HISTORICAL PERSPECTIVES FOR THE REFORM OF MATHEMATICS CURRICULUM: GEOMETRIC CURVE DRAWING DEVICES AND THEIR ROLE IN THE TRANSITION TO AN ALGEBRAIC DESCRIPTION OF FUNCTIONS

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

David Dennis

May 1995

HISTORICAL PERSPECTIVES FOR THE REFORM OF MATHEMATICS CURRICULUM: GEOMETRIC CURVE DRAWING DEVICES AND THEIR ROLE IN THE TRANSITION TO AN ALGEBRAIC DESCRIPTION OF FUNCTIONS

David Dennis, Ph. D.
Cornell University 1995

This thesis establishes the following three claims:

1). Experience with mechanical curve drawing devices played a role of fundamental historic and conceptual importance of in the development of analytic geometry, algebraic symbolism, calculus, and the notion of functions.

2). Two secondary students of mathematics benefited from their experiences with physical curve drawing devices, and that both the geometric and algebraic analysis of such devices raised, for them, crucial epistemic issues, the consideration of which led them to engage in a more balanced dialogue between the physical world and symbolic mathematical language.

3). A mathematical discussion of the tangents, areas, and arclengths associated with many curves need not be deferred until calculus, and that, quite the contrary, an understanding of the semiotic importance of calculus comes from seeing it as a language constructed from physical experiences. Such an understanding depends upon being able to correlate symbolic manipulations with independently verifiable geometric and physical experience. Such experience can be readily gained from the use of mechanical curve drawing devices, and from simulations of such devices using available dynamic geometry computer applications.

The first and third claims are established through a detailed analysis of the genetic epistemology of analytic geometry and calculus. This analysis proceeds from original historical documents mostly from seventeenth century European sources. Based upon historical curve drawing techniques, both physical replicas of devices and computer simulations are used to explore the properties of curves.

The second claim is established through a series of clinical interviews with two high school students. The students worked with replicas of three different physical curve drawing devices. Their conceptions, inventions, beliefs, and modes of expression are analyzed from videotapes made during these interviews.

HISTORICAL PERSPECTIVES FOR THE REFORM OF MATHEMATICS CURRICULUM: GEOMETRIC CURVE DRAWING DEVICES AND THEIR ROLE IN THE TRANSITION TO AN ALGEBRAIC DESCRIPTION OF FUNCTIONS.....i
Chapter One: Introduction 1
1.1 General Introduction 1
1.2 Curves, Graphs, Equations, Functions, and Parameters.....9
1.3 Educational Theory.....12
1.4 A Response to the Views of Some Professional Mathematicians19
1.5 The Role of History in Mathematics Education.....25
1.6 Educational Implications 31
Chapter Two: The Historical Role of Geometric Curve Drawing Devices 35
2.1 Opening Comments.....35
2.2 Apollonius and Conic Sections 42
2.3 Abscissas, Ordinates, and Functions of a Curve 59
2.4 Drawing Parabolas Using the Focus / Directrix Method 65
2.5 Drawing Ellipses from their Foci.....78
2.6 Drawing Hyperbolas from their Foci.....94
2.7 Drawing Hyperbolas from Asymptotes 105
2.8 Conchoids Generalized from Hyperbolas.....117
2.9 Drawing Ellipses with Trammels 122
2.10 Fermat's Quadratic Transformations 133
2.11 Newton's Organic Constructions.....138
2.12 Descartes' Geometric Means and a Construction of Log Curves149
2.13 Roberval: Cycloids and Sine Curves 167
2.14 Pascal and Leibniz: Sines, Circles, and Transmutations 177
2.15 Summary Remarks.....198
Chapter Three: Interviews with Students Exploring Elliptic Devices.....203
3.1 The Content and Purpose of the Interviews 203
3.2 The Structure and Method of the Interviews.....209
3.3 A General Comparison of Two Student Approaches.....213
3.4 Tom's First Interview.....218
3.5 Tom's Second Interview 232
3.6 Jim's First Interview 251
3.7 Jim's Second Interview 268
3.8 A Comparison of the Methods and Epistemology of Tom and Jim.....296
Chapter Four: Summary and Conclusions.....304
References 312

Chapter One: Introduction

1.1 General Introduction

One of the fundamental goals of secondary and undergraduate mathematics education is for students to develop an understanding of the concept of a function. If one discusses this idea with teachers and students, it is not so clear what this goal entails. Many textbooks contain in their opening sections a definition of the word "function" which is given in the language of sets (i.e. a special type of relation where every member of the domain is paired or associated with exactly one member of the range). This general definition usually disappears quickly from subsequent mathematical activities, suggesting that it is not useful within the context in which students usually find themselves. Perhaps it is not the appropriate conceptual tool with which students work in analytic geometry and calculus. When students are asked what a "function" is, they often reply that it is some kind of equation, involving x and y, which can be graphed. This answer better reflects their experience in mathematics classes where the emphasis is on equations which express real valued functions of one real variable. This is reflected in some textbooks when they discuss functions as numerical machines with inputs and outputs. But is a function the equation? or the graph? or an algorithm? or something else?

Students spend a lot of time learning to see the connections between an equation and its graph. From the beginning of analytic geometry in high school through several courses in college calculus, this remains one of the prevalent themes in mathematical education. Learning to move flexibly between equations and their graphs is seen as one of the most fruitful abilities that mathematical education can offer. Students are encouraged by teachers to perceive the existence of a dialogue between the algebraic structure of an equation and the shape of its graph. Many recent educational reform efforts (e.g. N.C.T.M. Standards, 1991) have suggested that a greater effort be placed on qualitative visualization skills that emphasize what can be deduced from or about the visual shape of a graph.

The presentation of this dialogue in our current curriculum is often unbalanced. Consider the phrase: "an equation and its graph." Implied here is that graphs are created by equations, and indeed this is usual experience of students in mathematics. For example, many current reform efforts stress the use of graphing calculators. This tool creates an environment where curves can exist only secondarily, as representations of algebraic statements or numerical data. The existence of algebraic equations is privileged over that of geometrically produced curves. The substitution of the word "graph" for the word "curve" implies that a coordinate system must exist before a curve can exist, which gives primacy to algebraic expression. In mathematics classrooms curves are usually created from algebraic equations or numerical data, and only rarely by physical or geometric actions (lines and circles being the only common exceptions).

There are many recent educational materials that attempt to broaden the dialogue between curves and numerical relationships. Many have focused on graphs of data that come from physical experiences (e.g. Confrey, 1994c; Hilbert, et al, 1994; Monk, 1992). This stronger use of contextual problems ties mathematics more strongly to important scientific phenomena. These recent reform efforts have also looked at the role of qualitative graphing concepts in the scientific thinking of students. These reform efforts are useful and important, but curves are still being created in reference to a pre-established coordinate grid.

There is a ordered progression of activities in both our current curriculum, as well as many of the reforms suggested by the use of experimental data and qualitative graphing.[1] In the usual curriculum that is, first the equation, second the coordinate grid, and last the plotting of points as a graph. In the reform curriculum, one often starts, instead, with a table of numerical data and then proceeds to a coordinate grid and then to a graph. Things come into existence in that order before any dialogue can begin. I want to claim that this important process could become more powerful in our classrooms if it were allowed to reverse itself and flow fully in both directions, thereby creating a more complete balance in the dialogue between curves and equations. Analytic geometry could be presented as a feedback loop between the tactile, geometric world of curves and the semiotic world of algebra and numerical data. Such a feedback loop could avoid hierarchy altogether (von Glasersfeld, 1978; 1990). The concept of a function could be seen as a cognitive process that moves flexibly between multiple representations none of which is given special status or definitional preeminence. Confrey (1992; Confrey & Smith, 1991) calls this an "epistemology of multiple representations." [2]

The question that then emerges is how can we construct the missing half of this feedback loop? What tools, activities and concepts would be required in order to create a mathematical environment where curves have a primary existence, and coordinates and equations come into existence as secondary facilitators. To investigate this I have looked carefully at the historical genesis of analytic geometry. For example, the most famous original source on the subject is René Descartes' *Geometry* (1952), which was first published in 1637. Never once did Descartes write down an equation and plot its graph. His book is devoted to generating curves with devices (usually mechanical linkages), and then imposing a relevant coordinate system, and lastly finding an

---

[1] Some educational researchers have used explorations of qualitative graphing in ways that begin to allow for the discussion of curves that have been created prior to numerical or algebraic symbolism (e.g. Rubin & Nemirovsky, 1991).

[2] It should be noted that Confrey emphasizes the role of tables as well as graphs and equations in this epistemology.

equation. Algebraic equations are used to create a taxonomy of curves, hence the name "analytic geometry." For Descartes, algebraic equations were codes for geometric constructions and actions which produce curves. Throughout his book curves have a primary existence prior to the imposition of any coordinate system. An arbitrary algebraic equation, which is not connected to a curve drawing device, has no epistemic significance for Descartes (Lenoir, 1979).

The main goal of this thesis is to indicate specific ways in which curve drawing devices and activities could be used to create a fruitful educational environment where profound connections between geometry and algebra could be explored by students without isolating or privileging either. My first arguments will be based on a combined investigation of conceptual historical genesis and a "rational reconstruction" (Lakatos, 1976) of historic ideas facilitated by dynamic geometry computer applications. These will be followed by an analysis of students working first with curve drawing tools and arriving at algebraic equations only secondarily.

Descartes' development of analytic geometry was as unbalanced as our modern curriculum, but in the opposite direction. During most of the seventeenth century, ancient traditions persisted that viewed geometry as the dominant form of mathematical thought. Algebra was considered secondary at best, and by some entirely untrustworthy (Cajori, 1929). In our own times this dominance relation has been reversed, especially in American public schools where relatively little time is spent on geometry, and what geometry is taught is mostly formal and isolated from other topics.

Curve drawing devices as a bridge between geometry and algebra have almost disappeared from our mathematical classrooms despite, or perhaps because of, their compelling physicality. The immense value placed by Western Culture on disembodied, decontextualized, abstraction has led to a general trend over this entire century to separate mathematics even from its roots in engineering and science, and certainly from its social and historical genesis. The notion of a "pure mathematics" divorced from its physical and empirical origins, and separate from "applied mathematics" indicates a disturbing lack of communication. This inability to communicate has troubling repercussions in education. It represents an inability to discuss mathematics as a cultural tool with a social historical genesis rooted in problems and activities.

This trend has been repeatedly challenged by those outside of professional mathematics such as educators and scientists (e.g. Confrey, 1993a). It has also been occaisionaly challenged by working mathematicians; for example, by Richard Courant (1888 - 1972), who made extensive use of physical phenomena, like soap bubbles, to teach advanced subjects like the calculus of variations. Courant, writing about the famous mathematician, Carl F. Gauss (1777-1855), said:

> He (Gauss) was never aware of any contrast, not even of a slight line of demarcation, between pure theory and applications. His mind wandered from practical applications, undaunted by required compromise, to purest theoretical abstraction and back, inspiring and inspired at both ends. In light of Gauss' example, the chasm which was to open in a later period between pure and applied mathematics appears as a symbol of limited human capacity. For us today, as we suffocate in specialization, the phenomenon of Gauss serves as an exhortation. The representatives of both camps should not be proud of their limitations, but should do everything to understand each other and to bridge the chasm. In my judgment, it is critical for the future of our science that mathematicians adopt this course, both in research and in education. (Courant, 1984, p 132).

Another leading figure in modern mathematics, Jon von Neumann (1903 -1957), was also troubled by this separation of pure mathematics from empirical and physical investigations. Von Neumann argued that when the formal procedures of mathematical discourse have become far removed from their empirical origins that:

> there is a grave danger that the subject will develop along the lines of least resistance, that the stream, so far from its source, will separate into a multitude of insignificant branches, and that the discipline will become a disorganized mass of details and complexities. In other words, at a great distance from its empirical source, or after much "abstract" inbreeding, a mathematical subject is in danger of degeneration..... Whenever this stage is reached, the only remedy seems to me to be the rejuvenating return to the source: the reinjection of more or less directly empirical ideas. (1984, p.234).

The views of Courant and von Neumann and others like them have gone largely unheeded in most classrooms, especially when it comes to analytic geometry and calculus. Certain iconic "word problems" are repeated in these courses, but seldom are physical activities profoundly linked to the development of formal language. Often attempts at "applied problems" have become subject to formalization and "abstract inbreeding," and have often been stripped of any strong physical sense that would bring them to life (see section 2.9 for a specific example).

The most powerful and lasting mathematical achievement of seventeenth century Europe was to establish the possibility of representing many aspects of physical geometry with the symbolic language of algebra. That this can be done with some degree of consistency is a fundamental mathematical belief upon which modern science is constructed. The work of Leibniz and Newton at the end of the century was a compelling demonstration of the extent to which physical geometry and semiotic algebra are compatible. That is to say, when one can construct two different ways of arriving at, say a certain ratio or area, then the two different approaches will agree.

Belief in this consistency did not come easily (Cajori, 1929), and yet is casually assumed in our mathematics classrooms without questions or demonstrations of either a physical, symbolic, or intellectual type. The mathematical language that was constructed in the seventeenth century (i.e. analytic geometry and calculus) remains, for most students, the dominant core of their secondary and undergraduate mathematical curriculum, and yet students are rarely put in an environment where they can either doubt or confirm the startling claim that geometry and algebra can express compatible things in different settings.

The first principle of Cartesian philosophy is that all well founded intellectual beliefs begin with profound doubt (Descartes, 1989). In a larger philosophical setting, the goal of this thesis is to suggest how students could be given a chance to experience that doubt through the interaction of physical and linguistic activities, and then go on to transcend that doubt; moving on to an experience of conviction. A variety of physical curve constructions, often taken directly from seventeenth century texts, will be analyzed. The material will be discussed in its original cultural setting, and will also be looked at in light of modern computer technology with the dynamic animation capacity that is readily available in high schools (e.g. the software, *Geometer's Sketchpad*, Jackiw, 1994).

I will not suggest that computer animations can take the place of physical experience. They can, however, provide rapid variation and extrapolation of experiments conducted by students who have first had some experience with actual curve drawing devices. Computer animations and simulations can provide intermediary environments which help to set up a dialogue between the physical experience of geometry and the symbolisms of algebra (Confrey, 1993a). I will discuss specific examples of how these intermediary tools could facilitate the multidirectional flow of the cognitive feedback loop which I see as analytic geometry.

This main content of this work will be organized into two parts as follows. Chapter 2 will examine in detail a variety of geometric curve drawing devices and their crucial role in the creation of analytic geometry and calculus. These devices will be explored and extrapolated with the use of *Geometer's Sketchpad*. There are many precisely drawn figures and curves in Chapter 2, but none were created by inputting an equation and plotting a graph. All of the figures were created by geometric actions. Chapter 3 will look, via clinical interviews, at the techniques, attitudes, and beliefs of two high school students as they confront three very different physical devices, which all draw ellipses. Thus I will examine the historical genesis of a broad set of mathematical ideas and then build, as an example, one specific environment where individual students can begin to experiment with different perspectives.

In summary the goals of this thesis are as follows:

1). To establish the fundamental historic and conceptual importance of curve drawing devices in the development of analytic geometry, algebraic symbolism, calculus and the notion of functions.

2). To show how two secondary students of mathematics benefited from their experiences with physical curve drawing devices, and that both the geometric and algebraic analysis of these devices raised, for them, crucial epistemic issues, the consideration of which led them to engage in a more balanced dialogue between the physical world and symbolic languages.

3). To show that a discussion of the tangents, areas, and arclengths associated with many curves need not be deferred until calculus, and that, quite the contrary, an understanding of the semiotic importance of calculus depends upon being able to correlate its symbolisms with independently verifiable geometric experience. Such experience can be readily gained from the use of physical curve drawing devices, and from simulations of such devices using available dynamic geometry computer applications.

1.2 Curves, Graphs, Equations, Functions, and Parameters

Section 1.1 opened with some questions concerning the definition of "function," and went on to raise certain issues about curves, graphs, and equations. One might ask what definitions will be used in this work. I want to say right from the start that no particular set of clear and precise logical definitions will be educationally advocated. Both the historical and the student investigations will indicate a variety of useful conceptions, but no fixed definitions will emerge as capable of encompassing all of these useful conceptions. Mathematical flexibility will demand a certain ambiguity when facing both the intractability of the material world, and the intractibility of human cognition to perceive the material world (von Glasersfeld, 1984).

To be as clear as possible, the term "graph" will be used to denote a curve that has been created from numerical information by plotting points in a pre-established coordinate grid. The more general term "curve" will be used to denote a path in the plane that can be traced by some sort of reproducible action. Most of the examples of curves discussed here can be drawn with some sort of mechanical device, although other methods, such as paper folding, will also be mentioned. I will not discuss random tracings, but I will discuss computer animations made from geometric inputs which simulate mechanical devices.

The general issue of how algebraic symbols will be viewed, both in historical and educational terms, will be discussed repeatedly in varying contexts. One main intention will be to use mathematical history to create a broad and flexible notion of how language evolved in response to activities and experience. The results of experimental activities were well known before symbols could be formulated capable of expressing those results. It is this sense of linguistic response to known experiences that I want to see brought into mathematics education. This is essential to a constructivist philosophy of education (see Section 1.3).

The issue of the difference between equations and functions must be seen in light of this broader epistemic issue. If physical activity is to play a larger role in the educational development of, say, analytic geometry then a certain ambiguity will be unavoidable. A strict definition of "function" will never entirely serve the needs of flexibility within a cognitive feedback loop. Historically, a broad concept of "function" evolved in response to many activities. In the classroom, this evolution must be allowed to take place in such a way that a variety of concepts, rooted in activities, are allowed to fruitfully coexist. It is the progressive elimination of both prior activities and prior concepts that curtails our curriculum and causes the "suffocation in specialization" that Courant warned against (1984).

For example, the geometric definition of "functions of a curve" given by Leibniz (see Section 2.3) is still a very useful tool for investigation, even while it tends to blur the difference between equations and functions. It is an excellent example of a lost phase of history that can provide a powerful conceptual approach to curves. It yields simple and elegant results in analytic geometry that are usually delayed until calculus, and it is particularly appropriate within the environment of dynamic geometry computer applications.

As one experiences different functional concepts, the use of specific "parameters" can help to untangle some of the ambiguities, but only after one has been immersed in the kind of curve drawing activities that gave rise to Leibniz's geometric definition in the first place. As the historic material in Chapter 2 is presented, at first parameters, such as arc length or time, and parametric equations are avoided, but as the material from the seventeenth century unfolds certain parameters come to play important central roles. Parameters demand to be considered if certain actions and activities are to be understood and analyzed (Sections 2.12, 2.13, and 2.14), but when one is just beginning to think about physical curve generating activities, they can sometimes be a needless burden. Both of the students in Chapter 3 mention possible parametric approaches to the devices they are investigating, although they eventually abandon them. The original use of the term "function" by Leibniz (Sec. 2.3) was purely geometric and did not involve parameters, although his usage can not be formally reconciled with any modern definition of function without the use of parameters. Thus, although several parts of both Chapters 2 and 3 suggest that parametric equations may be a very natural educational approach to functions in a geometric context, this will not be a major topic of investigation in this work

Many modern curricular reform agendas (e.g. N.C.T.M., 1991) enshrine the notion of "function" as central in the secondary curriculum, thus making functions a kind of umbrella concept under which to interpret large sections of mathematics. One regrettable consequence of this approach is that the only curves that enter into the curriculum are graphs, i.e. curves are presented only as secondary representations of data or equations. In both the historic material of Chapter 2, and the student interviews of Chapter 3, there emerges a notion of algebraic equations as secondary representations of curve drawing actions. An algebraic description of a function is clearly seen as a linguistic response to a dynamic curve drawing action. This complete feedback loop, with all of its attendant ambiguities, is what I will educationally advocate.

The central curricular role of the function concept will not be challenged here, provided that concept can serve dual, over and under, purposes. That is to say that the function concept can serve both as a unifying general concept, as well a set of specific tools for the unpacking of dynamic activities. In particular, it is important that functions can both represent and be represented by curves.

1.3 Educational Theory

My opening remarks have hinted at some of my theoretical and epistemological stances. I now wish to make these stances specific. I am attempting to apply the emerging theory of Jere Confrey as described in her recent paper "The role of technology in reconceptualizing functions and algebra," (1993a). For many years Confrey's research has proceeded using a radical constructivist philosophy, based originally on the ideas of Piaget and von Glasersfeld. Her work centered on careful listening to students and the design of software and contextual problems which were tailored to foster a variety of perspectives and multiple representations. In order to design problems and software which would allow students maximal flexibility of expression, she often investigated the history of mathematics in order to provide a broader framework with which to describe, examine, and legitimate student conceptions. Her earlier work viewed mathematical development from a largely constructivist, Piagetian framework. The construction of knowledge by the student was seen as a cyclic interaction between personal reflection, problematic situations, and the actions and tools available (Confrey, 1993b). Others were seen as important sources of interactions via communications and the posing of questions.

More recently Confrey's work includes some of the developmental theories of Vygotsky. These theories stress the importance of social interaction and cultural context in the construction of knowledge. Vygotsky worked within a Marxist tradition that emphasized the importance of tools in the transformation of labor. He extended the notion of tools beyond the physical to include language, symbols, and any semiotic environment in order to arrive at one of his most fundamental principles, i.e. that tools mediate knowledge. Vygotsky (1962) viewed thought and language as an Hegalian dialectic from which knowledge emerges. Confrey's work (1993a) links the theories of Vygotsky with the more modern radical constructivist work by viewing mathematics as formed from a dialectic between "grounded activity" and "systematic inquiry" (1993a, p. 51). Her dialectical structure is mediated by tools, both physical and semiotic.

Confrey (1993a) criticizes Vygotsky's writings for losing the dialectic between thought and language and ultimately privileging symbolic abstractions over what is learned through physical activity. Piagetian research has shown the rich and varied ways in which children construct knowledge by adapting to physical as well as semiotic environments (Confrey, 1993b; von Glasersfeld, 1982). Confrey's new dialectical structure gives equal roles to grounded activity and systematic inquiry. She views mathematics as comprised of an equal balance between physical, empirical activities and symbolic, linguistic investigations.

This perspective has led her to challenge the portrayal of abstraction in both Vygotsky and Piaget, and to question their assumptions about the necessity of detachment from bodily activity (Confrey, 1994e). She points out that these assumptions make it difficult to acknowledge the abstract understandings of expert carpenters, mechanics, engineers, and architects whose expertise lies in a unification of both mental and physical dexterity (Milroy, 1990).[3]

Like Vygotsky, Confrey wishes to focus on mathematics as kind of tool, however, she criticizes Vygotsky for having moved too completely to a focus on psychological tools as he viewed language or semiotic systems, and thus neglecting the interplay between physical and symbolic tools. As a result the focus on advanced technologies in today's schools tends to be on what they display rather than on how they work. The role of physical investigation and context in creating appropriate challenges and invitations for student thinking are lacking.

---

[3] During the seventeenth century, for example, a "geometer" was someone whose primary job was the design of fortifications, siege engines, canals, water systems, etc. In Descartes time this is what was taught in a University course on "Geometry."

In their historical work, Confrey and her associates used epistemological arguments to examine and critique standard historical descriptions of mathematics. In doing so, they sought out instances in which there were uses of multiple representations(e.g. Confrey & Smith, 1995; Dennis & Confrey, 1993; Smith, Dennis & Confry 1992). They used the study of the genesis of ideas to critique the practices of mathematicians that contribute to the mystification of the subject. Chapter 2 of this work will provide an extensive example of an historical investigation undertaken from this stance.

In their empirical work, Confrey and her associates have demonstrated how an unduly narrow definition of mathematics has led to the suppression of diversity in student methods. They have argued that the dominance of symbolic algebra at the secondary level has contributed to the alienation and discouragement of many students towards pursuing mathematics. In contrast, they have documented how may views suggested by students go unrecognized by teachers, but can lead to robust and epistemologically challenging mathematics. Chapter 3 of this work will provide examples of such student thinking in relation to the curve drawing tools.

Part of the epistemology implied by this dialectical view has been called by Confrey an "epistemology of multiple representations" (Confrey & Smith, 1991). This philosophy sees the construction of mathematical language as attaining its viability through the process of checking back and forth between different representations in order to solve interesting problems. For example, one might gather data from a geometrical situation and then enter numbers in a table. After manipulating and expanding table entries according to what makes sense within that representation, one then checks back to see whether these results have consistent geometrical implications. The focus is on recognizing the integrity of each representation's ability to contribute insight into the problem. It means acknowledging the differences as well as the commonalties in those contributions. (For an example of how students engage in this process see Confrey, 1993b; for an historical example see Dennis & Confrey, 1993.)

The curriculum and software developed by Confrey and her associates has centered on three representational forms; those being tables, equations, and graphs (Confrey, 1994b; Confrey & Smith, 1991). This work accomplished much towards providing students with a more flexible and diverse approach to functions and problem solving. Important in this work was the emphasis upon equal and independent status for each representation.[4] This thesis will suggest the use of curves as another representation of functions as well as the use of functions themselves as representations of curves. I have already indicated in my opening comments the difference between graphs and curves, and I will discuss, in Chapter 2, an abundance of examples. I will describe a variety of possible tools and activities which could allow students to work within a world of curves and to correlate those experiences with other representational forms. These tools will include both physical devices and computer simulations.

The direct mechanical generation of curves provides a form of grounded activity which can lead in a variety of profound directions. Using only hinged rods (linkages) one can draw any algebraic curve[5] (Artobelevskii, 1964). Linkages can also be used to generate arbitrarily dense sets of points on transcendental curves (e.g. log curves, see Section 2.12). If rolling wheels are included one can draw cycloids, sine curves and a host of other transcendental curves (see section 2.13). The tools involved are immediately physical and comprehensible in a way that many tools in science often are not. Even though the resulting action can be surprising, all the parts of a mechanical

---

[4] For example, her software Function Probe allows for the manipulation of graphs in terms of stretch, reflection, and translation, directly through mouse actions without requiring an algebraic manipulation.

[5] Algebraic curves are those whose coordinates can be represented by an equation of finite algebraic degree, as opposed to curves like cycloids and exponentials which require an infinite series for a representation of their coordinates.

curve drawing device are open to inspection. Unlike the mysterious electronics of computers, the actions that I will discuss are physically apparent and demonstrable to students.

From a direct analysis of the action used to draw a curve one can often determine an equation for the curve, after the fact, using geometric similarity. Confrey's work (1994a; 1988) with elementary school students indicates the need for a much stronger sense of ratio that is more tied to geometric similarity. In her splitting conjecture, she argued for the independence of the roots of splitting (multiplicative structures), and counting (additive structures), placing the roots of splitting in sharing, geometric similarity, and the iteration of these processes. These provide the precursors to ratio and proportion. The activities that I will describe will suggest some ways that these reforms could be extended to the secondary curriculum, and connected to a diverse, multi-representational view of functions. In fact, the idea for this thesis germinated when I was working on a research assignment, for Jere Confrey, to look for examples in mathematical history which displayed greater use of geometric similarity. This led me to an investigation of Descartes' hyperbolic device (see Section 2.7), which in turn led to me to a more detailed study of curve drawing devices.

Even before algebraic equations are found, one can often determine tangent lines, areas between curves, and arclengths of curves, all from an analysis of the actions which produced the curves.[6] Historically as the languages and notations of algebra and calculus were developed, they needed to be tested for viability against independently established examples. Curve drawing devices often provided the crucial tests for the viability of these new linguistic tools; they provided the critical experiments. It is educationally important to present this sense of mathematical

---

[6] If no coordinate system or unit of measure have been imposed, one can still discuss the ratios of areas and arclengths in a purely geometric setting. This will be seen in many of the historical examples presented in Chapter 2 (e.g. Section 2.13).

language and notation as a culturally viable tool that can be tested empirically against grounded activity. Tests for viability are more profound when they occur across as diverse a set of representations as possible.

The examples that I shall present concern the situation where geometry moves, and curves are created. I shall present material which will broaden the diversity of possible representations in the area of analytic geometry and calculus. This conceptual material is historically important, but even without a knowledge of history it opens a series of alternative pathways from which to approach analytic geometry and calculus. I believe that it could be helpful in promoting lively conjectural discussions in secondary classrooms, making it easier for teachers to listen to students in a constructivist way (Confrey, 1994a). An advantage of the curve drawing devices is that they present familiar material in a different way, and they easily lead to challenging open questions, especially when the generation of curves is extrapolated using computer animations (e.g. Section 2.11).

The students, whose interviews I will discuss in Chapter 3, were not told anything about the history of the devices with which they worked, until after they had finished with their own activities and analysis. The apparatus itself (e.g. Figure 2.9a) placed them in an alternative representation (dynamic geometry), and they had to find their own ways of connecting this with a more familiar representation (e.g. algebra). Both of these students experienced a strong sense of satisfaction when they achieved what, to them, was a satisfactory cross-representational consistency, although they differed in their relative faiths in each separate representation; one relied more on algebra, the other more on geometry.

The student interviews presented in Chapter 3, display both the diversity of student conceptions, and the power of tools which allow for multiple representations. Together with the historical analysis of Chapter 2, these interviews clearly illustrate the main tenets of Confrey's educational theory. By combining the historical analysis with the possibilities created by newly developed computer applications, these approaches and ideas suggest profound directions for broad educational reform.

1.4 A Response to the Views of Some Professional Mathematicians

Confrey states, "The contributions of Piaget and the constructivist program have documented time and time again that the formal presentation of mathematics by mathematicians forms an inadequate basis for engaging students in the process of learning mathematics" (1993b, p.307). Formal notational hierarchies as the main educational format in mathematics has lead to a "deafening silence" that pervades our secondary classrooms (Confrey, 1993b, p.305). The historical material that I will present is intended for educational use within an epistemology of multiple representations, yet how will mathematicians respond to such a philosophy in the classroom?

Many mathematicians have experienced mostly formal mathematics, and may know nothing about history, but I want to briefly recount a discussion that I had with Anil Nerode whose knowledge of both mathematics and history is vast. He objected to the use in the classroom of an "epistemology of multiple representations" saying that it was an anachronism which would "throw students back into the days before the Weierstrass revolution." He said that of course that was how Pascal, Wallis, Newton, Leibniz, and Euler proceeded, "all those guys were in the same boat." Nerode meant that an epistemology of multiple representations was the only way that these early mathematicians had to justify their new techniques, language, and notations, whereas the nineteenth century found new and more sophisticated language which added a whole new level of certainty and agreement to analytic geometry and calculus. This movement culminated in the work of Karl Weierstrass (1815 - 1897) and led directly to the formalisms of our time (e.g. the format of a "modern analysis" course).

I feel compelled to answer the objections of Nerode and others who have voiced similar opinions to me. I am not presenting a new series of proofs. Although much of the historical material in Chapter 2 could be presented as proofs that would satisfy most mathematicians, my goal is to display a heuristic process of experimentation which is necessary for the validation of any formal linguistic development. In Sections 2.12 and 2.13, I have presented some material that would be very difficult to formalize. Those sections especially may draw from some the reaction that "oh god, that's exactly the kind of thinking that we've got to keep out of calculus, that stuff leads to contradictions." That is to say that some mathematicians may strongly object to exposing students to the kind of empirical and heuristic arguments that were used in the seventeenth and eighteenth centuries, because such methods can lead to paradoxes when a larger repertoire of functional definitions is later introduced (e.g. definition via infinite series). I remain, however, firmly committed to an educational epistemology in mathematics that is both empirical, and finds its viability in checks between multiple representations. I shall justify this commitment in several ways.

Piaget and Vygotsky both espouse forms of "genetic epistemology" which compels educators to examine the historical, social, and cultural genesis of all knowledge (Confrey, 1993a). The mathematics that dominates our secondary curriculum stems almost entirely from the empirical science of the seventeenth century. This is no cultural accident since capitalism and the modern state both originated in the same period. A social history of the genesis of science and mathematics would explain a lot about our educational values, but that is beyond the scope of this work. I wish to state clearly, however, that I am not making an argument for empirical curve drawing from a belief that students must relive history in their educational development (i.e. that phylogeny recapitulates ontogeny). Certainly it is possible to learn a great deal of mathematics with no knowledge whatsoever of curve drawing devices. I was trained that way, and so were most of the current generation of mathematicians. I was able to learn the mathematics of the Weierstrass revolution without experiencing any of the actions or problems that occupied Descartes or Pascal, but many others were not so fortunate and I always had a suspicion that I was missing something.

Many secondary teachers of mathematics have only the vaguest notion of what the Weierstrass revolution is about. They often remember some painful course called "Analysis" or "Advanced Calculus" that they were required to take in college and have never used since then. One lasting result of these required courses seems to be that nearly all questions concerning tangents, curved areas, and arclengths (except for the circle) are avoided completely in the precalculus curriculum. The language of these college analysis courses has made such questions seem far more difficult than they are. These topics can then strike fear in the hearts of teachers and any questions from students concerning, say tangents, are often answered with "You'll find out about that when you get to calculus."

These teachers can not comprehend the language of Weierstrass nor the need for it, and yet they are also cut off from the earlier empirical epistemology which could provide them with satisfying answers and activities, which they could share and discuss with their students. Secondary teachers need better education, but that is not usually to be had from taking more traditional courses in a college mathematics department. The next chapter will provide many compelling examples of what I feel is needed. I will show examples of curves, tangents, areas, and arclengths, that could be discussed with secondary students at a very early stage. If such activities were developed, these students could then later view many of the early results of their calculus classes as linguistic reformulations of things that they had already experienced. The language of calculus would then, from the very beginning, be about something with which they already had some experience.

From a constructivist point of view, one must ask what problematic situation created the need for the Weierstrass revolution in language. I will not discuss this in detail, but, from almost any historical view, the problems and contradictions that directed the mathematical development of the late nineteenth century occur within a context that is far removed from that of a secondary classroom, or even a college calculus course. Euler, working in the eighteenth century within an epistemology of multiple representations, achieved very sophisticated results in analysis and differential equations that are the basis of much of the advanced engineering mathematics of today (Euler, 1988; Boyer, 1968). His empirical approach to convergence did occasionally lead to paradoxes, but Euler's attitude (1988) reminds me of an old joke. The patient says "doctor it hurts when I do this." The doctor responds, "don't do that." That is to say, he continued to experiment empirically until he arrived at a method which was meaningful within the contextual circle in which he worked.

Chapter 2 will indicate ways that many curves and their tangents and areas can be profoundly discussed by experiencing the actions that produce the curves and analyzing them using only the most basic language of geometry. Notions of tangency, area and convergence can be treated without formal definitions (see Section 2.2.). This material could be expanded in ways that could be useful in both teacher training courses and in secondary classrooms. With the aid of simple mechanical devices and computer software like *Geometer's Sketchpad*, these important historical topics could become standard curricular reforms that speak to readily available physical experiences, and lead to profound mathematics within an epistemology of multiple representations.

I do not mean to disparage the linguistic achievements of the Weierstrass revolution. This movement in mathematics certainly did create a subtlety of language in analysis that put an end to a variety of disturbing paradoxes and controversies (Boyer, 1968), but the examples that gave rise to these controversies can only be constructed using methods (usually infinite series) that are beyond the scope of most secondary classrooms, and certainly do not arise in mechanical curve drawing activity. The Weierstrass revolution is a linguistic response to a set of problematic activities that emerge in a setting far removed from the set of activities that gave rise to analytic geometry and basic calculus. As Lakatos (1976) might say, the monsters which lurked on the edges of calculus in the nineteenth century were not at all of the same breed as those which prompted the linguistic formulations of Descartes or Leibniz. Educationally, linguistic formulations must be made in response to the problematic activities experienced by the students. Symbolic language must earn its validity directly from student experiences.

The Weierstrass movement, towards a more formal mathematics, can still never achieve the absolute certainty desired by its originators. The Gödel theorem has shown the twentieth century that that is impossible once infinity is admitted into mathematics, as it must be if anything like classical geometry is to continue. Those who desire a formal hierarchical mathematics with certainty must confine themselves to finite mathematics as has been suggested by von Neumann and others in more recent times.

I do not think our society ought to restrict itself to finite mathematics, despite the allure of certainty, and the achievements of modern computer languages and logic. Geometry, with all of its infinite notions and their attendant uncertainties, is a cultural icon which most mathematicians want to retain. In this sense, the Weierstrass revolution can be viewed as another linguistic representation in an expanded epistemology of multiple representations. No matter what its level of logical subtlety, can anyone believe that this form of mathematics would ever have gained acceptance if it was not consistent with earlier forms of empirically established geometry? I hope that the reader will recall this question when he/she reads Christopher Wren's determination of the arc length of a cycloid (section 2.13). Arguments are one thing, but does the answer come out within reasonable bounds of empirical measurement? If it does not, then history shows us that it is the language that must be readjusted.

Another parallel comes to mind. Consider the formalization of logic by Frege, Russell, and others in the early part of this century. People did not need Russell's axioms to know how to make a logical argument. This had been known for centuries. Russell's formalizations did more to show what logic could not do, than what it could.

All of this is only interesting to someone who has had some direct experience with the making of logical arguments. The informal empirical methods of curve drawing and tangent finding will help students to see what they can accomplish using simple tools together with the linguistic methods that are appropriate to those experiences. The positive results of such investigations should precede the restrictive logical analysis of more formal methods.

1.5 The Role of History in Mathematics Education

Before embarking on my own interpretive view of some mathematical history, I wish to comment generally on the possible uses that history could have for mathematics education. What kind of investigations are desirable? Where and how should they be presented and discussed? What sort of curriculum reforms can history inspire? What kind of history, if any, should be presented directly to secondary students? or to teacher candidates? What part should history play in educational philosophy and epistemology? Recently various articles and curriculum materials have appeared that take very different views, and, before explaining my own stance, I would like to comment on a few of these.

First of all there is anecdotal, largely biographical history that some textbooks and teachers like to include to "spark the students interest" or "give mathematics a human face." The calculus textbook by Howard Anton (1988) is one of the better recent examples of this. I feel that such historical material is at best trivial and ineffective, and can sometimes be conceptually misleading. The actual mathematics that is being presented is entirely unaffected, and no historical problems or conceptions are discussed. Quite often, this sort of history only serves to perpetrate certain cultural mythologies. For example, pictures of Newton and biographical scraps often surround modern material and notation that comes mostly from Leibniz and Euler. Within such a flurry of anglophilia, the actual work and conceptions of Newton are almost never discussed, because they either depended on an intense involvement with geometry (see, for example, section 2.11), or because they depended on an empirical use of tables and interpolation which are now cloaked as algebra and analysis (Dennis & Confrey, 1993). With no sense of original conceptions, this type of history can leave students with the debilitating impression that mathematics sprang, complete and god-like, from the heads of divine icons.

Let me turn to some recent, more serious, educational attempts to utilize history in an attempt to create a theoretical perspective for the development of the function concept. Sfard (1992) discusses how students may or may not come to see functions as objects rather than processes, actions or algorithms. She is concerned with how students move from an operational to a structural concept of function. Sfard compares this process in students to historical developments in mathematics. I would claim that her view of both of these processes accepts a hierarchical view of mathematics, and promotes a progressive absolutist view of history (Confrey, 1980). That is to say, that history is seen only as a means to illuminate the march of progress towards the inevitable higher level of current mathematical practice.

Within this view of history Sfard thoughtfully discusses the difficulties that students encounter in the process of reification (from Latin res=thing). She shows that students can not regard functions as single objects (e.g. as points in a function space) unless they have reached a problematic level where such a view is required, and she parallels these discussions with historical controversies about the meaning of variables. Sfard uses history to stress that what may seem, to a modern mathematician, as a simple act of definition was once a source of considerable controversy. Her discussions about appropriate levels of functional language for students are compared to the historical formalization of analysis in the period after Weierstrass.

