JAMIE TAPPENDEN

PROOF STYLE AND UNDERSTANDING IN MATHEMATICS I:
VISUALIZATION, UNIFICATION AND AXIOM CHOICE[0]

To the memory of Heda Segvic

1. INTRODUCTION – A "NEW RIDDLE" OF DEDUCTION

Mathematical investigation, when done well, can confer understanding. This bare observation shouldn't be controversial; where obstacles appear is rather in the effort to engage this observation with epistemology. The complexity of the issue of course precludes addressing it *tout court* in one paper, and I'll just be laying some early foundations here. To this end I'll narrow the field in two ways. First, I'll address a specific account of explanation and understanding that applies naturally to mathematical reasoning: the view proposed by Philip Kitcher and Michael Friedman of explanation or understanding as involving the *unification* of theories that had antecedently appeared heterogeneous. For the second narrowing, I'll take up one specific feature (among many) of theories and their basic concepts that is sometimes taken to make the theories and concepts preferred: in some fields, for some problems, what is counted as understanding a problem may involve finding a way to represent the problem so that it (or some aspect of it) can be *visualized*. The final section develops a case study which exemplifies the way that this consideration – the potential for visualizability – can rationally inform decisions as to what the proper framework and axioms should be.

The discussion of unification (in sections 3 and 4) leads to a mathematical analogue of Goodman's problem of identifying a principled basis for distinguishing grue and green. Just as there is a philosophical issue about how we arrive at the predicates we should use when making empirical predictions, so too there is an issue about what properties best support many kinds of mathematical reasoning that are especially valuable to us. The issue becomes pressing via an examination of some physical and mathematical cases that make it seem unlikely that treatments of unification can be as straightforward as the philosophical literature has hoped. Though unification accounts have a grain of truth (since a phenomenon (or cluster of phenomena) called "unification" *is* in fact important in many cases) we are far from an analysis of what "unification" is. In particular, the degree of unification cannot be usefully taken to turn upon simple syntactic criteria such as counting axioms or argument patterns. I'll argue that existing unification – based accounts need to be supplemented by an account of qualitative distinctions between

1

homogeneous and heterogeneous theories, between "natural" and "artificial" predicates. I'll argue further that in both mathematical and broader scientific practice, rational distinctions between more and less natural properties are made systematically. The considerations brought to bear to rationalize these distinctions are typically more complex and varied than philosophical accounts tend to recognize. But though the principles we rely on to distinguish "natural" from "artificial" categories are varied and case-specific, they are no less rational for all that.[1] I'll emphasize one particular consideration among those that are important in practice: the "natural formulation" of a problem is typically expected to be *fruitful* (to generate further discoveries, "make things easier" and so on).[2] The later parts of the paper take up a specific case: one among the considerations that is sometimes adduced in practice as supporting the fruitfulness of a formulation is that it allows crucial problems or objects to be *visualized*.

The role of visual reasoning, as represented by Artin's *Geometric Algebra*, is complicated in some cases where mathematical axioms and basic categories are chosen. This section is meant also to serve a particular dialectical function. Considerations like "fruitfulness" and especially "visualizability" are sometimes classed as "merely subjective" or "merely psychological" or "merely pragmatic", and therefore not of genuine epistemological significance. This dismissive position can gain an initial force because, of course, some cases where diagrams are preferred, or where "fruitful" techniques are recommended, *are* epistemologically uninteresting. One point the concluding case study is meant to illustrate is that the most interesting cases are simply too systematic and far-reaching to be shrugged off in this way. "Visualizability" is a more intricate and more methodologically interesting theoretical property than it seems on a cursory analysis. A similar point holds for the more general case of "fruitfulness": on examination it turns out to be so extensively and systematically embedded in our theoretical practices that to dismiss it on the grounds that it is "psychological" / "subjective" / "pragmatic" is tatamount to a global skepticism of the "why should our best theories be true?" variety. (Since this is everyone's problem, it is not specifically a problem for the student of visualization or fruitfulness.)

The first section is devoted to spelling out why I think these questions are urgent for the student of methodology and why I am approaching it in just the way I am. First I'll note a key historical motivation that will otherwise serve as an unarticulated background point of reference: the mathematical revolution initiated by Riemann in the nineteenth century. The rest of the

section discusses some of the issues the Riemann's revolution in mathematical style raises, to bring out how systematically embedded in mathematical practice these methodological issues are.

2. UNDERSTANDING AND EXPLANATION IN MATHEMATICAL METHODOLOGY: THE TARGET

2.1. Riemann vs Weierstrass: a classic opposition

Toward the end of the nineteenth century, developing into the twentieth, and in some ways continuing to the present, a division emerged in the proof methods in complex analysis that raises compelling questions for mathematical methodology. One approach, whose driving force was Weierstrass, was broadly computational in its outlook: it aimed at finding explicit representations of functions and explicit algorithms to compute their values. The other, initiated by Riemann, has appropriately been called "conceptual": it aimed to describe functions in terms of general properties, and to prove indirect function existence results that need not be tied to explicit representations.[3] The ramifications of this split were (and remain) far-reaching. In the words of the only philosopher to have discussed this development in print: "Mathematics underwent, in the nineteenth century, a transformation so profound that it is not too much to call it a second birth of the subject – its first having occurred among the ancient Greeks..." (Stein (1988)) There were differences in what proof techniques were counted as acceptable, in what connections to physical applications were available, in which generalizations were natural, and even in what definitions were accepted for the basic objects of study.[4] The opposition persisted through several other changes in the mathematical scene and in some respects continues to this day.[5]

This work defined the mathematics of the late nineteenth century and onward. It was so central to the development of contemporary mathematics that it is a challenge to the philosopher, as an analyst of mathematical method, to give an account of what was at stake.[6] This prompts a second challenge: once it is granted that these developments are of philosophical importance, it requires work to find a philosophical niche to put them in.

The historical details and philosophical overtones in the development of complex analysis are too complicated to be addressed in a single paper; this paper is one installment in a divide-and-conquer strategy of slicing the problem into parts. This installment will deal with the questions induced by the following observations: A) among the points of separation between the Riemann and Weierstrass approaches are a cluster of considerations that could reasonably be described as constituting different ways of *understanding* the subject–matter of complex analysis.[7] B) Among the motives for the

Riemann style is that it proved, and has continued to prove, *fruitful* both in facilitating the solution of given problems and in unearthing new important problems. C) The fact that Riemann surfaces (the basic context for analysis in the Riemann style) allow complex functions to be easily *visualized* was, and remains, a contributor to the fruitfulness of the Riemann approach.[8] To keep the discussion relatively simple and independent of specialized background, I'll set the motivating case of complex analysis aside, and address the phenomenon of styles of understanding proving fruitful, with special reference to visualization as a contribution to that fruitfulness in connection with more elementary and tractable examples.[9]

2.2. *Understanding as an objective guiding mathematical investigation*

This methodological analysis is given a special urgency by the observation – widely accepted by those familiar with the events – that in Riemann's work an entirely new style of mathematical thinking appeared, one that (in different subspecies corresponding to different interpretations of the style) has come to be part of the core of contemporary mathematics. In this connection it will be useful to glance at a discussion of some distinctive characteristics of twentieth century mathematics, as they were perceived by Hermann Weyl (who, of course, was right in the middle of this research).

> We are not very pleased when we are forced to accept a mathematical truth by virtue of a complicated chain of formal conclusions and computations, which we traverse blindly, link by link, feeling our way by touch. We want first an overview of the aim and of the road; we want to understand the *idea* of the proof, the deeper context.
>
> . . . Minkowski contrasted the minimum principle that Germans tend to name for Dirichlet. . . with the true Dirichlet principle: to conquer problems with a minimum of blind computation and a maximum of insightful thoughts. It was Dirichlet, said Minkowski, who ushered in the new era in the history of mathematics. (Weyl, 1995, p. 453)[10]

In the essay Weyl contrasts the topological style for approaching Riemann's results – which he takes to be an orthodox descendant of the original Riemann function theory – with algebraic approaches to the subject.[11] (Both approaches are seen as possessing different, contrasting strengths.) Weyl echoes some of the themes noted above, singling out fruitfulness as a criterion marking the "natural" generalizations:

What is the secret of such an understanding of mathematical matters, what does it consist in? Recently, there have been attempts in the philosophy of science to contrast understanding, the art of interpretation as the basis of the humanities, with scientific explanation, and the words intuition and understanding have been invested in this philosophy with a certain mystical halo, and intrinsic depth and immediacy. In mathematics, we prefer to look at things somewhat more soberly.... I can single out, from the many characteristics of the process of understanding, one that is of decisive importance. One separates in a natural way the different aspects of a subject of mathematical investigation, makes each accessible through its own relatively narrow and easily surveyable group of assumptions, and returns to the complex whole by combining the appropriately specialized partial results. This last synthetic step is purely mechanical. The great art is in the first, analytic, step of appropriate separation and generalization. The mathematics of the last few decades has revelled in generalizations and formalizations. *But to think that mathematics pursues generality for the sake of generality is to misunderstand the sound truth that a natural generalization simplifies by reducing the number of assumptions and by thus letting us understand certain aspects of a disarranged whole.* Then it is subjective and dogmatic arbitrariness to speak of the true ground, the true source of an issue. *Perhaps the only criterion of the naturalness of a severance and an associated generalization is their fruitfulness.* (Weyl (1995) p. 454-455 emphasis mine)

Among the many things in these remarks is: Weyl associates understanding with the unification of "aspects of a disarranged whole". I'll revisit this in connection with theories of explanation later. In this section I'll set that point aside, and concentrate on the idea of understanding and the indicated connections to assessments of "naturalness" and "fruitfulness". It is of course a truism that in mathematical practice we seek understanding, not just logically cogent argument. A characteristic case is described by Michael Atiyah in these words:

I remember one theorem that I proved, and yet I really couldn't see why it was true. It worried me for years and years... I kept worrying about it, and five or six years later I understood why it had to be true. Then I got an entirely different proof...

> Using quite different techniques, it was quite clear why it had
> to be true. (Atiyah, 1984, 305)

A proof or proof sketch can give cogent grounds for believing a claim, but it might fail nonetheless to provide the sort of illumination we can hope for in mathematical investigation. It is not unusual, nor is it unreasonable, to be dissatisfied with a proof that doesn't convey understanding and to seek another argument that does. Sometimes one proof may be counted superior to a second even though both proofs are carried out within the same theoretical context (same definitions, primitive concepts, formal or informal axiomatic formulations, etc.) In other cases, notably those I want to consider here, the advantages of one argument over another appear to derive partly from the definitions and/or axioms in terms of which they are framed.[12]

One among the many reasons accepted in practice for preferring one formulation over another is that one way of framing and addressing a topic can be more fruitful than another. When a formulation (set of definitions or axioms, etc.) is found to be fruitful, this fact often is taken as evidence that the formulation in question is the "natural" one, that the problem or subject is properly understood when it is set up this way.[13] Moreover, it is generally a necessary condition on a principle or definition proposed as natural that it support interesting new proofs.[14] Though it is difficult to lay out with draftsman's precision what "fruitfulness" is, the effort to devise fruitful ways of setting up problems and topics is part of what constitutes mathematical activity and makes it valuable to us.[15]

Whether or not a framework is "fruitful" is something that must be borne out by the facts, in the long run. This was a test passed by the Riemann approach to complex analysis: as devices for visually representing complex functions Riemann surfaces were indeed helpful, but they would have been relegated to the marginal shelf labeled "mere handy tricks" had they not consistently, systematically, and in unexpected ways continued to facilitate understanding and discovery. When they continued to bear fruit in novel and unexpected ways they were stably accepted as the proper context in which to study the functions of interest to complex analysis, as "not merely a device for visualizing the many-valuedness of analytic functions, but rather an indispensable essential component of the theory; not a supplement, more or less artificially distilled from the functions, but their native land, the only soil in which the functions grow and thrive." as Weyl put it elsewhere.[16]

That is: if it is judged that some framework or theoretical context is "the right one" because it is judged to confer understanding in a way that facilitates the solution to important problems, the judgement is *revisable*. Subsequent investigation can reveal that the framework provided just the illusion

of understanding, if its advantages prove short lived or one – dimensional while another framework is more flexibly fruitful in the long run.[17] This observation helps address a worry that is especially favored by Wittgenstein enthusiasts: the worry that an assessment of "being right" will be empty if there is no independent criterion of success – if there is no difference between "being right" and "seeming to be right". That might seem to be a problem here if we concentrate solely on the psychological sense of "naturalness" that can be generated by placing a problem in its "natural" context/finding the "natural" formulation. But there is more to being the "proper" context/formulation than just generating that feeling: innovation and proof of important results must be facilitated *in practice*, over the long haul.

What is "natural" about a "natural" formulation typically must be *learned*: it will not be obvious on a simple inspection of the formulation.[18] The natural formulation need not be the one that is most initially attractive to the untutored. This runs counter to a suggestion that has been advanced, and which has some initial plausibility, that to understand a problem or concept is to reduce it to terms that are already familiar.[19] In many of the most interesting mathematical examples (Riemann surfaces in complex analysis, projective geometry for the study of (*inter alia*) conic sections, scheme theory in algebraic geometry, and much else…) what is counted as "understanding" is achieved by reformulating familiar problems in terms which are initially strikingly unfamiliar.

Another related point – one that will be crucial in section 3 and 4, and will be explored in connection with an example in section 5 – is that these judgements of "naturalness" and the like are *reasoned*. It is not just some brute aesthetic response or sudden, irrational "aha!" reaction that brings about the judgement that – for example – "the scheme is the more natural setting for many geometric arguments" (Eisenbud & Harris, 1992, 5) or that the Artin framework considered in section 5 is "the proper setting for many problems in linear algebra." (Hughes & Piper, 1973, 285).[20] Quite the contrary: elaborate reasons can be and are given for and against these choices. One job facing the methodologist of mathematics is to better understand the variety of reasons that can be given, and how such reasons inform mathematical practice.

This factual observation should be beyond controversy: reasoned judgements about the "proper" way to represent and prove a theorem inform mathematical practice. I have found that more contention is generated by the disciplinary classification of the study of these judgements and the principles informing them: is this *philosophy*, or something else, like cognitive psychology? It is hard to frame this worry precisely, since it inherits all the

unclarity of what, precisely, philosophy is. Much of the rest of this paper will tacitly address the issue, but in addition I should say a few words directly.

The point of contention is the classification of advantages of theories that appear to depend on psychological facts about the people who use the theories. There seems to be a relatively widespread view that a property of a theory must be "objective" to be the proper object of a philosopher's attention.[21] (I should note explicitly that I am not endorsing this view, but only reporting that it is widespread enough that it needs to be confronted.) The relevant sense of "objective" is hard to pin down precisely, but one aspect of the idea seems to be that for an advantage to be "objective" in this sense it must be an advantage for every possible inquirer. "Subjective" features of theories – it is maintained – are someone else's problem.[22]

This sort of objection might seem to be an obstacle here. No doubt extraterrestrial beings with different wiring in their brains might differ from us on what formulations of problems are "most natural", and would discover that different ways of setting up problems would be most fruitful, and better at facilitating discovery or "making things easier". And there are *some* cases where mathematicians display nearly uniform preferences that do indeed seem to be of interest only to the student of psychology and not the student of method. The use of "z" as the variable of choice for complex variables is so entrenched it can be distracting if a different letter is chosen, but that is hardly a deep and interesting point of method. Using yellow paper rather than white may reduce fatigue and hence foster creativity, but that is not the methodologist's concern.

But while acknowledging that *some* cases should be shunted into the "not philosophically interesting" scrap yard, it would be a mistake to ignore them all. The situation is similar to the (more widely recognized) one that faces us when evaluating the role of judgements of "simplicity" in general philosophy of science. The assessments of simplicity or fruitfulness we make would no doubt be different if our brains were wired differently, and this would affect the mathematics and science that we produced, but still the judgements we actually make are too systematically embedded in our actual practices to be simply shrugged off in studies of either scientific or mathematical method.[23] The advantages and shortcomings of the Riemann approach to complex analysis as opposed to the Weierstrass approach is just one among many concrete examples that illustrate and anchor the point.[24] The reasons that can be cited for the preferences are so far – reaching and systematic that to set the issue aside as philosophically insignificant would be to abandon altogether the hope of a philosophically satisfying account

of mathematical method. Given the thoroughness with which mathematical method is intertwined with physics, this would be tantamount to conceding defeat to skepticism of the global "Why should we think our best theories are true?" variety. Here and in the coming papers of this series I'll spell out why we need not make such a drastic sacrifice: if not a refutation of demon skepticism, we can at least get a better sense of what makes our best mathematical theories deserve the status of "best".

2.3. *Visualization as factor in mathematical reasoning*

> In my opinion, a well-balanced introduction to topology should stress its intuitive geometric aspect, while admitting the legitimate interest that analysts and algebraists have in the subject... I have followed the historical development where practicable, since it clearly shows the influence of geometric thought at all stages. This is *not* to claim that topology received its main impetus from geometric recreations like the seven bridges; rather it resulted from the *visualization* of problems from other parts of mathematics – complex analysis (Riemann), mechanics (Poincaré) and group theory (Dehn). (Stillwell (1993) vii emphasis in original)

One among the many reasons that can be (and sometimes is) cited for formulating a subject in one way rather than another is noted in the quoted remarks: the ability to visualize can be among the factors that shape a subdomain of mathematical practice. To avoid unnecessary controversy it is important to see that the importance of visualization in mathematical reasoning can be explored without making any assumptions about the nature of visualization itself. We don't have to answer the question of just what visualization is, or take a position in the now classic debate among cognitive scientists and philosophers of mind about how to make sense of talk of mental imagery.[25] We don't need to address what it could mean to literally "form a picture" of an infinite dimensional figure or a Klein bottle. All that will be needed here is the observation that we have enough of a naïve grip on the notion of visualization or pictorial representation to acknowledge that whatever such representation may ultimately amount to, in mathematical practice there are cases where people try to "visualize" (whatever we may be doing when we do that), that such visualization often is helpful, and that this often informs mathematical practice (more often in some fields than in others).[26] (This should not be philosophically contentious: it is just a factual observation about the methods and theoretical preferences of many working mathematicians and amateurs.)

