_p = 1 - \frac{\#\Omega_{p^m}}{\#X(\mathbb{Z}/p^m\mathbb{Z})} \in [0, 1], \qquad (1.3)$$

for any prime p. The following is our main result for quadrics.

Thing is, so much of what makes mathematics so incomprehensible to the outsider is just the notation and terminology – it's so specialised, it's really tricky to get a foot in the door. So this is me, looking at the introduction they've written and using Google to find definitions and so forth. That's one use of resources we should note, since we just need to know what the definitions of some words are!

So here goes, I'm going to attempt a natural language translation, so at least the language will be recognisable to us.

Let m be an element in the set of natural numbers (positive integers. For each prime p we suppose that we are given a non-empty set of residue classes

artofproblemsolving.com/wiki/index.php/Residue_class

Apps | Must Lists: 50 Scien... | Herb Profiles | Mathematical symb... | Birds of North Ame... | Radio Boise

Art of Problem Solving | AoPS Online | Beast Academy | AoPS Academy

AoPS Online

SIGN IN

ONLINE CLASSES BOOKSTORE COMMUNITY RESOURCES

▶ RESOURCES » AOPS WIKI » RESIDUE CLASS

Residue class

In modular arithmetic, a residue of an integer a in modulo n is the unique value of $0 \leq r \leq n-1$ such that $a = kn + r$. In the context of division, a residue is simply a remainder.

A residue class is a complete set of integers that are congruent modulo n for some positive integer n. In modulo n, there are exactly n different residue classes, corresponding to the n possible residues $\{0, 1, 2, 3, \ldots n-2, n-1\}$

Each residue class contains all integers in the form $kn + r$ where r is the corresponding residue.

So a residue class is a set of integers that give the same result when divided by something: like 1, 13, 25, 37, 49 all have remainder 1 when divided by 12. $kn + r \rightarrow 12k + 1$

So the non-empty set of sets of integers that give the same remainder when divided by n ... if n were 12, then possible values for r might be 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 so you could have some subset of those.

It's the set of residue classes, the set of sets of integers.

Original in colour

so the non-empty set of residue classes is named Ω_{p^m}, and is written: $\Omega_{p^m} \subseteq X(\mathbb{Z}/p^m\mathbb{Z})$

is a subset of or equal to

set of integers

Whole thing means: the set of residue classes of the integers modulo p^m i.e. the set of sets of numbers that give the same remainder when the integers are divided by p^m

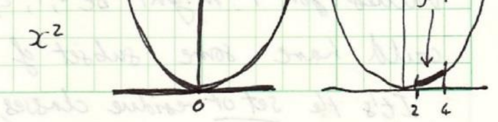
So this whole expression I think just means that the set of residue classes we're looking at is a subset of all the possible residue classes. What about that 'X'?

1.2. **Sieving on quadrics.** By restricting attention to quadrics we can show that quantitatively stronger versions of the previous results are possible. For the remainder of this section $X \subset \mathbb{P}^n_{\mathbb{Q}}$ is a smooth quadric hypersurface of dimension at least 3 over $\mathbb{Q}$ and $H : \mathbb{P}^n(\mathbb{Q}) \to \mathbb{R}$ is the standard exponential height function associated to the supremum norm. There is a natural choice of model $\mathcal{X}$ given by the closure of X inside $\mathbb{P}^n_{\mathbb{Z}}$; we shall abuse notation and write $X(\mathbb{Z}) = \mathcal{X}(\mathbb{Z})$ and $X(\mathbb{Z}/m\mathbb{Z}) = \mathcal{X}(\mathbb{Z}/m\mathbb{Z})$.

This is on the previous page. So I think that's defining X as a hypersurface, which is a subset of projective space ($\subset \mathbb{P}^n_{\mathbb{Q}}$). So when you get X(something) it means X defined across those things, like a function. So we're thinking of a subset of all the possible residue classes when divided by p^m as being a part of a hypersurface. This isn't super clear to me what that implies! But let's move on.

So: ~~the~~ our set of residue classes when dividing by p^m is a subset of or equal to the hypersurface defined across all the residue classes for all integers when dividing by p^m.

I guess looking at a part of a hypersurface is a little like looking at a section of a graph, e.g.:



Original in colour

Next sentence:

Our goal is to measure the density of points $x \in X(\mathbb{Q})$

is a member of — a hypersurface — defined across rational numbers

whose reduction modulo p^m lands in Ω_{p^m} for each prime p.

another way of saying residue

So they're trying to count (estimate?) how many integers give a certain set of remainders when divided by p^m. This is referred to as 'sieving', as in you make 'holes' at certain remainders and see which integers fall through!

So what might this do? If we had a number x, and 'sifted' for remainder zero when divided by every number smaller than x, then if nothing fell through we'd know that x was prime! It looks like there are applications like finding out how many twin primes there are, that is primes that are only 2 apart. Things like that.

So, from the introduction:

1. Introduction

Sieves are a ubiquitous tool in analytic number theory and have numerous applications. Typically, one is given a subset $\Omega_p \subset \mathbb{Z}/p\mathbb{Z}$ for each prime p and the challenge is to count the number of integers n in an interval for which $n \bmod p \in \Omega_p$ for all p. In general only upper bounds are realistic, and these can be obtained through a variety of means, the most successful being variants of the large sieve or the Selberg sieve, as explained in [13, Chapters 7–9].

The above set-up can be generalised in many ways, such as in the abstract version of the large sieve developed by Kowalski [14, §2.1] ...

This is all encouraging. And they're working with a kind of 'sieve' developed by Atle Selberg

which is this guy →

So... we learned a ton. But not so much why we're looking at p^m rather than p, and what M is. In our extract we get

④ For any positive integer M and each prime power factor $p^m \| M$ suppose that we are given a non-empty subset

in this context this notation means that m is the highest power of p that will divide M. Imagine M is 18. P is 3. If m is 2 then 3^2 will divide 18. If m is 3, 3^3 won't!

I had guessed that M and m were going to be related, since they're the same letter – sometimes you can intuit relationships based on the notation, since it's been chosen for a reason, and they wouldn't give you a RED HERRING!

So, we saw G saying 'Ω_{p^m} has not been defined in this section, only Ω_M' in edit 3, the sentence above wasn't in there yet. It looked like this:

Let $F \in \mathbb{Z}[x_1, \ldots, x_n]$ be an isotropic quadratic form with non-zero discriminant $\Delta_F \in \mathbb{Z}$. For any positive integer M, suppose that we are given a non-empty subset

$$\Omega_M \subseteq \{\mathbf{x} \in (\mathbb{Z}/M\mathbb{Z})^n : \gcd(\mathbf{x}, M) = 1,\ F(\mathbf{x}) \equiv 0 \bmod M\}. \quad (4.1)$$ ③

For $\mathbf{x} \in \mathbb{Z}^n$, we write $[\mathbf{x}]_d$ for its reduction modulo M. In this section we shall use the Hardy–Littlewood circle method to produce an asymptotic formula f...

This sentence is replaced by this

and this expression by this one

$$\Omega_{p^m} \subseteq \{\mathbf{x} \in (\mathbb{Z}/p^m\mathbb{Z})^n : p \nmid \mathbf{x},\ F(\mathbf{x}) \equiv 0 \bmod p^m\}. \quad (4.1)$$ ④

Note the changes

Looks like they're 'sieving' for exact division by p^m ie. remainder 0.

They also add this right after:

④ Put $\Omega_M = \prod_{p^m \| M} \Omega_{p^m}$.

The subset of residue classes mod M / is the product of / (wherever p^m is the highest power of p that divides M) / The subset of residue classes mod p^m

Gosh, they do pack a lot of meaning in to those very short expressions, don't they! We're taking pages and pages to unpack it all!

So... before they added in those statements about how Ω_M and Ω_{p^m} related, they were just jumping in to talking about Ω_{p^m}

product of — This means size of the set

... should not cause too many problems) and let Ω_M be as in (4.1). Then

③ $$\widehat{N}(B,\Omega_M) = \sigma_\infty(w)B^{n-2}\prod_{p\nmid M}\sigma_p \prod_{p^m \| M} \frac{\#\Omega_{p^m}}{p^{m(n-1)}} + O_{\varepsilon,F,w}\left(B^{n/2+\varepsilon}M^{n/2+\varepsilon}\right),$$

First appearances in this section

for any $\varepsilon > 0$. ♣♣♣ G [redacted] [Ω_{p^m} has not been defined in this section, only Ω ...
[It is too obvious to repeat here. I'll answer the remaini...

They say that mathematicians are just machines for turning coffee into theorems

So... if p^m divides M then Ω_M is a subset of Ω_{p^m}, right?

2' more things are divided by this to give remainder 0 — 4 than by this — so $\Omega_M \supset \Omega_{p^m}$

The above expression is written including $\prod$ (something involving Ω_{p^m}) so that's the ... product of sets. G says that 'the subset of a product of sets is not a product of subsets. So you cannot in general write Ω_M as a product of sets of the form Ω_{p^m}.' But they add in that earlier definition of Ω_M in terms of Ω_{p^m}, and that solves the problem.

In Email 2, H says 'right; i always need to be a product.' This is probably a typo, with the intended sentence being 'it always needs to be a product' or 'I always need it to be a product.' But it's tempting to think that the mathematician may have been thinking something like 'so I'm Ω_M, and I have to be a product'! Somehow puzzling through all of this makes the possibility of inhabiting Ω_M and Ω_{p^m} almost inviting. But perhaps I'm getting a bit weird from thinking about it too long!

Alright, so we've started to get a sense of where the sets and subsets are in the paper, but I still don't see what the problem might be with assuming that the subset of a product of sets is a product of subsets, and I've Googled those phrases like mad and found nothing. We might need to appeal to a real live mathematician for help. Happy for me to do that? I guess this is an important aspect of expertise: having the ability, from knowledge of all of the various rules of set theory, to answer very specific questions about it on demand, and know (perhaps more importantly) what the <u>point</u> of that question might be...

Best,

X

Original in colour

Sure, go ahead!

All best

Y

Original in colour

Original in colour

Right, so this is the relevant part of the reply I got!

THE PROBLEM WAS EXACTLY AS I SAID: THAT A SUBSET OF A PRODUCT OF SETS IS NOT A PRODUCT OF SUBSETS. IN OUR SETTING, WE ARE GIVEN AN INTEGER M AND ARE TRYING TO RELATE A SUBSET OF $(\mathbb{Z}/M\mathbb{Z})^n$ TO THE PRODUCT OF SUBSETS OF $(\mathbb{Z}/p^m\mathbb{Z})^n$ WHERE p^m RUNS OVER THE PRIME POWERS DIVIDING M. BUT THIS CAN'T BE DONE IN GENERAL. HERE $\mathbb{Z}/M\mathbb{Z}$ DENOTES THE SET OF INTEGERS MODULO M.

FOR COMPLETE CLARITY, TAKE M=6 AND n=1. ALSO I FORGET THE EQUATION $F(x)=0$ AS IT IS NOT RELEVANT HERE. THEN WE HAVE $6=2\times 3$. BY THE CHINESE REMAINDER THEOREM WE HAVE

$\mathbb{Z}/6\mathbb{Z} = \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z}$ (PRODUCT OF SETS).

EXPLICITLY, THE ELEMENTS OF $\mathbb{Z}/6\mathbb{Z}$ ARE $\{0,1,2,3,4,5\}$ AND THE ELEMENTS OF $\mathbb{Z}/2\mathbb{Z}$ AND $\mathbb{Z}/3\mathbb{Z}$ ARE $\{0,1\}$ AND $\{0,1,2\}$, RESPECTIVELY.

THE BIJECTION IS GIVEN BY AN ELEMENT x OF $\mathbb{Z}/6\mathbb{Z}$ MAPS TO THE PAIR (x MOD 2, x MOD 3). E.G. 4 MAPS TO (0,1).

NOW I TAKE THE SUBSET OF $\mathbb{Z}/6\mathbb{Z}$ GIVEN BY $X = \{0,1\}$. ASSUME THAT THIS IS A PRODUCT OF SUBSETS X_2 OF $\mathbb{Z}/2\mathbb{Z}$ AND X_3 OF $\mathbb{Z}/3\mathbb{Z}$. THE SIZE OF A PRODUCT OF SUBSETS IS EQUAL TO THE PRODUCT OF THE SIZES, SO THAT $2 = \#X = \#(X_2 \times X_3) = \#X_2 \times \#X_3$. AS 2 IS PRIME, IT FOLLOWS THAT EITHER $\#X_2 = 1$ OR $\#X_3 = 1$, I.E. EITHER X_2 CONSISTS OF A SINGLE ELEMENT OR X_3 CONSISTS OF A SINGLE ELEMENT. BUT THE IMAGE OF X INSIDE $\mathbb{Z}/2\mathbb{Z}$

IS GIVEN BY $\{0,1\}$ AND THE IMAGE OF X INSIDE $\mathbb{Z}/3\mathbb{Z}$ IS ALSO GIVEN BY $\{0,1\}$, SO THAT NEITHER OF THESE HAVE SIZE 1. IT FOLLOWS

THAT OUR ASSUMPTION WAS FALSE; X IS NOT A PRODUCT OF SUBSETS. SO WE CAN'T WRITE X AS A PRODUCT OF SUBSETS OVER THE PRIMES DIVIDING M.

OK. So G said, 'we are given an integer M and are trying to relate a subset of $(\mathbb{Z}/M\mathbb{Z})^n$ to the product of subsets of $(\mathbb{Z}/p^m\mathbb{Z})^n$.' Great, so this is about as we thought.

Then follows what seems like a 'toy' example.

'For complete clarity, take $M=6$ and $n=1$. ... Then we have $6 = 2 \times 3$. By the Chinese remainder theorem, we have $\mathbb{Z}/6\mathbb{Z} = \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z}$ (product of sets).'

set of residue classes mod 6 = set of residue classes mod 2 × set of residue classes mod 3

I've tried to at least understand the statement of this theorem even if understanding any proof is beyond me!

⇓

iF you know the residues of n mod x, y and z you should be able to Find n mod xyz (as long as no one prime will divide any two of x y and z)

So if we know the residues of $\mathbb{Z}$ mod 2 and 3 we can Find $\mathbb{Z}$ mod 6.

'the elements of $\mathbb{Z}/6\mathbb{Z}$ are {0, 1, 2, 3, 4, 5} and the elements of $\mathbb{Z}/2\mathbb{Z}$ and $\mathbb{Z}/3\mathbb{Z}$ are {0, 1} and {0, 1, 2}, respectively.'

these are the residue classes, enumerated

'the bijection is given by an element x of $\mathbb{Z}/6\mathbb{Z}$ maps to the pair (x mod 2, x mod 3). E.g. 4 maps to (0, 1).'

one-to-one correspondence.

OK, so iF I try to make a 'mapping' For each member of $\mathbb{Z}/6\mathbb{Z}$:

x (x mod 2, x mod 3)

0 : {0,0}
1 : {1,1}
2 : {0,2}
3 : {1,0}
4 : {0,1}
5 : {1,2}

Aah, there's one of each combination! Six of them! And that looks like the way you take the product of sets, right? (Google Cartesian product to see what I mean) So then we see that the bijection, the mapping, works.

Now I take the subset of $\mathbb{Z}/6\mathbb{Z}$ given by $X = \{0,1\}$ OK, so in my list that's 4.

Assume that that is a product of subsets X_2 of $\mathbb{Z}/2\mathbb{Z}$ and X_3 of $\mathbb{Z}/3\mathbb{Z}$.

So – the fact that we're beginning by making an assumption here makes me think that we're planning on a proof by contradiction, so we're assuming something that will turn out to be false and lead to an inconsistency. The inconsistency is how we learn it's false. I think I RECOGNISE THE APPROACH!

The size of a product of subsets is equal to the product of the sizes, so that $2 = \#X = \#(X_2 \times X_3)$.

the size of $\{0,1\}$ – 2 elements

the size of these two should multiply to 2

As 2 is prime, it follows that either $\#X_2 = 1$ or $\#X_3 = 1$ i.e. either X_2 consists of a single element or X_3 consists a single element. But the image of X inside $\mathbb{Z}/2\mathbb{Z}$ is given by $\{0,1\}$ and the image of X inside $\mathbb{Z}/3\mathbb{Z}$ is also given by $\{0,1\}$, so that neither of these have size 1.

OK now I'm super lost. So if we're looking at the elements of $\mathbb{Z}/2\mathbb{Z}$ and $\mathbb{Z}/3\mathbb{Z}$ in X – X is $\{0,1\}$, so the element from $\mathbb{Z}/2\mathbb{Z}$ is just 0 and the element from $\mathbb{Z}/3\mathbb{Z}$ is just 1! (that I think is what 'the image inside' means, just the elements from that in the set) So don't <u>both</u> have size 1?? Which is also wrong!

I'm stumped. Can you help me out? I also don't quite see how answering questions about $\mathbb{Z}/6\mathbb{Z}$ is going to help us...

Best,

X

Original in colour

Original in colour

Hey X!

Firstly - yes, you're right, one thing we have to do here is to think about the nature of the explanation G's given us - I have some thoughts on that, which I'll go into later. What we have is an expert looking at a problem in a certain way and typing some sentences about a particular, manageably 'small' example in order to help us to see things the right way. So it'll be interesting to work out what the expert mind is doing to connect the two.

Alright, so the example. I got pretty stuck with this too and eventually talked it through with a friend who used to study maths (a human, expert resource!) and I think what went wrong was a question of brackets. !!

So look back at what G wrote. '**the elements of ℤ/6ℤ are {0, 1, 2, 3, 4, 5}.**' That's in curly brackets. And when G talks about the content of the mapping, '**E.g. 4 maps to (0, 1)**': normal brackets. But when you wrote out your list of mappings, it looked like this:

0 : {0, 0}		0 : (0, 0)
1 : {1, 1}	when it	1 : (1, 1)
2 : {0, 2}	should	2 : (0, 2)
3 : {1, 0}	have been	3 : (1, 0)
4 : {0, 1}	this:	4 : (0, 1)
5 : {1, 2}		5 : (1, 2)

But there was something else that helped you to get confused:

Original in colour

normal brackets →

G took that example of 4 mapping to (0, 1), and shortly afterward was talking about '*the subset of $\mathbb{Z}/6\mathbb{Z}$ given by $X = \{0, 1\}$*' ← curly brackets and all that happened was that you saw those similar-looking things and assumed that they were the same thing! But they weren't! Curly brackets are for members of a set, normal brackets for things like the mappings we were getting. the different brackets tell you about the kind of thing you're looking at in a way that's really obvious to an expert but maybe won't even be noticed by a novice.

G was talking about {4} which maps to (0, 1)
and {0, 1} " " " (0, 0) and (1, 1)

You thought you were taking the right cue, noticing the repetition of '0, 1'. But actually you were missing an important cue, the brackets, because you didn't know what to pay attention to. You can even be reading the same language, the same symbols, as an expert, but because the expert has a certain know-how for how to pay the right attention to the right parts, you might still get very different interpretations.

That was the funny thing about discussing this with my friend: every time we went through it it just sounded the same as your version but with a different outcome, because for my friend, the brackets were a huge distinction, and I just wasn't even perceiving them as different yet!

So, I'll work through it again to make it clear. We take '**the subset of $\mathbb{Z}/6\mathbb{Z}$ given by $X = \{0, 1\}$**.' So that maps to (0, 0) (1, 1) (from our earlier list). We're assuming (I think you're right, we assume wrongly on purpose) that this is '**a product of subsets X_2 of $\mathbb{Z}/2\mathbb{Z}$ and X_3 of $\mathbb{Z}/3\mathbb{Z}$**.' So as G said, if we have (0, 0) and (1, 1) then the elements we have from $\mathbb{Z}/2\mathbb{Z}$ are $\{0, 1\}$ and the elements from $\mathbb{Z}/3\mathbb{Z}$ are $\{0, 1\}$. The size of each, then, is two elements. We assumed that X was a product of X_2 and X_3. '**The size of a product of subsets**,' G told us, '**is equal to the product of the sizes**.' So 2×2. But that gives 4, and we know that the size of X is 2! So it leads to an inconsistency, so we can learn that X does not equal the product of X_2 and X_3. We can see that we don't get the same X by
1. taking a subset of the product of $\mathbb{Z}/2\mathbb{Z}$ and $\mathbb{Z}/3\mathbb{Z}$
2. taking subsets of $\mathbb{Z}/3\mathbb{Z}$ and $\mathbb{Z}/2\mathbb{Z}$ and taking their product.

So... we've learned that. But what else have we learned? How does this apply to the STATEMENT of Theorem 4.1?

The thing we've learned is that sometiimes, the subset of a product of sets and a product of sets won't be equivalent. We don't have to show that it always isn't, just that in at least one case it isn't. We could pick an example for which all of that would work, like for example
$X = \{0, 1, 3, 4\}$

that gives $\begin{cases} 0:(0,0) \\ 1:(0,1) \\ 3:(1,0) \\ 4:(1,1) \end{cases}$ image of $\mathbb{Z}/2\mathbb{Z} = \{0,1\}$
image of $\mathbb{Z}/3\mathbb{Z} = \{0,1\}$

size is 4 — sizes are 2

So that would all behave nicely when treated as above.

But that clearly doesn't always work! And we've seen that. G picked an example where it doesn't work to show us that. Mathematics is very concerned with what is always/never the case, and that's what is meant by saying '**this can't be done in general**', it's a case of whether you can always know that it'll work, or not.

So G said that they were '**trying to relate a subset of $(\mathbb{Z}/M\mathbb{Z})^n$**' (that's Ω_M) '**to the product of subsets of $(\mathbb{Z}/p^m\mathbb{Z})^n$**' (the product of Ω_{p^m}). Theorem 4.1 writes Ω_M as some function on the product of Ω_{p^m}, but that isn't something that can be done in general it doesn't always work, it depends on the particular subset! So all they do is add that bit we saw after (4.1) saying something like 'this is an instance where this subset can be made by taking the product of these other subsets'.

This took me a really long time to get a grip on, and I'm still not clear on all the details, so I hope you basically agree! I originally had kind of a backwards picture of what G's explanation was intended to do. I thought at first that G

Original in colour

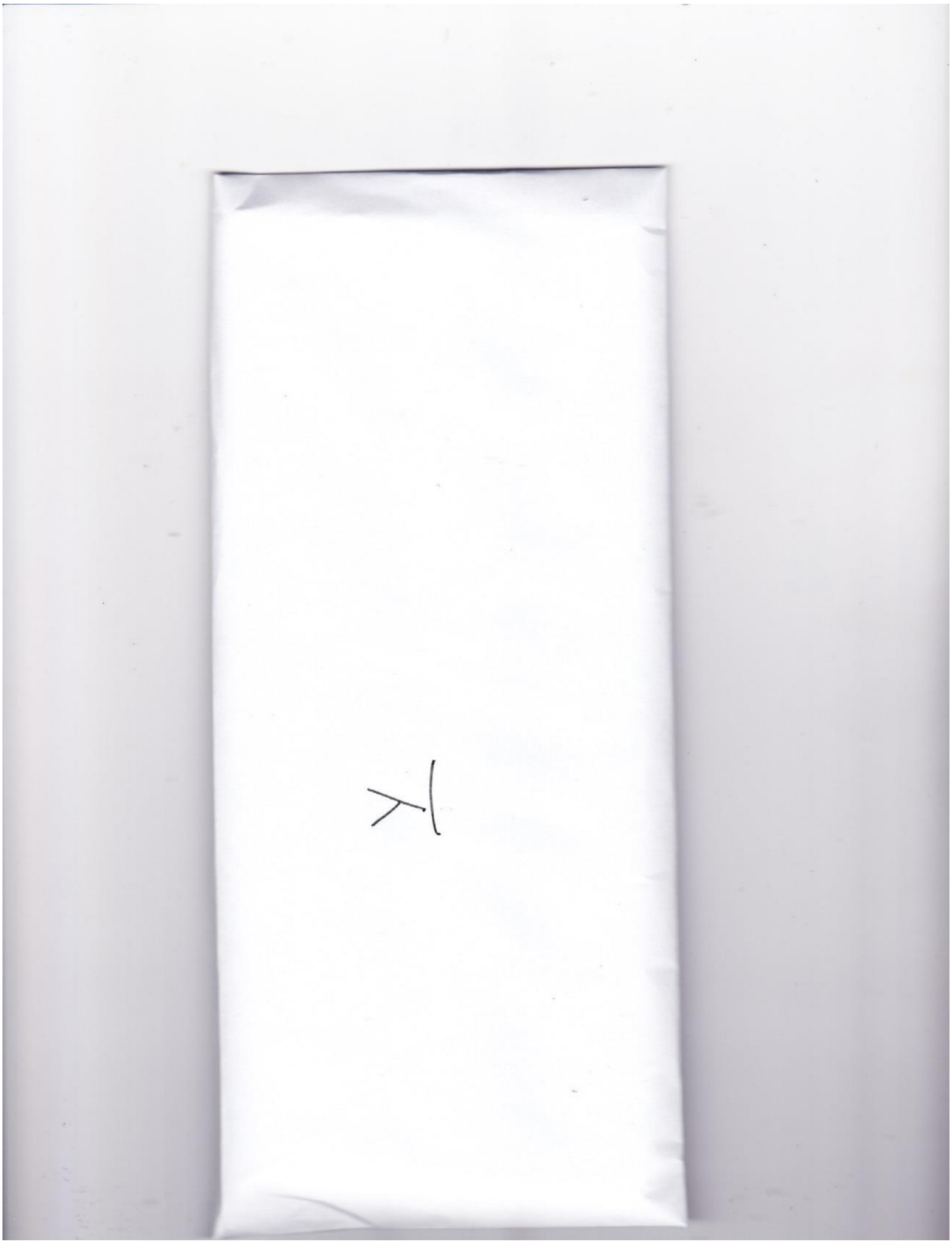
was demonstrating that a subset of a product of sets is never the same as a product of subsets, and I couldn't see why it would never work in the example (let alone in the context of the paper). But I kept staring at G's email and kind of RECOGNISED THE APPROACH, what G was intending to show us by describing it. Let's say a person is talking about something and whether it can always/sometimes/never be done. If you talk through just one example showing it doesn't work, that won't prove anything about whether it sometimes works or never works. But it does prove that it doesn't always work, because you have an example of it not working to point to! And I think that would be dead obvious to an expert – someone tells you some STATEMENT, but your experience in mathematical argumentation tells you which implications to pay attention to.

Whew! This is hard stuff. I feel I have a much shakier grasp on this one, perhaps because there are no pictures and perhaps because we're seeing the polishing stage not the ideas stage. I'm still finding the process of trying to make sense of it fascinating, though.

Brain achingly,

Y

Original in colour



Original in colour

Dear Y,

This is so great. You're totally right – and it just goes to show how subtle and important the effects of experience can be. It isn't just knowing what's right and what's a mistake, it's knowing <u>which things need to be paid attention</u> to know the difference. A person with experience might know

- Brackets are important. There are ones like this for members {} and ones like this for content ()()(). They aren't the same!
- Mathematical arguments proceed in certain structures, like proof by contradiction. If you have a certain set of STATEMENTS, there might be all kinds of things you can derive from them – RECOGNISING A STRUCTURE might tell you which to pay attention to.
- If you're talking about what's true <u>in general</u>, you might pay attention to a certain kind of evidence

etc.

This is why we have teachers, why a human person is so much a better explainer than a book. A person knows how to <u>use</u> a text, how to pay attention, and which symbols and implications are most important, and probably knows where an untrained mind might go astray in perceiving what's going on. Actually I told G about the way I'd misunderstood, and apparently the confusion about brackets is 'a common mistake made by 1st year undergraduates whilst they are still learning to do "real" mathematics'! I think that the above is part of why mathematical papers are hard to read, and people are so keen to meet in person. There is so much packed in to a single character, and true expertise involves knowing where to pay attention, and how to <u>use</u> a text, and the written record doesn't always offer much help with that – everything's so austere and succinct that it doesn't always wear its importance on its sleeve, so to speak! Other than occasional commentary. I love what you said about your friend's explanations just sounding like the same as what I was saying. Even with a human explainer, someone with expertise can be talking about two, separate concepts, and the listening newbie can still miss the distinction and think they're talking about just one thing. But the human explainer can notice that and tell the newbie to pay closer attention, right here.

Looking at your page I realised I could colour it just the way you did the paper – as EXHIBITED ACTION, AND COMMENTARY! You worked through something in front of me, and then gave commentary on the implications!

So I'll work through it again to make it clear. We take '**the subset of $\mathbb{Z}/6\mathbb{Z}$ given by $X = \{0, 1\}$.**' So that maps to (0, 0) (1, 1) (from our earlier list). We're assuming (I think you're right, we assume wrongly on purpose) that this is '**a product of subsets X_2 of $\mathbb{Z}/2\mathbb{Z}$ and X_3 of $\mathbb{Z}/3\mathbb{Z}$.**' So as G said, if we have (0, 0) and (1, 1) then the elements we have from $\mathbb{Z}/2\mathbb{Z}$ are $\{0, 1\}$ and the elements from $\mathbb{Z}/3\mathbb{Z}$ are $\{0, 1\}$. The size of each, then, is two elements. We assumed that X was a product of X_2 and X_3. '**The size of a product of subsets,**' G told us, '**is equal to the product of the sizes.**' So 2×2. But that gives 4, and we know that the size of X is 2! So it leads to an inconsistency, so we can learn that X does not equal the product of X_2 and X_3. We can see that we don't get the same X by
1. taking a subset of the product of $\mathbb{Z}/2\mathbb{Z}$ and $\mathbb{Z}/3\mathbb{Z}$
2. taking subsets of $\mathbb{Z}/3\mathbb{Z}$ and $\mathbb{Z}/2\mathbb{Z}$ and taking their product.

So... we're learned that. But what else have we learned? How does this apply to the STATEMENT of Theorem 4.1?

The thing we've learned is that sometiimes, the subset of a product of sets and a product of sets won't be equivalent. We don't have to show that it always isn't, just that in at least one case it isn't. We could pick an example for which all of that would work, like for example

$X = \{0, 1, 3, 4\}$

that gives $\begin{cases} 0:(0,0) \\ 1:(0,1) \\ 3:(1,0) \\ 4:(1,1) \end{cases}$ image of $\mathbb{Z}/2\mathbb{Z} = \{0,1\}$, image of $\mathbb{Z}/3\mathbb{Z} = \{0,1\}$

size is 4 — sizes are 2

So that would all behave nicely when treated as above.

But that clearly doesn't always work! And we've seen that. G picked an example where it doesn't work to show us that. Mathematics is very concerned with what is always/never the case, and that's what is meant by saying '**this can't be done in general**', it's a case of whether you can always know that it'll work, or not.

So G said that they were '**trying to relate a subset of $(\mathbb{Z}/M\mathbb{Z})^n$**' (that's Ω_M) '**to the product of subsets of $(\mathbb{Z}/p^m\mathbb{Z})^n$**' (the product of Ω_{p^m}). Theorem 4.1 writes Ω_M as some function on the product of Ω_{p^m}, but that isn't something that can be done in general it doesn't always work, it depends on the particular subset! So all they do is add that bit we saw after (4.1) saying something like 'this is an instance where this subset can be made by taking the product of these other subsets'.

This took me a really long time to get a grip on, and I'm still not clear on all the details, so I hope you basically agree! I originally had kind of a backwards picture of what G's explanation was intended to do. I thought at first that G

I've found it really interesting to see how G and H have exploited the properties of email and of comments on the text to understand one another. At times, they've responded to each other in situ to use text like a conversation, situated, localised, in place. As when they were both commenting on Theorem 4.1 in the document. Then, G was actually able to direct H's attention by switching to email, where they continued responding in a conversation-like way to each point in turn. That's one way that email is more versatile than letters, being able to quote! But one advantage <u>we</u> had was not being typographically limited the way they were, and what filled that gap was their mutual experience with the back end of LaTeX, so that Omega_M was just as familiar (well, almost) as Ω_M.

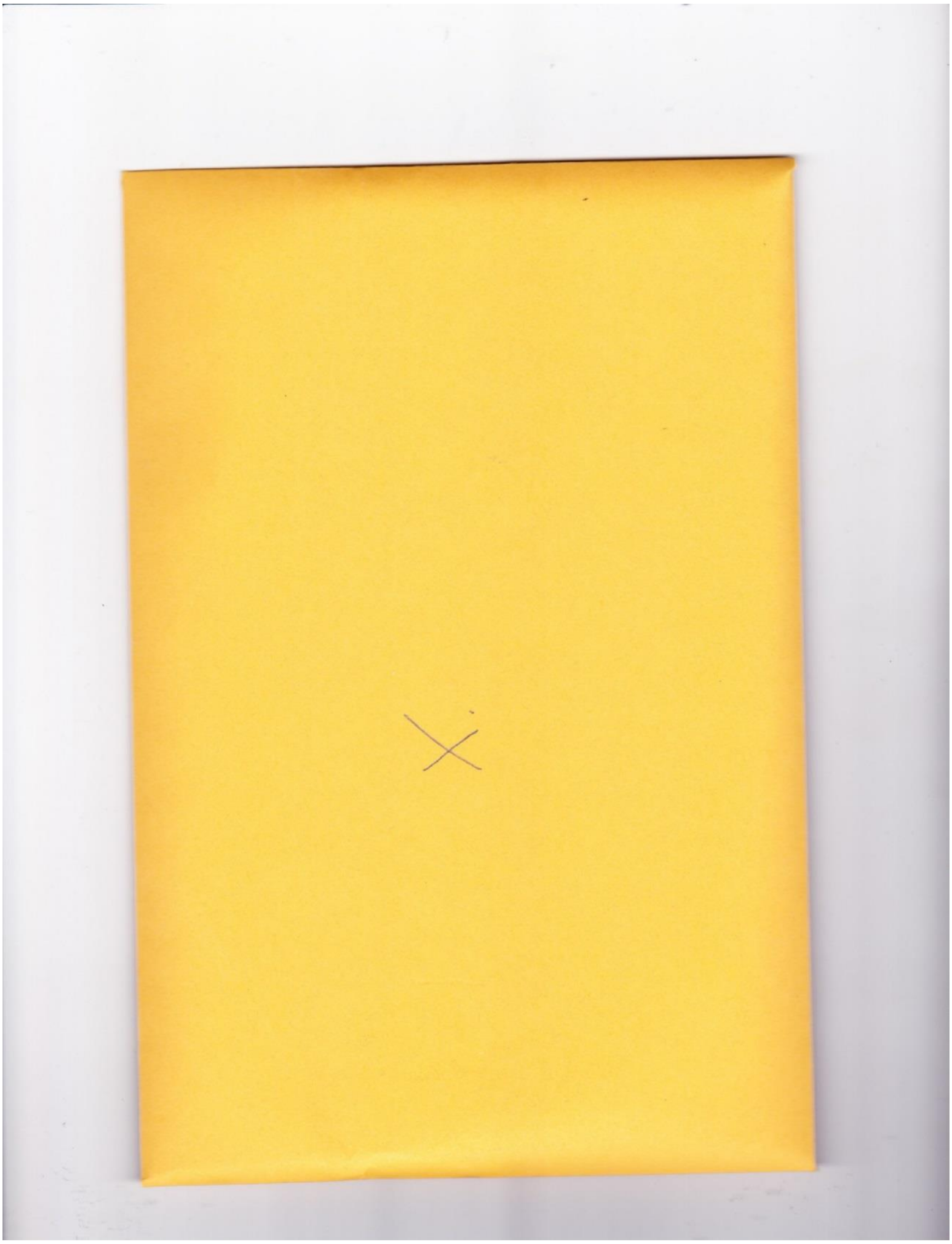
I really think, though, that one reason (among many) why we found this excerpt so much more difficult to penetrate was that we were mainly limited to text. We're not getting all the extra, human clues about how to pay attention from how/where someone's standing, their gestures, etc; we're somewhat at the mercy of how well we can find our way around the text. And even when you have the literal meaning down, that can still be hard! Being told that something is the case is one thing, but coming to see what it <u>means</u>,

how that knowledge should be *used*, is quite another!
And coming to understand what G was telling H in the email has also helped us to see what *kind* of expertise it was that H was making use of in understanding G — expertise that is not just about mathematical rules, or definitions, but also the *kind* of thing that can cause a problem. Which was hard for us to figure out!

We also learned that they were expecting readers of (4.1) to be looking for the purpose of the components of the STATEMENT they were reading. So again, we're getting a sense of how active a thinker is in reading a text, how much informed interpretive work is involved. Which reminds me, I've spoken to several people who say they generally read papers with pen in hand, scribbling on the page, working through examples. So it's like the explicit content of a paper is only half the story when coming to *understand* a piece of mathematics; the reader uses the text in a skilful way, in the course of constructing a working understanding. The text is maybe like a base, like footholds, and the expert mind knows something of how to climb them. Or like dance notation, and the expert knows how to turn the moves into a dance!

Best,

X



Original in colour

Hi X

Great! I like your summary. Here are a few more things:

We're talked about the process of refining a paper involving close attention to an INTERESTINGNESS VS. SPACE BALANCE, which seems like it connects up with some of what we saw as the group refined their STATEMENT in the last excerpt: making it precise, neat, succinct. Both paper and STATEMENT need to be as efficient and wieldy as possible, while maximising scope/usefulness. We're also seen how demanding this means texts can be on the reader, with so much hanging on a single character.

We discussed how papers seem to accommodate this by having both EXHIBITED ACTION AND COMMENTARY in the text, which might sit at different points between showing something and telling something, just like the difference between 'thinking out loud' in front of one another and addressing one another in the last excerpt. Both are being made intentionally available to be seen, and that's clear, but there are subtle differences in the nature of the evidence. In the case of EXHIBITED ACTION, I might wonder what the person is trying to do, but I could almost do that if covertly observing them writing. In the case of COMMENTARY, the writer is clearly considering my experience and trying to adjust it; so the intentions have to do with me, with changing my thinking.

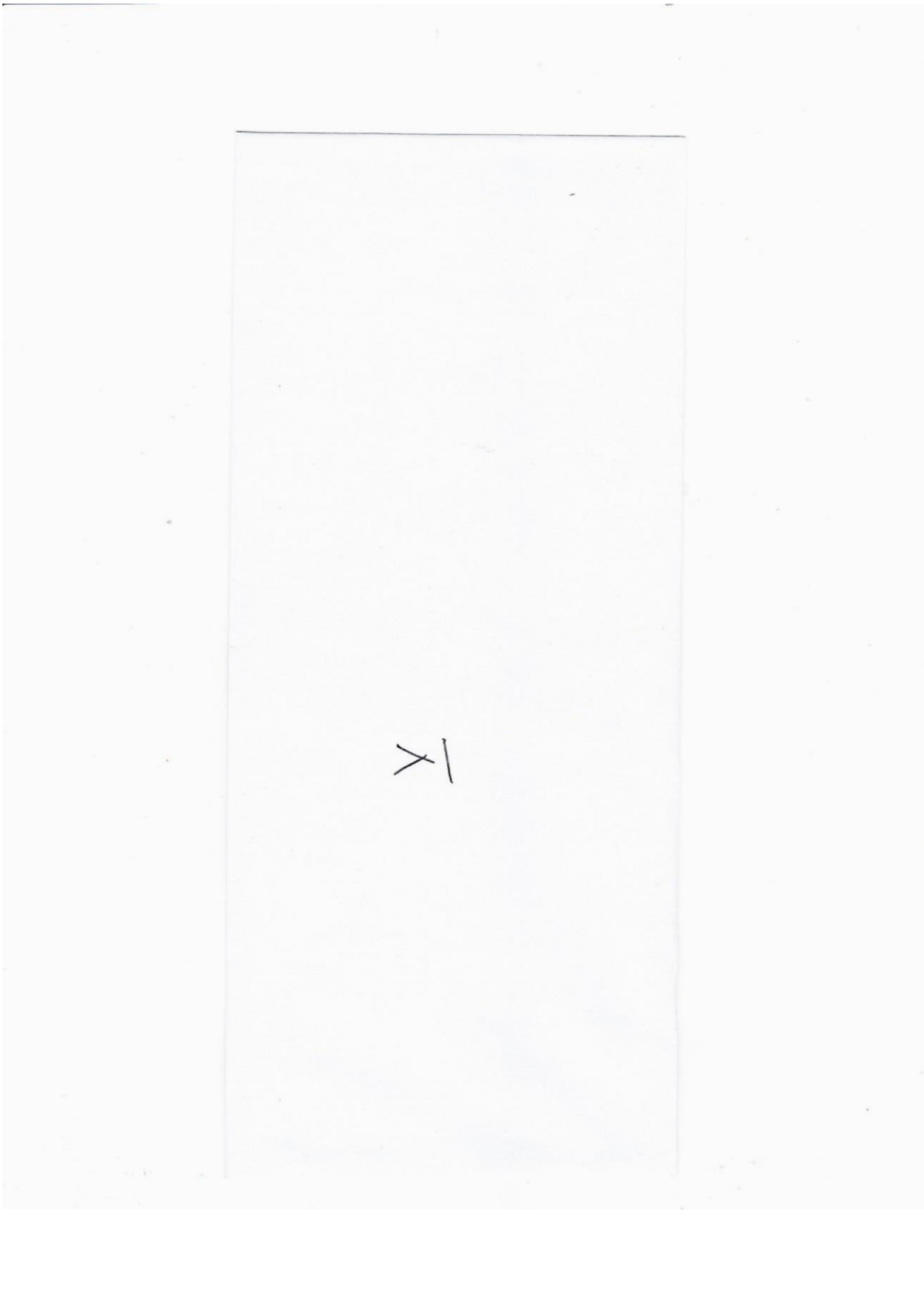
Original in colour

Thirdly, we decided that RECOGNISING THE SHAPE OF SOMETHING — a notation, a phrase, a type of proof, an argument — can help an expert to know where to pay attention: to this subscript, this character, this implication. So some part of becoming an expert is coming to recognise a situation and so how to use the text. And sometimes that's what the COMMENTARY can help with.

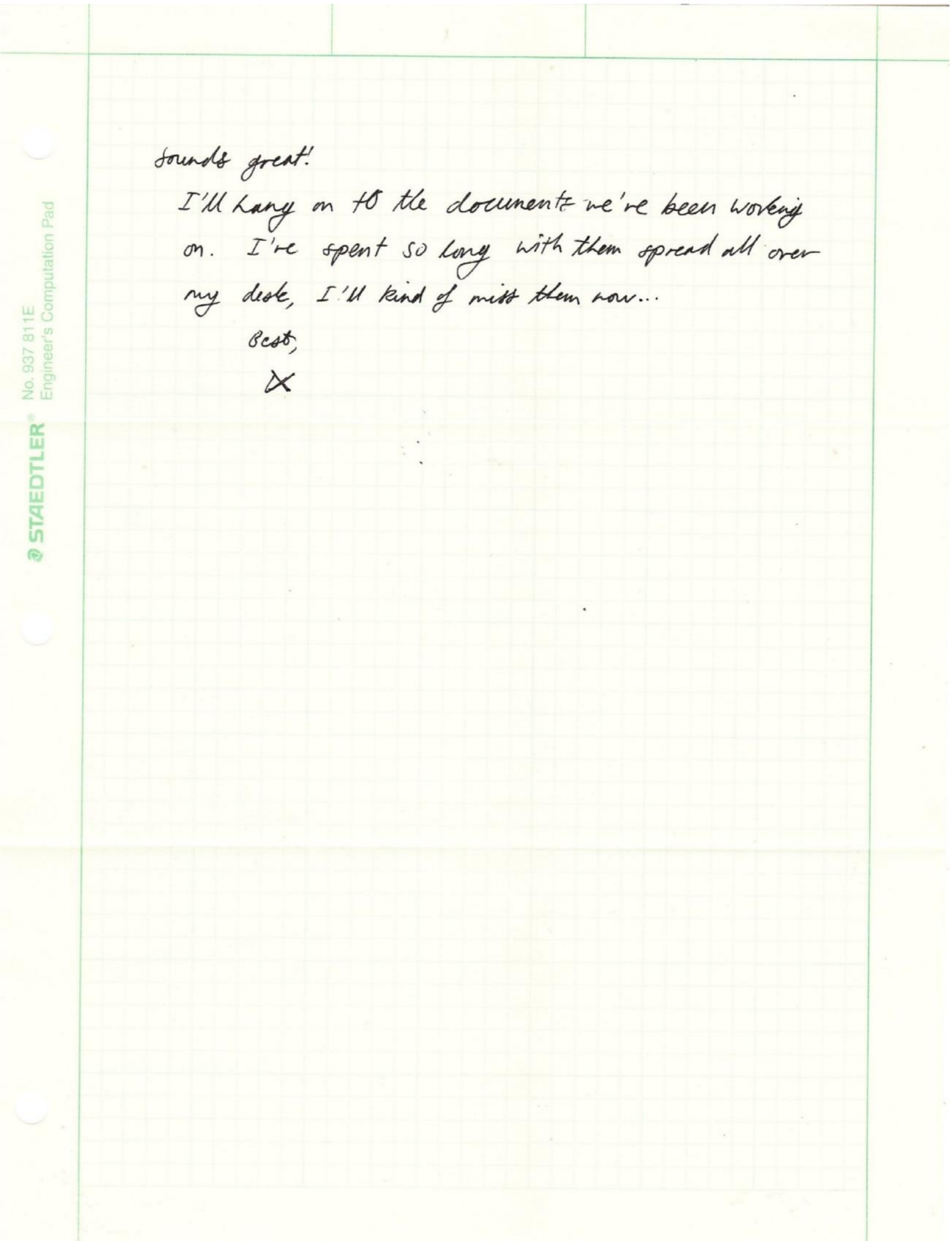
How does that all sound? I think we're at a natural stopping point, you?

All best

Y



Original in colour



STAEDTLER® No. 937 811E Engineer's Computation Pad

Sounds great!

I'll hang on to the documents we've been working on. I've spent so long with them spread all over my desk, I'll kind of miss them now...

Best,

X

Original in colour

4.4 Summary

This analysis is written in the form of a correspondence, in which two characters discuss the character of mathematical writing and the multiple functions that it fulfils for a reader. Again, the characters approach the material first from a lay perspective, then bringing in basic and then more advanced existing concepts from the world of mathematics.

A key observation from this chapter is that mathematical writing appears to fulfil quite different functions even within the same text. There is a sort of distinction between sections that seem to *exhibit action*, simply laying out sequences of expressions so that a reader can observe what is being done in the progression, and sections that seem closer to ordinary communication, giving commentary on what is going on. Both of these might be examples of ostensive-inferential communication, since both are kinds of evidence being openly and deliberately made available to a potential reader; in the first case, though, the reader is closely observing the progression to note what the author is doing with each move, and in the second, the reader recognises what the author intends that the reader recognises: a clearer case of *meaning* in the Gricean sense. In Chapters 6 and 7, I will discuss how this kind of variation is treated by Grice and relevance theory under the concepts of *showing* and *meaning*.

There is a much greater degree of explicitness in certain parts of the writing seen in the section of the paper than in the participants' more informal email exchanges (let alone the face-to-face communication of the last chapter), so an interesting question is whether the code model is better equipped to explain this more formal aspect of mathematical communication. In particular, the code used in symbolic expressions seems clearly well-defined enough to fully encapsulate meaning. However, what we see in the participants' email exchange is that in the process of writing and refining even the statement of a theorem, the authors fully expect readers to be making certain inferences about the intentions of the authors in formulating an expression in a certain way, and take this into consideration in order to avoid misleading readers. As such it seems reasonable to suppose that the comprehension of the content of a mathematical paper involves some kind of inference according to perceived intentions.

With this consideration in mind, the authors decide to include an explanatory note in the written commentary to explain why a condition is included. This decision can readily be explained in relevance theoretic terms: if a reader cannot see *why*, say, a particular condition is included, a slightly complicating factor that might restrict the implications of an expression and make it more complex to read (thus affecting both cognitive effort and effects), then the expectation of optimal relevance ought to prompt a search for justifying additional cognitive effects to justify this inclusion. The note in the text serves to guide this search.

Authors seem to pay close attention to the balance between *interestingness* and *space* when deciding whether to include something in the paper; the content needs to be deemed interesting enough to justify the space dedicated to it. The resemblance between this consideration and that of relevance is evident.

A final observation pertains to the nature of the difficulties experienced by a novice in interpreting mathematical texts. I, and so one of my fictional characters, ran into a problem when I failed to recognise the importance of a typographic detail: which brackets were being used, which is something that an expert would easily recognise as a meaningful detail. The kind of expertise that goes to shape an experience of *recognition* was a recurring theme through this analysis, the ability to recognise an approach, a type of expression, a type of proof and so on having a significant effect on ease of comprehension of these complex texts. Expertise, then, is something that can have an effect on cognitive effort in reading, or aptitude with dealing with, such mathematical texts simply by virtue of the familiarity a reader has with certain written forms and thus that reader's ability to distinguish which elements need to be paid attention. For an expert, the accessibility of a particular interpretation, one favoured by the mathematical community, might be significantly greater than for a novice; the same goes for the salience of the elements in an expression that might be expected to lead to important effects.