I find such a view of history unacceptable, and therefore I reject its educational implications. The historical march towards now is but one of many stories, and the fact that that one story has become the most dominant, is not due to its logical inevitability. If one forces mathematics and its history into such a hierarchy, a great many useful and beautiful conceptions are lost. I see little reason why education should set as its goal to push students as rapidly as possible towards reification of the function concept. I seek in history the broadest possible diversity of actions, and I want that conceptual diversity to blossom in the classroom. I hope that the examples that I present will display the otherness of history. Newer and more compact notation may streamline certain thoughts, but will also obscure other useful and beautiful views. Educationally valuable conceptions have been eradicated from the classroom by narrow views of history like the one taken by Sfard (1992).

Sierpinska (1992), like Sfard, uses mathematical history to investigate how students come to understand the notion of function, but she takes a somewhat broader view of history than does Sfard. She briefly describes a variety of historical conceptions of functions and gives some details from original sources. I would again claim, however, that her overall theory of history and its relations to education, remain progressive absolutist. For example, geometric actions which generate curves (the entire subject of this thesis) is mentioned by her, but it is seen as an epistemological obstacle. Although she initially states that such obstacles are not entirely negative, she states that "if we want to understand further, and <u>better</u> ..... we have to act against them" (1992, p.28). The tone of Sierpinska's writing directly implies that curve drawing was a tedious phase of development that mathematicians in the seventeenth century had to go through in order to progress to the more advanced levels that led to modern functional concepts. She concludes her discussion of "Functions and Curves" with the statement "a graph represents the function in an indirect, symbolic way" (1992, p.52). She thinks that one goal of education is to get away from the idea of a feedback loop where functions could also represent curves.

There is no indication in Sierpinska's article (1992) that by abandoning curve drawing something valuable might have been lost. History is not seen by her as source of conceptual diversity, but almost as a set of pitfalls to be avoided or overcome. She suggests that some sense of history can be useful in helping students to overcome these possible obstacles, but there is no indication that student investigations within a given historical conception might offer valuable insights that are obscured by modern conventions. At the end of her article she suggests that students be given a "broad spectrum of ways of giving functions" (1992, p.57), but the flexibility that she suggests is the usual narrow set of modern notations and graphs. Curve drawing and its attendant "functions of a curve" are mentioned by her (1992, p.51), but do not appear in her final scheme, and, therefore, seem relegated to become a relics of history.

A very different view of the role of history in mathematics education is taken by Jahnke (1994). He argues for the importance of seeing mathematics in a cultural setting. In order to accomplish this, he suggests that secondary students be exposed to some original historical sources that complement and conceptually diversify the traditional curriculum. Mathematics entirely stripped of its origins and cultural setting is called by Jahnke "fast food mathematics" and, he argues, it can not be fully appreciated nor entirely comprehended. He advocates for the exposure of students to the questions and problems that led to the genesis of mathematics, but he sees this exposure as an important, but separate part of curriculum that does not fundamentally change the standard approaches. In college teaching, recently some have carried these historical ideas further, and created college mathematics courses which are taught directly from original historical sources (e.g. Laubenbacher and Pengelley at New Mexico State University; or Cohen at Cornell University).

I share Jahnke's view of the importance of historical genesis and cultural interpretation for mathematics education. I feel that exposure to original source material can provide unique conceptual insights and that this would be extremely valuable especially for secondary teacher candidates. Often even courses in "history of mathematics" contain little or no original material, instead rewriting history in modern terms that eliminate conceptual diversity and social context. Such progressive absolutist history can do more harm than good since it reinforces the belief the mathematics is an inevitable march towards the current situation. More conceptual diversity is illuminated by the approach taken by Laubenbacher and Pengelley, where specific mathematical topics are investigated by students entirely from original sources.[7]

While I wholly support the movement towards greater exposure to original mathematical texts and their cultural origins, that is not the main direction of this thesis. I seek fundamental reforms in the secondary curriculum inspired by diverse historical conceptions that are then reinterpreted in light of tools, technology, and student activities. I am indebted to the work of historians who have preserved, translated, edited and made available the mathematical texts from which I have so richly benefited. I try to always keep in mind Jahnke's exclamation that "history of mathematics is difficult!" (1994, p.141).

Since all historical accounts involve interpretation, one might ask what sort of historical interpretation will be provided here? What sort of "rational reconstruction" (Lakatos, 1976) of mathematical history will I provide, and why is it valid? Keeping in mind my theoretical perspective (Section 1.3), I approached the historical material with following intentions:

1). To understand the genetic epistemology of the relations between curves, equations, and functions.

2). To see the history of the function concept in relation to an epistemology of multiple representations.

3). To seek out examples of the interrelationships between grounded activities and systematic inquiry.

The history that will be presented in Chapter 2 focuses on specific mechanical tools. When one builds a copy of one of these tools, and experiences its action, one is

---

[7] Laubenbacher and Pengelley have been experimenting with this kind of teaching at the University of New Mexico at Las Cruces. They hope to publish soon a collection of the source material that they have found most successful. In January 1995 at the joint meetings of the MMA and the AMS they organized a contributed paper session devoted entirely to this method of teaching.

recreating history at a sensory motor level. When I describe various forms of analysis of these tools that were carried out by seventeenth century mathematicians, I will try to communicate a strong sense of the original conceptions that shaped the genesis of these mathematical ideas. Complete authenticity can only be obtained form a reading of original sources (they are referenced), but even if a modern reader does not consult these original texts, a profound and valid experience is to be had from the physical experience of the tools themselves. Even crude cardboard replicas of these devices can convey much about their actions.

While researching the historical material for this thesis, I encountered many aspects of social and cultural history which provided various insights into how and why our mathematics curriculum has come to be what it is. This led me to search more aggressively for the roots of certain developments in mathematical thought that have come to privilege some forms of expression over others. I have sought to imagine how mathematics might be viewed in another light. The main focus of this investigation always remains on how the physical tools shaped experiences, and how language, and notation were forged from those experiences. Important cultural insights can indeed be gained from thinking about tools, and the activities and language which arise from them. It is the history that it is embedded in direct experience with the curve drawing tools which shall indicate and validate my educational claims.

1.6 Educational Implications

A variety of epistemic and conceptual issues are raised by a close study of the tools involved in the historical genesis of analytic geometry. I wish to argue that these issues have direct implications for curriculum reform. Historical details have been included in what follows in order to highlight some different conceptions that suggest specific activities and approaches to curriculum. Throughout Chapter 2, descriptions and analysis of important historical curve drawing devices have been interwoven with suggestions for physical experiments with linkages and paper folding. These descriptions often utilize original figures from the seventeenth century. Many other figures have been generated from computer simulations and animations which generalize and extrapolate these physical situations. When thinking about the educational implications of these historical ideas, the possibilities of new electronic tools will be presented and discussed.

When focusing on the issue of how language and symbols are created in response to activities, computer simulations will provide an intermediary environment which will facilitate the flow of the feedback loop which I will propose as an educational model. Computers lie somewhere in between the semiotic and the physical (Confrey, 1993a). In their internal workings, they are strings of symbolic code, but the experience of a user of an application like *Geometer's Sketchpad* can feel, at times, very close to physical experiment.[8] Curve drawing devices may have historically dropped out of mathematical discussions because of their physical limitations, and the increasing success of newly developed notation systems (e.g. Leibniz's calculus). Computer applications, however, are creating a new environment where this approach can be effectively revived with greatly extended capacity. The linkage discussed in Section

[8] Research in this area of cognitive science is currently being pursued by the Educational Development Center in Newton, Mass.

2.12, for example, has a very awkward motion that is easily halted by cumulative friction. Although conceptually it was very important to Descartes, it does not function well a mechanical device, but a computer simulation of that device is very satisfying.

In Chapter 3, students discuss their work using several physical curve drawing devices that placed them within an important historical conception of curves. Although these experiments with students did not include computer simulations, the issues that they raised using the physical tools alone were profound and manifold. As they investigated the actions which produced curves, and thought about different ways to analyze those actions, they confronted many issues about the relations between geometry, coordinates, and algebra. The situation allowed them to experiment in their own ways, and to express their beliefs in contexts that were meaningful to them.

Both of students who were interviewed for the material in Chapter 3 found the actions of the three devices with which they worked, physically straightforward, and yet subtle and perplexing when they attempted to fully analyze the actions. They were not in the habit of generating mathematical language from physical actions. Although they were surprised and, at first perplexed, this did not lead to stagnating frustration. Quite the contrary, because of the seemingly simple actions which they were directly experiencing, they were determined to reach a certain level of conclusive analysis, using whatever language seemed appropriate. Their approaches were quite different, and unlike many activities in mathematics classrooms, the linguistic form of their approaches was not predetermined. The actions of the tools were the constant focus. Although they were free to use any language and symbols they wanted, their discussions, guided by their activities, led constructively to material that is important in the secondary curriculum.

The historical material, computer simulations, and student interviews will combine to support the three claims made at the end of Section 1.1 in the following ways. Chapter 2 will mainly present a detailed reexamination of historical records, that will demonstrate tools and activities which formed a basis of knowledge which was indispensable for the growth of the linguistic structures of analytic geometry, calculus and the notion of functions. Most sections of Chapter 2 will also include computer experiments made using *Geometer's Sketchpad*, which will demonstrate how a variety of physical experiments can be extended in ways that allow elementary access to topics that are rarely discussed before calculus. Chapter 3 will provide detailed analysis of student conceptions that will provide extensive support for the first two claims.

Although the third claim concerning tangents, areas, and arclengths is not addressed by the student interviews of Chapter 3, the extensive material in Chapter 2, and the remarkably clear and enthusiastic response of the students to the basic ideas of curve drawing, suggests that this would be a very fertile direction for future research. Students with experience with both physical devices and dynamic geometry computer applications might engage in detailed analysis of a large number of problems that are routinely delayed until after calculus. It would be most interesting to see how they would respond to the formalisms of college mathematics if this larger experiment in constructivist education were carried out.

The educational and epistemological stances that I am taking can not be separated from the curricular activities that I will suggest. Genetic epistemology leads to a careful reading of historical sources which in turn immerses one in an epistemology of multiple representations. The required checking between multiple representations leads to a careful consideration of tools and activities, and they in turn mediate knowledge, and the language in which it can be expressed. This makes for a complete cognitive feedback loop. These educational views have been shaped directly by the work of Jere Confrey and through her, indirectly by von Glasersfeld, Piaget, and Vygotsky.

An empirical epistemology of multiple representations is historically significant, and through the use of curve drawing devices, I shall provide compelling examples of how this approach could invigorate classroom activities and discussions. Lakatos (1976) suggested this approach, and the extensive work of Confrey has established a theoretical framework for how such an epistemology could shape educational practice (Section 1.3). The theoretical pioneering work has been done. It seems to me, that it is now time for the constructivist educational approach to enter into a phase of baroque elaboration. This educational approach calls for an immense broadening of classroom activities at all levels of mathematics. As the following chapters will show, direct geometric generation of curves can make a fundamentally important contribution to this base of activities.

Chapter Two: The Historical Role of Geometric Curve Drawing Devices

2.1 Opening Comments

This chapter will present an investigation of a variety of physical curve drawing devices and techniques. All of these ideas and concepts came from important historical sources, mostly from seventeenth century Europe. The investigation, however, will not be solely historical. Along with descriptions of original conceptions, I will discuss various ways to combine and reinterpret these ideas in light of modern computer applications, and modern educational goals and theory. This rethinking of mathematical history will suggest more diverse approaches to the relations between algebraic equations and curves, as well some different views of the concept of a function and its relation to the notion of a parameter (see Section 1.2).

Section 1.3 discussed the work of Confrey, von Glasersfeld and others that view functions as a cognitive feedback loop that gains viability by moving between multiple representations. Confrey's studies of secondary and undergraduate student conceptions and her consequent software (1994d), and curriculum development projects (1994c), stressed three important functional representations, i.e. tables, graphs and algebraic equations. In her work with younger students she frequently mentions the need to develop a much stronger sense of ratio that is more closely tied to geometric similarity (1993c). This chapter will suggest some ways to extend this work on geometric similarity into the secondary and undergraduate curriculum via mechanics and dynamic geometry.

By generating curves geometrically as an initial activity the possibilities for coordination between multiple representations become even broader than what is available in the three functional representations that Confrey analyzed in her pioneering work (1994c). Although her computer software (1994d) allows for some geometric manipulation of graphs it does not allow curves to be generated geometrically as a primary action. Indeed, such activity goes beyond the usual mathematical notion of a function, however they may be represented.[9] The question becomes: What is a representation of what? Are curves visualizations of a functions? or are functions linguistic tools which can represent curves? The former view is taken in most mathematics curriculum. The tools that we give to students (e.g. graphing calculators) force one into this view. An investigation of what happens when the latter view is taken, can help to complete a cognitive feedback loop that has a rich history.

Functions that will be investigated here as part of a cognitive feedback loop that flows back and forth between curve generating actions and their representations in a symbolic language, i.e. algebraic equations. As was mentioned in Section 1.2, the difference between curves and graphs lies in their method of generation. All of the examples that are discussed in detail here concern the situation where curves are generated first, and coordinates and equations provide a secondary analysis. It is this part of the feedback loop, particularly the physical, grounded activity of curve drawing, that is underrepresented in our curriculum and often scarcely known even among trained mathematicians. This chapter will restrict itself entirely to plane curves, because there is so much there that can be said that is both profound, and rarely addressed in our usual mathematics curriculum.

Much of what will be discussed involves neglected aspects of geometry, but no attempt will be made to construct a set of formal mathematical proofs from a more geometric standpoint. Approaching analytic geometry as a cognitive feedback loop can avoid the establishment of any hierarchy of knowledge (von Glasersfeld, 1984, 1982). If there is an extensive use of geometry in what follows it is presented in the interest of conceptual diversity and as an antidote to the preponderance of algebra in our

---

[9] Parametric equations come closest.

mathematics curriculum. Much of the lost diversity of seventeenth century European mathematics concerns the dynamic, mechanical generation of curves.

This chapter will establish the importance and power of this form of activity from the beginnings of analytic geometry, through to the development of calculus. Many tangent, area and arclength properties, usually discussed only in a calculus class, will be demonstrated in a setting that requires little more than a knowledge of similar triangles. This material will suggest numerous ways to restore to our classrooms some of the grounded activity without which the development of the systematic language of calculus would not have been possible or necessary. This discussion will move, in an almost circular manner, between static geometry, dynamic geometry, and algebraic equations. Questions will be raised concerning the extent to which one can establish an empirical consistency between these various representations.

Mathematical experimentation, which first established a partial linguistic consistency between geometry and algebra, was a major topic of discussion in seventeenth century Europe, i.e. the period of the scientific revolution (Cajori, 1929). Because of this, a majority of the examples presented are taken from seventeenth century European sources. It is in the discussions from this period that one can regain both a sense of doubt, and a sense of wonder, at the physical activities that allowed people to believe that such a consistency was possible, and that it could be symbolically expressed. When presenting historical examples, I will usually try to stay as close as possible to the original conceptions of the writer, in an attempt to convey alternative mathematical concepts. It is these alternative pathways that enliven and enrich the feedback loop which connects curves with analytic geometry and functions.

I am <u>not</u> approaching these issues from a belief that ontogeny recapitulates phylogeny. This rethinking of the historical genesis of the links between curves and equations is important for two reasons. First, although our secondary curriculum is almost entirely concerned with seventeenth century European mathematics, much of the beauty and diversity of that era have been lost from our textbooks and even from books on the history of mathematics. Hopefully, the reader will look with an open mind and find something here that is beautiful, profound and useful in the classroom. Second, the presence of computers in the classroom and the availability of software packages that allow for dynamic geometric experimentation, make much of the seventeenth century study of curves quite accessible to many students with only the most basic computer skills. Most sections of this chapter will begin with a figure taken directly from a seventeenth century text, and then animate that figure using the software, *Geometer's Sketchpad* . Analysis of each example will include both history and modern educational possibilities.

My own investigation of curve drawing devices began with a careful reading of René Descartes' *Geometry* (1952), first published in 1637. The original goal of my reading was to trace the history of exponents and their ties to geometric similarity (see Section 2.12), but I was immediately struck by the fact that not once did Descartes create a curve by plotting points from an equation. His method was always first to devise physical or geometric devices for the drawing of curves, and then to analyze the dynamic actions of the devices until an equation emerged. I was so struck by these devices that I began building simple models of them using cardboard, thumbtacks and rubber bands. Often crude models took only a few minutes to build. They gave an immediate and comprehensible feeling for the actions which built the curves. Descartes intended each part of the equations which emerged in his work, to be seen as a code for a particular geometric invariance, most often a set of similar triangles (1952; Lenoir, 1979).

After playing around with cardboard models for several months, I was introduced to the software *Geometer's Sketchpad*. At first, I found this software entirely unnatural for the purpose of simulating moving hinged rods (usually called linkages). Every line segment in this software behaves essentially like a rubber band unless confined by a circle, or defined as a translation of another fixed length. After a while I learned to work around the limitations of the software, and used it to simulate a large number of linkages from a variety of sources. For a time, I gave up building physical models altogether, and fell under the spell of the computer simulations. They were a great aid in investigating the more difficult devices particularly those of Isaac Newton.

When I subsequently began to seriously consider how I might present some of these curve drawing linkages to students, I once again began building physical models and trying to improve on the materials and designs. I was surprised at how my understanding of some of the simpler devices was deepened by the task of building moving linkages. I began noticing a variety of details and interconnections between devices that were largely obscured by the computer simulations. Most noticeable of all was a feel for the related rates of motion between different links in each device. Physical instincts revealed a variety of properties that I had missed on the computer. When I watched students experiment with my improved linkages even more was revealed, but that is the subject of the Chapter 3.

Both my own experiences and those of the students in Chapter 3 confirm the assertions of Jere Confrey that mathematics can be viewed as a dialectic between grounded activity and systematic inquiry (1993a). Physical linkages provide a solid sense of experimentation with little mystery. Although linguistic analysis may be difficult, all the connections in a linkage are immediately tangible and comprehensible. Algebraic equations can represent these actions in a symbolic way, and, vise versa, all algebraic curves can be drawn with linkages (Artobolevskii, 1964). Computers provide an interesting intermediary tool environment. They allow one to rapidly generalize and extend the range of physical devices, yet they often fail to provide the fundamental intuitions which can lead to appropriate conjectures and genuine mathematical understanding. Computer simulations are also mysterious. A great deal is hidden from view and from touch.

My advice to the reader is never to depend entirely on computer animations. Although in the sections that follow I will make extensive use of computer simulations of curve drawing devices, one can often quickly build a simple cardboard model which can be quite revealing. Many of the figures that are included in this chapter were made using *Geometer's Sketchpad*. In some cases I will present a pair of figures together on a page that show a device in two different positions; using two frames from an animation to convey on paper some sense of motion. Preceding such figure, I have usually included a drawing from the seventeenth century of an actual linkage device. These drawings are beautiful and historical, but even more important is the sense of physicality that they convey. My favorite examples are taken from the work of a friend and contemporary of Descartes, Franz van Schooten (1615 - 1660).[10] The figures from van Schooten's work (1657) include drawings of human hands manipulating the devices. On paper, this is as close as possible to physicality.

Not every topic in this chapter is strictly historical, and the history that is given is not always strictly chronological. The intention is to build towards a concept of analytic geometry as a broad and diverse cognitive feedback loop, wherein students can conduct their own experiments which reveal the possible inter-consistency of geometry and algebra. I will occasionally jump around chronologically in order to investigate a particular issue. At a few points, inspired by historical examples, I have investigated certain issues in my own fashion, and built my own examples that take advantage of the experimental possibilities of computers (most notably Sections 2.12 and 2.14). Every attempt will be made to be clear about what is and is not strictly historical. Secondary sources and loose translations can be very misleading. Original sources have been consulted,

---

[10] The name "Franz van Schooten" is listed in references sometimes as "Schooten, Franz van" sometimes as "Van Schooten, Franz" and sometimes in Latin as "Schooten, Francisci á." Throughout this work I will use the first listing.

whenever possible, in an attempt to convey interesting mathematical conceptions that lie outside our mainstream curriculum.

2.2 Apollonius and Conic Sections

The intent of this chapter is to discuss the dynamic generation of curves using a variety of continuously moving devices. Most of the sources for this material are from seventeenth century Europe. The mathematicians of this period, however, all read Apollonius (262 - 200 BC.). Often original works of mathematics from this period were written as commentaries on the work of Apollonius or as reconstructions of supposedly lost works of Apollonius (e.g. Schooten, 1657). When it came to the study of curves they often saw themselves as attempting to extend and generalize the work of Apollonius, especially *The Conics* (1952; Heath, 1961). Important Arabic additions to *The Conics* from the tenth and eleventh century were also crucial to the mathematics of the scientific revolution (Joseph, 1991). In fact, even to this day, large parts *The Conics* of Apollonius do not exist in their original Greek form, and are known to us only through Arabic translations.

For several reasons, we must look briefly at some of the theorems and definitions from Apollonius. First, although his approach was entirely from static geometry, *The Conics* is a systematic approach to conic sections which takes the view of a conic curve as a set of parallel, ordered line segments from an axis. This approach will be discussed in detail shortly, and will be seen to be very close to the idea of coordinate geometry. For this reason certain concepts developed in *The Conics* were fundamental in the development of the function concept. Second, *The Conics* contains a variety of beautiful theorems that are very little known today. In particular, the theorems on symmetry, axes and conjugate diameters are well suited to secondary education, and could help connect ratio and similarity with coordinates and equations in a non-linear setting. These theorems were used extensively throughout the seventeenth century, for example in Newton's *Principia* (1952) for the discussion of orbits (See Section 2.5).

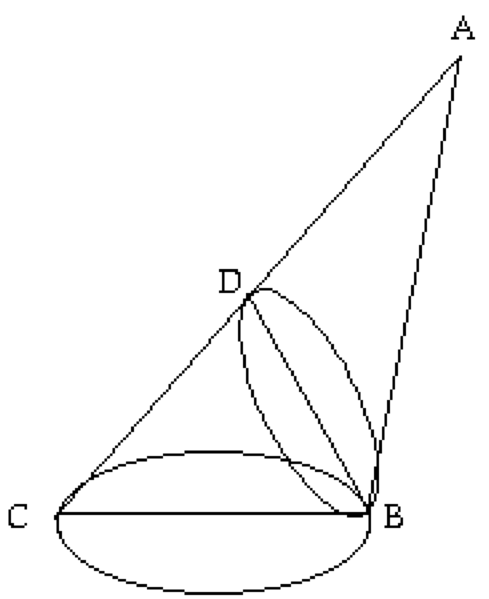
In this section I shall describe a few demonstrations from static geometry, and then summarize some concepts and results which I will explore later in the empirical setting of the dynamic curve drawing devices from seventeenth century Europe. There is strong evidence that most of the properties of conic sections from the first three books of Apollonius were known before the time of Apollonius but the earlier works, such as Euclid's books on conics, have all been lost (Heath, 1961). It was the work of Apollonius coming from Arabic culture that was influential in seventeenth century Europe. I shall make no serious attempt to reconstruct the history of curves in ancient Greek culture nor in medieval Arabic culture, although there is a great deal of material there could be of educational value.

For the reader who is intrigued by this section and wishes to explore the details of *The Conics* , I suggest looking at the wonderfully readable work by Heath (1961), but simultaneously consulting a more literal translation of Apollonius like that of Taliaferro (1952). Many of the original concepts and definitions are obscured by Heath in an attempt to make *The Conics* more palatable to a reader schooled in modern mathematics. Even though Greek mathematics is not the main subject of my investigation, the kind of rethinking of mathematics that gives historical investigations educational value requires, whenever possible, a close attention to original texts.

When conics are discussed in our secondary curriculum, it is mostly through the use of a set of standardized equations which are conceptually limiting in several ways. First, although some picture or model of a plane intersecting a cone is usually shown, there is usually no attempt made to demonstrate that the curves generated by the standard equations are actually the same as these conic sections. Second, the standard equations that are given are far from general, and students often fail to recognize conics in other algebraic forms, for example, that $y = \frac{5}{x} + 3x$ is the equation of a hyperbola.

This second issue will be addressed in latter sections of this chapter by looking at different drawing devices which naturally give rise to various forms of equations. I turn now to the first issue.

When I began reading Apollonius I was startled to find right at the beginning that there were some important basic aspects of cones that I had never thought about. A cone, for Apollonius, is the shape generated by considering all lines in space that pass through a fixed circle and a fixed point (the vertex) not in the plane of the circle (this really generates a pair of cones). The cone that most often comes to mind is the right cone. That is the case where the line from the vertex to the center of the fixed circle is perpendicular to the plane of the circle. In this case, all planes parallel to the fixed circle will have circular sections with the cone, and these are the only planes that make circular sections. If the line from the vertex to the center of the fixed circle is not at right angles to the plane of the circle, then there are <u>exactly two</u> families of parallel planes whose sections with the cone are circles. One of these families is parallel to the fixed generating circle and the other is called by Apollonius the subcontrary sections.



<u>Figure 2.2a</u>

If one looks at the triangles formed by the diameters of the circular sections (both families) and the vertex of the cone, they are all similar. That is to say in Figure 2.2a if the triangles ABC and ADB are similar then all sections of the cone parallel to either one of these circles will be circles. This is Book 1, Proposition 5 in Apollonius, and in all my years of studying mathematics I had never considered such a case. This is not an isolated oddity. Quite the contrary, it is the beginning of whole series of duality propositions about conics which are fundamental in many later approaches to curves and functions. As is often the case, similarity is at the heart of the issue.

I shall now demonstrate that when any cone is sliced by a plane which is parallel to one of its generating lines (but not containing a generating line); the resulting section is a parabola, in the sense that we know it as a curve, one of whose coordinates is proportional to the square of the other. This argument is taken from Apollonius (Book 1, Proposition 11), but was known much earlier being attributed by many ancient Greek writers to the early geometer Menaechmus (Coolridge, 1968). This argument does not depend on the cone being a right cone. All that is needed is some set of parallel circular sections.

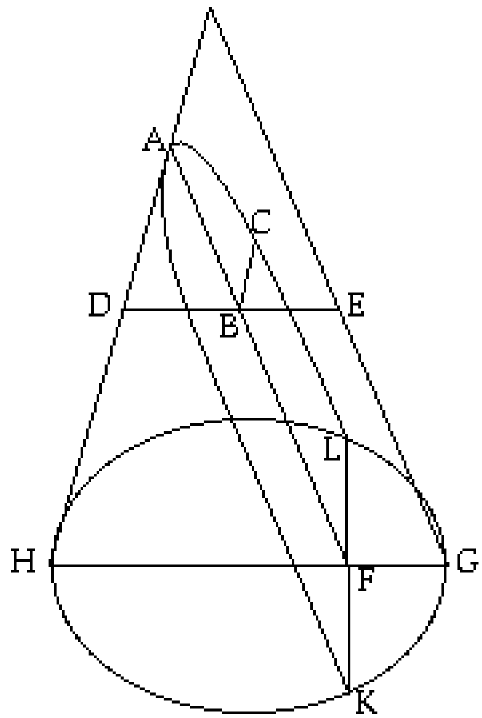


Figure 2.2b

In Figure 2.2b let the curve KAL be formed from slicing the cone with a plane parallel to the line $\overleftrightarrow{EG}$. Let GKH be a circular section and let D, C, and E lie on any parallel circular section (with C on the curve). Now LF is the geometric mean of HF and FG, i.e. $LF^2 = FH \times FG$. Likewise BC is the geometric mean of DB and BE, i.e. $BC^2 = BD \times BE$ (this is a fundamental property of the diameter of a circle and any chord perpendicular to it). Since the parabolic section is parallel to $\overleftrightarrow{EG}$, $BE = FG$. Since triangles ABD and AFH are similar to each other, we have the proportion $BD{:}BA = FH{:}FA$. If we think of sliding the parallel circular sections, and watching the two segments on the diameter, one of them is fixed ($BG = FE$), and the other (BD or FH) is proportional to the distance along the axis of the curve (BA or FA). Hence the distance along the axis of the parabola (BA or FA) is proportional to the square on the lateral distance out to the curve (BC or FL) by the geometric mean property of circles.

Introducing variables, we could let $AB = x$ and $BC = y$, and then say that there exists some constant of proportionality, p, such that $(p \cdot x) \cdot (BE) = y^2$. The constant, p, is the proportionality constant from the similar triangles just mentioned (i.e. $p = DB{:}AB$). It projects from the plane of the parabola (where x is measured) to the planes of the circular sections.

Of course Apollonius did not write an equation. Instead he showed how to construct a single line segment (called the latus rectum), and stated that the rectangle formed by this segment and $\overline{AB}$ is always equal to the square on $\overline{BC}$. This latus rectum (whose length equals $p \cdot BE$) works for any parallel circular section along this parabola. The latus rectum is a tangible line segment in his construction.

What I wish to stress is the view taken in this construction of a parabola as a changing set of geometric means where one of the segments is constant (in Figure 2.2b it is $BE = FG$ that is constant). In the cone, the circles which create these geometric means are not in the plane of the parabola, but the proportionality still holds there because of the similar triangles (i.e. $\Delta ABD \approx \Delta AFG$).

This view can be extended to the elliptic and hyperbolic cases by considering a linear proportional change in BE as we move down the cone. BE decreases in the elliptic case and increases in the hyperbolic case (see Figure 2.2b). In the elliptic case the equation coming from the geometric means on the series of circular sections would be: $y^2 = (p \cdot x) \cdot (p' \cdot (k - x))$; where y, x, and p are as before; k is the entire finite length of the axis along which x is measured; and p' is a constant of proportionality that projects $k - x$ onto the circular sections in the plane containing AHF ($p = p'$ only in the case where the ellipse is formed as a section of a cylinder). This equation comes from the cone in the same way as before except that BE is replaced by $p' \cdot (k - x)$. This equation can be rewritten as:

$\dfrac{y^2}{p \cdot p' \cdot \left(\frac{k}{2}\right)^2} + \dfrac{\left(x - \frac{k}{2}\right)^2}{\left(\frac{k}{2}\right)^2} = 1$ which is the standard form given in modern textbooks,

but with a translation because x is measured from a vertex at one end of the axis instead of from the center. For the hyperbola, simply replace $p' \cdot (k - x)$ with $p' \cdot (k + x)$ to indicate the increasing lengths of BE. k in this case is the distance between the vertex on the hyperbola and the vertex of the opposite branch. Again it should be noted that Apollonius did not write equations, but made these statements in ratio form with geometrically constructed adjustments to the latus recta. One should also note that Apollonius gave the converse of these arguments, i.e. that any curve with these geometric mean properties (equations) must be a conic section, because the geometric means can always be constructed from a series of circles whose radii are changing arithmetically (hence a cone).

A tenth century Arabic mathematician, Ibn Sina, wrote several commentaries on Apollonius in which he showed how to draw conics using ruler and compass (Joseph, 1992; Berggren, 1986). He simply flattened the picture, and put the series of circles in the same plane as the curve. His method was still the same. He built, for example, a parabola using a series of geometric means where one of the segments is held constant. Looking at Figure 2.2c, one sees a series of tangent circles all passing through the point S. A fixed vertical line at A constructs the geometric means between SA (constant) and a series of segments $\overline{AT}$, $\overline{AU}$, $\overline{AV}$. These segments are then plotted against the series of geometric means AX, AY, AZ, to give a set of points A, B, C, D, all of which lie along a parabola. In this picture the horizontal distances of the points (B, C, and D) from A is proportional to the square on their vertical distances. The latus rectum is SA.

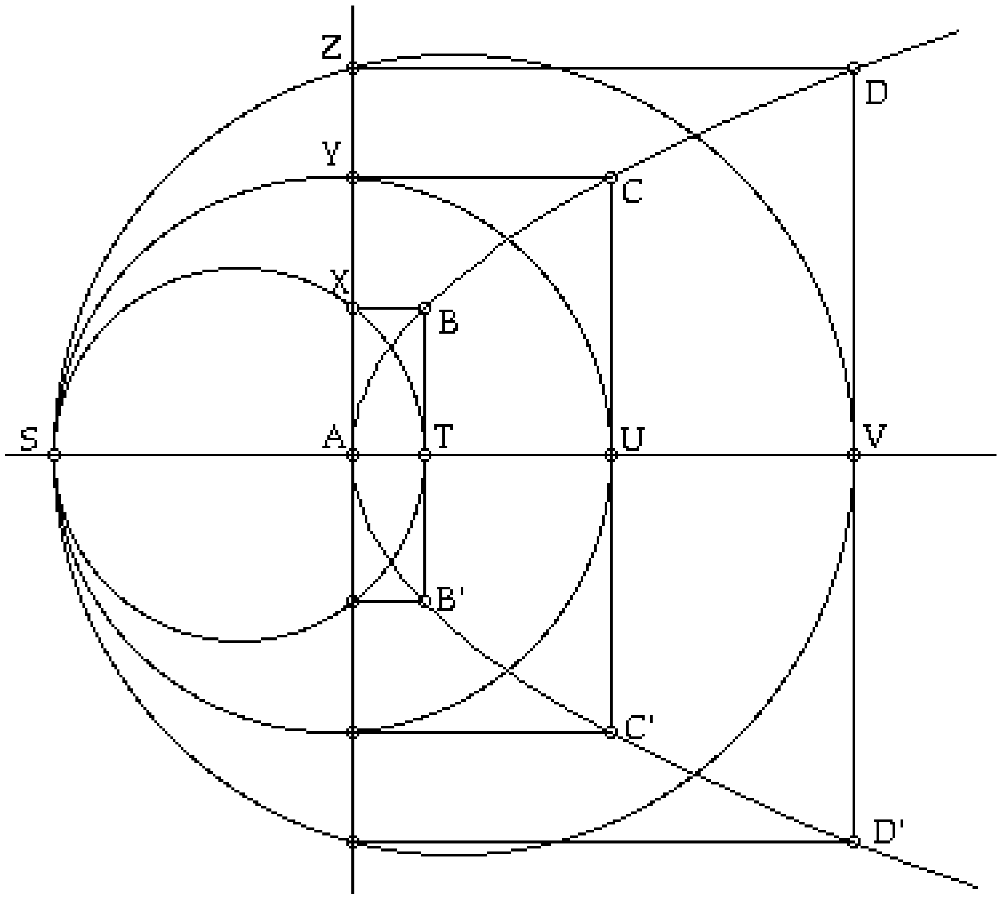


Figure 2.2c

This and other Arabic methods of drawing conics were static in the sense that they allowed one to plot as many points on the curves as desired, but they did not draw the curve in a continuous way. This particular construction of Ibn Sina, however, can be easily adapted as a design for a dynamic mechanism which will draw parabolas in a smooth continuous manner. Imagine a nail driven in at S, and a carpenter's square with the vertex of its right angle riding on the line $\overleftrightarrow{AX}$ and one arm resting against S. As we move the vertex X along the vertical line the other arm of the square will be intersecting the line SA at points T such that AX is always the geometric mean of SA and AT. Using a second carpenter's square XBT we can continuously locate the points on the parabola. Figure 2.2d shows two positions of an animation of such a device using *Geometer's Sketchpad*. The point X is being moved along the vertical line to drive the animation while the locus of the point B draws the parabola.

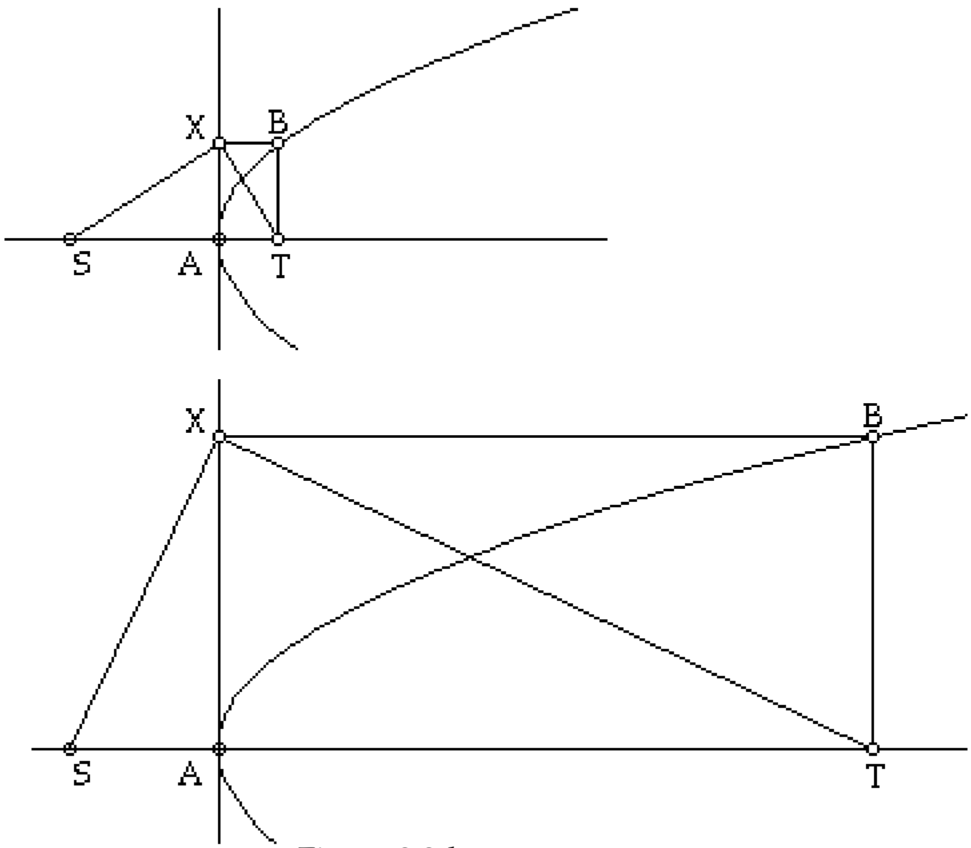


Figure 2.2d

This particular dynamic adaptation of Ibn Sina's construction was suggested to me by David Henderson, and was not taken from any historical source. It is similar in nature, however, to many of the devices from the seventeenth century that I will soon discuss. These opening examples provide a sense of one of the directions that this chapter will take, in that they allow the reader to see a transition from classical geometry to a dynamic situation where actions generate curves. I shall want to assert that an analysis of such actions provides an important groundwork for understanding the evolution of modern notions of analytic geometry and functions.

Arabic interest in drawing conic sections was spurred by their development of methods for solving cubic equations by intersecting two conic sections (e.g. by intersecting a parabola with a hyperbola). In the work of Omar Khayyam, a general geometric method for the solution of cubics is given (Joseph,1991). These methods had a profound impact on later Western mathematics (Katz, 1993). For example, one of the primary goals of Descartes' *Geometry* (1952) was to extend these Arabic methods by finding new ways to draw curves whose intersections would yield solutions to equations of arbitrary degree.[11] As we shall see in the following sections, the new views that transformed European mathematics in the seventeenth century were not just a move from geometry to algebra, but perhaps more importantly a move from a static to a dynamic geometry that involved time and motion.

Before discussing more examples of dynamic curve drawing actions, there are several more concepts from Apollonius which I wish to describe. In the opening definitions of *The Conics* the following definition appears:

> 4. Of any curved line which is in one plane I call that straight line the diameter which, drawn from the curved line, bisects all straight lines drawn to this curved line parallel to some straight line; and I call the end of that straight line (the diameter) situated on the curved line the vertex of the curved line, and I say that each of these parallels is drawn ordinatewise to the diameter (Apollonius, 1952, p. 604).