As mentioned above, this aspect of mathematical activity has tended to be ignored by philosophers, in part because of a sense that the phenomenon is accidental, or "pragmatic" or "subjective" or "psychological" in ways that make it philosophically uninteresting. This response is reinforced by the simple fact that a large number of examples in which visual representation in diagrams assists reasoning and problem-solving, really are uninteresting to the student of methodology. It has been well-known to memory system designers from medieval times that it can be easier to remember information if one can visualize or draw some table of pictures to list the information.[27] In such cases, the fact that visualization is involved might not be particularly interesting for the present purposes. Cases like this are ubiquitous in mathematics too. As an example, consider these two ways of representing the multiplication table for octonions.[28] Here is the multiplication table (with e_1 = 1):

	e_2	e_3	e_4	e_5	e_6	e_7	e_8
e_2	-1	e_4	$-e_3$	e_6	$-e_5$	$-e_8$	e_7
e_3	$-e_4$	-1	e_2	e_7	e_8	$-e_5$	$-e_6$
e_4	e_3	$-e_2$	-1	e_8	$-e_7$	e_6	$-e_5$
e_5	$-e_6$	$-e_7$	$-e_8$	-1	e_2	e_3	e_4
e_6	e_5	$-e_8$	e_7	$-e_2$	-1	$-e_4$	e_3
e_7	e_8	e_5	$-e_6$	$-e_3$	e_4	-1	$-e_2$
e_8	$-e_7$	e_6	e_5	$-e_4$	$-e_3$	e_2	-1

This is a lot to absorb. To assist the memory there is a standard picture. The connecting lines, including the curved lines mark out the products.[29]
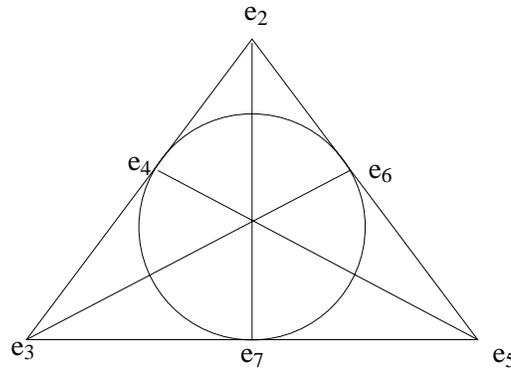
Unquestionably it is helpful to picture the multiplication operation of the octonions this way. However, the visual representation in a diagram might plausibly be said to function *solely* as a memory aid. That is, it might be that the sole advantage gained from the visual representation in this case is that most people just happen to find information easier to recall if it is encoded

into pictures of this kind than if it isn't. It isn't an implausible suggestion that it is an accidental feature of our cognitive makeup that memory can be thus facilitated by visualization. If so, there might not be any philosophically interesting consequences to the fact that people generally have a strong preference for the visual arrangement over the tabular one.[30]

I think that is the right thing to say in this case, and in many others like it.[31] The advantage gained by the diagram of the octonion table is one-off; it has no interesting systematic connections to anything else. Most significantly, the availability of the diagram doesn't correspond to anything deep about the basic concepts or proof methods used in the study of the algebraic structure represented. The point here will be that this case is a misleadingly simple exemplar. The function of visualization is much more intricate and systematic in cases like the Riemann approach to complex analysis. This, of course, is not so much an answer as a gesture in the direction of an answer: if the Riemann approach is different, we should be able to spell out how. This paper is a first step toward meeting that challenge.

One point that is worth noting is that the advantages of the diagrammatic representation of the octonions table seem to be *essentially* visual: the diagram helps because visual perception is an especially vivid mode of cognition, and the visual arrangement of the table allows it to be easily "taken in at a glance". The cases that interest us here are more complicated: though Riemann surfaces (for instance) admit of visual representation in a particularly straightforward way, the fruitfulness of this theoretical context for complex analysis persists even if diagrams are eschewed altogether. One might put the situation this way: there is a mode of organizing the subject which is especially natural, and which *happens* to connect to visualization in a direct way, but the mode of organization is theoretically valuable even without the essentially perceptual features like the vividness and immediacy of representation in diagrams. That is, the connection to vision is an intriguing and useful bonus, but the issues raised by the Riemann – Weierstrass opposition are of interest independently.

The distinction between the contexts of discovery and justification might also be suggested to apply here, since theoretical formulations count as "fruitful" at least partly because of how they facilitate the discovery of proofs.[32] It could be suggested – in line with the usual appeals to this distinction – that fruitfulness is only of interest in the "context of discovery" and that only what matters to the context of justification is relevant to epistemology, so once again these issues are someone else's concern. This issue is complicated and I can only indicate the shape of the answer here, deferring a more extensive discussion for elsewhere.[33] In brief, this is a case where the context

of discovery/context of justification distinction blurs. The preferred, fruitful formulation is *not* generally dispensable in the "context of justification" because there might not *be* any proofs available that do not use the fruitful formulation.[34] [35]

It will help us sharpen the issues to look for a philosophical niche served up by treatments of explanation and understanding in the natural sciences, since these have been extensively addressed. It will be especially helpful to consider the treatments of explanation and understanding as bound up with theoretical unification. I'll argue that these accounts are correct only with revisions and qualifications, but following out a presentation of the view and some of its obstacles will be a useful springboard. This will be the goal of the next two sections.

3. UNDERSTANDING, UNIFICATION AND EXPLANATION – FRIEDMAN

Efforts to explain and understand also guide investigation in the natural sciences, and mathematical theories and methods are integral parts of these investigations. This is a truism. Even if a sharp and principled divide between "applied mathematics" and "pure mathematics" can be marked out, it is unlikely that it would be so stark as to exclude unifying methodological themes. So it makes sense to try to compare explanation and understanding in science and mathematics.

Of course, there is a crucial *disanalogy*. The propositions to be explained in mathematics are not contingent, and so we can't appeal to the plausible and widely accepted suggestion that the explanation of an event provides part of the causal history of that event. If the plausible suggestion is right, then whatever is involved in mathematical explanation and understanding must differ from what is involved in at least *some* cases of scientific explanation and understanding.[36] This is not a problem here: I *grant* that the words "understanding" and "explanation" are probably not used unambiguously to pick out a single uniform phenomenon, and I agree that many things that are properly called explanations appeal essentially to causal ancestry. However, mindful of the fact that some explanations in physics and mathematics *do* seem to be governed by the same principles, I'll count it as an advantage of an account that it supports a uniform treatment of some mathematical and some physical explanations.

A promising candidate to support a uniform treatment of some pure mathematical cases and some non-mathematical ones is the treatment of explanation as unification as proposed in the seventies by Michael Friedman

and Philip Kitcher. In this section I'll pass through Friedman's now classic account in his (1974). Friedman embraces the suggestion that in at least some important cases, to provide an explanation of an event is to provide the resources to understand that event, and motivates his account accordingly. A further reason for considering Friedman's treatment here is that he states explicitly a requirement that we noted earlier: accounts of understanding must be *objective*. Scrutiny of the way that Friedman understands this condition will help clarify what is at issue.

This discussion aims to shore up these two points: a) The unifications that are regarded as valuable in practice involve some kind of *qualitative* advance, by producing a theory that is comparatively homogeneous to subsume two or more theories that had appeared heterogeneous. Mere reduction in the *number* of axioms or basic principles is typically not at issue. B) The conception of "objectivity" that motivates Friedman's account has to be relaxed in order to be true to the kinds of example that make unification accounts plausible. I'll suggest that Friedman's desiderata are actually better served by effecting something of a rapprochement with the view of Stephen Toulmin, which he examines and rejects.

3.1. Friedman: understanding as "objective"

The core of the Friedman approach is the compelling observation we've already noted, that often it is an important advance to bring about a unification of theories or hypotheses that had previously seemed disparate, by crafting a single theory or hypothesis that subsumes both. There are famous cases in which such unification appears to have been a clear advance: two prominent examples are Newton's unification of terrestrial and celestial mechanics and the unification first by Maxwell, and then within relativity theory, of electric, magnetic and optical phenomena. The prospect of a super theory encompassing quantum mechanics and gravitation is recognized as a motivating objective of current physical research. Looking back to the Riemann – Weierstrass contrast with which we began this study, we find different variations on this motif marking out Riemann's style. Riemann's approach to complex function theory admits of a variety of points of view in part because he effected a unification of complex function theory with the theory of complex curves and surfaces. Also, on a smaller scale, the "proper" choice of unifying definitions and basic principles was regarded by Riemann's contemporaries as one of the hallmarks of his work. A particularly clear expression appears in these words of Casorati:

> We believe that it is truly admirable with what assimilating
> power [Riemann] knew how to gather and establish in some

> compact, simple, and general theory, together with his own,
> all of the other studies that had an important relation to them.
> Of particular importance were Cauchy's many studies, spread
> over numerous publications, which were conducted with dif-
> fering purposes and often wrapped in a heterogeneous variety
> of terms and special notations. ... In paricular it is especially
> worth observing how Riemann always sets up his own conven-
> tions and definitions in such a way that every theorem can be
> stated as true without exception, or how many formulas and
> theorems ordinarily thought to be different from each other
> can be united in a single formula or theorem. (Casorati (1868,
> 140 note 2); quoted in (Bottazini, 1986, 218)).

Though it is open to dispute whether such advances are reasonably de-
scribed as "explanations", it is clear that they are successes, so I'll treat it as
uncontroversial that unification (whatever it may amount to in the final anal-
ysis) does represent an objective implicit in scientific practice.[37] What will
be at issue is the question of how to analyze these acknowledged successes.

Friedman argues that there is an irreducible advantage to reducing the
number of facts that have to be accepted as "brute facts".[38] Friedman faces
the elementary problem that we can reduce the number of axioms in any
theory generated by a finite collection of axioms $A_l \ldots A_n$ just by taking as
a single axiom the conjunction $A_l \& \ldots \& A_n$. But of course this wouldn't be a
theoretical advance. Friedman refines the idea of reducing basic premises by
introducing a characterization of premises as "independently acceptable". A
successful unification, on his account reduces the number of *independently
acceptable* premises.

Friedman's treatment doesn't seem to hang together unless a purely *a
priori* delineation of "dependent" and "independently acceptable" can be
worked out. Friedman's efforts to craft a technical definition of "indepen-
dently acceptable" foundered on straightforward counterexamples, the ini-
tial efforts at patching the theory foundered on further counterexamples, and
nothing much has been done to revive this part of the account in the last cou-
ple of decades.[39] In light of this I'll regard Friedman's theory in this specific
form as untenable. But it will be useful for diagnostic purposes to set aside
this problem: let us say for the sake of argument that we have managed to
fix which propositions are acceptable independently of which others. Even
so, Friedman faces the problem of distinguishing the *valuable* syntheses that
can plausibly be called "explanations" or "advances in understanding" from
those that cannot.

By unifying the theories of electricity and magnetism, Maxwell brought together two theories that had previously seemed to govern different *kinds* of phenomena. But there is no reason to expect that the new, more encompassing framework of electromagnetism had fewer axioms. Say that in some canonical axiomatisation, Maxwell's electromagnetism had twelve axioms and the two theories it replaced had just four each. So what? Would we cease to regard Maxwell's electromagnetism as a successful unification if we discovered, upon actually doing the count, that it had more axioms than the total axioms of the theories it replaced?[40] Similarly, if relativity theory and quantum mechanics were to be brought together into a single homogeneous theory, we would hardly reject the unified theory, or regard it as a regressive, non-explanatory step, if it had an independent axiomatic basis with more sentences than the theories it subsumed. In short, reducing the number of axioms need not increase understanding and increasing the number of axioms need not detract from understanding.[41] We need to focus not on the number of axioms but rather on what makes a given theory more homogeneous than just the conjunction of the two that it subsumes.[42]

Mathematical examples are worth adding, to be consistent with the theme that the relevant issues arise in both mathematics and natural science. In a classic paper (1882), Richard Dedekind and Heinrich Weber produced a unified theory of algebraic functions of one variable and algebraic numbers.[43] It would be hard to overstate the importance of this work; much of the subsequent development of German algebraic geometry in the first part of the twentieth century can be traced back to this source (in both content and style).[44] Much of the jolt of this work comes from the unification of a paradigmatically arithmetical subject (algebraic numbers) with another (one-variable algebraic functions) that (in the Riemannian tradition Dedekind occupied) had been seen as geometric. Key insights were (as we would now put it) that both algebraic numbers and algebraic functions can form a field, and that the elements of these fields can be analyzed in terms of ideals, and many of the crucial properties of both can be seen to depend just on this.[45] (In identifying the concepts of field and ideal as crucial, we can now recognize that Dedekind and Weber were backing winners. As the subsequent history has borne out, few concepts can rival these for fruitfulness within mathematics.) The uniform treatment of two apparently disparate subjects (achieved in this case by general argument patterns exploiting a "truly central" idea) gives a provisional credibility to the idea that the concepts involved, within the theoretical framework as a whole, really laid out what is going on. (The provisional credibility was, of course, ratified by subsequent research.) Once

again, this advantage is completely disconnected with any issue of how many axioms might be involved in the unifying theory.[46]

This is borne out by the account of Dedekind's algebraic work in the Weyl (1995) essay mentioned above. Weyl does say, in one quote given above that "a natural generalization simplifies by reducing the number of assumptions and by thus letting us understand certain aspects of a disarranged whole". But Weyl's subsequent discussion reveals that these remarks should not be understood in terms of a reduction of axioms. In the detailed analyses of Dedekind - style approaches to the theory of algebraic functions of one variable that form the main body of Weyl's essay, the idea of "reduction" in play is *conceptual*, in the sense that propositions incorporating a variety of apparently disparate concepts can be reduced to claims incorporating just one fruitful key idea: Dedekind's concept of ideal, that "runs through all of algebra and arithmetic like Ariadne's thread" (1995, p.649). So far as Weyl's analysis is concerned, a reduction of two propositions containing a range of heterogeneous concepts to more than two propositions framed solely in terms of ideals would be a unification contributing to understanding. Ideal theory is a natural framework for these problems, according to implicit and sometimes explicitly articulated criteria of "naturalness" that inform mathematical practice. We need to address the extra qualitative condition – that the theory doesn't just effect a bare, gerrymandered unification but that it does so in the right terms – before we can understand what gives the unification value and leads mathematicians to regard it as an advance in understanding.

An example from geometry will be especially helpful in connection with the discussion of section 5. A breakthrough in the understanding of general geometries, and the algebraic structures corresponding to geometries, was provided by Hilbert's axiomatization in Hilbert (1899) (refined in subsequent editions). This set out a single framework within which the range of geometries studied up to that point could be deductively developed and studied. However, this framework was remarkable in part for the *number* of axioms it contained. Far from reducing the number of axioms, Hilbert actually went out of his way to increase it: he aimed to isolate precisely what axioms a particular theorem depended on. To this end it was important to have available the axioms to support fine distinctions of deductive strength.[47] Consequently, even though the framework as a whole unified the study of diverse geometries, this was not unification according to Friedman's analysis; indeed, on Friedman's analysis it would be a regressive step.

The other mathematicians who set the tone at the turn of the century seemed also to be indifferent to number of axioms, though they were alert to differences between frameworks. Thus in his otherwise glowing review

of Hilbert's foundations, Poincaré was silent on the increased number of axioms, but he indicated one felt lack: unlike Lie's foundations, Hilbert did not set the idea of transformation in an important place. (Poincaré, 1903) For his own part, Hilbert complained of Lie's foundations that in Hilbert's view Lie's axioms were insufficiently "elementary". The points of comparison and contrast were solely qualitative; which of the frameworks had the fewest axioms didn't come into the debate at all.

I'll turn to the broader standpoint that informs Friedman's discussion. As noted above, Friedman suggests that his is the only proposal among those he considers to be both *objective* and to capture something that might fairly be called a notion of understanding.[48] Friedman appears to take "objective" to mean something in the ballpark of "ascertainable *a priori*", though it is not clear what positive account of objectivity he has in mind. However, it is clear what he thinks *isn't* objective in his sense: the view proposed in Stephen Toulmin's *Foresight and Understanding*, which Friedman calls (uncharitably, as Friedman acknowledges) "the intellectual fashion view." I think this rejection is a mistake. The spirit of Friedman's view seems to be salvageable only if something like Toulmin's position is grafted on, and it is unclear why we should take this to involve any sacrifice in objectivity.

We'll need a closer look. On Toulmin's telling, what counts as "understanding" is conditioned by a broader "ideal of natural order". The function of these ideals that is interesting for our purposes is that they tell us when something is in need of explanation and when it may stand unexplained. Similarly, it colors our view of what properties are "natural" and which "artificial". These ideals can change from one era to another. So for example, action-at-a-distance explanations stood for a time as absurd, then became acceptable for awhile, and then became unacceptable again. Friedman objects to incorporating this into an account of understanding because it represents understanding as "not objective" – as depending on capricious or subjective factors. But this is not obviously just: it depends on *what kind of reasons*, if any, are given to support the ideal. If we take something as basic, we need not be mute about *why* we take it as basic. Nor must the reasons we offer be bad reasons just because other people in different conditions might offer different reasons, supporting different conclusions. Quite the opposite: especially if the different conditions include different available data, as generally they do, we should *expect* different answers. We don't want to say that it is just fashion that leads us to change our minds when we *learn more*. Our judgements as to what is basic and what requires explanation, and so on, are not independent of what is known about the world, and what is appreciated about method, at a given time. It does not impeach this observation that there

will generally be no uniform algorithm for going from a theoretical situation and range of evidence to some fixed range of ideals and natural choices of basic axioms and properties. The best arguments may be highly case – specific. But they need be no less rational for all that: there is no way to tell in advance before we consider what the specific reasons are.

The argument of Friedman's paper has an instructive instability on this point. Friedman acknowledges explicitly that by Toulmin's account the choice of a particular "ideal of understanding" can typically be defended with reasons. This (appropriate) concession is in tension with both Friedman's core objection and the connotation of the label "intellectual fashion view" that ideals of intelligibility are arrived at arbitrarily or as caused by individual or group psychology rather than adopted as rationally justified. To be sure, Toulmin invites this uncharitable reading, as at times he appears to both assert and deny that reasons can be given for a preferred explanatory framework. For example, he writes:

> Those who build up their sciences around a principle of regularity or ideal of natural order come to accept it as self-explanatory. Just because (on their view) it specifies the way in which things behave of their own nature, if left to themselves, they cease to ask further questions about it. It becomes the starting point for explaining other things. Yet the correctness of a particular explanatory ideal (as we shall see) can never be self-evident, and has to be demonstrated as we go along. (Toulmin, 1961, 41-42)

The statement that an ideal "can never be self-evident, and has to be demonstrated as we go along" sits uncomfortably with the suggestion that this basis cannot be explained further:

There must always be some point in a scientist's explanations where he comes to a stop: beyond this point, if he is pressed to explain further the fundamental basis of his explanation, he can say only that he has reached rock-bottom. (Toulmin, 1961, 42)

I think it is the first of these that is truer to the facts: even when something is taken basic or fundamental, this does not preclude the possibility of arguments for and against *regarding it as* basic or fundamental, nor does it preclude arguments for and against its truth.[49] And in many of the interesting cases (such as Toulmin's example of magnetism in the nineteenth century) the changes in the categories taken as central and the principles taken as basic could be, and were, provided with a detailed rationale. Again, this is exemplified by the vicissitudes of German action-at-a-distance explanations, in their course from rejected to accepted to rejected again over the course of

a century.[50] This was a dispute over what to regard as the basic shape of an acceptable theory, but the proponents and opponents of action-at-a-distance were hardly reduced to inarticulate grunts. Elaborate reasons were given on both sides, involving appeal to reflections on general methodology and appeals to the advantages of the known available theories, as well as the range of known facts. It is hardly caprice or brute fashion if people come to prefer a different framework when they have gathered more data and learned more, or when they have formulated attractive theories that had previously eluded them.