Interlude 5. Tiers of notes

Several of my participants told me that they kept multiple tiers of notes as they worked, beginning with very scrappy, rough notes on loose paper, and increasing in neatness until eventually the work is typed up in LaTeX. Three of these agreed to let me photograph some of these multi-tiered notes.

Notes from subject K

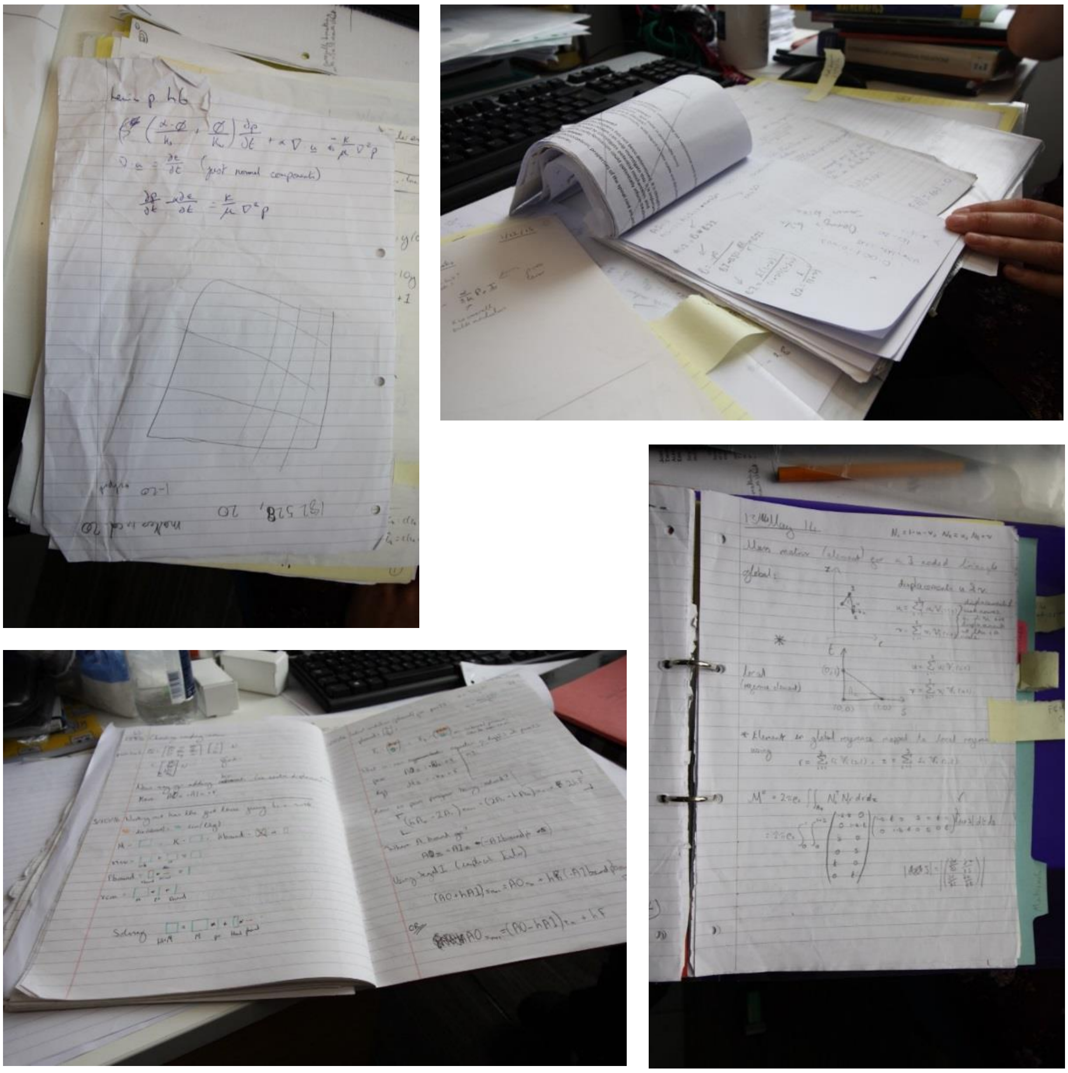
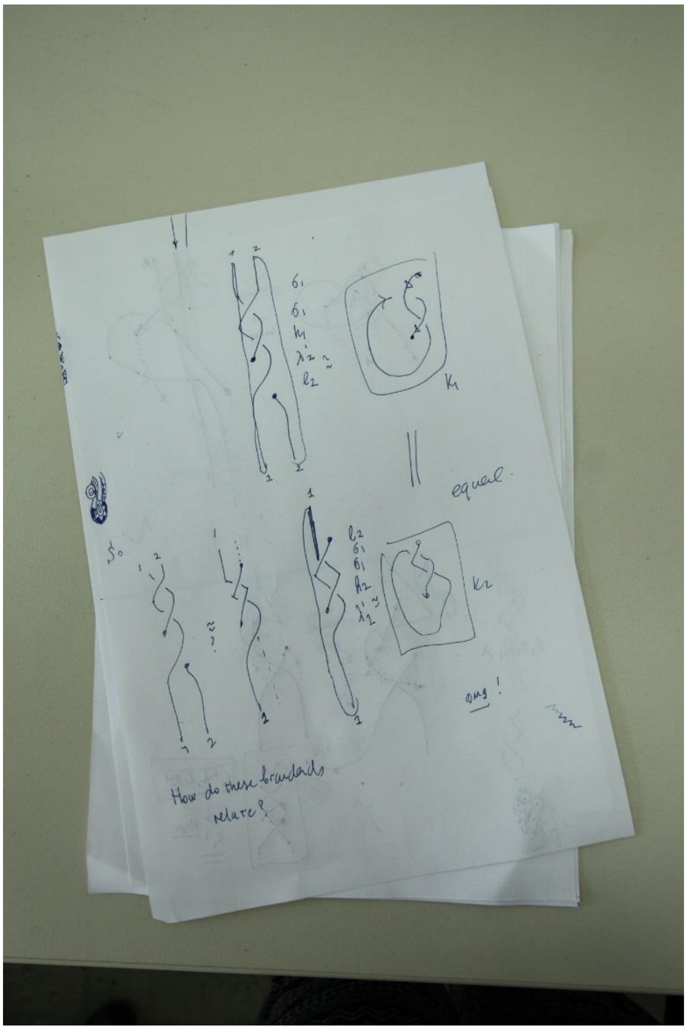


Figure 58. Clockwise in order of formality: notes on scraps of paper, notes in a home-made notebook, neater notes in a folder and large notebook. Original in colour.



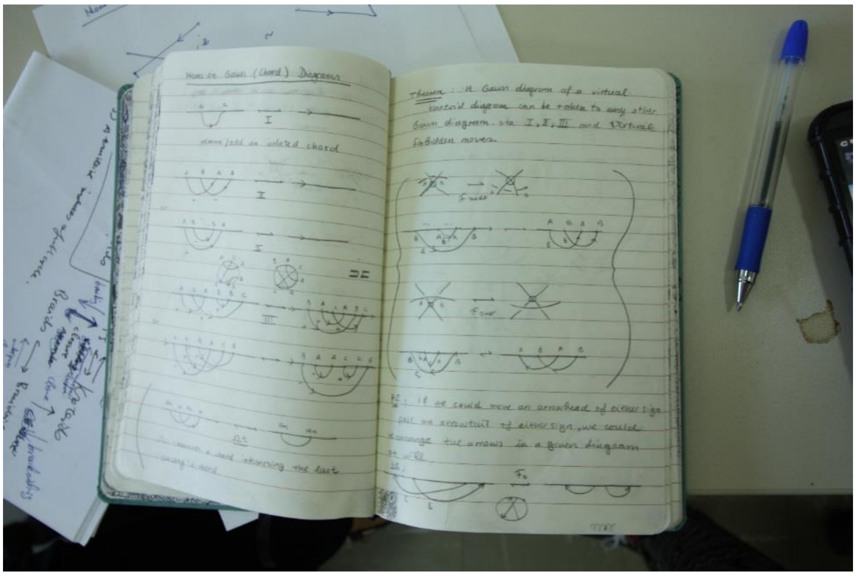


Figure 59. Initial thoughts and workings-out scribbled down on sheets of A4 paper, which develop into much neater, titled pages in a notebook. Original in colour.

0.12.57.000

So you would start out with kind of writing, scribbling it out on-

Always. Definitely by hand, yeah. And then I usually don't keep- somethings, um, if I'm doing something new I'll keep my little scribbles, like these are my working out some result. But I'll end up throwing this away. So I just can- I can tell when I have something better by if it has a date on it. […] And then once I write it up I'll throw this away.

Notes from Subject M

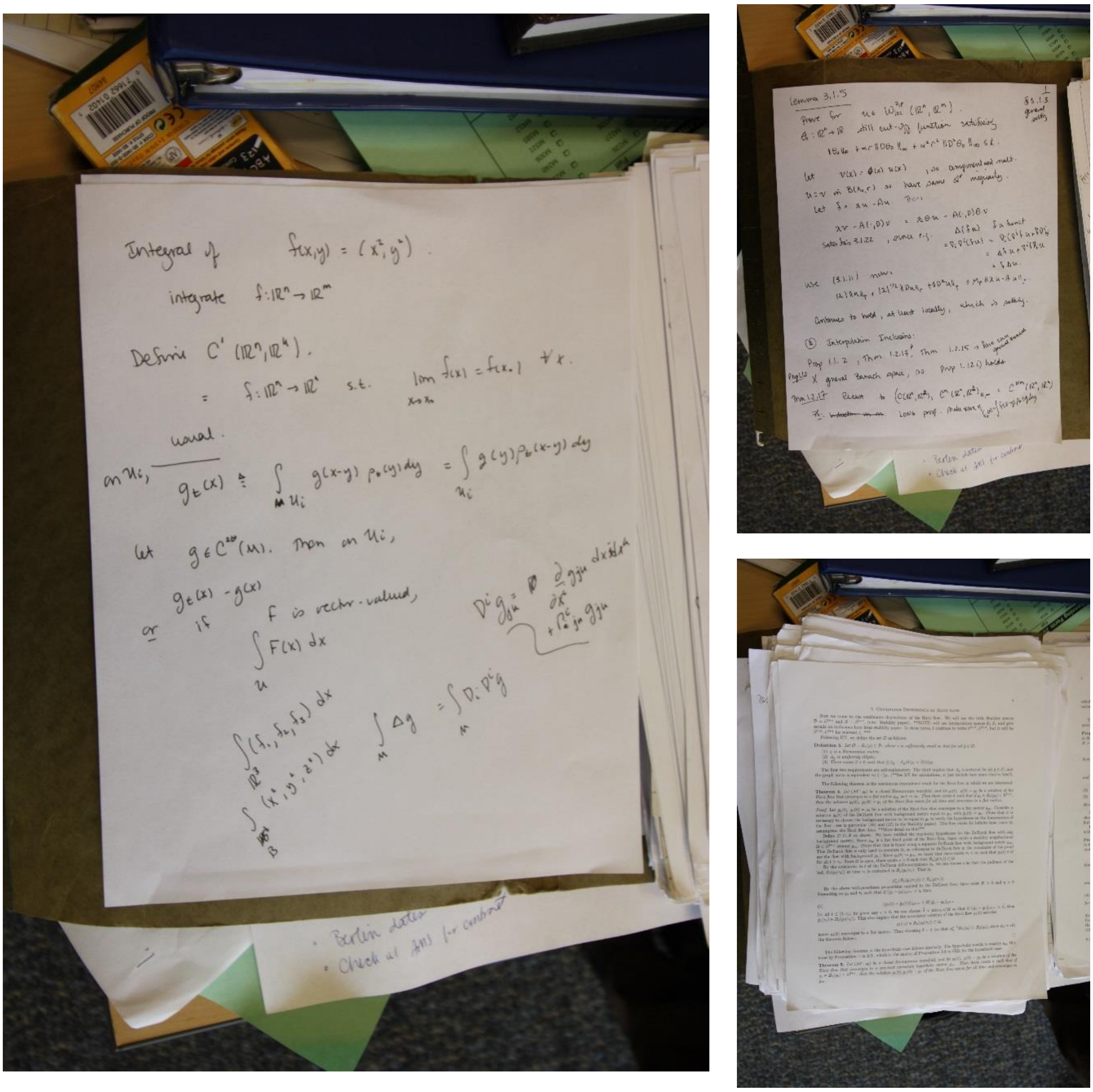


Figure 60. M described having three basic tiers: rough notes, neater notes (generally with a date and title) and typed-up notes which go to make up a written version. Original in colour.

Work site 5

Work site 5 was the desk space used by a PhD student in a large shared office at a UK university.

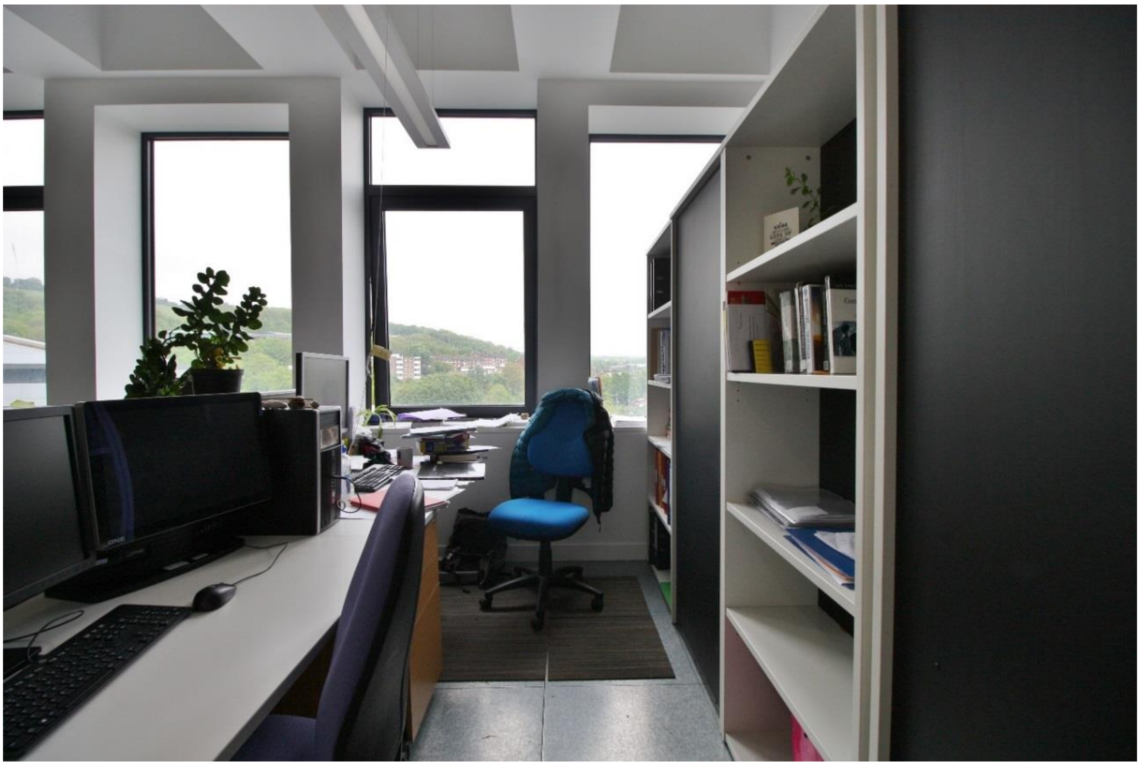


Figure 61. Desk space used by my participant, with bookshelves and a computer. Original in colour.

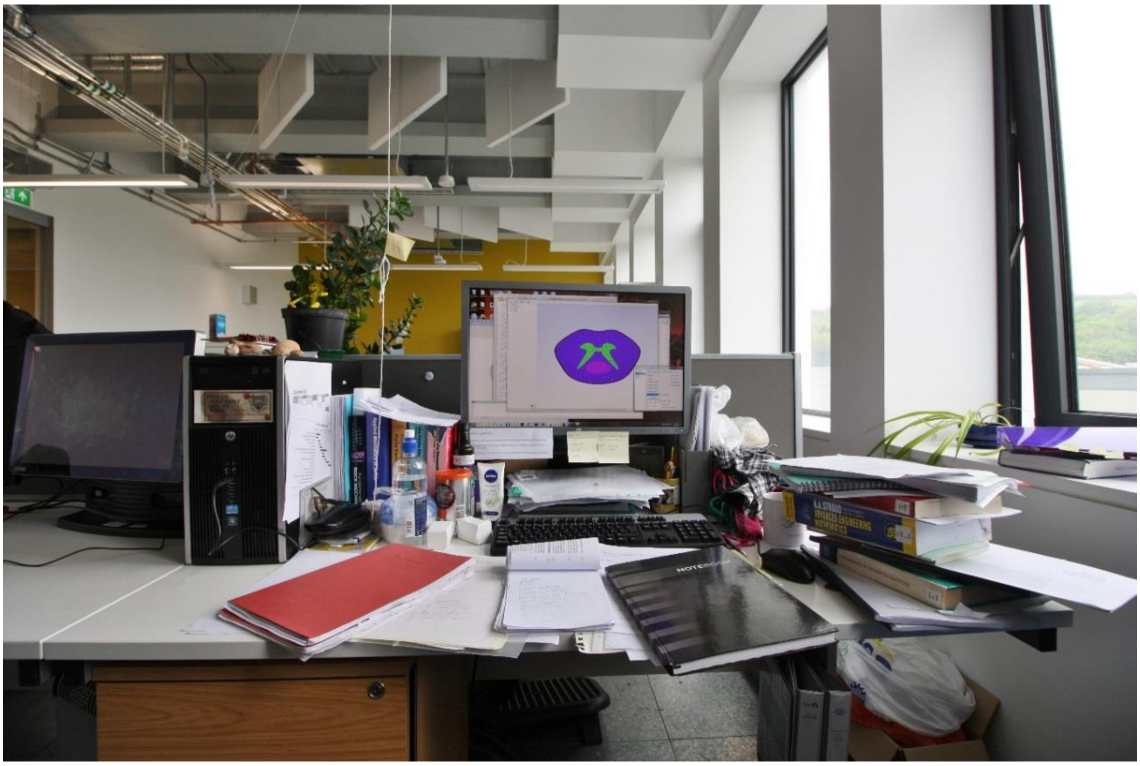


Figure 62. My participant's work space, with rough notes, notebooks and reference books. Original in colour.

5. Analysis of an excerpt from a mathematician's notes, leading to a section in a paper

The third excerpt is taken from a participant's notes and the section of a paper that they led to.

5.1. Record of a first time through

The excerpt analysed (see

5.3.1. Transcripts for the full transcript) is taken from notes in which a mathematician is working on a tablet, laying out a proof intended to be used in a paper. This particular participant reported that the really early ideas stage tended to happen on sheets of paper, and that tablet notes like these ones tended to come at a stage when the ideas were more clearly taking shape, making this an intermediate stage in the process of developing the mathematics. In the two pages examined, the participant lays out what appears to be a procedure and a proof by induction, which some examination of the paper that this work eventually led to allowed me to link with a procedure and proof laid out over the course of four pages. The participant is known by the letter I.

5.1.1. Carrying out the analysis

The observed material in this case was the written trace of a situation in which a mathematician sat with stylus and tablet, a configuration of thinker and inscription in search of clarification of thought. The mathematician may have been bowed over a desk in an office or café, or both; the inscriptions are portable, can serve as a constant in a thought process that might last several days. The mathematician may have been working from earlier notes, scraps of paper, using these as prompts or reminders, smoothing out the ideas and diagrams by the transition to digital marks.

The obvious route to follow was to dig in to this material by making notes in text and pictures. This was confirmed by a conversation I had with my mathematician supervisor at an early stage in my analysis, in which it was repeatedly suggested that I try to 'understand by drawing a picture', sketching experimental diagrams for myself until I felt I could see what was going on (see Figure 63). Inhabiting just one character, X, on this occasion (as appropriate to the solo nature of the notes being analysed), I took the pages and surrounded them with notes and sketches.

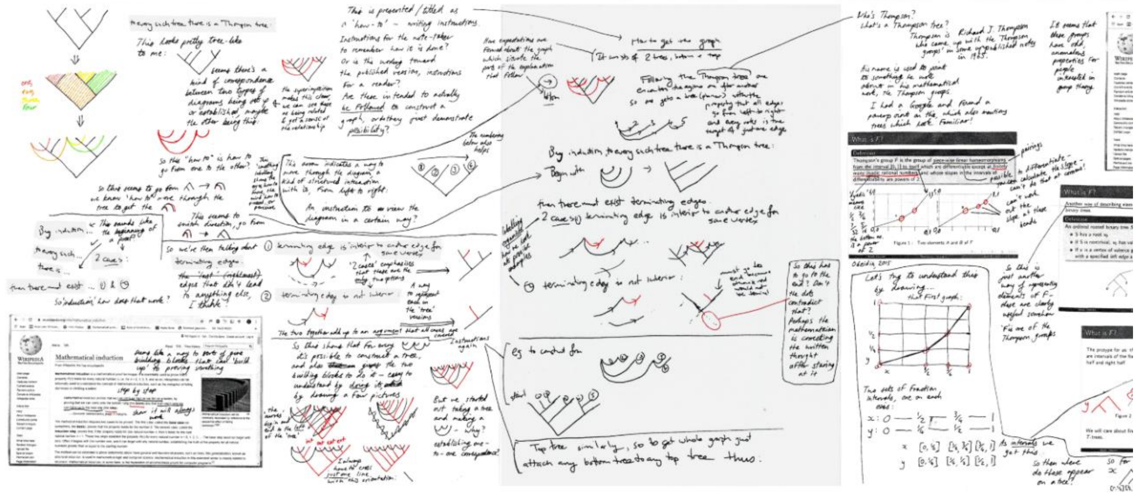


Figure 63. Notes. Original in colour.

I found that I understood the most by drawing, by working through examples myself and beginning to perceive the features of the situation in helpful ways. By drawing out a graph myself and coming across the bends or non-bends in the line I could put together a working understanding of how the graph, a tree diagram and a set of intervals might relate, focus on the relevant features, feel that by working through the example I must indeed have demonstrated that I understood it.

Approaching this excerpt as a ‘first time through’ posed new challenges. The participant used a variety of drawing techniques—colour, labelling, arrows—to indicate and bring out an ordered engagement with a set of diagrams, this in itself truly bringing order into being the first time through. On the other hand, another portion of the work I did to understand the mathematical content involved looking up certain terms that were offhandedly scribbled in the course of the notes, and making sense of these involved researching concepts as they are discussed in mathematical literature and taught to students. I had the sense that in the notes I was analysing, the mathematician was producing order anew in the structured engagement with the diagram in a very observable way; another aspect of the orderliness being produced was in the links being drawn between this structure and existing agreed structures in the mathematical community, something that was minimally referenced in the notes but clearly an integral part of the work being done. This latter in the notes was embodied by the experienced engagement of the mathematician, of being able to draw parallels from knowledge and perceive the significance and potential usefulness of the work being done from knowledge of the body of mathematical knowledge.

The mathematician was delineating a new way of representing the Thompson groups, three unusual infinite groups with properties that make them interesting to group theorists, as sets of a particular type of graph seen in knot theory. This new representation presents a way of making Thompson group elements from links (types of mathematical knot made up of more than one linked strand), and vice versa. This can give mathematicians new ways of finding links that are not equivalent to one another.

This contextualisation, largely implicit in the notes, is made much more explicit in the paper. As well as this, I found considerably more explicitness in the procedure described (see Figure 64) which I suspect is another effort to connect the present work with the body of mathematical knowledge.

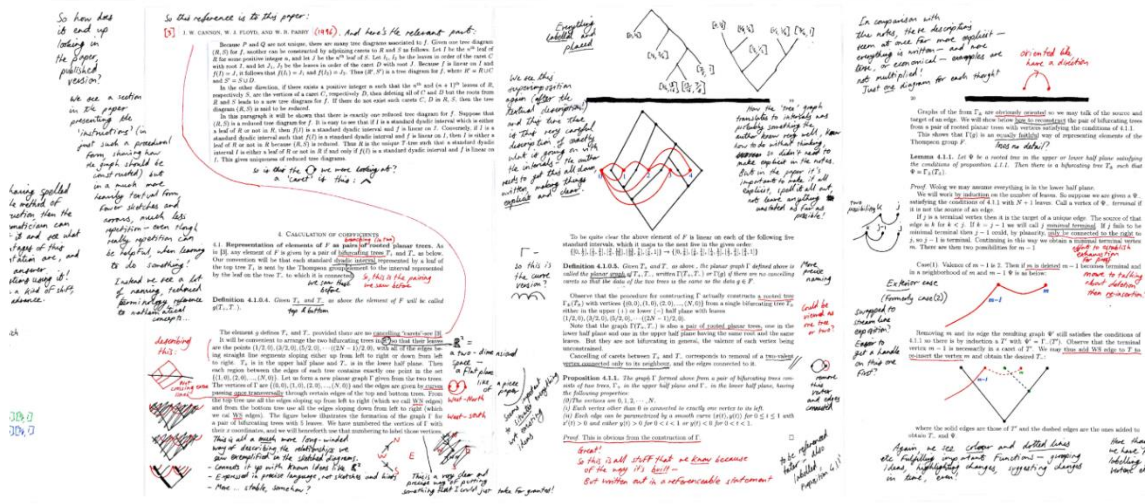


Figure 64. Notes. Original in colour.

As I came to be more familiar with concepts like the Thompson groups I noted that it was most crucial for me to feel that I had some kind of operational knowledge of them, like of how to build a member of the group. I also became comfortable with aspects that I did not know and lacked the resources to full explore by understanding some portion of the way that insiders would use them, their status as counter-examples seeming key.

5.2. Breaching experiment: bad diagramming

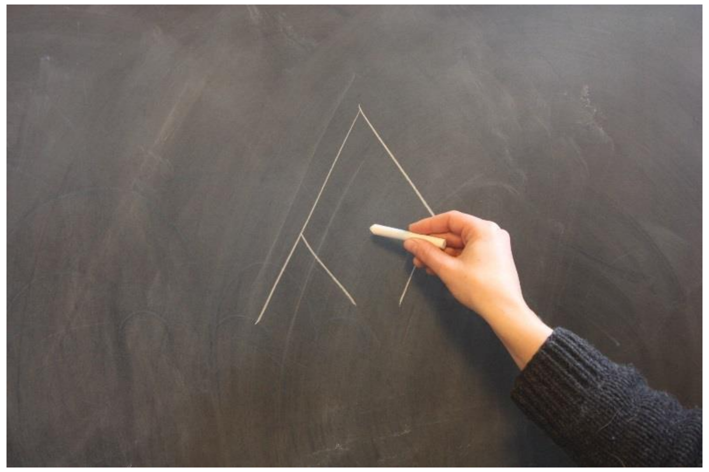
This breaching experiment tested the role and rules of drawing. While generally we would expect the form of the marks made to be the focal point, the chief vehicle for whatever is going on, in this experiment the drawn shapes were predetermined; I worked with a set of 64 possible subsets of lines from a basic tree shape similar to that in my observed material, manipulating just the order in which the subsets appeared.

I approached drawing the set of 64 with a different system each time to come to the full set, photographing each drawing as I went. It was a struggle to get the full set each time, to come up with a system that left me with no omissions, to iron out any blind spots.

This mode of drawing had me treat the mark-making in a systematic way that was even so operating according to a different logic than would be usual, the trees being constructed not to refer to something else but for completeness in themselves.

Looking back over the image sets I enjoy being able to 'read' the system of each one, to follow the sequence and recognise how I was trying to structure my drawing to catch all of the possible subsets. This kind of drawing, with a structure held in mind, produces a distracted kind of look to the drawings that is very characteristic of mathematical drawing, the quality of a line that is drawn when the mind is already elsewhere.

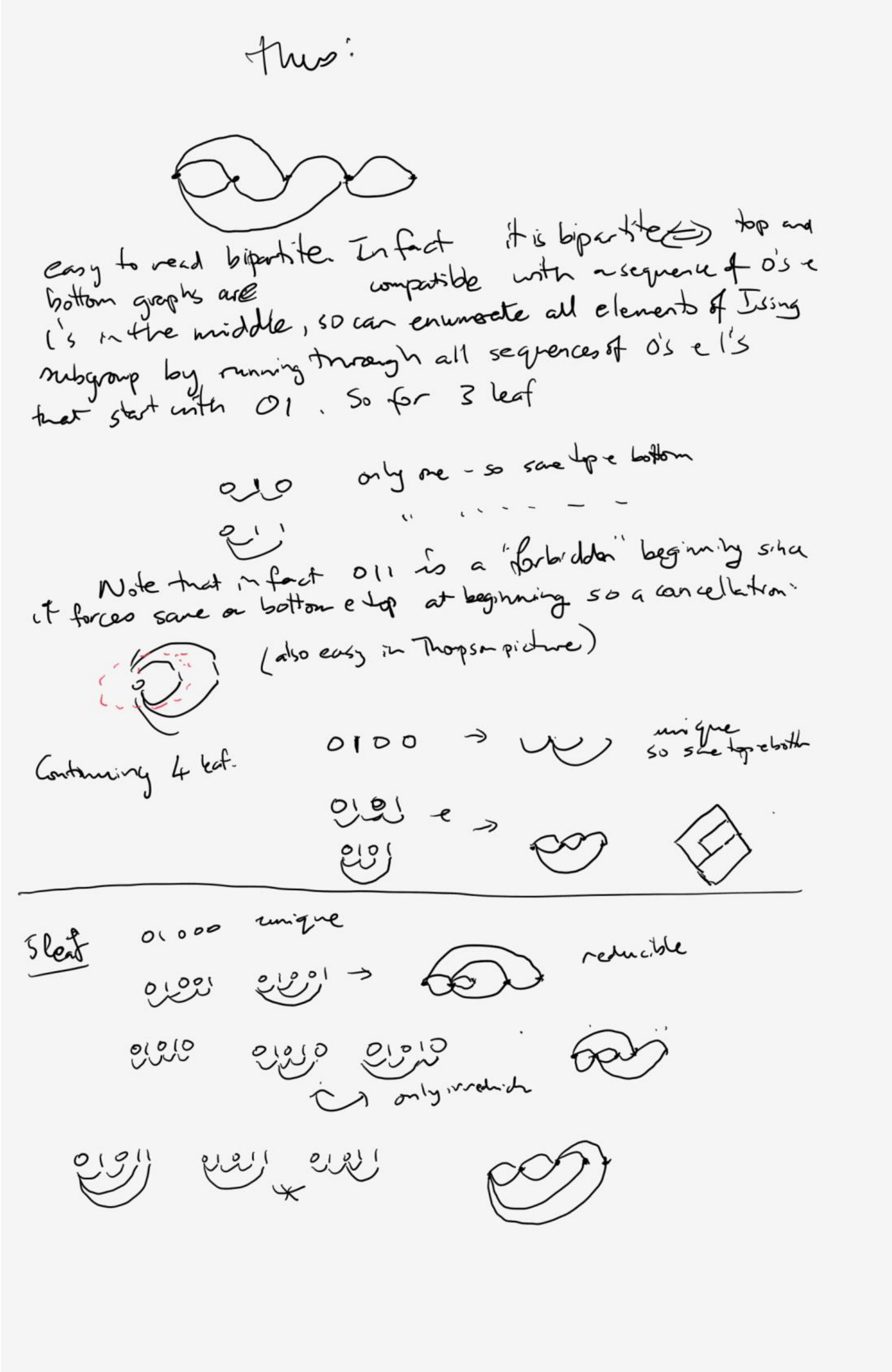
In one sense, the human agency in these sets of photographs slips outside the frame, existing in the spaces between the images. The reasoning mind can only be seen in the sequence, whereas the 64 trees in the images simply are what they are. But, the distracted, dusty details of the chalkboard drawings show something important in the human experience of mathematical drawing, of the confused relationship between hands and mind.



Original in colour.

5.3. Evidence

5.3.1. Transcripts



thus:

easy to read bipartite. In fact it is bipartite ⟺ top and bottom graphs are compatible with a sequence of 0's & 1's in the middle, so can enumerate all elements of Ising subgroup by running through all sequences of 0's & 1's that start with 01. So for 3 leaf

010 only one - so same top & bottom

011 " " " " " — "

Note that in fact 011 is a "forbidden" beginning since it forces same on bottom & top at beginning so a cancellation:

(also easy in Thompson picture)

Continuing 4 leaf. 0100 → unique so same top & both

0101 & →

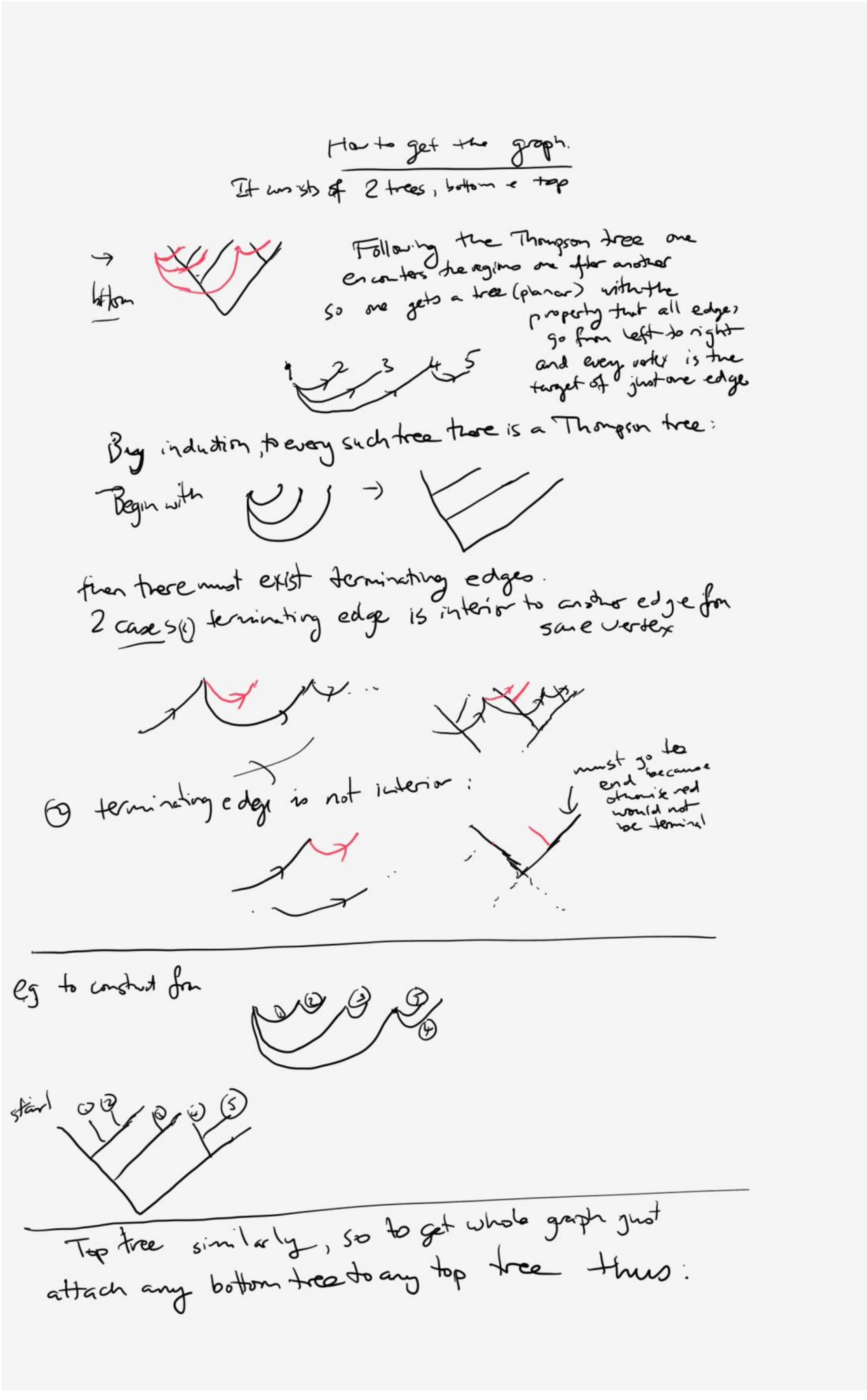
0101

5 leaf 01000 unique

reducible

only irreducible

Original in colour



How to get the graph.

It consists of 2 trees, bottom & top

→ bottom

Following the Thompson tree one encounters the regions one after another so one gets a tree (planar) with the property that all edges go from left to right and every vertex is the target of just one edge

1 2 3 4 5

By induction, to every such tree there is a Thompson tree:

Begin with →

then there must exist terminating edges.

2 cases (i) terminating edge is interior to another edge from same vertex

(ii) terminating edge is not interior:

must go to end because otherwise red would not be terminal

eg to construct from

1 2 3 5 4

start 1 2 3 4 5

Top tree similarly, so to get whole graph just attach any bottom tree to any top tree thus:

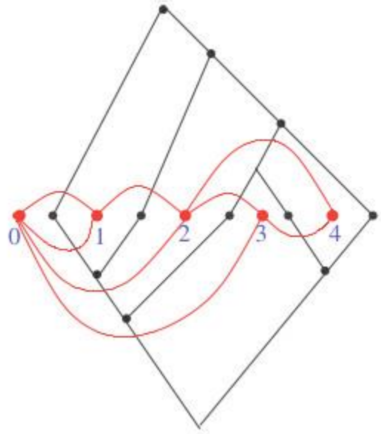
Original in colour

4. Calculation of coefficients

4.1. **Representation of elements of F as pairs of rooted planar trees.** As in [3], any element of F is given by a pair of bifurcating trees T_+ and T_- as below. Our convention will be that each standard dyadic interval represented by a leaf of the top tree T_+ is sent by the Thompson group element to the interval represented by the leaf on the tree T_- to which it is connected.

Definition 4.1.0.4. *Given T_+ and T_- as above the element of F will be called $g(T_+, T_-)$.*

The element g defines T_+ and T_- provided there are no cancelling "carets"-see [3].
It will be convenient to arrange the two bifurcating trees in $\mathbb{R}^2$ so that their leaves are the points $(1/2, 0), (3/2, 0), (5/2, 0), \cdots ((2N-1)/2, 0)$, with all of the edges being straight line segments sloping either up from left to right or down from left to right. T_+ is in the upper half plane and T_- is in the lower half plane. Then each region between the edges of each tree contains exactly one point in the set $\{(1,0), (2,0), ..., (N,0)\}$. Let us form a new planar graph Γ given from the two trees. The vertices of Γ are $\{(0,0), (1,0), (2,0), ..., (N,0)\}$ and the edges are given by curves passing once transversally through certain edges of the top and bottom trees. From the top tree use all the edges sloping up from left to right (which we call WN edges) and from the bottom tree use all the edges sloping down from left to right (which we cal WS edges). The figure below illustrates the formation of the graph Γ for a pair of bifurcating trees with 5 leaves. We have numbered the vertices of Γ with their x coordinates, and we will henceforth use that numbering to label those vertices.



To be quite clear the above element of F is linear on each of the following five standard intervals, which it maps to the next five in the given order:
$\{[0,\frac{1}{2}],[\frac{1}{2},\frac{3}{4}],[\frac{3}{4},\frac{13}{16}],[\frac{13}{16},\frac{7}{8}],[\frac{7}{8},1]\} \to \{[0,\frac{1}{8}],[\frac{1}{8},\frac{1}{4}],[\frac{1}{4},\frac{1}{2}],[\frac{1}{2},\frac{3}{4}],[\frac{3}{4},1]\}$

Definition 4.1.0.5. *Given T_+ and T_- as above , the planar graph Γ defined above is called the planar graph of T_+, T_-, written $\Gamma(T_+, T_-)$ or $\Gamma(g)$ if there are no cancelling carets so that the data of the two trees is the same as the data $g \in F$.*

Observe that the procedure for constructing Γ actually constructs a rooted tree $\Gamma_\pm(T_\pm)$ with vertices $\{(0,0),(1,0),(2,0),...,(N,0)\}$ from a single bifurcating tree $T_\pm$ either in the upper $(+)$ or lower $(-)$ half plane with leaves
$(1/2,0),(3/2,0),(5/2,0),\cdots((2N-1)/2,0)$.

Note that the graph $\Gamma(T_+,T_-)$ is also a pair of rooted planar trees, one in the lower half plane and one in the upper half plane having the same root and the same leaves. But they are not bifurcating in general, the valence of each vertex being unconstrained.

Cancelling of carets between T_+ and T_- corresponds to removal of a two-valent vertex connected only to its neighbour, and the edges connected to it.

Proposition 4.1.1. *The graph Γ formed above from a pair of bifurcating trees consists of two trees, Γ_+ in the upper half plane and Γ_- in the lower half plane, having the following properties:*
(0)The vertices are $0,1,2,\cdots,N$.
(i) Each vertex other than 0 is connected to exactly one vertex to its left.
(ii) Each edge can be parametrized by a smooth curve $(x(t),y(t))$ for $0 \le t \le 1$ with $x'(t) > 0$ and either $y(t) > 0$ for $0 < t < 1$ or $y(t) < 0$ for $0 < t < 1$.

Proof. This is obvious from the construction of Γ. □

Graphs of the from $\Gamma_\pm$ are obviously oriented so we may talk of the source and target of an edge. We will show below how to reconstruct the pair of bifurcating trees from a pair of rooted planar trees with vertices satisfying the conditions of 4.1.1.

This shows that $\Gamma(g)$ is an equally faithful way of representing elements of the Thompson group F.

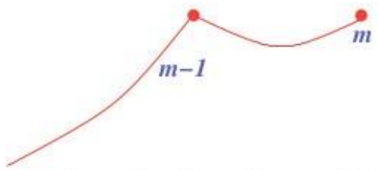
Lemma 4.1.1. *Let Ψ be a rooted tree in the upper or lower half plane satisfying the conditions of proposition 4.1.1. Then there is a bifurcating tree $T_\pm$ such that $\Psi = \Gamma_\pm(T_\pm)$.*

Proof. Wolog we may assume everything is in the lower half plane.

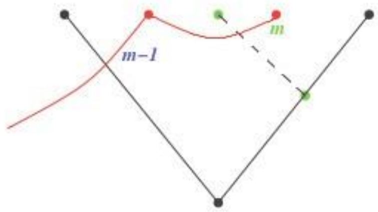
We will work by induction on the number of leaves. So suppose we are given a Ψ_- satisfying the conditions of 4.1.1 with $N+1$ leaves. Call a vertex of Ψ_- *terminal* if it is not the source of an edge.

If j is a terminal vertex then it is the target of a unique edge. The source of that edge is k for $k < j$. If $k = j-1$ we will call j *minimal terminal*. If j fails to be minimal terminal then $j-1$ could, by planarity, only be connected to the right to j, so $j-1$ is terminal. Continuing in this way we obtain a minimal terminal vertex m. There are then two possibilities for $m-1$

Case(1). Valence of $m-1$ is 2. Then if m is deleted $m-1$ becomes terminal and in a neighborhood of m and $m-1$ Ψ is as below:

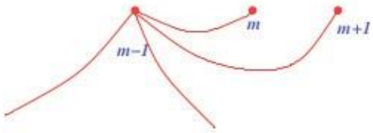


Removing m and its edge the resulting graph Ψ' still satisfies the conditions of 4.1.1 so there is by induction a T' with $\Psi' = \Gamma_-(T')$. Observe that the terminal vertex $m-1$ is necessarily in a caret of T'. We may thus add WS edge to T to re-insert the vertex m and obtain the desired T_-:

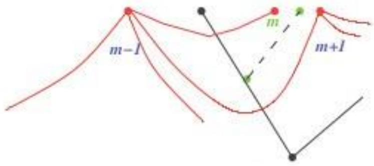


where the solid edges are those of T' and the dashed edges are the ones added to obtain T_- and Ψ.

Case(2) Valence of $m-1$ is >2. In this case there is an edge with source $m-1$ connecting it to a vertex k with $k>m$. By planarity there must be such an edge connecting $m-1$ to $m+1$. The situation near m is thus:



Removing m and its edge the resulting graph Ψ' still satisfies the conditions of 4.1.1 so there is by induction a T' with $\Psi' = \Gamma_-(T')$. There has to be a WS edge in T between $m-1$ and $m+1$ so we may add a WN edge to T' as in the figure below re-insert m and obtain the desired T_-:



where the solid edges are those of T' and the dashed edges are the ones added to obtain T_- and Ψ. □

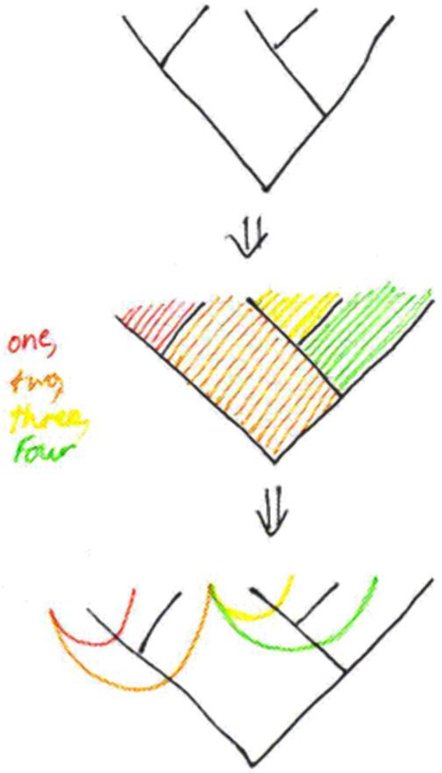
We have essentially given a bijection between two sets counted by the Catalan numbers which thus almost certainly exists in the literature. The result is so important to our examples that we have supplied a detailed description of what is an algorithm to obtain a Thompson group element from a pair of planar graphs with the same set of leaves and the same root.

Original in colour

5.3.2. Notes

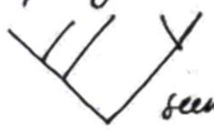
This analysis is carried out in the form of an extended continuous sheet of notes, working from left to right. In the interests of readability it is here divided into six smaller A3 sections on the following pages; a full scrollable pdf version is to be found on the accompanying CD.

Original in colour

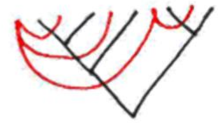


to every such tree there is a Thompson tree:

This looks pretty tree-like to me:



seems there's a kind of correspondence between two types of diagrams being set up or established, maybe the other being this:



the superimposition makes this clear, we can see these as being related & get a sense of the relationship

So this "how to" is how to go from one to the other?

So this seems to go from ⋀ → ⋂ we know 'how to' move through the tree to get the ⋂

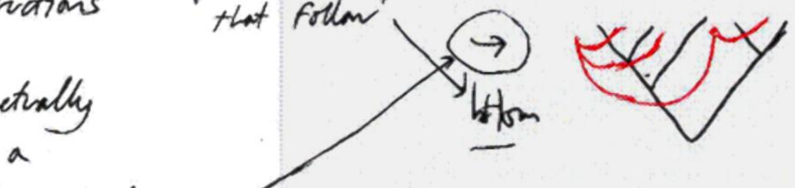
This 'smoothing' labelling shows the eyes how to move, the mind how to proceed, or perceive

This seems to switch direction, go from ⋂ → ⋀

This is presented / titled as a 'how-to' – writing instructions.
Instructions for the note-taker to remember how it is done?
Or is this working toward the published version, instructions for a reader?
Are these intended to actually be followed to construct a graph, or do they just demonstrate possibility?

Here expectations are formed about the graph which situate the parts of the explanation that follow

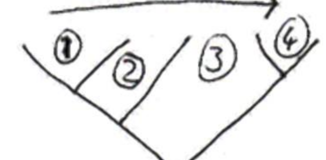
How to get the ...
It consists of 2 trees, bottom ...



Following the ... encounters the ... so one gets a tree ...

The numbering below also helps

This arrow indicates a way to move through the diagram, a kind of structured interaction with it, from left to right:



An instruction to view the diagram in a certain way?

By induction, to every such tree there ...
Begin with ... → ...
then there must exist terminating edge ...
2 cases (i) terminating edge is interior ...

labelling organises how we look at possible examples

(ii) terminating edge is not interior:

By induction... ← This sounds like the beginning of a proof!

to every such ... there is ...

2 cases:

then there must exist ... (i) & (ii)

So 'induction', how does that work?

So we're then talking about terminating edges.

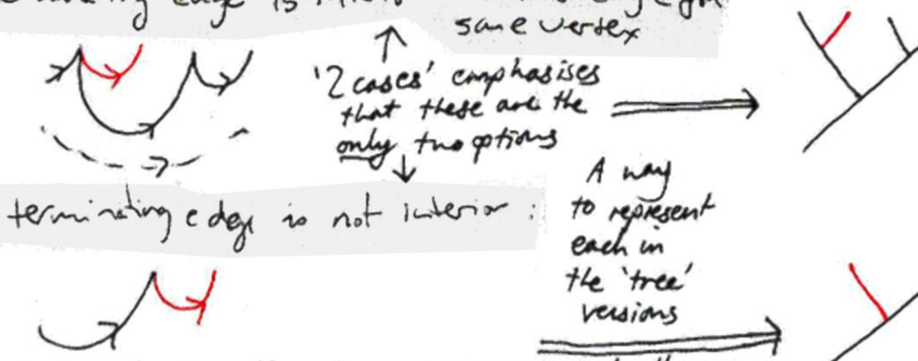
~~the 'last' (rightmost)~~ edges that don't lead to anything else, I think?