---

[11] One should keep in mind that equations to be solved were frequently stated first as geometry problems (e.g. doubling a cube, trisecting an angle, or finding the parameters of curves that meets certain tangent requirements). This was true for both Arabic and European mathematicians up to 1650.

When I first read Apollonius, I was totally perplexed by this definition, because it only makes sense if one is aware of some important symmetry and duality properties that hold for all conic sections.

It turns out that, in the case of circles, ellipses, and hyperbolas, the diameters are exactly the set of lines which pass through the center. In the case of the circle, the ordinates are always perpendicular to the diameter, but in the case of other conics this is not always so (see Figures 2.2e, 2.2f, 2.2g). The ordinatewise direction is the set of chords which are all parallel to the tangent line at the vertex of the diameter. It was surprising to me, at first, that such a family of bisected chords existed except along an axis of symmetry. This is never mentioned in modern discussions of conic sections, except in certain projective treatments. Apollonius defines an "axis" as a special case of a diameter where the family of bisected chords are perpendicular to the diameter.

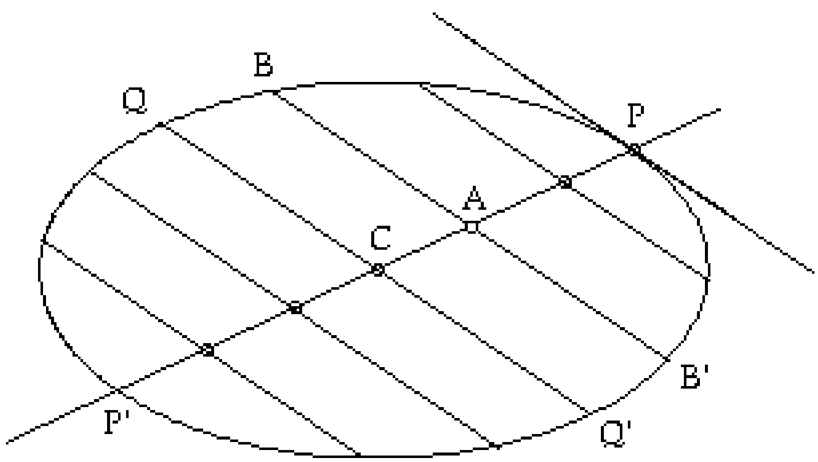


Figure 2.2e

Figure 2.2e shows an ellipse with center, C. Choosing an arbitrary point P on the ellipse, $\overleftrightarrow{PC}$ is a diameter according to the Definition 4, because all of the chords parallel to the tangent at P will be bisected by $\overleftrightarrow{PC}$. One can view this figure as a projection of a circle (although Apollonius did not explicitly do so). Another property of conics, that is demonstrated in Apollonius, is that the chord $\overline{QQ'}$ which passes through the center C has as its ordinates (i.e. bisected chords) all chords parallel to $\overleftrightarrow{PC}$. Hence the tangent at Q is also parallel to $\overleftrightarrow{PC}$. Apollonius called $\overline{QQ'}$ the conjugate diameter to $\overline{PP'}$. This is an example of a fundamental duality property of all conics. We shall return repeatedly to this concept of conjugate diameters in Sections 2.5 and 2.6, and see its relevance for both mechanics and planetary orbits.

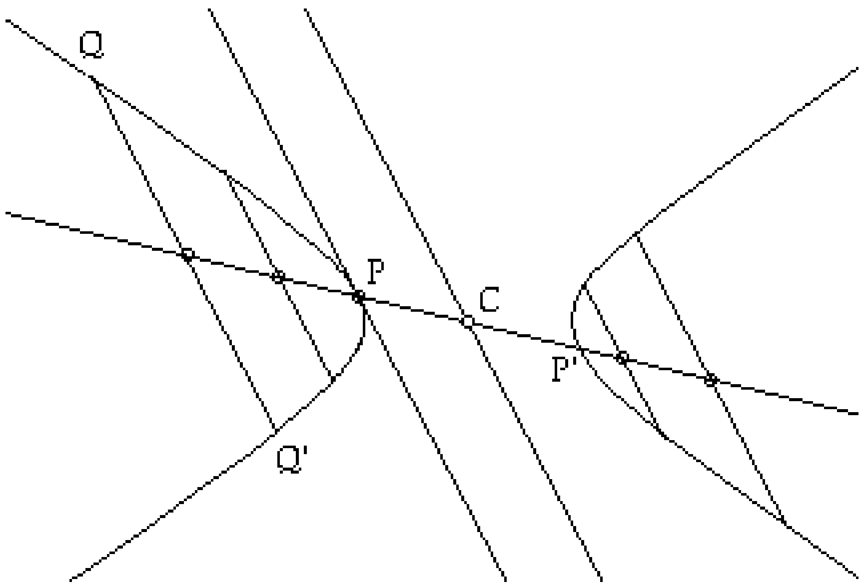


Figure 2.2f

Let us look next at these bisected chords and conjugate diameters on the hyperbola (see Figure 2.2f). It remains true that any line through the center, C, will bisect all chords (e.g. $\overline{QQ'}$) parallel to the tangent at the vertex P. There is a line through center which is parallel to the ordinates with respect to $\overleftrightarrow{PC}$, but it does not intersect the curve. It is still, however, a conjugate diameter in the sense that any chord parallel to $\overleftrightarrow{PC}$ will be bisected by this line through C. Such chords (e.g. $\overline{PP'}$) go between the two opposite branches of the hyperbola instead of being contained within one branch. Although this "conjugate diameter" does not intersect the curve, Apollonius gives a finite length centered at C so that it has the same ratio properties as in the elliptic case. This is a generalized version of what is usually done in classrooms when a specific rectangle is constructed in the middle of a hyperbola having the asymptotes as its diagonals.

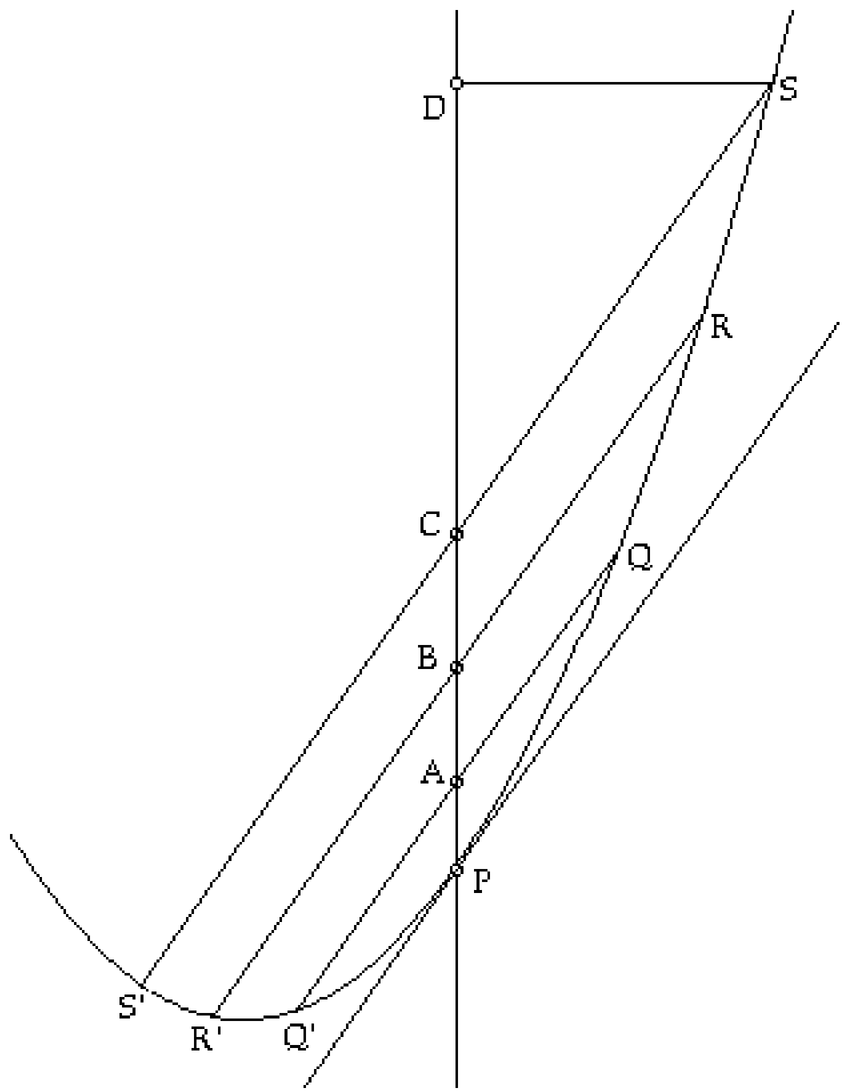


Figure 2.2g

Turning to the parabola adds more justification to Apollonius' definition of a diameter. A parabola has no center (it is at infinity) but it does have many diameters. Without projective methods, these diameters can not be defined as lines through the center. They turn out to be all lines which are parallel to the axis of symmetry (see Figure 2.2g). In this case none of the diameters intersect each other (unless one includes the point at infinity), so there are no pairs of conjugate diameters, but any vertical line in Figure 2.2g (e.g. $\overleftrightarrow{PC}$) will bisect all chords parallel to the tangent at its vertex P (e.g. $AQ = AQ'$, $BR = BR'$, and $CS = CS'$).

Apollonius demonstrated ratio properties for each conic section, that if translated into algebra, and applied to the axes (in the Apollonian sense), would yield our modern standard equations. He demonstrated, however, that these ratio properties are true for any of his diameters. In modern algebra this implies that the form of the equations of conic sections remains unchanged if we coordinatize the curve along different diameters, provided we always use the ordinate direction appropriate to the diameter (i.e. a non-perpendicular coordinate system using the direction of the bisected chords). Only the constants (e.g. the latus rectum in the parabolic case) will have to be adjusted. When we use our standard equations of conics (coordinatized perpendicularly along the axes of symmetry), we can choose either the positive or negative square root to find pairs of symmetrical points. This Apollonian bisection property shows that this can also be done using any diameter, even if it is not an axis of symmetry.

For example, in Figure 2.2g, if we use the diameter at P, then for any point C, we will have that $CS^2 (= CS'^2)$ will be proportional to PC. The right triangle CSD will give us the appropriate new proportionality constant, relative to the axis the of symmetry, by taking the ratio of the squares of CS to DS. For example let us assume that the parabola in Figure 2.2g has the equation $1 \cdot y = x^2$ with respect to the (axial) perpendicular coordinates with the origin at the vertex. Suppose we then construct the diameter (vertical line) through the point $P = (1,1)$ and use parallel chords all having a slope of $+2$ as our ordinates. Supposing that $P = (1,1)$ in Figure 2.2g, that is to say, let the coordinates of the point Q on the parabola, be given in the new system as $x' = AQ$ and $y' = PA$; or of the point S, as $x' = CS$, and $y' = PC$. The equation of the parabola in these coordinates will be $5 \cdot y' = x'^2$, because the ratio $CS^2 : DS^2 = 5:1$ (coming from a right triangle with legs in a $2:1$ ratio). For example, the points $(2,4)$ in the axis system becomes $(\sqrt{5},1)$ with respect to the diameter through P. The point $(-\sqrt{5},1)$ is on the curve at the other end of the bisected chord. Putting the constant (i.e. 5) with y preserves the dimensional integrity of the statement (i.e. it says that a rectangle equals a square). Apollonius would say that the latus rectum with respect to the new diameter is five times the latus rectum with respect to the axis. Such a statement avoids choosing a unit.

To demonstrate the statements in the previous paragraph, look back at Figure 2.2b. If the cone is oblique such that the plane containing A, H, and G is not perpendicular to the plane of the base circle $KHLG$, then AF will not be the axis of the parabola but it will still be a diameter (Apollonius, Book 1, Prop. 7). BC and FL will still be ordinates with respect to AF, but they will not be perpendicular to AF (they will still appear as in figure 2.2g) . The previous discussion of Figure 2.2b remains valid since it depended only on having a series of parallel circular sections which produce a series a geometric means. AB is still proportional to BC^2. The form of the equations remain the same. The same is true for ellipses and hyperbolas. Ratio and similarity are more flexible and general than they often appear in our usual curriculum.

For the moment, this is all I wish to say about *The Conics* (1952) of Apollonius. I have sketched a series of properties of conic sections, although I have no intention of discussing, in detail, the proofs and demonstrations that were given by Apollonius. Such a classic and readily available work speaks for itself, although some sense of the direction of this work is most helpful before one attempts to read *The Conics* (Boyer,1968; Heath, 1961).

Today even the few conic properties from Apollonius that I have mentioned are not well known, even among trained mathematicians. I think that this is because very few people have much direct physical or mathematical experience with conics or any other curves. Grounded activity in this case has almost disappeared from our mathematics curriculum. For example, the duality of the conjugate diameters, and the non perpendicular coordinate systems they imply, are only discussed very indirectly in a modern course on linear algebra, but in this highly generalized and abstract setting there is usually no mention of the geometric implications of the matrix algebra. Look, for example, at the discussion of bases and duality in a standard text in linear algebra like Hoffman & Kunze (1971).

I do not suggest reintroducing some form of the work of Apollonius into our secondary curriculum . Instead I will show in the following sections of this chapter how these and other properties of curves could be experienced through the dynamic activity of curve drawing. The following investigations will stress empirical experimentation with curves For example, the bisection properties of conic chords mentioned above might be explored quite simply in the following way. Draw a conic curve, and draw any chord, then draw another chord parallel to it. Next find the midpoints of these two chords and draw the line connecting them until it meets the curve. Check empirically that the tangent to the curve at this point is parallel to the chords. Check that any other chord parallel to the first two also has its midpoint on this line. Check that the center of the curve is on this line (if it has a center). Such an investigation of the tangent properties of conics could be applied to a variety of quadratic problems concerning maximum and minimums, mechanical devices, falling bodies, or planetary orbits. A full understanding of conic tangents can be had using only the most elementary geometric ideas.

Of course such an activity would first require ways to accurately draw such curves. Slicing cones is not the most convenient way, nor is it always the most revealing. Apollonius got away from cones quite rapidly and discussed conics as curves in the plane. I shall be investigating dynamic actions that produce curves, usually linkage drawing devices. These have several educational advantages. They are essentially two dimensional. They can often be easily built from cardboard, wood or metal. They can be simulated on readily available computer software. Most important of all, however, is that they provide a beautiful example of a situation where geometry and algebra meet in a set of grounded activities which could serve as wonderful introduction to equations, functions, and parameters. This interweaving of geometry and algebra can be seen both intellectually and historically as a platform from which to construct and validate a variety of language and notations, one important example being the calculus of Leibniz.

2.3 Abscissas, Ordinates, and Functions of a Curve

I must begin here with a confession. In my discussion of Apollonius' theory of diameters and ordinates I quoted the words "ordinate" and "ordinatewise" in my discussion knowing that they were conceptually misleading. These words appear in Heath's translation (1961) as well as in the more literal translation, although there a footnote is included (Apollonius, 1952, p. 604). Apollonius never used any one noun that could be reasonably be translated as "ordinate." Instead he always used phrases like "parallel line segments applied in order." Even in Descartes' *Geometry* (1952) the words "abscissa" and "ordinate" never appear except in English translations. He continued to use active phrases similar to those of Apollonius. Consult, for example, the bilingual, facsimile edition of *The Geometry* where the single English word "ordinates" is given as the translation of the phrase "*les pareilles lingnes sont appliquées par ordre*" (1952, p. 88). On this same page the word "diametre" (intended in the Apollonian sense) is routinely translated as "axis" which, with modern readers, tends to limit the generality of his arguments.

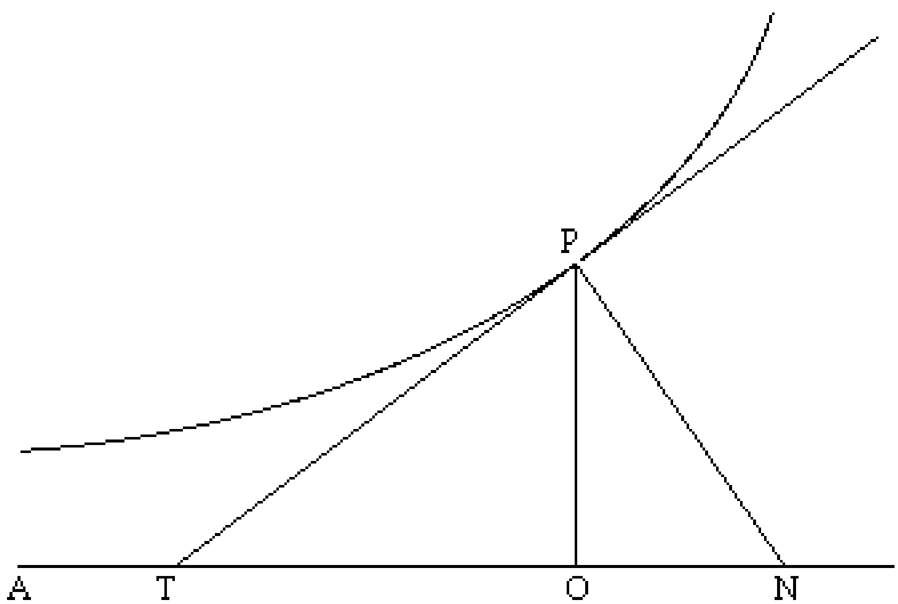
These kinds of translation imply more than just expedience. Phrases indicating action are here being replaced with single nouns. Modern scientific and educational values tend to move us as quickly as possible in the direction of objective abstraction. The actions which generate curves are replaced with a notion of a function as an object expressed in an abstract equation. I am not universally opposed to such conceptions, except when they are used to entirely cloak the physical actions which form the genetic roots of the function concept (Confrey, 1993c). I certainly do oppose the educational opinions of those, like Sfard (1992), who advocate moving as quickly as possible to a view of entire functions as objects, and actively seek to discourage students from a view of functions as actions. This glorification of the severing of mathematics from its experiential roots can only increase our "suffocation in specialization" (Courant, 1984, p.132). It is these values that I feel are at play in the issues of translation mentioned above.

When the notion of a function evolved in the mathematics of the late seventeenth century, the meaning of the term was quite different from the modern set theoretic definition, and also different from the algebraic notions of the nineteenth century. The main conceptual difference was that curves were thought of as having a primary existence apart from any analysis of their numeric or algebraic properties. Equations did not create curves, curves gave rise to equations. Equations, functions, and parameters were thus derived from curve generating actions. When Descartes first published The *Geometry* in 1638, he derived for the first time the algebraic equations of many curves, but never once did he create a curve by plotting points from an equation (1952).[15] Geometrical methods for drawing each curve were always given first, and then by analyzing the geometrical actions involved in the curve drawing apparatus he would arrive at an equation that related pairs of coordinates (not necessarily at right angles to each other). Descartes used algebraic equations to create a taxonomy of curves

[15] In the same year Fermat did approach the subject from the standpoint of graphing equations, but even in his work the study of equations made extensive use of geometric transformations of coordinates which almost amount to curve drawing, as I shall show in section 2.10.

(Lenoir, 1979). An isolated equation that was not a code for a geometrical action had, for him, no epistemic significance.

This tradition of seeing curves as the result of geometrical actions (as opposed to graphs) continued in the work of Roberval, Pascal, Newton, Leibniz, and others of that time. The term "function" was introduced into mathematics very late in the seventeenth century, by G. W. Leibniz (1646 - 1716), as a line segment or ratio that could be determined from each point on a curve relative to a given line or axis (Arnol'd, 1990). Descartes (1952) used the lower case letters x, y, z to represent variable lengths along curves, but he did not create any specific system of names nor did he ever use the term "function".



$\overline{PO}$ = ordinate $\overline{OA}$ = abscissa
$\overline{PT}$ = tangent $\overline{OT}$ = subtangent
$\overline{PN}$ = normal $\overline{ON}$ = subnormal

Figure 2.3a

Leibniz considered, among others, six different functions of a curve and gave them the following names: abscissa, ordinate, tangent, subtangent, normal, and subnormal. These six are shown in Figure 2.3a, for the point P on the curve, relative to the axis $\overleftrightarrow{AO}$. The line $\overleftrightarrow{PO}$ is perpendicular to $\overleftrightarrow{AO}$. The line $\overleftrightarrow{PT}$ is tangent to the curve at P, and the line $\overleftrightarrow{PN}$ is perpendicular to the tangent, $\overleftrightarrow{PT}$. Other geometers of the seventeenth century (e.g. Pascal, Barrow and Huygens) had begun discussing some of these objects connected with curves, and had given them names (e.g. Pascal used the term "ordonnées", and Barrow used the term "subtangent"), but their discussions focused on them one or two at a time (Struik, 1969). Leibniz considered them all as a system associated with each point on a curve, and so gave them the collective name: "functions" (Arnol'd, 1990).

It is important to note here that the curve and an axis must exist before these six functions can be defined, and that Leibniz was using the term "axis" in a more general way than did Apollonius He no longer required the curve to have the symmetry of bisected chord ordinates, but he did require the ordinate direction to be perpendicular to the axis. In this second sense he was still following the terminology of Apollonius concerning the distinction between an axis and a diameter, but Leibniz wanted to consider curves more general than conics, including those without symmetry. An axis here was a reference line for perpendicular ordinates, but it was no longer required to be an axis of symmetry, although this was done whenever possible.

In this definition, the abscissa and ordinate may at first seem to be a parametric representation of the curve, but this is not the case. No parameter, like time or arclength, is involved. The setting is entirely geometric. From the geometric point P, the line segments (functions) are defined relative to the axis $\overleftrightarrow{AO}$. *Abscissa* is Latin for "that which is cut off," i.e. a piece of the axis, $\overleftrightarrow{AO}$, is cut off. By cutting off successive pieces of the axis, the curve gives us a series of ordered line segments $\overline{PO}$. Hence the term *ordinate* . Leibniz used these two terms in the tradition of expressions that appeared in Apollonius.

It should also be noted here that all of these functions of a point, P, on a given curve were defined without reference to any particular unit of measurement. They are line segments. Leibniz of course, like Descartes, wanted to introduce quantification, and analyze the properties of curves algebraically, but since the definition of the functions is geometric, he could postpone the choice of a unit until a convenient one could be found with respect to particular properties of the curve at hand. Many illustrations of the advantage of this will be presented in latter discussions (e.g. Sections 2.4 and 2.12).

Since angles $\angle TPN$, $\angle POT$, and $\angle PON$ are all right angles, the triangles ΔTOP, ΔPON, and ΔTPN are all similar. This configuration will be familiar to geometers as the construction of a geometric mean between ON and OT, the mean being OP (i.e. the circle having diameter $\overline{TN}$ will pass through P). We have already seen how Apollonius employed such a configuration but more examples shall soon emerge.

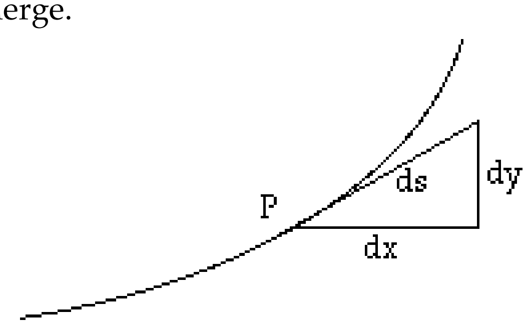


Figure 2.3b

Inspired by the work of Pascal, Leibniz saw a fourth triangle which was similar to the three mentioned above (Child, 1920; Arnol'd, 1990; Edwards, 1979). This was the infinitesimal, or characteristic triangle (see figure 2.3b), used by Pascal to integrate powers of the sin curve (see section 2.14). Leibniz viewed a geometric curve as made up of infinitely small line segments which each have a particular direction. He perceived the utility of this concept in Pascal's work, and it became one of the primary notions in his development of a system of notation for calculus. Although many modern formal mathematicians avoid this conception, it is still used as an important conceptual device by engineers. Figure 2.3b still appears in calculus books because it conveys an important meaning, especially to those who use calculus for the analysis of physical or mechanical actions.

Leibniz saw great significance in the triangles of Figure 2.3a because they were large and visible yet similar to the unseen characteristic triangle. This finding of large triangles which are similar to infinitesimal ones is a theme that runs through many of the most important works of Leibniz (Childs,1920; Edwards, 1979) and we shall investigate several examples in detail (e.g. Sections 2.4 and 2.14). From Figures 2.3a and 2.3b, the similarity relations tell us that: $\frac{dy}{dx}=\frac{PO}{OT}=\frac{ON}{PO}$. Leibniz considered ratios like this, as well, to be functions of the point P along the curve.

2.4 Drawing Parabolas Using the Focus / Directrix Method

Let us look at how this system of Leibniz works in the case of the parabola. We must first have a way to draw a parabola. Everything begins with the construction of a curve. Figure 2.4a shows a linkage which will draw parabolic curves (Schooten, 1657, p.357). This figure comes from the work of Franz van Schooten (1615 - 1660) whose extensive commentaries on Descartes' *Geometry* were widely read in the seventeenth century. Because his works supplied many of the details omitted by Descartes, they were more popular than the *Geometry* itself (Van Maanen, 1992).

The curve drawing device in Figure 2.4a comes from a set of commentaries on Apollonius by van Schooten (1657). It is built entirely from rigid, hinged and slotted rods. Such mechanical devices are referred to as linkages. Such devices can be built to draw any algebraic curve (Artobolevskii, 1964). Book 4 of van Schooten's work is entirely dedicated to linkages which draw conic sections (subsequent sections will display more figures from this work). My analysis of this linkage will take place within the framework of the functional system of Leibniz which emerged in the 1690's. I will also employ one observation from the beginning of the next century that comes from the early work of L. Euler (the greatest popularizer of the work of Leibniz). The curve drawing devices of van Schooten and others were carefully studied by mathematicians in the seventeenth century, including Leibniz. This example will illustrate the interaction between grounded activity and the systematic inquiry that followed.

Van Schooten's Book 4 (1657) has the title *Sive de Organica Conicarum Sectionum in Plano Descriptione* . Loosely translated this is *Organs for the description of conic sections in the plane.* Throughout the seventeenth century mechanical devices, such as the one in Figure 2.4a are described as "organic." Section 2.11 will look at Newton's tract on the organic construction of curves (1968). The word "organic" comes from the Latin word "organ" which refers to an instrument or device for a specific purpose, as in the musical instrument the "organ." The mechanical world view that was expounded in the seventeenth century scientific revolution extended the use of this word to parts of living bodies which were seen as machines for specific purposes. This led to our modern use of the word "organic" which now carries quite a different connotation except, perhaps, to those Cartesian scientists who still want to view the world as entirely mechanical (Merchant, 1980).

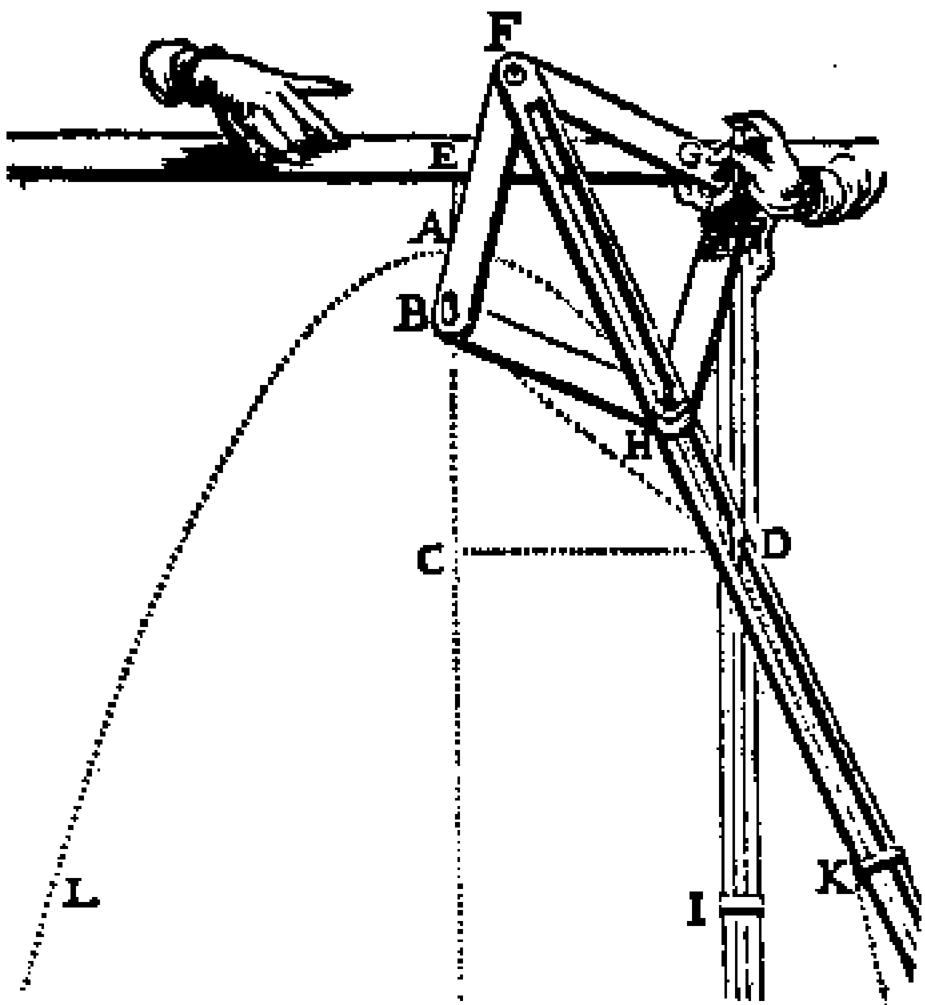


Figure 2.4a

The apparatus in Figure 2.4a constructs the parabola from the familiar focus/directrix definition. That is, the parabola is the set of point equidistant from a point and a line. The fixed ruler $\overleftrightarrow{GE}$ is the directrix and the point B is the focus. Four equal length links create a movable rhombus, $BFGH$, pinned down at B, which guarantees that $\overleftrightarrow{FH}$ will always be the perpendicular bisector of $\overline{BG}$, as G moves along the directrix $\overleftrightarrow{GE}$. $\overleftrightarrow{GI}$ is a movable ruler which is always perpendicular to the directrix $\overleftrightarrow{GE}$. The point D is the intersection of $\overleftrightarrow{FH}$ and $\overleftrightarrow{GI}$, as the point G moves along directrix. Hence at all positions $BD = GD$, and hence D traces a parabola with focus B and directrix $\overleftrightarrow{GE}$. The motion of G along the fixed horizontal line "directs" the drawing of the curve, hence the term "directrix."

This construction can be simulated on a computer using the software *Geometer's Sketchpad*. This software allows one to define a perpendicular bisector so the rhombus becomes unnecessary. One can either drag a point along the directrix or have the computer animate such a motion. Figure 2.4b was made using this software. The point F is the focus, and the point S is moving along the directrix. $\overleftrightarrow{BP}$ is the perpendicular bisector of $\overline{FS}$, $\overline{SP}$ is always perpendicular to the directrix, and the intersection point P traces a parabola.

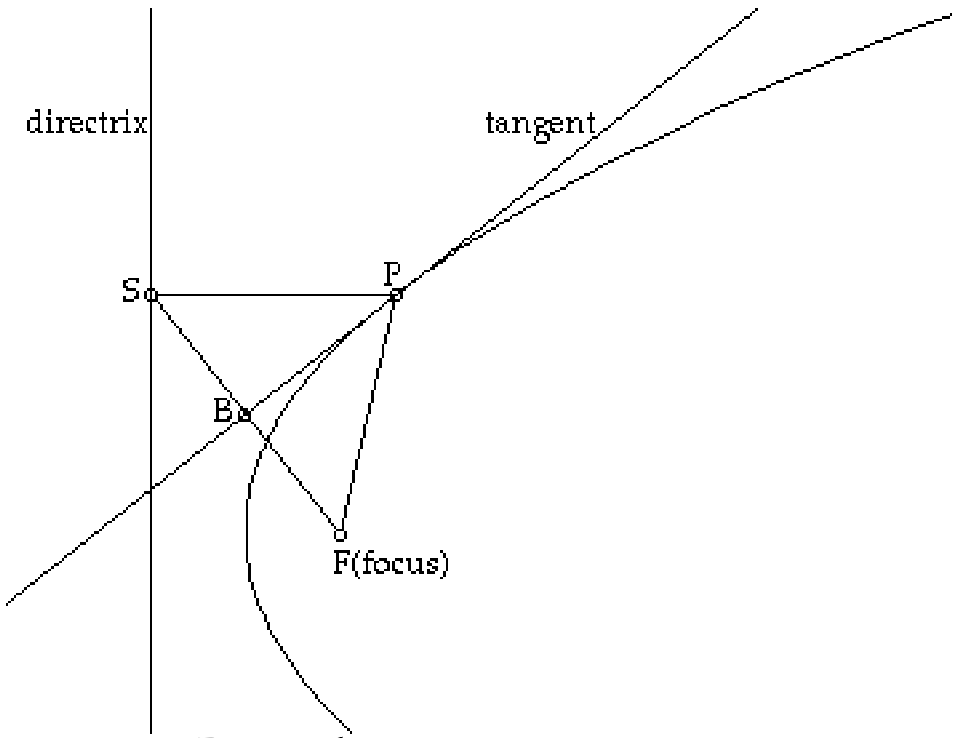
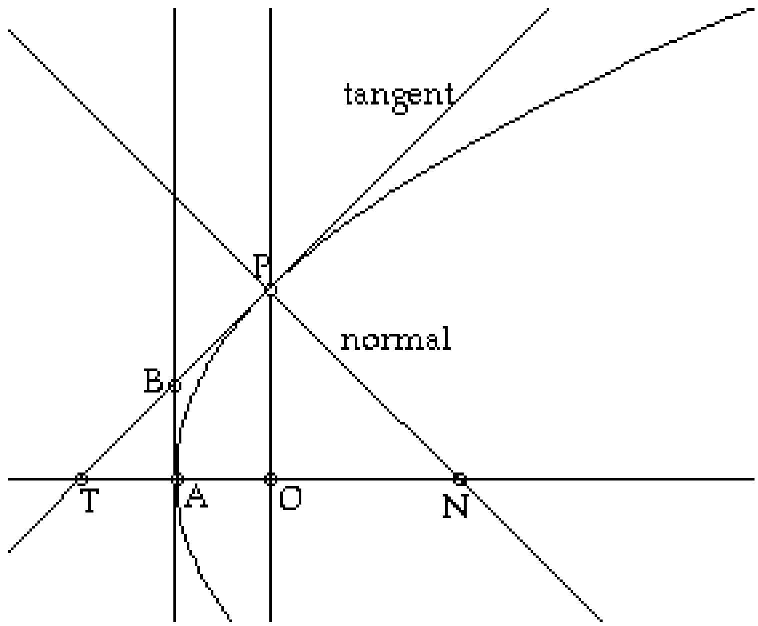


Figure 2.4b

One consequence of this construction that is immediately apparent to the eye is that at each point, $\overleftrightarrow{BP}$ is the tangent line to the curve at P. Curves can often be drawn by constructing a series of tangents to the curve. Such methods are sometimes called "envelope" constructions. This can also be done using strings or paper foldings (Row, 1966; Gardner, 1989). In order to fold a parabola as in Figure 2.4b, let one edge of a sheet of paper be the directrix and mark any point as the focus. Make a series of folds each of which brings some point on the directrix onto the focus. These folds will then be the perpendicular bisectors of the segments between these pairs of points and each one will be a tangent to the parabola. A series of these folds will outline the parabola. Using waxed paper will produce nicely visible folds.

In Figure 2.4c, I display the functions of Leibniz using the axis of symmetry of the parabola as the axis for abscissas and the vertex, A, as a starting point. Since the tangent line is part of the construction this can be readily accomplished with

Geometer's Sketchpad. It is impossible to convey the feel of this moving construction on paper, but Figure 2.4c shows the six functions of Leibniz for two different positions of P. The focus and directrix have been hidden in these figures so that one can concentrate on how the "Leibniz configuration" changes. A is the vertex of the parabola and B is same as in Figure 2.4b. The other labels are consistent with functional definitions given in Figure 2.3a.



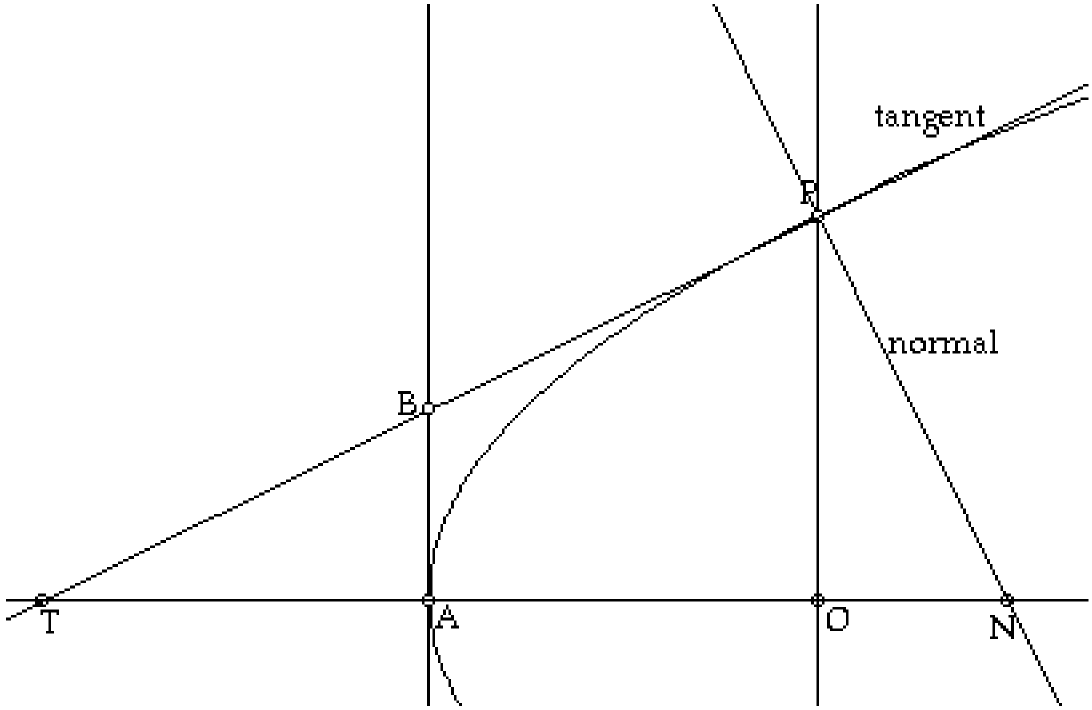


Figure 2.4c

What can be seen by watching the six functions in this dynamic setting? Perhaps it is not so easily seen from these static figures, but with the figure in motion and using color to highlight the six functions, two invariances become readily apparent. The first that most people notice is that the subnormal, $\overline{ON}$, has constant length. The second is that the vertex, A, is always the midpoint of the subtangent, $\overline{OT}$, for points O and T can be seen to approach and recede from point A symmetrically. These two invariances can be deduced from the geometry of the construction, but of greater significance is that they can be experienced empirically from the action of the construction. *Geometer's Sketchpad* allows for confirmation of ones visual experience by turning on meters which monitor these lengths empirically. Sure enough, $\overline{ON}$ has constant length, and the length of $\overline{AT}$ is always equal to the length of $\overline{AO}$.