Toulmin's (unnecessary and in my opinion regrettable) suggestion that the basic concepts and propositions are the bedrock at which the spade of argument is turned is (rightly) a sticking point for Friedman. It will be useful to quote and comment on a long stretch of Friedman's text to clarify what is at issue:

> There are many cases in the history of science where what seems explanatory to one scientist is a mere computational device for another; and there are cases where what is regarded as intelligible changes with tradition. However, it seems to me that it would be desirable, if at all possible, to isolate a common, objective sense of explanation which remains constant throughout the history of science; a sense of "scientific understanding" on which the theories of Newton, Maxwell, Einstein and Bohr all produce scientific understanding. It would be desirable to find a concept of explanation according to which what counts as an explanation does not depend on what phenomena one finds particularly natural or self-explanatory. In fact, although there may be good reasons for picking one "ideal of natural order" over another, I cannot see any reason but prejudice for regarding some phenomena as somehow more natural, intelligible, or self-explanatory than others. All phenomena... are equally in need of explanation, though it is impossible, of course, that they all be explained at once.
>
> Therefore, although the 'intellectual fashion' account may ultimately be the best that we can do, I don't see how it can give us what we are after: an objective and rational sense of 'understanding' according to which scientific explanations give us understanding of the world. (Friedman, 1974, 13)

There is much in these words I agree with. I am of one mind with Friedman in resisting the suggestion that "some phenomena should be seen as natural, intelligible or self-explanatory", if this is to mean that the status of

"natural" cannot be defended by good reasons. It was, I think, a misstep for Toulmin to cast the issue in these ways. But we do, in practice, see some categories and principles as "more natural, intelligible or self-explanatory" in the sense that they are methodologically basic: they are taken as the most reasonable principles to appeal to in addressing problems of a given kind, or the most natural categories to employ in connection with a given subject. That concepts and principles have this status is not inexplicable: part of the practice of mathematics and science – part of what makes it valuable to us – is that it incorporates reasons why the basic principles chosen are the good ones. There can be general changes of mind, because more facts are learned, or the theoretical situation comes to be better appreciated, or because new problems are confronted that prompt a reappraisal of accepted techniques, or just because people working in the field come up with some new ideas. (Such changes in what is taken as natural and basic can, of course, also be the result of irrational caprice, but they need not be.) If "natural" is understood as here, there need be no loss of "objectivity", and no appeal to caprice or irrationality, in "a concept of explanation according to which what counts as an explanation does [depend] on what phenomena one finds particularly natural or self-explanatory." Section 5 will deal with such a case, in which choices are made of axioms and basic categories for good reasons.

These observations do not conflict with Friedman's stated objective "to isolate a common, objective sense of explanation which remains constant throughout the history of science; a sense of "scientific understanding" on which the theories of Newton, Maxwell, Einstein and Bohr all produce scientific understanding." Indeed, what I've written above even strengthens the case for a (modified and weakened) version of the core theses that understanding is an objective of mathematical and natural scientific investigation and that in some interesting cases something reasonably described as "unification" is an important contributor to increased understanding. The quarrel is only with the details of the analysis of "unification": it isn't a reduction in the number of axioms, but something more complex. The bottom line is that we can't expect an account of the sort Friedman desires without concrete information about the reasons offered and accepted in scientific and mathematical practice for choosing what is to count as the "natural" or "reasonable" or "proper" primitive concepts and axiomatic formulations. If some of the features of theories that make them good examples of "unification" are more qualitative and (so to speak) "softer", that doesn't make them any less real, important or objective.

To nail down the point, let's reconsider the concepts of "field" and "ideal". Dedekind & Weber (1882) helped bring out what has subsequently become

patent: these concepts are mathematically central. By contrast, during roughly the same period, Frege developed the concept of "quantitative domain", which is similar to the concept of field in some respects, and which hasn't caught on at all.[51] The definitions alone don't make it obvious which of these would turn out to be the best choice of a framing concept: this has to be discovered by seeing what they can do. That field theory became central and the concept of quantitative domain sank below the waves was not because of academic politics or mob psychology or the whims of capricious fashion. Any working mathematician today who was introduced to Frege's definition could give a cogent rationale for the preference for the concept of "field" over that of "quantitative domain" as a framing concept.

The explanation will be intricate, though. It is not just that the concept of "field" is used to prove *more* theorems, since both concepts can be used for the same number of theorems (i.e. infinitely many). It is rather that field theory is needed for a striking number of "important" or "interesting" or "central" theorems. That a theorem deserves such an honorific designation can in turn be justified in terms of some combination of the connection to other mathematical subjects or physical applications, the importance of further theorems, qualitative observations about what is unexpected or surprising, and the like.

Other features of the concept can be invoked beyond just the results that it supports. For example, it can be pointed out that the concept of field is naturally refined into further fruitful subcases (zero vs. nonzero characteristic, finite and infinite fields). Once again, the preferences and distinctions that lead us to count one framework as more natural than another need not rest on brute, inarticulate preferences or transient fashion.

In sum, to make progress in clarifying what is at issue when theories are successfully unified, we need to learn more about qualitative features of theories: what makes a framework, and the categories in it, "natural" and "homogeneous" or whatever. Of course, there is only so far that a discussion of whether or not a set of reasons count as "objective" in the absence of the reasons themselves. This will of course depend on the specific case at issue. It will be the goal of section 5 to illustrate how involved these sorts of reasons can be in a case that is worked out in some detail. First I'll look more carefully at a specific touchstone: the problem of excluding gerrymandered predicates. This will provide the occasion to take up the potential for visualization, as a feature of theories that occasionally (though not always) contributes to an assessment of a category or framework as natural.

4. KITCHER: PATTERNS OF ARGUMENT

The account in Kitcher (1976, 1982, 1989) shares with Friedman the emphasis on unification as an animating objective in scientific inquiry. An important point of agreement between Kitcher's discussion and mine is in his appreciation of the affinities between some mathematical explanations and some natural scientific ones. He therefore sets this objective: an account that will accurately represent explanations in both domains. (1989 p. 482) Because of the above-noted technical problems in demarcating what is "independently acceptable", Kitcher sets aside the suggestion that reduction in the number of "independently acceptable" premises is the criterion for successful unification. Rather, the goal of unification is taken to be a reduction in the number of argument patterns. The background stance is that explanation involves seeing connections among phenomena, with the degree of unification depending on how economically this is achieved:

> Understanding the phenomena is not simply a matter of reducing the "fundamental incomprehensibilities" but of seeing connections, common patterns, in what initially appeared to be different situations... *Science advances our understanding of nature by showing us how to derive descriptions of many phenomena, using the same patterns of derivation again and again, and in demonstrating this, it teaches us how to reduce the number of types of facts we have to accept as ultimate (or brute).* (Kitcher (1989) p.432 emphasis his)

Kitcher's view begins with a set K, "the set of statements accepted by the scientific community", with a set of arguments deriving some members of K from other members of K a *systematization* of K. Argument patterns are represented as sequences of schematic sentences, with a restricted class (the "filling instructions") of acceptable substitutions into the schematic places.[52] (So for example, an argument pattern might have one place restricted to chemical substances, another to real numbers for arguments relating substances and atomic weights.) An explanation, for Kitcher, is an argument in the best systematization – which Kitcher designates E(K). Already – even before we consider the criteria of goodness of systematizations – an interesting stance on explanation is marked out. According to this view, whether or not an argument counts as an explanation is a global matter, depending on the overall structure of the theoretical framework. Explanations are arguments belonging to some class which has theoretical virtues as a class. This leaves open a range of different possibilities, depending on what criteria of "bestness" for systematizations are proposed.

For Kitcher, the best systematization of K is the one that "best satisfies" the two constraints of "minimizing the number of patterns of derivation employed and maximizing the number of conclusions generated." (p.432) There is, of course, an important point of agreement here; the goal of "maximizing the number of conclusions generated" is close to the goal of identifying fruitful formulations discussed above. Also as noted above, to make sense of comparing two sets of infinitely many conclusions, and to get closer to actual practice, we will have to introduce some refinements, such as placing special weight on "important" "interesting" or "deep" conclusions. But this point will not be at issue here. Unification is thus taken to have this goal: providing single argument schemes that apply to a variety of special cases, with a special premium placed on keeping down the number of schemata.

As in the case of axioms, the suggested emphasis on keeping the *numbers* down isn't a good fit with actual practice. Quite the opposite, it is reasonable and common to seek many different arguments for a single result, each argument exemplifying different principles and exploiting different techniques, and giving a different theoretical diagnosis. There is no shortage of examples; the search for novel proofs of already established results is a standard practice. In the more profound cases (the prime number theorem, say, or the Riemann – Roch theorem) entirely different subfields are induced by different proofs of one result.[53] This reflects the fact that successfully identifying unifying generalities is assessed not by counting the total number of patterns but rather by the quality of the patterns themselves: Are they the *right* ones (are they deep or fruitful or revealing or whatever?) Again, it is a challenge to clarify what these qualitative desiderata amount to, but we have to tackle them before we can count ourselves as having clarified the goal of unification in mathematics and science.

This section will address the following points. The first subsection will take up the suggestion that finding general patterns allows a reduction in the number of facts taken as brute; I'll argue to the contrary that the general patterns don't supersede the particular ones. In the second subsection, I'll argue that unification according to Kitcher's pattern does occur in important cases, but it is not an unconditional goal. Additional constraints – for example that the predicates employed in reasoning are not "gerrymandered" – come into play as well. To lay the groundwork for the final section I'll consider one case – graphic statics – which is especially favorable to Kitcher and in which visualizability is one of the contributing factors to the assessment of the naturalness of the formulation.

4.1. *Unification: General patterns and Brute Facts*

The identification of the general patterns of argument doesn't reduce the
number of "brute facts" because the general cases typically don't supersede
the special cases they generalize. The interaction of the special and general
cases is complicated. Kitcher emphasizes the economy of thought that is
gained by identifying these general patterns of reasoning, and he should, but
it is well to realize that generality of this sort is sometimes valuable for the
dual advantage that different special cases may have specific advantages, and
the ability to shift back and forth gives problem-solving advantages. Further-
more, when you have a single pattern of argument unifying two domains, the
pattern might be useful for different reasons in each: it might generalize in
different directions or admit different fruitful modifications in different in-
stances. That is, a unified general theory can be valuable in part because it
allows the systematic exploitation of residual differences.

A mathematical example – the duality of variety and ideal in algebraic
geometry – helps bring out this point.[54] The example rests on the "dictio-
nary" connecting ideals in simple algebra and varieties in elementary alge-
braic geometry. It is useful to note how the process of working out this
duality is described in an intermediate – level undergraduate textbook:

> In this chapter, we will explore the correspondence between
> ideals and varieties. . . . [The *Nullstellensatz*][55] will allow us to
> construct a "dictionary" between geometry and algebra, where-
> by any statement about varieties can be translated into a state-
> ment about ideals (and conversely). We will pursue this theme
> in ##3 and 4, where we will define a number of algebraic op-
> erations on ideals and study their geometric analogues. . . . In
> ##5 and 6 we will study [additional] more important algebraic
> and geometric concepts. . . notably the possibility of decom-
> posing a variety into a union of simpler varieties and the cor-
> responding algebraic notion of writing an ideal as an intersec-
> tion of simpler ideals. (Cox et al., 1992, 168)

Although the study of the "ideal" – "variety" duality (in contrast to the
duality in projective geometry that we'll consider in a few pages) is not di-
rectly framed in terms of linguistic schemata, it still stands as a striking sup-
porting example for Kitcher's picture of mathematical practice as pursuing
understanding and explanation by seeking out general argument patterns.
General arguments can be transformed into arguments in geometry or argu-
ments in algebra by systematic substitutions into general schemes. The ap-
parent qualitative difference between the subjects of algebra and (algebraic)
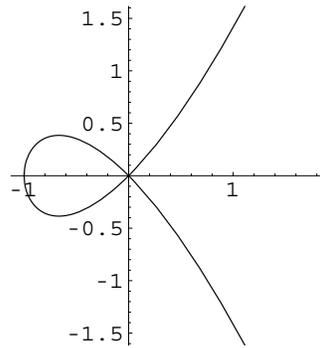
FIGURE 1. $y^2 - x^2 - x^3 = 0$.

geometry makes the discovered unity all the more compelling. However, the two subjects remain importantly different. It can be a subtle question whether a problem is more naturally addressed in one context or the other, and the ability to shift between the formulations is itself exploited as a prob-lem – solving strategy.[56]

Other, less involved examples are easy to come by. Theories of integra-tion in the plane are indifferent to what the underlying coordinates of the plane happen to be, but sometimes a careful choice of a specific set of co-ordinates can transform an integral from nasty to nice.[57] Another family of simple examples appears in the birational geometry of the plane (the study of properties of figures that are invariant under birational transformations).[58] Identifying curves that are birationally equivalent turns out to yield an in-teresting and useful theory, since certain key properties, such as the genus of a curve, are invariant under birational transformations. But the resulting generality does not mean that in studying these curves we should become indifferent to the specific details. Consider for example the resolution of singularities.[59] Say that in the plane we have the curve $(y^2 - x^2 - x^3) = 0$ (figure 1) which crosses the y-axis twice at a single point. (We count the origin *twice* because it is approached in two different ways by tangents.)

It is a bit of an irritation that the intersections coincide like this, and so it is helpful to exploit the fact that by a quadratic transformation – an especially
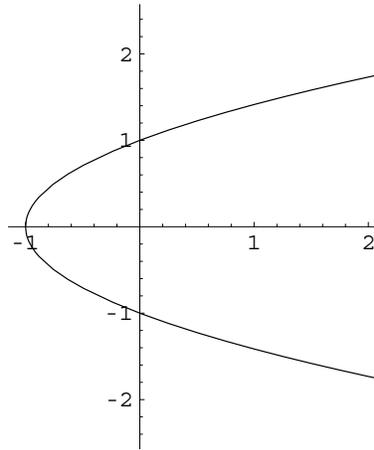
FIGURE 2. $y^2 - x - 1 = 0$.

simple birational transformation – the original curve can be mapped into the
parabola y^2 - x –1 = 0 crossing the y axis in two distinct places (figure 2).

So, as far as the properties of birational geometry are concerned, nothing
is lost if the more convenient representative stands as a proxy for the less
convenient one. This was a simple example; the gain in order and simplicity
is of course even greater when the zeros occur in clumps from cloverleaf
patterns and such. When we know we can "blow up" singularities in this
way, the fact that we have a general pattern of argument doesn't lead us
just to focus on the general pattern to the exclusion of specific details. The
general pattern also affords us a way of squeezing more information and
efficiency out of a good choice of special cases.[60]

4.2. Kitcher Unification in Practice: Projective Duality and the Gerrymandering Challenge

As we noted above, the goal of "minimizing the number of derivations" faces
a problem analogous to that faced by the candidate goal of "minimizing the
number of premises" in Friedman's treatment: the quantitative restriction, to
reflect actual practice, needs qualitative reinforcement. The derivations have
to be "the right kind", the unifying framework has to be "homogeneous"
and its basic categories "natural". The point also arises in connection with
Kitcher's account in two ways – one historical, one philosophical. One is an

analogue of the problem Friedman faced in defining "independently accept-able": some "unifications" (like logical conjunction) unify in an artificial way that doesn't advance understanding. Another concerns the history of some developments in mathematics and engineering that support Kitcher's account, but only part way. In the examples in question (general projective geometry, and its particular application to engineering techniques for analyz-ing the strength of components of physical structures) we find an a striking example of a mathematical theory being used for the explanation of physical events, in a way that reflects Kitcher's theory of explanation to a striking de-gree. The availability of dual patterns of argument is explicitly marked out as a theoretical virtue. But also, in this example, the advantages of patterns of argument that can be exploited in multiple ways is not treated as an un-conditioned objective: it is also constrained by the assessed "naturalness" of the basic categories, and the fruitfulness of the framework as a whole. I'll take up the second point after addressing the first.

The first problem is that unification will be too easy to achieve unless we can rule out "gerrymandered" properties as potential substitutions into argument schemes. If there is no constraint on what can count as a property then using a device well-known since Goodman's (1955) it is mere sport to come up with a theoretical unification of any two claims. Say we have two facts we want to explain/understand:

   a) A ball of uranium under conditions of extreme temperature never at-tains a radius of ten metres.
   b) Actors playing alongside chimps never win Oscars.

Here is an easy recipe to unify these theses. Let's define:

Pxy	iff	x is a ball of uranium in state y or x is an actor playing opposite y
Qx	iff	x is a state of extreme temperature conditions or x is a chimp
Rx	iff	x has a radius of 10 metres or x wins an Oscar.

Then, from the general proposition $(x)(y)(Pxy \ \& \ Qy \supset \sim Rx)$, and some ex-tra specifications (No actor is a ball of uranium...) we can derive both of our specific theses. It isn't difficult to set things up so that the derivations will be instances of a single argument scheme. This will give us a unified theory of critical mass and academy awards. We even get some unexpected, novel verifiable predictions (though not very interesting ones) like "No ball of uranium in a state of extreme temperature wins an Oscar". But I expect that it will be agreed on all sides that we haven't managed to explain or im-prove our understanding of either of the claims that we began with. In this

case the artificiality of the properties defined is evident, and their disjunctive character is written into the syntax of the definition. But of course we can't assume that spoiler predicates will always come with such a syntactic advertisement.[61] On what basis are we to distinguish the properties that support valuable unification from those that give us rubbish that isn't worth the effort? Kitcher is aware of this problem, and explicitly addresses a variation on it, but his response only makes our problem more urgent:

> We need some requirements on pattern individuation that will enable us to block the gerrymandering of patterns by disjoining, conjoining, tacking on vacuous premises, and so forth. The strategy sketched in the last paragraph attempts to disguise two patterns as one, and it does so by making distinctions that we take to be artificial and by ignoring similarities we take to be real. Thus the obvious way to meet the challenge is to demand that the predicates occurring in the schematic sentences [and playing other critical roles] all be projectable predicates of the language in which K is formulated. (Kitcher, 1989, p. 482)

Unfortunately, this answer loses one of the advantages of Kitcher's account that was most attractive to us here: the prospect of a unified treatment of mathematical and physical explanations. To the extent that we have any grip on the idea of projectibility at all, it has only been specified with reference to empirical predictions, and it remains to be seen how we should extend the idea to mathematical contexts.[62] So Kitcher's account of unification is incomplete: we need to supplement it with an account of how the range of acceptable substitutions is delineated in practice. This, of course, gets us back to our main theme, of how in practice we ascertain the methods of organization we will take as preferred and "natural".

It will give us a foothold if we turn to mathematical cases that support Kitcher's analysis in an interesting way, though only up to a point. It is true that uniformity of the kind he indicates has been sought, often quite self-consciously in the history of mathematics, science and even engineering. But there is always a bottom line: if the uniform patterns don't make things easier, if they don't support further discoveries, if they don't provide satisfying diagnoses, in short if they aren't fruitful, then they are set aside. A particularly illuminating example of this is embodied in the principle of duality in projective geometry and graphic statics in the nineteenth century and into the twentieth.