(1) terminating edge is interior to another edge from same vertex

'2 cases' emphasises that these are the only two options

(2) terminating edge is not interior:

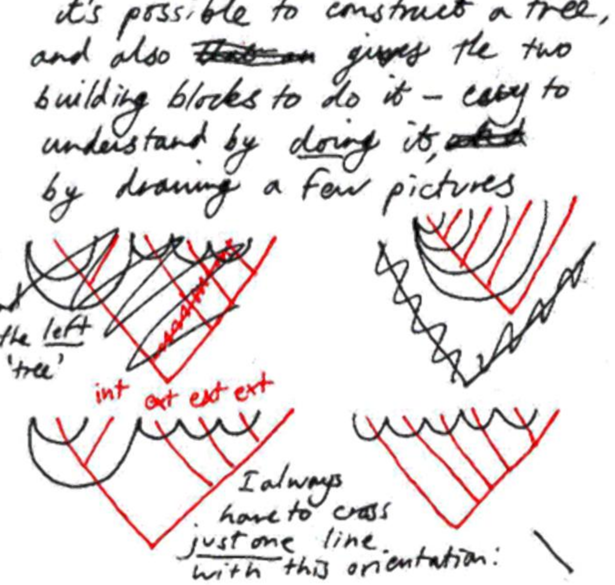
A way to represent each in the 'tree' versions



The two together add up to an argument that all cases are covered

So this shows that for every ⋃⋃⋃ it's possible to construct a tree, and also ~~that an~~ gives the two building blocks to do it – easy to understand by doing it, ~~and~~ by drawing a few pictures

Instructions again



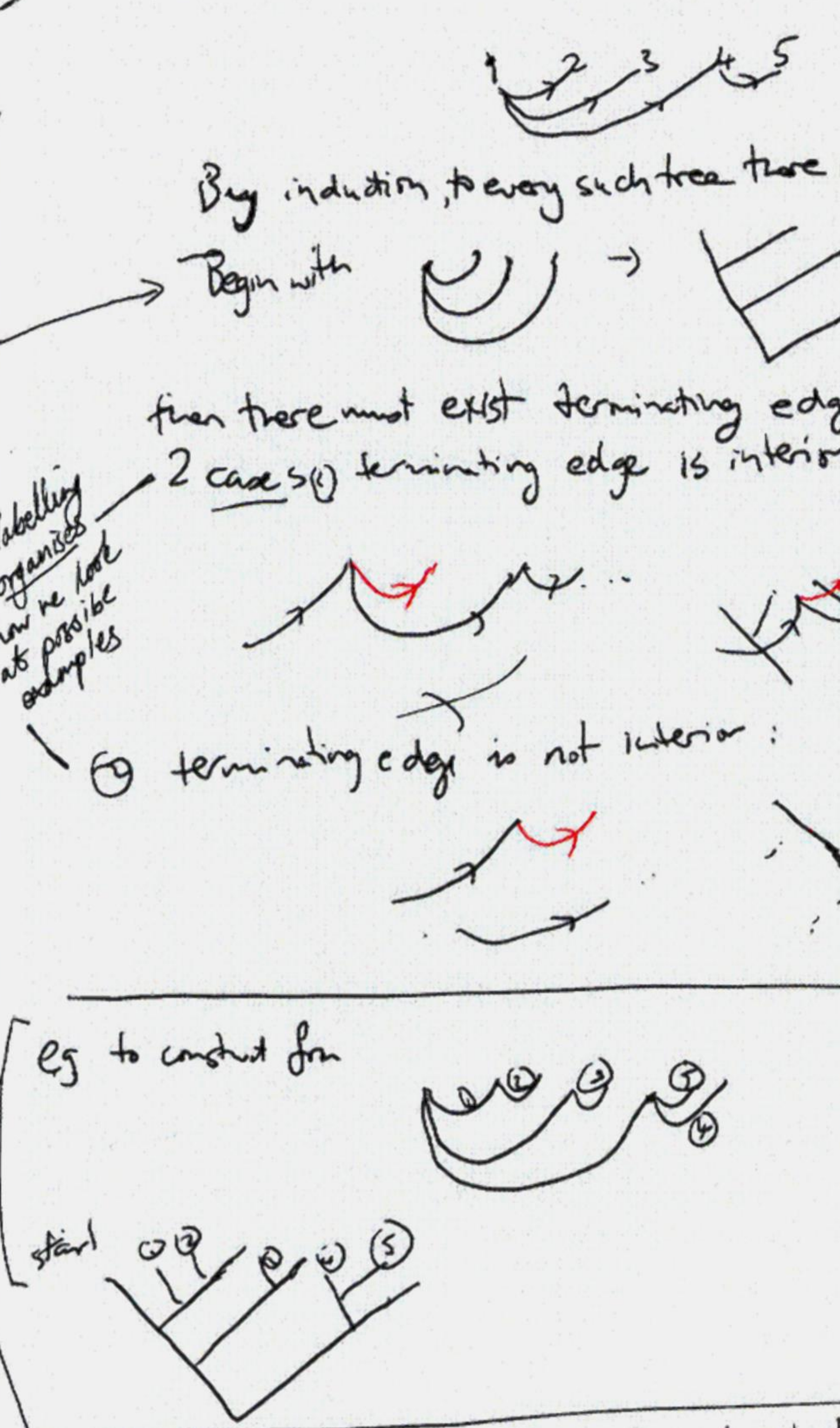
'the curves' begin and end to the left of the 'tree'

I always have to cross just one line with this orientation:

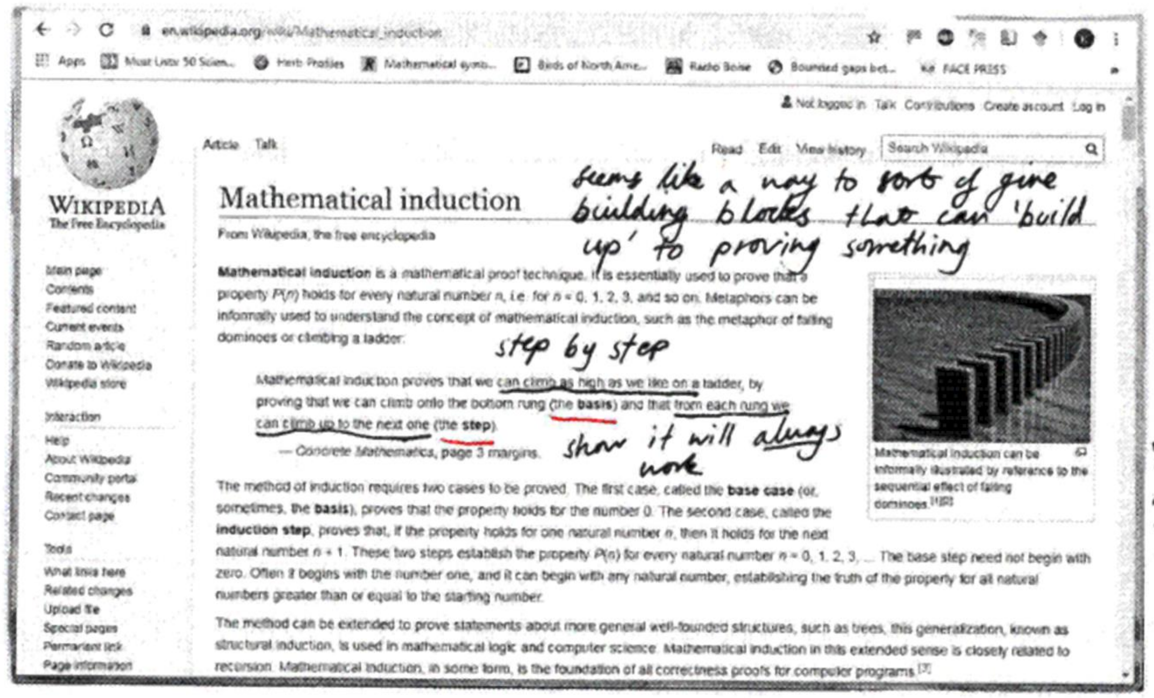
But we started out taking a tree and making a ⋃⋃ – why? establishing one-to-one correspondence?

eg to construct from ...

start



Top tree similarly, so to get whol...
attach any bottom tree to any top tr...


en.wikipedia.org/wiki/Mathematical_induction

Mathematical induction

From Wikipedia, the free encyclopedia

Mathematical induction is a mathematical proof technique. It is essentially used to prove that a property P(n) holds for every natural number n, i.e. for n = 0, 1, 2, 3, and so on. Metaphors can be informally used to understand the concept of mathematical induction, such as the metaphor of falling dominoes or climbing a ladder:

> Mathematical induction proves that we can climb as high as we like on a ladder, by proving that we can climb onto the bottom rung (the **basis**) and that from each rung we can climb up to the next one (the **step**).
> — Concrete Mathematics, page 3 margins.

The method of induction requires two cases to be proved. The first case, called the **base case** (or, sometimes, the **basis**), proves that the property holds for the number 0. The second case, called the **induction step**, proves that, if the property holds for one natural number n, then it holds for the next natural number n + 1. These two steps establish the property P(n) for every natural number n = 0, 1, 2, 3, ... The base step need not begin with zero. Often it begins with the number one, and it can begin with any natural number, establishing the truth of the property for all natural numbers greater than or equal to the starting number.

The method can be extended to prove statements about more general well-founded structures, such as trees; this generalization, known as structural induction, is used in mathematical logic and computer science. Mathematical induction in this extended sense is closely related to recursion. Mathematical induction, in some form, is the foundation of all correctness proofs for computer programs.

Mathematical induction can be informally illustrated by reference to the sequential effect of falling dominoes.

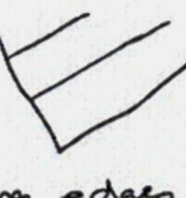
...et the graph.

..., bottom & top

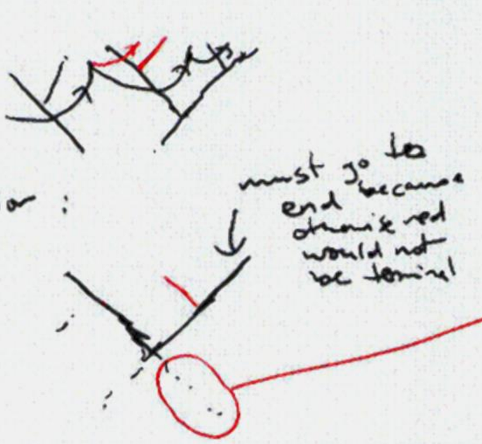
...g the Thompson tree' one ...the edges one after another ...to a tree (planar) with the property that all edges go from left to right and every vertex is the target of just one edge

...4 → 5

...e there is a Thompson tree:



...g edges. ...interior to another edge for same vertex



...et whole graph just ...top tree thus:

Who's Thompson?
What's a Thompson tree?

Thompson is Richard J. Thompson who came up with the 'Thompson groups' in some unpublished notes in 1965.

His name is used to point to something he wrote about in his mathematical work, the Thompson groups.

I had a Google and found a powerpoint on this, which also mentions trees which look familiar!

It seems that these groups have odd, anomalous properties for people interested in group theory

← → C en.wikipedia.org/wiki/Thompson_groups

Thompson groups

From Wikipedia, the free encyclopedia

This article is about the three unusual infinite groups F, T and V found by Thompson. For the sporadic simple group, see Thompson sporadic group.

In mathematics, the **Thompson groups** (also called **Thompson's groups**, **vagabond groups** or **chameleon groups**) are three groups, commonly denoted $F \subseteq T \subseteq V$, which were introduced by Richard Thompson in some unpublished handwritten notes in 1965 as a possible counterexample to the von Neumann conjecture. Of the three, *F* is the most widely studied, and is sometimes referred to as **the Thompson group** or **Thompson's group**.

The Thompson groups, and *F* in particular, have a collection of unusual properties which have made them counterexamples to many general conjectures in group theory. All three Thompson groups are infinite but finitely presented. The groups *T* and *V* are (rare) examples of infinite but finitely-presented simple groups. The group *F* is not simple but its derived subgroup [F,F] is and the quotient of *F* by its derived subgroup is the free abelian group of rank 2. *F* is totally ordered, has exponential growth, and does not contain a subgroup isomorphic to the free group of rank 2.

It is conjectured that *F* is not amenable and hence a further counterexample to the long-standing but recently disproved von Neumann conjecture for finitely-presented groups. It is known that *F* is not elementary amenable.

Higman (1974) introduced an infinite family of finitely presented simple groups, including Thompson's group *V* as a special case.

Contents [hide]
1 Presentations
2 Other representations

counter-examples can be very exciting in mathematics – can have large implications just by existing!

great names

(when viewed by an appropriately informed mind to draw out what it 'means'...)

What is F?

Definition

Thompson's group *F* is the group of piece-wise linear homeomorphisms from the interval [0, 1] to itself which are differentiable except at finitely many dyadic rational numbers and whose slopes in the intervals of differentiability are powers of 2.

pairings

possible

to differentiate – you can calculate the slope – can't do that at corners!

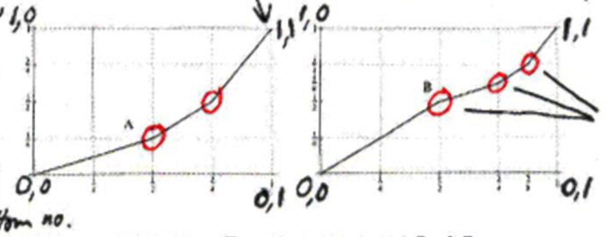


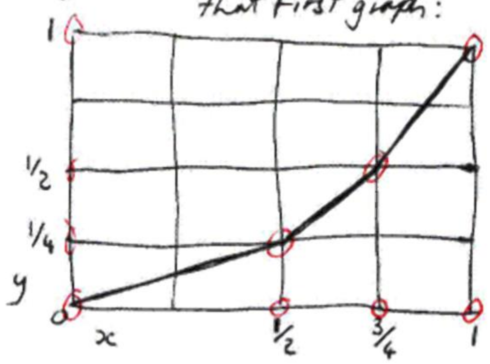
Figure 1: Two elements *A* and *B* of *F*

'dyadic' means like 1/2 3/4 7/32 5/16 – the bottom no. is a power of 2

can't work out the slope at these bends

Obeidin, 2015

Let's try to understand this by drawing... that first graph:



So this has to go to the end? Don't the dots contradict that? Perhaps the mathematician is correcting the written thought after staring at it

Two sets of fraction intervals, one on each axis:

x : 0 — 1/2 — 3/4 — 1
y : 0 — 1/4 — 1/2 — 1

As intervals we get this

x [0, 1/2] [1/2, 3/4] [3/4, 1]
y [0, 1/4] [1/4, 1/2] [1/2, 1]

So this is just another way of representing elements of F – these are clearly useful somehow

'F' is one of the Thompson groups

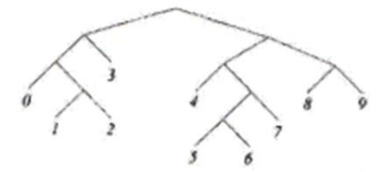
What is F?

Another way of describing elements of *F* is as pairings of *rooted, ordered binary trees.*

Definition

An ordered rooted binary tree *S* is a tree satisfying:

- *S* has a root v_0
- If *S* is nontrivial, v_0 has valence 2
- If *v* is a vertex of valence greater than 1, *v* has exactly two children, with a specified *left edge* and *right edge*



What is F?

The protype for us: the tree of *standard dyadic intervals*, whose vertices are intervals of the form $[\frac{a}{2^n}, \frac{a+1}{2^n}]$, where each vertex's children are its left half and right half

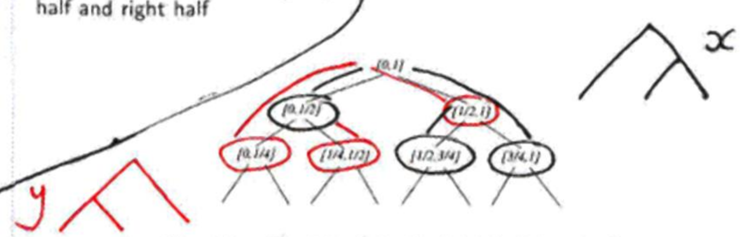


Figure 2: The tree of standard dyadic intervals, *T*

We will care about finite ordered subtrees of *T*, with the same root, called *T*-trees.

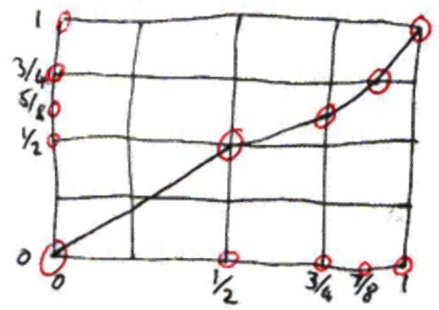
subsets – trees made up of some of these branches

So then where do these appear on a tree?

So for x

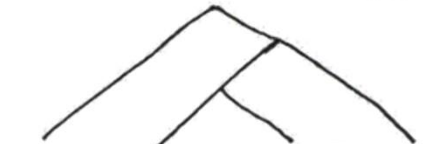
For y

Let's try the second example!



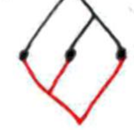
x : [0, 1/2] [1/2, 3/4] [3/4, 7/8] [7/8, 1]

y : [0, 1/2] [1/2, 5/8] [5/8, 3/4] [3/4, 1]



So this is a representation of an element of F! just as the graph we started with!

The two combined:



athematical symb... Birds of North Ame... Radio Boise Bounded gaps bet... FACE PRESS

Not logged in Talk Contributions Create account Log in

Read Edit View history Search Wikipedia

oups

opedia

ree unusual infinite groups F, T and V found by Thompson. For the sporadic simple group, see

on groups (also called Thompson's groups, vagabond groups or chameleon groups) are three $F \subseteq T \subseteq V$, which were introduced by Richard Thompson in some unpublished handwritten notes in ample to the von Neumann conjecture. Of the three, F is the most widely studied, and is sometimes group or Thompson's group.

in particular, have a collection of unusual properties which have made them counterexamples to many theory. All three Thompson groups are infinite but finitely presented. The groups T and V are (rare) y-presented simple groups. The group F is not simple but its derived subgroup [F,F] is and the quotient s the free abelian group of rank 2. F is totally ordered, has exponential growth, and does not contain a ee group of rank 2.

amenable and hence a further counterexample to the long-standing but recently disproved von y-presented groups: it is known that F is not elementary amenable.

infinite family of finitely presented simple groups, including Thompson's group V as a special case.

counter-examples can be very exciting in mathematics - can have large implications just by existing!

great names

(when viewed by an appropriately informed mind, to draw out what it 'means'...)

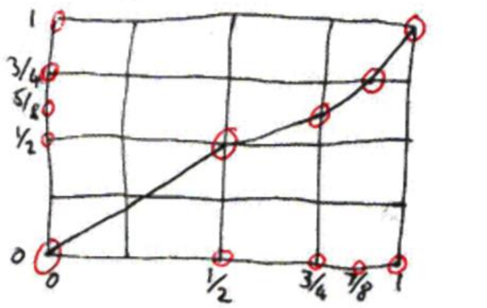
So according to all of this, ~~the~~ ~~procedure~~ building yourself a subtree and pairing it with another subtree with give you a member of F - which must be infinite, I think, if you can use any $\frac{a}{2^n}$?

So here the mathematician joins the top & bottom parts together, continuing from the previous instructions

then starts discussing some things that can be done with this new representation - noting what can be easily worked out or seen

ted, ordered

Let's try the second example!



And then the collection of these things have interesting properties - so how can we know that? Knowing how they're constructed?

I'll try differentiating:

change in x / change in y

$\frac{1}{2} / \frac{1}{2} = 1 = 2^0$

$\frac{1}{4} / \frac{1}{8} = 2 = 2^1$

$\frac{1}{8} / \frac{1}{8} = 1 = 2^0$

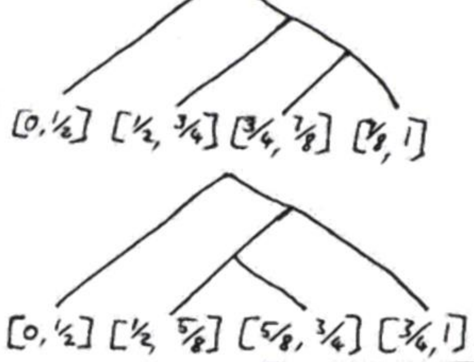
$\frac{1}{8} / \frac{1}{4} = \frac{1}{2} = 2^{-1}$

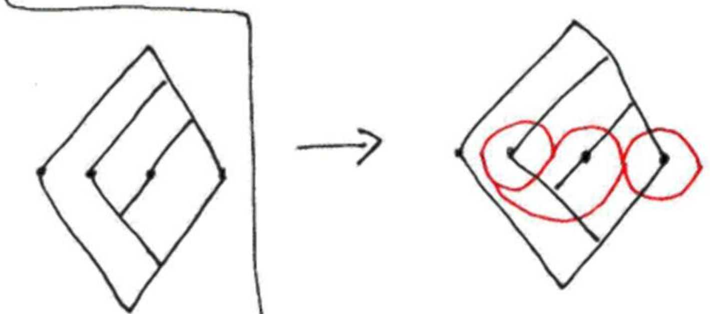
then later on the writing seems almost to be being itself used to do calculations (on top of recording ideas), to answer questions

wo children.

x: $[0, \frac{1}{2}]$ $[\frac{1}{2}, \frac{3}{4}]$ $[\frac{3}{4}, \frac{7}{8}]$ $[\frac{7}{8}, 1]$

y: $[0, \frac{1}{2}]$ $[\frac{1}{2}, \frac{5}{8}]$ $[\frac{5}{8}, \frac{3}{4}]$ $[\frac{3}{4}, 1]$





rvals, whose vertices
x's children are its left

tervals, T

h the same root, called
made up of
of these branches

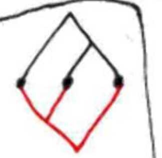
So this is a representation of an element of F! just as the graph we started with!

The page ends with this:

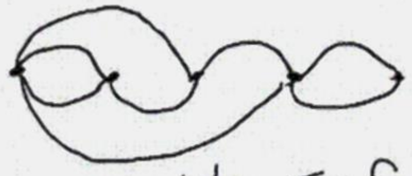
Top tree similarly, so to get whole graph just attach any bottom tree to any top tree thus:

So the next page continues as follows:

The two combined:



thus:



easy to read bipartite. In fact it is bipartite ⟺ top and bottom graphs are compatible with a sequence of 0's & 1's in the middle, so can enumerate all elements of Ising subgroup by running through all sequences of 0's & 1's that start with 01. So for 3 leaf

010 only one - so same top & bottom

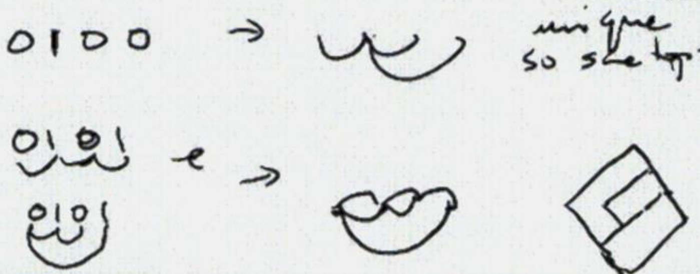
" " " " " " "

Note that in fact 011 is a "forbidden" beginning since it forces same on bottom & top at beginning so a cancellation:

(also easy in Thompson picture)

it seems to be a question of interest which representation shows you which thi

Continuing 4 leaf.

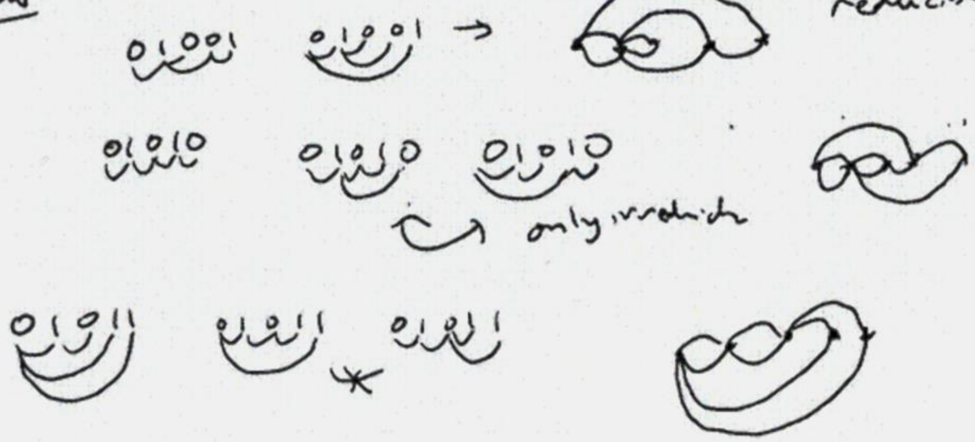


0100 → unique so same top & both

0101 & →

0101

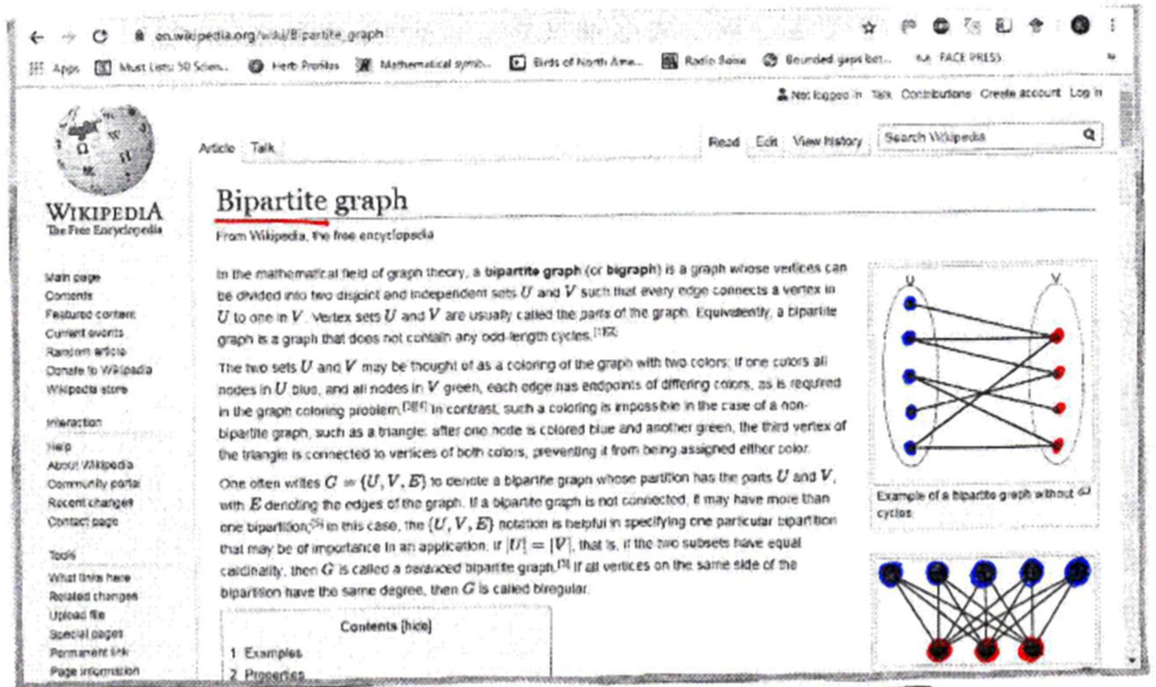
5 leaf 01000 unique



reducible

only irreducible

Original in colour


Bipartite graph

From Wikipedia, the free encyclopedia

In the mathematical field of graph theory, a **bipartite graph** (or **bigraph**) is a graph whose vertices can be divided into two disjoint and independent sets U and V such that every edge connects a vertex in U to one in V. Vertex sets U and V are usually called the parts of the graph. Equivalently, a bipartite graph is a graph that does not contain any odd-length cycles.[1][2]

The two sets U and V may be thought of as a coloring of the graph with two colors: if one colors all nodes in U blue, and all nodes in V green, each edge has endpoints of differing colors, as is required in the graph coloring problem.[3][4] In contrast, such a coloring is impossible in the case of a non-bipartite graph, such as a triangle: after one node is colored blue and another green, the third vertex of the triangle is connected to vertices of both colors, preventing it from being assigned either color.

One often writes $G = (U, V, E)$ to denote a bipartite graph whose partition has the parts U and V, with E denoting the edges of the graph. If a bipartite graph is not connected, it may have more than one bipartition;[5] in this case, the (U, V, E) notation is helpful in specifying one particular bipartition that may be of importance in an application. If $|U| = |V|$, that is, if the two subsets have equal cardinality, then G is called a balanced bipartite graph.[3] If all vertices on the same side of the bipartition have the same degree, then G is called biregular.

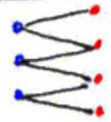
Example of a bipartite graph without cycles

Contents [hide]

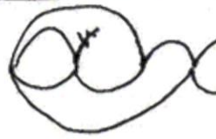
1 Examples

2 Properties

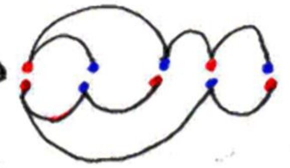
So a graph whose vertices can be sorted into two groups – each edge has endpoints in different groups



how about that example?



so these don't match!



I think that means it's actually not bipartite – so does that mean that

easy to read bipartite means – it's easy to read off whether the graph is bipartite?

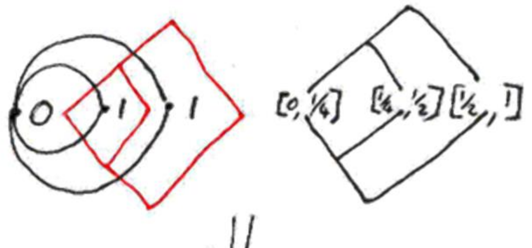
Just labelling makes it really easy to see either way

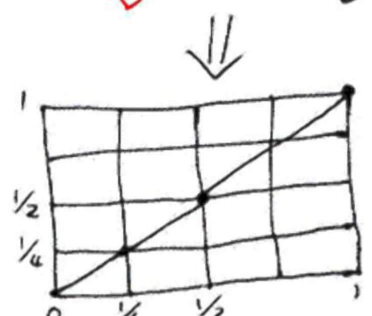
Can replace • and • with 0 and 1

So what is "forbidden" about forces save on bottom e top ?

Why is that bad?

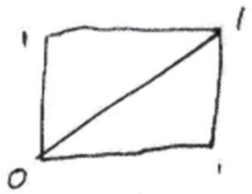
May be try an example and stare at it...





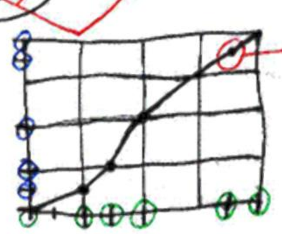
Just a straight line – there are no undifferentiable corners! It's like they arent there!

So the graph is the same as



same happens whenever the x and y co-ordinates are the same

So is this always the case when we see an 'eye' like this: ? that doesn't continue onto the right

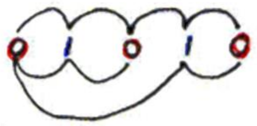


Same again! the makes no corner!

I don't understand the Ising subgroup part – there are things called Ising models in physics but I don't see how the literature connects up! I guess the members of the Ising subgroup are:

- irreducible
- Bipartite?
- Start with 0 1

So what might match, and not be the same?



The counting is partly done for you

So this talk of reducible and irreducible

might be which have this which can be removed

So having spelled out the method of construction, then the mathematician can look at it and see what the advantages of this representation are, and even answer questions using it! This is a kind of shift, an advance.

So how does it end up looking in the paper, published version?

We see a section in the paper presenting the 'instructions' (in just such a procedural form, showing how the graph should be constructed) but in a much more heavily textual form, fewer sketches and arrows, much less repetition – even though really repetition can be helpful, when learning to do something!

Instead we see a lot of naming, technical terminology, reference to mathematical concepts...

So this reference is to this paper:

[3] J. W. CANNON, W. J. FLOYD, AND W. R. PARRY (1996).

Because P and Q are not unique, there are ma
(R, S) for f, another can be constructed by adjoi
R for some positive integer n, and let J be the n^{t}
with root I, and let J_1, J_2 be the leaves in order
$f(I) = J$, it follows that $f(I_1) = J_1$ and $f(I_2) = J$
and $S' = S \cup D$.

In the other direction, if there exists a positi
respectively S, are the vertices of a caret C, respe
R and S leads to a new tree diagram for f. If th
diagram (R, S) is said to be *reduced*.

In this paragraph it will be shown that there
(R, S) is a reduced tree diagram for f. It is easy
a leaf of R or not in R, then $f(I)$ is a standard
standard dyadic interval such that $f(I)$ is a stand
leaf of R or not in R because (R, S) is reduced.
interval I is either a leaf of R or not in R if and
I. This gives uniqueness of reduced tree diagram

so is this the we m
a 'caret' is th

4. CALCULATION OF C

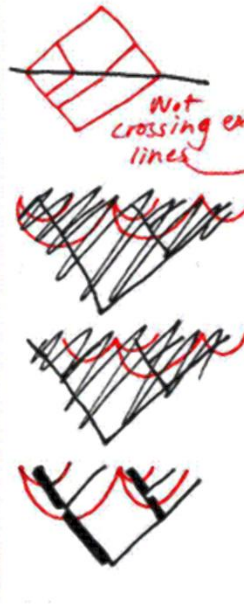
4.1. **Representation of elements of F as**
in [3], any element of F is given by a pair of b
Our convention will be that each standard dy
the top tree T_+ is sent by the Thompson grou
by the leaf on the tree T_- to which it is conne

we saw th before

Definition 4.1.0.4. *Given T_+ and T_- as a*
$g(T_+, T_-)$.

top & bottom

describing this:



The element g defines T_+ and T_- provided t
It will be convenient to arrange the two bifur
are the points $(1/2, 0), (3/2, 0), (5/2, 0), \cdots ((2N$
ing straight line segments sloping either up f
to right. T_+ is in the upper half plane and T
each region between the edges of each tree c
$\{(1,0), (2,0), ..., (N,0)\}$. Let us form a new pla
The vertices of Γ are $\{(0,0), (1,0), (2,0), ..., (N,$
passing once transversally through certain edge
the top tree use all the edges sloping up from le
and from the bottom tree use all the edges slo
we cal WS edges). The figure below illustrat
a pair of bifurcating trees with 5 leaves. We h
their x coordinates, and we will henceforth use t

This is all a much more long-winde way of describing the relationships w saw exemplified in the sketched diag

- Connects it up with known ideas lik
- Expressed in precise language, not sk
- More ... stable, somehow?

Original in colour

Original in colour

So this reference is to this paper:

[3] J. W. CANNON, W. J. FLOYD, AND W. R. PARRY (1996). And here's the relevant part:

Because P and Q are not unique, there are many tree diagrams associated to f. Given one tree diagram (R, S) for f, another can be constructed by adjoining carets to R and S as follows. Let I be the n^{th} leaf of R for some positive integer n, and let J be the n^{th} leaf of S. Let I_1, I_2 be the leaves in order of the caret C with root I, and let J_1, J_2 be the leaves in order of the caret D with root J. Because f is linear on I and $f(I) = J$, it follows that $f(I_1) = J_1$ and $f(I_2) = J_2$. Thus (R', S') is a tree diagram for f, where $R' = R \cup C$ and $S' = S \cup D$.

In the other direction, if there exists a positive integer n such that the n^{th} and $(n+1)^{\text{th}}$ leaves of R, respectively S, are the vertices of a caret C, respectively D, then deleting all of C and D but the roots from R and S leads to a new tree diagram for f. If there do not exist such carets C, D in R, S, then the tree diagram (R, S) is said to be *reduced*.

In this paragraph it will be shown that there is exactly one reduced tree diagram for f. Suppose that (R, S) is a reduced tree diagram for f. It is easy to see that if I is a standard dyadic interval which is either a leaf of R or not in R, then $f(I)$ is a standard dyadic interval and f is linear on I. Conversely, if I is a standard dyadic interval such that $f(I)$ is a standard dyadic interval and f is linear on I, then I is either a leaf of R or not in R because (R, S) is reduced. Thus R is the unique T-tree such that a standard dyadic interval I is either a leaf of R or not in R if and only if $f(I)$ is a standard dyadic interval and f is linear on I. This gives uniqueness of reduced tree diagrams.

so is this the ○ we were looking at?
a 'caret' is this: ^

...tion
...) (in
... procedural
...ing how
...hould be
...ed) but
...h more
...textual form,
...ketches and
..., much less
...ion – even though
... repetition can
...lpful, when learning
...to something!
...ead we see a lot
... naming, technical
...minology, reference
... mathematical
concepts...

4. Calculation of coefficients

4.1. Representation of elements of F as pairs of rooted planar trees. [branching (in two)] As in [3], any element of F is given by a pair of bifurcating trees T_+ and T_- as below. Our convention will be that each standard dyadic interval represented by a leaf of the top tree T_+ is sent by the Thompson group element to the interval represented by the leaf on the tree T_- to which it is connected.

we saw these before — So, this is the pairing we saw before

Definition 4.1.0.4. *Given T_+ and T_- as above the element of F will be called $g(T_+, T_-)$.*

top & bottom

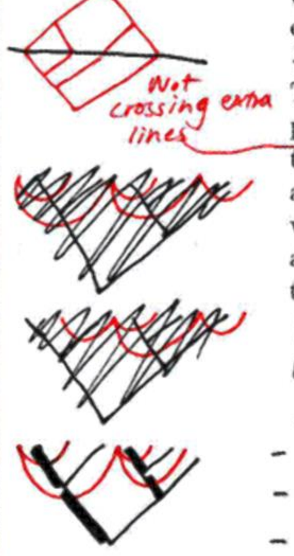
The element g defines T_+ and T_- provided there are no cancelling "carets"-see [3]. It will be convenient to arrange the two bifurcating trees in $\mathbb{R}^2$ so that their leaves are the points $(1/2, 0), (3/2, 0), (5/2, 0), \cdots ((2N-1)/2, 0)$, with all of the edges being straight line segments sloping either up from left to right or down from left to right. T_+ is in the upper half plane and T_- is in the lower half plane. Then each region between the edges of each tree contains exactly one point in the set $\{(1,0), (2,0), ..., (N,0)\}$. Let us form a new planar graph Γ given from the two trees. The vertices of Γ are $\{(0,0), (1,0), (2,0), ..., (N,0)\}$ and the edges are given by curves passing once transversally through certain edges of the top and bottom trees. From the top tree use all the edges sloping up from left to right (which we call WN edges) and from the bottom tree use all the edges sloping down from left to right (which we cal WS edges). The figure below illustrates the formation of the graph Γ for a pair of bifurcating trees with 5 leaves. We have numbered the vertices of Γ with their x coordinates, and we will henceforth use that numbering to label those vertices.

$\mathbb{R}^2$ = a two-dimensional space – a flat plane like a piece of paper

West-North

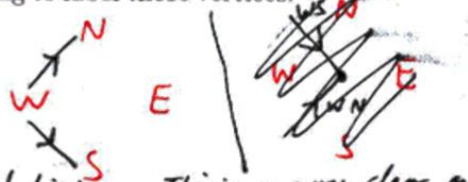
West-south

describing this:



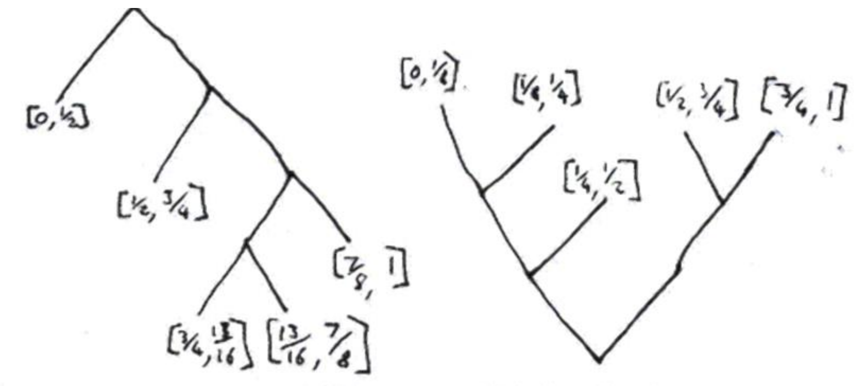
Seems important to situate everything wrt existing ideas

This is all a much more long-winded way of describing the relationships we saw exemplified in the sketched diagrams.
- Connects it up with known ideas like $\mathbb{R}^2$
- Expressed in precise language, not sketches and hints
- More ... stable, somehow?

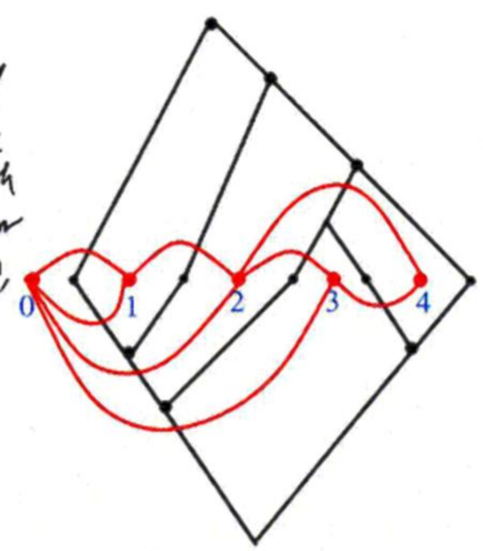


This is a very clear and precise way of putting something that I could just take for granted!

Everything labelled and placed



We see this superimposition again (after the textual description!) and this time there is this very careful description of exactly what is going on with the intervals – the author needs to get this all down, written, making things explicit and clear.



How the 'tree' graph translates to intervals was probably something the author knew very well, knew how to do without thinking, so didn't need to make explicit in the notes. But in the paper it's important to make it all explicit, spell it all out, not leave anything unstated as far as possible!

To be quite clear the above element of F is linear on each of the following five standard intervals, which it maps to the next five in the given order:
$\{[0,\frac{1}{2}],[\frac{1}{2},\frac{3}{4}],[\frac{3}{4},\frac{13}{16}],[\frac{13}{16},\frac{7}{8}],[\frac{7}{8},1]\} \to \{[0,\frac{1}{8}],[\frac{1}{8},\frac{1}{4}],[\frac{1}{4},\frac{1}{2}],[\frac{1}{2},\frac{3}{4}],[\frac{3}{4},1]\}$

Γ – so this is the curve version?

Definition 4.1.0.5. *Given T_+ and T_- as above , the planar graph Γ defined above is called the planar graph of T_+, T_-, written $\Gamma(T_+, T_-)$ or $\Gamma(g)$ if there are no cancelling carets so that the data of the two trees is the same as the data $g \in F$.*

More precise naming

Observe that the procedure for constructing Γ actually constructs a rooted tree $\Gamma_\pm(T_\pm)$ with vertices $\{(0,0), (1,0), (2,0), ..., (N,0)\}$ from a single bifurcating tree $T_\pm$ either in the upper (+) or lower (−) half plane with leaves $(1/2, 0), (3/2, 0), (5/2, 0), \cdots ((2N-1)/2, 0)$.

Note that the graph $\Gamma(T_+, T_-)$ is also a pair of rooted planar trees, one in the lower half plane and one in the upper half plane having the same root and the same leaves. But they are not bifurcating in general, the valence of each vertex being unconstrained.

Could be viewed as one tree or two?

Cancelling of carets between T_+ and T_- corresponds to removal of a two-valent vertex connected only to its neighbour, and the edges connected to it.

remove this vertex and edges connected

Proposition 4.1.1. *The graph Γ formed above from a pair of bifurcating trees consists of two trees, Γ_+ in the upper half plane and Γ_- in the lower half plane, having the following properties:*
(0)The vertices are $0, 1, 2, \cdots, N$.
(i) Each vertex other than 0 is connected to exactly one vertex to its left.
(ii) Each edge can be parametrized by a smooth curve $(x(t), y(t))$ for $0 \le t \le 1$ with $x'(t) > 0$ and either $y(t) > 0$ for $0 < t < 1$ or $y(t) < 0$ for $0 < t < 1$.

Proof. This is obvious from the construction of Γ. □

Great!
So this is all stuff that we know because of the way it's built –
But written out in a referenceable statement

to be referenced later – also labelled, Proposition 4.1.1!

In comparison with the notes, these descriptions seem at once far more explicit — everything is written — and more terse, or economical — examples are not multiplied! Just one diagram for each thought

oriented like have a direction

Graphs of the from $\Gamma_\pm$ are obviously oriented so we may talk of the source and target of an edge. We will show below how to reconstruct the pair of bifurcating trees from a pair of rooted planar trees with vertices satisfying the conditions of 4.1.1.

This shows that $\Gamma(g)$ is an equally faithful way of representing elements of the Thompson group F.

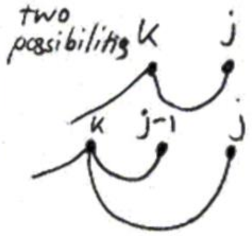
Loses no detail?

Lemma 4.1.1. *Let Ψ be a rooted tree in the upper or lower half plane satisfying the conditions of proposition 4.1.1. Then there is a bifurcating tree $T_\pm$ such that $\Psi = \Gamma_\pm(T_\pm)$.*

Proof. Wolog we may assume everything is in the lower half plane.

We will work by induction on the number of leaves. So suppose we are given a Ψ_- satisfying the conditions of 4.1.1 with $N+1$ leaves. Call a vertex of Ψ_- *terminal* if it is not the source of an edge.

two possibilities



If j is a terminal vertex then it is the target of a unique edge. The source of that edge is k for $k < j$. If $k = j-1$ we will call j *minimal terminal*. If j fails to be minimal terminal then $j-1$ could, by planarity, only be connected to the right to j, so $j-1$ is terminal. Continuing in this way we obtain a minimal terminal vertex m. There are then two possibilities for $m-1$

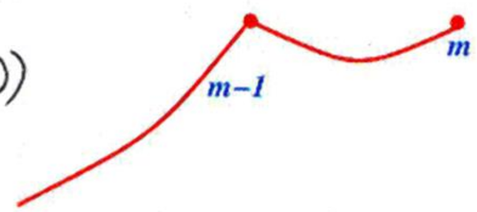
Minimal terminal is simple end

effort to establish exhaustion for proof

Case(1). Valence of $m-1$ is 2. Then if m is deleted $m-1$ becomes terminal and in a neighborhood of m and $m-1$ Ψ is as below:

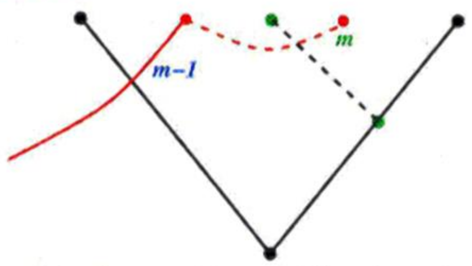
move to talking about deletion then re-insertion!

Exterior case (Formerly case(2))



swapped to stream line exposition? Easier to get a handle on this one first?

Removing m and its edge the resulting graph Ψ' still satisfies the conditions of 4.1.1 so there is by induction a T' with $\Psi' = \Gamma_-(T')$. Observe that the terminal vertex $m-1$ is necessarily in a caret of T'. We may thus add WS edge to T to re-insert the vertex m and obtain the desired T_-:



where the solid edges are those of T' and the dashed edges are the ones added to obtain T_- and Ψ.

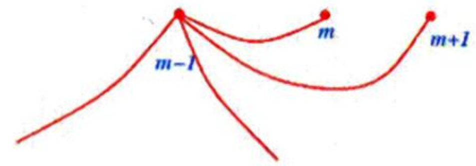
Again, we see colour and dotted lines etc fulfilling important functions — grouping ideas, highlighting changes, suggesting changes in time, even!

Here though, again, we have far more labelling of each vertex etc.