Postponing for a moment the geometrical proofs of these two statements, let us first look at what they tell us about the parabola. In the tradition of Descartes, variables are introduced after the curve is drawn. Let $x = AO$, and let $y = PO$, i.e. x is the length of the abscissa, and y is the length of the ordinate. Since triangles ΔTOP and ΔPON are similar, $\frac{PO}{OT} = \frac{ON}{PO}$. Since A is the midpoint of $\overline{OT}$, this becomes $\frac{y}{2x} = \frac{ON}{y}$, or $(2 \cdot ON) \cdot x = y^2$. Since ON is constant, this yields the equation of the parabola. The constant length $(2 \cdot ON)$ is known in geometry as the *latus rectum*, i.e. the rectangle formed by x and the latus rectum is always equal in area to the square on y. Being still free to choose a unit, one could choose $ON = \frac{1}{2}$. The equation then becomes $x = y^2$.

Using the similarity between the characteristic triangle and triangle TOP, one obtains: $\frac{dx}{dy} = \frac{OT}{PO} = \frac{2x}{y} = 2y$. Hence both the equation and the derivative can be found from considering the invariant properties of Leibniz's configuration under the actions which constructed the curve.

The choice of $ON = \frac{1}{2}$ gave the equation and derivative of the parabola in their best known form, but this is perhaps a little artificial from the geometric standpoint. The subnormal ON is the primary invariant of this curve-drawing action and can be seen as the natural choice of a unit for this curve. As it turns out, the subnormal ON is always equal to the distance between the focus and the directrix of the parabola. Thus it is a natural unit. Using the subnormal as a unit, the equation of the parabola becomes $x = \frac{y^2}{2}$, i.e. the common integral form of the parabola as the accumulated area under the line $x = y$. It is in this form that the parabola most often appears in the table interpolations of John Wallis and Isaac Newton (Dennis & Confrey, 1993).

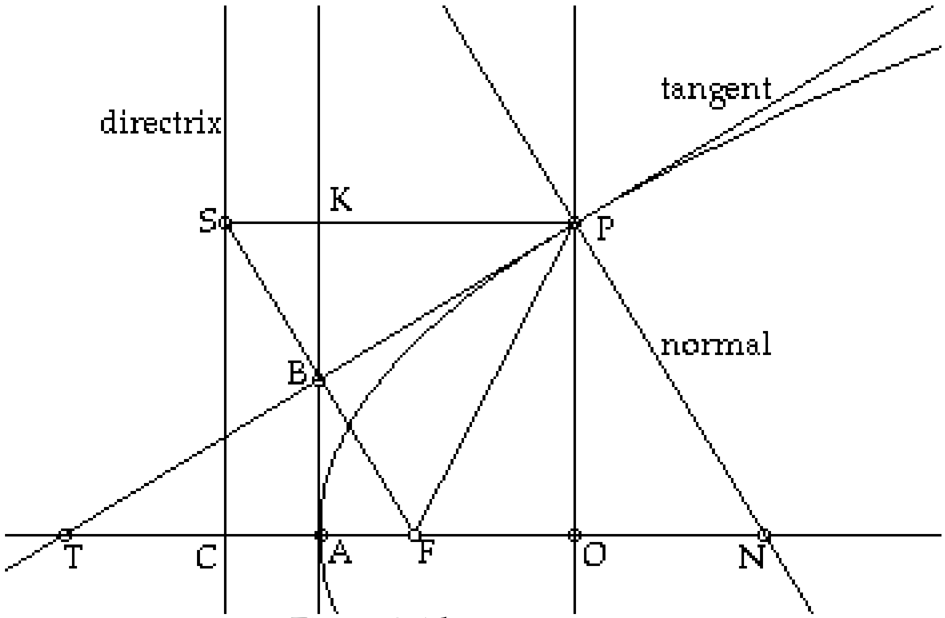


Figure 2.4d

One way to prove that the subnormal is always constant is to show that it always equals the distance between the focus and the directrix. Looking at Figure 2.4d, one sees that $\overleftrightarrow{SF}$ and $\overleftrightarrow{PN}$ are both perpendicular to $\overleftrightarrow{BP}$, so triangles ΔSCF and ΔPON are congruent; hence $ON = CF$.

In order to prove that the vertex A is always the midpoint of the subtangent $\overline{OT}$, one can establish that triangles ΔTBA and ΔPBK are congruent. They are clearly similar, but since B is the midpoint of $\overline{SF}$ it is also the midpoint of $\overline{AK}$, so they are congruent. Hence $TA = KP = AO$.

Lastly, one might ask: how can we be sure that the line $\overleftrightarrow{BP}$ is always tangent to the parabola? That is to say, how can one be sure that each instance of the line $\overleftrightarrow{BP}$ intersects the parabola in only one point? Let $Q \neq P$ be a point on $\overleftrightarrow{BP}$, and let R be the foot of the perpendicular from Q to the directrix $\overleftrightarrow{CS}$. Since R is the closest point to Q on the directrix, $QR < QS$. Since $\overleftrightarrow{BP}$ is the perpendicular bisector of $\overline{SF}$, $QS = QF$. Hence $QR < QF$ and Q cannot be on the parabola, being closer to the directrix than to the focus. Thus all such points Q lie on the same side of the parabola.[16] One could also check the tangency of $\overleftrightarrow{BP}$ analytically by writing the equation of the parabola and the line $\overleftrightarrow{BP}$ using the same coordinate system, and then solving the two equations simultaneously arriving at a quadratic equation with one repeated root. This is the method that Descartes developed for finding tangents, i.e. tangency occurs when repeated roots appear in the simultaneous solutions.

Having a method for drawing parabolas with their tangents at any point allows one to empirically investigate more of the properties of the curve mentioned by Apollonius. Draw a diameter of a parabola (i.e. any line parallel to the axis of symmetry) through some point A on the curve (see Figure 2.4e). Using the tangent at A, construct an ordinate $\overline{OP}$ with respect to this diameter (i.e. a line parallel to the tangent at A). Check that $OP = OP'$. Now draw the tangents at P and P' and discover that they both intersect the diameter at the same point T, and that $TA = AO$. Using various points P on the curve one discovers that this is always true. The property of the subtangent used in our previous discussion generalizes to any diameter and its ordinates. Any subtangent $\overline{OT}$ is always bisected by its vertex A even in the skewed coordinates of any Apollonian diameter.

---

[16] This argument for tangency is similar to those made by Apollonius. A reader might ask what definition of "tangent" is used in Apollonius? The answer is that he did not give a precise definition. He used the phrase "a line that touches a curve" in contrast to the phrase "a line that cuts a curve." A proof of tangency in Apollonius asserts that a line intersects a curve in only one point, and stays on the same side of the curve.

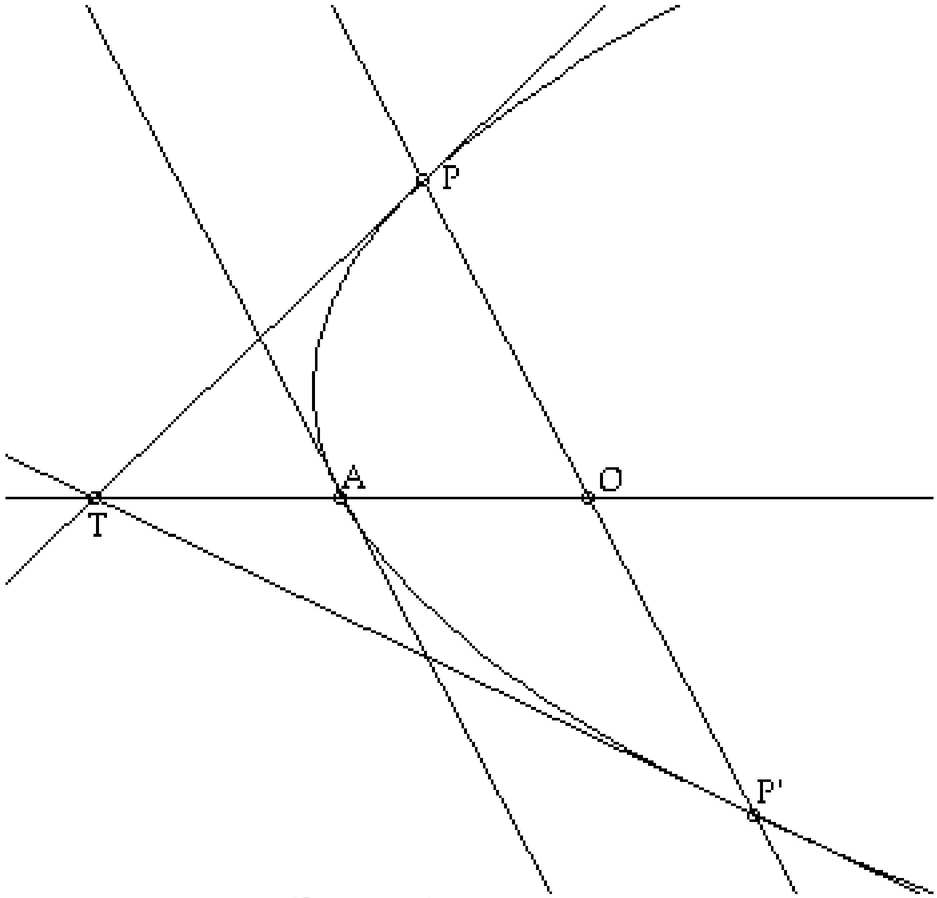


Figure 2.4e

In our previous discussion we used the bisected subtangents on the axis to find the equation of the parabola. In the more general case shown in Figure 2.4e, Apollonius made the argument in the other direction, deducing the general subtangent property from his knowledge of the ratios which imply the equation. It is this going back and forth between geometry and algebra that I find most suggestive for curriculum reform. From the sectioning of an oblique cone parallel to its side Apollonius knows that AO is proportional to the square on OP. In Book 1, Propositions 33 &35, he shows how this proportionality implies that $TA = AO$, and that constructing T such that $TA = AO$ is another way to find the tangents at P and P' from the curve. The arguments are simple but require a sharp sense of ratio. He used the theorem from Euclid that if a line segment is cut into two pieces which are then used as the sides of a rectangle, the largest area that the rectangle can have comes from bisection (Apollonius, 1952).

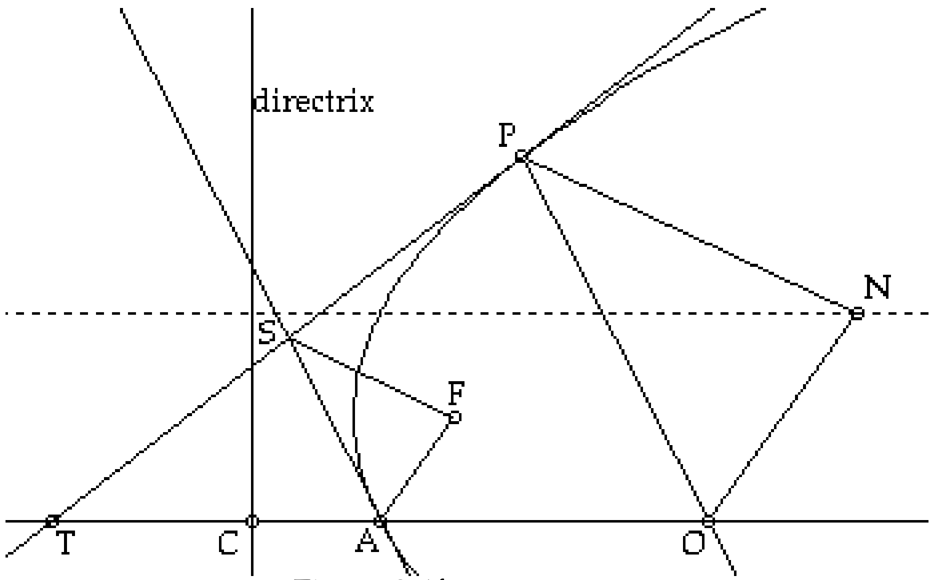


Figure 2.4f

As for the constant subnormal of a parabola along its axis of symmetry, here is my own generalization of this property to any Apollonian diameter done on *Geometer's Sketchpad*. Construct at each point P, a triangle ΔPON which is similar to ΔTOP (see Figure 2.4f). Although point N is not on the diameter $\overline{AO}$, as P moves along the parabola, $\overline{ON}$ will have a constant length, and N will move along a line (dotted) parallel to $\overline{AO}$ and equidistant from the focus F. Since the tangent at A bisects angle $\angle CAF$, $\overline{ON}$ will always be parallel to $\overline{AF}$, and twice as long as $AF = CA$, since triangle ΔAFS is similar to ΔONP and exactly one half as large. Although this configuration does not use right triangles as in the Leibniz configuration, the similarity still guarantees that the segment OP is the geometric mean of OT and ON, just as AS is the geometric mean of AT and AF. A fascinating special case occurs when P is positioned so that $\overline{PO}$ passes through the focus F. In that case T will fall on the directrix at C, and $OT = PO = ON$, i.e. this geometric sequence is constant. The geometric sequence is increasing on one side of this special case, and decreasing on the other.

Since triangles ΔTOP and ΔPON are similar, one obtains the equation in this skewed coordinate system as before from: $\frac{PO}{OT} = \frac{ON}{PO}$, or $\frac{y}{2x} = \frac{ON}{y}$, or $(2 \cdot ON) \cdot x = y^2$. The latus rectum in this coordinate system is still $2 \cdot ON$. As with the generalized bisection property of the subtangent, one can either deduce these properties directly from the geometry and use them to obtain the generalized equation, or, vice versa, deduce the these geometric invariances from the generalized equation obtained from the oblique conic section.

These two invariant properties of the parabola were never mentioned directly (so far as I know) in the published work of Leibniz. The fact that the vertex is the midpoint of the subtangent was demonstrated by Apollonius (1952). The fact that the subnormal is constant is credited to L. Euler, who expanded and popularized the ideas of Leibniz (Coolidge, 1968). They both appear in Book 2 of Euler's most famous textbook, the *Introduction to Analysis of the Infinite* (1990). This book, first published in 1748, was the first modern precalculus textbook and, along with its sequels on differential and integral calculus, did much to standardize curriculum and notation. Nearly all of the topics in our modern precalculus books are contained in Euler's book, but what is missing from our modern treatments is the bold empirical spirit of Euler's investigations, as well as most of his more advanced geometry and infinite series. Euler says in the preface to his text that he presents many questions which can be more quickly resolved using calculus. He insists, however, that students are rushing into calculus too rapidly, and that they will become confused because they lack the experiential basis (both geometric and algebraic) upon which calculus is built.

Already this one parabola example demonstrates how much can found using only basic geometry combined with empirical investigation. By letting the geometry move, one creates the situation where equations evolve naturally from geometry and vise versa. Too often in our schools we find our geometry curriculum static and isolated from other topics, especially algebra. Parabolas, for example, are generally not discussed in geometry, but only introduced as the graphs of quadratic polynomials. Two-column geometry proofs provide a shadow of Euclid, but they can not provide the dynamic experience that leads to an understanding of the relations of curves to equations, functions and calculus.

2.5 Drawing Ellipses from their Foci

In the previous sections I have presented two dynamic ways to draw parabolas, but how can one be sure that the two devices are really drawing the same family of curves? The ratios between the abscissas and ordinates provided the answers in the previous case and connected these devices to the parallel slice of the cone as well. Parabolas are all similar to each other (Apollonius, 1952), so they are all the same except for possible variations in one parameter (the latus rectum). As I mentioned in section 2.2, two parameters are at play in the elliptic case. An interesting introduction to ellipses might lie in a student discussion of how to tell when two ellipses are similar. In Section 3.6 a student brings up this issue on his own.

In this section I shall present the first examples of transformations on the curve drawing devices themselves. Three different devices for drawing ellipses will be presented, along with demonstrations that all three devices draw the same family of curves, via direct observations of the devices. No ratio properties or algebraic equations will be used. I shall show sequentially how to see each device acting within another device, and conclude directly that certain points in different linkages will have the same paths of motion. Take out a link here, replace it with another one there. I will use the direct tinkering approach that is at the heart of van Schooten's work (1657).[22]

I begin with the one curve drawing device (other than the compass) that still remains in our schools. That is the loop of string over two tacks, which draws an ellipse from its two foci by holding constant the sum of the two distances to the foci. Other than the circle, an elliptic equation derived from this device is usually the only instance, in our secondary curriculum of an equation being derived from a curve drawing action. Figure 2.5a shows van Schooten's illustration of such a device, with the foci tacked at *H* and *I* (1657, p. 326).

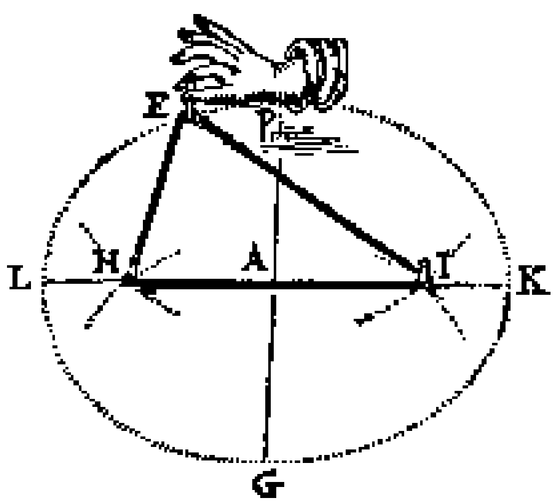


Figure 2.5a

The standard equation of an ellipse can be derived from this device using the Pythagorean theorem (i.e. the distance formula). This method gives an elliptic equation only with respect to the axis through the foci, and does not readily generalize that equation to an arbitrary diameter, as did the Apollonian approach via geometric means from the cone (see Section 2.2). Apollonius did prove (1952, Book 3, Prop. 52) that there

[22] This approach provides an example of a systematic inquiry taking place within a set of grounded activities with almost no reference to any system of codification (Comment by Jere Confrey).

are two points within an ellipse such the sum of the distances to any point on the curve is constant (and equal to the length of the axis, $HE + IE = LK$), but he did not give any special name to those points.

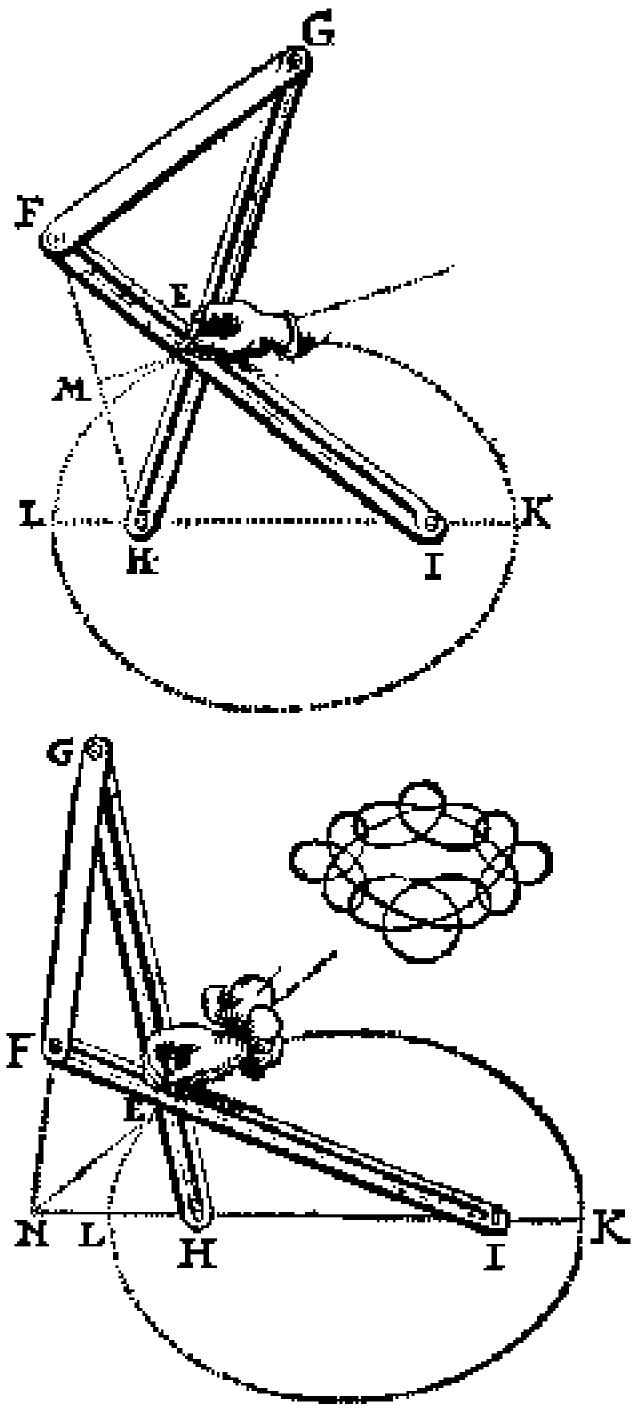


Figure 2.5b

Figure 2.5b shows two positions of a cross linked elliptic device (Schooten, 1657, p. 340, 341). As before the links are tacked down to the table at the points H and I. The requirements here are that the link FG be equal in length to the distance between the foci HI, and that the other two links, HG and IF, both be larger than HI, and equal to each other. I have included a double figure of this device to give some sense of its motion, since it was at first mysterious to me. While talking about this device one evening with a geometrically astute carpenter, he said that he felt sure that such a device would be rigid, since cross bracing was the standard technique in carpentry to achieve rigidity. We had to tack some scrap lumber together before he would believe that any motion was possible with such a device. A version of this device can easily be made from cardboard and thumb tacks (pointing down at H and I, and up at F and G). It is worthwhile experiencing its motion.

From the requirements of the device it can be seen that triangles ΔHEI and ΔFEG are always congruent at every position. Thus any point, E, on the curve is equidistant from H and F, and also from I and G. Hence the sum of the distances from E to the foci is constant, $HE + IE = HG = IF$. This device draws the same curves as the loop of string, and the equal links HG and FI are both equal to the length of the major axis LK. In general the study of three links which connect two stationary points is of great interest in mechanical engineering and robotics. If one violates some of the equalities in the requirements for this device, interesting families of fourth, sixth, or even eighth degree curves result (see Artobelevskii, 1964).

Besides having a very different feel from the loop on string this cross linked device tells us more about the ellipse. It gives us the tangent line at any point. Van Schooten's two drawings (Figure 2.5b) display two ways to find the tangent at the point E. Either draw $\overleftrightarrow{EM}$, where M is the midpoint of $\overline{FH}$ (top of Figure 2.5b), or extend $\overline{FG}$ until it meets the axis at N and then draw $\overleftrightarrow{EN}$ (bottom of Figure 2.5b). Either method will find the tangent line to an arbitrary point on the ellipse. The symmetry of the lines in Figure 2.5b will yield this, but I will postpone any proof of tangency until I transform this device into a third elliptic device.

Regardless of tangency, $\overleftrightarrow{EM}$ (or $\overleftrightarrow{EN}$) is clearly a line of symmetry for the links in Figure 2.5b, that is to say triangle ΔFEG is a reflection through $\overleftrightarrow{EM}$ (or $\overleftrightarrow{EN}$) of triangle ΔHEI. Hence $\overleftrightarrow{EM}$ (or $\overleftrightarrow{EN}$) bisects line segment $\overline{IG}$ (not drawn in Figure 2.5b). Van Schooten used this property to construct a new device which draws the curve by finding the point E as the intersection of $\overleftrightarrow{HG}$ with the perpendicular bisector of $\overline{IG}$. This new device constructs the ellipse as the set of points equidistant from a point, I, and a circle (i.e. the circle centered at H, and described by G as the link $\overline{HG}$ rotates around H). Van Schooten's drawing of this device is shown in Figure 2.5c (1657, p. 342). As in the parabolic device in Figure 2.4a, he used a flexible rhombus ($IOGP$) as a way to insure that $\overleftrightarrow{OP}$ will always be the perpendicular bisector of $\overline{IG}$, and hence $IE = EG$. The link $\overline{HG}$ is the same as in Figure 2.5b, but all the other links are new. The four links $\overline{IO}$, $\overline{OG}$, $\overline{GP}$, and $\overline{PI}$ can be any length as long as they are all equal. Their length does not affect the curve, which is completely determined by HI (focal distance) and $\overleftrightarrow{HG}$ (major axis). This device still draws an ellipse because $HE + IE = HG$ still holds.

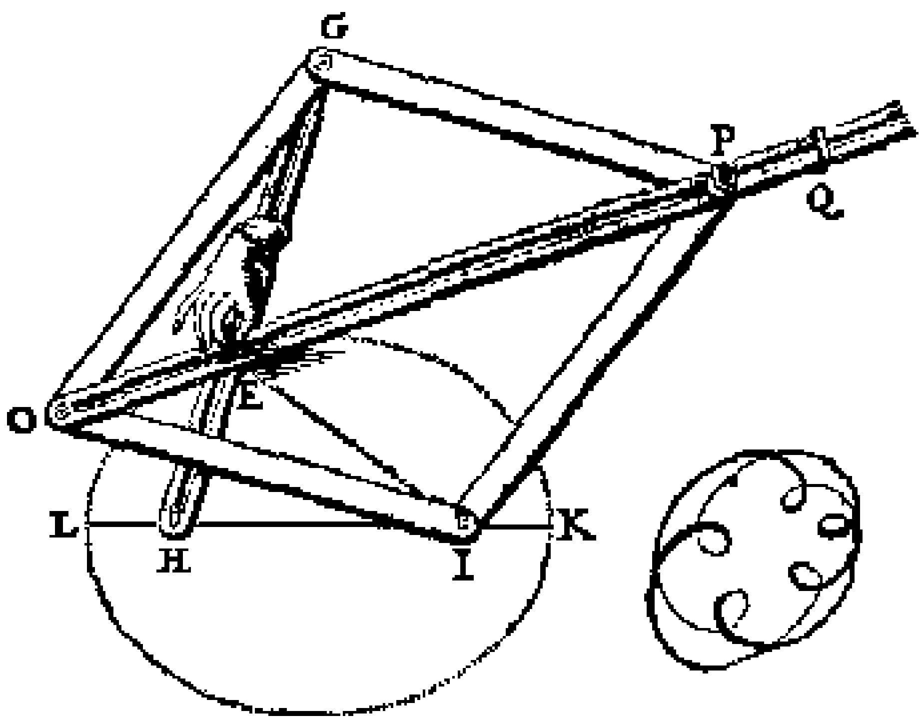


Figure 2.5c

Unlike the device in Figure 2.5b, this new physical device is not so easily constructed, and its motion is cumbersome. This new device, however, is much easier to simulate in *Geometer's Sketchpad* than the previous one. Start with a point G that moves on a circle centered at H (see Figure 2.5d). Choose any point, I, in the interior of the circle and connect I to G. One can dispense with the rhombus by directly defining the perpendicular bisector of $\overline{IG}$ (A is the midpoint of $\overline{IG}$). As the point G rotates around the circle, trace the locus, E, of the intersection of the radius $\overline{HG}$ with the perpendicular bisector of $\overline{IG}$. This point, E, traces an ellipse which is the set of points equidistant from the circle and the point I. The radius of the circle is the length of the major axis. The closer the point, I, gets to the circle the more elongated the ellipse becomes.

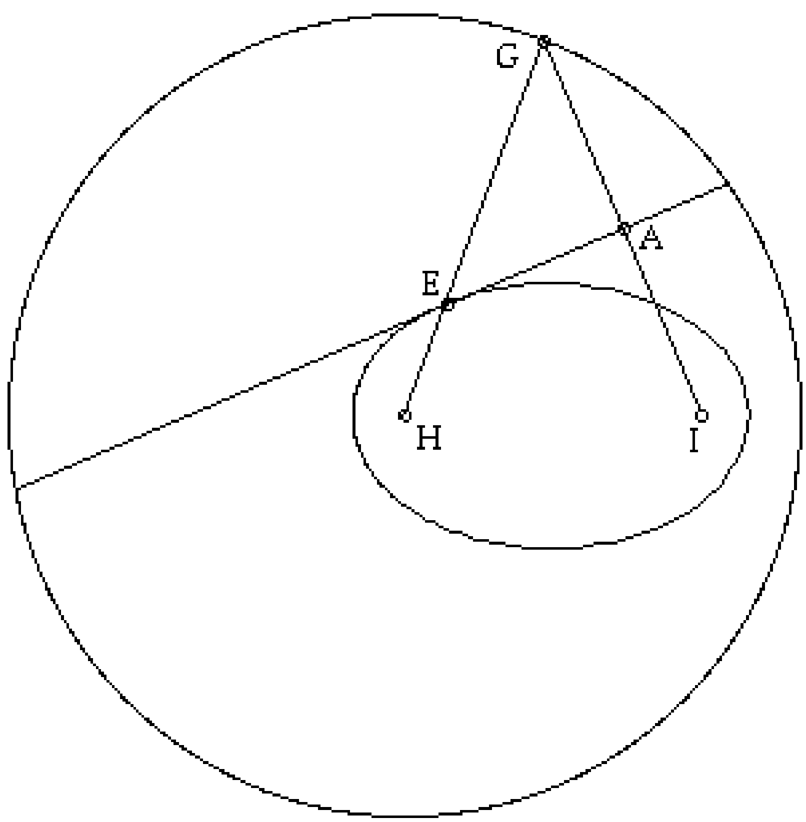


Figure 2.5d

Like the construction of the parabola in Section 2.4, this is an envelope construction. That is to say $\overleftrightarrow{AE}$ is always tangent to the curve at E. The proof of tangency that was given for the parabola remains valid here. Let $Q \neq E$ be a point on $\overleftrightarrow{AE}$, hence $QG = QI$. But since Q is not on the radius $\overline{HG}$, the distance from Q to the circle is less than QG. Hence the line $\overleftrightarrow{AE}$ intersects the ellipse only at E, and all points $Q \neq E$ on the line $\overleftrightarrow{AE}$ must be outside the ellipse.

Like the parabolic envelope construction of Section 2.4, this curve can also be constructed by paper folding. Cut out a circular piece of paper and mark any point, I, not equal to the center. By folding points on the edge of the circle so that they fall on I, one constructs the perpendicular bisectors $\overleftrightarrow{AE}$ which are tangent to the ellipse. These folds will outline the ellipse (Row, 1966; Gardner, 1989).

The equation of this ellipse with respect to the axis $\overline{HI}$, is known from the loop of string construction from which this construction evolved (i.e. $HE + IE = HG$). From my comments on geometric means in section 2.2, this equation connects these three devices back to the definition of the ellipse as a conic section. There is also a beautiful and direct geometric proof of this connection given in the nineteenth century by Dandelin, the so called "ice cream cone proof" (see Apostol, 1961).

Seeing the ellipse as the set of points equidistant from a point and a circle provides another unified way to see all conic sections. In the next section I shall show that the hyperbola can be seen this way as well, simply by placing the focus, I, outside the circle. The parabola is then seen as the border between these two cases where the directrix is seen as a circle of infinite radius, thus sending the other focus of the parabola (the center of the circle) to infinity. In the seventeenth century some mathematicians adopted this use of infinity to unify the conics (e.g. Pascal and Newton), while others followed Apollonius more closely and avoided it (e.g. Descartes and Van Schooten).

Having a direct way to draw tangents to ellipses allows one to empirically investigate a variety of their properties. The theorems of Apollonius on ordinate directions and on conjugate diameters can all be investigated directly (Figure 2.2e was made this way). If one uses the distance meters, and tabulations that are available on *Geometer's Sketchpad*, one can explore the standard equations of the ellipse along different diameters using oblique coordinates. Such data could then be further explored and analyzed by placing it in a spreadsheet or table setting like the one available in the software *Function Probe*, where a variety of rate and accumulation properties could be explored in a precalculus setting based on a direct empirical investigation of a geometric action.

Considering again the "functions of a curve" as defined by Leibniz, let us look at the ellipse. In the previous section, we saw that in the parabolic case, the vertex of any diameter was always the midpoint of the subtangent (i.e. the subtangent was always twice the abscissa). Can any such statement be made about the ellipse? Looking at some examples using the major axis, one can begin comparing the lengths of abscissas with corresponding subtangents. One finds that for ellipse, the vertex is always closer to the foot of the ordinate than it is to the foot of the tangent (see Figure 2.5e where $OL < TL$). This says, in a very physical way, that motion along an ellipse away from a vertex L curves inward towards the axis more sharply than does parabolic motion.

OK	2.487	2.766	3.315
OL	1.166	0.887	0.338
TK	6.879	5.376	4.067
TL	3.227	1.723	0.414
OK/OL	2.132	3.120	9.816
TK/TL	2.132	3.120	9.816

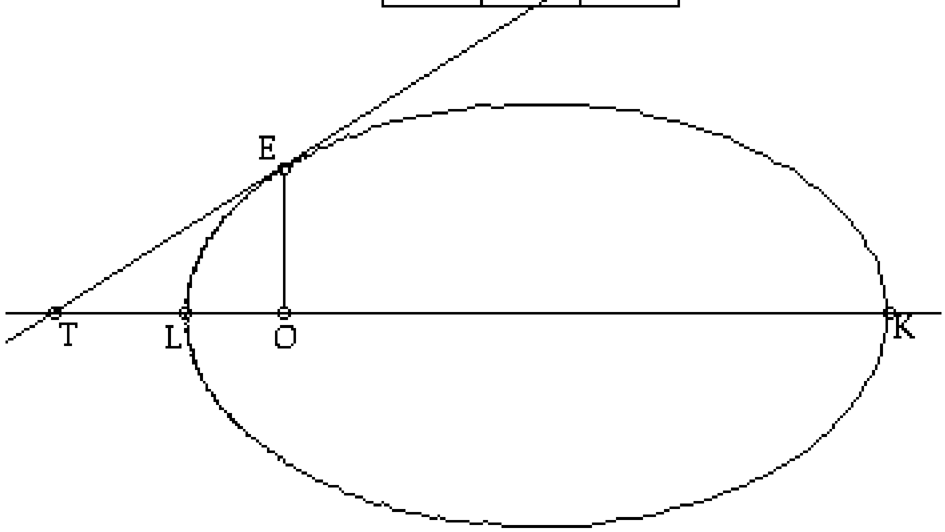


Figure 2.5e

There is an exact invariance here, and Apollonius considered it in Book 2 of *The Conics* (1952). The table in Figure 2.5e shows the lengths of four segments in inches together with two ratios, all for three different positions of E along the curve. *Geometer's Sketchpad* will display meters which continuously monitor lengths, and will automatically tabulate the readings of its meters in a table for different positions of a configuration. The ellipse and its tangents are being constructed as in Figure 2.5d with some of the lines and circles from the construction being hidden.

One sees in the table that T and O have the same ratios of distance to the vertices L and K at the ends of the axis. T is said to be the "harmonic conjugate" of O because T divides the major axis $\overline{LK}$ externally in the same ratio as O divides the axis internally. Strings of length TL and TK would play the same musical interval as strings of length OL and OK (given equal tension). Just as in the parabolic case of bisected subtangents, this harmonic property is true for any diameter and its ordinate direction. Since the equations in all these systems are same (up to proportional constants) algebraically this must be the case. It is fascinating to investigate this harmonic property in these oblique cases, but I leave that investigation for the reader.

One important use of ellipses is as a model for the orbits of planets. Most students are told in high school science classes that the planets follow elliptic orbits around the sun (Kepler's first law), and usually they are also told that a line between a planet and the sun will sweep out equal areas in equal times (Kepler's second law). These laws are rarely discussed in secondary mathematics classes. Even in calculus classes, Newton's conclusions that such motions are the result of a Force inversely proportional to the square of the distance from the sun are rarely discussed. Newton is popularly believed to have invented calculus, although what appears in calculus books is the work of Leibniz and his followers (e.g. Euler and Bernoulli). In fact, what appears in Newton's *Principia* is a lot of geometry that makes heavy use of the work of Apollonius, particularly conjugate diameters and their relations to the tangents of conic sections.

The previous construction of the ellipse contains a very direct way to see the velocity vector of a planet in an elliptic orbit. Let the point E represent a planet orbiting around the sun at point I, where the orbit is defined as the set of points equidistant from I and a circle centered at the other focus H (as before). We found E by intersecting the perpendicular bisector of $\overline{IG}$ with the radius $\overline{HG}$ as G moved around the circle. Now extend $\overline{IG}$ back in the other direction until it meets the circle at another point Q (see Figure 2.5f). This construction contains the tangent to the ellipse at E; the question is: what is the magnitude of the velocity vector at E? It turns out that, using Kepler's model, that magnitude is proportional to the length of $\overline{IQ}$ (the proportionality constant depends on one's choice of a unit for time). In Figures 2.5f and 2.5g, I have made the vector $\overrightarrow{ET}$ the same length as $\overline{IQ}$. These figures give a vivid sense of how a planet speeds up when it is closer to the sun and slows way down when in the more distant part of its orbit. These figures are two moments of an animation made on *Geometer's Sketchpad*.