In the mid-nineteenth century, the development of projective geometry is strikingly close to the pattern Kitcher describes.[63] After extending the
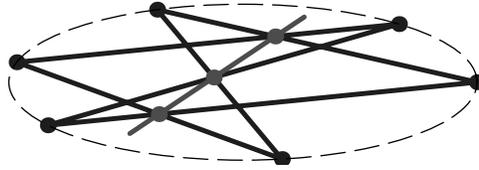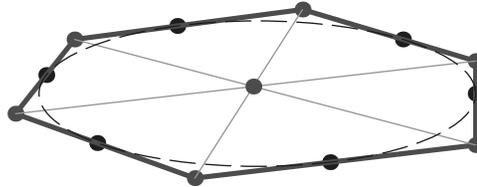
FIGURE 3. Pascal's Theorem.

FIGURE 4. Brianchon's Theorem.

Euclidean plane with points at infinity, reciprocal relations reveal themselves in the theorems of the extended system. It is possible to pair up vocabulary ("point" – "line"; "passes through" – "lies on" etc.; with induced pairings like "circumscribes"-"inscribes" etc.) so that given any theorem in projective geometry, the result of uniformly substituting each expression for its partner is also a theorem. This can yield quite striking pairs, as in the canonical examples of the Pascal and Brianchon theorems (see Figure 3 and 4):

Pascal's Theorem: Given a hexagon inscribed in a conic section, the points at which corresponding sides intersect all lie on a single line.

Brianchon's Theorem: Given a hexagon circumscribed about a conic section, the lines on which corresponding vertices rest all pass through a single point.[64]

This fact induces a quite general duplication of reasoning, as the substitutions also transform proofs into proofs, so that a single schematic argument delivers two proofs in one. This feature of general geometric reasoning became a fundamental aspect of the discipline in the late nineteenth century, to the extent that the standard convention in elementary textbooks and advanced research monographs alike was to write arguments in parallel

columns to display the dual arguments. This was an especially compelling example because the observed logical affinities among arguments were seen as more than just remarkable but idle epiphenomena. Duality came to be a cornerstone of general methodology for many pure and applied geometers in the mid-nineteenth century and subsequently. Producing theories that would issue in such dualistic patterns was seen by some geometers working in the area as a goal guiding the formulation of mathematical theories.[65]

Projective geometry and cognate fields, structured in this way in conformity to duality principles, are striking exemplifications of Kitcher's account.[66] But even in this highly favorable case, the issue is more complicated. The self-conscious focus on producing general schemata is not an unconditional goal. Once again it is important not only that the properties unify but that they are otherwise the "natural" or "right" ones.

A useful illustration here is the application of projective geometry to structural design, in the so-called "graphic statics" developed in the late nineteenth century by Maxwell and Culmann and developed further by Cremona. I will consider just the aspects of this rich history that are necessary to the issues we're addressing here. Fine details are available in secondary literature.[67]

Graphic statics was a theoretical formulation of techniques for analyzing engineering problems of structural reliability and strength of materials. Projective geometry is taken as the basic framework in the most ambitious and systematic formulation, presented in Culmann's *Die Graphische Statik*.[68] The problems involved a range of forces and pressures on hypothetical structures. One crucial early breakthrough from Maxwell (1864) was a technique for analyzing systems of forces in terms of *reciprocal diagrams*.[69] These worked by exploiting dualities to effect simplifications in the representations of forces acting on a structure. We won't need any further details here, except the key observation that this reciprocity allows complex stress diagrams to be reconfigured into diagrams that are easier to analyze, and which often display explicitly the desired information about stresses. Here is a relatively simple illustration:[70] (See figure 5.)

Say that the figure on the right represents a bridge in equilibrium with downward forces W_1, W_2 and W_3 and upward reactions A and B. The lengths of the lines represent the magnitude of the forces and the direction of the arrows mirror the direction of the forces. The stress diagram is the closed figure on the left. It will suffice for our purposes here to consider just one application to illustrate the technique. We can obtain the force on the particular support marked s by measuring the corresponding line s' on the stress diagram. One thing we can learn even without measuring is that the line t'
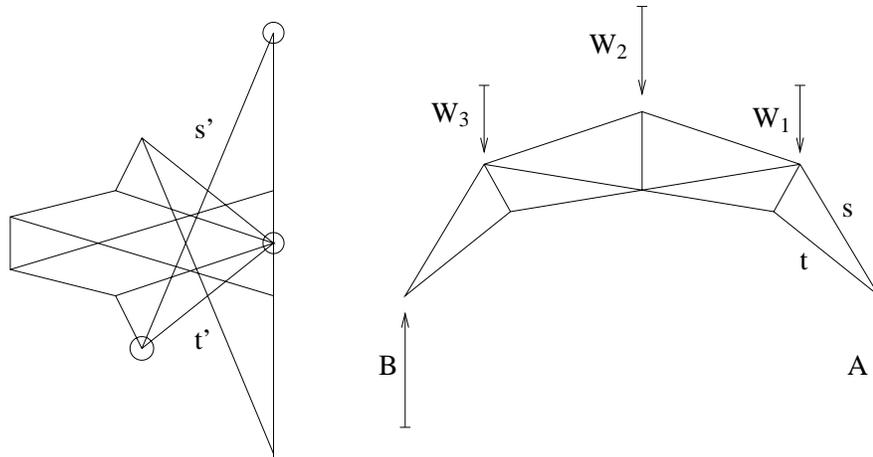
FIGURE 5.

corresponding to t in the original diagram is far shorter than s', reflecting the fact that the stress on s is much greater than that on t. (s' and t' are the lines between the circled points.) So if s and t are made of the same material and have no flaws, and the structure collapses because of s buckling, we can read off of the diagram a (defeasible) answer to the question "Why did s buckle rather than t?".

This example is satisfying from Kitcher's perspective not just because of the global role of duality in shaping the framework.[71] Note also that the preferred theoretical formulation doesn't distinguish between physical and mathematical situations. The account of (for instance) the stability of a configuration is the same whether we are concerned with an abstract vector sum or the stability of an actual bridge. It is of course an empirical question what frameworks are adequate representations of given physical situations, but solely mathematical/geometric criteria came into play in choosing which among the many equivalent frameworks is to be preferred as the representation of decomposition into component forces.

This yields a compelling example in which the theoretical virtues that led to the choice of a mathematical framework (and that consequently inform the ideas of understanding and explanation that the framework induces) influence the explanation of physical events as well. The fact that a rooftop can hold ten inches of dry snow, or that a cantilever bridge collapses, will be explained not just by appeal to familiar physical properties (the weight of the snow, the thickness of the roof, the weight of the girders, the temperature,...) but also by properties of the structure that are filtered through

the theoretical framework (the distribution of the load, the shear stresses at critical points...). What counts as an explanation here is shaped by global aspects of the theory, rather than merely by a picture of individual events with simple causal dependencies. So for example, the most renowned cantilever bridge failure of the time – the Quebec bridge collapse of 1907 – was set in motion when a single overstressed girder buckled.[72] The explanation of why that girder buckled of course required a grasp of not only the overall downward force and the strength of the materials but also an account of how the cantilever frame distributed the downward force through the structure as a whole, so as to indicate why precisely *that* girder was the one to go.[73]

For the last two decades of the nineteenth century and well into the twentieth, this was the dominant approach for studying the strength and stability of engineering structures.[74] Subsequently it was dislodged from its dominant position for a battery of reasons, among them that engineers came to confront problems of greater complexity than the graphic approach could easily address, and (more recently) because computers became more central to engineering practice.[75] However, our concern here is to address *why* graphic statics held sway over its analytical rivals during the time it *did* hold sway. Among the advantages that were noted, two are of special interest here: a) the theoretical formulation borrowed the fruitfulness of the general projective geometry that informed the treatments of Culmann and Cremona.[76] b) the visual representations in diagrams systematically conveyed the information in particularly vivid and effective ways; among the cited advantages were that the visual arguments make mistakes easier to catch,[77] that the visual presentation is easier to learn and teach without extensive mathematical training[78], and most importantly the graphic approach has that mysterious but crucial theoretical advantage: it just makes things easier.[79] Here too, the latter preference cannot be just shunted off into an incidental "context of discovery" since it shapes the terms in which justifications are given. The graphic framework was not set aside after it did the work of spurring creativity.

This brings us to a juncture similar to the one we reached above in the discussion of Friedman's account. It turns out to be true that dualities of the sort Kitcher isolates in his account of understanding have been taken to be contributors to the fruitfulness of mathematical formulations of problems. But even in the favorable case we are looking at, there is more going on. The value of the unifying account is not given *merely* by the fact that there are a variety of shared patterns (though in this case the sharing of argument patterns is important). The status afforded to graphic statics as the preferred way to address structural problems (during the period when it was so regarded)

didn't depend exclusively on how well the dualities of the background projective framework effected unification in Kitcher's sense. In addition – as the gerrymandering problem would lead us to expect – some additional constraints on the formulation have to be in place for the unified patterns of argument to be seen as worthwhile.

Any case like this will be complicated, with many factors involved, but we do have one foothold in the case of graphic statics since *one* of the principles taken to govern the theoretical formulation is stated unambiguously: it was taken to be a selling point that the representations of forces are *visualizable*. The interaction between visual representation and conceptual organization can be intricate. In particular, a survey of the textbooks of the time gives an interesting glimpse at a fact that will occupy us in the next section: some textbooks which were explicitly directed at laying out the valuable features of graphic statics didn't contain a single diagram.[80] The graphic framework remains valuable even if we do not directly exploit diagrams or vision *at all*. One reason for this is that rules for vector addition of forces is built into the principles for manipulating and interpreting the diagrams. At the time, the abstract versions of the ideas of vector space and vector sum were still imperfectly worked out and not well understood. Naturally the theoretical value of studying forces as vectors subject to rules of vector composition and decomposition goes well beyond the value that derives from the fact that they can be represented in diagrams. That is: some of the value of the visual presentation derives from features of the organization of information that are shared by the diagrammatic presentation and some analytic presentations. For these advantages the visual presentation – the fact that we can *see it* in the way we do – is incidental. This is true more generally of the projective framework that forms the background of the Culmann – Cremona treatment of graphic statics. Even in the abstract analytic presentation given by homogeneous coordinates, where diagrams or other visual representations need not be used, patterns of "geometric argument" are often judged to be especially acute.[81]

Let's review the state of play before moving on. We began with the challenge to clarify what was philosophically and methodologically at issue in the nineteenth century revolution in mathematical thought initiated by Riemann. We claimed a foothold in section 2 with the observation that one type of success is recognized in both mathematical and scientific reasoning, and counted as a contribution to understanding: Unifying apparently disparate phenomena within a single homogeneous framework. What was at issue was clarifying what the valuable unifications should be taken to be. The two candidate analyses we considered turned out to be at best incomplete, needing supplementation by an account of what reasons were given

for basic principles and axioms and how preferred ("non-gerrymandered")
categories were arrived at. This requires some appeal to an idea of "natural"
categories and principles, but contrary to what was argued by Friedman, we
do not have to set this aside as "non-objective" if the choice of candidate ax-
ioms and principles is based on good reasons. (This shifts the question of the
objectivity of the "natural" categories to the question of what the supporting
reasons are.) In classical projective geometry, especially in its application in
graphic statics, we found an example exemplifying, on a smaller scale, one
principle informing the hard case (Riemann's complex analysis) that we set
out to approach. In some cases, a contributor to an assessment of the "nat-
uralness" of a framework and its basic categories is that the arguments and
analyses of the framework can be visualized. Finally, we noted that some of
the advantages of the visualizable frameworks we considered persisted even
when they were formulated in non – diagrammatic terms, as systems of ab-
stract analytic geometry or vector addition. This helps narrow our search:
we need to get a better sense of how this sort of indirect connection to vision
can inform our choice of theoretical frameworks.

5. ARTIN AND AXIOM CHOICE: "VISUAL REASONING" WITHOUT VISION

Implicit in sections 3 and 4 was this answer to the problem of identifying
"gerrymandered" predicates: it may well be that there is no general *a priori*
principle that will divide categories into natural and artificial. But the ab-
sence of a general *a priori* answer doesn't indicate that everything is caprice:
in particular cases, good reasons can be given for the choice of one frame-
work as preferred. We also considered one basis that is cited in at least some
cases: a framework can be preferred if it has a desired kind of connection to
visual representation.

   A relatively tangible example of the choice of a framework is given by
the choice of axioms for a mathematical theory, which motivates the case
study of this section: the choice of axioms and basic concepts within Artin's
Geometric Algebra (Artin, 1957). Before engaging the details we need some
ground – clearing concerning the use of the word "axiom". Contrary to what
the expression may have meant in the past, in mathematical practice today
"axioms" are not "self – evident truths neither needing nor admitting proof."
Most of the axioms we'll see here are not self-evident, nor are they treated
as unprovable.[82] What makes them the right candidates for axioms is that a
good case can be made that they divide up the topic in the proper way.

   Artin's volume has a polemical aspect. He is striving to revive a "geo-
metric" style of presentation, as he notes in his preface. (This is not just

a casual remark: the preferences it indicates are followed out consistently throughout the volume.)[83]

> Many parts of classical geometry have developed into great independent theories. Linear Algebra, topology, differential and algebraic geometry are the indispensable tools of the mathematician of our time. It is frequently desirable to devise a course of geometric nature which is distinct from these great lines of thought... (Artin, 1957, p. vi).

The specific orientation this stance involves is indicated later when Artin indicates how an algebraic result should be restructured. Artin is discussing the isomorphism connecting the ring of homomorphisms of an n-dimensional vector space (over a field K) into itself and the ring of nxn matrices (with entries from K). This isomorphism introduces two different modes of presentation, a fact upon which he comments as follows:

> Mathematical education is still suffering from the enthusiasms which the discovery of this isomorphism has aroused. The result has been that geometry was eliminated and replaced by computations. Instead of intuitive maps of a space preserving addition and multiplication by scalars (these maps have an immediate geometric meaning), matrices have been introduced. From the innumerable absurdities – from a pedagogical point of view-let me point out one example and contrast it with the direct description.

> Matrix method: A product of a matrix A and a vector X (which is then an n-tuple of numbers) is defined; it is also a vector. Now the poor student has to swallow the following definition:

> A vector X is called an eigen vector if a number λ exists such that $AX = \lambda X$.

> Going through the formalism, the characteristic equation, one then ends up with theorems like: If a matrix A has n distinct eigen values, then a matrix D can be found such that DAD^{-1} is a diagonal matrix.

> The student will of course learn all this since he will fail the course if he does not.

> Instead one should argue like this: Given a linear transformation f of the space V into itself. Does there exist a line which is kept fixed by f? In order to include the eigen value 0 one should then modify the question by asking whether a

line is mapped *into* itself. This means of course for a vector spanning a line that

f(X) = nX.

    Having thus motivated the problem, the matrix A describing f will appear only for a moment for the actual computation of n. It should disappear again. Then one proves all the customary theorems without speaking of matrices and asks the question: Suppose we can find a basis of V which consists of eigen vectors; what does this imply for the geometric description of f? Well, the space is stretched in the various directions of the basis by factors which are the eigen values. Only then does one ask what this means for a description of F by means of a matrix in terms of this basis. We have obviously the diagonal form....

    It is my experience that proofs involving matrices can be shortened by 50% if one throws the matrices out. (Artin, 1957, 13-14).

There is much to comment on here. First, a basic observation: The structures of matrices and of homomorphisms *are* isomorphic but the differences between the structures are not, in this case, dismissed. Consider the opening question Artin floats (sticking to just two dimensions, for simplicity). Given a linear transformation of the plane with two independent eigenvectors is there a way to change the basis of the plane so that the transformation relative to that basis is representable as a diagonal matrix? There are two different ways to arrive at an answer. Artin's preferred approach sets up a visualizable situation and – only when needed – appeals to an algebraic representation of it. The second deals throughout with the computations that can be performed in the algebraic representation.

My informal canvassing has turned up the expected result that almost everyone is of one mind with Artin that the first of these approaches is preferable.[84] As noted earlier, this is echoed in print. In one example Hughes and Piper (1973) speak of Artin's framework as "the proper setting for many problems in linear algebra." (p.285) There are many reasons for this. First of all, as a pedagogical observation most people find his preferred approach much easier to grasp on first exposure, as Artin observes. There are also gains in the most elementary kinds of economy like proof length, if Artin's observation that proof length can often be shortened "by 50%" is correct, as I will take it to be. These are important advantages of Artin's perspective that need to be taken seriously. But there are deeper, more systematic advantages as well, which will come out as the picture unfolds in more detail.

However, the point cannot be that pictures are an essential part of "geometric" presentation, since Artin has hardly any![85] [86] There are only 6 diagrams in a book of over 200 pages (all but one of these are meant to clarify the axioms to be considered in a moment.)[87] Rather, we find that a "geometric language" (p.26) which facilitates certain "intuitive pictures" (p.26) and visual handwaving is developed and fleshed out with axioms, but the power of the framework lies principally in its systematic theoretical fecundity.

The core concept of the approach is that of a *transformation* (or *symmetry*): the idea of moving a point from one position to another, thereby tracing a line. (This perspective is actually much closer to some "philosophical" analyses of space and intuition than it might appear to those unfamiliar with the tradition, but since this point will take us too far afield, I'll leave it to be developed in other work.[88]) Two axioms are set down to ensure the basic structure of parallelism: I) given two distinct points there is a unique line connecting them II) Given a point P and a line l, there is a unique line parallel to l passing through P. Also there is an axiom that states that there are three distinct non-collinear points. Given this, as Artin puts it "We can hope for a 'good' geometry only if the geometry has enough symmetries." (Artin, 1957, 58). Hence the remaining axioms posit the existence of symmetries, where these are *dilatations*: mappings σ such that given a line l, the image $\sigma(l)$ is parallel to l.[89] A sub-class of the dilatations is distinguished: a *translation* leaves no point fixed. (That is: τ is a translation if there is no point P such that $\tau(P) = P$). The only dilatation leaving more than one point fixed is the identity, which leaves everything where it is, so every non-degenerate dilatation that isn't a translation or the identity leaves exactly one point fixed. (Among the reasons to regard these as reasonable choices as basic ideas are algebraic: the dilatations form a group with the translations as a normal subgroup.)

Thus we have the general framework: there are points and lines, and symmetries mapping lines into parallel lines. The axioms will take the form of statements as to the existence and properties of symmetries. Given what we are looking for, these are the choices that suggest themselves right away:

Desargues Axiom 1: Given any two points P and Q there is a translation P such that $\tau(P) = Q$.