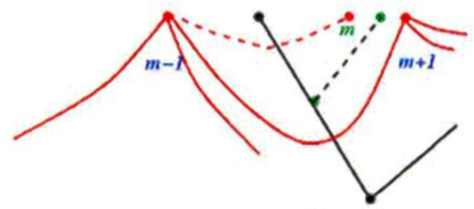
Interior case (Formerly case (1))

Again, the language of necessity and completeness

Case(2) Valence of $m-1$ is > 2. In this case there is an edge with source $m-1$ connecting it to a vertex k with $k > m$. By planarity there must be such an edge connecting $m-1$ to $m+1$. The situation near m is thus:



Removing m and its edge the resulting graph Ψ' still satisfies the conditions of 4.1.1 so there is by induction a T' with $\Psi' = \Gamma_-(T')$. There has to be a WS edge in T between $m-1$ and $m+1$ so we may add a WN edge to T' as in the figure below re-insert m and obtain the desired T_-:



where the solid edges are those of T' and the dashed edges are the ones added to obtain T_- and Ψ. □

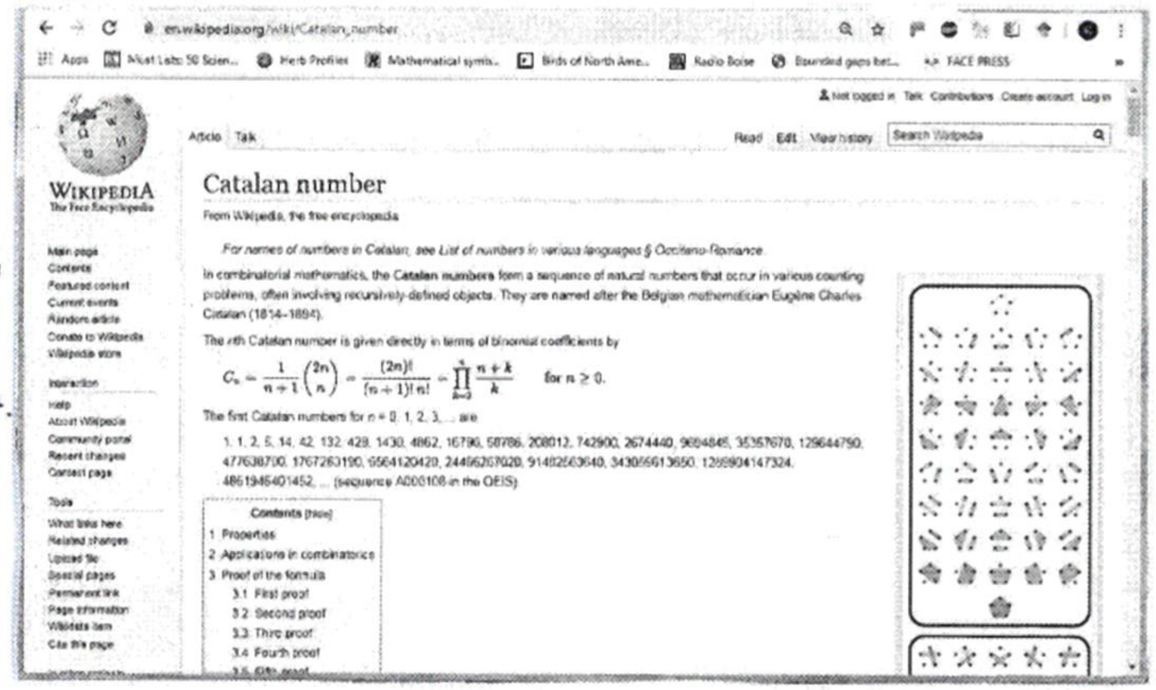
This means:- it is proved!

The author thinks that linking the types of graph isn't original, but has certain uses in mind, to follow.

We have essentially given a bijection between two sets counted by the Catalan numbers which thus almost certainly exists in the literature. The result is so important to our examples that we have supplied a detailed description of what is an algorithm to obtain a Thompson group element from a pair of planar graphs with the same set of leaves and the same root.

Again, situating things seems important

These are cool - there are all kinds of ways of representing them visually.



en.wikipedia.org/wiki/Catalan_number

Catalan number

From Wikipedia, the free encyclopedia

In combinatorial mathematics, the Catalan numbers form a sequence of natural numbers that occur in various counting problems, often involving recursively-defined objects. They are named after the Belgian mathematician Eugène Charles Catalan (1814–1894).

The nth Catalan number is given directly in terms of binomial coefficients by

$$C_n = \frac{1}{n+1}\binom{2n}{n} = \frac{(2n)!}{(n+1)!\,n!} = \prod_{k=2}^{n}\frac{n+k}{k} \quad \text{for } n \geq 0.$$

The first Catalan numbers for n = 0, 1, 2, 3, ... are

1, 1, 2, 5, 14, 42, 132, 429, 1430, 4862, 16796, 58786, 208012, 742900, 2674440, 9694845, 35357670, 129644790, 477638700, 1767263190, 6564120420, 24466267020, 91482563640, 343059613650, 1289904147324, 4861946401452, ... (sequence A000108 in the OEIS).

Contents
1 Properties
2 Applications in combinatorics
3 Proof of the formula
3.1 First proof
3.2 Second proof
3.3 Third proof
3.4 Fourth proof

5.4 Summary

This analysis was carried out in the form of a set of handwritten notes, as a character comes to understand the content through drawing. External resources are reproduced in the notes as they are used.

In this section, the thing that came to the fore was the range of ways in which picture-drawing can help a person to calculate, construct, organise material and lay out components in such a way as to make the answers to questions immediately obvious. Early in my analysis I was advised by a supervisor to see if I could 'understand by drawing a few pictures', and it became clear that the process of construction really was one of coming to make sense of the material, to understand how the parts fit together and operated. A representation could also be designed in such a way as to make the answers to certain questions about an object's properties available, such as for example the number or colouring of endpoints making evident whether a particular graph is bipartite. A person working with a pen is thus able to build representations that streamline certain cognitive tasks, these external representations thus extending what the thinking person is able to do.

Again, it is clear that the presentation of the material in the paper was considerably more explicit, with long, precise definitions given for elements that were represented in the scribbled notes without any such definition. The expert, of course, constructs and uses these notes and diagrams with a rich contextual knowledge of the mathematical field to which the scribbled representations are to be related; manipulating the representations does not necessarily mean keeping all of this context in mind, but the expert constructs the representations with the intention of referring to it, and if prompted, would be able to furnish that context. In this way, quite simple ad-hoc mark-making is imbued with the power to answer very complex questions. In the communicative context of a paper, however, the contextual knowledge of that particular expert cannot be depended upon; instead an effort is made to furnish the appropriate context and references within the text with as much explicitness as possible for the rich and broad possible cognitive effects to be more reliably accessible to a reader with an unknown level of expertise. (The extent to which it is realistic to attempt to produce a paper with truly complete referencing and definitions is something that will be considered in Chapter 7.)

What is interesting is that this presentation also produces a clunky text that is highly effortful to read. A reader must take all of that technical explanation and digest it, and perhaps even draw a few pictures in order to come to a relevant, operational understanding.

Subject G

0.48.52.070

G: ((talking about differences of opinion among collaborators)) People have different writing styles. […] [People end up rewriting things because they] think it'd be written more clearly if it was written a different way. Often it's very subjective. […] So I write it one way, and to me it's clear, and I can fill in all the steps in my head. But that's partly because of my mathematical background, say.

Subject P

0.32.09.000

P: I think what should be acceptable is that people should be able to expand [on a proven result], say well here's my interpretation of that result, somehow making that result a bit more widely accessible. Actually that does happen in my field […] there were some papers that came out in the 90s, and actually people are still debating about those papers now […] and it turned out there were maybe 5 or 6 papers, not copying that paper word for word but people giving their own interpretation of what they thought that approach and result was but applying it to a slightly different system.

Work site 6

Work site 6 was the office of a lecturer at a UK university.

Figure 65. Corridor space outside my participant's office. Original in colour.

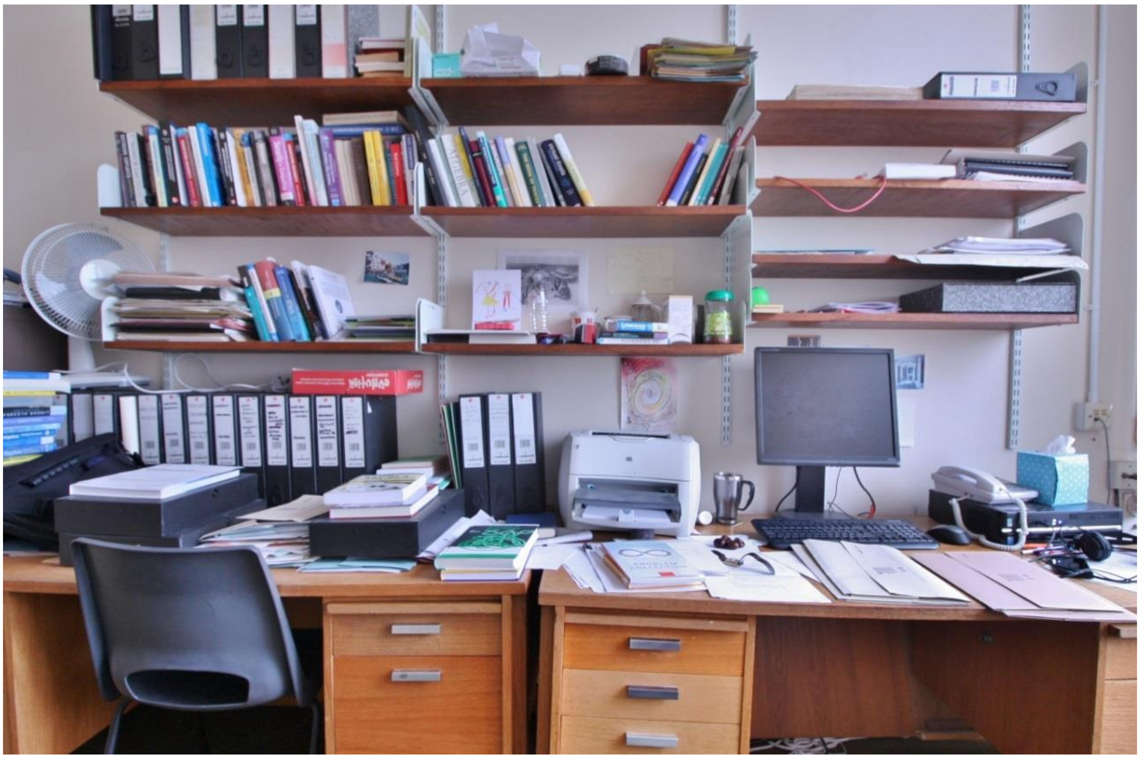


Figure 66. Workspace used by my participant, with folders of notes. Original in colour.

6. Discussion: Relevance and situatedness in the 'back' end of mathematics

As we saw in the introduction, it is a common habit to downplay the importance of the material practices of mathematics, and to view it as somehow separate from the 'real' mathematics. How is this idea supported? For this idea to make sense we might think of mathematics as essentially done through internal mental work and then written down; in that case the writing would just be an encoded summary of the work that has been done. How realistic is it to make sense of real mathematical situations in this light?

To answer this question we must first give careful consideration to the 'back' end of mathematics, and try to more closely examine how this work advances. To do so it would be as well to consider the role that mathematical writing is playing in situations of communication and thought, and to do that it is wise to give some thought to theoretical understandings of communication. If we are to look at mathematical writing as more or less just a conduit for pre-existing mathematical ideas, it would be well served by description in terms of what is known as the *code model* of communication.

Lemma 6.1. The code model of communication (and the mass of anthropological facts)

To begin, let us consider a possible characterisation of mathematical work. A mathematician thinks about a problem and sees the answer, writing it down. The mathematician thinks very hard and then writes down the symbols that encode the ideas.

Mathematicians work with concepts/entities and operations which are designated by certain symbols and systems of symbols. A mathematician may combine sets of symbols to indicate that certain entities should be operated upon in certain ways, and to indicate using those same systems of symbols what the outcome might be; the symbols can be read by another person who knows the definitions of each symbol to retrieve the message; to see what has been done and what the answer is. In this way, the discoveries of mathematics are shared.

This admittedly simplistic characterisation is based in a 'code model' picture of communication, in which each thinker has knowledge of a set of definitions, and so is able to translate concepts into and from symbolic representations. The code model, an idea that dates back to Aristotle but is today most often described in the terms originally delineated for information transmission by Shannon and Weaver (1949), describes communication as a process in which a thought is encoded, transmitted, received, and decoded to retrieve the original thought; it is this basic conception of communication that underlies the semiotic approaches of Peirce, de Saussure, and their followers, which they believed would make sense of communication, and beyond. Structuralists such as de Saussure expected that 'the laws discovered by semiology will be applicable to linguistics, and the latter will circumscribe a well-defined area within the mass of anthropological facts' (de Saussure *et al.*, 1974 p.12). Claude Lévi-Strauss also sought to find a common code or 'grammar', which underpinned, and would provide a solution to all kinds of aspects of culture. The mind-inscription interaction thus described is one governed by precise rules and straightforward translations, in which inscriptions are no more than a conduit.

It is worth revisiting certain criticisms that have been made of the code model and its ability to account for real-world communication in verbal contexts. Sperber and Wilson (Sperber & Wilson, 1986/1995) make an argument for the insufficiency of the code model in ordinary conversation that proceeds as follows: for a straightforward version of the code model to work, both parties must have, and know that they have, mutual knowledge of a language and a context (the Mutual-Knowledge hypothesis). In that case, they can use the mutually known language to encode a thought, and with the mutually known context, it can be reliably and correctly decoded. In conversation, it is not too difficult to argue that the correct portions of a language might be known by both parties in roughly similar ways, but to argue that a clear and limited context is identically known is harder. In ordinary conversation there is not even much of an attempt at establishing a restricted field of mutual knowledge, nor agreeing what in that field ought to provide the context,[10] and yet communication is generally successful; so, they argue, another model is needed.

The code model is the basis for semiotic approaches to culture, and can come up against similar limitations in these applications. A famously innovative analysis of mathematical papers was put forward by Brian Rotman, presented as a first attempt at a semiotics of mathematics (Rotman, 2000). This influential text describes the content of papers as consisting in a Code and meta-Code, the encoded meaning of mathematical notation and the meta-narratives that work above these. As described in a precursor to this paper, for Rotman:

> Proofs embody arguments -- discursive semiotic patterns -- that work over and above - before - the individual steps and which are not reducible to these steps; indeed, it is by virtue of the underlying story or idea or argument that the sequence of steps is the sort of intentional thing called a proof and not merely an inert string of formally correct inferences. (Rotman, 1998 p.65)

Rotman employs a semiotic model with a basis in what might be termed a 'code model' of communication, and accordingly his analysis proceeds by identifying different codes in mathematical papers, including posited characters invoked by different parts of a mathematical texts. These characters are not supposed to be explicitly considered by a reader but are described in order to

[10] For the question of mutual knowledge, law is invoked as a contrast case, since in that context there really is 'a serious attempt to establish mutual knowledge among all the parties concerned: all laws and precedents are made public, all legitimate evidence is recorded, and only legitimate evidence can be considered, so that there is indeed a restricted domain of mutual knowledge on which all parties may call, and within which they must remain' (Sperber & Wilson, 1995 p.19). It is to be noted that, given the debates about how to interpret precedent, an application of the code model is not entirely without problem even here, but the case made is that evidence of even an attempt is beyond that which is seen in verbal communication.

discern different kinds of content in a paper. On the other hand, he also argues that the real success of a paper depends on a reader's ability to recognise what he terms a meta-Code: some reading of intentions and an implicit narrative that allows a reader to understand how an argument fits together. This meta-Code represents a rather inferential-sounding augmentation to a simple code account, added to provide a more plausible characterisation of how a reader might become convinced. This is a very reasonable question to address, but an answer bound to code-model description is likely to run up against certain problems when it comes to explaining just how such a meta-Code might be accessed. The kinds of inscriptions seen in a mathematical paper do invite and tempt such a code-based description, girded as they are by definitions and expectations of explicitness and precision.

The alternative that Sperber and Wilson propose in the context of conversational communication is a process that involves both decoding and inferential processes (Sperber & Wilson, 1986/1995 p.3). While decoding with reference to a mutually known language and context is part of the story, they argue, it is less easy to explain how interlocutors understand one another in the very common event that mutual knowledge is not perfect, or that meaning is somehow underdetermined by the encoded meanings of the ostensive stimuli alone, or both. To answer this question, they posit that interlocutors are also able to make inferences about one another's intentions on the basis of Theory of Mind abilities, and to infer with reference to these a spectrum of more and less clearly intended conclusions.

While the case appears well made for everyday communication, mathematics would appear to be a better candidate for a domain in which the ways that its practitioners share ideas could be well explained by the code model, considering how well-defined the components of its inscriptions are, and the emphasis placed on precision and completeness by its practitioners. There is reason to take each communicative situation on its own merits, and with a critical eye that looks beyond the mere *existence* of a code, to the way that communication is actually achieved:

> It becomes an empirical question whether the code model can provide a full account of a given communication process. It is not enough to show that a code is being used; one must also be able to show that what is communicated is actually being encoded and decoded. Otherwise, all that can be reasonably maintained is that the use of a code plays some role in this particular communication process, without perhaps wholly explaining it. (Sperber & Wilson, 1986/1995 p.27)

What I set out to do is examine whether the way that mathematical communication works can be adequately explained within a code model framework. Following Sperber and Wilson, one question is whether the mutual knowledge hypothesis is really met in various cases of mathematical communication, since this is arguably a precondition. As a broader consideration, I will consider just what is achieved through mathematical writing, and whether it is the kind of thing that can be encoded—and if not, then what?

Remark 6.2. Not *saying* it, exactly

I will begin with the first of the excerpts analysed, in Chapter 3, in which we saw a group of mathematicians speaking around a whiteboard. In this example, the writing on the board was clearly the focal point of the conversation.

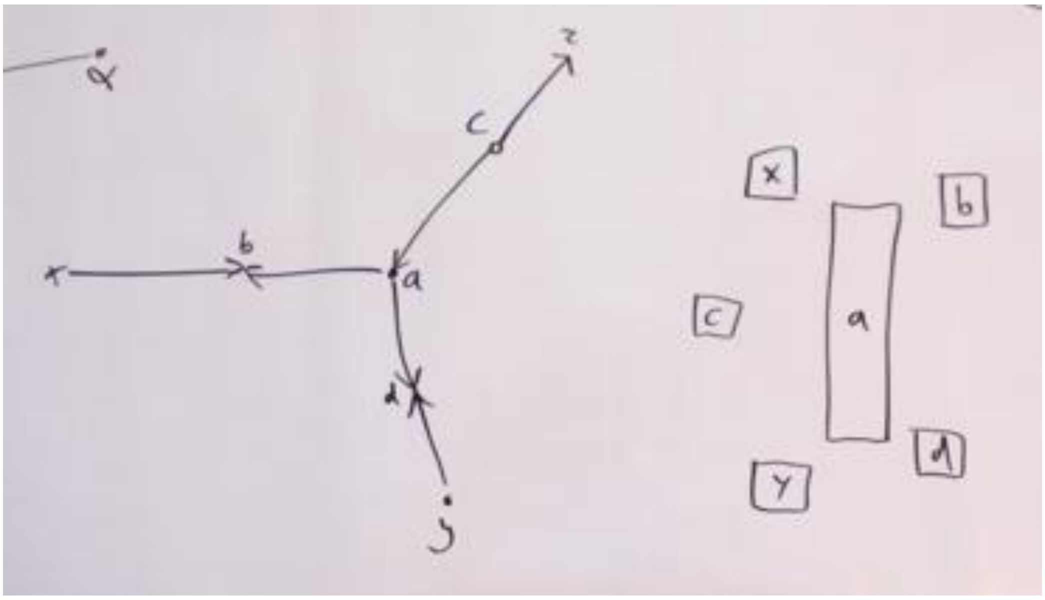


Figure 67. Board notes for the excerpt. Original in colour.

In this excerpt, the group has been focusing on the counter-example represented by the board notes shown in Figure 67. The group is working toward proposing a short, repeatable, technical description that will help them to stabilise a perspective on a situation; in this case, the group is seeking a statement that defines the situations in which an oriented graph (Figure 67, left) cannot be represented as a rectangle visibility graph (Figure 67, right), which they are referring to as what is 'bad'.

At one moment, B proposes such a statement, which is queried by C. A clarification is offered by F (Figure 68), and the form that this clarification takes shows us something of the way the board notes are being used during the excerpt.

01 B: alright so the thing that's [bad
02 D: [OK I'm with you
03 B: … is a vertex with three switching vertices
04 C: [what
05 A: ((nods)) [or more
06 B: or [more
07 C: [a - vertex with three switching vertices?
08 F: cos like [in that with that vertex a
09 E: [((unintelligible)) right there
10 F: … b is switching for a, [d is switching for a [and c is supposed to be switching for a
11 B: ((walks over, follows path xba)) [yup ((follows path zca))[yup
12 F: but we have nowhere to put – z
13 B: right

Figure 68. Transcript

F responds to C, and responds by talking through the counter-example in Figure 67 in such a way as to *exhibit* it as satisfying B's description of what is 'bad': 'a vertex with three switching vertices'. When asked for further clarification by a colleague, F's approach is simply to talk through the diagram, rather like giving C a guided tour of its features, with commentary. This commentary is quite minimal; F mentions three specific points, *b*, *d* and *c*, and their relationship to *a*. The reader may recall from Chapter 3 that a major shift in the group's thinking during this meeting had been from looking at the diagrams in terms of *switching paths*, paths that switch direction when considered as beginning from some origin and extending to some end point, to concentrating on *switching vertices*; the point at which switching occurs, with either two arrows pointing inward or two arrows pointing outward; B's definition was constructed in terms of *vertices*, reflecting this development. F's guided tour carefully focuses attention on the *vertices* in the diagram, localising the 'switching' at these points. F's description is also very rhythmic in a way that exhibits the fact that there are three such vertices that are all in a certain similar relationship to one vertex, *a*. F's description seems to behave rather like an attention-directing narration that shares a way of perceiving the diagram, that emphasises some features and downplays others, that brings forth a pattern and a way of looking at its structure that serves in some way to highlight B's definition as capturing well the particular feature that makes it the case that 'we have nowhere to put *z*'.

F's description *simultaneously* describes the tree graph *and* the process of constructing the RVG, or perhaps moves between the two. At the beginning, F refers to 'vertex *a'* in the tree graph, but the explanation ends with the moment at which the process of constructing the RVG breaks down, when 'we have nowhere to put *z*'—in the tree graph, of course, *z* is present and correct, so this last comment must be focused on the RVG representation alone. In this double-sided description F manifests the effect that a feature has on that process of construction, thus doing something like showing that paying attention to switching vertices will help you to see *why* a tree graph cannot be represented as an RVG.

Nowhere is this perspective *encoded* within the lines spoken. F could have given C a much more explicit affirmation of B's description, such as (1):

(1) B's description matches the features of the example on the board. In this example vertices *b*, *d* and *c* all meet the criteria for 'switching vertices' and each one is connected to vertex *a*. This shows that our problem example is described by B's definition.

Of course for the meaning to be fully explicit F would have to specify much more carefully what some of these terms refer to, such as which example is being discussed. More importantly though what is included in (1) is really not the equal of F's original rambling description. For one thing, F's clever double-speak is not included. We could add the following:

(2) If you were constructing an RVG for the tree graph in the example, you would be able to place *b, x, c* and *y,* but having placed *c* you would then be unable to place *z*.

This though still does not seem to capture what is conveyed by F. It would be possible to add a whole host of other elucidations in the interests of doing so, such as these:

(3) You need to pay attention to switching vertices and their relationship to other vertices

(4) You need to look for three-pronged constructions like this in other diagrams because this is the important feature highlighted by B's description

(5) If you pay attention to the switching vertices in turn as you attempt to construct an RVG for a tree graph then you will be able to pinpoint the moment at which the graph becomes unrepresentable

Yet even these do not seem to encapsulate the sum of what is conveyed by F. These attempts to make explicit that which F is conveying seem to say simultaneously much more and much less.

The question we have to answer, then, is how the participants come to understand one another in the course of the excerpt, if encoding is an insufficient explanation. F's answer to C is apparently sufficient—it is certainly accepted by B—but does not appear to explicitly answer the question. This answer serves to convey something that is subtle, very helpful, and seems even to resist being encapsulated in encoded form. How, then, is this 'something' conveyed and understood?

Proposition 6.3. Relevance theory

Relevance theory is a theory proposed by Dan Sperber and Deirdre Wilson, which explains utterance interpretation in terms of a combination of decoding and inference on the part of the hearer. The code model of communication holds that encoding and decoding messages is all that is needed to explain communication, but it has been argued to be insufficient to explain even the most straightforward instances of communication. I return to the following example, given in Chapter 1:

(4) Dehn solved one of Hilbert's problems

From which it might be possible to reach multiple valid interpretations:

(5) Dehn solved a problem in Hilbert's personal or professional life

(6) Dehn solved one of Hilbert's famously published list of 23 mathematical problems

Which of (2) and (3) would be selected as an interpretation by the hearer would be highly dependent on context, and could not be derived from the code alone (Sperber & Wilson, 1986/1995 p.13); however, in context, reaching the correct interpretation would be near-effortless. This is a problem with semiotics that is important to note: it has no way of explaining the gap between what is said and what is meant (the interpretation that we would all quickly and effortlessly reach), other than proposing yet more codes. This is ultimately unsatisfying; without reference to inference, the code model offers little account of how the codes come to be shared and unproblematically decoded even in such diverse circumstances and uses. By way of contrast, the inferential models proposed by philosophers such as Paul Grice and David Lewis see communication as achieved by producing and interpreting evidence.

The shift from considering sentences to considering utterances (that is, realisations of the phonetic representation of a sentence) is the shift from *semantics* to *pragmatics*. The question of how the hearer sets about narrowing down and choosing among possible interpretations is the business of pragmatics, and that must be some form of logical process on the part of the hearer that refers to certain relevant premises.[11] Relevance theory rests on two principles: a cognitive principle of relevance, in which they describe cognition as being geared toward maximising relevance, and a communicative principle of relevance that was influenced by Grice's Maxims of Conversation, through which he described communication as a cooperative activity. While relevance theory does away with his conception of communication as cooperative, it follows Grice by proposing that utterances raise expectations. The communicative principle takes it that utterances come with expectations of relevance, where relevance is defined in terms of a balance between cognitive effects, the results gained from an utterance, and the processing effort expended during their search for relevance. Relevance theory claims that the basic predictability of communicative situations makes it possible for hearers to have dependable expectations about the way that a speaker has chosen which utterance to make in terms of an expected level of *relevance*, and so to properly interpret utterances.

To the casual observer, it might appear that this debate ought not have much bearing on the study of mathematical work. For something as determined and unambiguous as mathematics, surely the code model is uniquely appropriate. My response is this: the content of a mathematical statement, rigorously written as part of a paper, may indeed be uniquely and clearly determined for those familiar with the language. But underdetermined aspects to mathematical communications creep in nonetheless. There is no clearly defined way for a proof to progress, and for example it is *not* clearly defined what might be intended when an inequality is expanded in one way rather than another. While the content of a statement may be clearly defined by a code, its implications will be dependent on the context of the problem under consideration, as well as subtle questions about the role being played in the paper. In-person communication between colleagues or collaborators is subject to the many moments of vagueness and adjustments common in any real-time conversation. In short,

[11] Sperber and Wilson actually offer a reconciliation between code and inferential models, stating that if the premises and process really were held completely in common between speaker and hearer, this would function as a code (for example, if speaker and hearer both knew—and knew that one another knew—that the Riemann hypothesis remains unproven, and the speaker then stated, 'if the Riemann hypothesis is still unproven then this question cannot yet be answered'); however they propose that such examples of clearly-defined mutual knowledge are rare, and so that other means must be sought to explain how these inferential processes are successful.

mathematical communication may be rather less code-like, and more based on the inferring of human minds, than it may appear.

When a speaker pronounces an utterance or, more broadly, behaves ostensively, the hearer becomes aware of an intention to communicate. Perceived intentions are central to this theory, which foregrounds comprehension as 'a variety of *mind-reading*, or *theory of mind* (the ability to attribute mental states to others in order to explain and predict their behaviour).' (Wilson & Sperber, 2005 p.276). Sperber and Wilson separate out two types of intention: the informative intention, which is the intention to inform an audience of something, and the communicative intention, which is the intention to inform the audience of one's informative intention. The recognition of an informative intention produces an expectation in that hearer that whatever the speaker is trying to communicate will have sufficient effects to justify the effort that the hearer is put to (Sperber & Wilson, 1986/1995 pp.29, 58, 61). The hearer therefore looks for the most relevant interpretation to satisfy the expectation of relevance that this recognition of intention promises, relevance being defined as follows:

> (1) Relevance of an input to an individual
> a. Other things being equal, the greater the positive cognitive effects achieved by processing an input, the greater the relevance of the input to the individual at that time.
> b. Other things being equal, the greater the processing effort expended, the lower the relevance of the input to the individual at that time. (Wilson & Sperber, 2005 p.252)

Because of the expectation that the speaker will select the most efficient route possible, the principle is that in interpreting an utterance, an audience looks for an interpretation that is optimally relevant; that is, an interpretation that will yield cognitive effects that will justify the effort needed to interpret it.

It is expectations like these in communicative situations that made the *breaching experiment* 'Dialogue without shared aims' in Chapter 3 so very difficult to write. In conversation, relevance theory would have it that we rely heavily upon mutual awareness of aims, intentions and beliefs and that our utterances are structured with these factors in mind. It was not the fact that the interlocutors had different aims that was the problem; conversations happen all the time between people who disagree with one another. It was the fact that the utterances did not build upon one another, did not properly respond to one another: were not relevant to one another. This transgression was so great as to make it seem that the conversation was not a real conversation at all.

Relevance theory is thus a theory that looks to our understandings of one another's minds to explain how we are able to understand one another. It also argues in terms of our cognitive architecture to posit forces that shape our communication strategies, such as a tendency toward brevity where possible, a consideration of effort, and the need for speaker and hearer to be aware of one another's abilities when communicating.

The interpretive process posited by relevance theory involves an element of decoding the explicit content of an utterance to find its *explicatures*, and also a process of inference with reference to contextual assumptions, which produces *implicatures*. The concepts of *explicature* and *implicature* are quite important. In mathematics, a conclusion might be considered to be implied by a set of axioms, but only made explicit by the theorem that demonstrates the development to that conclusion. In linguistics what is implicated need not have quite the same strong logical relationship with what is said, but the relationship is roughly analogous; a hearer might bring in all kinds of contextual knowledge, and will then use an expectation of relevance to select an interpretation. Implicatures can be more or less expected to be intended by the speaker, can be multiple, and can be anything from very strongly to very weakly implicated.

> Implicated conclusions are deduced from the explicatures of the utterance and the context. What makes it possible to identify such conclusions as implicatures is that the speaker must have expected the hearer to derive them, or some of them, given that she intended her utterance to be manifestly relevant to the hearer. (Wilson & Sperber, 2005 p.195)

The following example might explain this notion.

(7) (a) Peter: Would you use Rabin's probabilistic algorithm?
(b) Mary: I wouldn't use any plausibility argument.

It would be reasonable for Peter to understand this as an answer to his question, but only by bringing in some contextual assumptions of his own:

(8) Rabin's probabilistic algorithm is a plausibility argument.

If interpreted in a context with assumption (5), (4b) would yield the contextual implication (6)*:*

(9) Mary wouldn't use Rabin's probabilistic algorithm.

Here (5) is an implicated premise of (4b), and (6) is an implicated conclusion. Other premises and conclusions are perfectly plausible in a conversational context, but are less clearly intended by the speaker:

(10) People who refuse to use plausibility arguments do not believe them to be as good as a mathematical proof.
(11) Mary only accepts traditional proofs as constituting mathematical knowledge.

It is less clear that (8) ought to be described as an implicature. Sperber and Wilson might describe it as made *weakly manifest* by Mary's statement, concluding that 'there may be no cut-off point between assumptions strongly backed by the speaker, and assumptions derived from the utterance but on the hearer's sole responsibility' (Sperber & Wilson, 1986/1995 p.199). They thus open out the concept of implicature to include a spectrum of weak to strong, definitely intended and less definitely intended, conclusions.

It is important to note that relevance theory puts forward a view of communication as not simply a case of retrieving a message or acquiring knowledge. The cognitive effects mentioned can be thought of in terms of mutual manifestness, as a weaker and more subtle notion than that of mutual knowledge. For an assumption to be manifest, it need not be made conscious or entertained; it need only be made available to the hearer, a conclusion that might potentially be drawn. An individual's cognitive environment is the set of facts manifest to him, the set of potentially entertainable ideas, and this is described as being 'a function of his physical environment and his cognitive abilities' (Sperber & Wilson, 1986/1995 p.39). A speaker can hope, by ostensively modifying the hearer's physical environment through speech or other actions, to make certain assumptions more or less available to him, and to make a certain set of assumptions mutually manifest to the two of them.

Grice famously drew a distinction between what he called non-natural meaning (meaning$_{NN}$) and showing, distinguishing the kind of utterance interpretation that requires consideration of intentions, and that which can occur without reference to a speaker's intentions. Grice claimed that while an utterance of "St John is dead," *means*$_{NN}$ that that is the case, Herod's presentation of St John's head on a spike is rather an example of *showing*; after all, "Salome can infer that St. John the Baptist is dead solely on the strength of the evidence presented, and independent of any intentions Herod has in presenting her with his head" (Grice, 1989 p.218). Relevance theory challenged that hard distinction, accepting that in many cases, multiple types of evidence might come into play. Because of the presumption of optimal relevance, instances where a speaker simply adds some layer of ostension to something that already was evident can nonetheless encourage a person to do more interpretive work, and to expect additional cognitive effects. 'In many non-verbal cases (e.g. pointing to one's empty glass, failing to respond to a question), use of an ostensive stimulus merely adds an extra layer of intention recognition to a basic layer of information that the audience might have picked up anyway… in order to satisfy the presumption of relevance conveyed by an ostensive stimulus, the audience may have to draw stronger conclusions than would otherwise have been warranted' (Wilson & Sperber, 2005 p.260). Tim Wharton (2008) refined this with a range of examples that showed that the precise nature of the evidence provided might shift the balance of intention-recognising and non-intention-recognising interpretative work, making the case for a continuum of showing and meaning that includes multiple modes of linguistic and non-linguistic intentional communication. This makes it an ideal approach to use in analysing multimodal communication, as has been observed by Charles Forceville (2014).

Thinking in terms of showing and meaning can allow us to make sense of situations in which speakers appear to do little more than *deliberately act in front of* one another, and see them as nonetheless a form of communication. For example, a person interacting with a diagram deliberately and openly in front of collaborators is something that we can understand as *showing* the room a particular way of engaging with the material.

This theory gives us certain questions that we can ask in order to understand how in each case the people understand one another.

- What effects are being sought? What is the question at hand, and what kinds of assumptions are the interlocutors likely to be entertaining, that some utterance can serve to either strengthen or weaken?

- What is mutually manifest to the interlocutors? What set of resources is mutually available to them?

- Where is ostension being used to shift this mutually manifest landscape, to influence and adjust how manifest different aspects of that landscape are?

With these questions in hand, we have a clearer picture of how to make sense of what happens in the example discussed above. In Figure 68, C is looking for some clarification of B's definition with some set of ideas in mind about what is going on, and F's modification of the environment should make some set of assumptions more or less available to C, operating as efficiently as possible to be relevant. The diagram and preceding discussion are mutually manifest but F's ostensive behaviour in giving the 'guided tour' of the diagram directs C to attend to certain features of the diagram, *showing* a way of looking at it by increasing the manifestness of certain vertices and properties of the diagram such that certain assumptions about the way that the diagram and definition relate become more manifest to C.

Corollary 6.4. Metacognitive acquaintance

Opinions differ on the precise nature of that which is shared, whether it is a whole array of propositions that add up to a subtle mental state (Wilson, 2017) or a kind of inherently 'analogue mental state' that could not be digitised in this way (Pignocchi, 2018). Either characterisation recognises that one of the problems that relevance theory takes on, and offers insight into, is that of accounting for very subtle communicative situations in which what is conveyed is difficult to summarise. A well-known example from *Relevance* is as follows, a description of a situation in which ostensive behaviour in an environment serves to share an impression that could not be encapsulated in a sentence.

> Mary and Peter are newly arrived at the seaside. She opens the window overlooking the sea and sniffs appreciatively and ostensively. When Peter follows suit, there is no one particular good thing that comes to his attention: the air smells fresh, fresher than it did in town, it reminds him of their previous holidays, he can smell the sea, seaweed, ozone, fish; all sorts of pleasant things come to mind, and while, because her sniff was appreciative, he is reasonably safe in assuming that she must have intended him to notice at least some of them, he is unlikely to be able to pin her intentions down any further. Is there any reason to assume that her intentions were more specific? Is there a plausible answer, in the form of an explicit linguistic paraphrase, to the question, what does she mean? Could she have achieved the same communicative effect by speaking? Clearly not. (Sperber & Wilson, 1986/1995 p.55)

While Peter is recognising Mary's informative intention in this example, this is clearly a far cry from decoding an encoded proposition.

In a later paper, Sperber and Wilson describe the possible outcome of an interaction thus: '…as a result of the communicator's behaviour, the addressee may experience a certain change in his cognitive environment, and identify this change, or part of it, as something the communicator intended to cause in him and to have him recognise as what she intended to communicate. In this case, what is needed to identify the array is neither enumeration nor description, but merely metacognitive acquaintance' (Sperber & Wilson, 2015 p.140). The concept of 'metacognitive acquaintance' loosens up the kind of effect that might be identified; rather than enumerating new conclusions, we might say that what is shared is something like a way of seeing, perceiving or thinking. In the example above, we saw F sharing something that could well be described as a *way of seeing* a diagram, and while in one sense it conveys quite a definite idea (that the diagram fit the definition proposed), the change sought in C's cognitive environment was quite a subtle one, a shift in C's distribution of attention to the vertices and structure of the diagram. The sharing of such subtle ways of seeing can be seen even more pronouncedly in the continuation of the excerpt in Figure 69.

14	A:	I hav- I have a problem.
15	B:	oerr gahd
16	B:	OK what's your problem ((laughing))
17	E:	((laughing)) not again
18	B:	((speaking through laughter)) here comes another one of A's counter-examples
19	A:	it just seems… ((points at diagram on board)) isn't… wait, what happened to this… so
20	A:	so… *z* was switching. … [but what if it just continued this way and then switch ((follows path))
21	B:	[Yuh. I-
22	A:	but then a couple of things switched all at the- ((mimes two coming down from above))
23	B:	so if there was
24	A:	°s that a problem?°
25	B:	if there was… ((walks over to board, A steps away))
26	B:	the problem is not for *x*
27	A:	O↑H↓
28	D:	((unintelligible)) rectangle
29	B:	if there exists a: swi- a ↑vertex↓ from which three vertices are switching… then it's ↑bad↓. then
30	the whole graph's bad.	
31	A:	The whole graph's bad?
32	B:	the- the tree cannot be represented if the vertex ((trails off))

Figure 69. Excerpt

This is quite a subtle exchange, and what is most interesting about it is that, nonetheless, it is clear that on line 27 A experiences a moment of quite genuine revelation. Our task then is to understand just how this revelation was in fact reached.

On line 14, A responds to the previous exchange (in which B proposed a definition, which was further explicated by F) by saying, 'I have a problem'. With B's proposed definition hanging in the air, it is easy to conclude that A's 'problem' is in some way a problem with that proposed definition, and in relevance theoretic terms we might say that A is intending to weaken the assumption of the other participants that B's definition is accepted, or correct. On lines 20 and 22, A goes on to sketch what B describes as 'another… counter-example', or perhaps it should be called a counter-counter-example; this is an example that A thinks of as *contradicting* B's definition, which was a description of the original counter-example. To weaken the assumption that B's definition is correct, A is describing a case that fits B's definition and yet would nonetheless be representable.

If you are reading through lines 20 and 22 and finding the description a little hazy, that is not because you, the reader, aren't grasping it, or the mathematics, properly. What is interesting about this excerpt is that A really does not give this description of the counter-counter-example with much explicitness. The description is impressionistic, incomplete, and yet B appears to understand something important from it.

We seem to see A sketch a situation in which a lot of 'switching' happens a long way down a path, asking whether this would be a 'problem'. Really, much of the description is happening in A's physical interaction with the diagram, following a string of lines up and then miming two coming down from above (Figure 70).

See 3.3 Whiteboard animation of lines 20 – 22 on the accompanying CD: So z was switching…

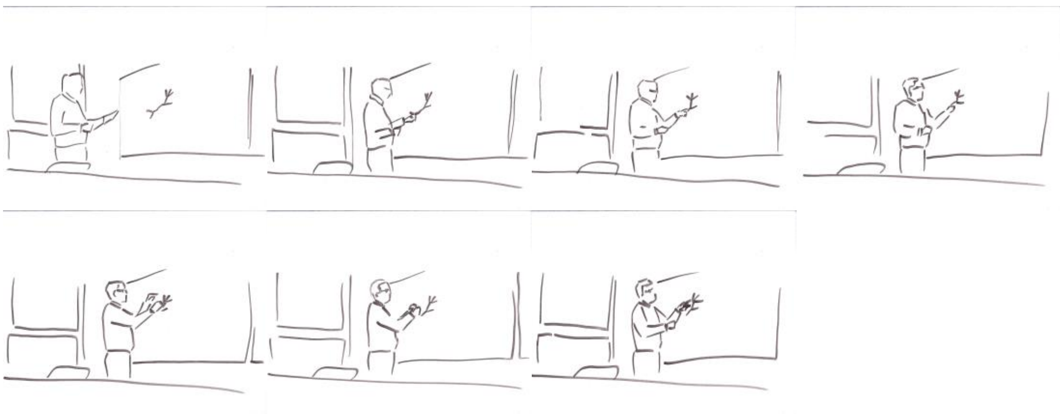


Figure 70. A's description

I recall the important shift discussed above, when the group moved from describing the diagrams in terms of switching *paths* to switching *vertices*. One thing that we can glean from carefully examining this description is that in some aspects A is still exhibiting a path-based understanding of the diagram; A asks the question 'what if it just continued this way…', following a path upward through various vertices, and continuing 'but a couple of things switched all at the-' while miming two arrows coming down from above; these arrows would produce two switching paths, true, but only one switching vertex, so we can see that there is some sense in which A's understanding and B's definition are not quite in alignment. Connectedly, A's description—where A begins following the path up, the implied question about the effect of additional switches further up the path—seems to 'count' switching paths/vertices relative to vertex x (see Figure 71). This was the vertex used as a 'base' to count from when the group were originally discussing switching paths, and for this reason was circled on the board at the time of the excerpt. However, an important aspect of B's description and the shift from paths to vertices was that these did not extend from any particular origin to endpoint, and rather the question was whether *any* given vertex had three switching vertices 'downstream' from it.

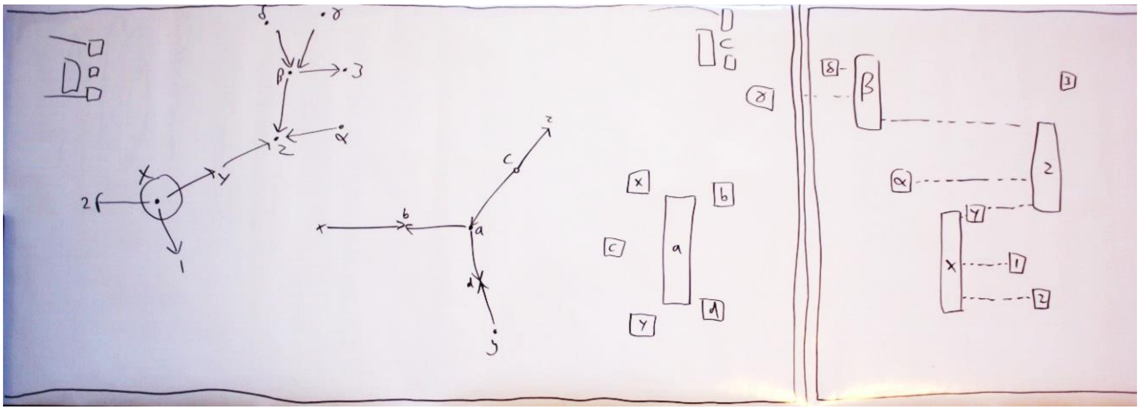


Figure 71. Board notes at the time of the excerpt. Original in colour.

With this in mind, then, we are in a position to understand B's response to A on line 26. Watching A's movements and listening to the description put forward, B is able to perceive that some of A's assumptions are not in alignment with the way the B is looking at the problem: A to some extent is still thinking in terms of paths, and 'counting' them relative to *x*. Importantly, this misalignment is somewhat subtle and multi-faceted; it is not simply that A has missed the shift to vertex-descriptions altogether. A's worry might be summarised as follows: that simply counting the first switching vertex will mean that all kinds of switching can happen downstream from that vertex and would not be counted by the definition (and yet would still make the diagram unrepresentable), which, since a feature of B's description is that it is possible to 'count' vertices from *any* given vertex, amounts to a blending of older and newer perspectives. B perceives this blending and gives a response that simply and minimally contradicts A's assumption that counting should be happening relative to *x*: B states on line 26, 'the problem is not for *x*.'

The perspectives shared here by A and B are in no way explicitly encoded in the words that they utter. Their comments are brief and impressionistic, and A and B's awareness of one another's aims and assumptions seem to have been key in making the inferences that allowed them to share these nuanced ideas about a complex topic. It is notable that some of the key evidence turned out to be A's gestural interactions with the diagram, gestures that cannot be said in a strict sense to be *encoded*.[12] Interestingly, A seemed to communicate in lines 20 and 22 by doing something like *intentionally providing evidence* of a certain *way of seeing* the diagram, or *showing* that way of seeing to the group: simply choosing to talk through the diagram in front of B in order to share a perspective on it. This kind of intentional revelation is at the heart of the relevance theoretic view of communication; Sperber and Wilson say, '[w]hat people do when they communicate is precisely to overtly reveal something of their own mind in order to bring about such changes of mind in their audience' (Sperber & Wilson, 2015 p.140). In the above, I have argued that mathematical verbal communication, for all its

[12] This is a question given interesting treatment in "Can Pictures Have Explicatures?" (Forceville *et al.*, 2014)

specialised technicality, is still best understood in the same inferential, pragmatic terms as ordinary verbal communication.

What seems worthy of remark, and further discussion, is the extent to which in these examples that revelation of mind was happening through interaction with an external, shared resource: the diagram. Through the creative data analysis process in Chapters 3-5 I adopted a multimedia re-enactive approach to take the data on terms that resisted the textualisation of textured, multi-media situations and recognised the particularity of the configurations of person and inscription in each situation. Each case was different and these differences non-trivial, the relationship between person and inscription in each case being in fact one of the questions that the participants themselves were examining and adjusting. In this last example, in which F provided a guided tour of a diagram, communication happened through the intentional exhibition of an interaction between person and diagram, this in the service of conveying a *way of seeing*, itself a particular relationship between mind and diagram.

Exhibited interactions with inscriptions are central to mathematics, chalkboard lectures being one of its most emblematic features; the chalkboard talk has been said to be highly important to mathematical communication for being a means of laying out ideas that allow viewers to see something of how the mathematics is actually *done* (Greiffenhagen, 2014; Barany, 2010; Lane, 2016) (see also my own participant's comments in Interlude 9). Something that indicates how writing has in mathematical communication is that we have no problem accepting representations whose actual manifestations diverge from certain important, intended properties (as with the RVGs in Chapter 3, which were often so hurriedly scribbled that the rectangles did not actually line up in the way required, but were nonetheless treated by the participants as though they did; see also the variation of circles seen in a lecture in Figure 72, their imperfections surely posing no problem to the audience, luckily, since a perfect circle is a notorious drawing challenge). The key is the part that it plays in communication, so that a representation need only be recognised as intentionally playing a part in communication for relevance heuristics to come into play in helping a viewer to infer how a diagram was intended to contribute to cognitive effects.

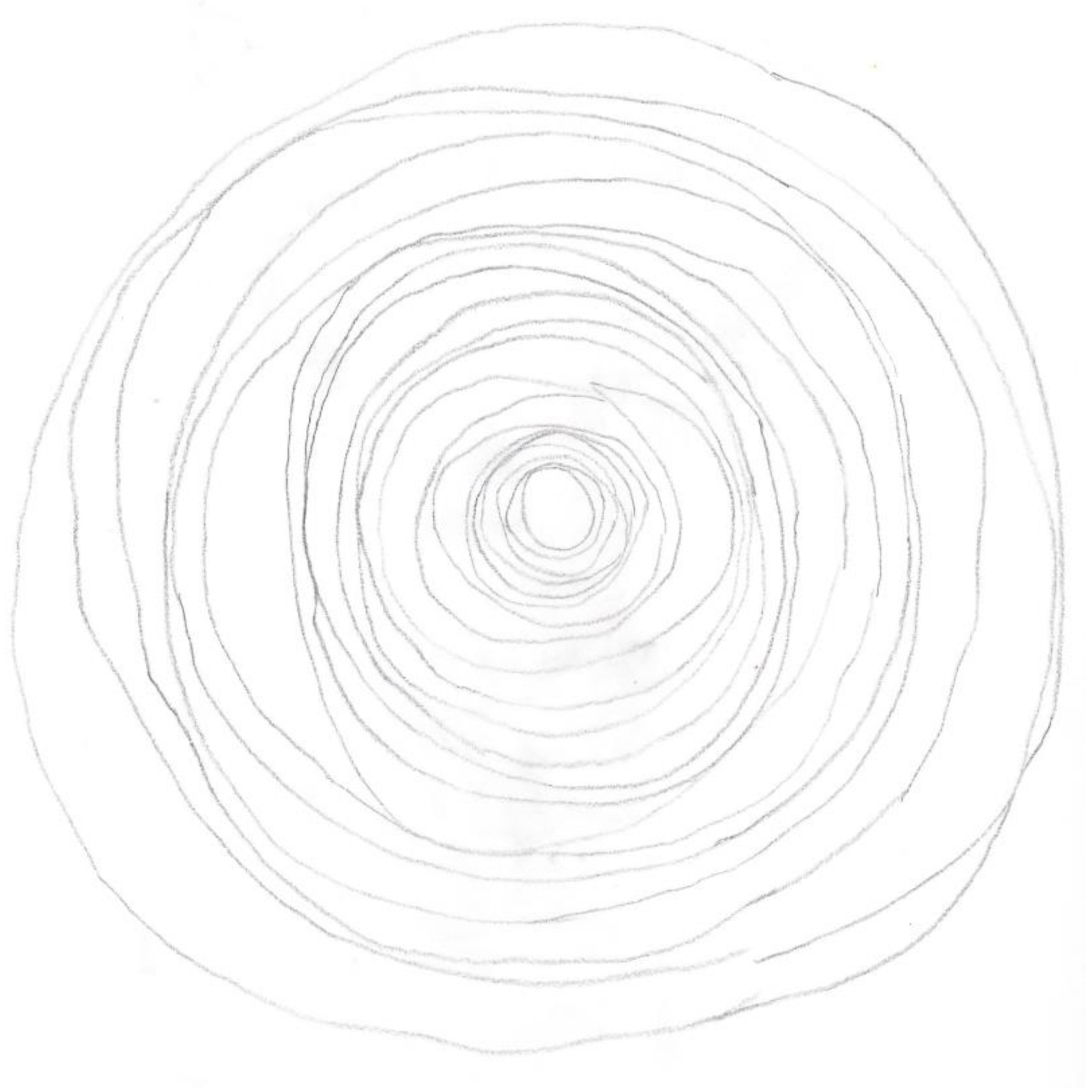


Figure 72. All of the circles drawn during a lecture on circle geometry, traced and nested inside one another.