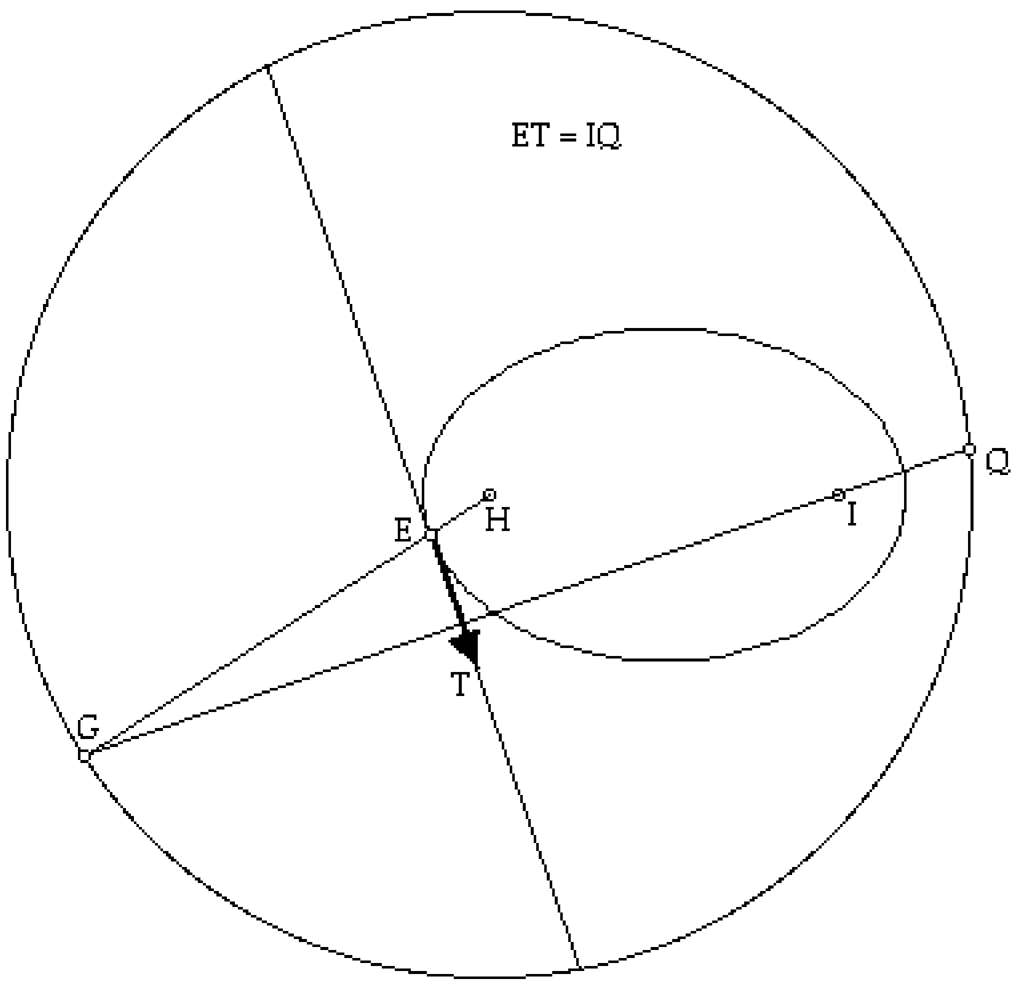


Figure 2.5f

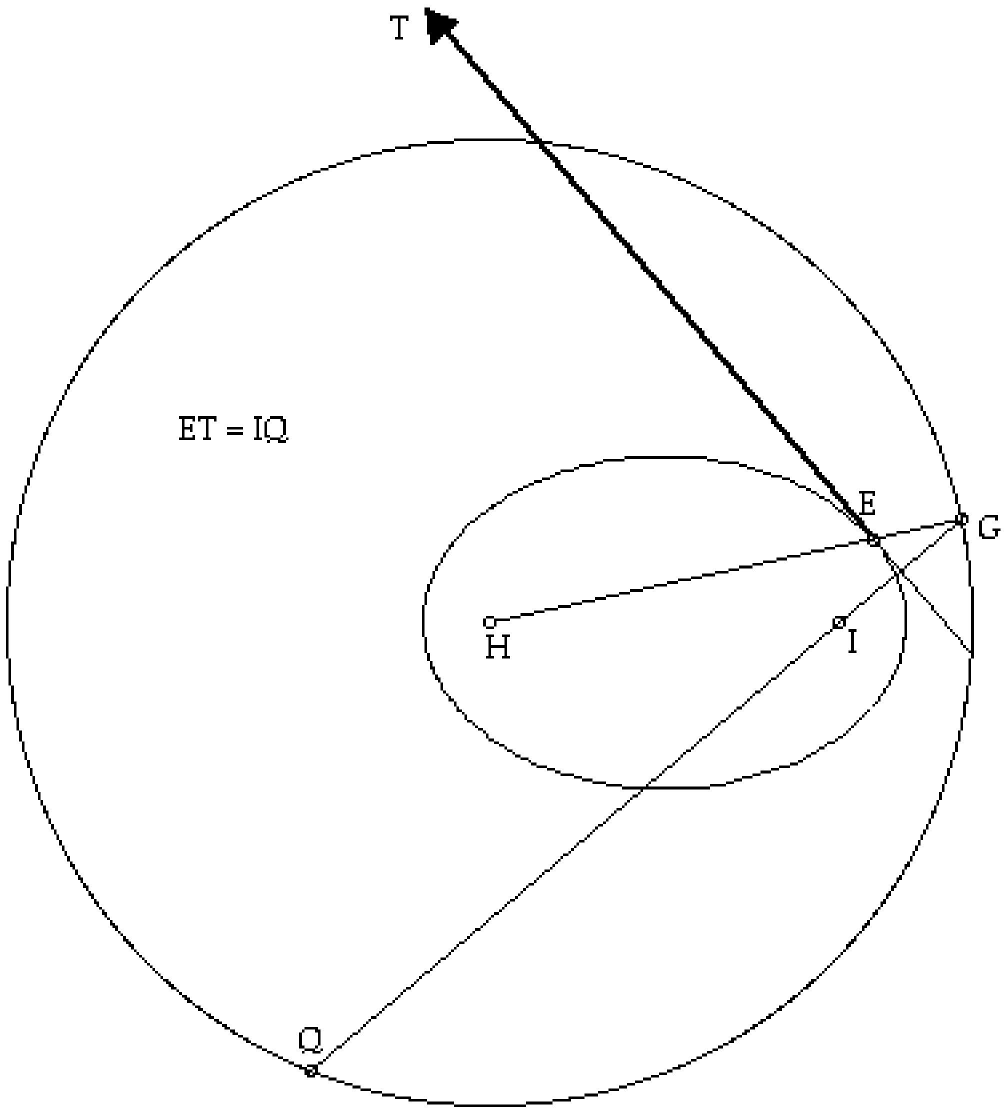


Figure 2.5g

For those who wish to work out on their own why a velocity vector proportional to *IQ* will produce orbital motion in accord with Kepler's second law, I offer the following helpful thought from Euclid (Book 3, Prop. 35). Since the point *I* is fixed, the product of *IQ* and *IG* will remain constant as *G* goes around the circle. For those who wish to see more details on the relations between this curve drawing device and planetary orbits, see the recent article by Andrew Lenard (1994). Lenard provides details and goes on to show how to geometrically construct the acceleration vector, and derive Newton's laws geometrically. It should be pointed out that Newton did not use this particular construction in his Principia (1952), but he certainly studied van Schooten carefully, and wrote commentaries on his work at the very start of his career (Newton, 1967). In his formal presentation of universal gravitation, Newton chose instead a more conservative route and stuck closely to Apollonius.

The early nineteenth century, Irish mathematician Sir William Hamilton seems to have been the first to have pointed out that if the velocity vectors of a planet's orbit are all placed with their feet at the sun, then their heads will trace out a circle (not centered at the sun). He called this circle the hodograph (from the Greek hodos for road or path) (Lenard, 1994). This hodograph is the circle in Figures 2.5f and 2.5g. These curve drawing devices often provide a strong sense of a parameter of motion which guides the construction of curves. In the previous example, it is this hodographic circle, which is geometrically related to the time parameter in an orbiting planet. Newton used the term "directrix" generally, to indicate any curve which directs the construction of another curve (see Section 2.11).

As a curve, the ellipse is easily drawn and is fundamental in a variety of important scientific investigations especially astronomy, yet its discussion in the secondary mathematics curriculum is scant. The empiricism of Kepler and the systematic inquiries of Newton have a mythological status in our society, and yet at the crux of their fame lies elliptic orbits which are rarely discussed even in calculus classes. I think this partly due to the fact that ellipses are curves but they are not "functions" according to the modern definition. The discussion of elliptic orbits is formally and algebraically inconvenient. The usual approach to geometry is as an exercise in logic that never even begins to discuss curves and tangents. It is my hope that this section provides some strong indications for curricular reform using dynamic geometry as a means to facilitate: first, activities in which students could explore ellipses, and second, discussions of orbits which could make the investigations of Kepler and Newton more than techno-sacred iconography.

2.6 Drawing Hyperbolas from their Foci

This section will parallel the previous one to such an extent that I will say less and try to let the figures speak for themselves. Every elliptic device and construction from the previous section can be altered or turned inside out, so that it produces a hyperbola. The properties and demonstrations almost entirely parallel what was given in Section 2.5. The only differences come from the fact that hyperbolas have two branches and a pair of asymptotes. For example, as a point on the curve approaches an asymptote, the angle between the diameter through that point and the corresponding ordinate direction (the conjugate diameter) becomes very small (or close to $180°$). The asymptotes are lines through the center but they are not diameters since at that position the ordinate direction collapses onto the same line. I will discuss coordinatizing a hyperbola along one or both of its asymptotes in Section 2.7. This arises naturally through the use of an entirely different curve drawing device that gives no indication of the foci.

Section 2.5 began with a loop of string over two tacks which held constant the sum of the two focal distances. Here I begin with a ruler hinged at one focus, F, and a string tied to a tack at the other focus, C. See Figure 2.6a taken from Van Schooten (1657, p. 338). The string is tied to the end of the ruler at point N, and the pencil is used to hold the string against the ruler. This guarantees that as the pencil moves along the curve, its distances from points C and F will be increasing by the same amounts. Thus at all points P along the curve there is a constant difference between the distances PF and PC, that difference being EK. Thus, $PF - PC = EK$ for all points on the left branch, where EK is determined by the length of the string.

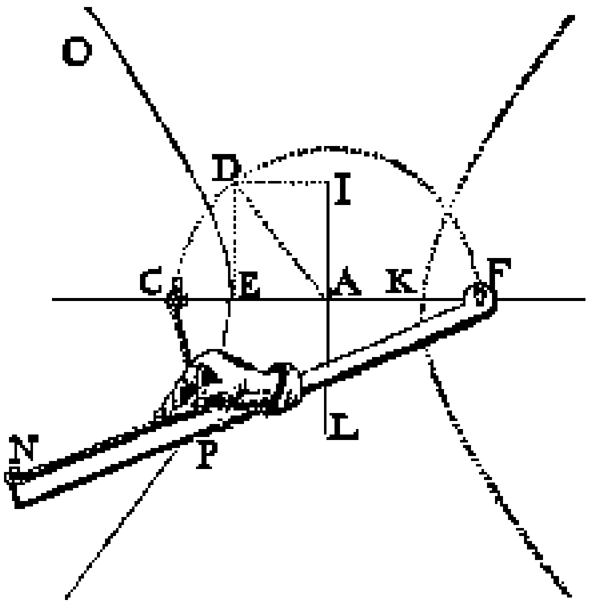


Figure 2.6a

Following the plan of section 2.5 we next set up three links between the two tacks at the foci (C and F) such that the middle one is equal to CF and the other two are arbitrary but equal to each other. That is to say in Figures 2.6b, 2.6c and 2.6d (Schooten, 1657, p.345-348), $DG = CF$, and $CD = FG$. This time, however, instead of crossing the links CD and FG as in figure 2.5b, one points them in opposite directions. The curve is still being drawn as the intersection of lines CD and FG but these lines must be extended in order to meet at the pencil point P. This apparatus maintains a constant difference between PF and PC, i.e. $PF - PC = FG = CD$. This can be seen, as before, through symmetry or through the congruence of triangles ΔPCF and ΔPDG (one of the triangles is made by connecting P to the foci, and the other by connecting P to the ends of the middle link). This constant difference is also equal to the axis EK. As in the elliptic case if one draws the line from P to the midpoint, H, of CG, this line is tangent to the curve at P. The symmetry of the picture is identical with Figure 2.5b.

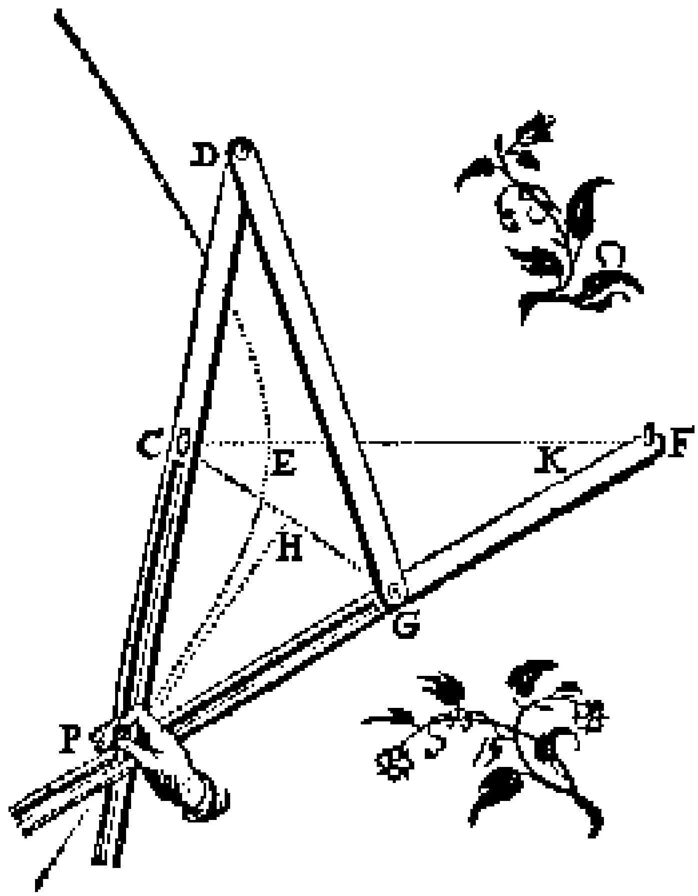


Figure 2.6b

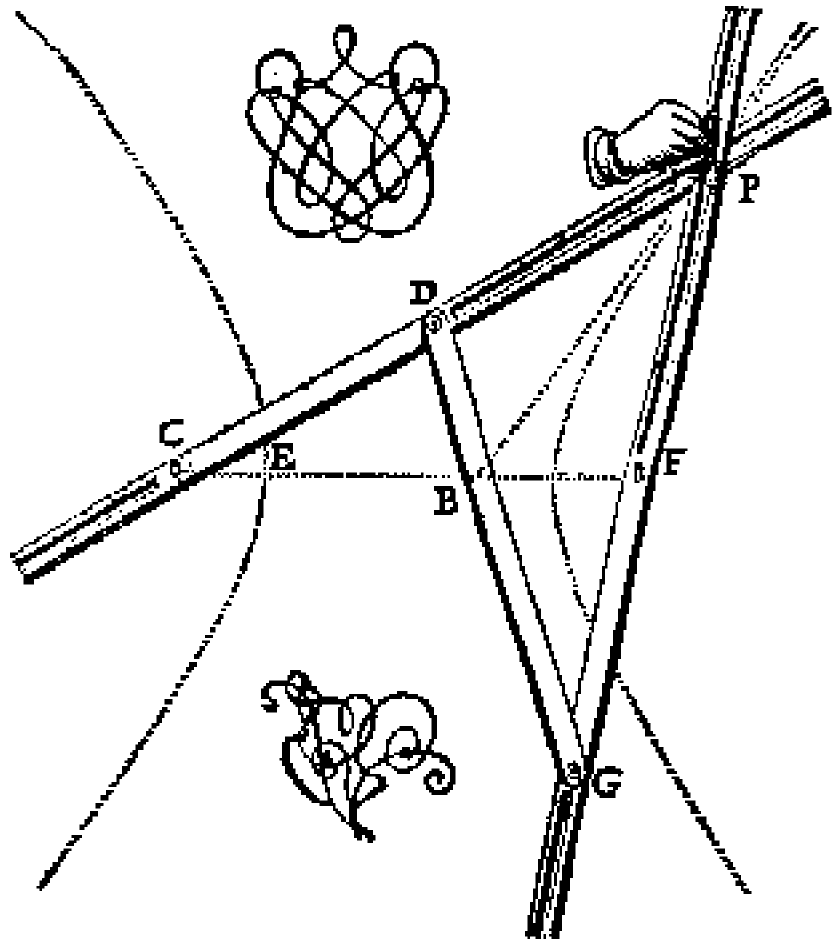


Figure 2.6c

Figure 2.6c (Schooten, 1657, p. 347) show this same device drawing the right branch of the curve. In this figure one sees that the point *B* where the link *GD* intersects the axis *CF* will also produce the tangent if connected with the point *P* on the curve.

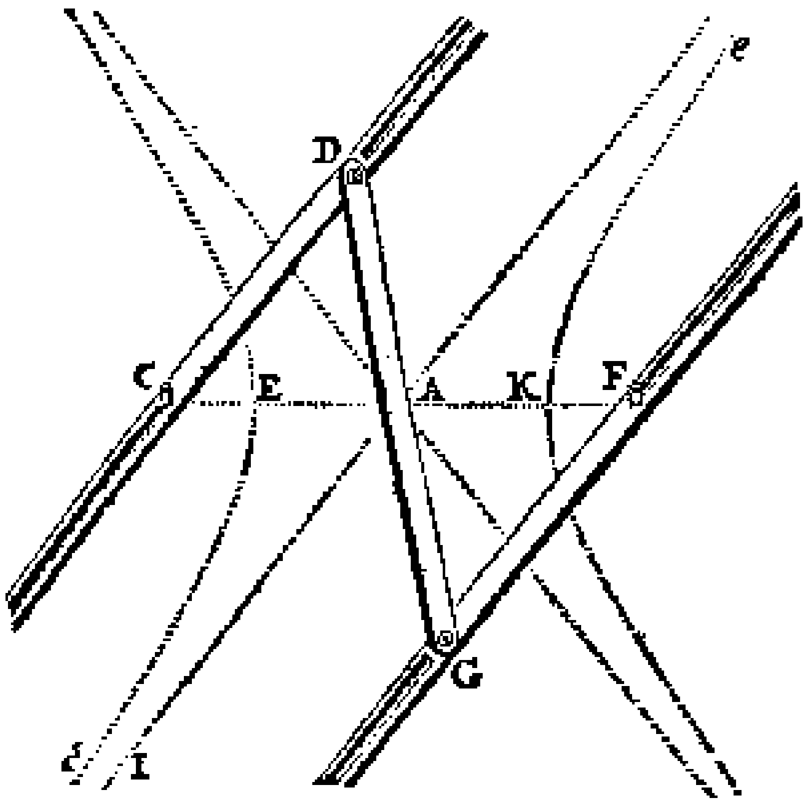


Figure 2.6d

Figure 2.6d shows this device in the asymptotic position where $\overline{CD}$ and $\overline{FG}$ are parallel. Keeping in mind that $DG = CF$ one sees that this is the only position where the link $\overline{DG}$ will pass through A, the center of the curve.

A proof of the tangency properties can be given which is identical with the arguments made in the elliptic and hyperbolic cases. As before, it can be seen from symmetry, or by transforming this device into one which draws the hyperbola as the set of point equidistant from a circle and point outside it. Once again van Schooten accomplished this through the use of a hinged rhombus ($LDMF$) which constructed the perpendicular bisector of FD. He then drew the curve as the intersection of this perpendicular bisector (LM) and an extended radius (CD) of the circle centered at the other focus (C). See Figures 2.6e, 2.6f, and 2.6g for three positions of this device (Schooten, 1657, p. 349 -352). As before this is an envelope construction where the LM is always a tangent to the hyperbola. This construction can also be accomplished through paper folding by repeatedly folding points on circle onto a fixed point outside the circle. These foldings will outline the curve with tangents (Row, 1966; Gardner, 1989).

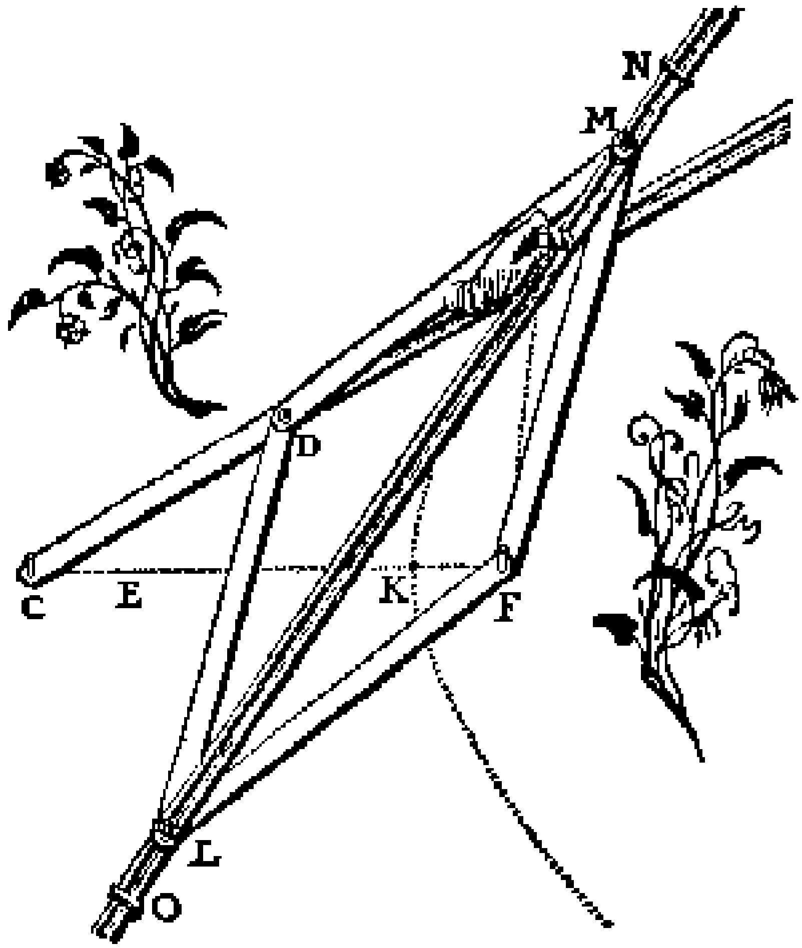


Figure 2.6e

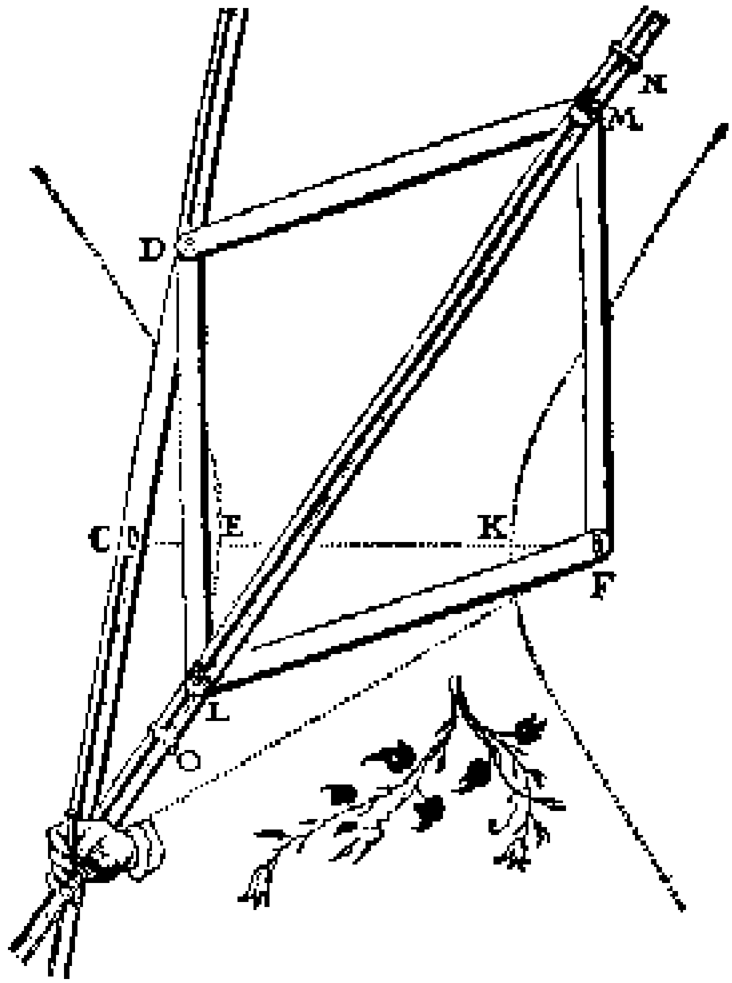


Figure 2.6f

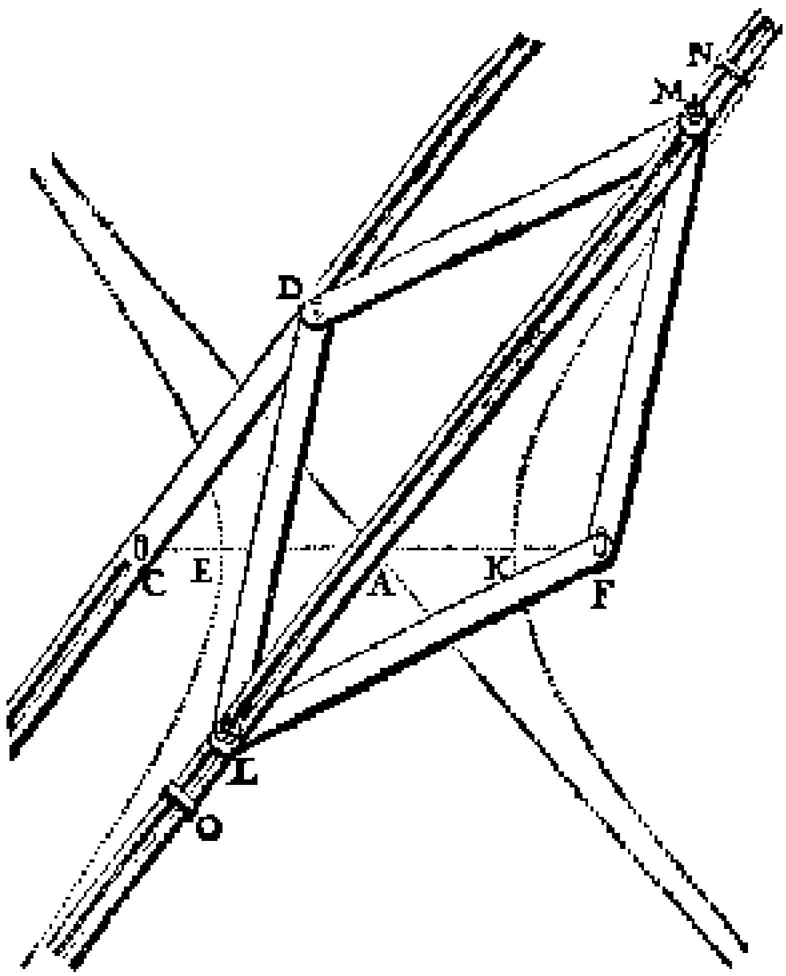


Figure 2.6g

Figure 2.6g shows the device in the asymptotic position. The lines *CD* and *LM* do not meet, so no point on the curve is defined, but in this position *LM* is the asymptote. A slight movement of the device in either direction will produce a point on the upper right or the lower left. This gives a strong sensation of connectedness between the branches of this curve. Descartes and other in the seventeenth century considered the hyperbola to be continuous because it can be drawn with one continuous action (Lenoir, 1979). This sense of continuity is not the same as adopting a projective view which sees the curve as connected at infinity, but is tied, instead, directly to an experience with a physical apparatus. In the seventeenth century the very term "geometer" most often referred to a person who designed and constructed machines and apparatus, such as siege engines, fortifications, canals, locks, bridges, etc.

As in the elliptic case this last, rather awkward device is actually the simplest of the three to animate on the computer using *Geometer's Sketchpad*. In fact one does not have to create a new construction at all. The elliptic one from Figure 2.5d becomes a hyperbolic one simply by dragging the off-center focus, I, outside the circle. I will change only the labels of the points in Figure 2.6h so that they are consistent with van Schooten's Figures 2.6e, 2.6f, and 2.6g. The circle (hodograph) is made by the rotation of the link $\overline{CD}$ around C, and one then traces the locus (P) of the intersection of the perpendicular bisector of $\overline{FD}$ with the extended radius $\overline{CD}$. The perpendicular bisector $\overleftrightarrow{AP}$ is the tangent line at P.

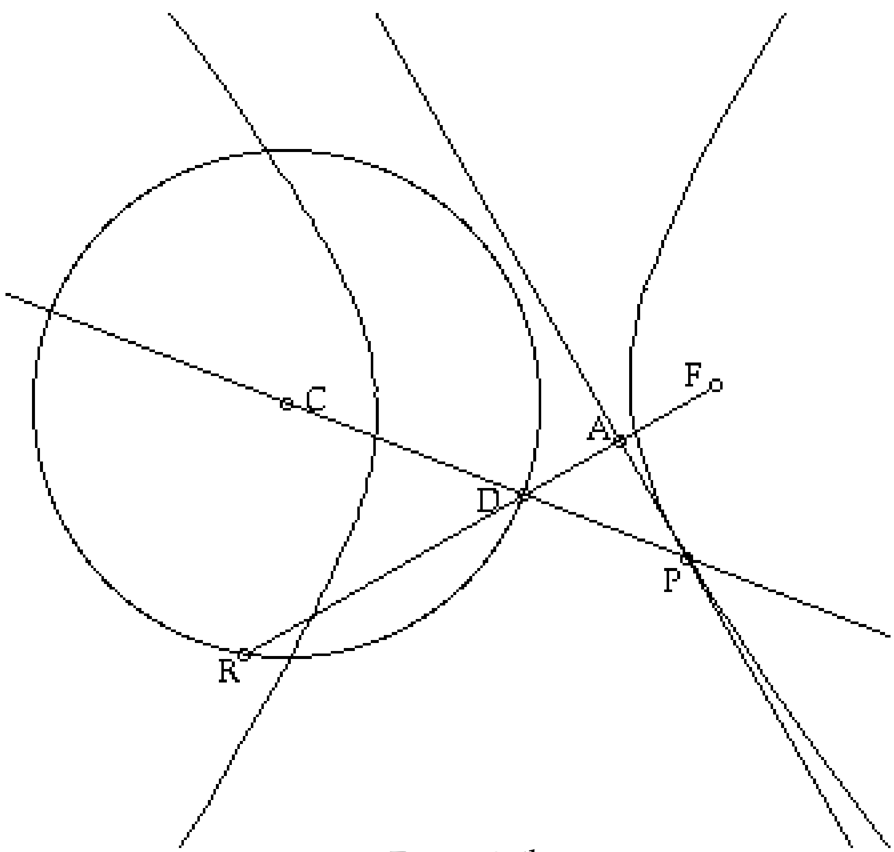
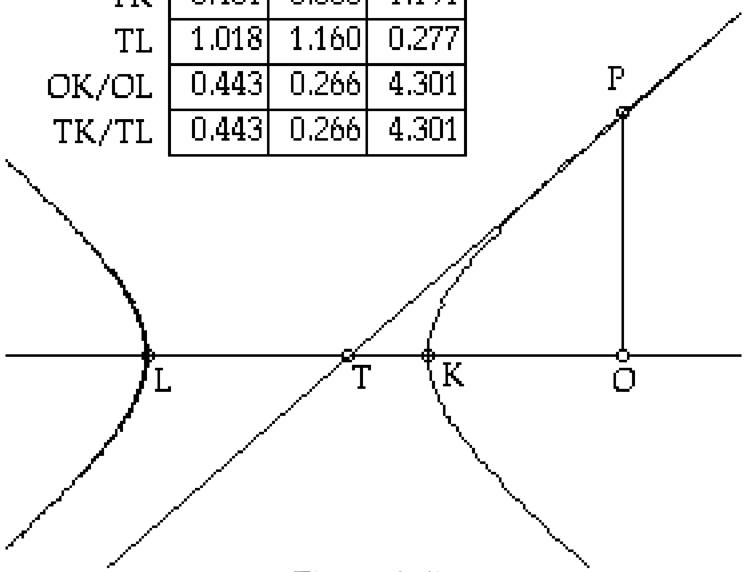


Figure 2.6h

If the point P is thought of as a planet in a hyperbolic orbit around the point F, once again the figure shows us not only the direction, but also the magnitude of its velocity vector. That magnitude will always be proportional to the length of the segment $\overline{FR}$, where the proportionality constant depends on the unit of time that is chosen. The geometry in the figure is equivalent to Kepler's second law, which says that a line between P and F will sweep out equal areas in equal times (Lenard, 1994). For those who wish to explore this on their own the same Euclidean hint applies; i.e. the product, $FD \cdot FR$, remains constant as D moves around the circle.



OK	1.168	0.532	1.913
OL	2.636	2.000	0.445
TK	0.451	0.308	1.191
TL	1.018	1.160	0.277
OK/OL	0.443	0.266	4.301
TK/TL	0.443	0.266	4.301

Figure 2.6i

The harmonic properties of the subtangent that were discussed for the ellipse in Section 2.5 continue to hold on the hyperbola. Dropping an ordinate from P to the axis, we find this time that the vertex K is always closer to the intersection of the tangent line T than to the foot of the ordinate O (see Figure 2.6i, where $TK < OK$). This shows, in an intuitive way, that the hyperbola opens more widely than does the parabola where TK would equal OK. The harmonic ratios reverse roles from the elliptic case, in that here T is between the vertices K and L, and O is outside. T divides $\overline{KL}$ on the inside in the same ratio that O divides $\overline{KL}$ on the outside. T and O are still harmonic conjugates of each other but the picture is inside out. Figure 2.6i tabulates values in inches for three different positions of P. The first two entries were made on the right branch of the hyperbola and the last entry where the ratios are greater than one was made with P on the left branch.

Once again I wish to point out the educational possibilities here for an empirical investigation of many aspects of curves and their tangents using linkages, paper folding and a computer. This can be done entirely with elementary geometry. Formal geometric proofs of these properties can be given (Apollonius, 1952), but that is not what I wish to stress. Solid intuitive experience with curves and tangents would make the transition to the language of analytic geometry and calculus so much more meaningful, since many of the initial statements made in these languages would be firmly grounded in a set of activities (Confrey, 1993a).

2.7 Drawing Hyperbolas from Asymptotes

In this section I will explore a device which comes directly from the first 1637 edition of Descartes' *Geometry* (1952). This was one of the first curve drawing linkage that I ever analyzed in detail. When I first read Descartes I could not believe that it worked. I immediately built a cardboard model in order to get the feel of it. I was excited by it and showed it to the other people on Prof. Confrey's mathematics education research project. We co-authored a paper dedicated entirely to this device, and the conceptual issues that it raised concerning the relations between curves, coordinates, and equations (Smith, Dennis, & Confrey, 1992). The issues discussed in that paper remain valid and will be mentioned here briefly, but at that time I had not carefully read Apollonius, and so I was unaware of the implications of this device concerning hyperbolic tangents. I was also unaware of how Descartes and his contemporaries saw their work in relation to Apollonius, and other classical, and Arabic geometers.

Descartes was concerned with bringing together and generalizing several mathematical directions. First, the algebraic notation of Vieté provided a way to write the ratio statements of Apollonius as equations (Klein,1968). Second, Arabic methods solved general cubic equations by intersecting conics, which were plotted point by point using ruler and compass (Joseph, 1991; Berggren, 1986). Descartes (1952) saw no reason to restrict himself to ruler and compass alone if he could provide exact means to draw a larger variety of curves. In the *Geometry* (1952) he solved a series of problems by intersecting various types of curves (see Section 2.12). These problems can all be stated as geometry problems, although many are discussed algebraically.

Descartes provided several curve drawing constructions which can be progressively iterated to produce curves of higher and higher algebraic degree. It is usually mentioned in histories of mathematics that Descartes was the first to classify curves according to the algebraic degree of their equations. This is not quite accurate. Descartes classified curves according to pairs of algebraic degrees, i.e. lines and conics form his first class (he used the term *"genre"*), curves with third or fourth degree equations form his second class, etc. (1952, p. 48). This is quite natural if one is working with linkages. With most examples of linkage iteration that I know of, each iteration raises the degree of the curve's equation by two, with special cases that collapse to an odd algebraic degree.[23] I shall mention several examples of this phenomenon in the following sections.

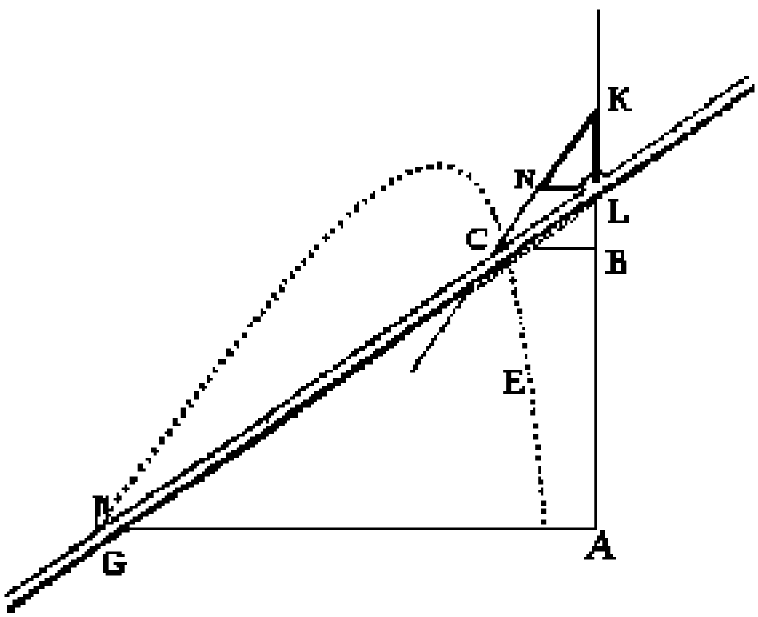


Figure 2.7a

[23] This same classification by pairs of degrees is used in modern topology in the definition of "genus." The "genus" of a non-singular algebraic plane curve can be thought of topologically as the number of "handles" on the curve when defined in complex projective space. In complex projective space, linear and quadratic non-singular curves have genus 0, and are topologically sphere-like. Similarly, curves of degrees 3 and 4 are topologically torus-like, and have genus 1. Curves of degrees 5 and 6 are topologically double-holed and have genus 2, etc. In the real model, (i.e. when considering only real solutions of one real equation in 2 variables) the genus 0 curves consist of at most one oval when you join up the aymptotes. The genus 1 curves will have 2 ovals, which is what you'd expect when cutting through a toric by a plane, etc. (This comment was made to me by Paul Pedersen.)

Figure 2.7a is taken from the original 1637 edition of Descartes' *Geometry* (1952, p. 50). Descartes described this curve drawing device as follows (note that his use of the term "diameter" refers to an asymptote which is not a "diameter" in the Apollonian sense, although it will serve as one of the coordinate directions):

> Suppose the curve EC to be described by the intersection of the ruler $\overline{GL}$ and the rectilinear plane figure NKL, whose side $\overline{KN}$ is produced indefinitely in the direction of C, and which, being moved in the same plane in such a way that its diameter $\overline{KL}$ always coincides with some part of the line $\overleftrightarrow{BA}$ (produced in both directions), imparts to the ruler $\overline{GL}$ a rotary motion about G (the ruler being hinged to the figure $CNKL$ at L). If I wish to find out to what class this curve belongs, I choose a straight line, as $\overleftrightarrow{AB}$, to which to refer all its points, and on $\overleftrightarrow{AB}$ I choose a point A at which to begin the investigation. I say "choose this and that," because we are free to choose what we will, for, while it is necessary to use care in the choice, in order to make the equation as short and simple as possible, yet no matter what line I should take instead of $\overleftrightarrow{AB}$ the curve would always prove to be of the same class, a fact easily demonstrated. (Descartes, 1952, p.51)

Descartes went on to find the equation of the curve in Figure 2.7a as follows. Introduce the variables (Descartes used the term "unknown and indeterminate quantities") $AB = x$, $BC = y$, (i.e. in modern notation $C = (x,y)$), and then the constants ("known quantities") $GA = a$, $KL = b$, and $NL = c$. Descartes routinely used the lower case letters x, y, and z as variables, and a, b, and c as constants, and our modern convention stems from his usage. Descartes, however, had no convention about which variable was used horizontally, or in which direction (right or left) a variable was being measured (x is measured to the left here). There was, in general, no demand that x and y be measured at right angles to each other. The variables were tailored to the geometric situation. There was a very hesitant use of negative values (often called "false roots"), and in most geometric situations they were avoided.

Continuing with the derivation, since the triangles ΔKLN, and ΔKBC are similar, we have: $\frac{c}{b} = \frac{x}{BK}$, hence $BK = \frac{b}{c}x$, hence $BL = \frac{b}{c}x - b$.

Now we have: $AL = y + BL = y + \frac{b}{c}x - b$.

Since triangles ΔLBC and ΔLAG are similar, we obtain: $\frac{BC}{BL} = \frac{AG}{AL}$, hence:

$$\frac{x}{\frac{b}{c}x - b} = \frac{a}{y + \frac{b}{c}x - b}$$

$$x\left(y + \frac{b}{c}x - b\right) = a\left(\frac{b}{c}x - b\right)$$

$$xy + \frac{b}{c}x^2 - bx = \frac{ab}{c}x - ab$$

(2.7-1) $$x^2 = cx - \frac{c}{b}xy + ax - ac$$

Descartes left the equation in this form because he wished to emphasize its second degree equation. He concluded that the curve is of the first class and a hyperbola. Descartes, however, was assuming that his readers were well acquainted with Apollonius.

If one continues to let the triangle ΔNLK rise along the vertical line, and keeps tracing the locus of the intersection of $\overleftrightarrow{GL}$ with $\overleftrightarrow{NK}$, the lines will eventually become parallel (see Figure 2.7b), and after that the other branch of the hyperbola will appear (see Figure 2.7c).