Desargues Axiom 2: For any points P, Q and R, there is a dilatation that holds R fixed and such that $\sigma(P) = Q$.[90]

Just by inspection we can say that these axioms certainly *seem* natural, in this framework. A deeper point, which comes out when we consider the reasons for regarding the axiom as a good axiom candidate, is that these axioms *ought to* seem natural: it is a strength of the framework that these axioms come out as natural – seeming as they do. First, though, it should be noted that the framework itself – representing the subject in terms of transformations of objects – has a compelling rationale of its own. It would require a separate paper – a long one – to even begin to develop the manifold ways that it has proven to be valuable to formulate a subject in terms of transformations and invariants. From physical theories of space and time to classifications of general geometries in the Klein program, to Galois theory and the theory of Lie groups, and untold other areas, pure and applied, this framework has shown itself to be a good one to choose, and the Artin framework of geometric algebra inherits these bona fides.

This point is worth emphasizing in connection with efforts to cut through the gerrymandering challenge by emphasizing elementary syntactic features of predicates – that they are "disjunctive" for example – as reasons to exclude them. This example illustrates a fairly general moral: whether or not something admits of a simple expression is going to depend upon global features of the framework it is studied in. In this particular case, the broader framework of studying geometries in terms of symmetries makes the Desargues axioms simple and immediate; the fact that we should treat especially seriously things that look simple in this particular framework is not justified by any *a priori* argument employing purely philosophical or linguistic criteria or appeals to principles of basic metaphysics. The justification of the framework has to appeal to the details of the subject – matter, including our amassed experience with frameworks of this type.

I'll return to the axioms themselves. For orientation it will be helpful to consider their classical forms:

(Classical) Desargues axiom 1:

If l_1, l_2 and l_3 are parallel lines in the (affine) plane and P_1, P_1', P_2, P_2', and P_3, P_3' be points on l_1, l_2 and l_3 respectively. Say that the line P_1P_2 is parallel to P_1'P_2' and P_1P_3 is parallel to P_1'P_3'. Then P_2P_3 is parallel to P_2'P_3'.

Diagram:
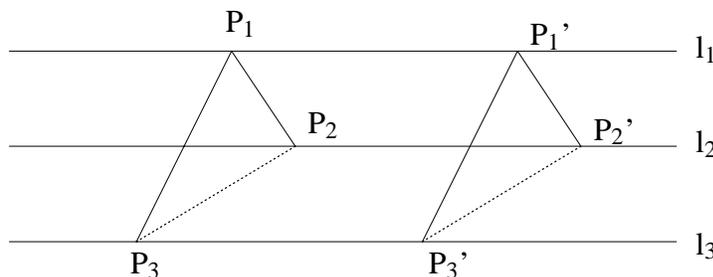
(Classical) Desargues axiom 2:

If l_1, l_2 and l_3 are lines in the (affine) plane intersecting in a point P* and P_1, P_1', P_2, P_2', and P_3, P_3' be points on l_1, l_2 and l_3 respectively. Say that the line P_1P_2 is parallel to P_1'P_2' and P_1P_3 is parallel to P_1'P_3'. Then P_2P_3 is parallel to P_2'P_3'.
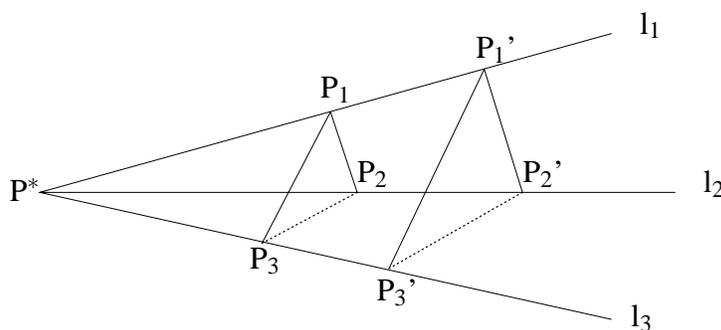
There are canonical reasons why these are good axioms.[91] A simple point is that the first Desargues axiom is equivalent to the uniqueness of a vector sum.[92] A more intricate consideration derives from the structure of familiar school analytic geometry. It is possible to assign coordinates to any collection of objects and introduce functions on those coordinates, so long as we are not too picky about what properties the coordinates and functions themselves have. If we assign coordinates in a general way to the objects of our geometry, the first Desargues property is equivalent to the statement that we can introduce operations of plus and times on the coordinates so that the equations of lines will be the familiar linear equations from school: $ax + b = y$.[93] A further, distinct consideration arises from the relations between planes and space: (relative to a reasonable axiomatisation of the geometry of space) the second Desargues axiom is equivalent to the thesis that the plane can be embedded in three dimensional space. The second Desargues axiom corresponds to further algebraic conditions on the addition and multiplication (the distributive law for the multiplication and addition operations on

the coordinates) to form a skew field. (i.e. a field minus commutativity of multiplication)[94] These reasons for regarding the Desargues axioms as distinguished draw on the fact that they are "robustly" central, in that when we reformulate our theory in other terms, with quite different structures and motivations, the Desargues axioms in their new forms remain rationally defensible as natural axiom candidates.

A further consideration that tells in favor of the Desargues theorems as axiom choices is the interest and richness of the divides they mark. Geometries in which the Desargues theorem *fails* have proven in practice to be a class of uncommon interest, sustaining extensive, interesting programs of research. On the positive side, the theorem itself is of considerable intrinsic interest, both for the consequences it supports and for the depth and intricacy the theorem reveals under more detailed study.[95] There is more that can be said to support the claim that the Desargues theorem really does deserve to be granted a distinguished status as axiom, but what has been said so far will suffice to illustrate the key point: Artin's choice of transformations as a basic category and of Desargues' Axioms as basic principles can be rationally defended by appeal to a range of different considerations. Even in this context, where empirical predictions are not directly in the offing, the distinction between natural and artificial/gerrymandered properties can be objectively made out.

In connection with the issues we have been concerned with, here is where we have arrived:

  i) The theory developed by Artin does have a deep and important connection to visual reasoning but
 ii) as a means of organization of the subject – matter it has value independent of the connection to vision and furthermore
iii) The basic details of the framework – its fundamental concepts and its axioms – admit of extensive justifications. That something is "basic" or an "axiom" is not bedrock at which the spade of explanation and argument is turned. Some of the reasons for shaping the framework as it is shaped may seem to be immediate brute responses or appeals to the brevity and simplicity of the expressions used (for example: the Desargues axioms in their symmetry forms just "look natural" in this context, and the statements don't invoke "disjunctive" properties or other funny looking constructions) while others reach far afield even to applications (for example: the "rightness" of the framework taking transformations as basic extends to physics).

Of course, these judgements are defeasible. We might decide, when we come to learn more about the subject, that this framework is not the right

setting for the problems addressed within it. But of course, the fact that in a different epistemic situation we would call different frameworks and categories "natural" for different reasons doesn't impeach the reasons we actually have, in the epistemic situation we are actually in, for our actual decisions as to what we regard as natural and what we don't. Much of the reasoning that goes into this decision is "quasi – empirical": among the information that the decision about what is a natural formulation or a good axiom choice draws on is information about what is fruitful, about what works and what doesn't. This makes it especially unlikely that a purely philosophical criterion of "gerrymanderedness" will suffice to exclude artificial, Goodman style properties. Our decisions about how to formulate the mathematical theories that we apply don't rest on *a priori* philosophical bedrock, and it appears unlikely that any *a priori* "rational reconstruction" could reproduce our best mathematical practice on abstract philosophical or logical grounds alone.

6. SUMMARY - THE "NEW RIDDLE OF DEDUCTION"

The paper began with two related questions. What philosophical niche can we find for a discussion of what was at stake in Riemann's revolution in mathematical method? What significance for general methodology should we grant to the role of visual representation as a mode of organization coloring some mathematical reasoning? We've arrived at a kind of mathematical analogue to Goodman's problem, but without the direct connections to causation and empirical prediction that are often taken to ground answers in the more familiar gruesome cases. To find a place for unification as a scientific and mathematical success, as it is treated in practice, we need to clarify certain qualitative features of theories and the properties they deal with. Which classes and theories are homogeneous and which are heterogeneous? Which classifications and properties are natural and which artificial? We need to be clear about what sorts of considerations are brought to bear, in deciding what formulations are the right ones to use. The conclusion suggested here, especially as exemplified in the case of *Geometric Algebra,* is that these distinctions are, in practice, made out in a way that is rationally justifiable, but also that they appeal to details of mathematical and scientific practice that are more involved and case-specific than philosophical accounts of explanation as unification have appreciated. This suggests that we reorient our conception of the methodology of mathematics in a "bottom up" direction: we can't hope to understand what mathematics contributes to our overall view of the world by shuffling philosophical abstractions alone; we need to get our hands dirty with the details of mathematics as it is done. To invert a

famous Kripkean slogan, in this case there is also no philosophical substitute for mathematics.

NOTES

[0]Special thanks for comments and criticism to Ian Proops, Larry Sklar, Colin McLarty, Jim Joyce, Rich Thomason and Peter Railton. An early version of some of this material was presented at a conference in the University of Toronto; Thanks to the audience, especially Jan Zwicky, Ian Hacking, Margaret Morrison, Jim Brown, Francis Sparshott and Achille Varzi for helpful reactions and/or encouragement. Thanks too to the audience at the Roskilde conference, especially Paolo Mancosu, Marcus Giaquinto, Reviel Netz, Jim Brown again, and Karine Chemla. Also special thanks for a range of assistance to Klaus Jørgensen. Thanks also to my extra-departmental colleagues Andrzej Nowak and Karen Smith for patient answers to questions on civil engineering and algebraic geometry respectively. Finally, a very early version of some of this material was presented at Princeton and a meeting of the Association of Symbolic Logic many years ago; thanks to those audiences, especially Gil Harman, Gideon Rosen, David Hilbert, Paul Benacerraf, Neil Delaney, Ed Nelson, Phil Ehrlich and Pen Maddy.

[1]This negative point – that there is no principled general delineation of the "essentially general" predicates – is argued in Railton (1993). This paper may be seen as a follow – up: if we accept that there is no general *a priori* account that will segregate the "essentially general" from the "gerrymandered" predicates, our attention naturally turns to working out the details in specific examples, with an eye to identifying defeasible heuristics and shared patterns that may be displayed in a range of cases. We can learn a great deal so long as we don't obsessively cling to an unrealistic picture of how simple a "philosophical" account of uniformity is allowed to be.

[2]I'm not, of course, suggesting that this idea of "fruitfulness" is clear or sharply defined, or even that it is a single uniform phenomenon, but only that it is a consideration that is in fact is appealed to in practice (under a variety of names). Explaining more clearly what "fruitfulness" amounts to is of course one of the jobs that has to be done.)

[3]The label "conceptual" is adopted for Riemann's innovative style in Laugwitz (1999) among others. The broader change in mathematical style that emerged in Göttingen in the mid-nineteenth century is explored with special reference to Dirichlet in the superb philosophical essay Stein (1988).

[4]The core details of the split in the approach to complex analysis have been well-explored by historians in recent years. An illuminating presentation of the Riemann stance is in Laugwitz (1999). On Weierstrass, Laugwitz (1992) is a helpful counterpoint. Good presentations of both sides of the split are Bottazini (1994) and Neuenschwander (1980) and (1981). Currently Jeremy Gray and Umberto Bottazini are carrying out joint work which promises to shed further light on the situation. I explore some of the philosophical ramifications of this split, and its historical

connections to the development of the foundations of mathematics, in a manuscript in progress.

[5]Contemporary texts in complex analysis tend to be unreflectively Riemannian in outlook. One historically sensitive text that is self-consciously Riemannian is Remmert's textbook of function theory. (Remmert (1991) and (1998)). Textbooks that are avowedly Weierstrassian in general outlook are harder to find, but they do exist: Abhyankar (1964) is one example. The fact that this division of styles has been robust enough to persist this long reinforces the point that more than merely transient "spurs to discovery" are at issue.

[6]By contrast, the historical details of the events constituting this contrast have been reasonably well – explored in recent years, and current research promises to push our historical understanding even deeper. In my own work I am exploring some of the philosophical overtones of the mathematical developments (with special reference to Frege) (see Tappenden (2001c)). The Frege connection appears because Frege was trained in the Riemann tradition (then a minority stance) and continued to work in that vein in his subsequent teaching and research. This can be seen to have colored his methodology in several respects, such as his stance on the "Caesar problem", his definition of magnitude, and his regular criticism of Weierstrass.

[7]I am sure that only the loosest family resemblance unites all the things that we call "understanding". I am certainly not setting out to provide anything like an *analysis* of "understanding" in the sense of a set of necessary and sufficient conditions such that all and only persons who satisfy those conditions understand something. However, even in the absence of an analysis of the notion of understanding, it is possible to isolate aspects of what we commonly associate with the idea, and work out their significance for epistemology and logic.

[8]I do *not* mean to suggest that the visualizability of Riemann surfaces is the sole advantage, or the most important one. It just happens to be the one feature of Riemann's approach that I am addressing here. In fact, it is one aspect of the revolutionary character of Riemann's research that there have been, historically, so many different ways of cashing out what is important in it. Just to consider the point at issue here: some Riemann students – notably Dedekind – strove to purge Riemann's results of their visual character, while others (Felix Klein, the Italian tradition of algebraic geometry, much contemporary theory of functions of one complex variable) embraced the visual character and strove to exploit it. (For an especially forceful discussion of the importance of visual intuition to this mathematical tradition see Segre (1904) (especially p. 454–455)).

[9]So as not to leave the impression that these issues pertain solely to the nineteenth century, it is worth mentioning another example exemplifying the themes discussed here: the concept of scheme in algebraic geometry. Though scheme theory is an extremely compelling example in the current connection, and I will consequently refer to it from time to time, it is also complicated enough that I will have to defer a sustained treatment for some later part of the project, when Colin McLarty's work on Grothendieck is ready to circulate. (Why do today what someone else is going to do tomorrow!) However, to reinforce the connection of the issues discussed here

to contemporary mathematics as well as that of the nineteenth century, I'll carry on a running commentary of scattered remarks about schemes in the footnotes.

Among the many reasons why scheme theory is especially interesting here is that one of its acknowledged virtues is that it supports a unification of number theory and algebraic geometry. It is an interesting question of methodology whether this theoretical unification is analogous to the benchmark unifications in physics, like Maxwell's or Newton's or in any important way different. In this connection it is worth noting that Grothendieck himself reportedly had equally grand hopes for the theory of "motives", envisioning a potential unification of Galois theory and topology. (cf. Cartier (2001, 405)).

[10]For smooth exposition I'm (inessentially) fudging some distinctions between Weyl's essay and my project, but it is worth a footnote to avoid leaving a misimpression. In my studies of the mid-century Göttingen revolution in mathematical methodology I have emphasized Riemann and his successors. Weyl (like Stein (1988)) emphasizes Dirichlet, Riemann's predecessor as professor at Göttingen; the reference to the Dirichlet principle could indeed be taken as a bit of a jab at Riemann's lack of contemporary rigor. For the issues I am most concerned with, Riemann is a better representative, and he is more important as a figure in Frege's intellectual environment. (Or at least he can be more easily documented to be a significant figure in Frege's Jena context.) But both figures represent, in different ways, the "conceptual" reorientation: Dirichlet was far more rigorous at the level of detail in argument, while Riemann's contributions to the stylistic innovations were more profound (full of what Ahlfors calls Riemann's "cryptic messages to the future.") though less rigorous. The work of both exemplified, in different ways, the style that in retrospect was a critical revolution laying the support for the twentieth century.

That it isn't distorting to take Weyl's words about Dirichlet's mathematics as remarks about Riemann's is borne out by the subsequent discussion in Weyl's essay: most of the mathematics he uses as his illustrations of contemporary work traces back to Riemann rather than Dirichlet.

[11]I speak about "algebraic approaches" in the plural here because Weyl runs together here what I count as two very different traditions and styles with the label "algebraic": a computational tradition of Kronecker and a distinctive "structural" algebraic approach exemplified by Dedekind. There is no reason, in the present essay, to refine Weyl's classification further, but I don't want to leave a misimpression.

[12]A useful illustration of this point is the article Harris (1992). There the development of algebraic geometry in the twentieth century is framed by the observation that "progress in algebraic geometry is measured more by its definitions than its theorems." (Harris, 1992, 99)

[13]I'm not suggesting fruitfulness is the only relevant consideration that bears on whether a formulation will be taken as "natural". One point of section 5 is to illustrate just how intricate the reasons for such judgements of "naturalness" can be.

[14]These remarks from Zariski give a typical statement of the mathematician's pragmatism in this regard:

> There is no doubt that the introduction of the concept of "schemes"
> due to Grothendieck was a sound and inevitable generalization of the
> older concept of "variety" and that this generalization has introduced
> a new dimension into the conceptual content of algebraic geometry.
> What is more important is that this generalization has met with what
> seems to me to be the true test of any generalization, that is, its
> effectiveness in solving, or throwing new light on, old problems by
> generalizing the terms of the problem (for example: the Riemann-
> Roch theorem for varieties of any dimension)....(Zariski, 1978, p.
> xvii)

[15]I do not know of any systematic studies of fruitfulness as a guiding criterion in mathematics or elsewhere, but the basic observation has some antecedents. Frege makes some fragmentary but rich remarks in *Grundlagen* which tie his logic and his account of "extending knowledge" to what he calls "the truly fruitful concepts". (I develop this observation about Frege in my (1995).) Thomas Kuhn observes that a crucial guide in practice to theory choice in the natural sciences is that the theory be "fruitful of new research findings: it should, that is, disclose new phenomena or previously unnoted relationships among those already known." (1977, 322) I don't know anywhere that Kuhn, or anyone else, expands on this bare observation. In the paper cited, Kuhn does not expand on his observation beyond the footnote remark: "The last criterion [listed], fruitfulness, deserves more emphasis than it has yet received." (How true.)

[16](Weyl, 1955, VII) My attention was originally drawn to this passage by (Wilson, 1992, p. 111). It should be noted in connection with these words that in context Weyl is not endorsing them unequivocally. Rather he is describing an attitude he expressed when, as a young man he wrote Weyl (1913), which the older Weyl spoke of as revealing a certain youthful naiveté. "Even more than the text, the enthusiastic preface betrayed the youth of the author." (p.VII)

[17]Once again, the history detailed in Harris (1992) is a useful illustration. Throughout the twentieth century there was a sequence of better and better candidates for the natural context for algebraic geometry. The reasons for one candidate succeeding another was never merely that the preferred candidate "seemed right" but that it in fact facilitated the solution to key problems.

[18]One simple example is the use of homogeneous coordinates/projective space ("the unifier" in the words of Clemens (1980, 5) in the study of curves, especially over the complex numbers ("the great unifier" – (Clemens, 1980, 7)). This is accepted as the right context for a range of problems, and it does indeed bring out forcefully many properties of (for example) conic sections. But it requires work to see this; it is not obvious at first sight. People who have had a standard North American mathematical education find ordinary Cartesian coordinates over the real plane so natural as to be almost inescapable; to come to see complex projective space as the natural context requires re-education.

[19]See for example the venerable Bridgeman ((1938, 37 and *passim*) and (Campbell, 1919/1957, 113 and *passim*)). I am grateful to John Norton for pressing me on the "explanation/understanding as reduction to the familiar" line. I do think that in an important range of cases we count something as an explanation, or as contributing to understanding, when it effects a reduction to the familiar, but the cases I am looking at aren't like that.