In the next section I will give some examination to the way that mathematical writing is used outside of strictly communicative situations, to better understand the way that mathematicians work through and in dialogue with these external representations.

Remark 6.5. Thinking with inscriptions

In Chapter 5, I examined the notes written by a mathematician in the course of developing ideas for a paper. In these notes there was evidence of writing fulfilling important, interactive roles, far from simply being a means to record existing thought.

The mathematician's work was concerned with building planar graphs, curve graphs of the kind seen in Figure 73, which are combined in pairs, top and bottom, to give a new representation of elements of the Thompson group F. A key question is whether there exist any 'eye' shapes—curves mirrored top and bottom that do not go on to connect to further vertices to the right—since these can be 'cancelled' or 'reduced', making effectively no difference to the construction. In Figure 73 the mathematician is investigating 'bipartite' graphs, graphs whose vertices can be divided into two distinct groups. The mathematician begins by following a sequence of 0s and 1s to produce a corresponding set of curves, each with a different digit at each end. The intention is to see whether multiple such solutions can be found, and whether once combined the graph so produced included these reducible elements, or whether the graph is irreducible. To find the solutions the mathematician first works systematically through a representation from right to left, following the 0s and 1s to decide where the possible curves can go. The mathematician then imaginatively 'flips' one set of curves and redraws it to produce the combined version, which can then be examined in search of reducible 'eyes'. In this way, by enacting a set of simple, organised interactions with and manipulations of these representations, the mathematician is able to answer a complex question: which 5-leaf bipartite representations of members of Thompson group F are irreducible. In this example, it seems right to say that writing takes on a role that is less like a recording device and more like a means of calculation.

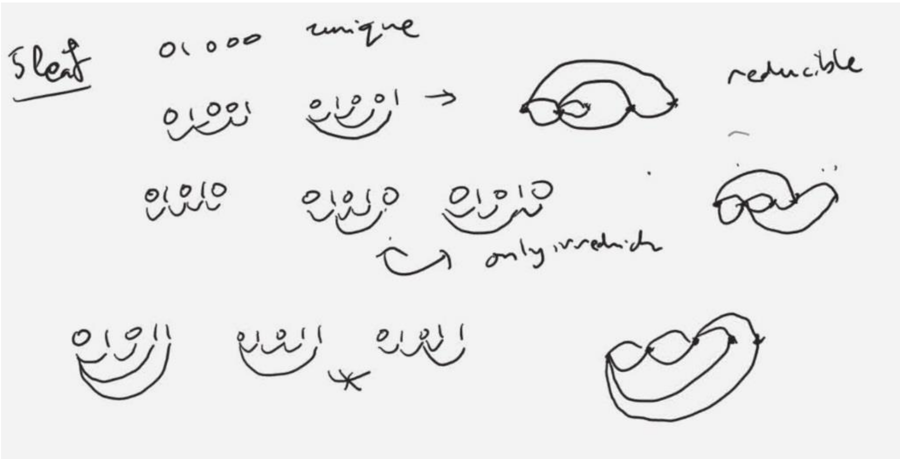


Figure 73. Excerpt from the material examined in Chapter 5

The intimate coupling of writing and thought in mathematics is an oft-noted fact (Barany, 2010; Lane, 2016). Brian Rotman describes 'scribbling/thinking' as the only action taken to fulfil the imperatives in a mathematical paper (Rotman, 2000 p.12), and argues that the irreducibility of this grouping of writing and thought undermines the philosophical explanations of mathematics given by both formalism (which focuses on the writing) and intuitionism (which focuses on the mind). In the data analysis I developed the idea of a thinking assemblage consisting of a person and environment, or a mind and inscription, in which the thinking person's interactions with external representations serves to extend cognitive capabilities.

What is key here is the *relationship* between mind and diagram, and the way that this organises action and cognition. A useful reference point here is J. J. Gibson's concept of *affordances*, a way to think about how a physical environment and an organism with its particular aims and needs relate to one another; an example might be that a coffee mug, looked at by a human, can be perceived as *affording* both the containment of liquid, and comfortable holding by the handle (Gibson, 1966; Gibson, 2014). Gibson's most famous definition is as follows: 'The affordances of the environment are what it offers the animal, what it provides or furnishes, either for good or ill. […] It implies the complementarity of the animal and the environment' (Gibson, 2014 p.179). These diagrams not only *afford* certain kinds of cognitive engagements in their static form, they can be actively annotated and manipulated to offer yet more useful affordances. The usefulness of these kinds of inscriptions is best understood in terms of the mind-diagram relationship, and also the deeply interactive process that shapes the marks on the page.

Organised engagements with writing can play a significant part in mathematical reasoning. The benefits of writing for cognition, in particular working with notation, have been recognised in the domain of mathematics (MacColl, 1880; Moktefi, 2017; Giardino, 2013; Villani, 2012). Hugh MacColl in 1880 described symbolic notations as 'enabling any ordinary mind to obtain by simple mechanical processes results which would be beyond the reach of the strongest intellect if left entirely to its own resources' (MacColl, 1880 p.45). Valeria Giardino describes both diagrams and notation as making use of a manipulative imagination based in our experience with the physical world, arguing for a 'moderate' application of the embodied mind perspective to explain the usefulness of well-designed notations and diagrams (Giardino, 2018).[13] The mind is itself shaped by the interactions; this 'manipulative imagination', Giardino says, is improved and strengthened by the practice. This kind of characterisation of cognitive tasks as completed by entangled mind-environment combinations is the province of situated cognition thinking.

Lemma 6.6. Situated cognition

In Clark and Chalmers' (1998) example of Inga and Otto, two people, one a patient of Alzheimer's disease, want to visit a museum and need to know the location. Inga remembers the address of the museum, whereas Otto keeps notebooks with him at all times which serve as a kind of extended memory. The kind of move made by situated cognition thinkers is this: to see these external strategies not as essentially separate from the work of the mind, but as importantly bound up with and shaping of cognition as we know it. This notion will help us to understand what 'thinking with inscriptions' can mean.

In *Cognition in the Wild* (1995), Edwin Hutchins explains how the navigation team offloads all kinds of difficult computations on their instruments, referencing Herbert Simon's famous line, a tentative suggestion for a way to characterise theorem proving: 'solving a problem simply means representing

[13] For the latter point, Giardino cites the Landy and Goldstone study on grouping in mathematical expressions (Landy & Goldstone, 2007), which shows that grouping elements together with (meaningless) visual embellishments can inhibit or enhance correct application of the mathematical order of operations in evaluating expressions; their subsequent discussion develops an account of symbolic reasoning as depending on the perception and manipulation of notational formalisms (Landy *et al.*, 2014). Giardino's perspective rests on these external representations having certain *affordances,* making the point that apparently subtly different notations can have quite profound effects on this manipulative imagination work (for example, whether an under-strand in knot theory is indicated by an occlusion or by dots, each of which will favour different manipulative imaginings).

it so as to make the solution transparent' (Simon, 1981 p.153). He compares these tools to the slide rule in a way that is very pertinent to this project.

> The slide rule is one of the best examples of this principle. Logarithms map multiplication and division onto addition and subtraction. The logarithmic scale maps logarithmic magnitudes onto physical space. The slide rule spatially juxtaposes logarithmic scales and implements addition and subtraction of stretches of space that represent logarithmic magnitudes. In this way, multiplication and division are implemented as simple additions and subtractions of spatial displacements. The tasks facing the tool user are in the domain of scale-alignment operations, but the computations achieved are in the domain of mathematics. (Hutchins, 1995 p.171)

This characterisation of the slide rule, in which a simple engagement with the environment is imbued with the ability to answer significantly complex questions, proves extremely helpful when considering the ways in which mathematicians make use of their diagrams and notation.

A key point of the book is that the team, in coordination, achieves what he considers to be cognitive tasks that no single member could achieve alone, and it is by virtue of these structures and their participation within them that these tasks are possible. Modern mathematics has become enormously diverse and specialised, and specialisation is such that even a mathematician collaborating on a paper may not fully understand the sections contributed by the other collaborators. Hutchins makes a relevant observation about an exchange between recorder and plotter: that sense can only be made of the very minimal queries and answers traded between them with reference to the whole structure that surrounds it, from the chart itself to the interlocutors' mutual awareness of the equipment and their preceding interactions. Hutchins observes, '[m]eanings seem to be in the messages only when the structures with which the message must be brought into coordination are already reliably in place and taken for granted' (Hutchins, 1995 p.238). We may understand a scratchy diagram in a set of notes, an equation, a lemma or even an entire paper as meaningful only inasmuch as it is situated in the whole structure of modern mathematics.

In Figure 73, the mathematician's manipulations of the diagrams are imbued with power in a way reminiscent of Hutchins' example of the slide rule, in which a simple engagement with the environment is imbued by a writing technology with the ability to answer complex questions. Hutchins observes that 'multiplication and division are implemented as simple additions and subtractions of spatial displacements. The tasks facing the tool user are in the domain of scale-alignment operations, but the computations achieved are in the domain of mathematics' (Hutchins, 1995 p.171). In this way, a simple engagement with the environment is imbued with the ability to answer significantly complex questions. The slide rule by its embedding within a navigational system takes part in precise and complex operations; these simple marks are written in a carefully organised way and embedded in a system of representations familiar to a highly skilled thinker such that that mathematician can wield them in operations on the domain of mathematical knowledge.

In the case of Figure 73, the mathematician is able to enact a set of simple, organised interactions with the scribbled representations, and by labelling them with 0s and 1s, to make immediately evident certain properties of the graph and the relationships of the endpoints. Then it is possible to enact another very simple engagement, following the sequences of 0s and 1s to build graphs with which the first can be paired, and thus to construct a bipartite 5-leaf representation of a member of the Thompson group F. Finally, the mathematician can scan the representation constructed for certain salient visible features, and in this way to answer the question of which of these are irreducible. By their embedding in a rich and complex system, then, these representations allow the mathematician to enact very simple cognitive tasks and in so doing reach complex mathematical outcomes.

The kind of manipulation seen above relatively straightforwardly displays mind-environment interactions that achieve a cognitive task. This task might indeed be achieved by other means, including purely 'mental' ones. However the ways in which this method facilitates and streamlines cognitive work, the affordances of the means of representation and the ability developed in the perceiving mind, bring the task down to quite different proportions and so bring a vista of new manipulations into reach. Indeed the very nature of the project—deploying a new and inventive means of building representations of the Thompson group *F*—demonstrates the enormous significance of developing a new system of inscriptions to connect with a body of work.

Corollary 6.7. Perception

In Figure 73, it is important to recognise the refined nature of the mathematician's interaction with the representation. The mathematician is enacting the simple task of seeking an 'eye' in the combined diagrams, but the path that brought the mathematician to that point was shaped by engagements with the mathematical community, with the other papers published and mathematicians working on the Thompson groups, and the work of other mathematicians on cancelling 'carets' (such as Cannon *et al.*, 1996), which the mathematician might have read or seen in a talk. In this way, the mathematician is able, through a judicious, deeply informed decision-making process, to carry out a simple perceptual task—seeking 'eyes'—and imbue it with deep mathematical meaning, as an indicator of an important kind of equivalence. The mathematician's lifelong engagements with that community served to shape the decision to seek out those 'eyes', to develop the system for perceiving equivalences between certain types of form and conceptual entities from the body of mathematical knowledge that allows this simple perceptual task to answer complex questions. In this way, simple engagements with writing can be linked to the complex work of other mathematicians, imbuing these curved lines with what we think of as mathematical significance.

Situated cognition thinking has an important component in the theory of perception. Perception is, according to philosopher Alva Noë, 'not something that happens to us, or in us. It is something we do.' (Noë, 2004). An internalist, representationalist view might see cognition as a process of taking in perceptual information, encoding it in some internal representation, reasoning about it, and then translating this into action (where the internal reasoning process may be seen as anything from reflexive to conscious), whereas situated cognition thinkers talk about active, responsive perception that is informed by and feeds in to expected action. To challenge the internalist/representationalist view, reference is often made to current advances in robotics which have made it clear that successful stable walking of the kind that humans and dogs are capable of requires action to be constantly adjusted in on-line perception-action tunings. The refining of imaginative engagements with diagrams described above by Giardino (2018) might be explained in terms of adjustment of perception and distinguishing of different situations and responding accordingly (Noë, 2004; Dreyfus, 2002), coming to see an inscription in a particular way, assigning divisions and meaning in a way that accords with assignations used by others in the community of experts and so connecting one set of mind-environment interactions with a vast, highly developed body of work.

Writing, importantly, is something that very often allows us to develop and refine our perception in useful ways. Figure 74 shows another excerpt from the notes in Chapter 5. In this excerpt, two distinct types of diagram are identified. The mathematician draws two sets of pictures to distinguish the two cases, indicating by use of dots and hints at undrawn extensions that these are intended to exemplify a whole range of different possible diagrams that share those particular characteristics. Numbers have been given to the two cases, the number (1) seeming to be added as something of an afterthought once the describing sentence had been written, perhaps when the mathematician began the following case.

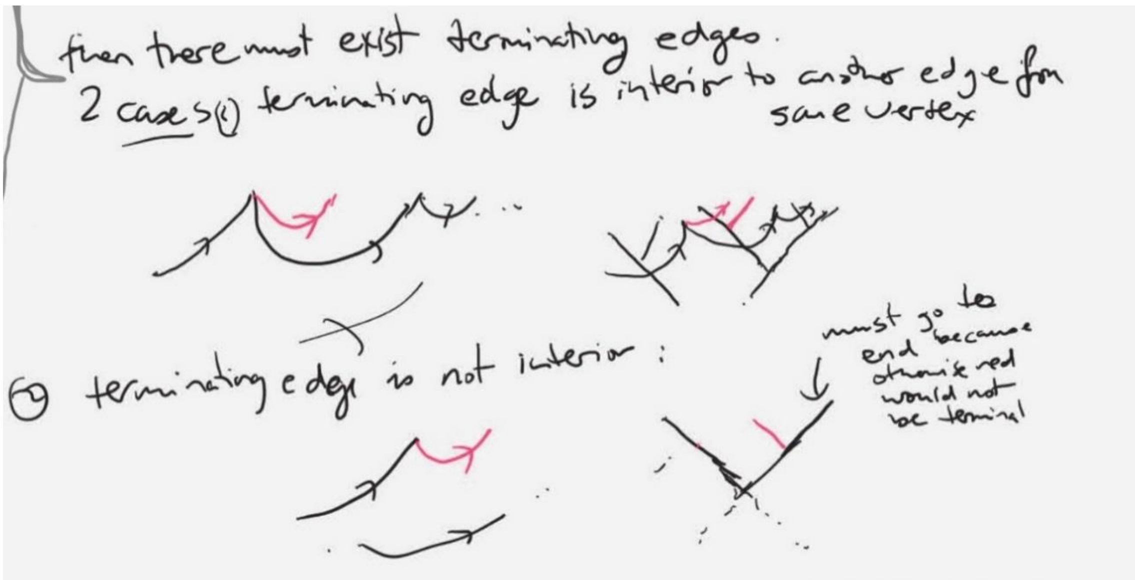


Figure 74. Two cases delineated in Chapter 5. Original in colour.

These drawings are the inscriptions made by a mathematician in the course of outlining a way of looking at the features of a diagram such that it becomes clear that they will all fit into one of two categories, which can then be treated in two identifiable ways. This adds up to a convincing argument that every feature of every diagram can be treated in one of these ways, and so that the treatment can always be given in this way. This is not simply a mathematician enacting certain mechanical verifications; this is a mathematician using writing to identify and develop a way of seeing that will be instrumental in achieving, and later in sharing, conviction that the proof is correct. The mathematician uses colour to highlight certain relevant features, a decision that can help that mathematician to mentally highlight (perceive certain features) of other possible diagrams. The mathematician also corrects certain details in the course of writing; the diagram on the bottom right is drawn to extend rightward as well as leftward, but the mathematician has appended a note that contradicts this detail, adding 'must go to end because otherwise red would not be terminal.' In the course of drawing and writing, the mathematician is developing and refining a way of looking at possible diagrams, of giving the cases clear labels, of recognising the shape that the diagram section ought to have in (2); the notes and colour changes help the mathematician to become adept at distinguishing certain situations, in perceiving certain features as the relevant, important ones. Also, by dividing up and organising the cases and managing them in this way the mathematician can become convinced that all cases are covered. The simple act of drawing out the two possibilities and labelling them as such, as the two options, allows the mathematician to examine and assess the reasoning, to evaluate its organisation, to decide whether it is correct—and how to lay it out for another person. Later, the mathematician will produce diagrams very like these, accompanied by a much expanded technical explanation, in a published paper, written to help others to develop this ability to distinguish situations.

To clarify the kind of impact that handwritten notes might have on cognitive work, we might briefly consider the cognitive-psychological concept of *cognitive load,* and the kinds of factors that are said to impact upon it. Cognitive load refers to the working memory resources used at a particular time, and a question asked in its study is which of these might be reduced by the adjustment of, for example, *design* of resources like learning materials. Research on working memory focuses not just on numbers of items to be remembered, but also what to emphasise and what to suppress (Baddeley, 2010; Engle & Kane, 2004; Engle, 2002; Barrouillet *et al.*, 2004; Cowan, 2016). The way that notes are laid out, the spatial arrangement, can also help us to not have to think about which parts are important and which are associated. The aesthetic and spatial resources available when using a piece of paper are many, such as tone of line, colour, placement, and so on. These aspects begin to organise the material, to shape perception of it. As well as the content of the notation, a mathematician is interested in developing a way of looking at it, of categorising, of perceiving the important parts, of recognising different types of situation, and knowing what might be done in each. As the distinctions become more habitual, they become simply part of the expert's experience and knowledge of the material.

Another aspect of reasoning with writing is that it has the potential to allow a person to make use of cognitive resources that are associated with taking part in a dialogue. What I mean by this relates to the confirmation bias that has been noted by many researchers, and that Hugo Mercier and Dan Sperber argue is a result of the essentially argumentative origins of human reasoning (Mercier & Sperber, 2011). If we learned to reason so as to be able to argue a certain position with an interlocutor, it is not so surprising that, when engaged in the kind of verbal tasks used to test reasoning, we are broadly speaking disposed to pay closer attention to evidence and arguments that support our position rather than contradict it. Mercier and Sperber say that it is possible for reasoners to somewhat overcome this bias inasmuch as they are able 'to distance themselves from their own opinion, to consider alternatives and thereby become more objective' (Mercier & Sperber, 2011 p.72); externalising ideas and then examining them might help a writer to see them as another might see them, to consider counter-examples, to look for flaws as our mathematician does. Of course the most reliable mechanism for producing good reasoning is, for Mercier and Sperber, reasoning in a community; 'the achievements [of human thought] are all collective and result from interactions over many generations' (Mercier & Sperber, 2011 p.72), and this is where the mathematician above moves next with this project: publication and peer review of the work. At this early stage, though, the mathematician benefits from constructing representations and critiquing them, noticing flaws and thus developing the definitions needed for the formal presentation and the ability that makes this expert able to reason quickly and confidently with the representations.

These examples of mind-inscription interactions demonstrate the important back-and-forth taking place as a mathematician works through a problem. Recognising these external representations as manifesting thinking in action help us to understand how communication sometimes works in informal settings of mathematical work, where, for example, collaborators are at work around a whiteboard.

Proposition 6.8. Thinking out loud

In *Cognition in the Wild*, Hutchins describes a navigation plotter, the member of the navigation team responsible for plotting the course, using his finger to locate a bearing in the bearing record book. In the instance described, the finger is used a tool in directing his own attention to locate the correct entry in the book, which serves the team as an external memory. As well as a part of his private cognitive process, this pointing also has a communicative role, since as an external gesture it is visible and available to the rest of the team. Hutchins notes that '[s]ome kinds of media support this kind of externalization of function better than others' (Hutchins, 1995 p.236); I am immediately reminded of the chalkboards and whiteboards that line the offices, corridors and classrooms of mathematics departments, external sites where the public writing acts of workings-out can be made available to colleagues and students.

With the above discussion of the external, material aspects of thinking in mind, we find ourselves with a new way to conceive of the metacognitive acquaintance described by Sperber and Wilson. If we understand that external work with representations as very genuinely representing a part of cognition then we have a way to think of mathematicians' work at the chalk- or whiteboard as intentionally exhibiting cognition in a very direct way, providing colleagues with a fascinating kind of access to the state of their thinking at the time. Thus we can see in a new light Sperber and Wilson's suggestion that '[w]hat people do when they communicate is precisely to overtly reveal something of their own mind in order to bring about such changes of mind in their audience' (Sperber & Wilson, 2015 p.140).

Such a perspective will even help us to make sense of some of that vastly underdetermined communication in Chapter 3, since we can think of this as people *seeing* one another think. Let us look again at the very first example (Figure 68). F intentionally interacts with the diagram in a way that *overtly reveals something of F's mind* in order to bring about changes in the mind of C. The guided tour is a way of giving C a kind of access to F's way of seeing the problem, the way that F is perceiving the relevant and less relevant features of the diagram, the *ability to distinguish* certain pertinent situations that F has developed. If C does come to perceive the diagram as F does, then the outcome is that C will see it as fulfilling that description. Another, more important outcome is that C will know which features should be paid attention to ascertain whether another diagram satisfies B's condition, or not. With all of this in mind we can begin to see this kind of mathematical writing, the informal, handwritten, 'backstage' kind, not as a simple tool used to record finished thought but as a technology that is deeply entangled in social and cognitive doings.

Alessandro Pignocchi, writing about art as communication, puts forward a reading of relevance theory that begins with a criticism of Sperber and Wilson's acceptance of the idea that human thinking is language-like. Instead Pignocchi emphasises the possibility of analogue mental states, argued for by proponents of grounded cognition (Barsalou, 2008), that do not exist in the form of language but of simulations of experience.

The question of how Theory of Mind abilities are acquired has been given a variety of treatments, and at this point it is worth mentioning the simulation explanation argued for by Vittorio Gallese (Gallese, 2007). Gallese's account rests on '[e]mbodied simulation and the mirror neuron system underpinning it provid[ing] the means to share communicative intentions' (Gallese, 2007), as a mind, presented with some evidence about what another mind might be experiencing, simulates for itself that experience in order to understand. He suggests the existence of a continuum between attribution of intention and simulation of action through mirror neurons, simulation that helps a hearer to generate beliefs about what a speaker was intending to do or say by drawing upon their own experience of similar actions. Thinking in these terms gives us a possible perspective on how even very subtle actions occurring as a dialogue between person and environment might serve as subtle aids to an observer to simulate that other person's understanding.

In the course of this chapter, I have made the case that in the informal, improvised uses of mathematical writing seen in communication and reasoning in the 'back' end of mathematics, mathematicians engage in processes of 'thinking out loud' that allow them to streamline and share cognitive work. These we can make sense of using relevance theory and situated cognition, and in so doing, see how mathematical work is not so distant from our other, everyday situated practices. Mathematicians use material and social resources in a rich and nuanced way in the course of manipulating and developing sophisticated ideas, using their ability to 'see' one another's thinking to work in cognitive teams, and using external representations to make complex ideas tractable. Mathematical writing in the 'back' end of mathematics is so knitted into our interactions and thought as to appear quite alive, constantly evolving and in turn shaping the writer's next move. What happens as mathematical writing becomes more formal is another question, and that will be the subject of the next chapter. What about when the writing mathematician is taken out of the equation, so to speak, and the mathematical text becomes something that has to stand alone?

Interview with subject J

00.05.23.000

It's interesting that you mention working from books—I guess I don't really know whether those are written in a different style than papers, when something's been turned into book form.

J: It depends on the age of the book, as well. Old books are harder, I find. New books seem to be catered to people with a short attention span and I find that easier ((laughs)) short sentences

That's funny Can you explain what an old book might be like, then?

J: Longer sentences, more long-winded ways of saying things, rather than getting to the point.

More prose?

J: Um… *long* prose. And sentences which are three or four lines long rather than, say, one and a half lines or something, that kind of thing. And I think also with maths textbooks especially the notation conventions have changed, not significantly but enough to make it look a bit- 'Ooh, I'm not familiar with that,' when you first look at different textbooks. So that's something else. I can still read it, it just takes me a bit longer and I'm not as comfortable.

And when you say old books, how old do you mean?

J: Mm, more than about 20 years, 30 years, that kind of thing.

0.06.02.000

C: ((in a teasing tone)) Yeah A thought a 1962 article would help us

A: Yeah… yeah I did. Until I got to this sentence, and I read: ((raises both index fingers)) an equivalent but more transparent formulation of the problem is obtained if we take what is known in algebraic topology—uh oh ((all laugh))—as the one-dimensional skeleton ((shakes hands in acquiescence)) of the *nerf* of the family. ((room erupts in laughter)) The *nerf*.

D: How's nerf spelled?

A: n-e-r-f.

F: huh

A: The *nerf*. I dunno but I asked one of my students recently what does it mean to *nerf*, you know, *Thor* in Marvel's clandestine champions. And he says, you know, that means that it's an overpowered character and they nerf it now they- they reduce it ((unintelligible)). BUT, I don't think that's related to this 1961 usage of the term from algebraic topology ((room erupts in laughter again))

O: I used books from the 1980s on [a project], and papers from the 1970s. […] I think it was the same author. […]

And using papers from the 1970s, was there any kind of- was that hard to read?

O: Yes. That's why I used the book. I think it was still put in LaTeX and things like that, but it- it hadn't been, sort of, that well explained. Well, it might have been well explained at the time, but it- it was using sort of older- maybe older ideas and things, so different app- ideas which might have been further developed now. And I think the author of those papers had then written books, and then people had written books on his books, things that were published in the 90s.

Work site 7

Work site 7 was the shared office space of a researcher at a UK university.

Figure 75. Office space shared by three researchers. Original in colour.

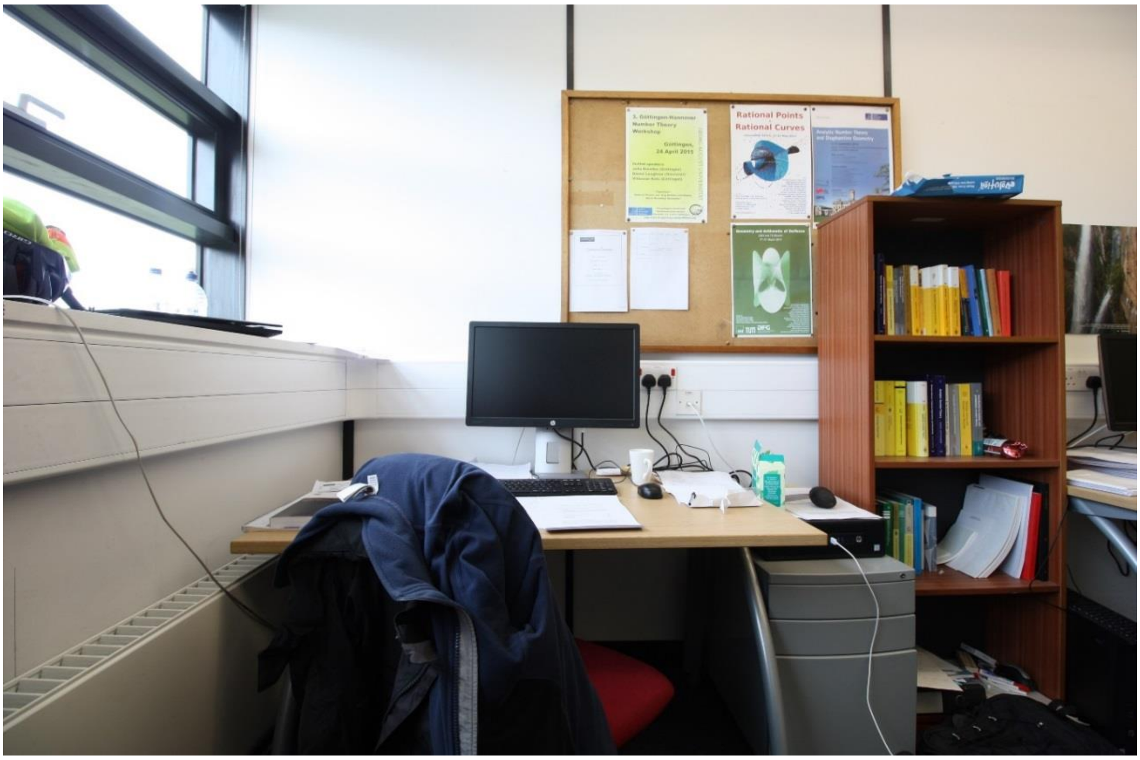


Figure 76. Workspace used by my participant, with reference books and computer and pad of paper for notes. Orginal in colour.

7. Moving toward the 'front' end of mathematics

Formal mathematical writing would appear to be quite a different beast from the flexible, changing inscriptions seen in the 'back' end of mathematics. The mathematical paper, certainly, appears a very stable, unchanging thing of a very different character. To begin considering this other side of mathematical writing, let us first look at a moment of transition from one mode of expression to the other.

In Figure 77 we see the beginnings of the kind of language that characterises the 'front'. It became clear in the course of this meeting that the participants saw their work as culminating in a written output, not least since the terminology ('2-bad') they chose to use even while discussing it aloud was ambiguous when spoken but unambiguous written (Chapter 3). Though the writing used during the course of the meeting was minimal and most important for the interactions that it facilitated, all signs point to a written version of the sentence being considered the outcome or goal of the work.

01 B: alright so the thing that's [bad
02 D: [OK I'm with you
03 B: … is a vertex with three switching vertices
04 C: [what
05 A: ((nods)) [or more
06 B: or [more
07 C: [a - vertex with three switching vertices?
08 F: cos like [in that with that vertex a
09 E: [((unintelligible)) right there
10 F: … b is switching for a, [d is switching for a [and c is supposed to be switching for a
11 B: ((walks over, follows path xba)) [yup ((follows path zca))[yup
12 F: but we have nowhere to put – z
13 B: right
33 A: I hav- I have a problem.
34 B: oerr gahd
35 B: OK what's your problem ((laughing))
36 E: ((laughing)) not again
37 B: ((speaking through laughter)) here comes another one of A's counter-examples
38 A: it just seems… ((points at diagram on board)) isn't… wait, what happened to this… so
39 A: so… *z* was switching. … [but what if it just continued this way and then switch ((follows path))
40 B: [Yuh. I-
41 A: but then a couple of things switched all at the- ((mimes two coming down from above))
42 B: so if there was
43 A: °s that a problem?°
44 B: if there was… ((walks over to board, A steps away))
45 B: the problem is not for *x*
46 A: O↑H↓
47 D: ((unintelligible)) rectangle
48 B: if there exists a: swi- a ↑vertex↓ from which three vertices are switching… then it's ↑bad↓. then
49 the whole graph's bad.
50 A: The whole graph's bad?
51 B: the- the tree cannot be represented if the vertex ((trails off))

Figure 77. Excerpt analysed in Chapter 3

In the rest of the excerpt in Chapter 3, we saw this definition repeated, quizzed, and restated with adjustments, a process that continued throughout the course of the meeting and seemed to elevate this revisited sentence somehow above the level of the utterances surrounding it, as a somewhat stabilised, examined, adjusted product of the group's work. Though it was never written on the board, as the definition was repeated it became familiar to the participants, an increasingly stable landmark of the group's evening's work. This stabilisation pinned down and made operational the way of seeing that the group was employing to successfully distinguish cases. We might recognise this progression by terming this a STATEMENT, a sentence put together as a stable expression of an idea with a view to recording and sharing. The outcome of the meeting could be summarised as the group's having collectively developed that ability and stabilised it by representing it in the form of a STATEMENT, which could be recorded and referred to if any member were to need a reminder and eventually play a part in making the group's work available to the mathematical community. The form that the STATEMENT took at the end of the meeting was as follows (admittedly something I have pieced together from a short back-and-forth to give a complete version, but our focus will be on the beginning):

> *If there are no k-bad vertices for k greater than or equal to three then [a certain algorithm works to convert this orienting path cover into a rectangle] (where 'k-bad' means a vertex with k switching vertices).*

Let us make a comparison with B's description on lines 01 and 03.

> *the thing that's bad … is a vertex with three switching vertices*

This sentence was already working toward the ideal of a STATEMENT, but we can see that B's description is more closely focused on the particulars of the counter-example before the group, focusing on *three* switching vertices. The group at the time certainly knew that any additional such vertices would also be a problem yet were focusing on the special case of three with this awareness implicit, taking particular cases and examining them to develop the STATEMENT. The definition is formed by linking this description with 'the thing that's bad,' a vague and foggy term whose meaning is clear to the group in the context of their shared aims and intentions, their ongoing work to define the property that makes a tree unrepresentable as an RVG. Here, the STATEMENT is situated according to a rich, mutually manifest landscape shared by the group of aims and intentions, and examples shared on the whiteboard.

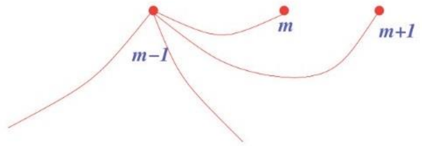
Outside of that context the STATEMENT needs a different kind of contextualisation, and this is the direction in which the final version of the STATEMENT moves. In the broader context of the aims of the group and the mathematical community the interesting question is when it is that tree diagrams *can* be represented as RVGs, and so this final version is constructed in terms of when the algorithm to do so *will* work. The ties to the specific examples have been loosened, this time the STATEMENT explicitly saying that what are excluded are cases where the number of switching vertices is greater than or equal to three. This new, improved version is longer, clunkier, and uses more words in pursuit of explicitness and accuracy. This would seem to raise some questions if we were to expect relevance theoretic principles to apply to this, quite different form of writing. Is an ostensive-inferential explanation still our best ally, and what about the appropriateness of the code model? We shall first look a little more closely the principles and concerns that shape more formal mathematical writing, and then consider our theories of communication and what they might have to say about it.

Remark 7.1. Portability and examination

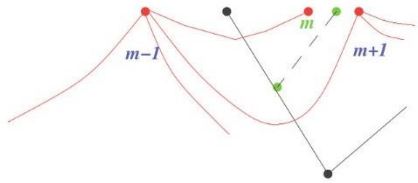
Making the formulation of a STATEMENT more explicit, even though clunky, seems geared toward making explicit some of what is implicit in the group's earlier, more impressionistic communications. Expressed as it is in natural language, which is widely known, the STATEMENT is expressed in a code that is better established than, say, a scribbled diagram using purpose-built annotations whose significance might be difficult to remember in a week or two when the group have moved on to thinking about other things. What's more, the STATEMENT has been tweaked so that the terms of its expression are more suitable for a broader audience than those present in the room that night, an adjustment that makes the formulation more stably interpretable in a mass communication context. The group has gradually moved toward something that is less heavily dependent on the cognitive environment that they shared during their meeting that night. Thinking about exactly what this shift entails will help us to get a clearer perspective on what it means to move from 'back' to 'front'.

In Chapter 5, it is possible to compare the way that a proof is written in a mathematician's own notes with the eventual presentation in a published paper. The presentation of 'Case(2)' in the paper eventually published on I's work includes far more textual expansion than was seen in the notes (compare Case (1) in Figure 74 with Case (2) in Figure 78—the numbering of the cases has changed places). While the inscriptions of the early-stage work, playing a part in a mind-inscription thinking system engaged in developing a certain distinction, operate in a context in which the mathematician has a rich understanding of the inscriptions that connects with I's broad mathematical experience and knowledge, for these inscriptions to maintain that richness in a way that is less dependent on the particular mind and set of abilities that I brings to the table, they must travel along with some baggage.

Case(2) Valence of $m-1$ is > 2. In this case there is an edge with source $m-1$ connecting it to a vertex k with $k > m$. By planarity there must be such an edge connecting $m-1$ to $m+1$. The situation near m is thus:



Removing m and its edge the resulting graph Ψ' still satisfies the conditions of 4.1.1 so there is by induction a T' with $\Psi' = \Gamma_{-}(T')$. There has to be a WS edge in T between $m-1$ and $m+1$ so we may add a WN edge to T' as in the figure below re-insert m and obtain the desired T_{-}:



where the solid edges are those of T' and the dashed edges are the ones added to obtain T_{-} and Ψ. □

Figure 78. 'Case(2)', as described in the paper. Original in colour.

These accompanying sentences describe the characteristics of and operation of 'Case(2)' in very clear terms. These sentences could be described as making explicit on the page that which was implicit in Figure 74. Alternatively, we might see this change as reducing dependence on the reader's experience and knowledge, placing specifications on the page in such a way as to simulate the mathematician's precision of distinction, using precise notation such that this is less dependent on the individual cognitive perspective and more contained within what is on the page. This is a point at which mathematical writing attempts to avoid the apparent underdetermination that often characterises utterances in everyday conversation, putting each thought in relatively more complete terms. To accommodate this desire for completeness, mathematicians develop a great many shorthands and innovative symbolic expressions to specify precise ideas. Where the notes in Figure 74 simply state that the 'Terminating edge is interior to another edge from same vertex,' the expansion in Figure 78 first specifies the valence of that vertex, then the existence of another vertex and connecting edge that this implies, then the relationship that this implies between this edge and another edge, each named by letters and expressions that contribute to placing them in the discussion. The mathematician provided a rich cognitive environment of recently entertained ideas and with this in mind was able to 'see' the completeness of identifying an interior and a not-interior case from the most minimal of scribbles. The paper-writer is obliged to specify much more fully how each thought leads to another, and uses a variety of orienting markers to place each vertex relative to one another; in this way, much more meaning is placed in the encoded form, and the encoding becomes much more specialised and particular to the terms and norms used by the mathematical community. It is one thing to convince yourself, with scribbles and a rich contextualisation in mind, but producing a text that can play a part in the mathematical community is something different.

What about the question of how it is understood by its audience? Can this more heavily defined style mean that the code model can account for the way that mathematical papers are read and understood, given a highly trained audience, or will we turn again to relevance theory?

Lemma 7.2. The code and intentions

In comparison with everyday language, that of mathematics is extremely formal, precise, logical. In theory, every term has a definition, every step appears totally logically justified. Surely, then, it would seem that understanding such a text should require nothing more than decoding: if the aim is to make everything completely explicit, then surely the reading of a text is nothing more than decoding a stimulus that is formal, progresses according to a limited number of known rules, and is fully referenced and completely well defined.

The beginning of one response to this argument is the observation that in the mathematical world these texts are seen as relatively *informal*. It might seem strange to say, but even these highly specified symbolic aspects of mathematical communication, many say, are rife with leaps and jumps in their logical progression. We might look at the difference between the kind of proof seen in a mathematical paper, and what is known as its formalisation. A traditional mathematical proof is actually relatively informal in comparison with its translation into a formal logic, with rigorously defined axioms and strict rules of transformation. Only in these formalisations, it is held, is the progression of a proof truly made explicit.[14] This seems a relevant point to bring up since it is an easy instinctive assumption to make that the logical nature of mathematical argument is watertight. Indeed it seems to be almost the pattern we use for our descriptions of strict, well-defined reasoning; and yet its practitioners regretfully admit that there are vast leaps in the logic, according to their, rather stricter standards.

While these formalisations are sometimes used to confirm a result, they are considered too unwieldy to be in any way communicative, or to approach sharing conviction (De Millo *et al.*, 1979). It is sometimes said that an informal proof is a kind of prediction that a formal proof *could* be constructed (Hamami, 2014; Larvor, 2016), but this kind of view is often vigorously resisted on the grounds that what is more important is to share ideas, to convince others, to advance practically, and while formal proofs do none of these, informal ones seem to ably fulfil those functions (Moktefi, 2017; De Millo *et al.*, 1979; Rotman, 2000; Rotman, 1998; Hersh, 1997). If such a mode of inscribing mathematical ideas exists but is not taken up as a form to be used for communication, this suggests that

[14] To give a sense of the scale of the question of total logical formalisation of mathematical propositions, I cite the somewhat infamous example of Bertrand Russell and Alfred North Whitehead's *Principia Mathematica* (1963), an intimidating multi-volume attempt to establish firm foundations for mathematics in terms of symbolic logic. In the first volume, Russell and Whitehead spend seven hundred pages establishing a basis for the proof of the validity of the proposition 1+1=2.

communication of mathematics is not, in fact, well served by an encoding-decoding mechanism alone.

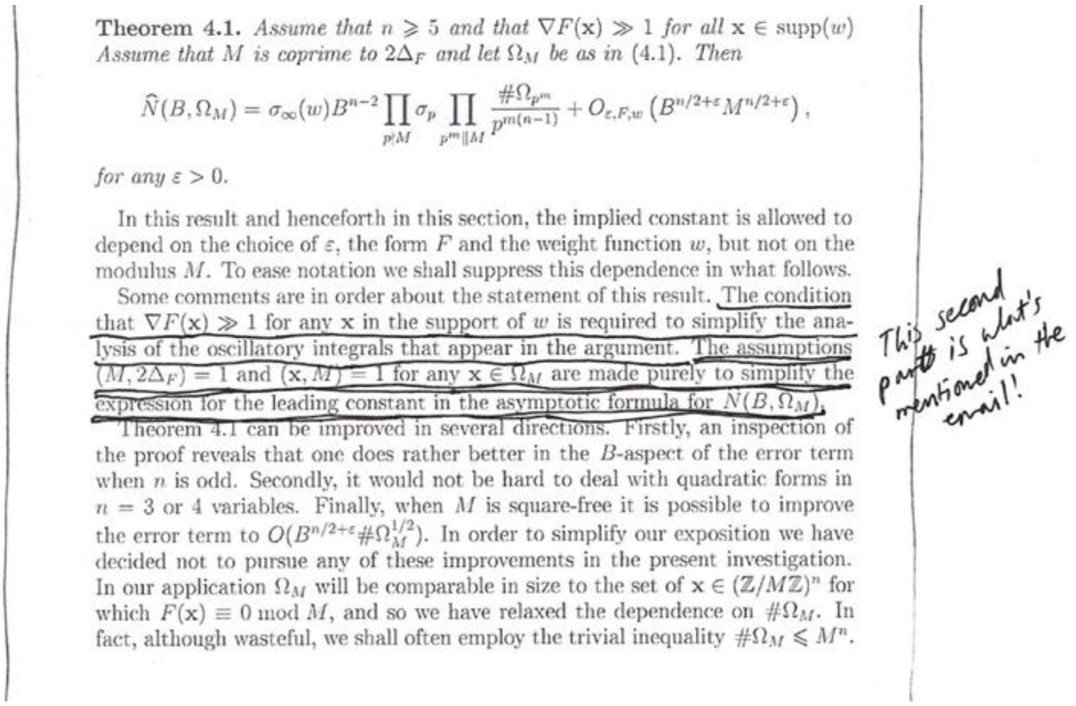
We need not necessarily adopt such stringent standards exactly, of course, and it seems reasonable to claim that a human reader would have enough in these wonderfully well-defined formulations to at least grasp without ambiguity what the situation is being described; surely from there that reader can make some well-defined inferences using knowledge of the agreed rules of mathematics to fill in any and all logical leaps, and this be sufficiently well defined all around as to still rest within the realm of decoding. In *Relevance: Communication and Cognition,* Sperber and Wilson describe what would be needed for inference to double as decoding in a situation of comprehension as follows: that two interlocutors 'share the tacit premise […] they must share the inference rule; and […] they must use that premise and that rule to the exclusion of any other tacit premise or inference rule at their disposal' (Sperber & Wilson, 1986/1995 p.14). The problem lies in that final clause. How should the reader know, Sperber and Wilson might ask, exactly which rule to use in filling in the gaps, when so many are shared by those schooled in the field? There is nothing necessary about the move from one step to another in a proof; it might take all manner of routes, and the reader is often looking for the structure of the argument, the intentions, how the logic of it all fits together.

In the *breaching experiments* in Chapter 3 and Chapter 5, I experimented with making drawings using constrained elements that followed certain rules, and pushing intentionality to the margins or the spaces in between. Even in the latter experiment where the images themselves were entirely predetermined there seemed to be plenty of scope for the apprehension of an intentional system, for the reading of a mind. In just this way, I believe, there is ample scope for the deploying of Theory of Mind abilities to infer intentionality in the progression of steps in an argument, however clearly encoded the content of each step.

Imagine two mathematicians who like to play a game in which they communicate only in mathematical STATEMENTS, wherein all of their terms and rules are well defined beforehand. On reading a new email response, a player is excited not to check all of the moves but to guess at where the other player is going with this latest move, to guess at the motivations for it, rather like a chess player looking for the strategy of an opponent. An excessively modest move might be scrutinised until the reader is able to guess at the overarching strategy that it builds towards. The reader who is able to recognise certain strategies can direct interpretive work in that direction. It seems there is still space, then, for creative inferential work in the reading of a proof, for a reader to look for intentions, to guess at what is happening in the author's mind.

In Chapter 4, we saw a pair of collaborators correcting and refining the presentation of their work in a paper. One of the changes that they made was to add a note in the text (see Figure 79) to explain why certain conditions were included; including these conditions at this stage served to simplify the statement of a later section, but since this was not obvious when reading this earlier section, the authors judged that it was best to include commentary in the text to let the reader know the reasons.

Their inclusion of this commentary indicates an important point about the consumption of mathematical texts; that even when reading highly technical and refined symbol-heavy STATEMENTS, a reader is looking not just at the explicit content of what is stated but for the *purpose* behind each component, the way that the argument fits together.



Theorem 4.1. *Assume that $n \geqslant 5$ and that $\nabla F(\mathbf{x}) \gg 1$ for all $\mathbf{x} \in \mathrm{supp}(w)$ Assume that M is coprime to $2\Delta_F$ and let Ω_M be as in (4.1). Then*

$$\widehat{N}(B, \Omega_M) = \sigma_\infty(w) B^{n-2} \prod_{p \nmid M} \sigma_p \prod_{p^m \| M} \frac{\#\Omega_{p^m}}{p^{m(n-1)}} + O_{\varepsilon,F,w}\left(B^{n/2+\varepsilon} M^{n/2+\varepsilon}\right),$$

for any $\varepsilon > 0$.

In this result and henceforth in this section, the implied constant is allowed to depend on the choice of ε, the form F and the weight function w, but not on the modulus M. To ease notation we shall suppress this dependence in what follows.

Some comments are in order about the statement of this result. The condition that $\nabla F(\mathbf{x}) \gg 1$ for any $\mathbf{x}$ in the support of w is required to simplify the analysis of the oscillatory integrals that appear in the argument. The assumptions $(M, 2\Delta_F) = 1$ and $(\mathbf{x}, M) = 1$ for any $\mathbf{x} \in \Omega_M$ are made purely to simplify the expression for the leading constant in the asymptotic formula for $N(B, \Omega_M)$.

Theorem 4.1 can be improved in several directions. Firstly, an inspection of the proof reveals that one does rather better in the B-aspect of the error term when n is odd. Secondly, it would not be hard to deal with quadratic forms in $n = 3$ or 4 variables. Finally, when M is square-free it is possible to improve the error term to $O(B^{n/2+\varepsilon} \#\Omega_M^{1/2})$. In order to simplify our exposition we have decided not to pursue any of these improvements in the present investigation. In our application Ω_M will be comparable in size to the set of $\mathbf{x} \in (\mathbb{Z}/M\mathbb{Z})^n$ for which $F(\mathbf{x}) \equiv 0 \bmod M$, and so we have relaxed the dependence on $\#\Omega_M$. In fact, although wasteful, we shall often employ the trivial inequality $\#\Omega_M \leqslant M^n$.