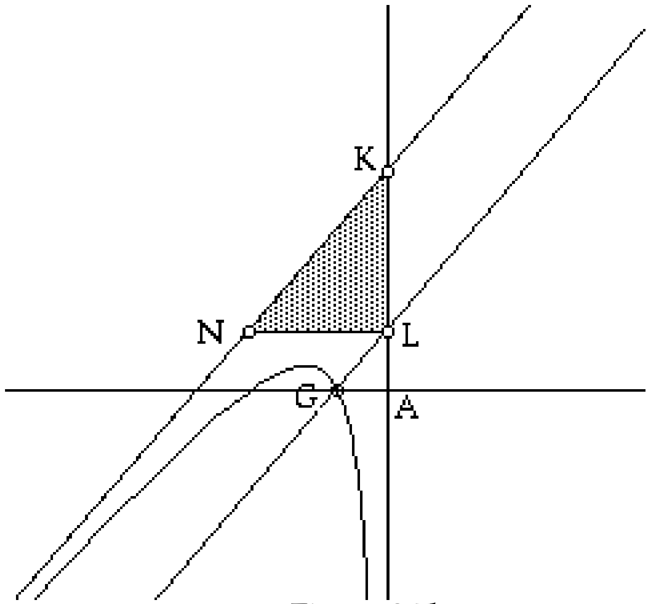


Figure 2.7b

These figures were made with *Geometers Sketchpad*, although I have altered slightly the values of the constants a, b, and c from those in Figure 2.7a. In Figure 2.7b, the line $\overleftrightarrow{KN}$ is in the asymptotic position. I will hereafter refer to this particular position of the point K, as point O. In this position triangles ΔNLK and ΔGAL are similar, hence the length of $AK = AO = \frac{ab}{c} + b$. The slope of the asymptote is the same as the fixed slope of $\overleftrightarrow{KN}$, i.e. $\frac{b}{c}$ (recall that $KL = b$, $NL = c$, and $GA = a$).

Let me rewrite equation 2.7-1 using A as the origin in the conventional modern sense, with x measured positively to the right. To do that I must substitute $-x$ for x. Making this substitution and solving Equation 2.7-1 for y one obtains:

(2.7-2) $$y = ab\frac{1}{x} + \frac{b}{c}x + \left(\frac{ab}{c} + b\right)$$

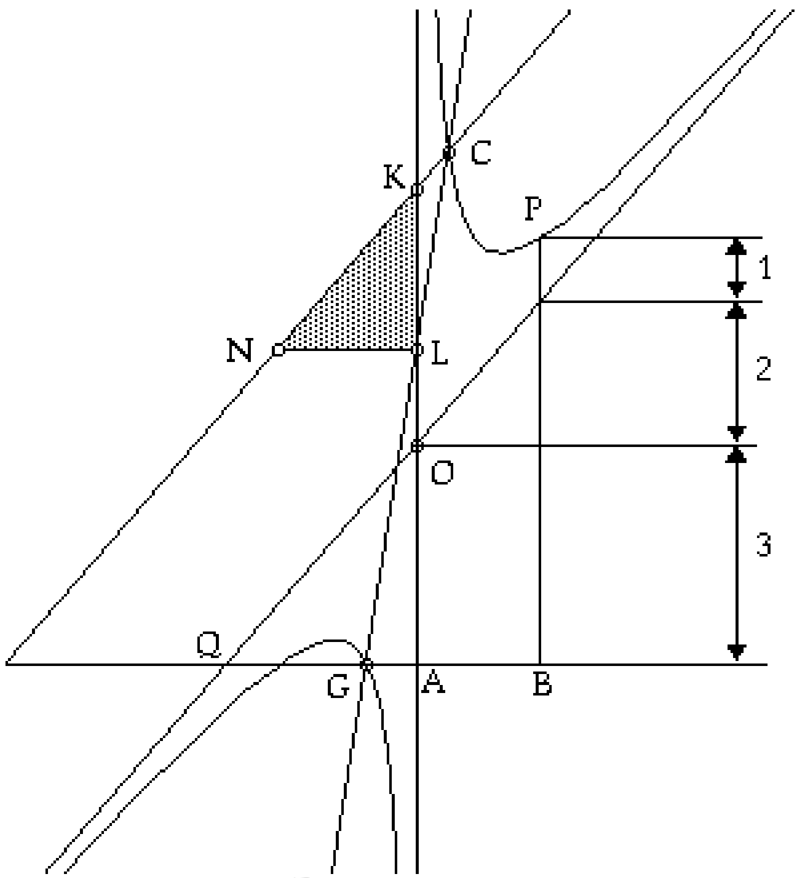


Figure 2.7c

One sees in Equation 2.7-2 that the linear equation of the asymptote appears as the last two terms of the equation. In Figure 2.7c, I have shown, to the right, the lengths that represent respectively the values of the three terms in Equation 2.7-2, for the point P (#1-inverse term, #2-linear term, #3-constant term). Term #3 accounts for the rise from the x-axis to the level of point O (the intercept of the asymptote. Adding term #2, raises one to the level of the asymptote, and term #1 completes the ordinate to the curve.

As a geometric construction, the hyperbola is drawn from parameters which specify the angle between the asymptotes ($\angle NKL$), and a point on the curve (G). If I change the position of the point N without changing the angle $\angle NKL$, the curve is unaffected as in Figure 2.7d. The derivation of the equation depends only on similarity, and not on having perpendicular coordinates. As long as $\overleftrightarrow{GA}$ is parallel to $\overleftrightarrow{NL}$, the derivation of the equation is the same except for the values of the constants $NL = c$, and $GA = a$, which have both become larger. Of course this equation is in the oblique coordinate system of the lines $\overleftrightarrow{GA}$ (x-axis) and $\overleftrightarrow{AK}$ (y-axis). It is the same curve geometrically, with the same form of equation, but with new constant values that refer to an oblique coordinate system. As long as angle $\angle NKL$ remains the same, and G is taken as any point on this curve, the device will draw the same curve. This form of a hyperbolic equation, as an inverse term plus a line, depends only on using at least one of the asymptotes as an axis.

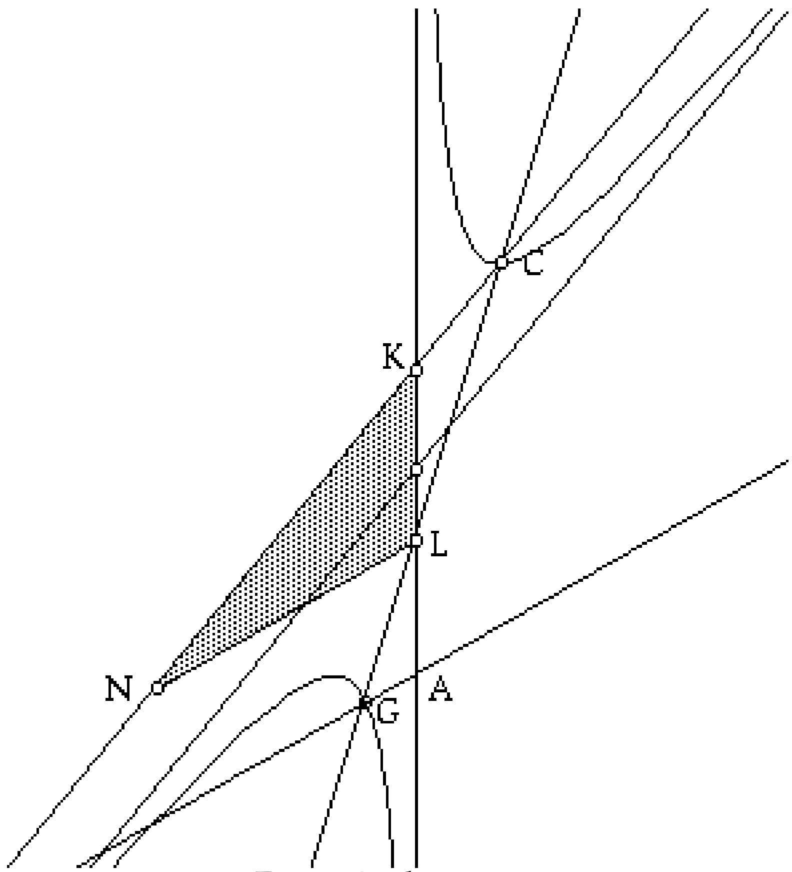


Figure 2.7d

I have encountered many students over the years who are well acquainted with the function $y = \frac{1}{x}$, and yet are entirely unaware that its graph is an hyperbola. Descartes' construction can be adjusted to draw right hyperbolas.

Consider the special case when the line $\overleftrightarrow{KN}$ is parallel to the x-axis (see Figure 2.7e). The point G is on the negative x-axis. Let $KC = x$, and $AK = y$ (i.e. $C = (x, y)$), $AG = a$, and $KL = b$. Now $AL = y - b$, and since triangles ΔLKC and ΔLAG are similar, we have:

$$\frac{KC}{KL} = \frac{AG}{AL}, \text{ or } \frac{x}{b} = \frac{a}{y-b}, \quad \text{hence the equation of the curve is:}$$

(2.7-3) $$y = ab\frac{1}{x} + b$$

A vertical translation by b would move the origin to the point O, and letting $a = b = 1$, would put G at the vertex $(-1,-1)$, yielding a curve with an equation of $y = \frac{1}{\text{x}}$.

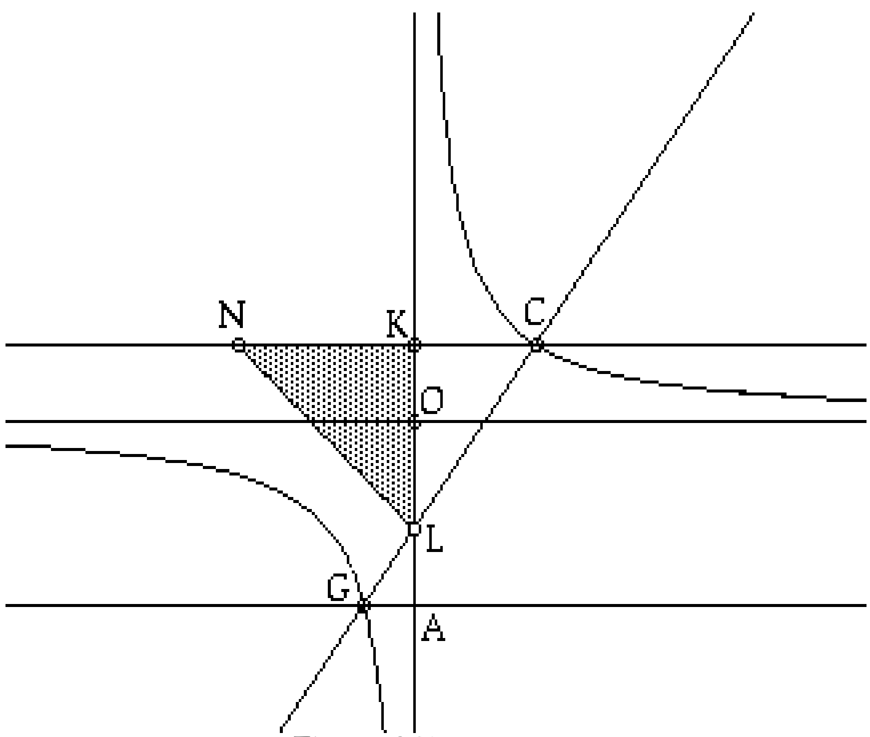


Figure 2.7e

Equation 2.7-3 can be seen as a special case of Equation 2.7-2 obtained by substituting ∞ for c, where c is thought of as the horizontal distance from L to the line $\overleftrightarrow{KN}$. In this case the linear term disappears. All translations and rescalings of the multiplicative inverse function can be seen as special members of the family of hyperbolas, via this construction.

One might ask how we know that these curves are in fact hyperbolas. Descartes said that this is implied by Equation 2.7-1. In his commentaries on Descartes, van Schooten gives us a bit more detail (Descartes, 1952, p. 55, note [86]). Once again these mathematicians assumed that their readers were familiar with a variety of ratio properties from Book 2 of Apollonius (1952; Heath 1961) that are equivalent to Equation 2.7-1. My intention here is not to give a set of proofs, but to suggest a means whereby students could explore these relations on their own.

Several of the theorems of Apollonius concerning the relations between tangents and asymptotes are beautiful, and easily explored in this setting. Using the asymptotes of a right hyperbola as edges to define rectangles, one sees that the points on the curve define a family of rectangles which all have the same area (see Figure 2.7f). Letting M and N be any two points on the curve, Equation 2.7-3 implies that the areas of $OPMS$ and $OQNR$ are both equal to $a \cdot b$, the product of the constants used in the drawing the curve. Another geometric property of interest is that the triangles ΔTSM and ΔNQU are always congruent. This congruence provides one way to dissect and compare these rectangles in a completely geometric way (see Henderson, in press).

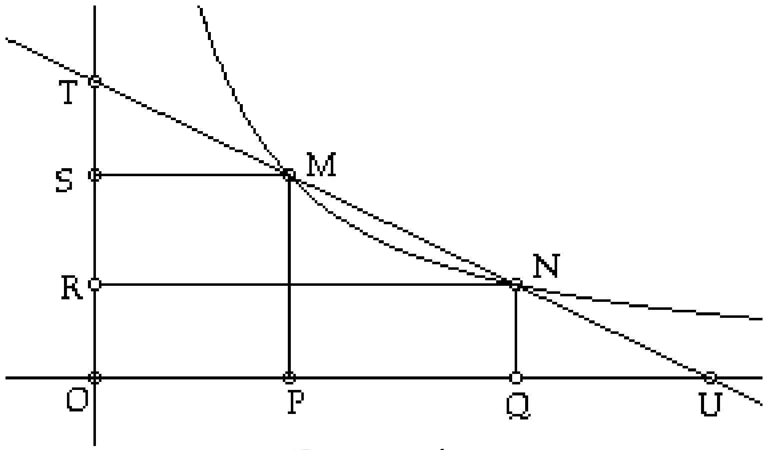


Figure 2.7f

Approaching these questions analytically, assume that the curve in Figure 2.7f has the equation $x \cdot y = k$ (using O as the origin). Let $M = (m, k/m)$ and $N = (n, k/n)$, i.e. $OP = m$ and $OQ = n$. Writing the equation of the line through M and N, one obtains: $y = \frac{-k}{mn}x + \left(\frac{k}{m} + \frac{k}{n}\right)$. Hence $TO = \frac{k}{m} + \frac{k}{n}$, and since $SO = \frac{k}{m}$, this implies that $TS = \frac{k}{n} = NQ$. Since triangles ΔTSM and ΔNQU are similar, this shows that they are congruent and that $TM = NU$. Now let the points M and N get close to each other. The line $\overleftrightarrow{MN}$ then gets close to a tangent line, and one can perceive the theorem of Apollonius, that the segment of any tangent line to a hyperbola, contained between the asymptotes, is always bisected by the point of tangency to the curve. This simple and beautiful theorem immediately implies, for example, that the derivative of $\frac{1}{x}$ is $\frac{-1}{x^2}$, by simply looking at the congruent triangles.

This bisection property of hyperbolic tangents is not restricted to the right hyperbola. Looking back at Figure 2.7-c, and Equation 2.7-2, one sees that any hyperbola coordinatized along its asymptotes will always have an equation on the form $x \cdot y = k$ for some constant k. To see this, subtract off the linear and constant terms from the y-coordinate, and then multiply the x-coordinates by a constant factor that projects them onto the asymptote $\overleftrightarrow{OQ}$. In the general case the curve can be seen as the set of corners of a family of equi-angular parallelograms which all have the same area. In Figure 2.7g, for any two points on the curve, M and N, the parallelograms $OQNR$ and $OPMS$ have equal areas. Since the triangles ΔTSM and ΔNQU are congruent, by letting M and N get close together one sees that any tangent segment $\overline{TU}$ is bisected by the point of tangency (M or N).

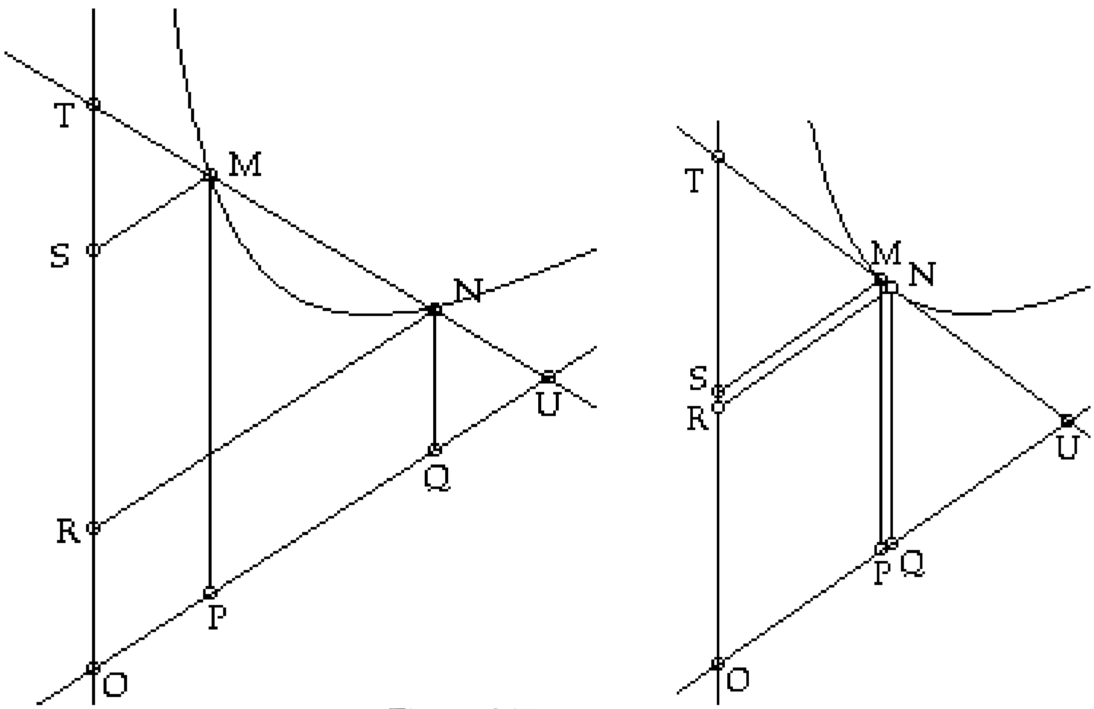


Figure 2.7g

An alternative view of the situations just described is to imagine any line parallel to $\overline{TU}$ meeting the asymptotes, and the curve in corresponding points T', M', and U'. Then the product $T'M' \times M'U' = TM \times MU$. That is to say, parallel chords between the asymptotes of a hyperbola are divided by the curve into pieces with a constant product. This follows from our discussion, because the pieces are constant projections of the sides of the parallelograms just discussed. It is this form of the statement that was most often used by van Schooten, Newton, Euler and others of that period (1650 -1750). This statement (from Book 2 of Apollonius (1952)) was traditionally used as an identifying property of hyperbolas. This constant product was given as a proof by van Schooten that the curve drawn by Descartes' device was indeed a hyperbola (Descartes, 1952, p. 55).

In this way it is possible to make a fully circular investigation of hyperbolas, using both geometric and algebraic representations. Neither representation is being used as a foundation for proof, instead one is lead to a belief in a relative consistency between certain aspects of geometry and algebra, through a checking back and forth between multiple representations. A calculus derivation of the derivative of $y = 1/x$ becomes, in this setting, a limited special case of the bisection property of hyperbolic tangents. It can be very satisfying for students to see symbolic algebra arrange itself into answers that are consistent with physical and geometric experience. Students can then experience the elation of Leibniz, as they build up a vocabulary of notation that becomes viable, because it can checked against independently verifiable, physical experience. Mathematical language is then seen as a code for aspects of experience, rather than as a dictator of truth. This is a revealing example of what is meant by "genetic epistemology."

2.8 Conchoids Generalized from Hyperbolas

Descartes generalized the previous hyperbola construction method by replacing the triangle ΔKLN with any previously constructed curve. For example, let a circle with center L be moved along one axis and let the points C and C' be the intersections of the circle with the line $\overleftrightarrow{LG}$, where G is any fixed point in the plane and $\overleftrightarrow{LG}$ is a ruler hinged at point L (see Figure 2.8a). Then C traces out a curve of degree four, known in ancient times as a conchoid (Descartes, 1952, p. 55). The two geometric constants involved are the radius of the circle r, and the distance a between the point G and the axis, along with the axis along which L moves.

Figure 2.8a shows three examples of conchoids for $a > r$, $a = r$, and $a < r$. If the curve is coordinatized along the path of L ($\overleftrightarrow{OL}$, a true Apollonian axis), and a perpendicular line through G ($\overleftrightarrow{OG}$), then its equation can be found by looking at the similar triangles ΔGOL and ΔCXL (top of Figure 2.8a). Since $GO = a$, $LC = r$, $CX = y$, $OX = x$, $XL = \sqrt{r^2 - y^2}$, one obtains the ratios of the legs in the triangles as: $\frac{\sqrt{r^2 - y^2}}{y} = \frac{\sqrt{r^2 - y^2} + x}{a}$, which is equivalent to: $x^2 y^2 = (r^2 - y^2)(a - y)^2$, which is of fourth degree, or of Descartes' second class. The squared form of the equation has both branches of the curve, above and below the axis, as solutions.

This example demonstrates Descartes' claim that, as one uses previously constructed curves to draw new curves, one gets chains of constructed curves that go up by pairs of algebraic degrees. Descartes called the conchoid a curve of the second class (i.e. of degree three or four). Dragging any rigid conic-sectioned shape along the axis, and drawing a curve in this manner will produce curves in the second class. Dragging curves of the second class will produce curves of the third class (i.e. degree five or six), etc. Descartes demonstrated this general principle by example, but he did not offer anything like a formal proof, neither geometric nor algebraic. His definition of curve classes was justified by his geometric experience.

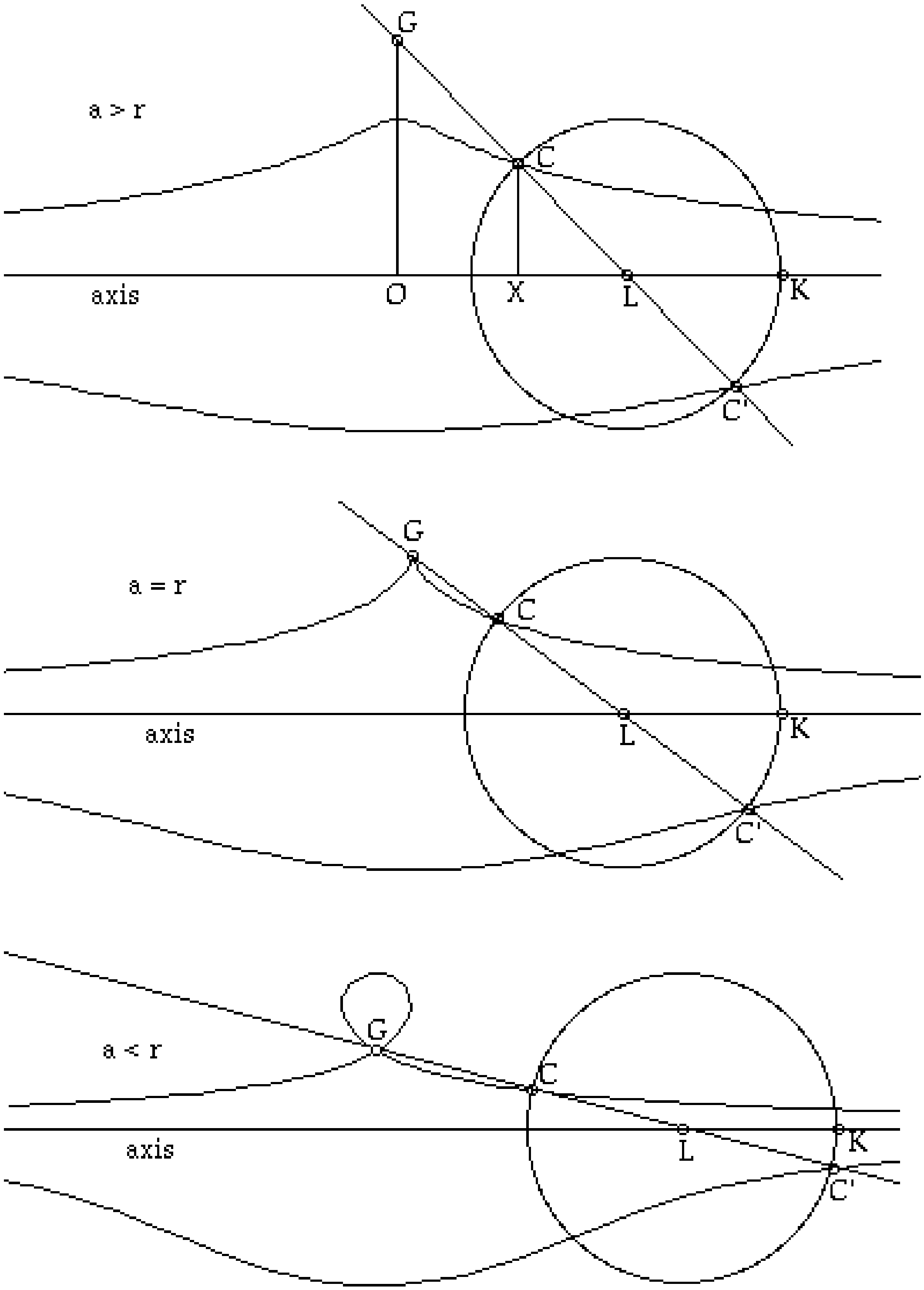


Figure 2.8a

Notice that when $a \leq r$, the point G becomes a cusp or a crossover point. When singularities like cusps or crossover points occur, these tend to occur at important parts of the apparatus, like pivot points (e.g. point G), or at a point on an axis of motion. Other examples of this phenomena will be displayed using Newton's organic devices (Section 2.11). I am not stating any particular or explicit mathematical theorem here. This observation is based upon my own empirical experience with curve drawing devices. There are probably several ways to make this observation into an explicit mathematical statement, subject to proof (Newton attempted several, 1968), but I wish to make the educational claim that students will benefit enormously from such empirical experience regardless of the extent to which they eventually formalize that experience in algebraic or logical language. An instinctual sense of where curve singularies might occur is fundamentally useful in many sciences.

One might ask here, why should one care about curves such as these conchoids? These curves and others generated from moving circles and lines (e.g. the cissoid, quadratrix, and spiral) were used by geometers to solve a number of important algebra and geometry problems (Descartes, 1952, p.40). They have been eliminated from our secondary curriculum, I feel, because they raise all of the difficult issues concerning the limitations of our standard notion of a function, and our belief that that notion can serve as universal umbrella concept under which all curves can be discussed. Our curriculum is governed not by what is useful in engineering or science, but by what is convenient in terms of the formal algebraic conventions of our standardized mathematical language.

The graphs of polynomial and rational functions do not occur spontaneously in terms of motion and mechanics. By focusing on these curves what is largely eliminated is the investigation of cusps and cross-over points which are very important in science (Arnol'd, 1990). If one goes on to study advanced calculus one repeatedly encounters the proviso on theorems that "any curve can be piece-wise defined by functions." It becomes a kind of litany that one learns to ignore, but hidden in this statement in the implicit belief that the function concept is the big idea which includes all others. When studying curves generated by physical motions, functions are most effective as secondary linguistic tools.

I want to suggest that even before students can deal effectively with the equations of such curves, that they be exposed to the curves that occur as simple geometrical motions. Functions might well be introduced gradually as tools for the analysis of increasingly complex motions, rather than immediately taking center stage. I hereby confess to having fond memories of my childhood hours spent with a curve drawing toy known as the *Spirograph*™. I was saddened as a child by the lack of interest of my teachers in the mathematics of that toy (Hall, 1992). Such devices belong in our mathematics classrooms.

2.9 Drawing Ellipses with Trammels

I begin this section by describing a device for drawing ellipses which is very old, and is used by carpenters around the world. This device is mentioned in American vocational educational programs, and was recently found in use among South African carpenters with little or no formal education (Milroy, 1990). It consists of placing a straight link (the trammel) within the right angle of a carpenters square, and then tracing the curve drawn by any point on the trammel as its ends move against the legs of the square. See Figure 2.9a, which is again taken from van Schooten (1657, p. 325). This device is quite natural for use by artists and crafts people, because it creates an ellipse, not from foci, but from the lengths of its axes. It also utilizes the most common tool of builders: the square.

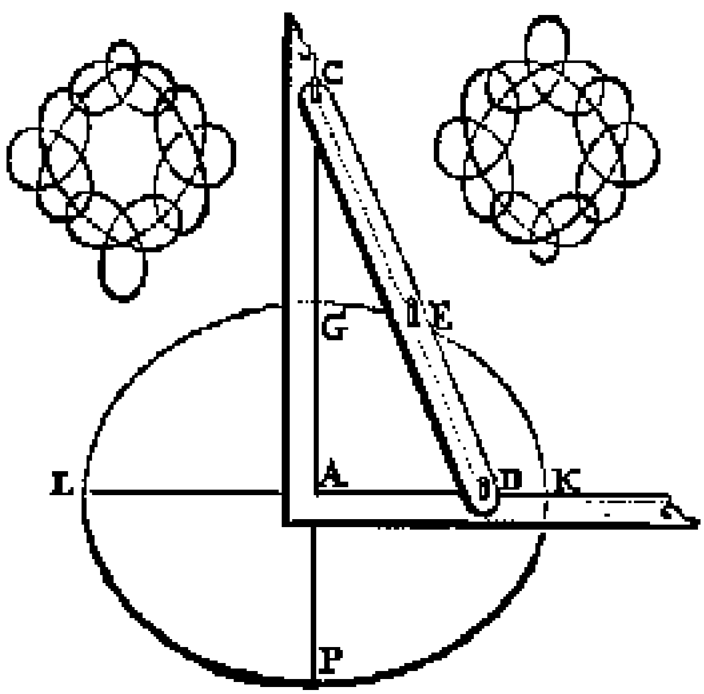


Figure 2.9a

It can be seen from the action of the device that $AK = CE$, and that $AG = DE$, so that the trammel size and the position of the point E will determine the lengths of the axes, but how does one know that this device is tracing a true ellipse? Let me first look at the special case where the pencil at E is riding on the midpoint of the trammel $\overline{CD}$. I claim the curve being drawn in this case is a circle. To see this drop the vertical line from E to meet $\overline{AK}$ at O (see Figure 2.9b). Since triangles ΔCDA and ΔEDO are similar, and E is the midpoint of $\overline{CD}$, then O is the midpoint of $\overline{AD}$. Hence triangles ΔEOD and ΔEOA are congruent (by SAS), and therefore $AE = ED = CE$. Since this is true for all positions of the trammel, midpoint E is tracing a circle of radius $CE = ED$.

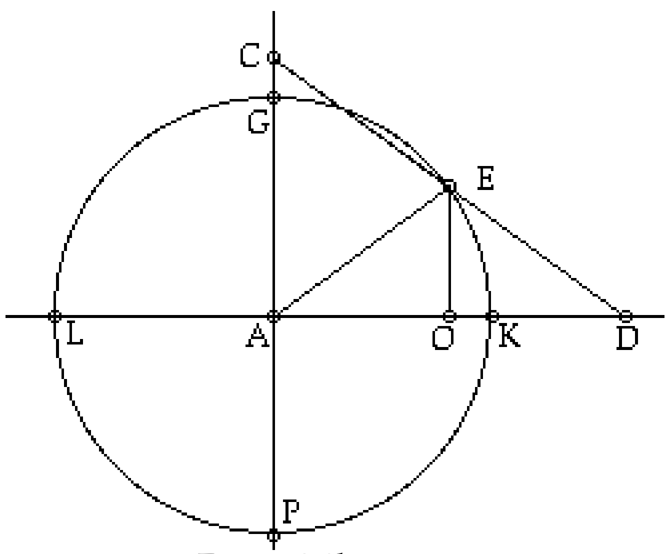


Figure 2.9b

Returning to the general case, let $CE = a$, and $ED = b$. I wish to show that point E on the trammel $\overline{CD}$ traces an ellipse with a horizontal axis of length $2a$, and a vertical axis of length $2b$. While drawing this curve I will simultaneously draw a circle with radius a, using another trammel $\overline{SR}$ of length $2a$ with its midpoint at P (see Figure 2.9c). Keeping the two trammels parallel to each other as they draw, one sees that triangles ΔPOR and ΔEOD are similar, hence at each point $\frac{PO}{EO} = \frac{PR}{ED} = \frac{a}{b}$. If one coordinatizes using A as the origin, this says that for every abscissa $x = AO$, the ordinate of the circle $y = PO = \frac{a}{b} \cdot EO$. Since the equation of the circle is $x^2 + y^2 = a^2$, letting $y' = EO$ (the ordinate of the curve), the equation of the curve is: $x^2 + \left(\frac{a}{b} \cdot y'\right)^2 = a^2$, which is equivalent to: $\frac{x^2}{a^2} + \frac{y'^2}{b^2} = 1$. The equations are mentioned here for a modern reader, but they are not necessary in order to prove that the trammel draws an ellipse. If one views ellipses as curves proportionally related to circles (as did Apollonius and van Schooten; see Section 2.2), the proof would then end a few lines earlier.

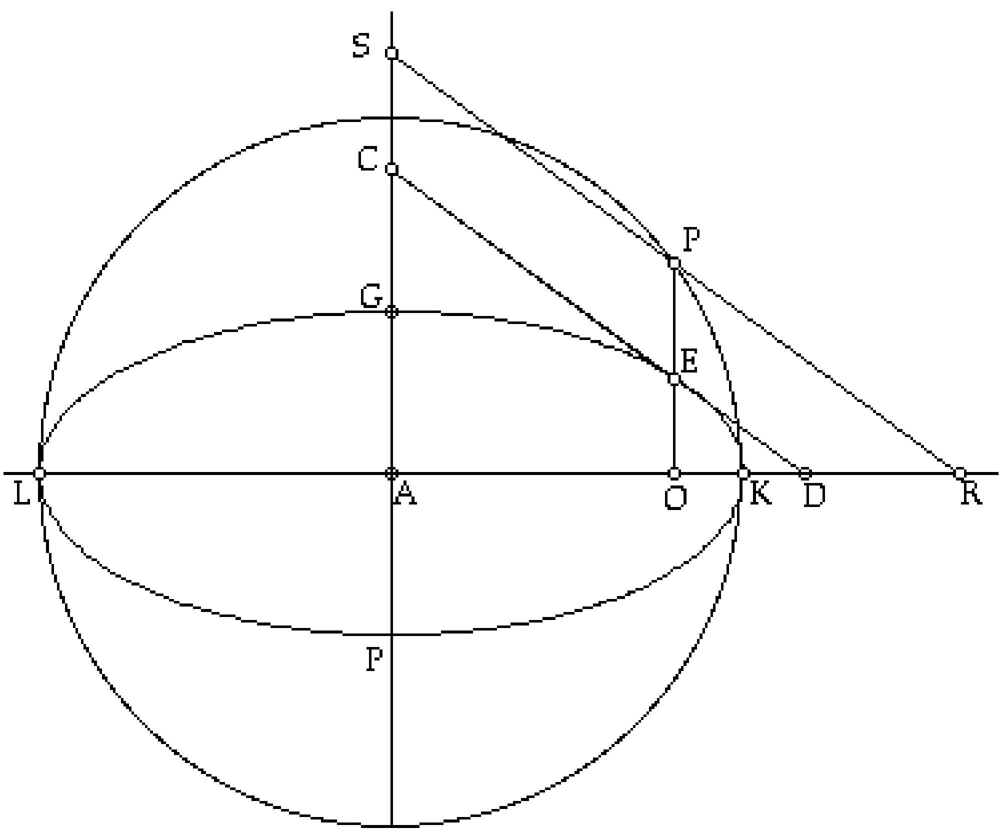


Figure 2.9c

The previous construction yields more than just a guarantee that the curve is an ellipse, and an equation. It gives the area of the ellipse relative to the circumscribed circle. It is nice to think of area here in the sense of Cavalieri's notion of the motion of a variable line segment, or in a more physical sense as a stack of cards. I have shown that each card in the deck (ordinate) that forms the ellipse is $\frac{b}{a}$ times the corresponding card in the deck (ordinate) that forms the circle of radius a. Hence the area of the ellipse is $\frac{b}{a}(\pi a^2) = \pi ab$.

This ratio holds not only for the area of the entire ellipse, but also for the area over any piece of the axis. For example, the area of the circle over the segment $\overline{OK}$ could be found by subtracting the area of triangle ΔAOP from the area of the circular sector AKP.[24] Multiplying by $\frac{b}{a}$ would then give the area of the ellipse over the segment $\overline{OK}$.

Many users of computer graphics have experienced this view of ellipses as circles with a uniform compression or expansion in one dimension. Novice computer users are often disturbed when they enter the equation of a circle, and the graph window shows an ellipse because they do not have identical scales on both axes. This phenomenon first led me to the disturbing realization that some computer pixels are not square, but rectangular.

All of these curve drawing devices can give rise to a host of interesting related rate problems. This particular device is indirectly mentioned in nearly every calculus book as the "ladder sliding down the wall problem." Usually one end of a trammel is assumed to be moving at a constant rate, and questions are then posed about the variable rate of the other end, or of points in between. I have talked with many calculus teachers, all of whom were familiar with this problem, but I have yet to meet one who was aware that the points on the sliding ladder follow elliptical paths. I myself discussed this problem for years, and was never

[24] In section 2.14, I will describe how Leibniz used another curve drawing device to generate an infinite series for the integration of circles.

aware of this. This points out how routinely teachers talk about situations in math classes with which they have no physical or empirical experience at all. Several calculus teachers have told me how most students in their classes detest the "ladder sliding down the wall problem." My response is that students must be allowed to do it. From the active experience far more emerges than just the solution to this standard calculus rate problem.

Returning to the trammel I wish to point out that any point on the line $\overleftrightarrow{CD}$ will trace an ellipse, even if it does not lie between C and D. See Figure 2.9d from Van Schooten (1657, p. 324). The lengths of the axes are still $2 \cdot CE$ and $2 \cdot DE$. Some people may be familiar with this form of the device as a popular desk top toy known as the "BS grinder."