[20]It will be useful to give another example of broad motivating remarks concerning the concept of scheme. Though nothing can substitute for an analysis of the use of the concept in practice, some examples of what those mathematicians who use the concept say about its importance may serve as a temporary buffer until such an analysis is worked out. Here is how David Mumford puts it in a retrospective introduction (in 1988) to the publication of his by then already classic 1960's *samizdat* introduction to schemes in algebraic geometry. This contains a (long) book-length defense of the thesis that scheme theory provides the "natural language" of algebraic geometry. The reasons on which the defense rests include the ability of the framework to cleanly express results in a variety of different other conceptual frameworks, its connection to "geometric intuition", and its ability to support new and very exciting results":

> It may be of some interest to recall how hard it was for algebraic geometers, even knowing the phenomena of the field very well, to find a satisfactory language in which to communicate to each other. At the time these notes were written, the field was just emerging from a twenty-year period in which every researcher used his own definitions and terminology, in which the "foundations" of the subject had been described in at least half a dozen different "mathematical languages". Classical style researchers wrote in the informal geometric style of the Italian school, Weil had introduced the concept of specialization and made this the cornerstone of his language, and Zariski developed a hybrid of algebra and geometry... But here was a general realization that not all the key phenomena could be clearly expressed and a frustration at sacrificing the suggestive geometric terminology of the previous generation.
>
> Then Grothendieck came along ... [with] the new terminology of schemes as well as with a huge production of new and very exciting results. These notes attempted to show something that was still very controversial at that time: that schemes really were the most natural language for algebraic geometry and that you did not need to sacrifice geometric intuition when you spoke "scheme". (Mumford (1988) p. V – VI)

The attitude that "scheme" was a thematically proper generalization and that the test of this was effectiveness in problem solving was echoed even by members of the old guard, as indicated by the quote from Zariski given in footnote 14.

[21]One illustration, especially pertinent here, appears in Steiner (1978) when he dismisses the suggestion that something counts as a mathematical explanation only if it can be visualized, on the grounds that such a condition would make mathematical explanation "subjective". (p. 139) I think he is right to regard the proposed criterion as inadequate, but here I'll be concerned also to spell out some ways that visualization in mathematical practice is more intricate and systematic than it might seem at first view. (This is in accord with another remark of Steiner's, which is that any "satisfactory theory of mathematical explanation must show why [the "explaining is making visual" thesis] is plausible." (p.139) I am indebted to Bertrand Guilliou here.

[22]Sometimes "pragmatic" is also used as a pejorative with the connotation "on to the 'not philosophically interesting' scrap heap with this one". So for example in his interesting account of the contributions of asymptotic explanations to our understanding of physical systems, Batterman (2002, 44) uses 'pragmatic' to frame a point with affinities to the one I am making in the text.

[23]One refinement is needed here. I'm not suggesting that no cases admit of analysis in "objective" terms. In some cases the advantages of a particular formulation can be analyzed in terms which are indisputably independent of psychological peculiarities of human reasoners. A paradigm of this sort of work is the analysis in Pratt & Lemon (1997). There certain advantages of diagrammatic reasoning are analyzed in the tangible terms of the computational complexity of algorithms. This work is extremely revealing and interesting, and I look forward to learning from further research of Pratt, Lemon and their collaborators. The attitude motivating the current work is not incompatible with that work, but rather complementary, studying some aspects of the choice of theoretical frameworks (especially in connection with the potential for visualization) that don't obviously admit an analysis in tangible complexity – theoretic terms.

(Clearly there are interesting cases where an analysis in terms of computational complexity is not going to help us much, even when what we gain are advantages in facilitated *practical* computation. One simple example is the use of homogeneous coordinates in computer modeling. The advantages of homogeneous coordinates over Cartesian coordinates are, I gather, well-established in practice, despite the (non-robust) complication of an extra parameter and the initial unfamiliarity of the framework (for most students). However, it is unlikely that computational complexity theory will support an analysis of the advantages of homogeneous coordinates, since the transition from homogeneous to Cartesian coordinates and back can be accomplished by operations that are insignificant from a complexity-theoretic point of view. And indeed, when advantages are explicitly noted for homogeneous coordinates in visual modeling they are qualitative advantages rather than complexity theoretic ones. (For example: the existence of dualities or the simplification of theoretically important expressions. See Reisenfeld (1981) and Stolfi (1991) especially ch. 1.)

[24]Another is the formulation of contemporary algebraic geometry in terms of the concept of scheme. See for example Eisenbud & Harris (1992), or the extended version Eisenbud & Harris (2000), which is largely devoted to explaining why "The scheme is... a more natural setting for many geometric arguments." (1992, 5); (2000, 8) In formulations of arguments in algebraic geometry in terms of schemes rather than antecedents like Weil's concept of "specialization", the concept of scheme is taken by some algebraic geometers to represent an advance because it is counted as "geometric". Another reason it is counted as an advance is that it supports a *unified* theory of key parts of algebraic geometry and number theory. The default assumption is surely that this unifying function is no more to be dismissed by the student of method as a "psychological" phenomenon than the unifying function of Maxwell's electromagnetic theories or relativity should be so dismissed.

Some functions of schemes are more complicated, and whether or not they are appropriate concepts depends on what questions are being addressed. In particular, the concept of scheme initially arises as an effort to extend a basic duality that occurs between restricted classes of rings, which appear in algebra, and varieties (loci of zeros of polynomials) that appear in algebraic geometry. This is a simpler version of a basic correspondence ("dictionary" in the words of (Cox et al., 1992, 168)) between ideals in simple algebraic settings and varieties in elementary algebraic geometry. This gives rise to a circumstance where two frameworks (algebraic and geometric) that are – in some important sense – equivalent are also – in another equally important sense as different as chalk and cheese. The philosophy of mathematics has emphasized the first sense, according to which the frameworks are the same if they are deductively equivalent, over the second. One could see the point of this paper as arguing that the sense in which the equivalent frameworks are crucially different also needs to be clarified before we can take ourselves to have made sense of the principles informing successful mathematical practice.

[25]The classic papers in this debate are assembled in Block (1981).

[26]The point here is not that the use of visual representation is uniform among mathematicians: it isn't. The point is rather that the preference is sufficiently widespread to make a mark on mathematical method.

[27]This has long been known. (See for example Yates (1966) for some history.) More recently it has been well-studied by cognitive psychologists. For some early research into the mnemonic advantages of imagery, see Bower (1972).

[28]Here is a quick explanation of what these are. There are familiar extensions of the real numbers gained by adding additional square roots of 1. The complex numbers are obtained by adding i and closing under $+$ and $\times$. The quaternions are obtained by adding three new roots of -1: i, j, k. For this extension to be adequately specified we need to say more than just that $i^2 = j^2 = k^2 = -1$. The results of multiplying the new elements must be stated too: $ij = -ji = k$ for example, and further details need to be made explicit. The octonions are the numbers obtained with eight roots of -1.

[29]So for example the curve from e_4 to e_7 extends to the product $e_6 = (e_4 \cdot e_7)$. The sign of the product depends on the direction of travel: counter-clockwise products are negative, clockwise are positive.

[30]To avoid misunderstanding, I should make explicit that my discussion of this example is meant only to draw a contrast. Hence the discussion gives away several points for dialectical reasons. For example, I don't mean to grant more than provisional credence to the idea that some advantages are "solely" mnemonic. The facts about how memory interacts with reasoning are quite involved. Nor indeed do I want to assert that the preference for diagrammatic representations would be devoid of philosophically interesting consequences even in cases where the preference turned out to have purely mnemonic advantage. My point only that there is an at least *prima facie* plausible case to be made for these suggestions. In presenting this material, that *prima facie* plausibility has had a sufficiently strong pull for sufficiently many people that it is worthwhile to make explicit that the case for the philosophical importance of visualization in the cases I am studying here can be made out even if these points are granted.

In fact, I think the question of how to separate the methodologically interesting from "accidental" uses of visual representation is complicated. Even cases that might seem to use vision in a philosophically uninteresting way, such as when pictures are used as memory aids, can be surprisingly involved. I concede that there are some uninteresting cases, but this doesn't mean that I want to say that all cases that might appear uninteresting in this way really *are* uninteresting in this way. In cases where visual memory aids are well–developed and systematic, as in the elaborate medieval memory systems studied in Yates (1966) and Rossi (2000) it is surprisingly hard to make out sharp boundaries between visual coding as an accidental concomitant of artificial memory techniques and visual coding that facilitates memory in virtue of being embedded in an broader system of reasoning. The intricacy of the interweaving of systems of thought and systems of visual representation that was involved in the medieval arts of memory is especially emphasized throughout the uncommonly illuminating studies Carruthers (1990, 1998). Gaukroger (1995) (p.160-164 and *passim*) points out that this perceived connection between visual imagery and thought informs Descartes' *Regulae* in striking ways. Conversations with Terri Palmer, Ian Hacking and Raviel Netz have helped me here.

Moreover, lest my frame of reference and choice of examples (stressing axiom choice and de-emphasizing actual diagrams, pictures and mental images) leave a misimpression, I should emphasize that I think that the study of the details of concrete visual representation (as in diagrams, etc.) and the manipulation of actual diagrams is extremely interesting to the philosopher of mathematics. By approaching the topic of visualization and geometry as I am, I am in no way meaning to slight those who have approached the topic of reasoning with actual diagrams. Quite the opposite, I regard the recent richness of work on the reasoning with diagrams and its connection to mathematical reasoning as of the greatest interest. (This work has been advanced from different perspectives and with reference to diverse problems by Barwise and Etchemendy (and their students), Oliver Lemon, Ian Pratt (and others

in the Manchester group studying visual reasoning) Marcus Giaquinto, Jim Brown, Robert Lindsay, and others. Also of genuine interest in this connection is the work of Achille Varzi and Roberto Casati ((1994) etc.) on the logical structure of theories of spatial structure. Readers interested in such work can choose a beginning among the papers in Allwein & Barwise (1996), or Glasgow et al. (1995) and follow out the references. Also helpful for jump-off points from additional perspectives is Pratt & Lemon (1997). An unusual and stimulating investigation of overlapping themes by two mathematicians is in Carbone & Semmes (2000). Also worth mentioning in this connection is Hartshorne's (2000) masterful reexamination of Euclid's elements.

[31] Several books by Tufte on the visual arrangement of information (see, for instance, his (1983) and (1997)) are good collections of examples. What we find here are visual representations (tables, graphs, maps) that are clearly the most effective and forceful ways to present the information they present. In these cases, the function of visual representation appears to be important solely from a "pragmatic" point of view – in the sense of "pragmatic" that seems to connote 'not deserving philosophical attention' according to some philosophers' usage. If there is philosophical interest in such examples, it will be of a different kind from what we're exploring here.

[32] This is especially pressing in the case of the contrast of Riemann and Weierstrass since just this comparison of the two was made long before Reichenbach introduced the distinction into general methodology: "The method of Riemann is above all a method of discovery; the method of Weierstrass one of proof." (Poincaré, 1898, 7)

[33] I treat this point further in Tappenden (2001b).

[34] This was, for example, true of the concept of scheme. When Grothendieck introduced it, one clear testament to its fruitfulness was that it opened the way to a proof of the Weil conjectures. But to apply the context of discovery/context of justification distinction makes no sense here. Not only did Grothendieck (and subsequently Deligne) prove the Weil conjectures using his newly introduced scheme theory, but he provided what still remain as the *only* proofs available despite extensive attempts in some circles to find proofs that avoid the Grothendieck machinery. (Here I am indebted to correspondence with Colin McLarty and a conversation with Karen Smith.)

[35] This is, of course, an example of a widespread phenomenon in studies of confirmation: the familiar debates presuppose some language or framework remaining fixed. When this can't be taken for granted, many further assumptions break down. Examples of this phenomenon are revealingly treated in Earman (1992) chapter 5 and chapter 8.4.

[36] There is one alternative that is worth attention, but I will have to leave it for another place. In a paper (Kim, 1994) that touches on some of the issues addressed here, Jaegwon Kim suggests that the causal component of causal explanations might

be derivative from a prior idea of *dependence*. If the idea of metaphysical dependence can be made out, it could in principle be extended to mathematical explanations as well, to provide a unified treatment. Michael Strevens has developed an account with this shape in unpublished work.

The idea that mathematical explanation turns on an idea of logical or metaphysical priority over others was a feature of Bolzano's account of mathematical explanation as Paolo Mancosu has pointed out in recent work. (Mancosu, 1999)

[37] The thesis of Morrison (2000), which is the best general descriptive treatment of scientific unification I know, is that unification is often an objective of scientific inquiry but it has little if anything to do with explanation. I agree with much of this, but my final position is a bit more concessive to the idea of unification as understanding: sometimes (but not always) we count ourselves as having understood or explained some phenomena because we have set them in a unified framework, though generally unification alone is not enough unless the framework has other attractive features. (I'll add in this connection that I'm completely in agreement with this upshot of Morrison's perceptive treatment: theory unification is far more complicated in practice than it often is taken to be in the literature.)

[38] I should note that the positions Friedman takes in this early work need not be preserved in his more recent writings. Indeed, his most recent work on scientific theories and the "relativized a priori" has obvious affinities with the "rapprochement between (early) Friedman and Toulmin" that I suggest is necessary.

[39] The most effective display of counter – examples is in Kitcher (1976). The whole controversy is given a retrospective postmortem in (Salmon, 1989, 94 – 101).

[40] To simplify the discussion I am assuming that the theories we are dealing with are given to us already rendered into axiomatic form. In practice, of course, this can't always be assumed. (For example, continuum mechanics was studied for many years before Noll provided an axiomatization, as Clifford Truesdell often pointed out. (Truesdell, 1984, p. 137 and *passim*)) But since I am just using the assumption to simplify the formulation of this negative point against Friedman, the assumption is harmless here. Clearly if we are dealing with an unaxiomatized theory, the "reduction in the number of basic principles" account is even harder to defend.

[41] A variation on this point has already been made effectively by Humphries (1993), who pointed out the disconnect between number of axioms and understanding with respect to various axiomatizations of propositional logic.

[42] Readers familiar with the scholarly literature on William Whewell will recognize echoes of the ideas here in Whewell's notion of "consilience". Though there are no specific points at which this essay was informed by this literature, I do owe a general debt to the papers (Butts, 1973) and Morrison (1997).

[43] Algebraic numbers are real number solutions to polynomials $x^n + a_1 x^{n-1} + \ldots + a_n$ where the coefficients a_i are rational. Algebraic functions result when the a_i are themselves one-variable functions. Useful treatments of this material in the secondary literature are W. Geyer (1981) and throughout Dugac (1976). (Dieudonné (1985) is a good, though brief English – language discussion of the content of

Dedekind & Weber (1882). A clear (mid-twentieth century presentation) of general versions of results of Dedekind & Weber (1882), in a broadly Dedekindian style, plus a nod to other styles of proving the same results, is in (van der Waerden, 1991, Ch. 19). Another textbook presentation of the theory of algebraic functions of one variable that is avowedly in the spirit of Dedekind & Weber (1882) is Chevalley (1951).

[44]In retrospect, Dedekind & Weber (1882) appears as one of a handful of papers of the nineteenth century that inaugurated distinctive styles marking the twentieth century in mathematics. Dieudonné (1985) lauds the paper for originality and importance, and counts it as second only to Riemann's work in its "introduction of a series of notions which have become fundamental in the modern era."

[45]A field is a collection of objects with two associative, commutative operations defined on the whole collection. Relative to one of the operations (addition) there is an identity element 0 (one for which $a + 0 = 0 + a = a$) and every element a has an inverse a^{-1} such that $a + a^{-1} = 0$. Relative to the other operation (multiplication) there is an identity and inverses for the collection consisting of every element but the additive identity 0. Distributive laws hold.

To convey what an ideal amounts to, I'll define a special case (though the more general definition exploits the concept of "ring", which is weaker than "field"). A field I contained in another field F is an ideal, if given any a in I and any b in F the product ab is in I.

For the precise definition see any university level text, such as Jacobson (1974).

[46]Again we find this emphasis on the "qualitative" unification in contemporary discussions in algebraic geometry. This is not just true of the schemes and motives discussed in earlier footnotes; another instance – not at all exceptional – is the discussion in Smith et al. (2001) of the pre-Grothendieck work by Weil and Zariski as distinguished by how it brought out "deep connections between previously separate areas of mathematics, such as number theory and the theory of Riemann surfaces" (p.2) Once again, it is hard to see that these connections would lose any value, or be any less unifying, if they turned out not to reduce the number of brute facts in Friedman's sense.

[47]A historical aside: this is a point over which Frege and Hilbert simply stood at cross–purposes. Frege held as Friedman does (mistakenly, I think) that there is an intrinsic advantage to be gleaned from reducing the number of axioms, and indeed he held that the value of an explanation was directly proportional to the reduction. (1979, 36) A hint of this difference shows up in Frege's reaction to his first viewing of Hilbert's foundations. Frege states that he (Frege) believed he could have made do with fewer primitives. (cf. (Frege, 1980, 35))

[48]For example, Friedman argues that the well-known deductive – nomological account Hempel proposes falls afoul of the second requirement – it fails to connect explanation with something plausibly called "understanding" – though Friedman counts it as appropriately objective.

[49]I leave aside the expression "self-evident", as it has epistemological resonance that I want to shed. A proposition can be a reasonable choice as an axiom without being obvious, and a category can be a reasonable choice as natural without it appearing natural on first encounter. We'll see some examples in section 5.

[50]An accessible discussion of the basic points of this historical evolution, with glimpses at more recent showings of the action-at-a-distance view like the Feynman – Wheeler incarnation, is in (Hesse, 1961, ch. VII and VIII and p. 279-289). The discussion throughout Darrigol (2000) is illuminating on the give – and – take between action-at-a-distance accounts and rivals. A. Assis (1994) develops one of the nineteenth century theories in considerable detail, from a contemporary perspective, with an illuminating systematic comparison between theories in the Gauss – Weber style and the Maxwell – inspired theories that dominate today.

[51]See (Frege, 1980, 57) and *passim*. Also (Frege, 1903, 160) and *passim*.

[52]There are additional details in Kitcher's subtle analysis, but they will not be relevant to the points I'll be making here. For the full account of Kitcher's presentation of "patterns of argument" see his (1989, p. 432-435)

[53]More on the first example: the Erdös – Selberg ("elementary") proofs of the prime number theorem contrast with the ("analytic") proofs exploiting the Riemann zeta function following the path blazed by Hadamard and De Vallée-Poussin. The former have the advantage that they use only "elementary" techniques, while the latter, though presupposing much more analytic machinery, seem to be widely held to better "go to the heart of the matter." (Even setting aside such a suggestion as potentially too loaded it is clear that the analytic proofs are shorter, far less intricate and more easily understood.) For a textbook presentation of both styles of proof (presupposing only high school mathematics) see (Apostol, 1976, ch. 4 and ch. 13).