Figure 79. Note added to the text of the paper in §3.2 to explain the conditions that are included in the statement of (4.1) in order to simplify a later section

It would seem, then, that some kind of inference with reference to awareness of human minds is likely to be at play in the reading of a mathematical text. Let us consider, then, whether an inferential explanation will turn out to be of use to us.

Lemma 7.3. Communicative relevance?

In the example above, a relevance theoretic explanation jumps readily to the mind. A surprising condition might inhibit understanding of the flow of the argument since the reader has reason, given the added complexity, to stop and think about the reason for including the conditions: extra effects that will justify the extra effort that it takes to read and digest. This would indeed seem to be a case where a reader would be making inferences, with reference to intentions. But it is important to be clear about what kind of intentions we are talking about.

As we saw above, while a paper might involve a lot of notation, the writer is also including a lot of detail about how the argument is to be followed. The progression of steps in a logical argument like a proof are by no means pre-ordained and while the statements therein are quite encoded, the way in which the argument progresses is something that a reader will need to infer, considering the mind of a writer and in what direction that person might be aiming, as well as the basic principle that if it is being included in a communicative form like a published paper, it must be relevant. In Chapter 4 we saw that readers of even a theorem STATEMENT will be looking not only for the content of the statement but also for where the author was going by stating it, to the extent that this sometimes needs to be addressed in the text. In this case, an audience recognises multiple kinds of intentions: one, the writer's intention to communicate an idea, and two, the writer's intentions and beliefs in relation to the work being done. To avoid confusion we might henceforth call these intentions$_1$, for the communicative intention, and intentions$_2$, for the broader range of intentions. All of the material in a mathematical paper is intentionally$_1$ shared, for the purpose of making it available to a reading audience. Much of the inferential work described above though has the purpose of guessing at a writer's intentions$_2$, in the sense of goals and aims. These kind of intentions$_2$ are something that a person might guess at in a clearly non-communicative situation, such as if they were covertly watching a person work and figuring out what they were trying to do. We can still recognise the decision to publish these actions as a decision to communicate, and appeal to the *showing* end of the spectrum to understand it as such.

In Chapter 4, I noted that different parts of a paper seemed to consist of *exhibited action*, a series of steps with cognitively relevant expressions laid out before the reader, and other parts seemed to give *commentary* on that action. A mathematical paper seems required to do multiple things at once: to provide the cognitively streamlined expressions but also to give some of the gloss to guide a new mind through using them. We might understand the text of a mathematical paper as providing both the inscriptions, the cognitive tools, *and* the training material needed for a reader to arrive at a skilled engagement with those tools. We might then divide mathematical writing into two (sometimes overlapping) functions: one, the footholds are provided for a process of thinking and becoming convinced, the external artefacts that will help a mind to reach a certain cognitive experience; two, some evidence of another mind's such experience is shared, in the form of commentary or other kinds of evidence, shared with the intention of guiding another human mind through using those tools.

We might better understand this by drawing a link with the kind of *showing* behaviours discussed in Chapter 6. In those examples, mathematicians deliberately interacted with notation in front of one another and in so doing, were able quite effectively to share a way of thinking through the material, which could then be discussed directly by the group. In those cases, the written representations provided footholds for thought and the mathematicians were very concerned with sharing a particular *way of seeing* the content. As the proof of a Lemma or Theorem is laid out, the sequence of notations provides a kind of partial account of a progression that leads to a conclusion, with the reader inferring the reasoning process of which this notation is a part, or for which it provides the footholds. This is not quite a reconstruction of the process by which the writer became convinced, being a somewhat idealised retelling. Understanding the notation as laying out the footholds for a desired chain of inferences, rather than addressing the reader in any direct way, allows us to see why subsequent commentary might be useful.

I recall Brian Rotman's search for an 'underlying story or idea or argument' to render a proof meaningful and 'not merely an inert string of formally correct inferences' (Rotman, 1998 p.65). If we interpret such a string in the terms of Chapter 6, as something like an author *thinking out loud* in a deliberate and ostensive way so as to share a subtle mental state, a state of conviction, then perhaps we can propose an alternative to Rotman's meta-Code explanation. In the progression of the core of an argument, a reader is *shown* a kind of record or trace of a reasoning process, that then the reader is welcome to work through and reconstruct (as in Interlude 3). The heavily encoded statements are barely half of the story; what is important is the intentional making available of evidence or a *trace* of a thinking process which a reader can then actively engage in reconstructing.

This conception of the purpose of formal mathematical text helps us to answer a tricky question when it comes to providing a relevance theoretic account of formal mathematical writing. In Remark 7.1. Portability and examination I described two examples in which a piece of mathematics is 'readied' for formal presentation, and commented that the formalised presentation gains a somewhat clunky quality, one that is heavy with definitions and difficult to read. If a formalised STATEMENT is difficult to read, and we are to look at it in terms of our usual expectations of optimal relevance, then how can this extra effort be justified in terms of effects?

In some ways, it might appear that this more heavily defined style of mathematical writing is simply adjusted to be relevant to a greater audience. Care is taken to make links with existing research and address existing aims and questions, ensuring broad cognitive effects. An effort is also made to express everything in as well-defined terms as possible, which to a certain extent serves to make successful interpretation possible for a broader audience. In this case we might indeed entertain the idea that optimal relevance is being pursued.

However, unlike other cases of mass communication, such as, say, advertisements (Forceville, 1998), this writing is not geared toward instant relevance to a large number of people. This style of writing is difficult and demanding to read, requires a broad contextual knowledge of notations and terms that few readers will actually have, might require a reader to puzzle through a complex expression to make sense of it. Even decoding the content of a simple STATEMENT is a process that balloons enormously, encapsulating complex ideas that have built upon one another for millennia, and understanding the significance of a STATEMENT for the field goes far beyond that. Mathematical notation could still be said to be pursuing certain principles of relevance, with its emphasis on maximising *interestingness* for minimised space. It is not, however, always so to a particular hearer with limited knowledge and experience, who might have to invest quite some cognitive effort to make sense of the long string of symbols. The kind of relevance being pursued seems rather more like optimal relevance only to some kind of optimal reader, a style of writing that in all of the right circumstances will give statements with absolutely as broad implications as possible, while maintaining succinctness. While this can be framed as a kind of pursuit of optimal relevance, the parameters are somewhat unfamiliar, greatly changed from ordinary communicative situations.

An alternative proposition may simply be that this type of STATEMENT is not intended to address a reader and communicate relevantly; it has other functions to fulfil. While a group of collaborators could depend on their rich shared history (and clarify for one another as needed), with an unknown reader, a lack of explicitness runs the risk that the cognitive effects may be entirely inaccessible due to obscurity. For a piece of reasoning to really become part of the field of mathematical knowledge there is a sense in which it needs to become more stable, and so precision and explicitness becomes an important aim even at the cost of communicative relevance to an individual. The writing becomes an artefact to be made use of by a community more than a means to address a human reader.

That would not, of course, be to say that there is nothing in a paper that is addressed to a reader; as mentioned before, papers include commentary, and there are parts of mathematical communication explicitly geared toward guiding a reader's engagements with its prompts. A paper does not just consist of STATEMENTS; there is a structure with section headings, an introduction, remarks in the text, a lot of commentary of the kind seen in Chapter 4. Section headings give a certain kind of evidence of what the author is intending$_2$ to achieve with a particular section; a theorem, for example, is more interesting (has more implications, is more potentially relevant) than a proposition, which is more interesting than a lemma; a lemma is taken to be relatively uninteresting and is labelled as such to allow the reader to appreciate it as a necessary building block as the paper progresses. A conjecture is something believed to be true but not proven in the paper, declared as such to preserve trust in the writer (see Figure 80, and the list of headings in a real paper in Figure 81 and Figure 111). As with the graded certainty in statement types noted by Latour and Woolgar (Latour & Woolgar, 2013), there are no strict procedures for assigning these labels. By recognising what an author is intending$_2$ to do with a section a reader is given an indication of what to expect in terms of effects. A reader might recognise two of the five names given in a précis in the introduction and, knowing that nothing but the key literature would be mentioned here, make a quick inference about the kind of topic being addressed; with that in mind, her familiarity with similar questions might lead her to look for a certain kind of proof. Reading the first few lines, in which a hypothetical assumption is made, she might infer that the author is planning a proof by contradiction, and look for the inconsistency that the assumption leads to.

Figure 80. A physical manifestation of the layout of a paper, with the sections marked and explained. Original in colour.

Introduction Definition 1.1. Theorem 1.2. Definition 1.3. Theorem 1.4. Definition 1.5. Theorem 1.6. Theorem 1.7. Theorem 1.8. Example 1.9. Theorem 1.10. Corollary 1.11. Theorem 1.12. Layout of the paper. Notation. Acknowledgements. Lemma 2.1. Proof. Definition 2.2. Proposition 2.3. Proof. Corollary 2.4. Proof.	Remark 2.5. Lemma 3.1. Proof. Lemma 3.2. Proof. Definition 3.3. Example 3.4. Remark 3.5. Definition 3.6. Lemma 3.7. Proof. Proof of Theorem 1.2. Proposition 3.8. Proof. Proposition 3.9. Proof. Lemma 3.10. Proof. Corollary 3.11. Proof. Proof of Theorem 1.4. Proof of Theorem 1.6. Lemma 3.12.	Proof. Theorem 4.1. Lemma 4.2. Proof. Lemma 4.3. Proof. Lemma 4.4. Proof. Lemma 4.5. Proof. Lemma 4.6. Proof. Lemma 4.7. Proof. Lemma 4.8. Proof. Lemma 4.9. Proof. 5.2. Proof of Theorem 1.8. 5.3. Proof of Theorem 1.10. 5.4. Proof of Theorem 1.12.

Figure 81. List of the standard heading types used in a paper

Following this train of thought, it may be that the main concern when producing formal mathematical writing is not to be relevant to each member of the reading audience but to *manifest* ideas as completely and yet succinctly as possible. We might still call this a pursuit of an optimally *efficient* sentence, just as in communicative situations, but one that is likely to be *relevant* not to the vast majority of individuals in its reading audience but to an optimal mind with perfect knowledge—or perhaps optimally relevant to the mathematical community as an entity rather than its individual members. One way of thinking about this might be to call this pursuit of optimal efficiency one of *cognitive* rather than *communicative* relevance; we might even playfully term it *super-relevance*, for its exaggerated nature. The aim, then, may not be to guarantee relevance to a particular individual, but to pursue optimal relevance assuming optimal conditions. Rather than hoping to communicate relevantly to a set of individuals, the symbolic expression is intended to fix large amounts of precision in a written form so that it can in theory be grasped, examined and manipulated effectively. For this reason, it skirts the rules of communication but also moves outside of them, operating not as a straightforward means of communication between two minds but as a resource for facilitating thought that is intentionally shared.

Thinking of a STATEMENT less as a communicative utterance and more as a shared, optimally efficient resource for thought might help to explain something of the way that papers are actually used. The only time when a paper is really expected to be read from beginning to end is when it is under review by a journal, at which point reviewers will check through a theorem with some thoroughness before it will be accepted as part of the body of mathematical knowledge. By contrast, a mathematician reading a paper is highly unlikely to read it from beginning to end. The structure of a paper is designed to facilitate a 'dipping in' approach and it is common to only read certain results or certain sections, perhaps reading the theorem statements, perhaps digging in to a particular proof; if a mathematician really wants to understand a section, it will be with pen and paper in hand, working through examples. The STATEMENT has been laid out and shared in order to be used.

In 'Magic Words: How Language Augments Human Computation' (1998), Andy Clark considers various ways in which the use of words can be said to have increased humans' cognitive capacities. As I have in this thesis, Clark considers words both in their communicative capacity and for their role in more ostensibly private cognitive work, such as the situation in which a person internally repeats a sentence to act as a reminder or a shaper of action. Discussing the latter Clark suggests that deliberately forming a sentence in this way renders it 'an object for both ourselves and for others' (Clark, 1998 p.43), whether in fact shared or otherwise; Clark's position is that linguistic formulation renders a thought in a newly stable structure,[15] that can be subjected to evaluation and criticism. While interpretation of utterances as encountered in natural communication is a highly contextual endeavour, as we have seen, Clark here focuses on the technology of words themselves as repeatable units that bring useful properties by virtue of providing a stable code. Speaking about the 'fixing' qualities of public language Clark states the following:

> [T]his […] involves the development of a type of code which minimizes contextuality (most words retain more-or-less the same meaning in the different sentences in which they occur), is effectively modality-neutral (an idea may be prompted by visual, auditory or tactile input and yet be preserved using the same verbal formula), and allows easy rote memorization of simple strings. By freezing our own thoughts in the memorable, context-resistant and modality-transcending format of a sentence we thus create a special kind of mental object -- an object which is apt for scrutiny from multiple different cognitive angles, which is not doomed to alter or change every time we are exposed to new inputs or information, and which fixes the ideas at a fairly high level of abstraction from the idiosyncratic details of their proximal origins in sensory input. Such a mental object is, I suggest, ideally suited to figure in the evaluative, critical and tightly focused operations distinctive of second order cognition. […]. Language stands revealed as a key resource by which we effectively redescribe our own thoughts in a format which makes them available for a variety of new operations and manipulations. (Clark, 1998 p.44)

I quote this at length because although this description is intended for the transition between the hazier world of thought and formulation in speakable sentences, it can be read as an apt description

[15] Clark does not seem to subscribe to a language-like thought position.

of the transition from hazy, contextual notes and chalkboard scribblings to the more formal, stable, abstracted writing found in mathematical papers. These latter formulations we have seen are shaped to more readily stand alone, being less dependent on the contextual knowledge of a colleague's explanation or the detail of a problem a person has been tossing around for days; is modality-neutral in the extreme, the ideal being that all ideas can be encapsulated by purely logical notation; is geared heavily toward efficiency of expression, resulting in highly compressed expressions of fiendishly complex ideas that the well trained mind can thus grasp and manipulate readily. These features, Clark argues, enhance our ability to scrutinise ideas, to recognise flaws, more clearly assess the reliability of instinctive judgements, and better appreciate the logical transitions in our thought.

I recall Mercier and Sperber's emphasis on reasoners' ability 'to distance themselves from their own opinion, to consider alternatives and thereby become more objective' (Mercier & Sperber, 2011 p.72), for them the barest beginnings of the far more reliable benefits of collective reasoning. Mathematics, then, is formalised to be shared more widely; no particular surprises there, although it is interesting to begin to think of this in terms of increasing portability, or rather, decreasing dependence on the messy, situated details of a small-scale reasoning situation and increasing dependence on textual technologies, on definitions, on a well-defined code and on the shared aims of the mathematical community of a whole. As mathematical writing is stabilised in published form, it is expressed in terms that are known to and that link it to the work of a wide community, is expressed in terms of the aims of that community; in short, it is directed to a quite different audience. The mathematical writing is still intricately connected with a context of humans and materials but rather than five people in a room with a whiteboard it is thousands of people, thousands of talks and papers, and the terms and notations they use to coordinate themselves.

Formal mathematical writing, then, seems to fulfil a variety of functions. These range from the clearly communicative, for example in the introductory and explanatory prose sections of a paper, to the somewhat stranger form of the mathematical STATEMENT, which at times seems best described as making available the external component of a process of thinking-with-inscriptions for a reader to utilise. Formal mathematical writing, then, we might see as ranging from straightforwardly having *meaning* for an intended reader, to *showing* a reader something that might aid them in a particular process of thinking. I will leave open the question to what extent it is the pursuit of optimal communicative relevance, exactly, or some cousin thereof, that guides the construction of the mathematical STATEMENT; it seems fruitful to consider the STATEMENT in communicative terms but perhaps it is best thought of as hovering somewhere on the boundaries of communication, its construction guided by principles that are not-so-distant relatives of conversational communicative relevance, and yet remain somewhat apart. This family resemblance could simply be of a principle similar in kind to that of everyday communicative relevance but shaped by atypical expectations about the audience; alternatively it could be distinguished by something like the particular open-endedness of the intended cognitive effects in a reader, the STATEMENT being made available to facilitate a parallel cognitive journey to that of the writer, but without a particular expectation that it should be taken up. In any case, both the prose and the STATEMENTS of mathematical writing facilitate the sharing of mathematical work between mathematicians, and as such both play essential parts in a complex, collective cognitive system.

Remark 7.4. Expertise and aptitude

What, then, of the human components of that system, and the undoubted skill and expertise that they bring to their work? Is there still a place to talk about individual expertise in this collective cognitive system? In the case of in-person collaborative work, we have seen how important a sophisticated, shared cognitive environment is as collaborators share and develop ideas. In the course of moving toward publication, we have seen how the inscriptions become more explicit and specialised, demanding a different kind of expertise, first in simple knowledge of the code, and then with a kind of experience working with it, knowing where to direct attention and having the familiarity to make quick, useful inferences with it. A particular kind of expertise that is highlighted when we talk about mind-inscription interaction is this kind of *aptitude with inscriptions*, the particular abilities involved in efficiently using these potentially cognitively super-relevant representations.

In the above example from Chapter 3, what is being shared or adjusted are the details of the way a mind is using and thinking through an inscription. The group work together to develop and then to solidify a *way of seeing* potential diagrams in a new inscription, a short statement that would guide a new mind through finding that *way of seeing* anew. In Chapter 4, we saw how a pair of collaborators agreed to include commentary to guide a reader through interpreting a string of statements, recognising that part of a viewer's work is to infer how the statements add up to conclusions with relevance to the mathematical community. In Chapter 5, the mathematician wrote to develop and eventually share first a *way of seeing* possible diagrams of two types so as to perceive a correspondence between them, and second a way of organising perception of that correspondence and the features relevant to establishing it that demonstrated that every possible diagram of one type had an equivalent in the other. In each case, we see evidence of both work thinking-with inscriptions and also work done to refine those mind-inscription interactions. The picture I have been building up of mathematical work has such skilled interactions at its very centre and so a key task in seeing that mathematical work is successfully and effectively shared is to guide other minds into achieving these skilled engagements; a knowledge how to work with a highly developed cognitive technology.

A notation can be good or bad for the interpretations that it encourages or inhibits, which can play a part in which notation becomes standard (Pimm, 1987; Brown, 2008; Villani, 2012). In this way the context packed in to a STATEMENT can become something that is not consciously entertained, and is more like a question of experience. This experience can even mean that an experienced reader recognises a particular notation and context and knows which parts of the expression are going to contain the most important information, like seeing an Σ for a summation and looking to see what kind of function follows it, and what the limits on the summation are set to above and below its symbol (Figure 82).

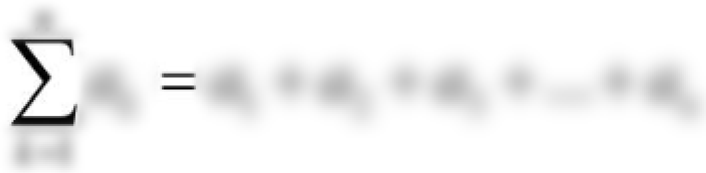


Figure 82. Recognising the form: an experienced reader looks for certain expected and useful information in an expression

Mathematical papers are read by *doing*, read to be enacted, read with pen in hand. Several of my participants reported that their reading habits when it came to papers they really wanted to understand involved sitting down and actively working through sections, manipulating the inscriptions for themselves and coming to a working knowledge or knowledge-how in relation to them (see Interlude 3). A mind must learn its way around mathematical writing, around the relevant features, how to perceive it and how to manipulate it, and this is best learned by *doing*.

When I struggled to recognise the importance of the different types of bracket in Chapter 4, one of my participants commented that this was ‘a common mistake made by 1st year maths undergraduates whilst they are still learning how to do “real” mathematics.’ The fact is that there is a lot packed in to mathematical notation, and a trained mind knows where to look for it: the benefit of experience can be to know where cognitive effects might be found. If novices like me are looking at a text without that kind of guidance, we very well might totally miss things. So much is packed in to mathematical notation that is meant to be succinct that it doesn’t always wear its importance on its sleeve. One thing we can note though is that if communication produces expectations of relevance then that can be exploited to share those expectations (for example, if a teacher ostensively highlights which brackets are being used while writing them, and so directs pupils’ attention). These kinds of strategies can be used in the classroom or presentation, and also help us to understand why mathematicians put so much emphasis on in-person communication (as in Interlude 2), where the bandwidth for subtle attentional nudges is so much greater.

A way to understand an expert's engagement with a mathematical text is to recognise the difference in the accessibility and salience of different interpretations for expert and novice. These notions were introduced for utterance interpretation in Chapter 1, but are common ideas throughout cognitive psychology. If a particular assumption is often 'activated' then it is likely to jump readily to mind; if a particular interpretation is more commonly used in mathematical work then an expert will more readily leap to it, and if a particular component of an expression tends to be very significant then that expert will know that it ought to be paid careful attention.

Should mathematical expertise then be understood in terms of aptitude with representations, developed in such a way as to agree with the usage favoured by the mathematical community of experts? In the example above, the group produce a STATEMENT which is a stabilisation of an ability to distinguish between situations. But a newcomer will also have to apply intelligent behaviour to that STATEMENT itself for it to achieve its cognitive relevance, will need to be familiar with the potentially clunky 'greater than or equal to' phrasing, to feel at home with the if-then structure of the STATEMENT to readily interpret and work with it. The user of these cognitive tools is faced with the task of learning to make use of them, distinguishing which parts are important and deserving of attention, knowing what to look for in the sequence of steps that make up an argument.

This characterisation would allow us to understand at least some portion of mathematical expertise as knowledge-how rather than knowledge-that. Knowledge-how is best characterised as an ability, like riding a bike. The fact that this knowledge-how is a part of work that produces the clear propositional statements of mathematics is no obstacle to its being knowledge-how; positions such as the *interrogative capacity* view have been developed to reconcile intellectual and anti-intellectual views on knowledge-how to account for its functioning as an *ability to generate* propositions (Habgood-Coote, 2019). While this kind of skill can be described and even shared as a set of instructions, a person wishing to learn the skill would have to do considerable work to interpret those instructions and put them into action, learning a great deal more in the form of additional, subtle detail about how action and perception should be organised. As well as being told how to do something, a person can be *shown* how, an intentional enacting in front of a person in order that they see how the action should proceed, which can make available quite a range of data that the shower may not even be sufficiently consciously aware of, to explicitly mention. These characterisations seem a good fit for the mathematical situations examined in this thesis, as well as the observations that I have mentioned that mathematicians favour talking in person to share ideas, and take very active approaches to reading papers (Interludes 2 and 3).

In Figure 77, the STATEMENT that the group eventually wrote was itself a kind of guide to working with inscriptions. The sentence itself was a stabilisation of the group's work of developing a certain way of seeing any potential future example, a sentence that included all of the building blocks for them to pay attention to a diagram as F did and 'see' it in a discerning way. While B's definition on lines 01 and 03 takes quite a definite and brief form, it is clear from F's expansion that the function it can fulfil for a hearer is quite complex and nuanced, and one that goes beyond what is explicitly stated to a whole array of inferences, which the hearer may or may not reach. F's deliberately exhibited interaction with B's definition and the diagram served to guide C in drawing out a whole array of such inferences, which amount to something like an *ability to discern* situations (in which the diagram does or does not meet the conditions) or a *way of seeing*. The real usefulness of the outcome of the meeting is in the *way of seeing* that a reader might be able to derive from it; it contains certain indications of a stabilisation of a way of looking at a potential diagram—seeing a diagram in terms of numbers of vertices in a certain relationship to one another—pinned down in language so that it can be shared.

One view on what intelligent behaviour itself *is,* according to Hubert Dreyfus' development of Merleau-Ponty's notion of *maximum grip* (Dreyfus, 2002), is the ability to distinguish between types of situation with greater and greater sophistication, and respond accordingly. This is the idea that a person develops an ability to identify certain selected relevant features in a whole range of situations and that this allows the person to respond to those situations in the best way; Dreyfus' position is that learning to make these distinctions is constitutive of intelligent behaviour. A person learning to play chess, for example, is first able to distinguish basic threats, then more complex strategies, with increasing levels of sophistication, and perceiving which situation is which is what allows a person's responses to be increasingly apt and effective. In this case, F guides C through a learned engagement with a diagram that could be applied to any number of diagrams, to distinguish one situation from another. In just such a way, some significant portion of what is means to be an expert mathematician consists in learned discernment when working with external representations like diagrams and notation.

We now have a sense of formal mathematical writing as part of an extended cognitive system, providing important, refined cognitive tools which human actors must learn to manipulate and become skilled in using. The final point that I want to make is that while this formal writing takes this heavily refined, precise form, it is still just possible, with the right kind of eyes, to see the traces within it of the part it plays in a very human system of reasoning.

Remark 7.5. Recognising the particularity in mathematical expression

There are two details of the STATEMENT in Figure 77 that I would like to highlight. This formalised version of the statement describes vertices as 'k-bad'. As we saw in Chapter 3, that term has a perfectly technical definition; in fact mathematical texts often use words like 'bad' or 'nice', with appropriately technical definitions. But the term is also a charming throwback to the whole history of endeavour and setback in the meeting, to the group's frustration and laughter. Second, this clarification is using the language of 'switching'. A reminder: the 'switching' description made the most sense when the group were talking about switching paths, paths that switched direction, before the shift of focus to vertices. There are plenty of other ways to talk about these vertices with arrows going all in or all out, for example that they're double-headed, or double-ended, or (to invent a Greek term, from *homo* for same, and *akri* for tipped) 'homoakrinous'. But the 'switching' language remains, another residue of the history of ways in which the group were looking at the diagrams.

This work has not yet been published, and it is possible that neither of these little traces in the statement will persist when the work is published in a paper; either term might be replaced by something else. But it is not impossible that they might persist, and the excitingly vibrant bits of language that pepper mathematical papers demonstrate that they sometimes do. The *breaching experiment* in Chapter 4 brought to the fore some of the surprisingly emotive language of endeavour and failure to be found in mathematical papers, language that might go unnoticed but still survives as a record of the indirect paths of difficult mathematical work. And if that language does persist, then so much the better! On one level, even a subtle acknowledgement of the winding path that research can take might help a new researcher to recognise the commonality of that experience. On another level, these traces could serve to help a reader to reach a particular insight.

The group began with the 'switching path' way of seeing because it was intuitive, and then made the switch to vertices. If the 'switching vertices' terminology retains some trace of that progression then it might even help a reader to direct interpretive work in that direction, to do something of a reconstruction of the that particular route taken in coming to understand.

Clark particularly notes the capacity of writing to guide new minds through the routes needed to access complex ideas, taking into account what is known as *path dependency* in human learning. Path dependency, which has been much investigated in research on Artificial Neural Networks (Elman, 1993), refers to the kind of intellectual journey taken in something like formal education, wherein minds are taken through a progression of ideas that build upon one another (even beginning with ones now thought of as 'incorrect'), each priming the mind for 'a finer grained truth' (Clark, 1998 p.40). In this way, textbooks are wonderful resources to build and shape a path to a tricky idea. At a smaller scale, a mathematical paper can serve to do the same, to bring a mind through premises and argument to reach a conclusion. At a smaller scale yet, these traces of old ways of seeing baked in to even a sentence could perhaps lead a mind through different conceptions of the problem to arrive at just the place that the authors did. And yet, without examining the progression in the way that we have, these details would be all but invisible in the improvised yet precise world of mathematical language.

If there are traces of these kinds of messy processes of discovery baked in to the technical language and terminology used, the language does not somehow fail to be sufficiently technical. Rather, this simply is a feature of technical language and the way that it develops; it is something that is born of a stepwise refinement of language from situated, impressionistic ways of communicating gradually toward something more general and portable that can become part of the community's work. It could be expressed otherwise. But the very small details of its expression might help or hinder the creation of shared understanding in subtle and generally unseen ways.

Philosopher and historian of mathematics Madeleine Muntersbjorn notes how easy it is to assume that different language points to essentially the same thing, but powerfully makes the case that these representational differences have been part of real and substantial differences in the concepts and their situation within the field of knowledge at the time. She describes mathematical objects as tacit insights that aid in problem-solving across different situations, and emphasises that innovations in notations, diagrams and terminology often serve to make implicit features of successful mathematical reasoning explicit (Muntersbjorn, 2003; Muntersbjorn, 1999). Her position is that it is the use of cleverly designed representations in the essentially collaborative activity of mathematical reasoning that brings the sophisticated mathematical objects that we know into being, and that the culture of mathematical writing reflects this: 'Mathematicians instinctively recognize the connection between symbolic manipulation and object reification and so rely heavily on chalkboards, scratch pads and computers when actively engaged in innovative mathematical reasoning. […] Our sensible intuition can only construct new mathematical results *via* the intentional manipulation of visual imagery. Neither intention without imagery, nor imagery without intention, is enough to get the job done' (Muntersbjorn, 1999 p.196). Again, we see the importance of the relationship and interaction between intending mind and sophisticated inscription.

Looking at mathematical artefacts in this way allows us to see them not as perfect objects used by imperfect minds but as objects produced by and productive of human reasoning, themselves the manifestation of minds' efforts to reason most effectively. As a representation is used and passed around in a community it is shaped by that usage; notations that facilitate contemporary usage tend to replace older ones that are clunkier for the purposes they are being put to. This is important since even details like the spacing of an expression can have substantial effects on how readily minds can reason about it (Jansen *et al.*, 2003; Jansen *et al.*, 2007). The details of a notation can facilitate or not facilitate certain kinds of thought, can offer different affordances, can be adjusted to open up new conceptual vistas, and once a notation is chosen, it shapes which paths are more likely to be followed in subsequent research.

In this chapter, I have considered formal mathematical writing from multiple communicative perspectives, and concluded that its function is not entirely encapsulated by communication, and made the case that even formal mathematical writing ought to be seen as a component in a cognitive system: one that encompasses not just a single mathematician or a group of collaborators, but the entirety of the mathematical community. I have written about how writing is sometimes used to share footholds for complex cognitive tasks, and about the importance of *aptitude* in working with these inscriptions as a significant part in mathematical expertise.

This thesis has aimed throughout to manifest the ideas it describes, and the importance of knowledge-how has been important in highlighting the importance of practice-based research. In Chapter 8, I will document some of the ways in which I have approached these themes through practical experiments in generating and working with inscriptions.

Interlude 8. What is not said

Subject L

0.41.40.000

L: In the first paper […] I realised that also the referees comment that I skipped many details, or not details, for example I'd say: OK, it's clear that this follows from this, but I need to write that. Even though people know that. In a mathematical paper it should be very very clear. […] You need to prove every statement that you are writing, even though it's not the statement of a theorem. […] I would write, you know, 'It is this.' I was thinking that it was a little bit insulting to the reader to write an explanation for that because I was thinking it's a really well known fact. But I learned that […] I need to explain in a good way, not an elementary way. […] And because of this you need to be very patient when writing mathematics. Sometimes you know, I don't explain this, it's easy, straightforward. And I see math preprints full of like, 'it's obvious,' 'it's trivial,' 'straightforward,' things like that. […] [When other people do that sometimes] it makes me lost.

Research meeting

00.17.32.000

M: So do you still have- um – so he does he prove that you have the Holder spaces o-of- into a vector bundle

N: Uh, Rafe you mean?

M: yeah I mean does anybody actually prove that those are Banarch spaces and if you look at … you know the the what would be the natural definition somehow of a Holder space of sections into a vector bundle that those that really is a Banach space

N: Like like so you mean uh so so I assume you mean actually check…

M: Yeah

N: So no nobody writes that down so everybody just says that's the…

M: I mean is that obvious

N: it's an elementary exercise is what they say

M: OK

Subject G

0.56.50.410

G: You don't write every individual step down. You always leave some steps to the reader to fill in by themselves. Because you always assume the reader has a PhD in mathematics, for one thing […] and you assume that they're knowledgeable about your area. […] so then it's like, how many gaps can you leave in the argument, kind of thing. So I write it one way, and to me it's clear, and I can fill in all the steps in my head. But that's partly because of my mathematical background, say.

It gets more extreme than that. Like, some mathematicians have zero respect for the reader. They write their papers so badly. To them- if it's clear to them, then they assume it's clear to the reader. […] they don't think about the reader at all.

So 'zero respect for' means not being helpful?

G: Yeah like- they don't assume the reader is like, a person who doesn't understand everything they know, who doesn't have the same brain as them. And they can be quite rude about it, and say, you know, 'it's trivial.' […] So they'll write the paper almost just for themselves.

Some famous mathematicians think they don't need to write complete proofs because it's clear to them, and they can get away with it because they're famous and everyone just takes their word. But then sometimes there are mistakes.

Work site 8

Work site 8 was the office of a lecturer at a university in the USA.

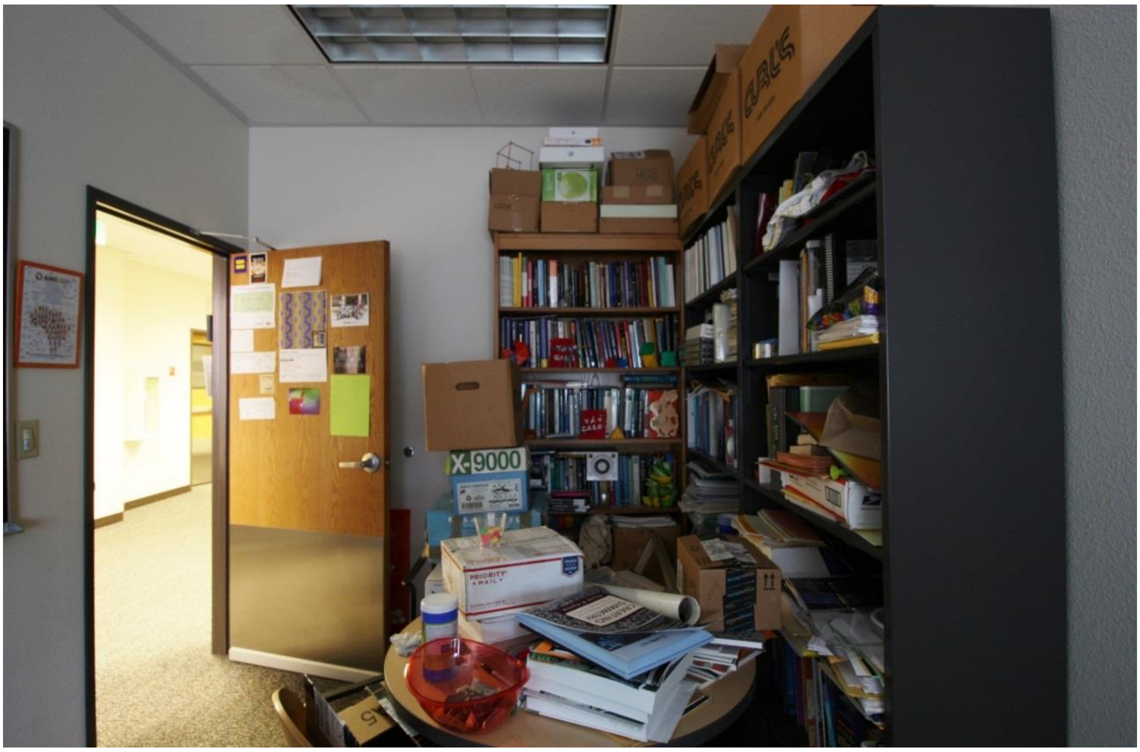


Figure 83. Bookshelves and table in my participant's office. Original in colour.

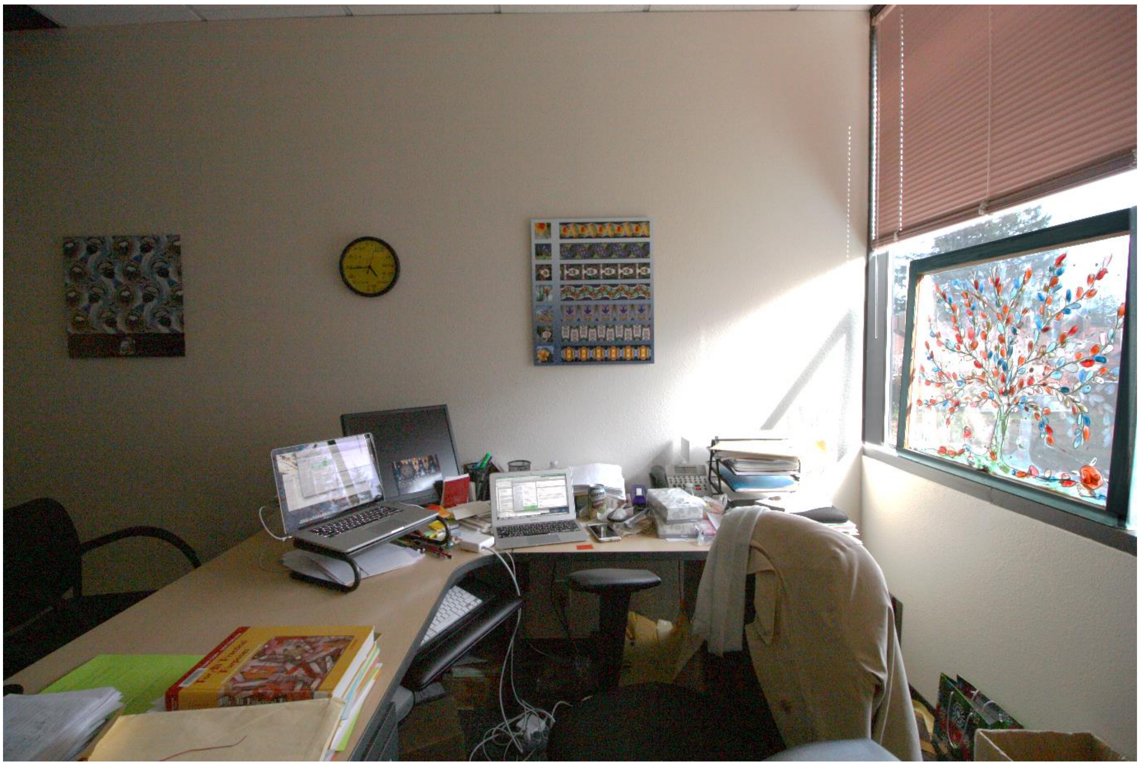


Figure 84. Workspace in my participant's office, with computers and tablets, and prints of mathematical art on the wall. Original in colour.

8. Discussion: doing to understand

In the previous chapters, I put forward a picture of mathematical writing as playing a part in a complex, distributed cognitive system, playing an essential part in both communication and reasoning, and demanding of its users a high degree of knowledge-how in order to be used effectively. In this chapter, I present some of the practical experiments I developed to explore the uses and effects of working with external representations for reasoning and communication. These experiments are lighthearted and playful, intended to suggest some unlikely way of seeing or behaving. The reader may notice a strategy of deploying something like a mathematical approach in unlikely or unfamiliar situations, to test out what that approach looks like in an unfamiliar context.

These experiments constitute the continuation and conclusion of the practical experiments that have been seen earlier in the thesis, which I will briefly summarise below.

8.1. Artistic ethnography

Throughout this research one of my aims was to develop a practice that fit the strengths of genuine artistic ethnography laid out in Chapter 1. These, in summary, were to approach material in a holistic, multimodal way, and employ means of investigation and presentation that go beyond text; to generate insight by putting forward possible worlds or worldviews that are somehow other to or incommensurable with existing ones, but nonetheless demonstrate their consistency; to experimentally examine and challenge systems of perception that shape our picture of the world around us. Catherine Elgin's view that art essentially functions by a kind of exemplification begs the question of what exactly is exemplified, a question that exposes the strength of that characterisation for the range of responses that could be given, and could suit different kinds of art: a sensation; an atmosphere; a perspective; a transgression; an organised thought process; and so on.

In the data analysis chapters of this thesis I used certain creative and practical methods of data analysis, which also informed the way that the research was presented. I conducted my data analysis by imitation, mimicking the forms in which my data came as I analysed it in order to gain a richer, more nuanced understanding of the media observed. I used in my analysis process a practice of exemplifying ways of working with representations in order to examine them, and documented these through records of that process that function like the adjusted retelling of a process of discovery seen in scientific papers.

In the *breaching experiments* found in these chapters, I dug more deeply into the methods used by my participants by imitating them with deliberate, transgressive alterations to change the parameters of these practices and see what the outcome would be. These experiments often had unexpected or strange outcomes and proved quite informative. The intention was to exemplify a way of working that paralleled but diverged from the real-world ones that I was observing.

As I began work on synthesising the overarching arguments of this thesis, this kind of imitation-with-alterations extended into a kind of mapping practice, a process of passing material through different representative media with each representation making aspects more or less visible, which existed as a research method above and beyond the data analysis. In Appendix 1 I document some key aspects of this practice.

As the key points made in this research came into focus, I began looking at ways to explore by experiment and to exemplify for readers of this thesis the kinds of mind-inscription interactions that I had come to describe. These efforts resulted in the proposal of a number of practical exercises designed to foment certain practices and experiences, such as that of using representational technology to extend and abstract away from concrete experience, coming to a working knowledge of a notation system by experimenting with examples, and coming to invent notations that extend thought in ways particular to their affordances. 8.2. Experiments outlines the practice outcomes of this thesis: three main experiments, and documentation of their deployment among an interested audience.

8.2. Experiments

I devised three experiments designed to elicit organised mind-environment interactions. A key focus came to be what happens as a decision-making mind comes to shape and be shaped by an object. These interactions are what make up the sophisticated interactions with material resources that we saw playing a part in complex cognitive tasks in the previous chapters.

I present to the reader some of the results when I have enacted these experiments with human participants. In keeping with the direct access to data that has been a guiding principle of this research, I have also designed these experiments to be something that the reader can try, to be another source of direct data and experience that I hope will broaden the kind of knowledge that the thesis serves to share. I strongly encourage the reader to try these experiments, to find a pen and paper or a conversation partner and experience the results first-hand.

8.2.1. A Solid

The first such experiment explored the capacity of engagements with representative media to build concepts that extend out of our everyday experience. In this experiment, the representations are a given, set up as an aid in an imaginative thought experiment. The experiment takes the form of a booklet with a set of progressing instructions, diagrams, and examples, with reader-writers asked to work through the instructions, and space provided to record the results if the reader-writer wishes.

The instructions ask a reader-writer to imagine something from everyday experience that could serve as a 'point', such as the experience of jumping at a sudden noise; the task is then to extend this in various directions, unifying an imagined range of situations under an imagined conceptual 'block' of experience. Diagrams are key to this exercise. While the instructions could be followed alone, it is the set of progressing diagrams that make the conceptual extension feel intuitive (Figure 85), and the notion of viewing this experiential range as in some way an entity. In this way a reader-writer can come to conceptualise something as unitary and unique as an experience in an extended, abstracted way.

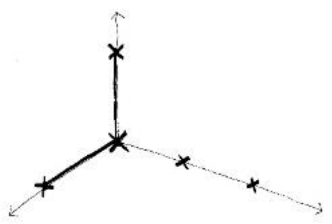


Figure 85. Diagram

Below I reproduce the set of instructions from the booklet.

Choose a 'point'

Example:

Jumping at a sudden noise

Imagine another point with a common axis

Example:

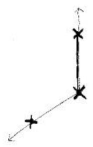
Wincing at a sudden noise

Extend the first point to the second to make a 'line'

Example:

Responding to a sudden noise in a way that encapsulates a whole range of responses, from jumping to flinching to wincing

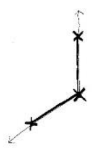
Imagine another point on a different common axis



Example:

Jumping at a sudden noise, rather later after it happened

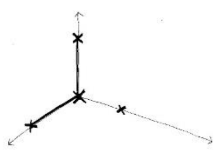
Extend the first point to this one to form another 'line'



Example:

The jump at a sudden noise, in which a person's muscles tense and propel them into the air, heart beating fast, with senses heightened, lasts, continuously and without change, for several seconds

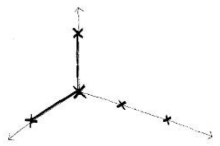
Imagine another point on a different common axis



Example:

Jumping at a sudden tap on the shoulder

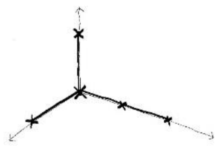
Imagine another point farther along that common axis



Example:

Jumping at a sudden salty taste

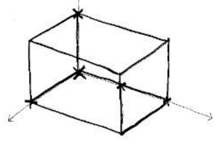
Extend the first point to this one to form another 'line'



Example:

Jumping at a range of sudden sensory stimuli, from a sudden sound to a sudden tap to a sudden salty taste

Starting from your point, extend it to form your first line, then extend along your second axis to form a surface, then extend along your third axis to form a solid. Describe your solid.



Example:

Responding to a range of sudden sensory stimuli, from a sudden sound to a sudden tap to a sudden salty taste, in a way that encapsulates a whole range of responses, from jumping to flinching to wincing, and this response, in which a person's muscles tense and propel them into the air, heart beating fast, with senses heightened, and their face twists into a strange expression, and their eyebrows *raise and arms raise, lasts, continuously and without change, for several seconds.*

Figure 86. Instructions from the booklet

Below are some extracts from some of the responses I had.

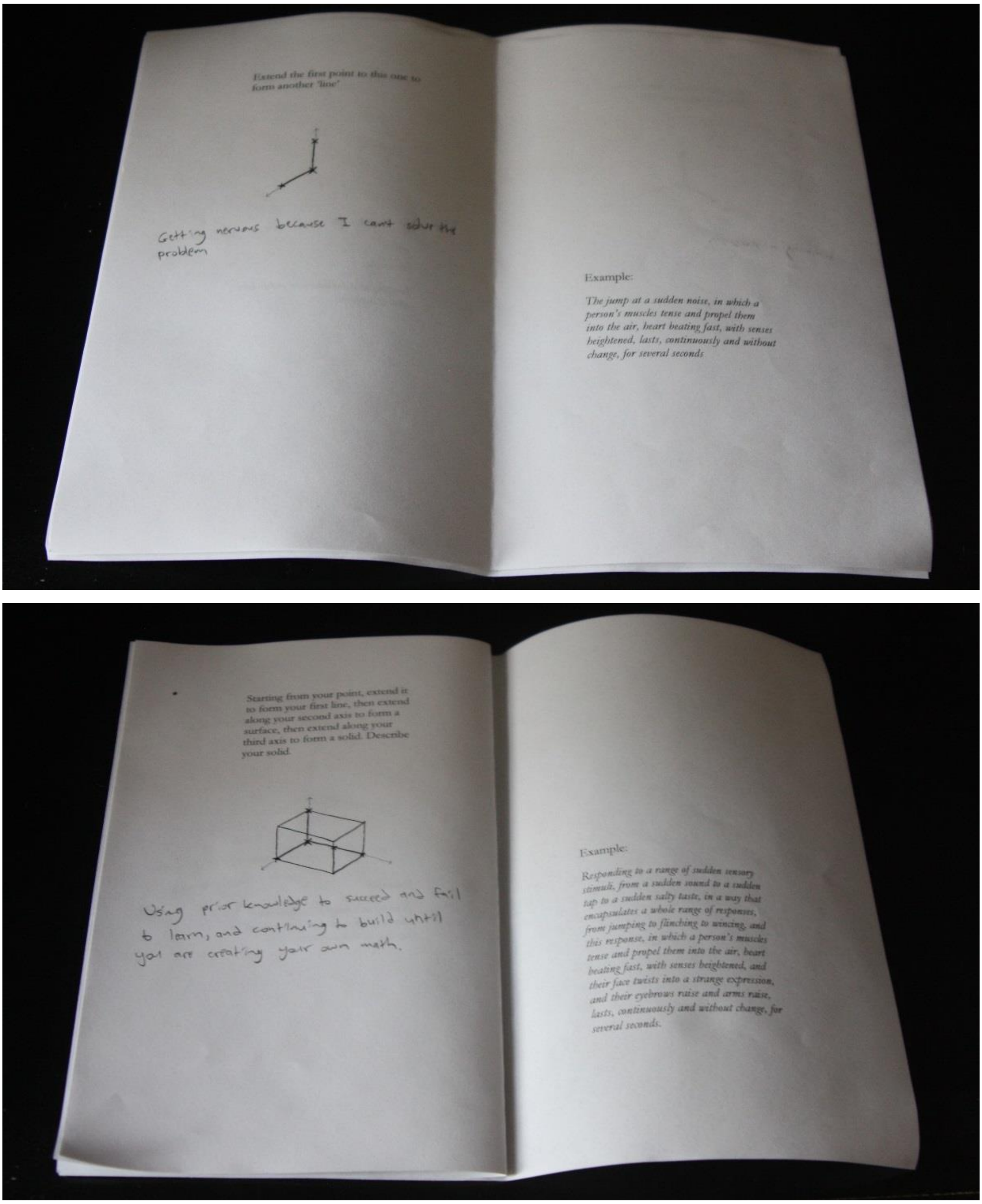


Figure 87. Booklets as filled in by participants

The stepwise instructions prompt the bringing together of an abstracted viewpoint with an imagined real-world situation. Working step by step with the diagrams encourages a kind of imaginative engagement that might be surprising but still proceeds according to a rational structure. In this way a written representation can guide a mind through novel imaginings and deploy a logical-feeling structure in quite unexpected ways.