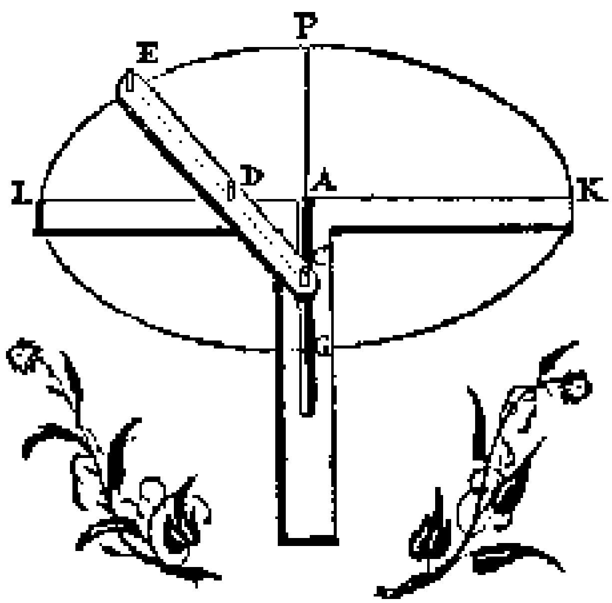


Figure 2.9d

What happens if one traces curves using a trammel whose ends are moving along lines which are not at right angles? (See Figure 2.9e.) The answer is that one still traces ellipses, but now neither the path of C nor of D are axes of symmetry for the curve. Their intersection, however, remains at the center, and hence the ends of the trammel still move along Apollonian diameters. Finding the standard equation of the curve along an axis can be difficult here, but, using Apollonius, we can see what its equation will be by using one track as a diameter, and finding its appropriate ordinate direction. See Figure 2.9e from Van Schooten (1657, p. 329). When $\angle CDA = 90°$, then point E will be at its maximum height, and therefore the tangent there will be horizontal. Hence at this point, $\overline{AE}$ is the conjugate diameter to $\overline{AL}$. The angle □ $\angle EAD = \alpha$ can be determined from the fixed angle $\angle CDA = \beta$ of the rulers, since the ratio of their tangents is equal to $\frac{a+b}{b}$. Now half the length of the conjugate diameter $\overline{AE}$ will equal $\frac{b}{\sin\alpha}$, and so an equation of the curve along the diameter $\overline{AL}$, using its ordinate direction $\overline{AK}$, will be $\frac{x^2}{a^2} + \frac{(\sin^2\alpha)y^2}{b^2} = 1$

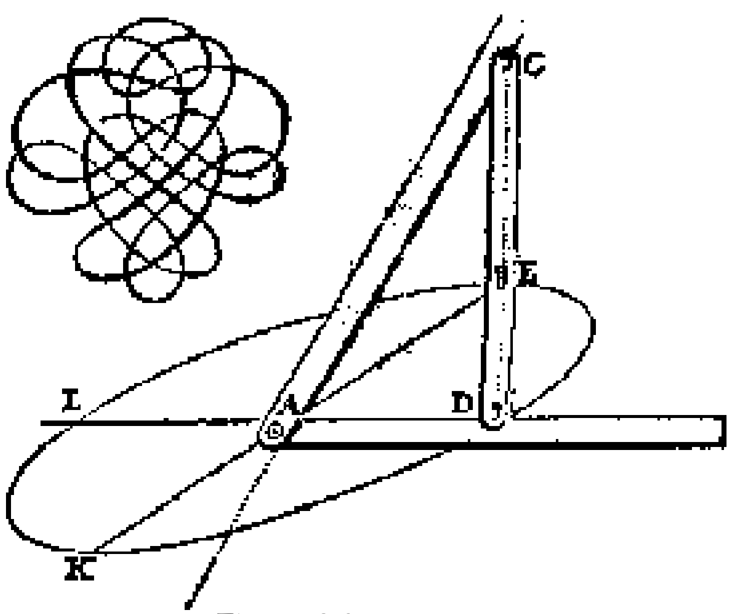


Figure 2.9e

Another variation on the trammel device (see Figure 2.9f) is to replace the linear trammel with any rigid triangle (ΔCDE), and let C and D move along perpendicular paths as before. The point E will once again trace out an ellipse skewed to the paths of C and D. This device was mentioned several times in the work of by Leonardo Da Vinci (for more details, see Pedoe, 1976). I have never discovered exactly how or why Leonardo employed such a device, but after playing with one I have a guess as to why an artist might find this method of drawing ellipses useful.

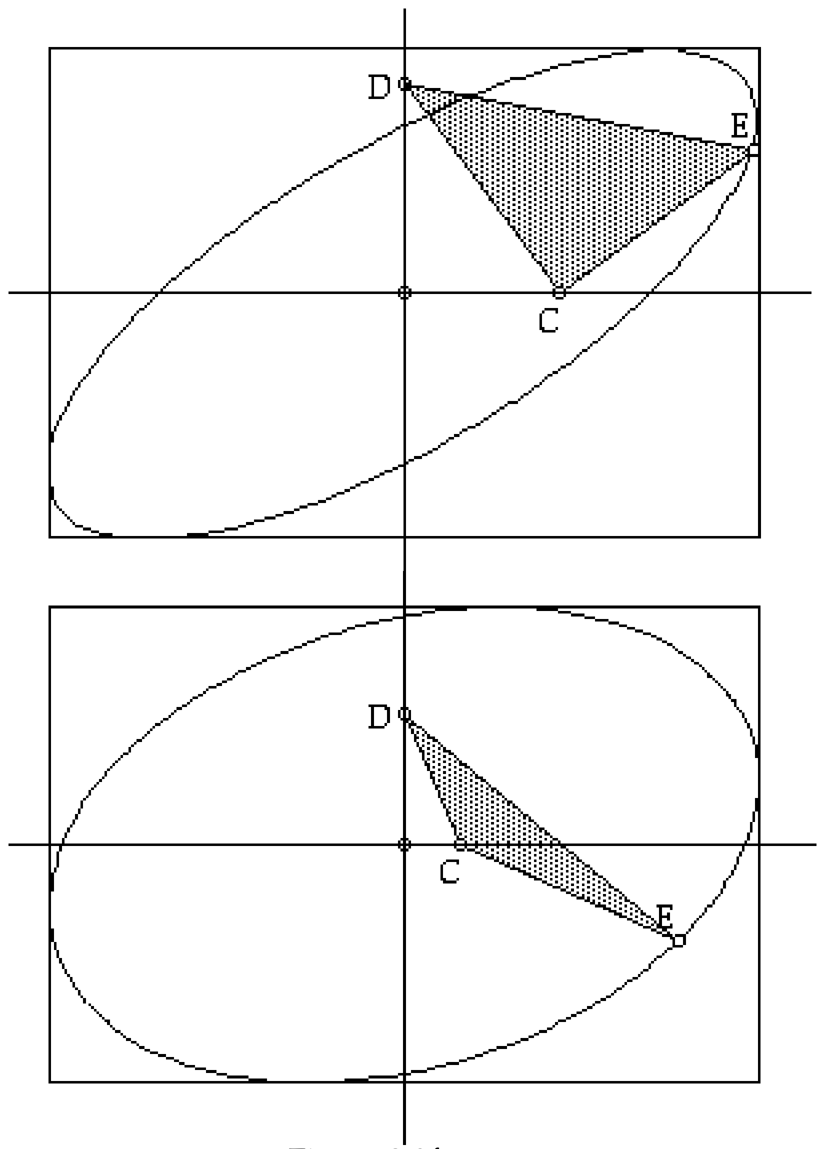


Figure 2.9f

The sides of the triangle $\overline{DE}$ and $\overline{CE}$ will determine the size of the rectangle which circumscribes the ellipse, since they are the maximum horizontal and vertical displacements of the point E. By leaving these two sides of the triangle fixed, and varying the length of $\overline{CD}$, one can draw the entire family of possible ellipses which can be circumscribed in a fixed rectangle. Using ellipses as projected circles, and fitting them into prescribed rectangular spaces is exactly the problem that often comes up in drafting. Templates of such families of ellipses were common drafting tools until the advent of computers.

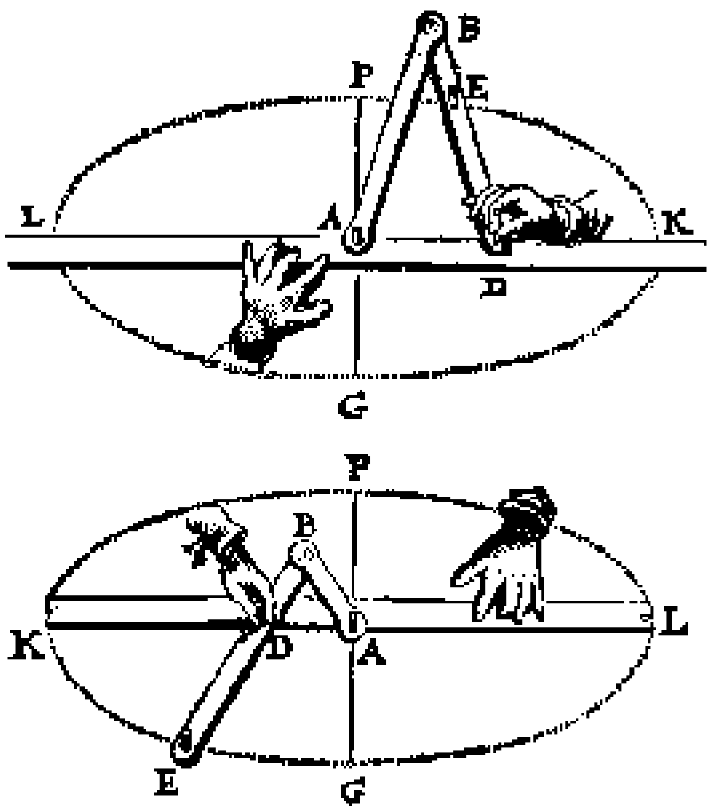


Figure 2.9g

Returning to the original trammel in Figure 2.9a, van Schooten transformed this device into another as follows. Let B be the midpoint of $\overline{CD}$. Since the point B describes a circle around A, one could connect it to point A with a hinged link half the length of $\overline{CD}$. One could then cut off half of the trammel $\overline{CB}$, and the motion of the other half ($\overline{BD}$) would remain unchanged by this transformation. See Figure 2.9g from van Schooten (1657, p. 322). The curves described by any point E on $\overline{BD}$, as D moves along the horizontal remain the same as when E rode on the trammel. The semi-major axis is $AB+BE$ and the semi-minor axis is $AB-BE$. I like to call this device the folding straw device, because I have demonstrated its action over lunch to many people by folding a plastic straw in half and piercing it with a pen. More discussion of this device, in conjunction with the trammel, will be found in the student interviews in Chapter 3.

The transformation of the trammel into this device is a proof that it draws ellipses but it depends on having $AB=BD$. If these two lengths are not equal, the device will draw ovals which are egg shaped, i.e. having a tighter curve at one end than the other. If AB and BD are extremely unequal it will draw curves shaped like kidney beans. This family of curves falls into Descartes' second class (i.e. degree four with possible special cases of degree three). An example is shown in Figure 2.9h. I was surprised by this since the other distortions of the original trammel device continued to draw ellipses.

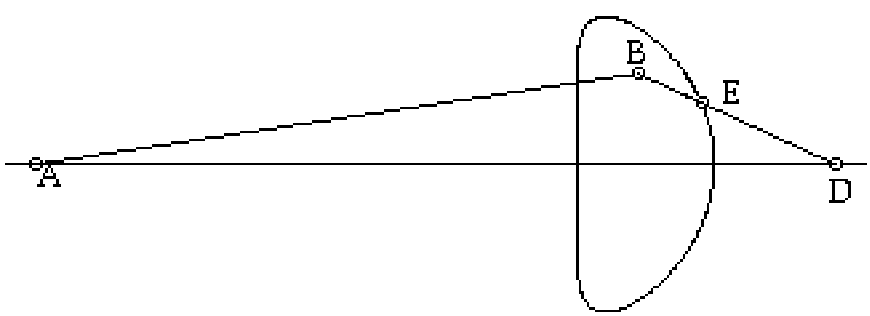


Figure 2.9h

Interesting variations on the trammel device can be obtained by letting one end of a trammel move along a line, while the other end moves along a circle. See Figure 2.9i for an example. Archimedes used such a trammel to trisect an arbitrary angle (Courant & Robbins, 1941), and Newton generalized Vieté's description of such a device as a general method for geometrically constructing the solutions of cubic equations (1967, p.72).

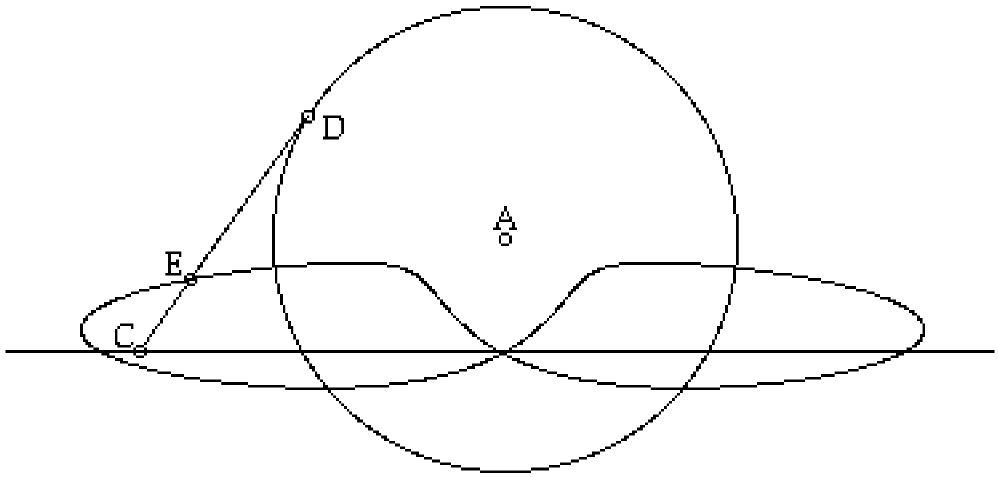


Figure 2.9i

The point *D* of the folding straw in Figure 2.9g can also be linked to a circle in which case, thinking of the radius of the circle as a third link, one returns to the general problem of three links between two fixed points mentioned in Sections 2.5 an 2.6. Another way to think of the three link problem is to imagine a trammel that moves between two circles. Thinking of the centers of the circles as the fixed points, and their radii as the two other links, this situation is equivalent to tracing the path of a point on the middle link of three links between two points.

The three link situation is of fundamental importance in mechanical engineering, and the curves that it generates provide a situation in which to investigate a family with a remarkable set of variations (Artobolevskii, 1964). Although these curves come from very simple linkages, and are easy to draw, it is very challenging to analyze them from an algebraic standpoint. Many geometric descriptions and conjectures are sure to arise from a investigation of these curves. Eliminating such topics solely because they present algebraic difficulties leads students to a false dependency. If they have only seen curves with easy algebraic representations, then they will begin to think that algebra will always be the best approach. These curves, which have a simple method of generation, will quickly disabuse one of over dependence on algebra. Figure 2.9j shows one of my favorite examples of this family of curves, most of which tend to have a figure-8 sort of shape. I suggest an investigation of these curves as a group project for students. It is sure to lead to some surprises.

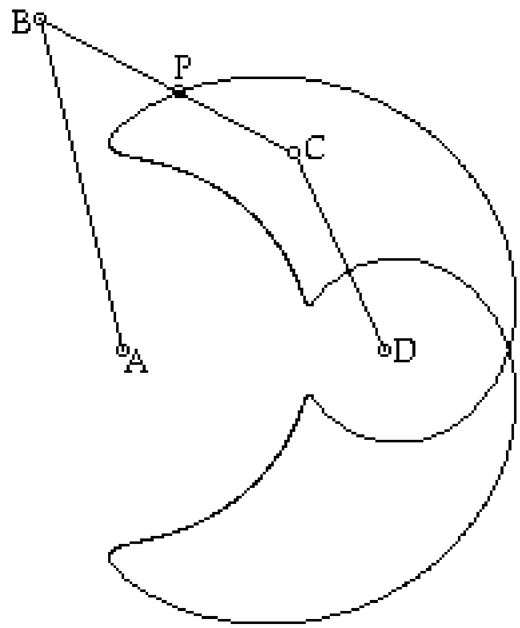


Figure 2.9j

2.10 Fermat's Quadratic Transformations

One of the first mathematical works by Pierre de Fermat was his *Introduction to Plane and Solid Loci,* which was sent to his friends in Paris in 1637, at almost the same time that Descartes was reviewing the galley proofs of his *Geometry*. Fermat worked within the school of analysis laid out by Vieté. His notations, unlike Descartes', adhere quite strictly to those used by Vieté, and his system bears a great similarity to the language of, for example, Galileo. Fermat was not trying to create a new language or system, but was trying to push forward with the analytic program that had been laid out by Vieté and his followers (Mahoney,1973; Klein, 1968). He was a well established provincial lawyer, and government official who pursued mathematics in his spare time.

Fermat considered his work *Introduction to Plane and Solid Loci* to be a commentary on Apollonius, but in it he developed his own system of analytic geometry. His approach and style differ considerably from that of Descartes. As I have shown, Descartes was concerned with finding equations for curves that he first constructed geometrically. Fermat was mainly concerned with classifying equations according to the shape of the curve that is produced by geometric interpretations of the variables. His approach was to create the curves from equations, and thus he studied graphs of equations in way that is much closer to analytic geometry as it currently taught. It is interesting to note that Descartes' background was in military engineering, while Fermat began his career working for the government on tax and monetary policy. Mechanical engineering is alien to Fermat's thinking, instead he worked extensively on financial mathematics, and eventually developed the first systematic techniques for solving maximum and minimum problems (Mahoney, 1973). His work pioneered the study of graphical representations of numerical and algebraic phenomena.

Fermat's method of interpreting algebraic equations is to consider a fixed line and a line segment of variable length, one of whose endpoints moves along this line. As in Apollonius, all of the line segments make the same angle with the fixed line, but this angle need not always be a right angle. Thus his graphing of equations is more general than what is usually taught in our curriculum.

The main theorem of his *Introduction* is that any equation of second degree, in two variables, graphed at any fixed ordinate angle, will always produce a conic section. This is the converse of Descartes' assertion that all conic sections will have second degree equations no matter how they are coordinatized. Fermat created seven standard equations whose loci he demonstrated, and then proceeded to use transformations to reduce all other second degree equations to one of these seven types. These seven standard equations are essentially algebraic interpretations of the standard ratio properties from Apollonius.

Fermat, for example, showed that $b^2 - x^2 = y^2$ produces a circle of radius b when segments of length y are erected at a right angles, at points along an axis where x measures distances from some fixed point on that axis. That is to say the endpoints of the segments trace out a circle. Similarly he showed that the equation $b^2 - x^2 = ay^2$ produces an ellipse, but here the fixed angle between the ordered segments (ordinates) and the axis (abscissas) is arbitrary.

Although Fermat was working from a largely algebraic point of view, his transformations of equations utilized methods that, although not mechanical, are almost equivalent to curve drawing devices. As an example, I wish to describe one of Fermat's reductions of a more complicated equation to one of his general forms (I will use modern notational conventions which are nearly identical to those of

Descartes[25]) (Mahoney, 1973, p. 89). Starting with the equation $b^2 - 2x^2 = 2xy + y^2$, Fermat began by rewriting the equation as:

$b^2 - x^2 = (x+y)^2$. Now let N be a fixed point on a fixed line $\overleftrightarrow{NM}$ and let Z be a moving point on that line whose distance from N is x. Constructing a perpendicular segment $\overline{ZI}$ at Z whose length was $x+y$, Fermat knew from the second equation that the endpoint I traced out a circle of radius b (see Figure 2.10a).

Now from the point N construct a line $\overleftrightarrow{NR}$ at 45° to the fixed line $\overleftrightarrow{NM}$. Extend each ordinate of the circle $\overline{ZI}$ to meet this line at a point O. Now define the point V so that $OV = ZI = x+y$. As Z moves along the horizontal axis, the point V will now trace out a new locus as shown. This construction can be carried out dynamically using *Geometer's Sketchpad* by defining the locus from a the point I revolving on the circle.

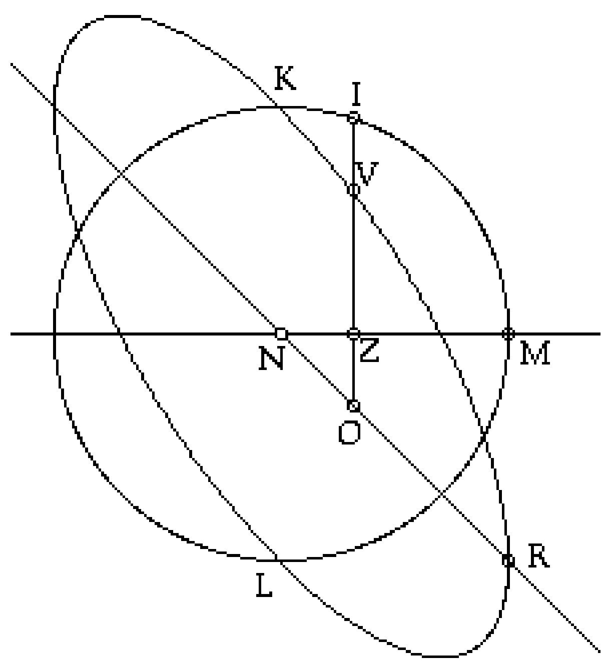


Figure 2.10a

---

[25] The notation of Fermat was not hugely different. For example, he used "A" and "B" as his varibles and his exponents were written differently.

Fermat chose the 45° angle so that $NZ = OZ = x$. Since $OV = x + y$, then $ZV = y$ by subtraction, and therefore $\overline{ZV}$ acts as the ordinate which represents our original equation in the system of perpendicular ordinates from the line $\overleftrightarrow{NZ}$. But now consider V as the endpoint of the moving segment $\overline{OV}$, as point O moves along $\overleftrightarrow{NR}$. In this system the fixed ordinate angle is 45°. Let $u = NO$ and $v = OV$; then $u = \sqrt{2}x$, and $v = x + y$. Substituting these into the original equation, it becomes: $b^2 - \frac{u^2}{2} = v^2$, or $2b^2 - u^2 = 2v^2$. From this equation Fermat knew that the point V traced out an ellipse because this is one of the standard equational forms coming from Apollonius. That is to say that $\overline{NR}$ and a vertical line through N are conjugate diameters of the ellipse. Note that the tangent to the ellipse at R is vertical and the tangents at L and K (where $x = u = 0$) are parallel to $\overleftrightarrow{NR}$.

This is one of Fermat's more imaginative transformations, but the technique is typical of his entire approach to analytic geometry. He shifts the geometric framework until the unknown locus (graph) of an equation becomes recognizable. Fermat did not really use coordinate geometry as we know it. His approach has been called ordinate geometry or axial geometry because he worked at various angles from one axis. It should also be stressed that Fermat was not providing what we would now call complete proofs. The example described above is, however, completely general. There is nothing special about the twos in the original equation. By choosing the appropriate angle for the axis $\overleftrightarrow{NR}$, any equation of this form can be reduced to an ellipse measured in a standard Apollonian system (i.e. seen to have a standard equation).

It is interesting to vary this angle on *Geometer's Sketchpad* and observe its effect on the resulting ellipse. For example, when the angle $\angle MNR = 60^\circ$, the ellipse is much more eccentric. As before $OV = ZI$, but here the new system $\overline{NO}$,

$\overline{OV}$, has a fixed 30° angle (see Figure 2.10b). For more details on Fermat and his mathematics, see Mahoney's book on the subject (1973).

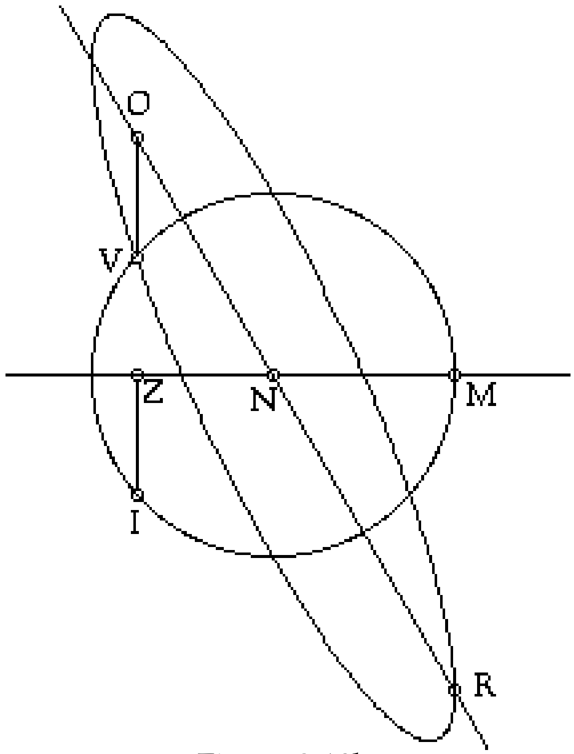


Figure 2.10b

2.11 Newton's Organic Constructions

Among the mathematical papers of Isaac Newton (1642 - 1727) there are several investigations of curve drawing devices, and one type in particular is discussed repeatedly in his papers (1968). Newton published very little of these researches, but a letter containing some of his thoughts was circulated early in his life, and was read by Leibniz. Although Newton published very little of the details of his mathematics in his lifetime, his private papers reveal great detail as to the directions of his thoughts. Thanks to the monumental editing work of D. H. Whiteside (Newton, 1967; 1968) all of these original papers are now generally available (in 8 volumes).

Van Schooten had explicitly searched for a single uniform method by which he could draw all conic sections using continuous uninterrupted motions. As was shown in Sections 2.5 and 2.6, he came very close to succeeding. Newton carefully studied both Descartes and van Schooten, and extended the work of both in the years 1667 and 1668 (see Whiteside's introduction to Newton, 1968). He produced the first systematic classification of cubic curves (30 varieties) (Newton, 1968).

I shall describe one type of linkage device that was invented by Newton as a solution to van Schooten's quest. He seems to have claimed at various points in his papers that this one device, when iterated, is capable of drawing all possible algebraic curves. His proofs, however, are incomplete, and I have been unable to complete or contradict them. This device is quite simple to construct and can be very easily animated on *Geometer's Sketchpad* . It produces an abundance of clear examples of Descartes' principle that as one iterates a device, the algebraic degree of the equations jumps up by two, and also that special points on the device often become special points on the curve (i.e. cusps, crossover points, etc.).

Imagine a pair of rulers hinged together at a point with a locking screw that will fix the angle between them. Now pin the pair of rulers to a plane so that the rulers can rotate together around their intersection point A, always preserving the fixed angle between them. Now make a second such pair of rulers (having a possibly different fixed angle), and pin them down to rotate at a second fixed point B in the same plane. Newton called these two points (A and B) the "poles" of the rulers.

The construction then worked as follows. Let one of rulers from A and one from B meet at the point D. Now move D along any previously constructed curve, and trace the locus of the intersection P of the other two rulers (one from A and one from B). Newton called the rulers that meet at D the "directing legs," and the ones that meet at P the "describing legs." The previously described curve along which point D moves, he called the "directrix." Thus his device draws new curves from old ones. Newton's use of the term "directrix" is far more general than our modern usage, but gives more sense to the meaning of the word as a path which "directs" the construction of a curve. In Section 2.4 the directrix of the parabola did indeed direct the drawing of the curve via van Schooten's device, but that is only a special case of what Newton had in mind. Throughout this section I will use the term "directrix" in the broader Newtonian sense.

To begin building curves, one starts by using a line as a directrix. As the point D (the intersection of the directing legs) moves along this given line the point P (the intersection of the describing legs) will move along a conic section. The type, size and position of the conic will be determined by the placement of the poles, A and B, the sizes of the two fixed ruler angles, and the position of the directrix. In each of the figures in this section I have marked the fixed angles rotating around A and B, with bold line segments. The directrix is always labeled with the points D and R. Color coding on *Geometer's Sketchpad* is quite helpful here.

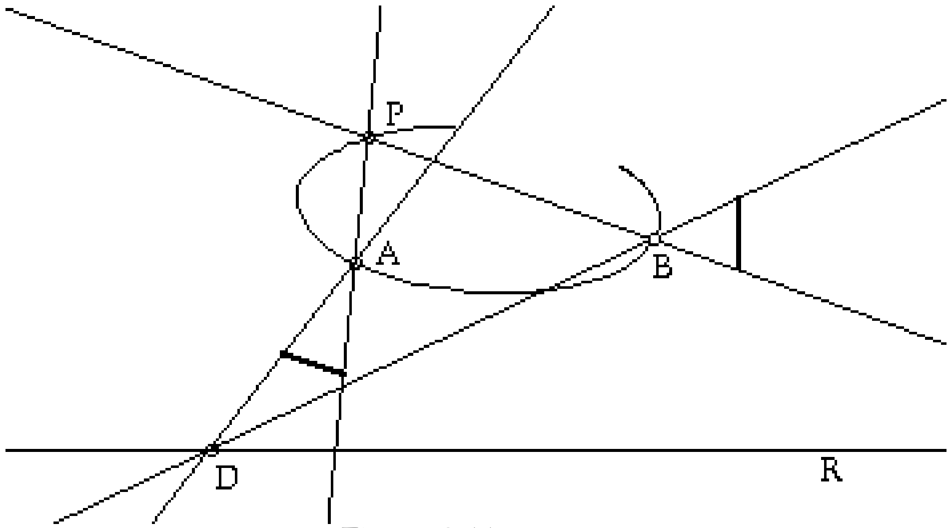


Figure 2.11a

Figure 2.11a shows Newton's device drawing an ellipse. The ellipse is not complete because this device always sets up a one to one correspondence between the points on the directrix (line $\overleftrightarrow{DR}$) and the points on the curve traced by P. I have shown the portion of the curve of the curve that is drawn by letting D move a few inches off the figure on both the right and the left. The missing piece of the curve will be drawn, at increasingly slow rates, by positions of D along the rest of the infinite line of the directrix. The motion of point D along the directrix provides a very explicit, physical sense of a parameter which governs the generation of points on the curve. The introduction of functional language into this example would most naturally be expressed in terms of this physical parameter.

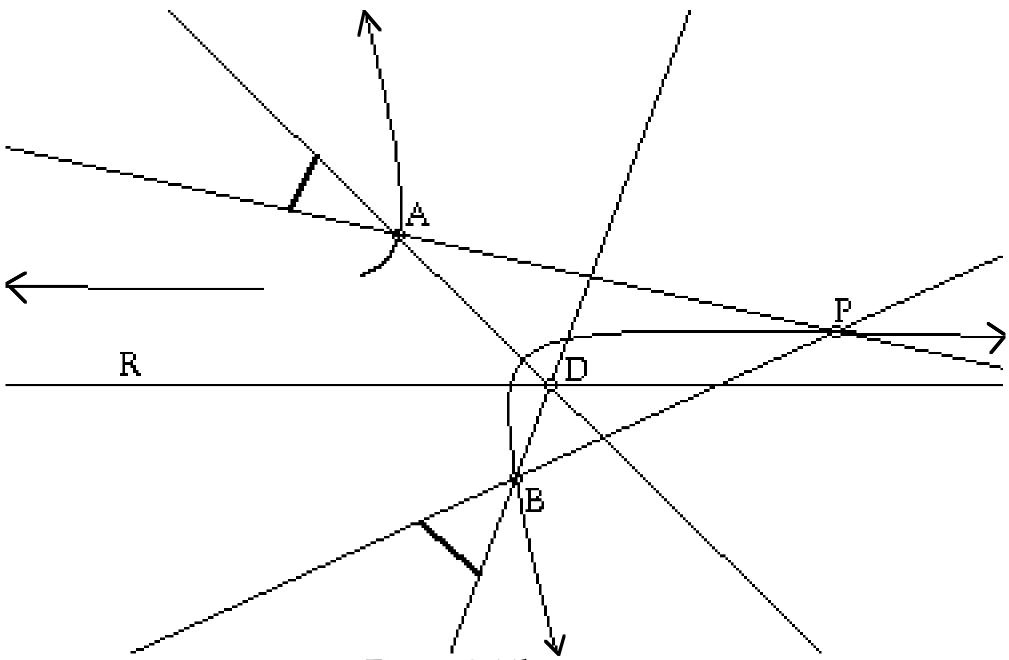


Figure 2.11b

Figure 2.11b shows Newton's device drawing a hyperbola. Again there is a missing piece of the curve that corresponds to positions of D far to the right and to the left. It is interesting to locate the position on the curve that corresponds to D at ∞. Both $D = \pm\infty$ correspond to the same point on the curve where the directing legs become parallel.

With Descartes' hyperbolic device in Section 2.7, one of the ∞-points on the hyperbola at the ends of one of the asymptotes was smoothly connected by a small motion of the device and the other ∞-point corresponded to the ∞-point of the linear motion of the device. Here the ∞-point of the directrix does not correspond to either one of the ∞-points on the hyperbola at the ends of the asymptotes. Both asymptotes are smoothly connected by small motions of D along the directrix. Descartes avoided any direct mention of ∞-points, but Newton used them quite freely. These curve constructions lend themselves quite naturally to a discussion of curves in projective geometry, a subject that would flourish in England for the next two centuries (Richards, 1988).

Newton considered the case where a ruler pivot point (a pole) is allowed to be placed at infinity. This "rotation" of two rulers around a point at infinity means that the two rulers move parallel to each other, in a fixed direction, at a fixed distance apart. Figure 2.11c draws a parabola by moving the pole B to infinity, and placing the rulers through B at right angles to the "directrix" $\overleftrightarrow{DR}$. Note that the line $\overleftrightarrow{DR}$ does not usually correspond with the modern restricted meaning of the term "directrix" even for parabolas. In Figure 2.11c, $\overleftrightarrow{DR}$ is a "directrix" for the parabola but it is not the "directrix" in the modern sense. Note that in Figures 2.11a and 2.11b, the poles A and B were on the curves. In Figure 2.11c the ends of the parabola are approaching the pole B at infinity.

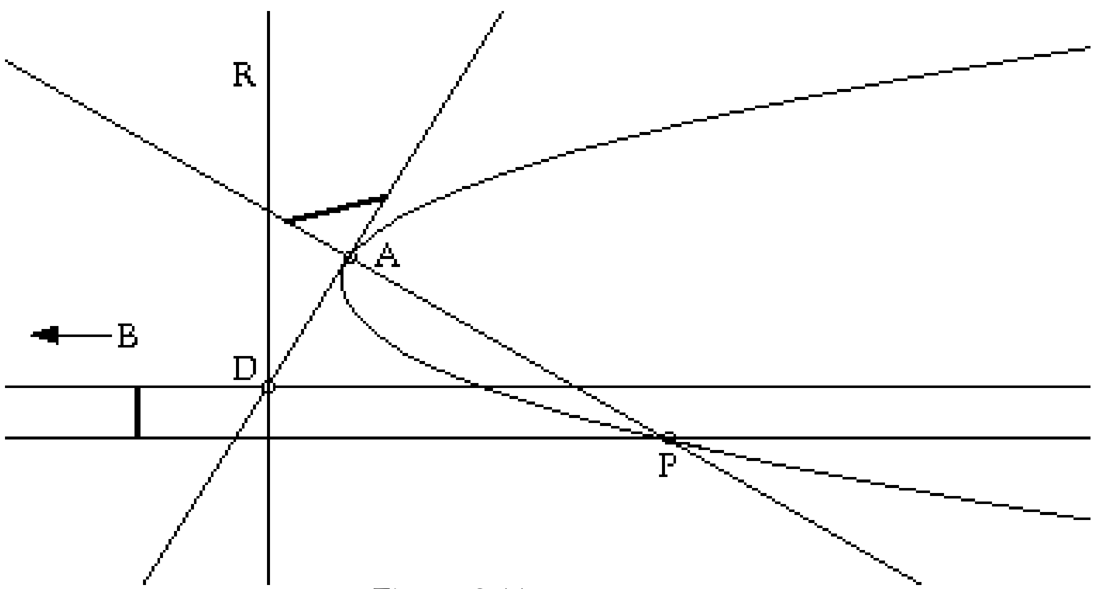


Figure 2.11c

Newton emphasized the utility of this device for solving the problem of drawing a conic section through any five points in a plane (1968, p. 119). Let the five given points be A, B, C, E, and F. Place the hinged sets of rulers at A and B, and set the fixed angles so that the rulers at pole A have angle $\angle CAB$, and the rulers at pole B have angle $\angle CBA$. This places the intersection of the describing legs at C, just when the two directing legs coincide along the line $\overleftrightarrow{AB}$. Having fixed these ruler angles now rotate both pairs so that the intersection of the describing rulers P falls at the next point E. Now mark the position of the intersection of the directing rulers D_1. Do the same for the last point F, to find another point D_2. Let the directrix be the line through these two marked points D_1 and D_2. The device will now be set to draw a conic section passing through all five points A, B, C, E, and F. Newton suggested that:

> after you have described the curve one way, you may test it in others whether the description is accurate: precisely, by fixing the rules in other angles (say $\angle EAB$, and $\angle EBA$ and so on) or by taking other points for the poles of the rules"
> (Newton, 1968, p. 119).

When one follows Newton's suggestion, and interchanges the roles of the five points in this construction, one draws the same curve, but in a very different way and at very different rates along the same sections of the curve, i.e. using very different parameters of motion.

Newton's organic curve drawing device is capable of describing curves from a variety of other given sets of geometric prerequisites, without recourse to analytic methods. For example, one could construct a conic passing through four specified points and having a specified tangent at one of those points, or a conic passing through three points with two specified tangents, or a conic through four points with a specified line parallel to an asymptote, etc. (Newton, 1968, p. 123).

Newton went on to give ways to draw third and fourth degree curves (Descartes' second class of curves) from given geometric prerequisites. In these cases, one must use a conic section as a directrix. In Figures 2.11d, 2.11e, and 2.11f, I have used *Geometer's Sketchpad* to draw some examples of fourth degree curves using, each time, a circle as a directrix. I have varied the positions of the poles (A

and B) with respect to the circle. Starting with A and B both outside the circle, I have drawn a figure-8 shaped curve (Figure 2.11e). Point A is not on the curve, but B is a singular point where the curve crosses itself.

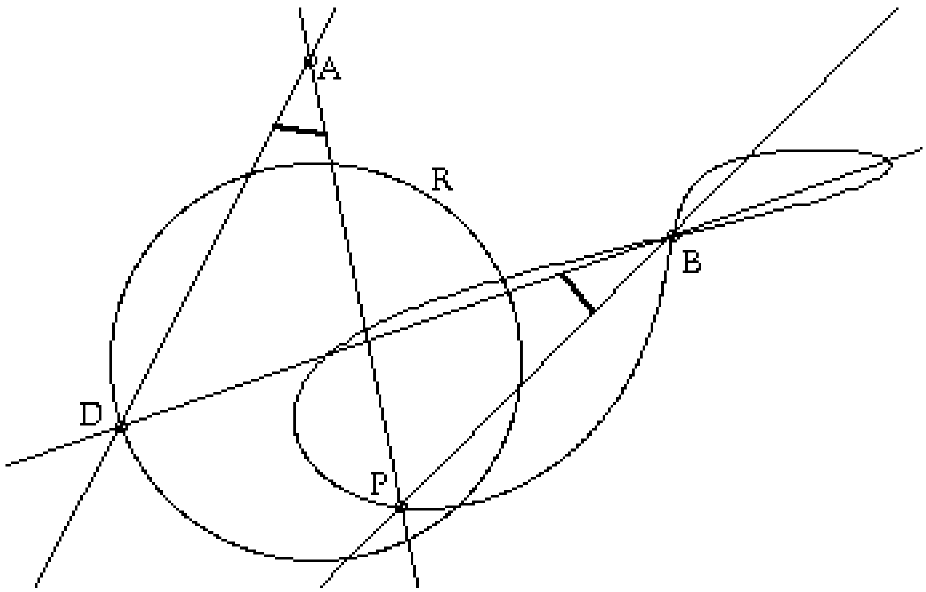


Figure 2.11d

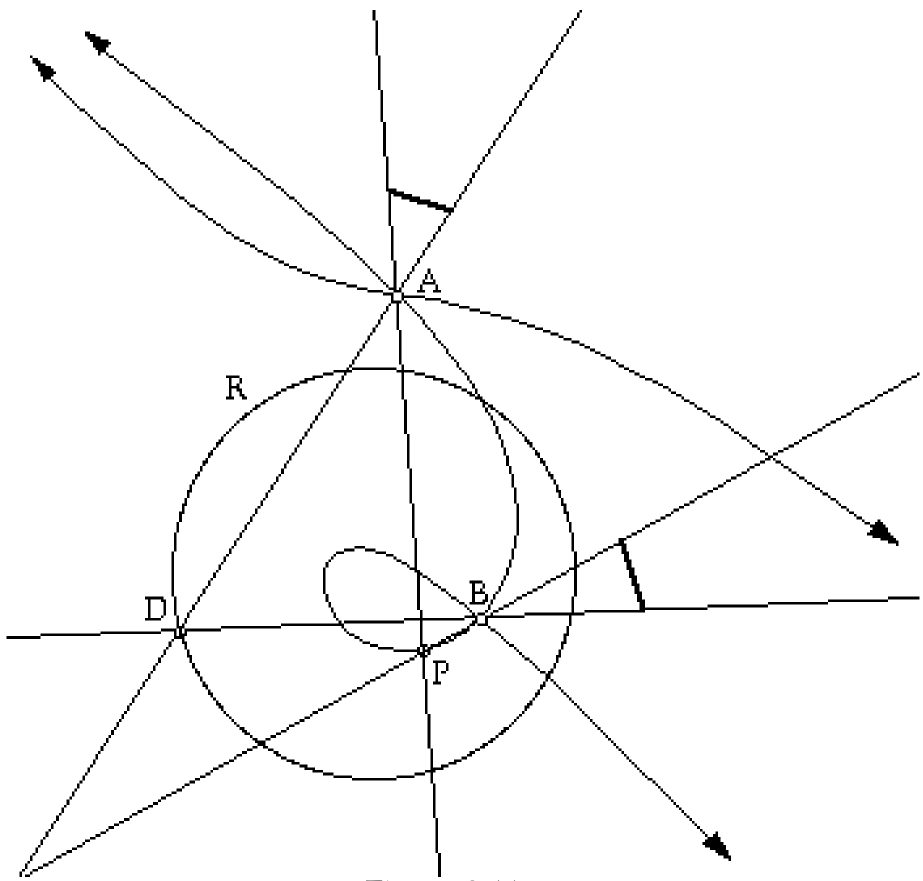


Figure 2.11e

In Figure 2.11e, I have moved the pole B inside the circle. There are now two asymptotes. The pole at A has now become a cross over point as well, but different branches of the curve cross at point A. By moving A inside the circle as well, both A and B become the intersection points of different branches, and the loop which once passed through B is now outside the circle (see Figure 2.11f). Newton's curve drawing device can provide many interesting and beautiful examples of the trends we have seen thus far. Iterating devices tends to jump the algebraic degree of the curves up by two, and pivot points in constructions tend to become singularities (Newton called them "principal points").