The many proofs of the Riemann – Roch theorem serve up a more complicated story, which I hope to discuss in further work. The early history is illuminatingly discussed in Gray (1998).

[54]It is worth pointing out as well that this is a twentieth-century example, which raises issues that remain alive in current work. Indeed, preserving this duality in a general setting is one of the more elementary functions of the concept of (affine) scheme. I mention this to reinforce the point that the issues raised by the development of projective geometry in the nineteenth century are not confined to some distant time, irrelevant to mathematics as it is currently practiced.

[55]You don't have to know what the *Nullstellensatz* is. I retained reference to it in the quote just as a benchmark for discussions in future work. (The point is that where the *Nullstellensatz* doesn't generally hold (in finite fields, for example), a new concept is needed to retain the algebra – geometry correspondence. This is part of the work that the concept of scheme does for us.)

[56]A similar point noted in a different connection by Batterman (2000, 233), who complains that the unification account "fails to respect the individuality of problems" (a neat phrase he attributes to Mark Wilson). I don't see this as an objection to

the unification account as much as it is a further subtlety that the account should incorporate. Unification is an important goal in scientific practice, but a multi-faceted one.

[57]Anyone who reflects on the miraculous solution of the integral $\int_{-\infty}^{\infty} e^{-x2} dx$ effected by recasting the question into polar coordinates will know what I mean.

[58]Birational transformations are 1-1 transformations that can be represented as fractions in which both the numerator and denominator are polynomials; in the plane these are also called Cremona transformations. An old-fashioned, concrete presentation of the topic is in Coolidge (1959). A presentation in more contemporary terms is in (Smith et al., 2000) see especially ch. 7.

[59]I'm grateful to Karen Smith for help with this example.

[60]Another example is the use of "reciprocal diagrams" to be considered in a few pages. Here too we have a device which both creates a dual pattern of reasoning that is interesting both because it isolates a significant general pattern and also because it allows the exploitation of shifts from special case to special case.

Naturally we don't see this interaction between general case and special just in mathematics: it shows itself whenever one physical realization of a general theory is used as a tangible model for another.

[61]There are very simple examples of properties that appear disjunctive in one context but which are revealed not to be disjunctive in the "natural setting". Whether or not a property is "disjunctive" can depend upon ontology – on what objects there are in the domain. The simplest example is perhaps the idea of "intersection" in the projective plane. In the Euclidean plane, arguments typically have annoying special cases that arise when two lines are parallel. By expanding the plane with "points at infinity" where parallel lines intersect, the artificial predicate "intersecting or parallel" becomes simply "intersecting", thus eliminating the special cases. (This motive for introducing points at infinity is discussed in many introductory level discussions; see for example (Courant & Robbins, 1941, 180ff).)

[62]One point is worth noting in passing here, though it is sufficiently complicated that I'll have to set it aside here; I will be developing it in future work. There is this much of an anchor to the idea of "projectibility" in mathematics, in that a judgement to the effect that a definition or principle is fruitful incorporates a prediction that results of desired kinds will in fact be produced in the future by those who adopt the definition or principle as part of their working repertoire. The connections between these previsions of future discoveries and judgements of plausibility in mathematics are quite involved, and bear some affinities to versions of the problems of old evidence that are familiar in the study of Bayesian methodology. Both seem to arise from a common root, in which assessments of likelihood depend crucially on expectations that the empirical event of the discovery of a necessary truth occurs.

[63](I will concentrate just on the plane for simplicity – similar patterns emerge in higher dimensions.)

[64]Some expressions – conic section, hexagon – are self-dual. In these theorems, "hexagon" is understood more broadly than we learned in school.

[65]Notably Chasles, in an influential essay of (1837) who called it a "general principle of science". For a statement of the general methodological significance of duality in the English literature, J. Booth (1873) is an especially unrestrained example. See especially p. xi - xiv

[66]Of course, it is important for the plausibility of Kitcher's account that we should find at least *some* cases where theories are consciously designed in a way that conforms to his general picture. If Kitcher is right that unification as he characterizes it is a governing objective of scientific and mathematical practice, we should expect that sometimes the pursuit of the goal would be self – conscious. So it is reassuring that in the case of nineteenth-century duality, the pattern Kitcher presents *was* self-consciously pursued.

Another case in which a variation on Kitcher's picture was a self-consciously adopted methodological guide was in the early nineteenth century debate over the adoption of the Leibnizian notation in Great Britain. Babbage, in his essay "On the Influence of Signs in Mathematical Reasoning" (1827) spells out examples of how a careful choice of notation can unify a proof that consists of several distinct arguments in Newton's *Arithmetica Universalis* into one single pattern. I discuss this in more detail in Tappenden (2001b).

[67]A good short overview of the subject is Scholz (1994). For further details, see Benvenuto (1991), Timoshenko (1953) and Charlton (1982). The role of duality considerations in the development of graphic statics is especially well brought out in Scholz (1989). Dubois (1877) is an English – language textbook of the time which gives a good glimpse into the subject and the attitudes toward it. Also helpful is Graham (1887) which contains extended contrasts and comparisons of analytic and graphic methods.

[68]Culmann (1865); the projective background is made more explicit and systematic in the second edition (1875).

[69]These are sometimes called *Cremona diagrams* because Luigi Cremona popularized the technique in his widely used textbooks Cremona (1872) and (1874). (English translations in (Cremona, 1890).)

[70]This particular example is taken from Ziwet (1904) p. 226; I have chosen this example both because it is a good illustration of the point and also for a somewhat sentimental reason. The long – dead Ziwet has been a great help in my current projects (in ways that it would take too long to explain) so I'm happy to grasp the opportunity to cite him in some way. But similar examples are analysed in sources that are easier to obtain today: so for example there are several examples like this one worked out with characteristic clarity in the Schaum's outline on statics and strength of materials. (Jackson & Wirtz, 1983, 117 - 135).

[71]It should be noted, though, that graphic statics is not a perfect illustration in one respect: there is only an indirect connection between the dualistic patterns that best exemplify Kitcher's picture of explanation and the applications to engineering. (This is, of course, a price that has to be paid for choosing actual examples: the real world rarely serves up events that are as clean as the thought experiments that can be crafted at will in the thought laboratory.) On the abstract side, in general

projective geometry, the parallel patterns of argument were, and were taken to be, an important feature of the theoretical structure. In its applied form in graphic statics, dual diagrams played a central role, but these dualities didn't translate into a line-by-line parallelism of proofs and arguments, as in the abstract case. (I have only found one engineering textbook – Crotti (1888) – where the formal duality of argument is spelled out explicitly and represented in the "dual columns" format of projective geometry textbooks. (According to Charlton (1982, 155)) Crotti's text was unique in this regard.)) What we have in graphic statics is a case in which a framework explicitly informed by Kitcher – type patterns of multiple argument is applied to concrete problems, thereby coloring what counts as explanations of these concrete problems. This is good enough for the present purposes, though the example would of course be cleaner if the dualities of argument that shape the abstract mathematical investigations figured more prominently in the engineering applications.

[72]On the Quebec bridge collapse see (Ferguson, 1992, 172-178). An illuminating glimpse into the patterns of explanation characteristic of turn-of-the-century engineering can be found in the pages of the professional weekly *Engineering News* during the months after the disaster, where candidate reasons for the collapse are dissected and discussed at length.

[73]I am indebted here to Nancy Cartwright (1983, 56-67) who advances the similar point that against the background of realism about forces, patterns of vector addition and decomposition may involve reference to theoretical fictions. My point here is different – I am setting aside any questions of ontology – but Cartwright's discussion was helpful in nudging my thoughts at a crucial juncture.

[74]An illustration of its importance is that graphical statics was taken to deserve a massive (90 page) chapter to all to itself in the Physics volume of Klein's *Encyklopädie der Mathematischen Wissenschaften*. See Hennenberg (1903).

[75]Though it might be noted that a residual nostalgia for the older techniques persists. In his discussion of graphic statics, Ferguson remarks: "Even though digital computers are making graphical methods seem both old-fashioned and insufferably slow, a few younger engineers, along with the old fogeys, are beginning to understand that speed has sometimes been bought at the cost of understanding." (Ferguson, 1992, 152)

In this connection it is worth noting further that some of the old results of graphic statics have recently been revived and generalized. (See for example (Crapo & Whiteley (1982)) and Whiteley (1985). On this work I am grateful to Walter Whiteley for email correspondence and to Branko Grünbaum for sending me a copy of his unpublished lectures Grünbaum (1976).) Here too the interest of the results does not depend exclusively upon the visualizable character of the represented structures, though the visual flavor of the work is still important.

[76]See for example (Dubois, 1877, iv) (Cremona, 1890, 121,123-4,131-137) (Culmann, 1875, vii–xv etc.)

[77]So for example, the engineering professor Rankine remarks that an advantage of the graphical methods compared to analytical methods that they make mistakes

much easier to catch. ((Rankine, 1869, 411); quoted with tacit agreement in (Cremona, 1890, 133)). The point was reiterated in these terms in more recent days by a contemporary engineer defending the virtues of the old ways: "When nearly all engineers carried out structural analysis using ... graphic statics [and similar methods]... the advantages of visually monitoring one's calculations (Does it look right? Are the numerical answers reasonable?) were built into the graphical mathematics they used." (Ferguson, 1992, 152)

[78] As one textbook of graphic statics puts it, with a quaint Victorian flair: "... the power conferred by the graphical method is to a large extent at the disposal of those who have had but little mathematical training. The writer once had occasion to explain a practical application of the triangle of forces to a class of working men, who seemed at once to grasp and appreciate it." (Clarke, 1888, v)

[79] A characteristic opinion is expressed in Rankine's discussion of reciprocal figures:

> When compared with algebraic methods, the simplicity and rapidity of execution of the graphic method is very striking... If this is the case when the loads are uniform or symmetrical, the advantage is much more strikingly in favour of the graphic method when the loads are not symmetrical, and when they are inclined... or as in such cases as the framed arch and suspension bridge. In fine, the diagram once drawn acts as a sort of graphic formula for the strain on every part of the bridge or roof, and it is a formula which can hardly be misapplied. ((Rankine, 1869, 441); part of this passage is quoted with tacit agreement in (Cremona, 1890, 133))

[80] One example is Dubois (1877).

[81] This is an oft-repeated theme in the literature on analytic projective geometry; I'll mention just two examples that illustrate the point. Referring in particular to the nineteenth century analytic geometer Plücker, Felix Klein commends his style of argument in these terms:

> In Plücker's geometry, the bare combination of equations is translated into geometrical terms, and the analytic operations are led back through the geometric. Computation is avoided as much as possible, but by doing this, a mobility heightened to the point of virtuosity, of inner intuition, of the geometric interpretation of given analytic equations, is cultivated and extensively applied. (Klein, 1926/79, 110)

Bear in mind that Klein is here discussing someone who bucked the trend of the then – dominant synthetic geometry in favor of streamlined analytic methods. The praise is not for the use of diagrams but rather for a certain way of organizing the analytic methods so as to gain an elegant means of addressing the subject. This is reinforced by the Klein's subsequent illustration of his remarks with "an example of Plücker's way of thinking" (p.110). He presents a device ("abridged notation") Plücker used to systematically manipulate and transform analytic equations so as to

better fix on the crucial parts of the underlying geometric situation. Even without any connection to vision, the symbolic technique was, and remains a valuable (if now somewhat old-fashioned) tool.

A more recent commentator on essentially the same phenomenon of Plücker's style of analytic argument is Dieudonné:

> One of the attractions of [nineteenth century] complex projective geometry is its relative independence from algebra and the formal independence of its results, in contrast to the massiveness of most of the coordinate calculations of the preceding century.... Möbius, Plücker, and Cayley give projective geometry a solid base by the use of homogeneous coordinates accompanied by a harmonious choice of indexing notation that maintains a symmetry and a clarity in the calculations so that they closely follow the geometric argument. (Dieudonné, 1985, 9)

Here again an advantage of the "geometric" framework is taken to include an elegant way of formulating the subject matter, which happens to have an important tie to visual representation but which is valuable independently of it.

[82]A particularly charming illustration of this point appears in Coxeter's textbook *The Real Projective Plane* where Desargues' theorem is adopted as an axiom of projective geometry. After providing one proof of Desargues' theorem, Coxeter remarks "Since we will eventually take Desargues' theorem as an axiom, it seems worthwhile to give an alternative proof." (1992, 7) and he proceeds to give it.

[83]It should be noted that Artin's use of "geometric" is somewhat idiosyncratic. Artin was one of the greatest forces propelling the abstract turn of twentieth century algebra, and even when working in a self-consciously "geometric" vein, his tastes tilt to the algebraic. This makes his treatment especially useful for present purposes: "geometry" for Artin turns out to have an exceedingly indirect connection to *vision*.

This peculiarity of Artin's attitude hasn't gone unnoticed. A noted algebraic geometer told me in conversation that in his opinion Artin's text was "not really geometric" (except in the sense emerging from the Klein program of characterizing geometries with groups of transformations). We also find a variant of this opinion in a review of Artin (1957):

> Most of this book is devoted to the study of algebraic structures arising from various geometries. The approach is algebraic rather than geometric...
>
>   ...
>
> In Chapter II [the focus of this article], affine geometry is introduced axiomatically and then coordinatized. Even here the approach is algebraic. (Jans, 1957, 604)

[84]Not everyone, though. The approach to matrices in Edwards (1996) is motivated by an explicit preference for computations that is apparently as strong as Artin's animus.

[85]Another advantage, mentioned in some textbooks that adopt the Artin approach, is that it allows a smooth introduction of coordinates, in contrast to the "messy" algebraic approach of (for example) Hall (1943). See (Hartshorne, 1967, 101) and (in nearly identical language) Kadison & Kromann (1996) (p. 105 – 106) Here too there is nothing specifically "visual" about this advantage.

[86]In a treatment of overlapping material, with some ideological affinities to Artin's book (Dieudonné, 1969) the absence of diagrams is principled. After emphasizing the importance of material which can be represented in visual intuition (p.12), Dieudonné continues: "I have taken the liberty of omitting all diagrams from the text, if only to show that they are unnecessary" (p.13) Though Dieudonné doesn't present himself as resurrecting a geometric presentation as Artin does, the approach is nonetheless "geometric" in Artin's sense to the extent that I) the concept of mapping rather than computation with coordinates is explicitly marked out as basic and systematically developed. (p.13-14 and *passim*) II) the core intuitions are spun out from a consideration of intuitive maps on linear varieties. (see especially chapter III) So we can draw the same conclusion: that this material admits of representation in diagrams is valuable, but it doesn't exhaust the value of this particular framework for organizing information.

[87]Though I should note that the point about the absence of pictures is slightly softened by Artin's instruction to the reader to draw pictures while reading. (Artin, 1957, 52) But even with this qualification it is clear from Artin's discussion that he sees the role of pictures as secondary in his "geometric" presentation.

[88]This historical point should be flagged, though it would represent too much of a digression to work it out here. Artin's approach to rendering the idea of geometric intuition rigorous has a distinguished pedigree: many debates in the mathematics of the nineteenth century are illuminated if this is recognized. In particular, as Michael Friedman points out in a superb article, (Friedman, 2000) Helmholz interpreted Kantian "intuition" in terms of transformations of space, setting aside the idea of "construction in intuition". (Friedman credits Robert DiSalle for key observations in this connection.) This forged a bridge with the new geometry that was then emerging.

The Friedman-DiSalle insight that the Kantian idea of intuition was undergoing a metamorphosis among scientifically informed students of geometry is of broad significance for our understanding of mathematicians' talk of intuition at the time. In particular, the insight allows one to flesh out some gnomic remarks Frege makes in *Grundlagen* about intuition and geometrical knowledge, and fit them into other features of his mathematical environment. I develop this point further in Tappenden (2001a).

[89]This version of the definition leaves aside a degenerate case that will be of no interest here.

[90]The second axiom implies the first.

[91]It deserves mention, but I won't expand on this point here, that classic versions of the Desargues theorem have historically been quite important in applied geometry, especially in the development of the theory of linear perspective in painting and

architecture. A good general source is Field (1997). On the geometry and theory of perspective of the historical Desargues see Field and Gray (1987). On Desargues' theorem and the historical development of perspective, Field (1987) and (1988) are detailed and helpful.

[92]This point is discussed in (Blumenthal, 1961, 81–84).

[93]Artin notes this fact on p. 51. It is discussed more extensively in this symmetry-based context by (Kadison & Kromann, 1996, 116–120).

[94]The additional constraint that the multiplication operation be commutative (i.e. that the coordinates are a field) corresponds to a further geometric axiom with a classical pedigree: the Pappus theorem. The Pappus theorem is another illustration of the themes of this section, but I will leave discussion of it for another time.

[95]An interesting discussion of the Desargues theorem from this point of view is in (Rota, 1997, 140–146) especially p. 141 on the "zen ideal" combined with many applications and 145 on the "horizon of possibilities" the Desargues theorem opens up. Rota's account of reasons for regarding the Desargues theorem as central draws on a venerable analysis of Desargues' theorem in connection with the underlying combinatorial situation (the "Desargues configuration") detailed at length in Baker (1929). This, incidentally, gives yet another point of view from which the Desargues axioms turn out to represent a natural carving point: it also corresponds to deep and rich facts in finite combinatorics. I won't be exploring this perspective further here: for those who are interested a contemporary introduction to this point of view, in which the Desargues configuration shows itself prominently, is Batten (1997).

REFERENCES

Abhyankar, S. (1964). *Local Analytic Geometry*, Academic Press, New York.

Allwein, G. & Barwise, J. (1996). *Logical Reasoning with Diagrams*, Oxford University Press, Oxford.

Apostol, T. (1976). *Introduction to Analytic Number Theory*, Springer, Berlin.

Artin, E. (1957). *Geometric Algebra*, Wiley Interscience, New York.

Assis, A. (1994). *Weber's Electrodynamics*, Kluwer, Dordrecht.

Atiyah, M. (1984). Interview with Michael Atiyah, *Mathematical Intelligencer* **6**(1). Page references to the reprinting in (Atiyah, 1988).

Atiyah, M. (1988). *Collected Works vol.1*, Oxford University Press, Oxford.

Babbage, C. (1827). On the Influence of Signs in Mathematical Reasoning, *Transactions of the Cambridge Philosophical Society* **II**: 325 – 377.

Baker, H. (1929). *Principles of Geometry vol. I Foundations*, Cambridge University Press, Cambridge.

Batten, L. (1997). *Combinatorics of Finite Geometries*, 2nd edn, Cambridge University Press, Cambridge.

Batterman, R. (2000). A 'Modern' (Victorian) Attitude Towards Scientific Understanding, *The Monist* **83**(2): 228 – 257.

Batterman, R. (2002). *The Devil is in the Details: Asymptotic Reasoning in Explanation, Reduction and Emergence*, Oxford University Press, Oxford.

Benvenuto, E. (1991). *An Introduction to the History of Structural Mechanics Part II: Vaulted Systems and Elastic Structures*, Springer - Verlag, Berlin.

Block, N. (1981). *Imagery*, MIT Press, Cambridge Mass.