8.2.2. Three operations, imagined

The second experiment was in generating a shared understanding of a subtle idea by active engagement with and manipulation of an interactive environment. I built a notation system out of a selection of everyday objects and hand gestures, and developed three distinct operations whose inputs and outcomes could be represented in that system. I then built an interactive website in which a visitor could click through each of the functions and experiment with them in a variety of different ways in order to get a sense of what each one consisted in, and then to decide for themselves how it might be extended and what the results would be. The website offers no instruction, just an environment that will respond to clicks so that a visitor can test out actions and see the results, and in so doing come to a working knowledge that they then are prepared to extend. The web address is www.situatingmathematics.com/0Intro3.

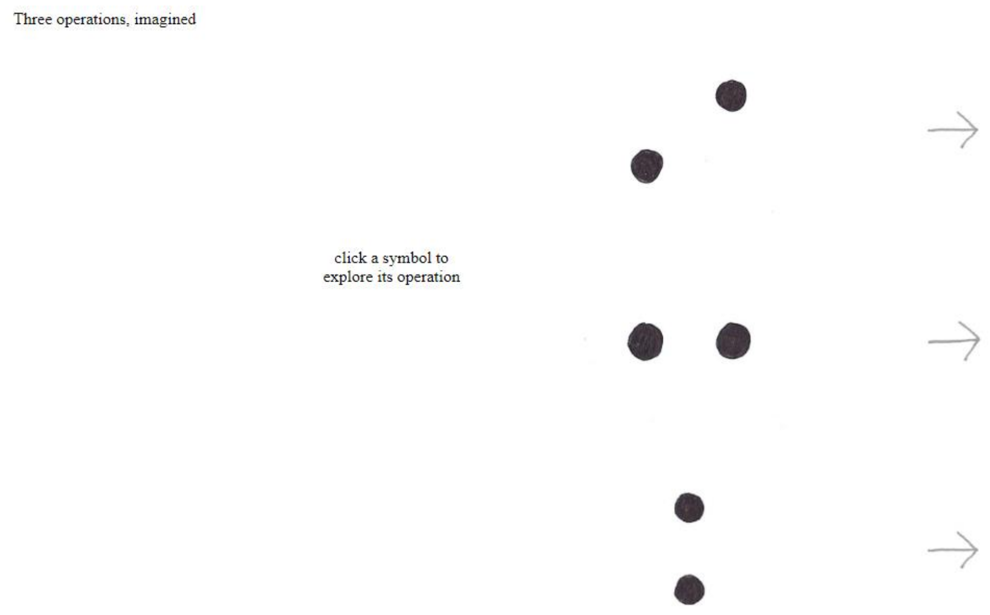


Figure 88. Introduction

The three 'functions' were developed not with any verbal definition in mind but by manipulating and experimenting with a variety of materials and coming to a particular understanding of the way that they might relate. At the end of this process, I was able to come up with verbal descriptions of each function, but their genesis was a process of experimentation, manipulation and imagination, and hence the interactive environment invites a visitor in to a similar process.

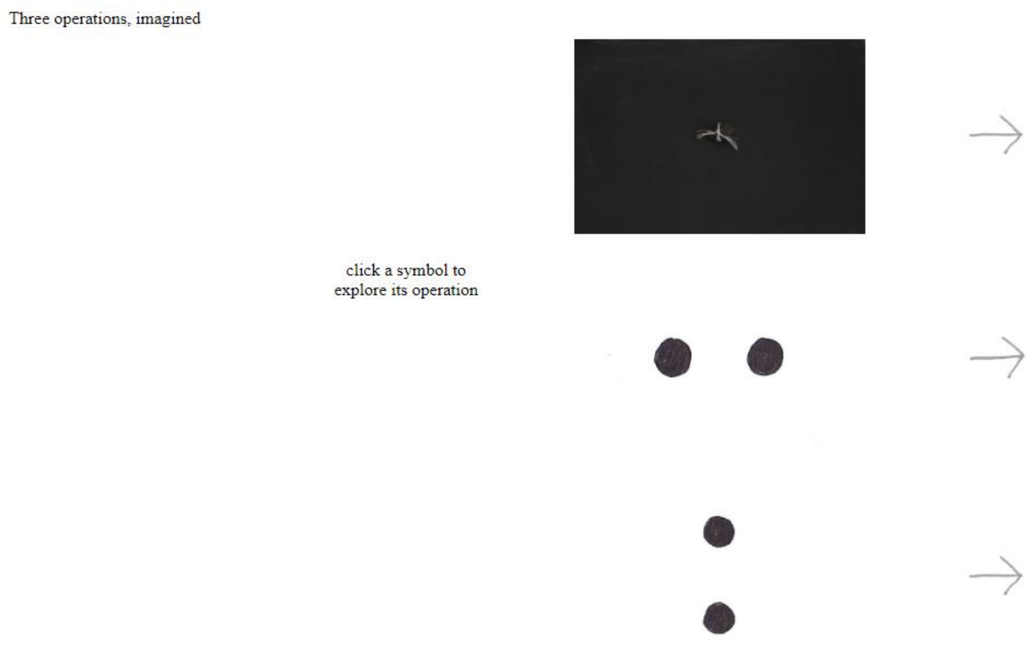


Figure 89. Clicking through one of the 'functions' from the introduction, after which a viewer is offered an arrow to click to experiment with it further

Each one of the 'functions' offers a route for exploration which moves through various stages before asking the visitor to imaginatively extend its functioning. In the first stage, a visitor simply clicks through a variety of manipulations, rather like watching a lecturer work through a few examples and so getting a first sense of how an operation proceeds.

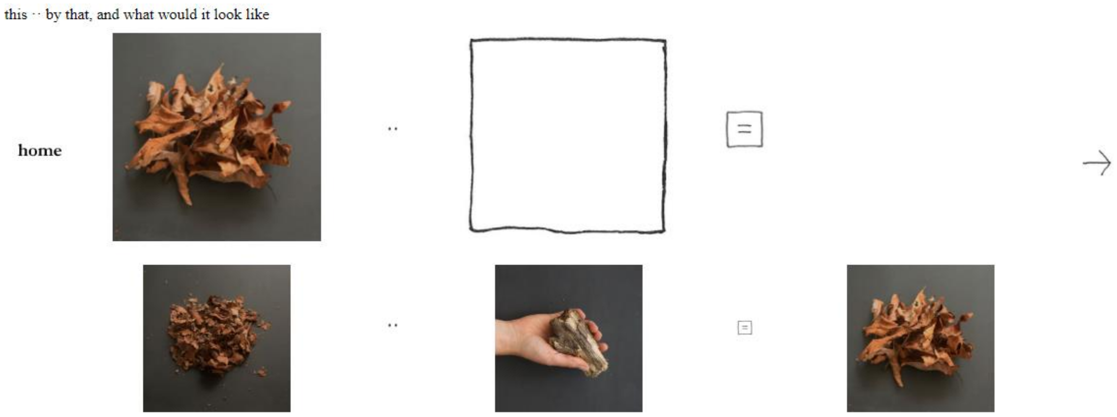


Figure 90. Experimenting with the function and its outcomes to get a sense of the implementation of the function '..' that is being proposed. Original in colour.

In the second, a visitor is invited to experiment with the function in a variety of different ways, choosing examples and finding out the results. As a visitor to the website experiments with the function, there is no verbal account of what it does. The only texts to be found on the website are a set of fragmentary ruminations on choice and imagination, which frame the decisions made as speculative, open to development, a matter of choice by a perceiving individual. A visitor is invited to click to experiment with the pictures and through that interactive process to build up a sense of the perspective that the maker's mind is proposing on these operations, get a feel for what each implementation does, and how the elements relate to one another. The presentation through objects resists easy paraphrase. A visitor is then invited to begin extending the field of imagined operations, and evaluating the result, and at this stage the inputs become verbal, a way to draw upon a person's imaginative experience and also to deliver a decision-making autonomy, an openness in terms of slight aesthetic decisions that might shift the outcomes in any number of different directions.

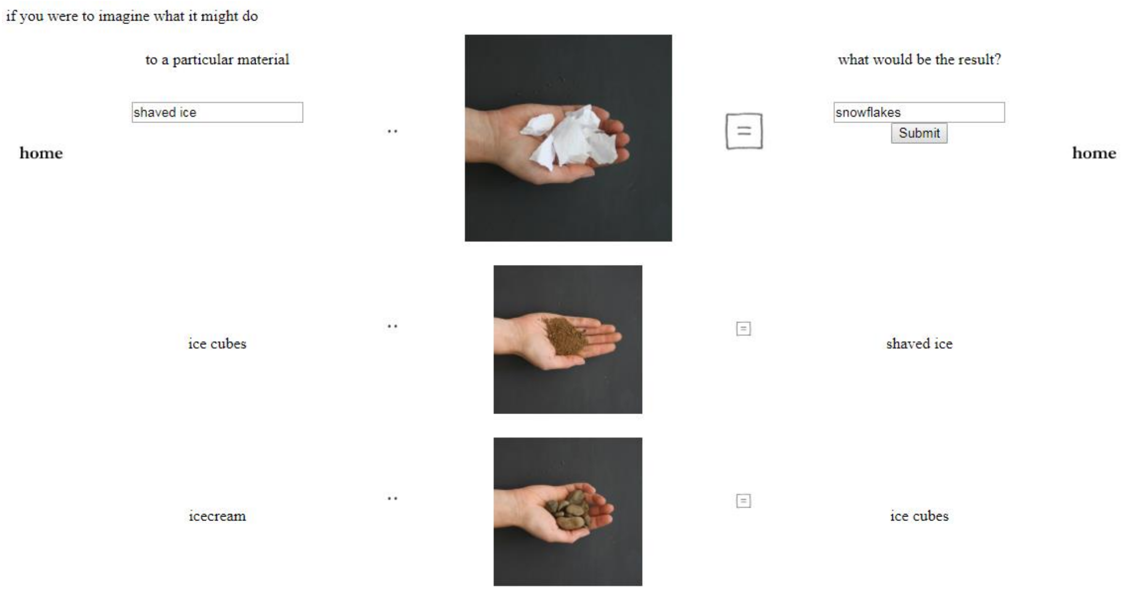
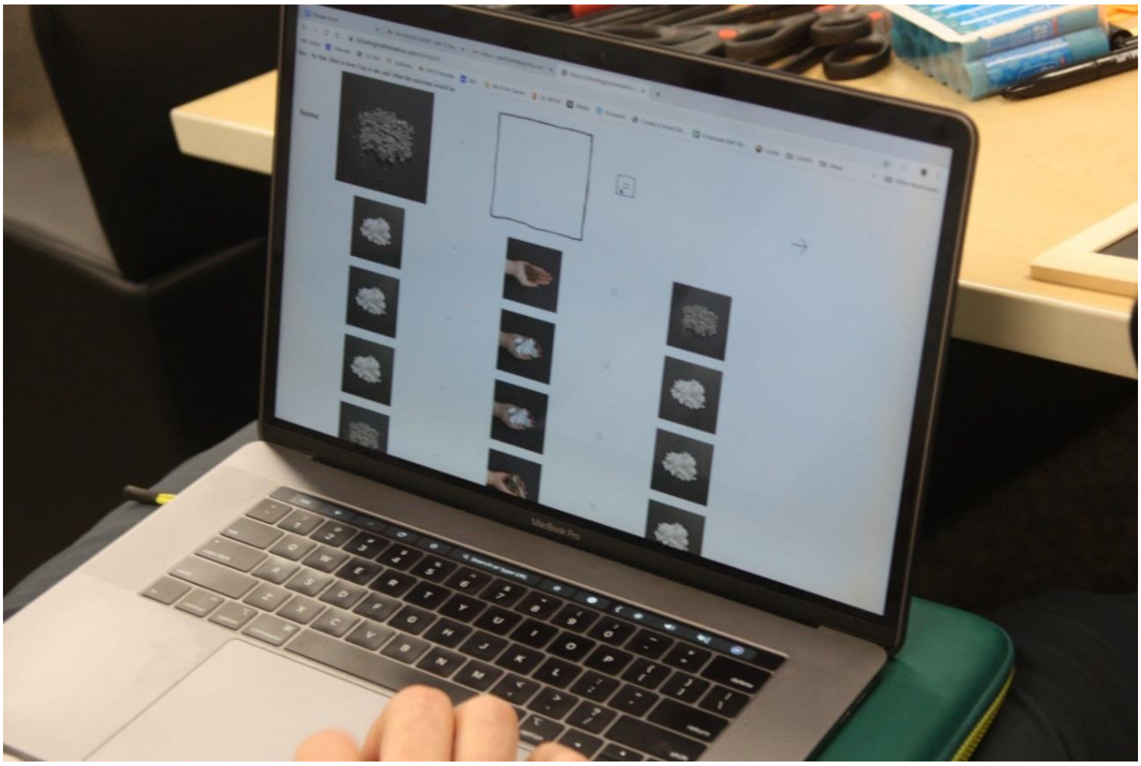


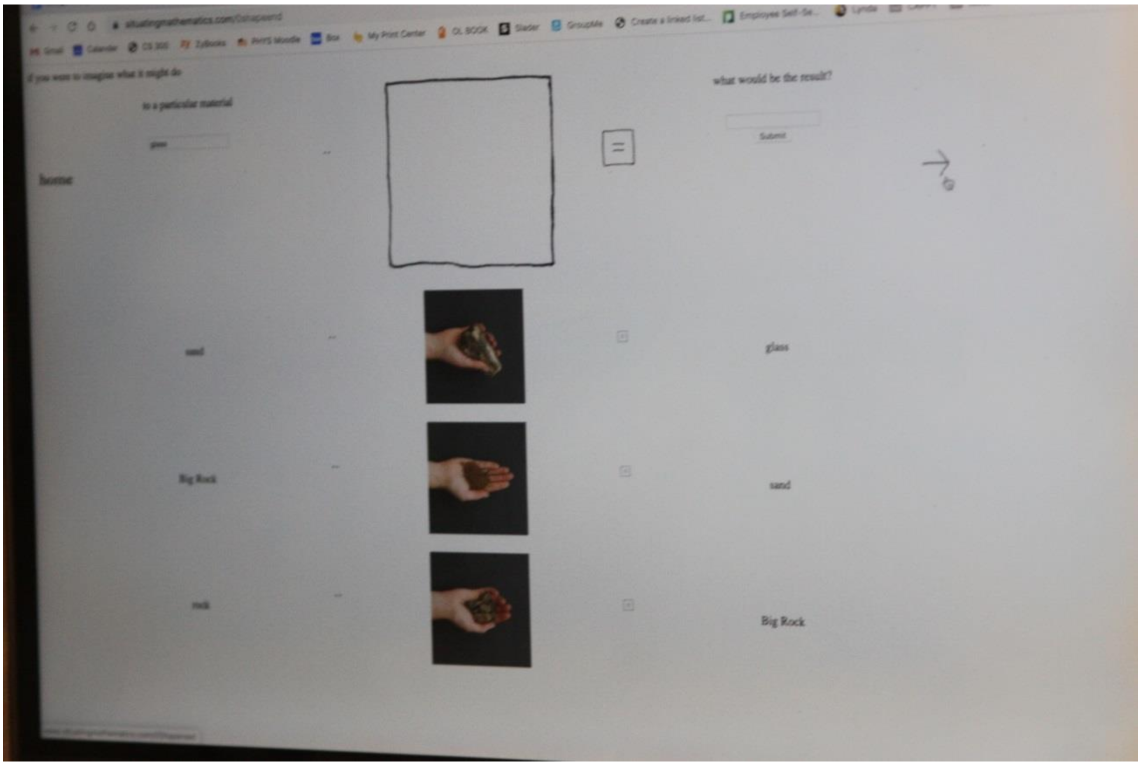
Figure 91. Inputting new values and evaluating the result. Original in colour.

Finally, a visitor is invited to begin extending the function, changing first just the inputs and then any and all of the variables.

Figure 92. Choosing an input and what it is operated upon by, and evaluating the result. Original in colour.

In this way the website is designed to manifest the particular experience of coming to know by first observing, then doing, and then being able to go on, to contribute an idea of what might come next.




what would be the result?
home
glass
sand
Big Rock
Big Rock

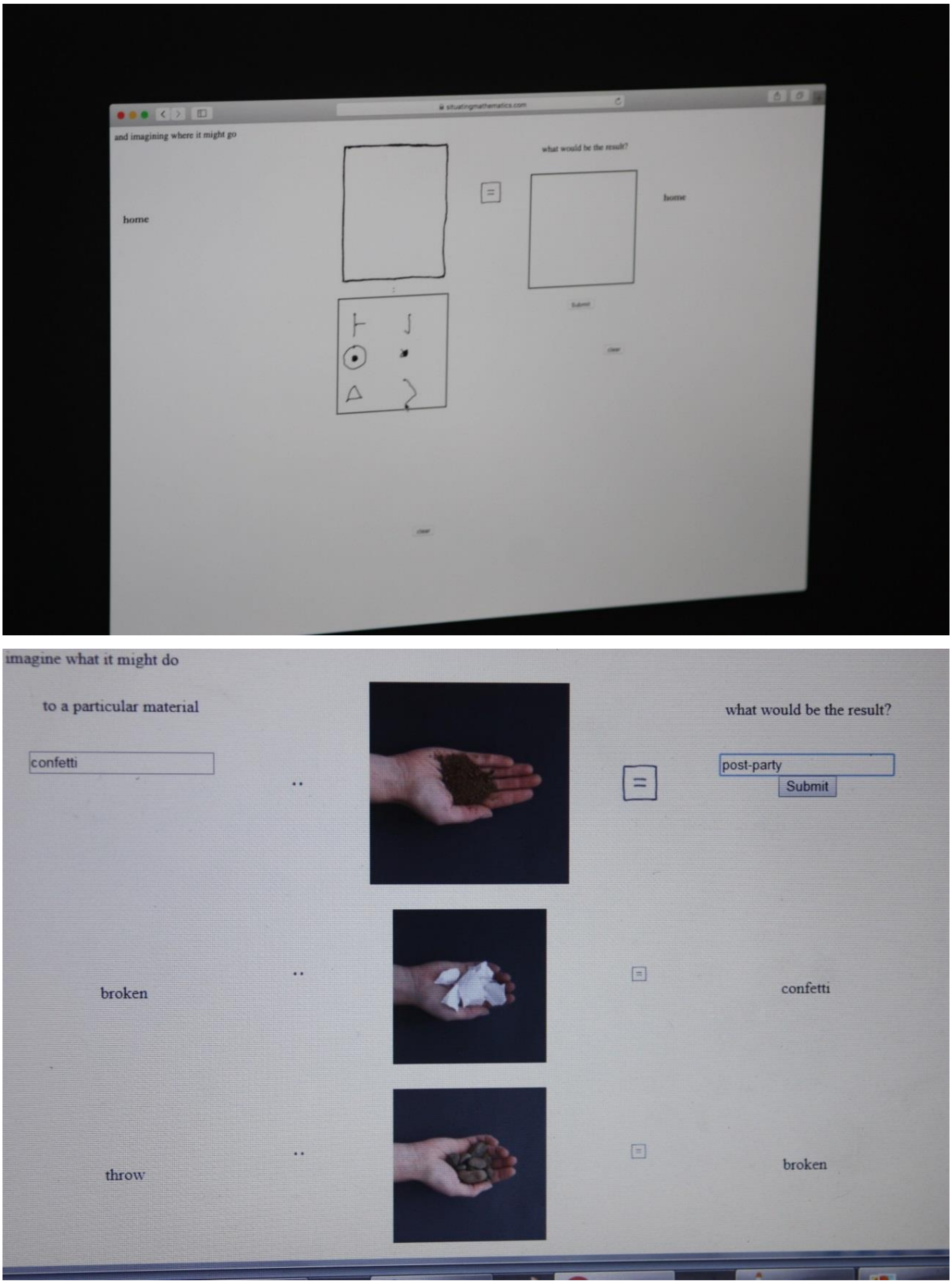


Figure 93. Participants' inputs to the website. Original in colour.

8.2.3. Notation workshops

I would characterise my goal with these practical experiments as seeking to construct the kind of interactive, empowered engagements with inscriptions that I have been describing over the course of the thesis. The previous two experiments invite an audience into imaginative engagements with representations that offer new conjectural pathways and, in the latter case, a sense of another perceiving mind. This third experiment is a workshop that constructs a more participatory and openly creative situation. As the *Dialogue without shared aims* demonstrated, and as is generally argued in post-Gricean pragmatics, it is very difficult to make sense of human interactions without recognising the simple fact of cooperativity as a guiding force. In this experiment I designed three tasks that ask participants to work together and with a variety of representations on tasks that orbit a list of phrases that are very close in meaning but different in wording. The nearness of essential meaning pushes the site of activity away from the concepts and toward the particulars of representation and the cooperative negotiations of the participants.

The tasks that were proposed are as below. These experiments were put forward as prompts with diversions and amendments from participants eagerly welcomed. This was run as a workshop at Pacific University with a group of mixed students and faculty from Art and Mathematics departments.

[1]	Inside out
[2]	Through and through
[3]	From top to bottom
[4]	Up and down
[5]	From beginning to end
[6]	To the end
[7]	In toto
[8]	To the nth degree
[9]	All the way
[10]	In entirety
[11]	To the max
[12]	Without omission
[13]	In full

Figure 94. List of phrases

The tasks each have to do with communicating or representing one of a list of phrases. The phrases are very close in meaning but expressed in quite different terms, meaning that to recognise the distinctions between them, close attention must be paid to the way that they are represented.

These results presented with many thanks to participants at the *To the nth degree* workshop at Pacific University in Oregon:

Kae Christopher
Timmy Brown
Heather Fleischer
Shi Wen
Alyssa Watson
Roman Stein
Chad Farias
Wyatt Ma'a

Christine Guenther
Waverly Sudborough
Erin Dahl
Olivia Chau
Jillian Lamb
Kara Putman
Michael Timmerman
Jordan Zweifel
Erin Melia
Christian Cloke
Nick Slenning
Luke Davis
Ian Besse
Octavio Garcias-Mejía
Tui Tuitele
Nancy Ann Neudauer
Michael Sentman
Sierra Wolfe
John Paul Takacs
Xallan Wilson
Angela Sims

Task 1: Interactivity

Participants work in pairs or small groups. One has a pen and chalkboard, and begins drawing a representation of one of the phrases in the list of phrases—no letters or words The other tries to guess what the phrase is.

Participants work in pairs or small groups. One has a pen and chalkboard, and begins drawing a representation of one of the phrases in the list of phrases. The guesser(s) describe(s) what they are seeing, speaking constantly, until they are ready to guess what the phrase is.

Figure 95. A participant's representation of 'to the nth degree'. Original in colour.

The groups came up with a wide variety of innovative ways to represent these phrases, often using repetition or development in time as a part of this. One of the groups commented that they felt much more able to use this aspect of public drawing in the second, speaking condition; the guesser, rather than waiting for the drawing to reach a 'finished state' that could then be guessed, was empowered to pick up and comment on developments in a way that in turn allowed the drawer to respond and make use of them in diachronic, developing representations. This observation I felt reflected some of the difference between the changing, developing state of mathematical writing in a chalkboard proof and the static, polished form of the writing in a printed one.

Figure 96. A participant's representation of 'from top to bottom'. Original in colour.

Figure 97. A participant's representation of 'through and through' and 'from beginning to end'. Original in colour.

Task 2: Thinking through a representation

In this task, the pair or group all work together. Choose three of the phrases on the list and build representations of those out of modelling clay. You could represent them as a whole or make them into strings of physical 'notation', using modelling clay to build shapes; for example a symbol for each word, or syllable, or tiny conceptual bit. Try going for some really similar ones.

Next, try to represent the difference between two of them. Make a representation of that You might also try focusing on what they have in common.

Figure 98. A group's representation of 'through and through'. Original in colour.

The groups' representations of the phrases were wildly varied. Many, again, used a certain orientedness to reflect the linearity of text in the original phrases; for example, one group had a long string looping through a circle once and then twice, as a representation of the phrase 'through and through'. In this way, the new representations reflected some of the properties of the old.

In the group photo below, a group's representation of the phrase 'up and down' can be seen on a desk on the bottom right (see Figure 99). In the following photo is the same group's representation of the phrase 'to the nth degree', which they have represented as a string of growing spheres (Figure 100).

Figure 99. A group discusses a plan of action. Original in colour.

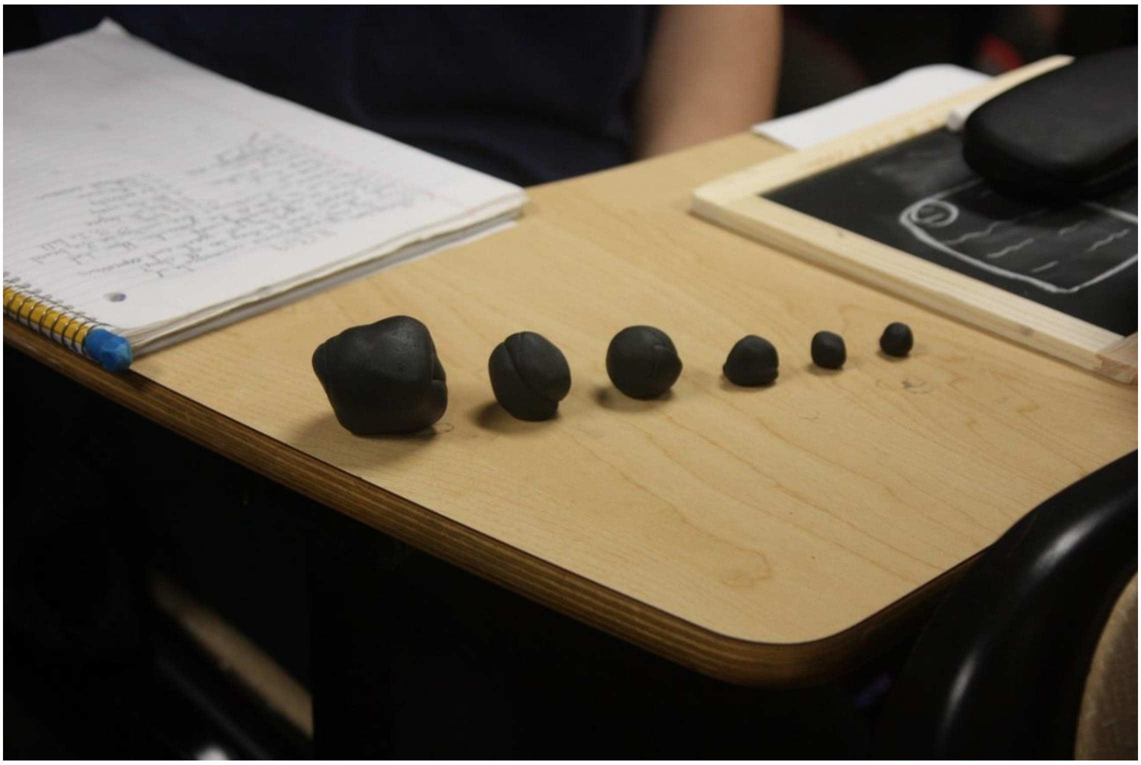


Figure 100. That group's representation of 'to the nth degree'. Original in colour.

Considering the differences and commonalities between their representations, the group noted the iterative, repetitive format that they held in common, and opted to extend that by joining the two together (Figure 101).

Figure 101. Two representations combined, emphasising the repetition in each. Original in colour.

Another group recognised that they had used orientation in each of their representations, and reflected this commonality with an arrow.

Figure 102. A group's representation of 'through and through' and 'from beginning to end', with orientedness emphasised. Original in colour.

Another group noticed that one of their representations was a structure consisting in one element, but the other came in two parts: an active, looping part, and a ring that served as a base for it to loop through, like the set for a play around which the action unfolds. Their ‘difference’, then, was the fact of an additional medium, a given base around which to work.

Figure 103. A group's representation of 'through and through'. Original in colour.

The ‘differences’, then, were often found in the particulars of a given representation, in features that were barely consciously chosen but that then proved useful in answering new questions. By way of contrast the ‘commonalities’ seemed to adhere to a certain theme that could be traced back through the development of the models from earlier textual representations. They seem to hint at an oriented mode of communication that develops through time, each word building upon the former, moving from the beginning of the idea to the end.

Summary

Each of these experiments was designed to prompt experimentation with representative forms, and to highlight the ways that representations play an intricately embedded part in systems that include speech, interaction and perception. Enacting each of these as experiments highlighted how important stepwise progression was in building up a new or unfamiliar idea. In each one, a decision-making mind comes to shape and be shaped by an object, and it is by developing those interactions one small step at a time and then building upon the last that interesting conclusions are reached.

These interactions are what make up the sophisticated interactions with material resources that we saw playing a part in complex cognitive tasks in the previous chapters, but in these experiments the parameters were different, the contexts more unfamiliar. Mathematics operates in very abstract contexts, often very distant from any kind of application, and topics are chosen for that elusive quality of *interestingness*, and often with a remarkable playfulness; so perhaps to a mathematician, these redeployments might not be so surprising at all, just a step in a slightly different direction (and with fewer strictly mathematical conclusions to draw).

Interlude 9. Chalkboards

Subject P

0.43.56.000

P: So I find that in pure maths talks if people use- write things on the board, they always tend to go down better with an audience. […] At least if I'm in the audience, it's easier for me to follow and take notes and see how the argument is structured.

Work site 9

Work site 9 was the office space of a lecturer at a university in the USA.

Figure 104. Bookshelves and a table and chair used for visitors in my participant's office. Original in colour.

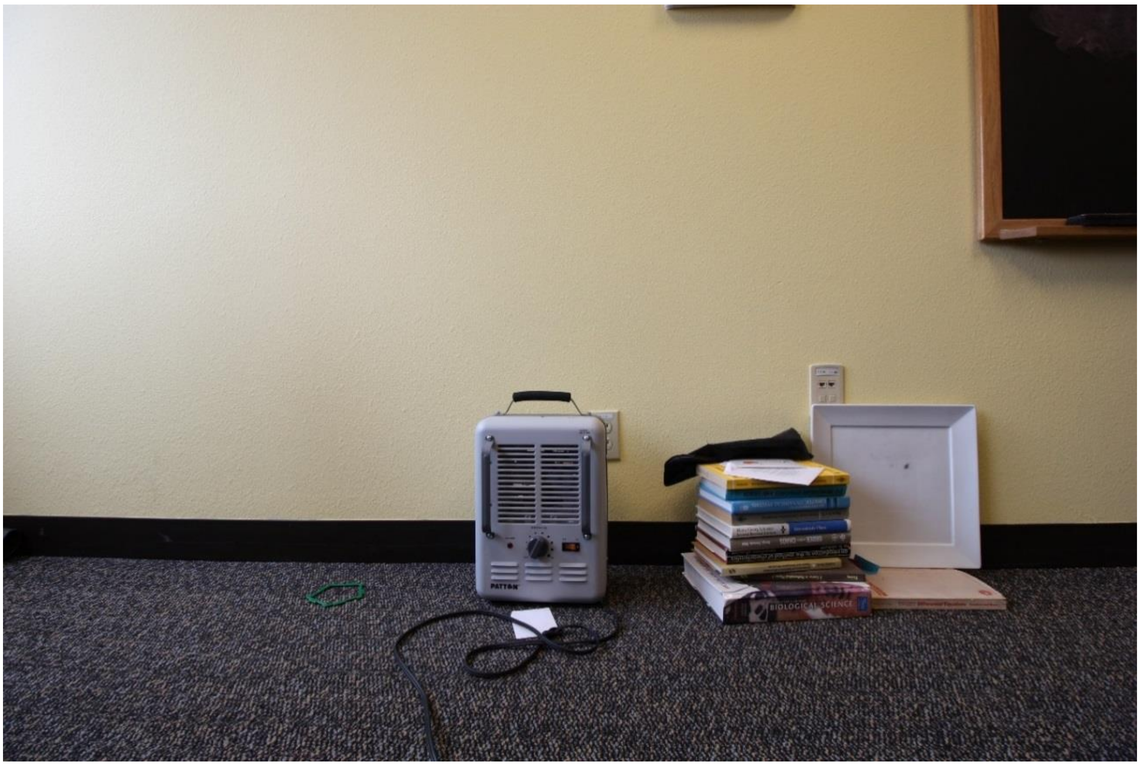


Figure 105. Miniature heater, reference books and frame by the wall in my participant's office. Original in colour.

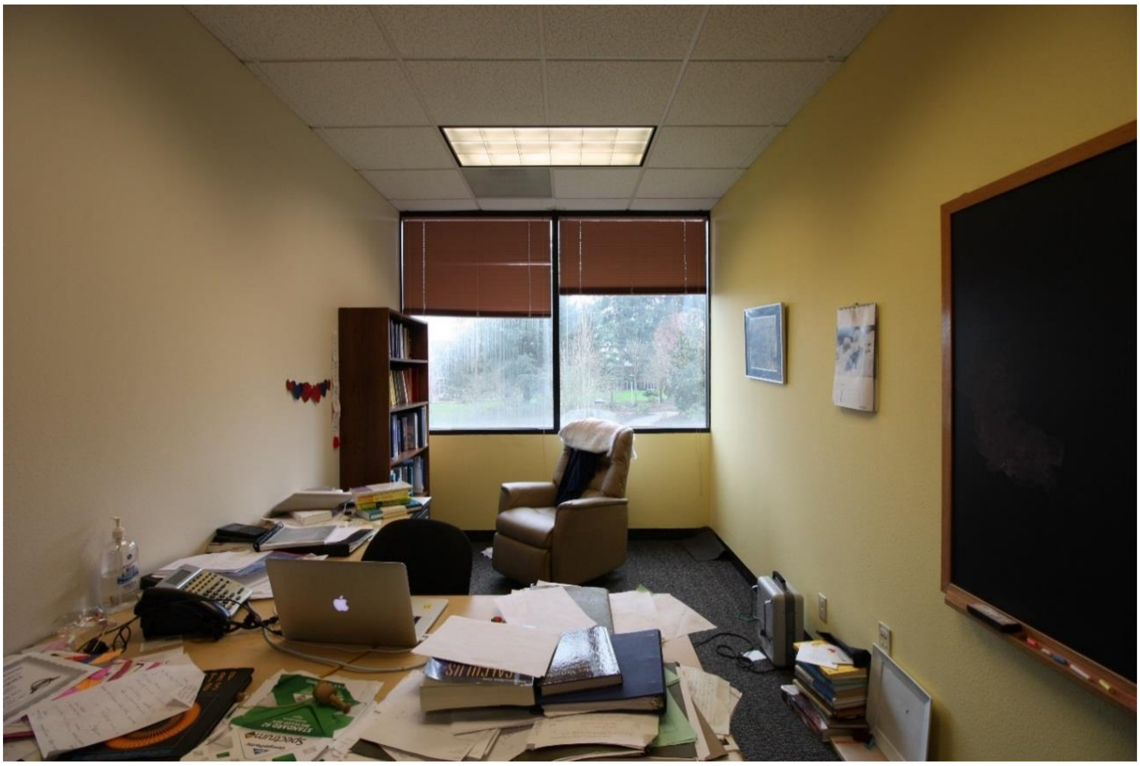


Figure 106. Desk with computer and notes on loose paper, and comfortable chair with bookshelf. Original in colour.

9. Conclusion

In the course of this thesis, I have made the case that mathematical writing should be seen as a living, active thing, constantly changing, constantly influencing and influenced by its users, an essential part in a complex, collective cognitive system.

I return to the research questions that I stated in Chapter 1, as a means to frame the responses that I have formulated over the course of this thesis, and the conclusions that I will here summarise.

1. What role does mathematical writing play across different situations of communication and reasoning, and what are the cognitive and communicative forces at work in shaping that writing?
2. What can ethnographic data show us about how mathematics advances through private and collaborative reasoning, and the resources used in the course of that reasoning?
3. How do mathematicians reach consensus, and what is involved in gaining mathematical expertise?
4. What kind of research approach and presentation will give a reader access to the observed data, including the mathematical content, examine it in a truly multimodal way, and prioritise the aim of breaking out of tenacious ideological assumptions and subjecting the data to alternative ways of seeing?

According to question 4, I set out to bring the everyday realities of mathematical work to the forefront of the research, to provide as direct and accessible evidence as I could, and to put forward some way of looking at those mathematical realities that brings them together with the complexity and abstraction that characterises mathematics and sets it apart. The research design brought together a diverse range of ideas in search of a fresh way of conducting ethnography, a particularly necessary and daunting task given the peculiar status of its subject of study. Adopting a creative practice approach informed by ethnomethodology in my research methods and eventual presentation of conclusions was my answer to the tricky methodological question of how to see mathematics anew, consider it in a way that is truly multimodal, and bring the basis in data to the reader in as complete a form as possible.

Throughout the data analysis and in Chapters 6 and 7, I gave consideration to the roles played by mathematical writing in situations of communication and reasoning, and the kinds of processes by which the field advances and consensus is reached. Unsurprisingly, communication was an important theme throughout the thesis, as was cognition.

Even though the mathematical world appears very distant from that of everyday conversational communication, I found that there is serious insight to be gained by looking at mathematical writing and communication more generally with principles from relevance theory, an ostensive-inferential theory of communication. It proved helpful to see these communications through the lens of intention-reading, cooperativity and the sharing of subtle mental states. This shows that while mathematics in many ways seems set apart from other aspects of human endeavour, it is built up from the same basic, adaptive practices that govern many other aspects of our lives, refined and skilfully managed in a way that allows us to achieve arguably some of our most impressive cognitive tasks.

The most important practice that I have highlighted is that of working with inscriptions, the intelligent back-and-forth between person and representation that has made it possible to achieve dizzying cognitive tasks. The situated cognition paradigm has proved extremely useful even in this cerebral, abstracted area of human endeavour; providing a way to understand how external resources can couple with thinking minds and make complex, sophisticated ideas easily tractable.

In the informal, improvised uses of mathematical writing seen in communication and reasoning in the 'back' end of mathematics, mathematicians use material and social resources in a rich and nuanced way in the course of manipulating and developing sophisticated ideas, using their ability to 'see' one another's thinking to work in cognitive teams, and using external representations to make complex ideas tractable. Mathematicians engage in processes of 'thinking out loud' that allow them to streamline and share cognitive work through *metacognitive acquaintance*, and so to work in close cognitive collaboration in order to develop incredibly sophisticated ideas.

I considered the more formal mathematical writing of the 'front' of mathematics from multiple communicative perspectives, and, concluding that it is in part communicative and in part a shared cognitive tool, made the case that even formal mathematical writing ought to be seen as a component in a cognitive system: one that encompasses not just a single mathematician or a group of collaborators, but the entirety of the mathematical community. I have written about how writing is sometimes used to share footholds for complex cognitive tasks, and about the importance of *aptitude* in working with these inscriptions as a significant part in mathematical expertise.

Early in the thesis, I examined the question of Hersh's 'front' and 'back' of mathematics, and exactly what could be gained from exposing this 'back' to an unaccustomed audience. I find myself agreeing with Greiffenhagen and Sharrock that really the two are not so distant as they might appear, though my conclusions are to be stated in somewhat different terms; I think that there are importantly different principles at work that shape the 'front' into a different kind of being, but that with the right kind of eyes and with very close attention, it is possible to see that the two are nonetheless of a piece, intimately bound up together, each shaping and shaped by the other. This is the case even in the apparently depersonalised form of mathematical publications, in which traces persist of the messy, situated, emotional work of discovery, and inferences about human minds are essential to their successful interpretation.

Exciting as it is to perceive the continuities between mathematical work and other aspects of human life, it would be perverse not to also consider just what it is that does make mathematics so different from other activities. If through this thesis I have attempted to demystify the mathematical world to some extent, it has all been in the service of recognising what profound and interesting things can be done with the means at our disposal. According to the framework I have proposed, I believe that it is precisely the level of refinement and development that has been given to interactions and cognitive tools that makes its products both so strange-seeming and exclusive and so very effective. It does not seem a leap to say that pure mathematics is largely pursued for its own sake, for the sake of reaching cognitive heights, and of answering questions simply because they are there. As such the engagements with representations that are seen in this field are pursued and refined in an atmosphere of intellectual development for its own sake, and as we have seen, refined interactions with external means play no small part in cognitive pursuits. As such mathematics might be seen not only as the height of our cognitive achievements but also of a certain kind of skilled work with an environment, that allows us to fiddle with chalk and blackboard, graphite and paper, and achieve astonishing things. Mathematics, in my opinion, ought to be recognised as a skilled practical endeavour, with chalk and pencil the defining tools of work.

Glossary

code model of communication – the idea that communication is essentially a case of taking a thought and encoding it in some transmissible format, which can then be decoded by a receiver

ethnography of communication – the study of communication in the context of the beliefs and practices of a particular community

ethnomethodology – a branch of ethnographic study that is dedicated to examining the methods that people use to establish order in everyday life, and how sense is made of the world and one another

inference – reaching conclusions on the basis of evidence and reasoning

inferential model of communication – the idea that communicators provide one another with not a fully encoded thought but instead a kind of evidence, from which they are expected to make certain inferences about what the other was intending to convey

inscriptions – a Latourian term for the writing practices of a discipline

metacognitive acquaintance – an addressee experiencing a certain change in his cognitive environment, and identifying this change as something that a communicator intended to cause in him and to have recognised as intentional

ostension – an act of deliberately and openly pointing out or exhibiting something, making clear that it is intentional

RVG – Rectangle Visibility Graph, a kind of graph wherein what is important is which rectangles have a clear line of sight to one another in a particular direction

situated cognition paradigm – a research programme of the last few decades that views cognition as importantly comprised of a whole person's actions in and interactions with the world

Thompson groups – three groups, introduced by Richard Thompson, with unusual properties which make them counterexamples to many general conjectures in group theory

List of References

Anderson, G., Buck, D., Coates, T. and Corti, A. (2015). Drawing in Mathematics: From Inverse Vision to the Liberation of Form. *Leonardo*, 48: 439–448.

Anderson, M., Meyer, B. and Olivier, P. (eds) (2002). *Diagrammatic Representation and Reasoning*. London: Springer London. http://link.springer.com/10.1007/978-1-4471-0109-3. Accessed 30 August 2016.

Asprey, W. and Kitcher, P. (1988). *History and Philosophy of Modern Mathematics*. Minneapolis: University of Minnesota Press.

Baddeley, A. (2010). Working memory. *Current Biology*, 20: R136–R140.

Barany, M. J. (2010). *Mathematical Research in Context, MSc by research*. Edinburgh: University of Edinburgh.

Barany, M. J. and MacKenzie, D. (2014). Chalk: Materials and Concepts in Mathematics Research. In: C. Coopman, J. Vertesi, M. Lynch & S. Woolgar (eds) *Representation in Scientific Practice Revisited*. The MIT Press. pp.107–130. http://mitpress.universitypressscholarship.com/view/10.7551/mitpress/9780262525381.001.0001/upso-9780262525381-chapter-6. Accessed 27 February 2017.

Barrouillet, P., Bernardin, S. and Camos, V. (2004). Time constraints and resource sharing in adults' working memory spans. *Journal of Experimental Psychology: General*, 133: 83.

Barsalou, L. W. (2008). Grounded cognition. *Annu. Rev. Psychol.*, 59: 617–645.

Bernard, H. (1974). Scientists and policy makers: An ethnography of communication. *Human Organization*, 33: 261–276.

Bloor, D. (1987). The Living Foundations of Mathematics E. Livingston (ed.). *Social Studies of Science*, 17: 337–358.

Borgdorff, H. (2006). *The debate on research in the arts*. Kunsthøgskolen i Bergen Bergen, Norway.

Bourguignon, J.-P. and Casse, M. (2012). *Mathematics, A Beautiful Elsewhere* (1 edition). Paris : New York: Thames & Hudson.

Bourriaud, N. (1998). *Relational Aesthetics*. Dijon: Les Presse Du Reel.

Brown, J. R. (2008). *Philosophy of Mathematics: A Contemporary Introduction to the World of Proofs and Pictures* (2 edition). New York: Routledge.

Cannon, J. W., Floyd, W. J. and Parry, W. R. (1996). Introductory notes on Richard Thompson's groups. *Enseignement Mathématique*, 42: 215–256.

Carston, R. (2009). The Explicit/Implicit Distinction in Pragmatics and the Limits of Explicit Communication. *International Review of Pragmatics*, 1: 35–62.

Cicourel, A. V. (1987). The Interpenetration of Communicative Contexts: Examples from Medical Encounters. *Social Psychology Quarterly*, 50: 217–226.

Clark, A. (1998). Magic Words: How Language Augments Human Computation. In: *Language and Thought: Interdisciplinary Themes*. Cambridge: Cambridge University Press. pp.33–51.

Clark, A. and Chalmers, D. (1998). The extended mind. *analysis*, 58: 7–19.

Clark, B. (2013). *Relevance theory*. Cambridge: Cambridge University Press. http://capitadiscovery.co.uk/brighton-ac/items/1332621. Accessed 3 May 2017.

Cowan, N. (2016). *Working Memory Capacity: Classic Edition*. Routledge.

De Millo, R. A., Lipton, R. J. and Perlis, A. J. (1979). Social Processes and Proofs of Theorems and Programs. *Commun. ACM*, 22: 271–280.

Descartes, R. (1641). Meditations on First Philosophy. In: *The philosophical writings of Descartes*. Cambridge University Press. pp.1–62.

Dreyfus, H. L. (2002). Intelligence without representation–Merleau-Ponty's critique of mental representation The relevance of phenomenology to scientific explanation. *Phenomenology and the cognitive sciences*, 1: 367–383.

Duff, P. A. (2010). Language socialization into academic discourse communities. *Annual review of applied linguistics*, 30: 169–192.

Elgin, C. (2017). Nature's Handmaid, Art. In: O. Bueno, G. Darby, S. French & D. Rickles (eds) *Thinking about Science, Reflecting on Art: Bringing Aesthetics and Philosophy of Science Together* (1 edition). London ; New York: Routledge.

Elman, J. (1993). Learning and Development in Neural Networks: The Importance of Starting Small. *Cognition*: 71–99.

Engle, R. W. (2002). Working Memory Capacity as Executive Attention. *Current Directions in Psychological Science*, 11: 19–23.

Engle, R. W. and Kane, M. J. (2004). Executive attention, working memory capacity, and a two-factor theory of cognitive control. *Psychology of learning and motivation*, 44: 145–200.

Feyerabend, P. K. (1975). *Against Method* (4th edition). London ; New York: Verso.

Forceville, C. (1998). *Pictorial Metaphor in Advertising* (1 edition). London: Routledge.

Forceville, C. (2014). *Relevance Theory as a model for analysing multimodal communication, in Visual Communication, David Machin (ed.).* Walter de Gruyter GmbH & Co KG.

Forceville, C., Clark, B., Forceville, C. and Clark, B. (2014). CAN PICTURES HAVE EXPLICATURES? *Linguagem em (Dis)curso*, 14: 451–472.

Foster, H. (2004). An Archival Impulse. *October*, 110: 3–22.

Foster, H. (1996). The Artist as Ethnographer? In: *The Return of the Real*. Cambridge, Mass.: MIT Press. pp.302–308.

Francis, D. and Hester, S. (2004). *An invitation to ethnomethodology: Language, society and interaction*. Sage.

Galison, P. (1997). Three laboratories. *Social Research*: 1127–1155.

Gallese, V. (2007). Before and below 'theory of mind': embodied simulation and the neural correlates of social cognition. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 362: 659–669.

Garfinkel, H. (2002). *Ethnomethodology's program: Working out Durkheim's aphorism*. Rowman & Littlefield Publishers.

Garfinkel, H. (1967). Studies in ethnomethodology.

Garfinkel, H., Lynch, M. and Livingston, E. (1981). I. 1 The Work of a Discovering Science Construed with Materials from the Optically Discovered Pulsar. *Philosophy of the social sciences*, 11: 131–158.

Giardino, V. (2013). A Practice-Based Approach to Diagrams. In: A. Moktefi & S.-J. Shin (eds) *Visual Reasoning with Diagrams*. Springer Basel. pp.135–151. http://link.springer.com/chapter/10.1007/978-3-0348-0600-8_8. Accessed 24 August 2016.

Giardino, V. (2018). Manipulative imagination: how to move things around in mathematics. *Theoria*, 33: 345–360.

Gibson, J. J. (2014). *The ecological approach to visual perception: classic edition*. Psychology Press.

Gibson, J. J. (1966). The senses considered as perceptual systems.

Goldman, A. I. (2012). *A moderate approach to embodied cognitive science. Review of Philosophy and Psychology, 3 (1), 71–88*.

Greiffenhagen, C. (2014). The materiality of mathematics: Presenting mathematics at the blackboard. *The British Journal of Sociology*, 65: 502–528.

Greiffenhagen, C. (2008). Video Analysis of Mathematical Practice? Different Attempts to 'Open Up' Mathematics for Sociological Investigation. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research* (on-line), 9. http://www.qualitative-research.net/index.php/fqs/article/view/1172. Accessed 15 October 2015.

Greiffenhagen, C. and Sharrock, W. (2011). Does mathematics look certain in the front, but fallible in the back? *Social Studies of Science*, 41: 839–866.

Grice, H. P. (1975). Logic and conversation. In: P. Cole & J. Morgan (eds) *Syntax and semantics*. New York: Academic Press. pp.41–58.