Blumenthal, L. M. (1961). *A Modern View of Geometry*, W. H. Freeman and Co., San Francisco.

Booth, J. (1873). *A Treatise on Some New Geometrical Methods*, Vol. I and II, Longman Green, London.

Bottazini, U. (1986). *The Higher Calculus: A History of Real and Complex Analysis from Euler to Weierstrass*, Springer, Berlin. W. van Egmund (trans.).

Bottazini, U. (1994). Three Traditions in Complex Analysis: Cauchy, Riemann and Weierstrass, *in* I. Grattan-Guinness (ed.), *Companion Encyclopedia of the History and Philosophy of the Mathematical Sciences*, Vol. I, Routledge, London, pp. 419–431.

Bower, G. (1972). Mental Imagery and Associative Learning, *in* L. Gregg (ed.), *Cognition in Learning and Memory*, Wiley, New York.

Bridgeman, P. (1938). *The Logic of Modern Physics*, McMillan, New York.

Butts, R. (1973). Whewell's Logic of Induction, *in* R. Giere & R. Westfall (eds), *Foundation of Scientific Method: Nineteenth Century*, Indiana University Press, Bloomington, pp. 53–85.

Campbell, N. (1919/1957). *Foundations of Science*, Dover, New York. 1957 reprint of 1919 original.

Carbone, A. & Semmes, S. (2000). *A Graphic Apology for Symmetry and Explicitness*, Oxford University Press, Oxford.

Carruthers, M. (1990). *The Book of Memory: A Study of Memory in Medieval Culture*, Cambridge University Press, Cambridge.

Carruthers, M. (1998). *The Craft of Thought*, Cambridge University Press, Cambridge.

Cartier, P. (2001). A Mad Day's Work: From Grothendieck to Connes and Kontsevich, the Evolution of the Concepts of Space and Symmetry, *Bulletin of the American Mathematical Society* **38**(4): 389 – 408.

Cartwright, N. (1983). *How the Laws of Physics Lie*, Clarendon Press, Oxford.

Casati, R. & Varzi, A. (1994). *Holes and other Superficialities*, MIT Press, Cambridge Mass.

Casorati, F. (1868). *Teorica delle funzioni di variabili complexi*, Pavia.

Charlton, T. (1982). *A History of the Theory of Structures in the Nineteenth Century*, Cambridge University Press, Cambridge.

Chasles, M. (1837). *Aperçu historique sur l'origine et le développement des méthodes en géométrie, particulièrement de celles qui se rapportent à la géométrie moderne, suivi d'un mémoire de géométrie sur deux principes généreaux de la science, la dualité et l'homographie*, M. Hayez, Bruxelles.

Chevalley, C. (1951). *Introduction to the Theory of Algebraic Functions of One Variable*, American Mathematical Society, Reinhart and Wilson, New York.

Clarke, G. (1888). *The Principles of Graphic Statics*, 2nd edn, E. & F. N. Spon, London.

Clemens, C. (1980). *A Scrapbook of Complex Curve Theory*, Plenum Press, New York.

Coolidge, J. (1959). *A Treatise on Algebraic Plane Curves*, Dover, New York.

Courant, R. & Robbins, H. (1941). *What is Mathematics?*, Oxford University Press, Oxford.

Cox, D., Little, J. & O'Shea, D. (1992). *Ideals, Varieties and Algorithms: An Introduction to Computational Algebraic Geometry*, Springer, Berlin.

Coxeter, H. (1992). *The Real Projective Plane*, 3rd edn, Springer, New York.

Crapo, H. & Whiteley, W. (1982). Statics of Frameworks and Motions of Panel Structures: A Projective Geometric Introduction, *Structural Topology* **6**: 43 – 82.

Cremona, L. (1872). *Le figure reciproche nella statica grafica*, Hoepli, Milan.

Cremona, L. (1874). *Elementi di calcolo grafico*, Paravia, Torino.

Cremona, L. (1890). *Graphical statics. Two treatises on the graphical calculus and reciprocal figures in graphical statics*, Oxford University Press, Oxford. English translation of (Cremona, 1872) and (Cremona, 1874) by Thomas Beare, with a new preface by Cremona.

Crotti, F. (1888). *La teoria dell'elasticità ne' suoi principi fondamentali e nelle sue applicazioni pratiche alle costruzioni*, Hoepli, Milan.

Culmann, K. (1865). *Die Graphische Statik*, Meyer and Zeller, Zürich.

Culmann, K. (1875). *Die Graphische Statik*, 2nd edn, Meyer and Zeller, Zürich.

Darrigol, O. (2000). *Electrodynamics from Ampère to Einstein*, Oxford University Press, Oxford.

Dedekind, R. & Weber, H. (1882). Theorie der algebraischen Funktionen einer Veränderlichen, *Journal für Reine und Angewante Mathematik* **92**: 181 – 290. Dated 1880, page reference to the reprinting in Dedekind's Werke vol. I 238 - 249.

Demopoulos, W. (ed.) (1995). *Frege's Philosophy of Mathematics*, Harvard University Press, Cambridge.

Dieudonné, J. (1969). *Linear Algebra and Geometry*, Hermann, Paris.

Dieudonné, J. (1985). *History of Algebraic Geometry*, Wadsworth Books, Monterey. J.Sally, (trans.).

Dubois, A. (1877). *The Elements of Graphical Statics and their Application to Framed Structures*, Wiley and Sons, New York.

Dugac, P. (1976). *Richard Dedekind et les Fondements des Math ématique*, Vrin, Paris.

Earman, J. (1992). *Bayes or Bust: A Critical Examination of Bayesian Confirmation Theory*, MIT press, Cambridge.

Edwards, H. (1996). *Linear Algebra*, Birkhäuser, Boston.

Eisenbud, J. & Harris, J. (1992). *Schemes: The Language of Modern Algebraic Geometry*, Wadsworth and Cole, Pacific Grove, Ca.

Eisenbud, J. & Harris, J. (2000). *The Geometry of Schemes*, Springer, Berlin.

Ferguson, E. (1992). *Engineering and the Mind's Eye*, MIT Press, Cambridge Mass.

Field, J. V. (1987). Linear perspective and the projective geometry of Girard Desargues, *Nuncius Ann. Storia Sci.* **2**(2): 3–40.

Field, J. V. (1988). Perspective and the Mathematicians: Alberti to Desargues, *in* C. Hay (ed.), *Mathematics from Manuscript to Print, 1300–1600*, Oxford Sci. Publ., Oxford Univ. Press, New York, pp. 236–263.

Field, J. V. (1997). *The Invention of Infinity*, Oxford University Press, Oxford.

Field, J. V. & Gray, J. J. (1987). *The Geometrical Work of Girard Desargues*, Springer-Verlag, New York.

Frege, G. (1903). *Grundgesetze der Arithmetik*, Vol. II, H. Pohle, Jena.

Frege, G. (1980). *Philosophical and Mathematical Correspondence*, University of Chicago Press, Chicago. Gabriel, G. Hermes, H. Kambartel, F. Thiel, C and Veraart, A. Abridged from the German edition by McGuinness, B. Translated by Kaal, H.

Friedman, M. (1974). Explanation and Scientific Understanding, *Journal of Philosophy* **73**: 250 – 261.

Friedman, M. (1979). Truth and Confirmation, *Journal of Philosophy* **76**: 361 – 382.

Friedman, M. (2000). Geometry, Construction and Intuition in Kant and his Successors, *in* G. Sher & R. Tieszen (eds), *Between Logic and Intuition*, Cambridge University Press, Cambridge, pp. 186–218.

Gaukroger, S. (1995). *Descartes: A Critical Biography*, Oxford University Press, Oxford.

Geyer, W. (1981). Die Theorie der Algebraischen Funktionen einer Veränderlichen nach Dedekind und Weber, *in* W. Scharlau (ed.), *Richard Dedekind, 1831-1981: eine Würdigung zu seinem 150 Geburtstag*, F. Vieweg, Braunschweig, Wiesbaden, pp. 109–133.

Glasgow, J., Narayanan, H. & Chandrasekaran, B. (eds) (1995). *Diagrammatic Reasoning*, MIT Press, Cambridge Mass.

Goodman, N. (1955). *Fact, Fiction and Forecast*, Bobbs - Merrill, Indianapolis. 1965 Reprint.

Graham, R. (1887). *Graphic and Analytic Statics in their Practical Application to the Treatment of Stresses in Roofs, Solid Girders, Lattice, Bowstring and Suspension Bridges, Braced Iron Arches and Piers, and other Frameworks*, 2nd edn, Crosby Lockwood and Co., London.

Gray, J. (1998). The Riemann-Roch Theorem and Geometry 1854 - 1914, *International Congress of Mathematicians Proceedings*, Vol. III, pp. 811–822.

Grünbaum, B. (1976). Lectures in lost mathematics. Unpublished mimeographed notes.

Hall, M. (1943). Projective planes, *Transactions of the American Mathematical Society* **54**: 229 – 277.

Harris, J. (1992). Developments in Algebraic Geometry, *Proceedings of the AMS Centennial Symposium*, American Mathematical Society Publications, Providence.

Hartshorne, R. (1967). *Foundations of Projective Geometry*, Benjamin Press, Cambridge. Lecture Notes Harvard University.

Hartshorne, R. (2000). *Geometry: Euclid and Beyond*, Springer Verlag, Berlin.

Hennenberg, E. (1903). Die graphische Statik der starren Körper, *Encyklop ädie der Mathematischen Wissenschaften*, Vol. IV(1), pp. 345–435.

Hesse, M. (1961). *Forces and Fields*, Thomas Nelson and Sons, London.

Hilbert, D. (1899). *Die Grundlagen der Geometrie*, 1st ed. in *Festschrift zur Feier der Enthüllung der Gauss-Weber-Denkmals in Göttingen*,Teubner, Leipzig.

Hughes, D. R. & Piper, F. C. (1973). *Projective Planes*, Vol. 6 of *Graduate Texts in Mathematics*, Springer-Verlag, New York.

Humphries, P. (1993). Greater Unification equals Greater Understanding?, *Analysis* **53.3**: 183 – 188.

Jackson, J. & Wirtz, H. (1983). *Schaum's Outline of Theory and Problems of Elementary Statics and Strength of Materials*, McGraw - Hill, New York.

Jacobson, N. (1974). *Basic Algebra*, Freeman, San Francisco.

Jans, J. (1957). Review of (Artin, 1957), *American Mathematical Monthly* pp. 604 – 605.

Kadison, L. & Kromann, M. (1996). *Projective Geometry and Modern Algebra*, Birkhäuser, Boston.

Kim, J. (1994). Explanatory Knowledge and Metaphysical Dependence, *Philosophical Issues* **5**: 51 – 69. Truth and Rationality.

Kitcher, P. (1976). Explanation, Conjunction and Unification, *Journal of Philosophy* **73**: 207 – 212.

Kitcher, P. (1982). Explanatory unification, *Philosophy of Science* **48**: 507–531.

Kitcher, P. (1989). Explanatory Unification and the Causal Structure of the World, *in* P. Kitcher & W. Salmon (eds), *Scientific Explanation*, University of Minnesota Press, Minneapolis.

Klein, F. (1926/79). *Development of mathematics in the 19th century*, Vol. IX of *Lie Groups: History, Frontiers and Applications*, Math Sci Press, Brookline, Mass. With a preface and appendices by Robert Hermann, translated from the German by M. Ackerman.

Kuhn, T. (1977). Objectivity, Value Judgement and Theory Choice, in *The Essential Tension*, University of Chicago Press, Chicago.

Laugwitz, D. (1992). 'Das letzte Ziel ist immer die Darstellung einer Funktion': Grundlagen der Analysis bei Weierstraß 1886 historische Wurzeln und Parallelen, *Historia Mathematica* **19**: 341 – 355.

Laugwitz, D. (1999). *Bernhard Riemann: 1826 - 1866 Turning Points in the Conception of Mathematics*, Birkhäuser, Boston. A. Schnitzer (trans.).

Mancosu, P. (1999). Bolzano and Cournot on Mathematical Explanation, *Revue d'Histoire des Sciences* **52**: 429 – 455.

Maxwell, J. C. (1864). On the Calculation of the Equilibrium and Stiffness of Frames, *Philosophical Magazine* **27**: 294 – 299.

Morrison, M. (1997). Whewell and the Ultimate Problem of Philosophy, *Studies in the History and Philosophy of Science* **28**: 315 – 324.

Morrison, M. (2000). *Unifying Scientific Theories*, Cambridge University Press, Cambridge.

Mumford, D. (1988). *The Big Red Book of Varieties and Schemes*, Springer - Verlag, Berlin. Springer Lecture Notes in Mathematics: **1358**.

Neuenschwander, E. (1980). Riemann und das 'Weierstrass'schen Prinzip' der analytischen Fortsetzung durch Potenzreihen, *Jahresbericht der Deutschen Mathematiker-Vereinigung* **82**: 1 – 11.

Neuenschwander, E. (1981). Studies in the History of Complex Function Theory II: Interactions among the French School, Riemann, and Weierstrass, *Bulletin of the American Mathematical Society* **5**: 87 – 105. New series.

Poincaré, H. (1898). L'Oeuvre Mathmmatique de Weierstrass, *Acta Mathematica* **22**: 1–18.

Poincaré, H. (1903). Review of Hilbert's Foundations of Geometry, *Bulletin of the American Mathematical Society* pp. 1 – 23. E. Huntington, trans., original, published 1902.

Pratt, I. & Lemon, O. (1997). Spatial Logic and the Complexity of Diagrammatic Reasoning, *Machine Graphics and Vision* **6**(1): 89 – 108.

Railton, P. (1993). Essentially General Predicates, *in* P. French, T. Uehling & H. Wettstein (eds), *Philosophy of Science*, Vol. XVII of *Midwest Studies in Philosophy*, University of Minnesota Press, Minneapolis, pp. 166–176.

Rankine, T. (1869). *On the Practical Application of Reciprocal Figures to the Calculation of Strains on Frameworks*, Vol. XXV of *Transactions of the Royal Society of Edinburgh*.

Reisenfeld, R. (1981). Homogeneous Coordinates and Projective Planes in Computer Graphics, *IEEE Computer Graphics and Applications* **1**(1): 50 – 55.

Remmert, R. (1991). *Theory of Complex Functions*, Springer Verlag, Berlin. R. Burkel (trans.).

Remmert, R. (1998). *Classical Topics in Complex Function Theory*, Springer Verlag, Berlin. L. Kay (trans.).

Rossi, P. (2000). *Logic and the Art of Memory: the Quest for a Universal Language*, University of Chicago Press, Chicago. S. Clucas (trans.).

Rota, G.-C. (1997). *Indiscrete Thoughts*, Birkhäuser, Boston.

Salmon, W. (1989). Forty Years of Scientific Explanation, *in* P. Kitcher & W. Salmon (eds), *Scientific Explanation*, University of Minnesota Press, Minneapolis.

Scholz, E. (1989). *Symmetrie Gruppe Dualität*, Birkhauser, Basel.

Scholz, E. (1994). Graphical Statics, *in* I. Grattan-Guinness (ed.), *Companion Encyclopedia of the History and Philosophy of the Mathematical Sciences*, Vol. II, Routledge, London, pp. 987–93.

Segre, C. (1904). On Some Tendencies in Geometrical Investigations, *Bulletin of the American Mathematical Society* pp. 412 – 468. J. Young (Trans.), Italian original published 1891 in *Revista di Matematica*.

Smith, K. E., Kahanpää, L., Kekäläinen, P. & Traves, W. (2000). *An invitation to algebraic geometry*, Universitext, Springer-Verlag, New York.

Smith, K., Kahanpää, L., Kekäläinen, P. & Traves, W. (2001). *An Invitation to Algebraic Geometry*, Springer, Berlin.

Stein, H. (1988). Logos, Logic and Logistiké: Some Philosophical Remarks on the Nineteenth-Century Transformation of Mathematics, *in* W. Aspray & P. Kitcher (eds), *History and Philosophy of Modern Mathematics*, University of Minnesota Press, Minneapolis, pp. 238–259.

Steiner, M. (1978). Mathematical Explanation, *Philosophical Studies* **34**: 135 – 151.

Stillwell, J. (1993). *Classical Topology and Combinatorial Group Theory*, 2nd edn, Springer Verlag, Berlin.

Stolfi, J. (1991). *Oriented Projective Geometry: A Framework for Geometric Computations*, Academic Press, Boston.

Tappenden, J. (1995). Geometry and Generality in Frege's Philosophy of Arithmetic, *Synthése* **102**: 319 – 361.

Tappenden, J. (2001a). "Geometrical Invariance and Mathematical Practice in Frege's Foundations of Geometry". In preparation.

Tappenden, J. (2001b). "Proof Style and Understanding in Mathematics II: Notation". In preparation.

Tappenden, J. (2001c). *A Reassessment of the Mathematical Roots of Frege's Logicism I: The Riemannian Context of Frege's Foundations*. (Working title) manuscript in progress.

Timoshenko, S. (1953). *History of the Strength of Materials*, Dover Publications, New York. 1983 reprint.

Toulmin, S. (1961). *Foresight and Understanding*, Harper and Row, New York.

Truesdell, C. (1984). *An Idiot's Fugitive Essays on Science*, Springer-Verlag, New York.

Tufte, E. (1983). *The Visual Display of Quantitative Information*, Graphics Press, Cheshire, Conn.

Tufte, E. (1997). *Visual Explanations*, Graphics Press, Cheshire, Conn.

van der Waerden, B. (1991). *Algebra*, Vol. II, Springer, New York. Translated by J. Schulenberger.

Weyl, H. (1913). *Die Idee Der Riemannschen Fläche*, Teubner, Leipzig.

Weyl, H. (1932). Topologie und Abstrakte Algebra als zwei Wege mathematischen Verständnisses, *Unterrichtsblätter für Mathematik und Maturwissenschaften* **38**: 177 – 188. Gesammelte Abhandlungen vol. 3 pp. 348 - 358.

Weyl, H. (1955). *The Concept of A Riemann Surface*, Addison - Wesley, Reading, Mass. English translation of 3rd revised edition of (Weyl, 1913), G. Maclane, trans.

Weyl, H. (1995). Topology and Abstract Algebra as Two Roads of Mathematical Comprehension, *American Mathematical Monthly* pp. 453 – 460 and 646 – 651. Abe Shenitzer, trans., A translation of (Weyl, 1932).

Whiteley, W. (1985). The Projective Geometry of Rigid Frameworks, *in* C. Baker & L. Batten (eds), *Finite Geometries*, Marcel Dekker, New York.

Wilson, M. (1992). Frege: The Royal Road From Geometry, *Noûs* **26**: 149 – 180. Page references to the reprinting in (Demopoulos, 1995).

Yates, F. (1966). *The Art of Memory*, University of Chicago Press, Chicago.

Zariski, O. (1978). *Collected Papers*, Vol. III, MIT Press, Cambridge, Mass. Edited by M. Artin and B. Mazur.

Ziwet, A. (1904). *Theoretical Mechanics*, Norwood Press, Norwood, Mass.