Grice, H. P. (1967). Logic and conversation. William James lectures.

Grice, H. P. (1957). Meaning. *The philosophical review*, 66: 377–388.

Grice, H. P. (1989). *Studies in the Way of Words*. Harvard University Press.

Gustafson, D. L., Parsons, J. E. and Gillingham, B. (2019). Writing to transgress: knowledge production in feminist participatory action research. In: *Forum Qualitative Sozialforschung/Forum: Qualitative Social Research*. DEU. p.25.

Habgood-Coote, J. (2019). Knowledge-How, Abilities, and Questions. *Australasian Journal of Philosophy*, 97: 86–104.

Hamami, Y. (2014). Mathematical rigor, proof gap and the validity of mathematical inference. *Philosophia Scientiæ. Travaux d'histoire et de philosophie des sciences*: 7–26.

Hersh, R. (1991). Mathematics has a front and a back. *Synthese*, 88: 127–133.

Hersh, R. (1997). *What is Mathematics, Really?* Oxford: Oxford University Press.

Hofstadter, D. R. (1980). *Gödel, Escher, Bach: an eternal golden braid:[a metaphorical fugue on minds and machines in the spirit of Lewis Carroll]*. Penguin Books New York.

Hutchins, E. (1995). *Cognition in the Wild*. MIT Press.

Hutchins, E. (2010). Cognitive Ecology. *Topics in Cognitive Science*, 2: 705–715.

Hymes, D. (1964). Introduction: Toward Ethnographies of Communication1. *American Anthropologist*, 66: 1–34.

Jansen, A. R., Marriott, K. and Yelland, G. W. (2003). Comprehension of algebraic expressions by experienced users of mathematics. *The Quarterly Journal of Experimental Psychology Section A*, 56: 3–30.

Jansen, A. R., Marriott, K. and Yelland, G. W. (2007). Parsing of algebraic expressions by experienced users of mathematics. *European Journal of Cognitive Psychology*, 19: 286–320.

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

Kuhn, T. S. (1962). The Structure of Scientific Revolutions Vol.

Lakatos, I. (1976). *Proofs and Refutations: The Logic of Mathematical Discovery*. Cambridge University Press.

Landy, D., Allen, C. and Zednik, C. (2014). A perceptual account of symbolic reasoning. *Frontiers in Psychology* (on-line), 5. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4001060/. Accessed 13 March 2017.

Landy, D. and Goldstone, R. L. (2007). How abstract is symbolic thought? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33: 720–733.

Lane, L. (2016). The Bridge Between Worlds: Relating Position and Disposition in the Mathematical Field. Edinburgh: University of Edinburgh.

Larvor, B. (2016). Why the naïve derivation recipe model cannot explain how mathematicians' proofs secure mathematical knowledge. *Philosophia mathematica*, 24: 401–404.

Latour, B. and Woolgar, S. (2013). *Laboratory Life: The Construction of Scientific Facts*. Princeton University Press.

Laudan, L. (2004). The epistemic, the cognitive, and the social. *Science, values, and objectivity*: 14–23.

Levinson, S. C. and Levinson, S. C. (1983). *Pragmatics*. Cambridge University Press.

Livingston, E. (1999). Cultures of Proving. *Social Studies of Science*, 29: 867–888.

Livingston, E. (2006). The Context of Proving. *Social Studies of Science*, 36: 39–68.

Livingston, E. (2015). The Disciplinarity of Mathematical Practice. *Journal of Humanistic Mathematics*, 5: 198–222.

Livingston, E. (1989). The ethnomethodological foundations of mathematics.

MacColl, H. (1880). Symbolical Reasoning. *Mind*, 5: 45–60.

Mancosu, P. (ed.) (2008). *The Philosophy of Mathematical Practice* (1 edition). Oxford ; New York: Oxford University Press.

McCallum, K., Mitchell, S. and Scott-Phillips, T. (2019). The Art Experience. *Review of Philosophy and Psychology* (on-line). https://doi.org/10.1007/s13164-019-00443-y. Accessed 30 October 2019.

Mercier, H. and Sperber, D. (2011). Why do humans reason? Arguments for an argumentative theory. *Behavioral and brain sciences*, 34: 57–74.

Merz, M. and Cetina, K. K. (1997). Deconstruction in a `Thinking' Science: Theoretical Physicists at Work. *Social Studies of Science*, 27: 73–111.

Moktefi, A. (2017). Diagrams as scientific instruments. In: A. Benedek & A. Veszelszki (eds) *Virtual Reality - Real Visuality: Visual, Virtual, Veridical*. Frankfurt: Peter Lang GmbH. pp.81–89.

Muntersbjorn, M. M. (1999). Naturalism, Notation, and the Metaphysics of Mathematics. *Philosophia Mathematica*, 7: 178–199.

Muntersbjorn, M. M. (2003). Representational Innovation and Mathematical Ontology. *Synthese*, 134: 159–180.

Murdoch, I. (2001). *The Sovereignty of Good* (2 edition). Princeton, NJ: Routledge.

Murdoch, I. (1959). The Sublime and the Good. *Chicago Review*, 13: 42–55.

Noë, A. (2004). *Action in Perception*. MIT Press.

Noë, A. (2000). Experience and experiment in art. *Journal of Consciousness Studies*, 7: 123–136.

Noë, A. (2009). *Out of our heads: Why you are not your brain, and other lessons from the biology of consciousness*. Macmillan.

Noë, A. (2015). *Strange Tools: Art and Human Nature*. New York: Hill & Wang.

Numarkee (ed.) (2015). Appendix Transcription Conventions in Conversation Analysis. In: *The Handbook of Classroom Discourse and Interaction*. John Wiley & Sons, Inc. pp.527–528. http://onlinelibrary.wiley.com/doi/10.1002/9781118531242.app1/summary. Accessed 8 May 2017.

Numberphile *The World's Best Mathematician (*) - Numberphile*. https://www.youtube.com/watch?v=MXJ-zpJeY3E. Accessed 3 October 2017.

Ogden, L. A. (2011). *Swamplife: people, gators, and mangroves entangled in the Everglades*. University of Minnesota Press.

O'Riley, T. (2011). A Discrete Continuity: On the Relation Between Research and Art Practice. *Journal of Research Practice*, 7: 1.

Pandian, A. and McLean, S. J. (2017). *Crumpled paper boat: Experiments in ethnographic writing*. Duke University Press.

Pignocchi, A. (2018). The Continuity between Art and Everyday Communication. *Advances in Experimental Philosophy of Aesthetics*: 241.

Pimm, D. (1987). *Speaking Mathematically: Communication in Mathematics Classrooms*. Routledge.

Rancière, J. (2013). *The Politics of Aesthetics*. A&C Black.

Rawls, A. W. (2008). Harold Garfinkel, Ethnomethodology and Workplace Studies. *Organization Studies*, 29: 701–732.

Richardson, L. and St Pierre, E. (2008). A method of inquiry. *Collecting and interpreting qualitative materials*, 3: 473.

Rotman, B. (2000). *Mathematics as Sign: Writing, Imagining, Counting*. Stanford University Press.

Rotman, B. (1998). The Technology of Mathematical Persuasion. In: *Inscribing Science: Scientific Texts and the Materiality of Communication, ed. Lenoir, Timothy*. California: Stanford University Press.

Russell, B. and Whitehead, A. N. (1963). *Principia Mathematica*. Cambridge: Cambridge University Press.

Saussure, F. de, Sechehaye, A., Bally, C. and Reidlinger, A. (1974). *Course in general linguistics* (Rev. ed.). London: Fontana. http://capitadiscovery.co.uk/brighton-ac/items/120622.

Saville-Troike, M. (2008). *The ethnography of communication: An introduction*. John Wiley & Sons.

Shannon, C. E. and Weaver, W. (1949). *The mathematical theory of communication*. Urbana, IL: University of Illinois Press.

Shapiro, S. (2001). *Thinking About Mathematics: The Philosophy of Mathematics*. New York: Oxford University Press, U.S.A.

Simon, H. A. (1981). *The sciences of the artificial.* MIT press.

Sperber, D. and Wilson, D. (2015). Beyond Speaker's Meaning. *Croatian Journal of Philosophy*, 15: 117–149.

Sperber, D. and Wilson, D. (1995). *Relevance: Communication and Cognition*. Wiley.

Stretcher | Features | Nicolas Bourriaud and Karen Moss. http://www.stretcher.org/features/nicolas_bourriaud_and_karen_moss/. Accessed 9 March 2017.

The Individual and the Organisation: Artist Placement Group - Announcements - e-flux. https://www.e-flux.com/announcements/33678/the-individual-and-the-organisation-artist-placement-group/. Accessed 21 February 2019.

Traweek, S. (1988). *Beamtimes and lifetimes*. Harvard University Press.

Tsing, A. L. (2015). *The mushroom at the end of the world: On the possibility of life in capitalist ruins*. Princeton University Press.

Tymoczko, T. (1998). *New Directions in the Philosophy of Mathematics: An Anthology*. Princeton University Press.

University of Bristol *Some random thoughts*. https://www.youtube.com/watch?v=nHMpYYDZv60. Accessed 2 October 2017.

Villani, C. (2012). L'écriture des mathématiciens. *É. Guichard*: 199–212.

Wharton, T. (2008). "MeaningNN" and "showing": Gricean intentions and relevance-theoretic intentions. *Intercultural Pragmatics* (on-line), 5. http://www.degruyter.com/view/j/iprg.2008.5.issue-2/ip.2008.008/ip.2008.008.xml. Accessed 4 May 2016.

Wilson, D. (2017). Wilson, D. (2017). Communication, comprehension and 'non-propositional effects'.

Wilson, D. and Sperber, D. (2005). Relevance Theory. In: *The Handbook of Pragmatics* (1st Edition). Oxford: Wiley-Blackwell.

Bibliography

Anderson, G., Buck, D., Coates, T. and Corti, A. (2015). Drawing in Mathematics: From Inverse Vision to the Liberation of Form. *Leonardo*, 48: 439–448.

Anderson, M., Meyer, B. and Olivier, P. (eds) (2002). *Diagrammatic Representation and Reasoning*. London: Springer London. http://link.springer.com/10.1007/978-1-4471-0109-3. Accessed 30 August 2016.

Asprey, W. and Kitcher, P. (1988). *History and Philosophy of Modern Mathematics*. Minneapolis: University of Minnesota Press.

Baddeley, A. (2010). Working memory. *Current Biology*, 20: R136–R140.

Barany, M. J. (2010). *Mathematical Research in Context, MSc by research*. Edinburgh: University of Edinburgh.

Barany, M. J. and MacKenzie, D. (2014). Chalk: Materials and Concepts in Mathematics Research. In: C. Coopman, J. Vertesi, M. Lynch & S. Woolgar (eds) *Representation in Scientific Practice Revisited*. The MIT Press. pp.107–130. http://mitpress.universitypressscholarship.com/view/10.7551/mitpress/9780262525381.001.0001/upso-9780262525381-chapter-6. Accessed 27 February 2017.

Barrouillet, P., Bernardin, S. and Camos, V. (2004). Time constraints and resource sharing in adults' working memory spans. *Journal of Experimental Psychology: General*, 133: 83.

Barsalou, L. W. (2008). Grounded cognition. *Annu. Rev. Psychol.*, 59: 617–645.

Bernard, H. (1974). Scientists and policy makers: An ethnography of communication. *Human Organization*, 33: 261–276.

Bloor, D. (1987). The Living Foundations of Mathematics E. Livingston (ed.). *Social Studies of Science*, 17: 337–358.

Borgdorff, H. (2006). *The debate on research in the arts*. Kunsthøgskolen i Bergen Bergen, Norway.

Bourguignon, J.-P. and Casse, M. (2012). *Mathematics, A Beautiful Elsewhere* (1 edition). Paris : New York: Thames & Hudson.

Bourriaud, N. (1998). *Relational Aesthetics*. Dijon: Les Presse Du Reel.

Brown, J. R. (2008). *Philosophy of Mathematics: A Contemporary Introduction to the World of Proofs and Pictures* (2 edition). New York: Routledge.

Cannon, J. W., Floyd, W. J. and Parry, W. R. (1996). Introductory notes on Richard Thompson's groups. *Enseignement Mathématique*, 42: 215–256.

Carston, R. (2009). The Explicit/Implicit Distinction in Pragmatics and the Limits of Explicit Communication. *International Review of Pragmatics*, 1: 35–62.

Cicourel, A. V. (1987). The Interpenetration of Communicative Contexts: Examples from Medical Encounters. *Social Psychology Quarterly*, 50: 217–226.

Clark, A. (1998). Magic Words: How Language Augments Human Computation. In: *Language and Thought: Interdisciplinary Themes*. Cambridge: Cambridge University Press. pp.33–51.

Clark, A. and Chalmers, D. (1998). The extended mind. *analysis*, 58: 7–19.

Clark, B. (2013). *Relevance theory*. Cambridge: Cambridge University Press. http://capitadiscovery.co.uk/brighton-ac/items/1332621. Accessed 3 May 2017.

Cowan, N. (2016). *Working Memory Capacity: Classic Edition*. Routledge.

De Millo, R. A., Lipton, R. J. and Perlis, A. J. (1979). Social Processes and Proofs of Theorems and Programs. *Commun. ACM*, 22: 271–280.

Descartes, R. (1641). Meditations on First Philosophy. In: *The philosophical writings of Descartes*. Cambridge University Press. pp.1–62.

Dreyfus, H. L. (2002). Intelligence without representation–Merleau-Ponty's critique of mental representation The relevance of phenomenology to scientific explanation. *Phenomenology and the cognitive sciences*, 1: 367–383.

Duff, P. A. (2010). Language socialization into academic discourse communities. *Annual review of applied linguistics*, 30: 169–192.

Elgin, C. (2017). Nature's Handmaid, Art. In: O. Bueno, G. Darby, S. French & D. Rickles (eds) *Thinking about Science, Reflecting on Art: Bringing Aesthetics and Philosophy of Science Together* (1 edition). London ; New York: Routledge.

Elman, J. (1993). Learning and Development in Neural Networks: The Importance of Starting Small. *Cognition*: 71–99.

Engle, R. W. (2002). Working Memory Capacity as Executive Attention. *Current Directions in Psychological Science*, 11: 19–23.

Engle, R. W. and Kane, M. J. (2004). Executive attention, working memory capacity, and a two-factor theory of cognitive control. *Psychology of learning and motivation*, 44: 145–200.

Feyerabend, P. K. (1975). *Against Method* (4th edition). London ; New York: Verso.

Forceville, C. (1998). *Pictorial Metaphor in Advertising* (1 edition). London: Routledge.

Forceville, C. (2014). *Relevance Theory as a model for analysing multimodal communication, in Visual Communication, David Machin (ed.)*. Walter de Gruyter GmbH & Co KG.

Forceville, C., Clark, B., Forceville, C. and Clark, B. (2014). CAN PICTURES HAVE EXPLICATURES? *Linguagem em (Dis)curso*, 14: 451–472.

Foster, H. (2004). An Archival Impulse. *October*, 110: 3–22.

Foster, H. (1996). The Artist as Ethnographer? In: *The Return of the Real*. Cambridge, Mass.: MIT Press. pp.302–308.

Francis, D. and Hester, S. (2004). *An invitation to ethnomethodology: Language, society and interaction*. Sage.

Galison, P. (1997). Three laboratories. *Social Research*: 1127–1155.

Gallese, V. (2007). Before and below 'theory of mind': embodied simulation and the neural correlates of social cognition. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 362: 659–669.

Garfinkel, H. (2002). *Ethnomethodology's program: Working out Durkheim's aphorism*. Rowman & Littlefield Publishers.

Garfinkel, H. (1967). Studies in ethnomethodology.

Garfinkel, H., Lynch, M. and Livingston, E. (1981). I. 1 The Work of a Discovering Science Construed with Materials from the Optically Discovered Pulsar. *Philosophy of the social sciences*, 11: 131–158.

Giardino, V. (2013). A Practice-Based Approach to Diagrams. In: A. Moktefi & S.-J. Shin (eds) *Visual Reasoning with Diagrams*. Springer Basel. pp.135–151. http://link.springer.com/chapter/10.1007/978-3-0348-0600-8_8. Accessed 24 August 2016.

Giardino, V. (2018). Manipulative imagination: how to move things around in mathematics. *Theoria*, 33: 345–360.

Gibson, J. J. (2014). *The ecological approach to visual perception: classic edition*. Psychology Press.

Gibson, J. J. (1966). The senses considered as perceptual systems.

Goldman, A. I. (2012). *A moderate approach to embodied cognitive science. Review of Philosophy and Psychology, 3 (1), 71–88*.

Greiffenhagen, C. (2014). The materiality of mathematics: Presenting mathematics at the blackboard. *The British Journal of Sociology*, 65: 502–528.

Greiffenhagen, C. (2008). Video Analysis of Mathematical Practice? Different Attempts to 'Open Up' Mathematics for Sociological Investigation. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research* (on-line), 9. http://www.qualitative-research.net/index.php/fqs/article/view/1172. Accessed 15 October 2015.

Greiffenhagen, C. and Sharrock, W. (2011). Does mathematics look certain in the front, but fallible in the back? *Social Studies of Science*, 41: 839–866.

Grice, H. P. (1975). Logic and conversation. In: P. Cole & J. Morgan (eds) *Syntax and semantics*. New York: Academic Press. pp.41–58.

Grice, H. P. (1967). Logic and conversation. William James lectures.

Grice, H. P. (1957). Meaning. *The philosophical review*, 66: 377–388.

Grice, H. P. (1989). *Studies in the Way of Words*. Harvard University Press.

Gustafson, D. L., Parsons, J. E. and Gillingham, B. (2019). Writing to transgress: knowledge production in feminist participatory action research. In: *Forum Qualitative Sozialforschung/Forum: Qualitative Social Research*. DEU. p.25.

Habgood-Coote, J. (2019). Knowledge-How, Abilities, and Questions. *Australasian Journal of Philosophy*, 97: 86–104.

Hamami, Y. (2014). Mathematical rigor, proof gap and the validity of mathematical inference. *Philosophia Scientiæ. Travaux d'histoire et de philosophie des sciences*: 7–26.

Hersh, R. (1991). Mathematics has a front and a back. *Synthese*, 88: 127–133.

Hersh, R. (1997). *What is Mathematics, Really?* Oxford: Oxford University Press.

Hofstadter, D. R. (1980). *Gödel, Escher, Bach: an eternal golden braid:[a metaphorical fugue on minds and machines in the spirit of Lewis Carroll]*. Penguin Books New York.

Hutchins, E. (1995). *Cognition in the Wild*. MIT Press.

Hutchins, E. (2010). Cognitive Ecology. *Topics in Cognitive Science*, 2: 705–715.

Hymes, D. (1964). Introduction: Toward Ethnographies of Communication1. *American Anthropologist*, 66: 1–34.

Jansen, A. R., Marriott, K. and Yelland, G. W. (2003). Comprehension of algebraic expressions by experienced users of mathematics. *The Quarterly Journal of Experimental Psychology Section A*, 56: 3–30.

Jansen, A. R., Marriott, K. and Yelland, G. W. (2007). Parsing of algebraic expressions by experienced users of mathematics. *European Journal of Cognitive Psychology*, 19: 286–320.

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

Kuhn, T. S. (1962). The Structure of Scientific Revolutions Vol.

Lakatos, I. (1976). *Proofs and Refutations: The Logic of Mathematical Discovery*. Cambridge University Press.

Landy, D., Allen, C. and Zednik, C. (2014). A perceptual account of symbolic reasoning. *Frontiers in Psychology* (on-line), 5. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4001060/. Accessed 13 March 2017.

Landy, D. and Goldstone, R. L. (2007). How abstract is symbolic thought? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33: 720–733.

Lane, L. (2016). The Bridge Between Worlds: Relating Position and Disposition in the Mathematical Field. Edinburgh: University of Edinburgh.

Larvor, B. (2016). Why the naïve derivation recipe model cannot explain how mathematicians' proofs secure mathematical knowledge. *Philosophia mathematica*, 24: 401–404.

Latour, B. and Woolgar, S. (2013). *Laboratory Life: The Construction of Scientific Facts*. Princeton University Press.

Laudan, L. (2004). The epistemic, the cognitive, and the social. *Science, values, and objectivity*: 14–23.

Levinson, S. C. and Levinson, S. C. (1983). *Pragmatics*. Cambridge University Press.

Livingston, E. (1999). Cultures of Proving. *Social Studies of Science*, 29: 867–888.

Livingston, E. (2006). The Context of Proving. *Social Studies of Science*, 36: 39–68.

Livingston, E. (2015). The Disciplinarity of Mathematical Practice. *Journal of Humanistic Mathematics*, 5: 198–222.

Livingston, E. (1989). The ethnomethodological foundations of mathematics.

MacColl, H. (1880). Symbolical Reasoning. *Mind*, 5: 45–60.

Mancosu, P. (ed.) (2008). *The Philosophy of Mathematical Practice* (1 edition). Oxford ; New York: Oxford University Press.

McCallum, K., Mitchell, S. and Scott-Phillips, T. (2019). The Art Experience. *Review of Philosophy and Psychology* (on-line). https://doi.org/10.1007/s13164-019-00443-y. Accessed 30 October 2019.

Mercier, H. and Sperber, D. (2011). Why do humans reason? Arguments for an argumentative theory. *Behavioral and brain sciences*, 34: 57–74.

Merz, M. and Cetina, K. K. (1997). Deconstruction in a `Thinking' Science: Theoretical Physicists at Work. *Social Studies of Science*, 27: 73–111.

Moktefi, A. (2017). Diagrams as scientific instruments. In: A. Benedek & A. Veszelszki (eds) *Virtual Reality - Real Visuality: Visual, Virtual, Veridical*. Frankfurt: Peter Lang GmbH. pp.81–89.

Muntersbjorn, M. M. (1999). Naturalism, Notation, and the Metaphysics of Mathematics. *Philosophia Mathematica*, 7: 178–199.

Muntersbjorn, M. M. (2003). Representational Innovation and Mathematical Ontology. *Synthese*, 134: 159–180.

Murdoch, I. (2001). *The Sovereignty of Good* (2 edition). Princeton, NJ: Routledge.

Murdoch, I. (1959). The Sublime and the Good. *Chicago Review*, 13: 42–55.

Noë, A. (2004). *Action in Perception*. MIT Press.

Noë, A. (2000). Experience and experiment in art. *Journal of Consciousness Studies*, 7: 123–136.

Noë, A. (2009). *Out of our heads: Why you are not your brain, and other lessons from the biology of consciousness*. Macmillan.

Noë, A. (2015). *Strange Tools: Art and Human Nature*. New York: Hill & Wang.

Numarkee (ed.) (2015). Appendix Transcription Conventions in Conversation Analysis. In: *The Handbook of Classroom Discourse and Interaction*. John Wiley & Sons, Inc. pp.527–528.

http://onlinelibrary.wiley.com/doi/10.1002/9781118531242.app1/summary. Accessed 8 May 2017.

Numberphile *The World's Best Mathematician (*) - Numberphile*. https://www.youtube.com/watch?v=MXJ-zpJeY3E. Accessed 3 October 2017.

Ogden, L. A. (2011). *Swamplife: people, gators, and mangroves entangled in the Everglades*. University of Minnesota Press.

O'Riley, T. (2011). A Discrete Continuity: On the Relation Between Research and Art Practice. *Journal of Research Practice*, 7: 1.

Pandian, A. and McLean, S. J. (2017). *Crumpled paper boat: Experiments in ethnographic writing*. Duke University Press.

Pignocchi, A. (2018). The Continuity between Art and Everyday Communication. *Advances in Experimental Philosophy of Aesthetics*: 241.

Pimm, D. (1987). *Speaking Mathematically: Communication in Mathematics Classrooms*. Routledge.

Rancière, J. (2013). *The Politics of Aesthetics*. A&C Black.

Rawls, A. W. (2008). Harold Garfinkel, Ethnomethodology and Workplace Studies. *Organization Studies*, 29: 701–732.

Richardson, L. and St Pierre, E. (2008). A method of inquiry. *Collecting and interpreting qualitative materials*, 3: 473.

Rotman, B. (2000). *Mathematics as Sign: Writing, Imagining, Counting*. Stanford University Press.

Rotman, B. (1998). The Technology of Mathematical Persuasion. In: *Inscribing Science: Scientific Texts and the Materiality of Communication, ed. Lenoir, Timothy*. California: Stanford University Press.

Russell, B. and Whitehead, A. N. (1963). *Principia Mathematica*. Cambridge: Cambridge University Press.

Saussure, F. de, Sechehaye, A., Bally, C. and Reidlinger, A. (1974). *Course in general linguistics* (Rev. ed.). London: Fontana. http://capitadiscovery.co.uk/brighton-ac/items/120622.

Saville-Troike, M. (2008). *The ethnography of communication: An introduction*. John Wiley & Sons.

Shannon, C. E. and Weaver, W. (1949). *The mathematical theory of communication*. Urbana, IL: University of Illinois Press.

Shapiro, S. (2001). *Thinking About Mathematics: The Philosophy of Mathematics*. New York: Oxford University Press, U.S.A.

Simon, H. A. (1981). *The sciences of the artificial*. MIT press.

Sperber, D. and Wilson, D. (2015). Beyond Speaker's Meaning. *Croatian Journal of Philosophy*, 15: 117–149.

Sperber, D. and Wilson, D. (1995). *Relevance: Communication and Cognition*. Wiley.

Stretcher | Features | Nicolas Bourriaud and Karen Moss. http://www.stretcher.org/features/nicolas_bourriaud_and_karen_moss/. Accessed 9 March 2017.

The Individual and the Organisation: Artist Placement Group - Announcements - e-flux. https://www.e-flux.com/announcements/33678/the-individual-and-the-organisation-artist-placement-group/. Accessed 21 February 2019.

Traweek, S. (1988). *Beamtimes and lifetimes*. Harvard University Press.

Tsing, A. L. (2015). *The mushroom at the end of the world: On the possibility of life in capitalist ruins*. Princeton University Press.

Tymoczko, T. (1998). *New Directions in the Philosophy of Mathematics: An Anthology*. Princeton University Press.

University of Bristol *Some random thoughts*. https://www.youtube.com/watch?v=nHMpYYDZv60. Accessed 2 October 2017.

Villani, C. (2012). L’écriture des mathématiciens. *É. Guichard*: 199–212.

Wharton, T. (2008). “MeaningNN” and “showing”: Gricean intentions and relevance-theoretic intentions. *Intercultural Pragmatics* (on-line), 5. http://www.degruyter.com/view/j/iprg.2008.5.issue-2/ip.2008.008/ip.2008.008.xml. Accessed 4 May 2016.

Wilson, D. (2017). Wilson, D. (2017). Communication, comprehension and ‘non-propositional effects’.

Wilson, D. and Sperber, D. (2005). Relevance Theory. In: *The Handbook of Pragmatics* (1st Edition). Oxford: Wiley-Blackwell.

Appendices

Appendix 1. Exploratory mapping in creative practice research

Throughout this research one of my aims was to develop a practice that fit the strengths of genuine artistic ethnography laid out in Chapter 1. These, in summary, were to approach material in a holistic, multimodal way, and employ means of investigation and presentation that go beyond text; to generate insight by putting forward possible worlds or worldviews that are somehow other to or incommensurable with existing ones, but nonetheless demonstrate their consistency; to experimentally examine and challenge systems of perception that shape our picture of the world around us. Catherine Elgin's view that art essentially functions by a kind of exemplification begs the question of what exactly is exemplified, a question that exposes the strength of that characterisation for the range of responses that could be given, and could suit different kinds of art: a sensation; an atmosphere; a perspective; a transgression; an organised thought process; and so on.

As I began work on synthesising the overarching arguments of this thesis, my data analysis method of imitation-with-alterations extended into a kind of mapping practice, a process of passing material through different representative media with each representation making aspects more or less visible. The map, as they say, is not the territory; each instance of mapping puts forward a possible way of seeing the world, a picture that carries implicit in it a particular way of seeing or regime of sense that shapes what is registered and represented and what is not. This mapping practice existed as a research method above and beyond the data analysis. Below I document some key aspects of this practice.

My mapping began with analysing some of the situations I was observing and pushing some of the organisation I was beginning to do of different factors and ideas into the representative medium. For example, Figure 107 shows the progression of a conference call as a constructed environment made up of a set of resources, the colours showing what kind of knowledge or resource is deployed at various point during the exchange. The physical flattening of what was a real situation in an office with human actors is emblematic of the process a researcher must go through to begin to make sense of ethnographic materials; a situation saturated with uncompressible detail must somehow be reduced to be made manageable, limited to certain key questions and factors about which observations can be made. Constructing the scene from components cut from coloured card meant mentally viewing all of the complexities of a moment in a room as cut from a limited few essential types: physical inscription, projections of past and future events, mutual social understanding, and so on.

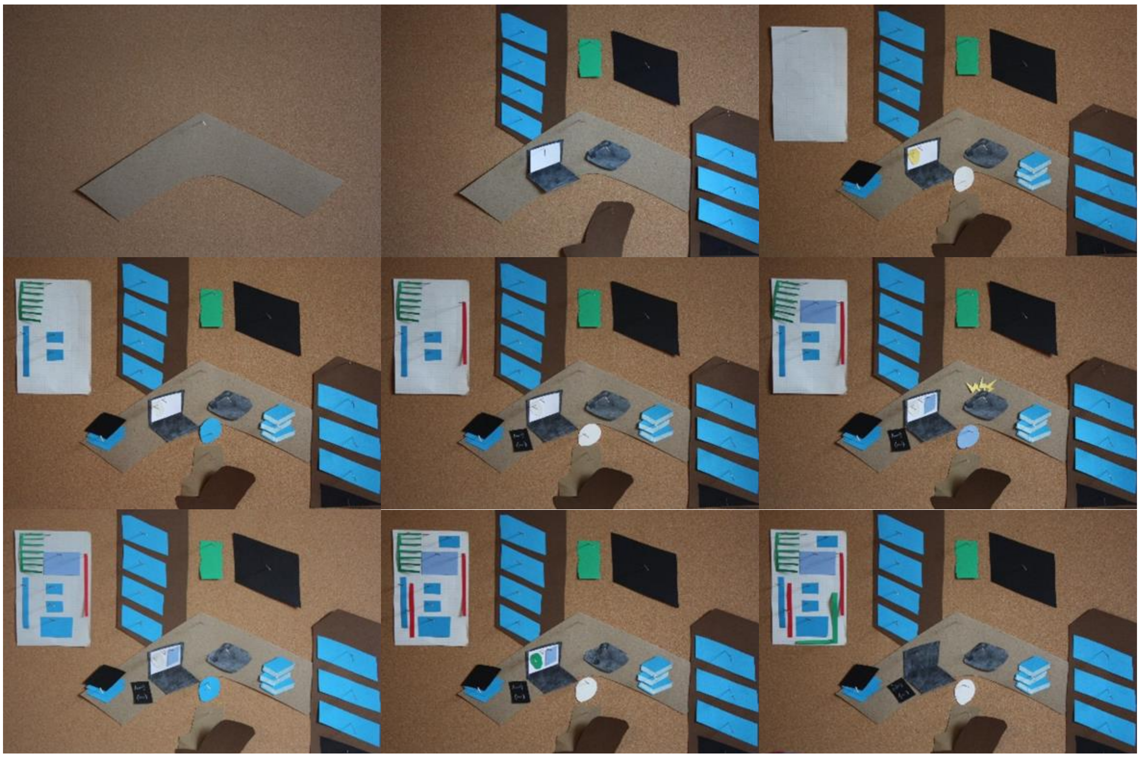


Figure 107. Analysing resources in a conference call. Original in colour.

As I was carrying out my observations, I maintained a chalkboard mapping practice to keep track of the themes emerging in the early research and their situation in the literature (Figure 108). This allowed me to lay out and divide up ideas using not colour but lines and boxes, coming back to these maps day after day to add to and adjust them.

Figure 108. Maps of the research. Original in colour.

These went to shape the ongoing organisation of the materials I was gathering and producing—texts, transcripts, deconstructions and representations—under an archive (Figure 109). I found it interesting and not coincidental that as I observed mathematicians working away on chalkboards in offices filled with stacks of papers, folders and books, my investigative environment followed that aesthetic closely, itself filling with sheets of diagrams and dusty chalkboards.

Figure 109. Archive of my investigative materials. Original in colour.

This archive was housed in a set of shelves and housed a variety of records, from collected texts to records of new experiments, and also the conjunction of the two, a variety of practices of 'laying out' ideas and representations in space. A part of my investigation was to take some of those paper artefacts and expand them, building representations that were laid out in space and focused on one particular aspect above others. For example in Figure 110 I built maps of the internal referencing used in various mathematical papers, each line a reference from one page to another, to build a picture of the structure of the argument and concentration of key points in the documents. Figure 111 shows the map that I built of the various sections of a particular paper, each section heading type (such as Lemma, Proposition, Conjecture) represented by a component of a particular shape.

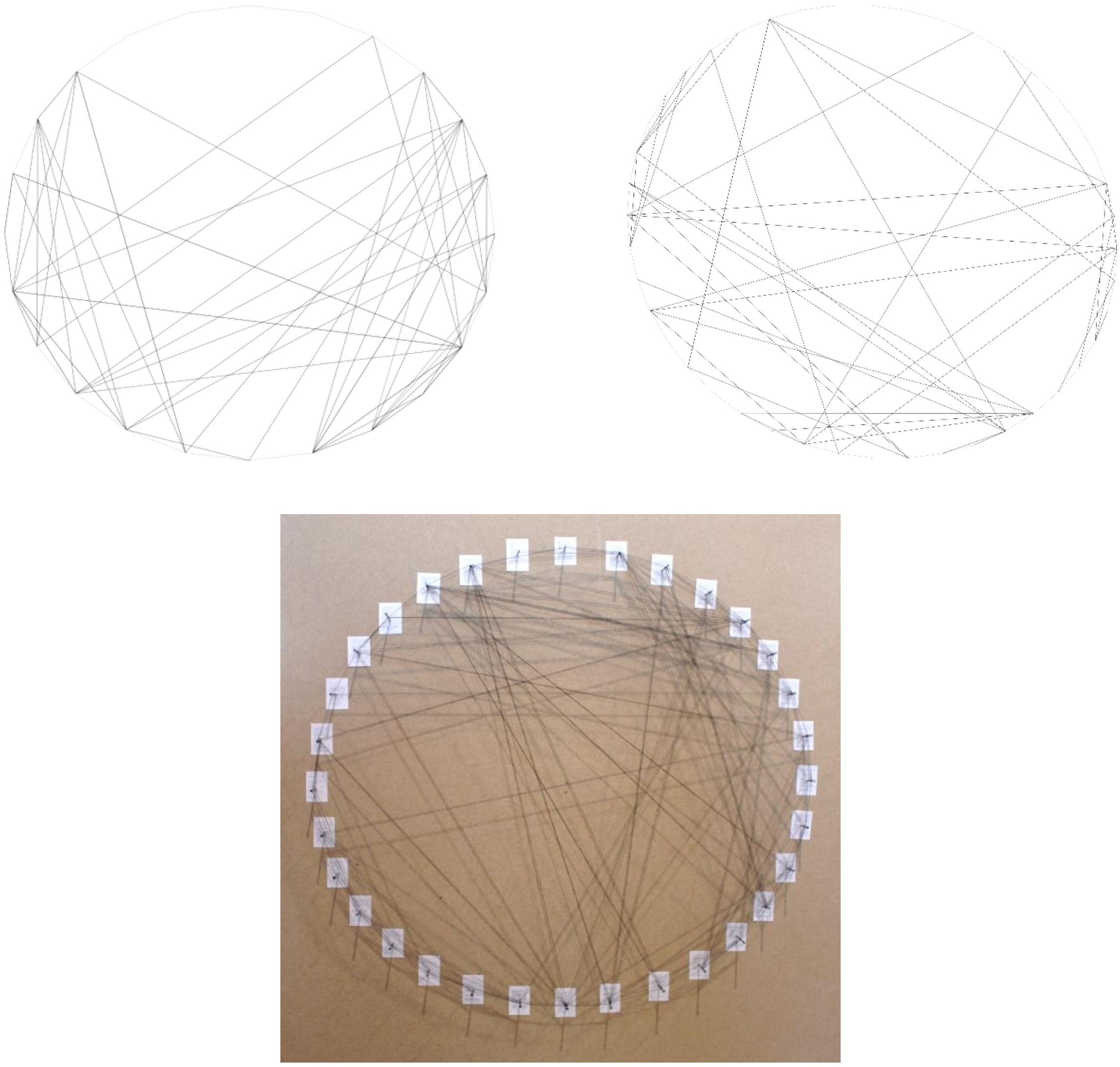


Figure 110. Maps of referencing in papers. Original in colour.

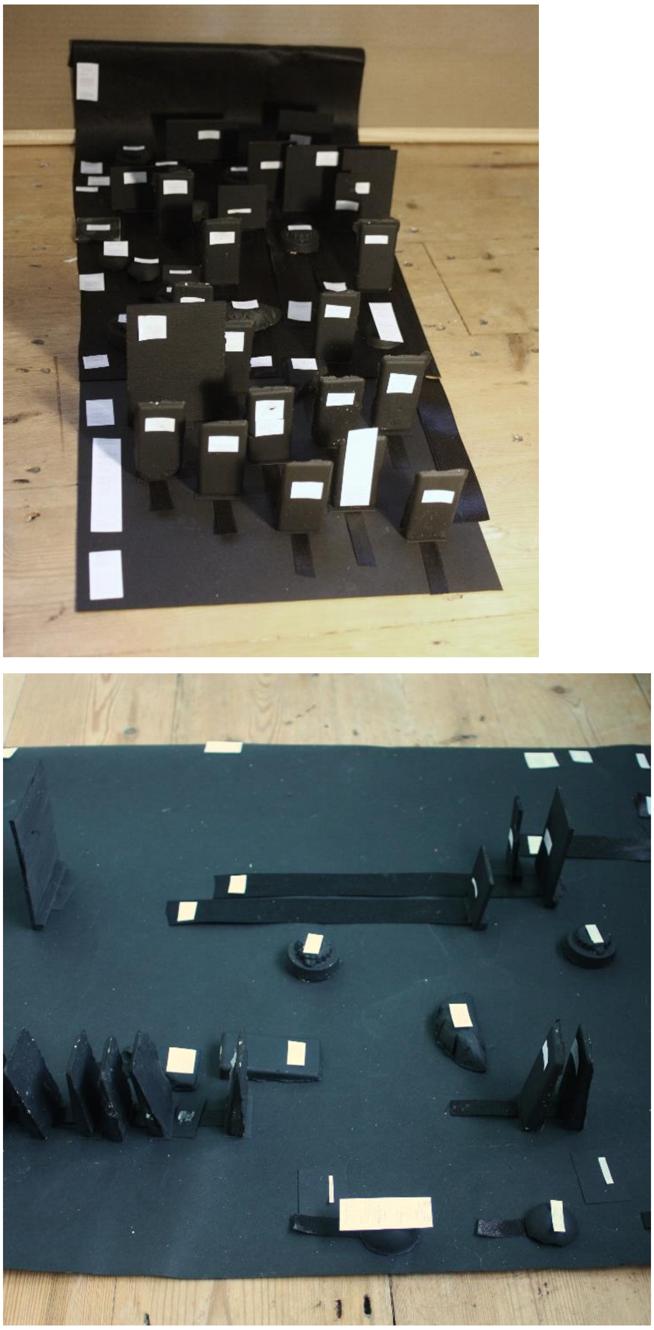


Figure 111. Model of the components of a paper. Original in colour.

These maps show what it looks like to think of a mathematical artefact in some particular limited way, to care only about which parts are referenced over and over again as the argument develops, or about the conceptual structuring that labelling decisions effect on the body of a text. Each also constitutes a raft of representational decisions that turn a paper into a graph or constellation or forest. Unlike textual description, modelling in this way produced physical objects with their own affordances, with the capacity to become invitations to interact according to particular parameters.

Some of the experiments in the archive were exercises that I set myself in producing organisation by mark-making. These exercises were simplified models of the kind of organised activity that Alva Noë describes as allowing us to order and develop our world through back-and-forth with some medium (see Chapter 1), roadmaps of a simplified mind-environment interaction that produced objects that manifested production by a working hand as well as structure that is generative and facilitating of a newly ordered outlook. In school mathematics, the square grid of the exercise book is the essential background to all mathematical work, the baseline from which all mathematical operations can proceed. This grid was something I used as an emblem of the general background of mathematical thinking, something that is so basic as to be almost invisible but that might itself be built and that work made visible. I set myself exercises in manually producing grids by hand, building up that baseline from scratch, wobbly hand-drawn lines produced in aggregate adding up to something with an air of organisation (Figure 112, Figure 113 and Figure 114). In another experiment I made space for decision-making in these organising drawings, finding multiple ways to render an environment in gridlines according to different principles (Figure 115).

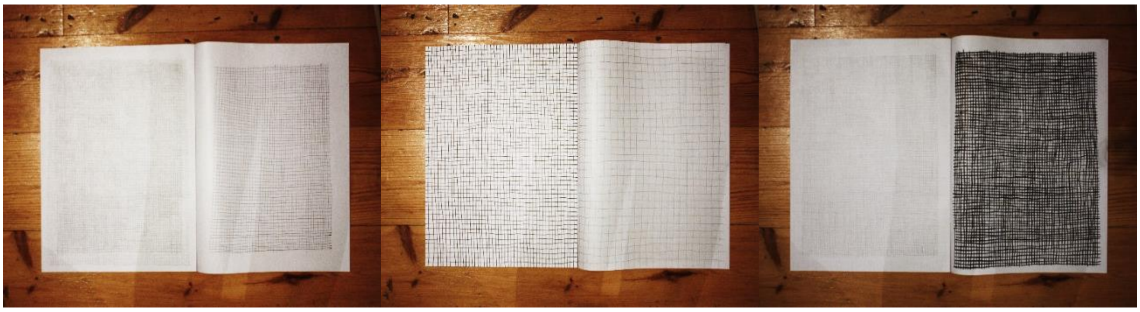


Figure 112. Drawing grids. Original in colour.

Figure 113. Drawing grids. Original in colour.

Figure 114. Handmade exercise book. Original in colour.

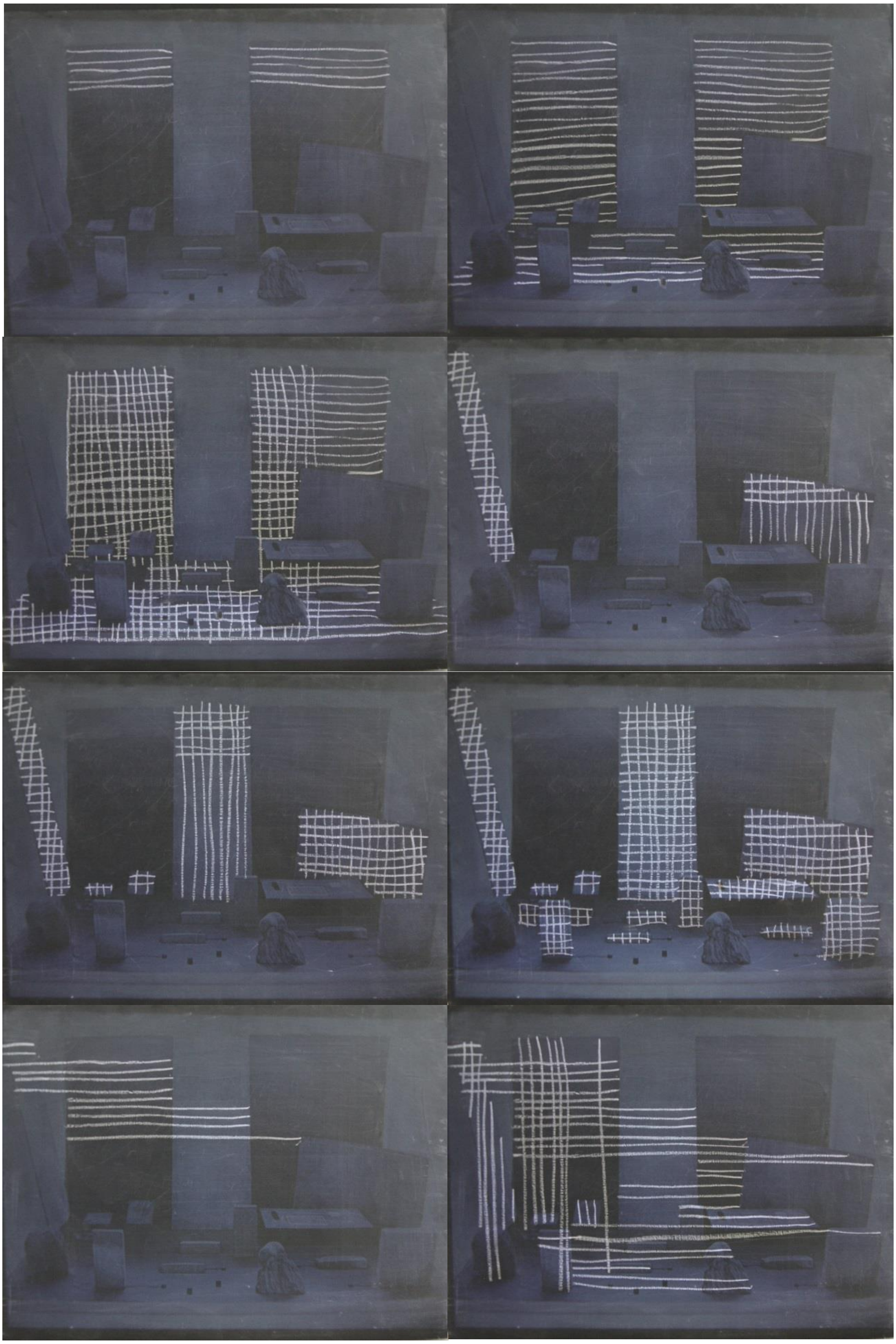


Figure 115. Grid drawings. . Original in colour.

Partway through the research I exhibited some of my ongoing analytic work in the mathematics department of the University of Brighton. One of the aims of the research stated in Chapter 1 was to make the research relevant to the contemporary work of mathematicians and this exhibition aimed to engage that audience and make available some of the exploratory practice of the early research. The exhibit took the form of a set of investigative 'stations', reminiscent of the interactive exhibits found at science museums (Figure 116). Each 'station' corresponded to a particular situation of mathematical work (a person working alone, a collaboration, a publication, a presentation), and were each a simple box with cut-out sections that accommodated a variety of practical experiments related to that situation, such as testing how short a chalk stroke could be and make a mark, or how comprehensible a sequence of mathematical expressions would be with all spatial organisation removed. Each station also included some kind of map of the situation it represented. The 'museum exhibit' presentation suggested itself as an intuitive counterpart to the archiving practice I had been engaging in, a way to lay out the records of a set of engagements with ethnographic material, to build a space for an audience to stand back from and gaze at the everyday practices of mathematical work.

Figure 116. Museum exhibit. Original in colour.

The setup of the exhibition referenced the interactive stations seen in science museums in which visiting children (and adults) are encouraged to engage in activities to see the truth of some scientific insight for themselves. In setting the 'stations' up in this way I hoped to provoke a kind of active, experimental mindset, asking visitors not to carry out some particular task but to view mathematics as something open to experimentation, but at the same time pushing viewers to the position of an interested observer. The intended audience were insiders to the field, so a challenge was to make strange that which, to them, must be very familiar. Some of the components were set up in such a way as to invite imaginative interaction, the chalk used to make the marks still present, the components standing ready; yet no particular task was proposed. Asking mathematicians to write with chalk would see them doing just what they always do. Asking them to consider the notion of writing with chalk might have a different effect.

Figure 117 shows *Collaboration Booth*, a map of the essential features of collaborative infrastructure. Two actors are placed on the scene by two empty chairs, though they themselves are not represented.

Before them is a shared writing space that is angled between their two 'bases', a venue for shared representations to be built and amended. Below it is an avenue for direct communication in which the actors can address one another, sending notecards back and forth. These notecards can also be affixed to the shared writing space, anchoring the commentary to particular aspects of the shared representations. A space is set up for imaginative interaction but it cannot be inhabited, the two chairs being half size for an average adult. The viewer is thus shown the components of interaction but kept on the observing outside, forced into the ethnographer's position on their own world.

Figure 117. Collaboration Booth. Original in colour.

The problem with using the language of a museum exhibit is that any audience is all too accustomed to being placed in the position of passive observers, having wisdom imparted by an unseen, didactic authority. While the archival spaces of Thomas Hirschhorn, Tacita Dean and Sam Durant were described by Hal Foster as setting up utopian spaces out of the records of past and even failed endeavours (Foster, 2004), a museum exhibit seems doomed to an essential inertness, a failure to admit the possibility of action on the part of an audience. A map should made to be used, an object that has been built to facilitate organised action, and while the situation set up in the *Collaboration booth* was one suggestive of action it functioned more as a representation of action, a picture of action, than as a true component in an invited engagement. The natural thing to seek, then, was a way to bring objects into serving as active components in explorations, by setting up mind-environment interactions in different configurations.