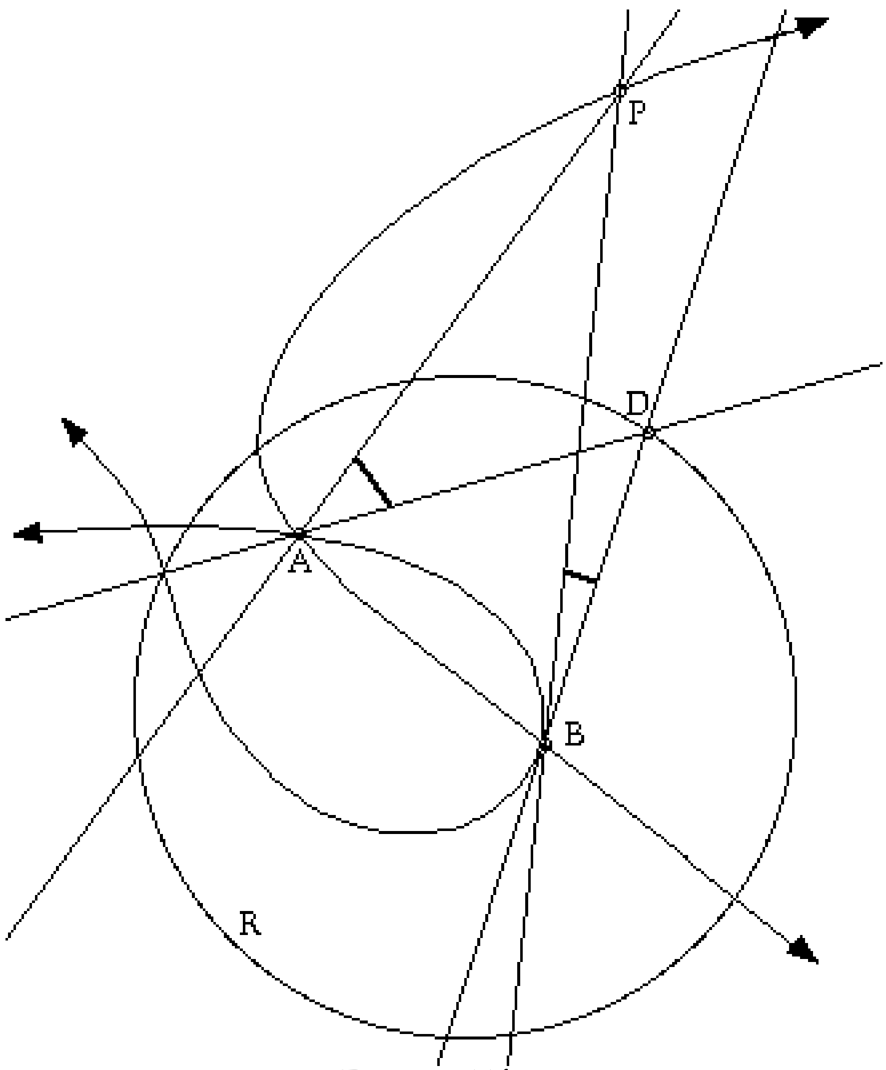


Figure 2.11f

This method of drawing curves readily solves a number of problems geometrically, but does not lend itself very easily to analytic methods. Writing equations from the geometric parameters in these devices can be quite difficult. Newton's first attempt, in 1668, at an algebraic analysis of this device ended in a hopeless tangle of equations (Newton, 1968, p. 152)[26]. It was not until many years later that he was able to give a complete analytic proof that when the directrix is a straight line, the curve described is a general conic (Newton, 1968, vol. 5). The

[26] For a constructivist educator, who wishes to work historically, these rambling worksheets of Newton's, filled with mistakes and comments could prove fertile material. They provide evidence that is almost the equivalent of a clinical interview.

main problem seems to be that the device depends primarily on a pair of fixed rotations, and rotations are difficult to deal with analytically within a fixed rectangular coordinate system.

There is an educational trend to discuss transformations earlier in our secondary curriculum, but this is usually restricted mostly to reflections, translations and changes of scale. If one is tied to a rectangular coordinate system as an epistemological prerequisite, then rotations are just too hard to deal with.[27] Curricular fixation on a narrow concept of function excludes general curve constructions and their natural parameters. This fixation leads to a general reluctance to discuss rotations, since when one rotates the graph of function, it is no longer the graph of a function in this narrow sense. This policy, that if something can't be discussed algebraically, then it's better not to mention it, limits our curriculum unjustifiably. Rotations are simple and natural physical actions, and they should be discussed early on in mathematics classes. Both Newton's device and those that defined ellipses and hyperbolas as the set of points equidistant from a point and a circle (Sections 2.5 and 2.6), depended on actions that involved two rotations. They were easily animated on *Geometer's Sketchpad and* revealed a variety of important aspects of the curves (e.g. the modeling of orbits or projective properties).

Newton's organic construction of curves leave many open mathematical questions. He never resolved many of his claims, but his private papers indicate the extent to which his experience with drawing curves provided a foundation of grounded activity from which many of his later researches benefited. I think that students who played with these devices might easily come up with a large number of

---

[27] For an example of how far this trend has gone, consider the entirely unnatural definition of a "rotation" given in the newly reformed *Geometry* textbook issued by the University of Chicago School Mathematics Project: "Definition: A rotation is the composite of two reflections over intersecting lines" (Coxford, et. al., 1993, Lesson 6-3, page 267).

interesting mathematical conjectures which could spur wonderful discussions. Lack of closed algebraic solutions should not deter such discussions, especially when students have ready access to computers with dynamic capacity for experimentation. Exposure to difficult and intriguing open mathematical question in this intuitive physical setting could create an open secondary classroom atmosphere like the one described by Lakatos in *Proofs and Refutations* (1976).

2.12 Descartes' Geometric Means and a Construction of Log Curves

This section contains two investigations. First, I will describe Descartes' linkage device for finding any number of geometric means between any pair of lengths. This device produces a family of curves which I will analyze briefly. Second, I will employ Descartes' device as part of a construction that will find any number of points, as densely as desired, on any logarithmic or exponential curve. This construction of log curves closely parallels some of the later work of Descartes, but uses a more modern setting. This modernization combines historical ideas with several theories and ideas of Jere Confrey concerning the use of computers, covariation, and multiple representations (1993a; 1992), and her work on exponential functions (1988).

In the *Geometry* (1952), Descartes considered the problem of finding n mean proportionals (i.e. geometric means) between any two lengths a and b (with $a < b$). That is to find a sequence of lengths, beginning with a and ending with b, such that the ratio of any two consecutive lengths is constant. In modern algebraic language, that is to find a sequence $x_0, x_1, \ldots, x_{n+1}$ such that for some fixed ratio r, $x_k = a \cdot r^k$, and $x_{n+1} = b$. Hence the terms of the sequence have a constant ratio of r, and form a geometric sequence beginning with a and ending with b.

Descartes began, as always, with a geometric construction. He imagined a series of rulers with square ends sliding along and pushing each other creating a series of similar right triangles. See Figure 2.12a, which is reproduced from the original 1637 edition of the *Geometry*, (1952, p. 46). Let Y be the origin with A and B on a circle of radius a. As angle $\angle XYZ$ increases, C moves further out the x-axis. The vertical from C then intersects the line $\overleftrightarrow{XY}$ at D which is still further from the origin. The triangles ΔYBC, ΔYCD, ΔYDE, $\Delta YEF, \ldots$ etc., are all similar, being right triangles all containing the angle $\angle XYZ$. Hence we have $\frac{YB}{YC}=\frac{YC}{YD}=\frac{YD}{YE}=\frac{YE}{YF}=\cdots$ etc. Therefore the sequence of lengths $a=YB,\ YC,\ YD,\ YE,\ YF,\ldots$ etc., form a geometric sequence.

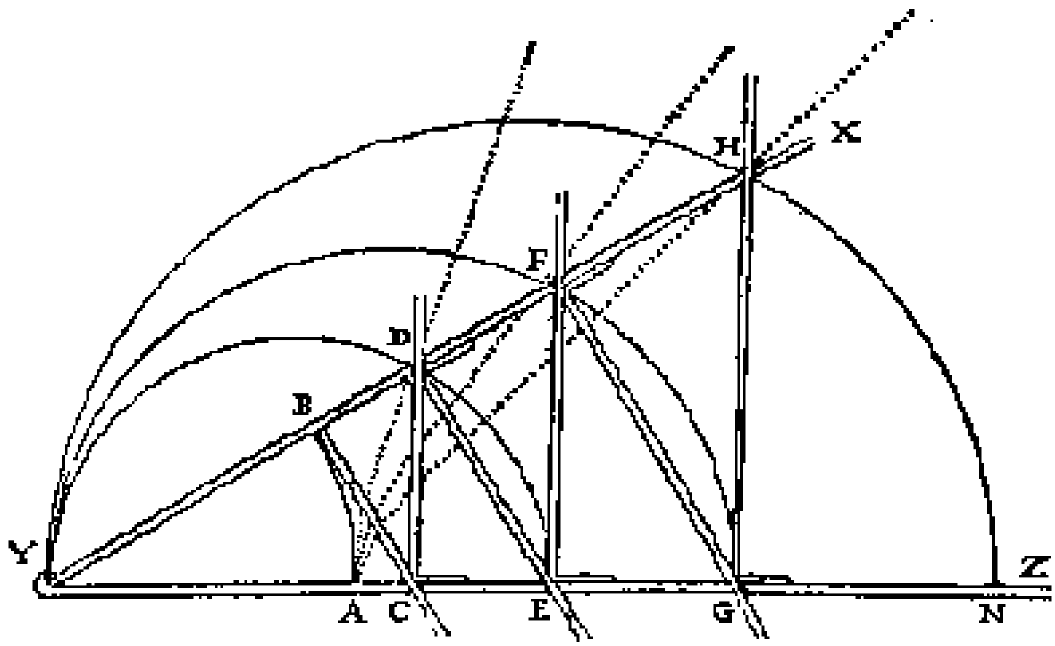


Figure 2.12a

If one lets $a=1$ and angle $\angle XYZ=60°$, one forms the sequence 1,2,4,8,16,32, ... Letting angle $\angle XYZ=45°$, one forms the sequence $1, \sqrt{2},\ 2,\ 2\sqrt{2},\ 4,\ 4\sqrt{2},\ 8,\ \ldots$, which is a refinement. As angle $\angle XYZ$ decreases, one obtains increasingly dense geometric sequences. In modern terms, the relationship between the constant ratio r and angle $\angle XYZ$ is: $\sec(\angle XYZ)=r$. This relationship is never mentioned in the *Geometry*. Descartes instead emphasized the curves traced by the points $D,\ F$, and H, shown in Figure 2.12a by the dotted lines. These curves all have algebraic equations, as opposed to the secant, which can only be computed with some infinite process.

To solve the original problem of finding n mean proportionals between a and b, Descartes suggested using the curves drawn by the device. If two mean proportionals are sought, mark off length $b=YE$ on the line $\overleftrightarrow{YZ}$. Next construct the circle having diameter $\overline{YE}$, and find its intersection D with the first of these curves (see Figure 2.12a). Then drop the vertical line from that point D to $\overleftrightarrow{YZ}$ to locate the point C. YC and YD

will then be the desired mean proportionals. This method uses the curve drawn by D to determine the appropriate angle of the device so that the point E will fall on any specified length b.

The equations of the curves traced by D, F, and H can all be found by successively substituting into the similarity relations upon which the device was built. To find these equations one can proceed as follows (this derivation comes from commentaries on Descartes published in 1730 by Claude Rabuel (Descartes, 1952, p.47)). Let $Y=(0,0)$, $D=(x,y)$, and let $YD=z$. Now $z^2=x^2+y^2$, but one also knows that $\frac{z}{x}=\frac{x}{a}$. Hence $z=\frac{x^2}{a}$, and therefore by substitution one obtains, for the path of point D, the equation: $x^4=a^2(x^2+y^2)$.

Now let $F=(x,y)$, and let $YF=z$. Now $\frac{z}{x}=\frac{x}{YD}$, hence $YD=\frac{x^2}{z}$. One also knows that $\frac{x}{YD}=\frac{YD}{YC}$, so substituting and solving for YC one gets: $YC=\frac{x^3}{z^2}$. Lastly one knows that $\frac{YD}{YC}=\frac{YC}{a}$, and hence: $\frac{ax^2}{z}=\frac{x^6}{z^4}$. Solving for z one obtains: $z=\sqrt[3]{\frac{x^4}{a}}$. As before, $z^2=x^2+y^2$, so $\sqrt[3]{\frac{x^8}{a^2}}=x^2+y^2$. Cubing both sides one obtains, for the path of point F, the equation: $x^8=a^2(x^2+y^2)^3$.

In a similar fashion, one can find that an equation of the curve traced by the point H is: $x^{12}=a^2(x^2+y^2)^5$. Note that all of these curves pass through the point $A=(a,0)$, and that as one moves from one of these curves to the next the degree of the equation always increases by four (on both sides of the equation), which is an increase of two of Descartes' classes. The curve traced by D is of the second class; the curve traced by F is of the fourth class; etc. To get from any one of these curves to the next one involves two perpendicular projections, each of which raises by one the class of the curve.

Descartes, after stating that, "there is, I believe, no easier method of finding any number of mean proportionals, nor one whose demonstration is clearer," (1952, p. 155) goes on to criticize his own construction for using curves of a higher class than is necessary. Finding two mean proportionals, for example, is equivalent to solving a cubic equation, and can be accomplished by using only conic sections (first class), while the curve traced by D is of the second class. The solution of cubics by intersecting conics had been achieved by Omar Khayyam, and was well known in seventeenth century Europe (Joseph, 1991, Berggren 1986). Descartes spent much of the latter part of the *Geometry* discussing the issue of finding curves of minimal class which will solve various geometry problems (1952).

Descartes expounded an epistemological theory which sought after a universal structural science of measure which he called "mathesis universalis" (Lenoir, 1979). Fundamental to his program was his classification of curves in geometry. He wanted to expand the repertoire of curves that were allowed in geometry beyond the line and circle, but he only wanted to include curves whose construction he considered to be "clear and distinct" (Lenoir, 1979). For Descartes this meant curves which could be drawn with linkages and classified by his system according to pairs of algebraic degrees. These curves he called "geometrical" and all others he called "mechanical."

This distinction is equivalent to what Leibniz would latter call "algebraic" and "transcendental" curves. Descartes viewed "mechanical" (i.e. transcendental) curves as involving some combination of incommensurable actions. Examples that he specifically mentioned are the spiral, quadratrix, and cycloid. These curves all involve a combination of rotation and linear motion that can not be connected and regulated by some linkage. The drawing of such curves involves rolling a wheel or the unwinding of string from a circle.

This is not to say that Descartes did not address himself to problems concerning these curves. I shall mention some of his thoughts on the cycloid in the next section. I turn now to logarithmic curves. Two years after the publication of the *Geometry*, Descartes addressed a problem that was sent to him by De Beaune, which asked for the construction of a curve in a skewed system where the ratios of the subtangents to the ordinates are everywhere equal to the ratio of the ordinates to a fixed segment, i.e. a type of logarithmic curve; De Beaune's requirement being, in modern terms, a first order differential equation (Lenoir, 1979). Descartes generated a method for point-wise approximation of this curve, and also provided a detailed study of how the curve could be drawn by a combination of motions with particular progressions of speeds. He then stated:

> I suspect that these two movements are incommensurable to such an extent that it will never be possible for one to regulate the other exactly, and thus this curve is one of those which I excluded from my *Geometry* as being mechanical; hence I am not surprised that I have not been able to solve the problem in any way other than I have given here, for it is not a geometrical line.
> (Descartes, quoted in Lenoir, 1979, p. 362)[28]

I will now proceed to construct point-wise approximations of logarithmic curves, but I will not follow the particular example discussed in Descartes' letters to De Beaune. That example turns out to have been a transformation of a logarithm added to a linear function. I will instead construct standard logarithmic curves using both the device of Descartes' shown in Figure 2.12a, together with the original conception of logarithms by John Napier, as pairings of geometric and arithmetic sequences (Smith & Confrey, 1994; Edwards, 1979; Boyer, 1968).

Although the theme of this entire thesis is curve drawing actions, this construction depends on a covariational view of functions (Confrey & Smith, 1995). This is essentially a view from tables of data that looks for methods to simultaneously

---

[28] For a fascinating social and philosophical analysis of why Descartes would adopt such an attitude see the article by Lenoir (1979). It certainly had nothing to do with his ability to contend with such problems.

extend or interpolate values in both columns rather than looking for a rule which calculates between different columns. This approach to functions was central in the thinking of Leibniz (Child, 1929), and has been shown to be important in the thinking of students (Rizzuti, 1991; Confrey, 1994a).

Napier and others in the early seventeenth century made tables of logarithms by placing arithmetic sequence alongside geometric sequences. They devised various ways to make these tables dense (Edwards, 1979). These early approaches to logarithms were entirely tabular and calculational, and did not involve curves or graphs. When Descartes constructed a curve as a solution to De Beaune's problem, he did not view the curve as a logarithm. A more flexible view that could go back and forth between curves, graphs and equations did not begin to evolve until the end of the seventeenth century, especially with respect to transcendental curves whose general coordinates could only be found using series expansions (Dennis & Confrey, 1993).

The following construction resembles a graph in the sense that points are being plotted, rather than continuously drawn. However, the points are plotted, not from numerical inputs or equations, but instead from a pair of continuous geometric actions. The numerical coordinates of the points come from after the fact measurements. The distinctions between curves, graphs, and functions are both poignant and slightly ambiguous here. The analysis of the curves and their tangents will follow in the system of Leibniz (Section 2.3), but the covariational construction of the points bends the concept in a slightly different direction.

The primary aim here is to provide modern students with a hands on way to build logarithmic and exponential curves through a series of simple geometric constructions using *Geometer's Sketchpad*. From the standpoint of covariation there is little difference between exponentials and logarithms. The pair of actions which builds one also builds the other. I have constructed the following curves as logarithms, but the same construction could be viewed as exponential by simply repositioning the constructive actions.

I start by building a simulation of Descartes' device for the construction of geometric sequences with $a = 1$ (see Figure 2.12b). Let O be the origin and let H be any point on the unit circle. By moving H around the circle, the distances of the labeled points from the origin will form geometric sequences with any common ratio. That is if $r = OG_1$, then $r^2 = OH_2$, $r^3 = OG_3$, $r^4 = OH_4$, etc.

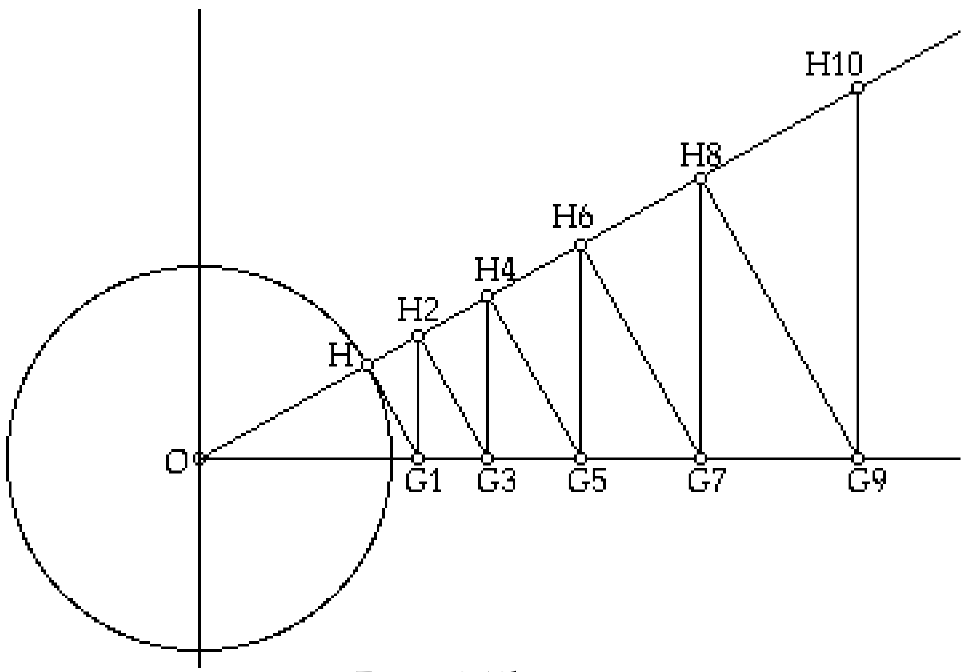


Figure 2.12b

This construction can also be extended to the interior of the unit circle to obtain segments whose lengths are the negative powers of r. Once again, as with the preceding construction, the odd powers of r are on the horizontal while the even powers of r are on the line $\overleftrightarrow{OH}$. This can be seen, as before, by considering the series of similar triangles with common vertex O. See Figure 2.12c.

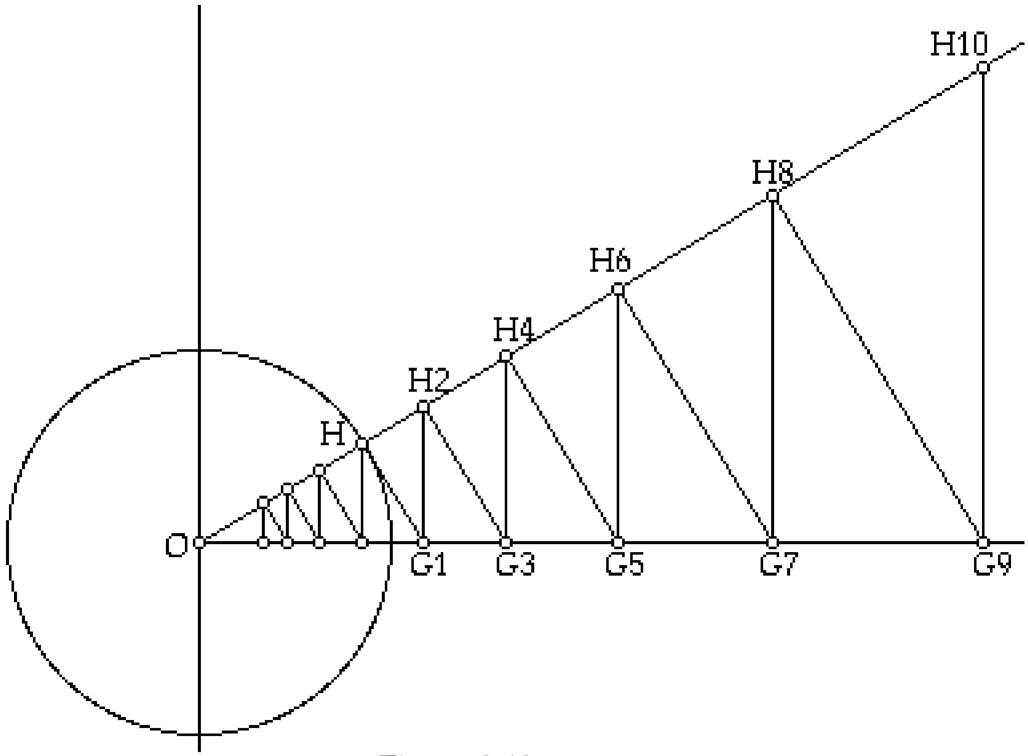


Figure 2.12c

In order to get the entire geometric sequence on one line, I will now translate the lengths marked on the line $\overleftrightarrow{OH}$ down to the horizontal by using circles centered at O. See Figure 2.12d. Thus I now have a geometric sequence, laid out on the x-axis, whose common ratio or density can be varied as the point H is rotated. The point where the circle through H_2 intersects the x-axis I will name G_2, likewise for H_4, H_6, etc. The points on the x-axis that are inside the unit circle I will call G_{-1}, G_{-2}, … etc. where the subscripts correspond to the powers of r which represent their distances from O.

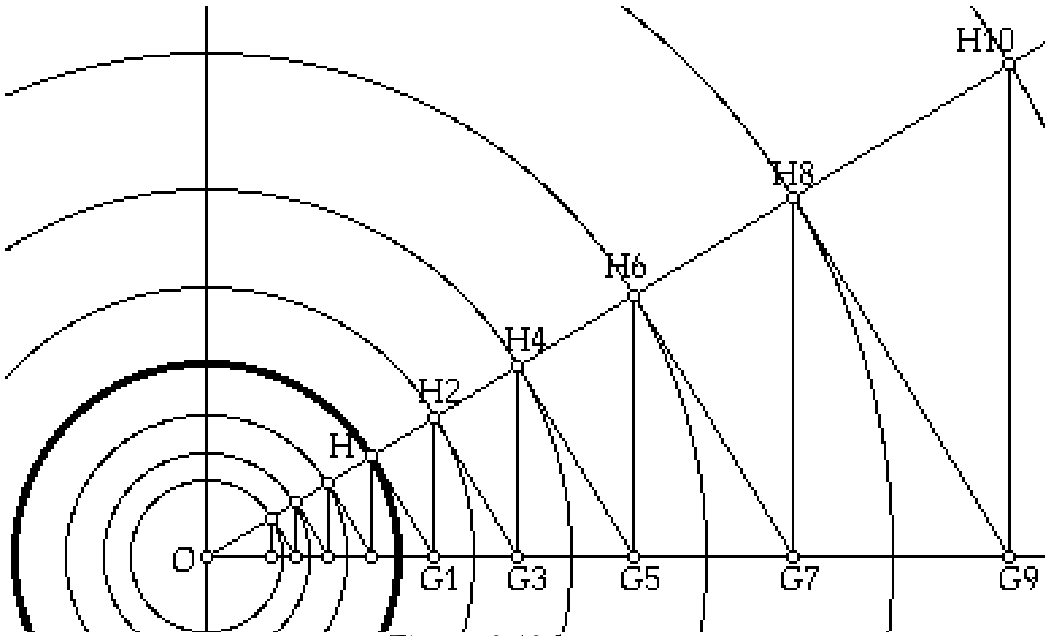


Figure 2.12d

In order to construct logarithmic curves, we must now construct an arithmetic sequence A_i on the y-axis with a variable common difference. This can be achieved in a variety of ways. The common difference d in the arithmetic sequence will be adjusted here by moving A_{10} along the y-axis. The points in the geometric sequence are then each vertically translated by lengths corresponding to consecutive points in the arithmetic sequence, thus creating the points (G_i, A_i). These new points then lie on the graph of a logarithm. I then connect the new points by line segments (bold in Figure 2.12e) to approximate a log curve. Using the meters available in *Geometer's Sketchpad*, one can continuously monitor the lengths in both the sequences, and hence the coordinates of the points (G_i, A_i) on the log curve. Some of the construction lines have been hidden for greater visual clarity.

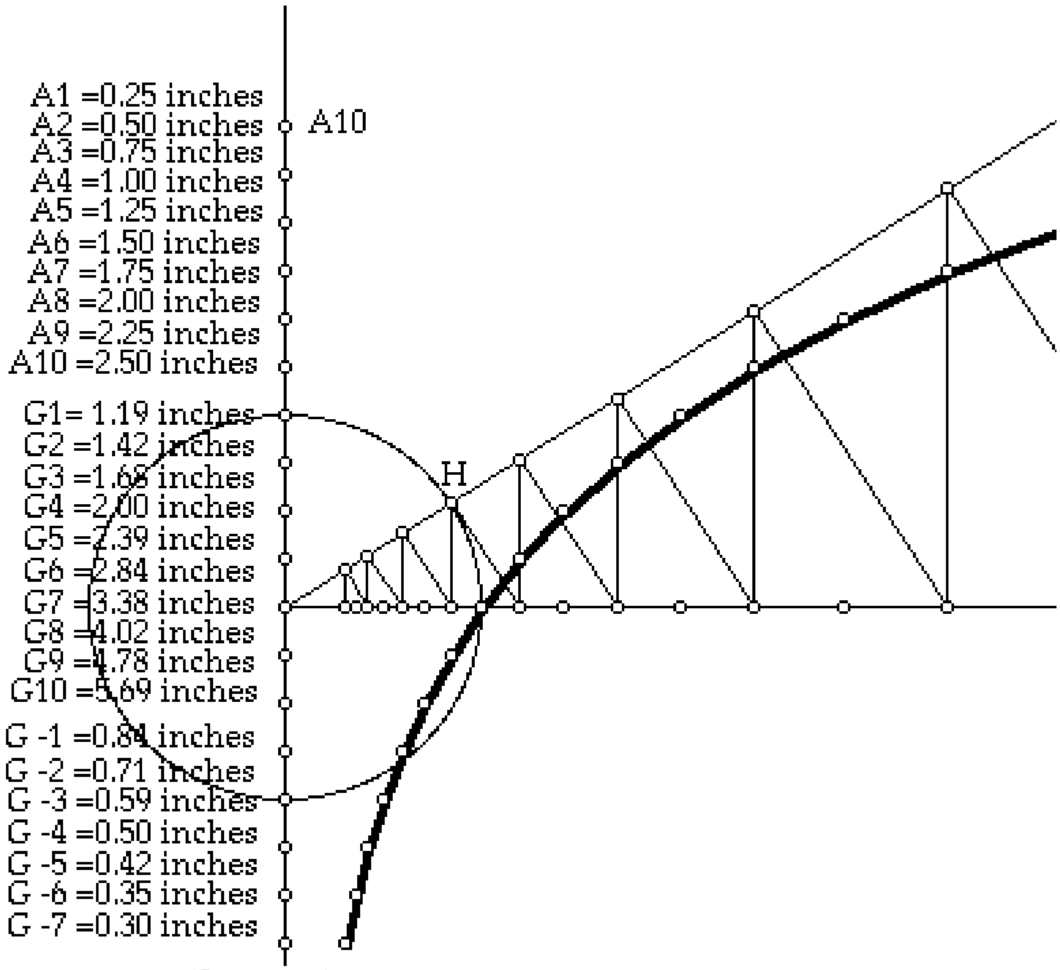


Figure 2.12e

This construction yields an adjustable curve. By moving H around the unit circle, or A_{10} along the y-axis, one can map any geometric sequence against any arithmetic sequence. In Figure 2.12e, the point H is adjusted so that $G_4 = 2$, and the point A_{10} is adjusted so that $A_4 = 1$. Hence this curve follows a graph of the log base 2. By readjusting A_{10} so that $A_8 = 1$ the curve shifts dynamically to become a graph of the log base 4. See Figure 2.12f.

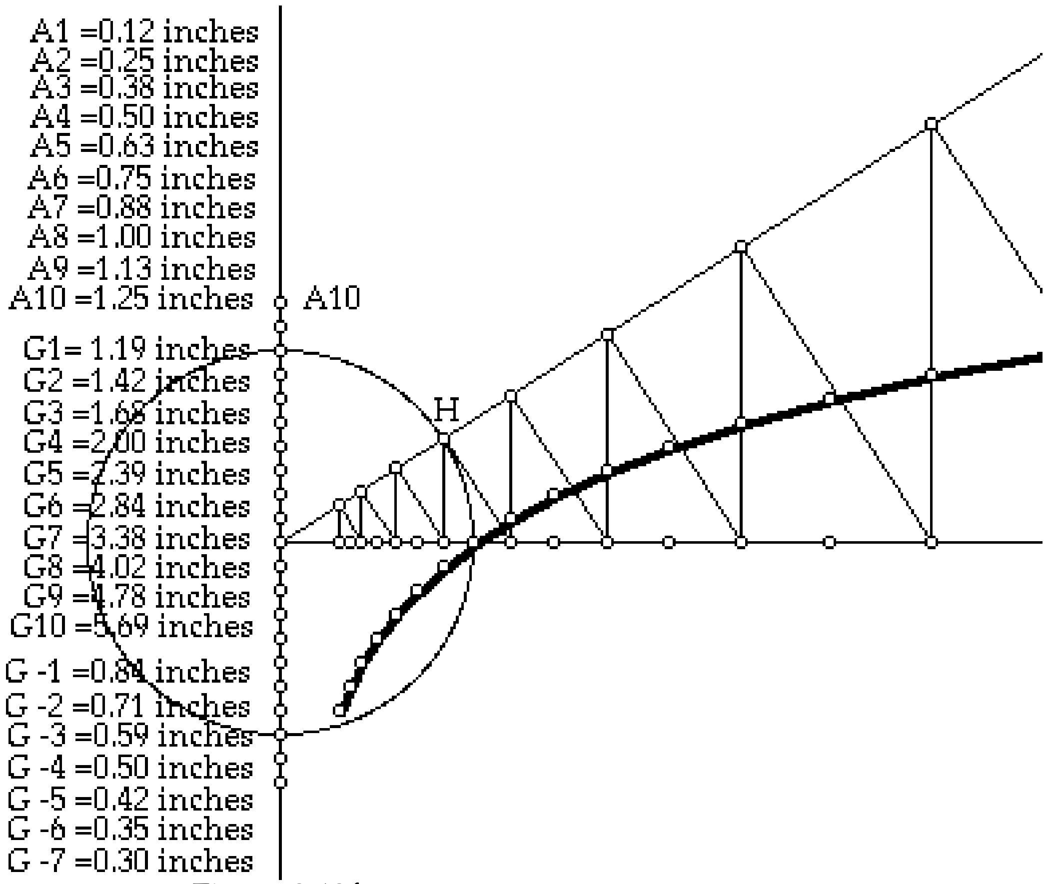


Figure 2.12f

By readjusting the point H so that $G_8 = 3$, one obtains a curve that follows the graph of the log base 3. Since the monitors are measuring to the hundredth of an inch, which is smaller than a pixel, it is not always possible to get exactly the numbers desired (e.g. G_8 reads 2.96 in Fig. 8). See Figure 2.12g. One can get around the pixel problem by using the appropriate rescaling window commands, but for a first experience this would decrease the sense of a direct physical approach, which I feel is important for students.

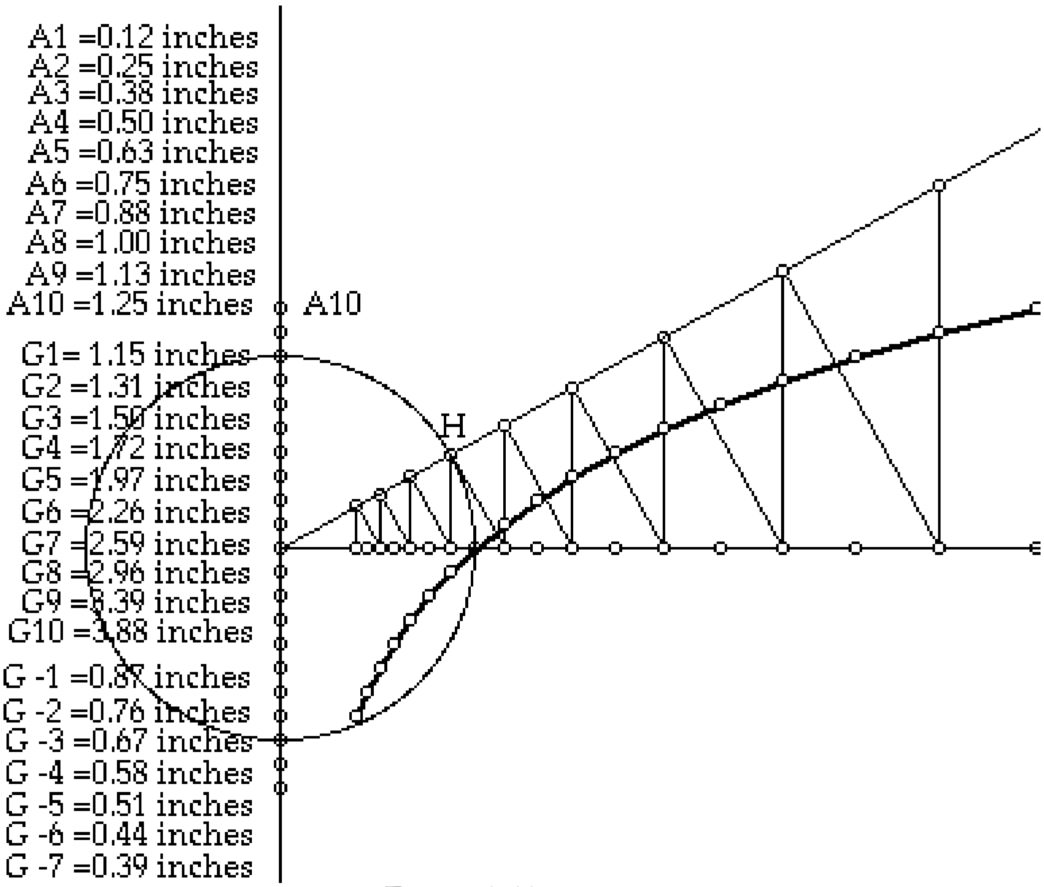


Figure 2.12g

It is fascinating to watch these curves flex and bend as the arithmetic and geometric sequences are manipulated. Even when the points are quite broadly spaced, as in Figure 2.12e, the graphs look very smoothly curved, even though they are made up of line segments. When the angle of H is increased the geometric sequence spreads out rapidly off of the screen. By scanning far to the right it is instructive to see just how incredibly flat log curves become.

When the arithmetic and geometric sequences are both spread out the graphs can eventually become "chunky" since the points are being connected with line segments. However, by manipulating both sequences it is possible to increase the density of points on any particular log graph without changing the base. For example, we could create another curve that follows the graph of the log base 2 by setting $A_8 = 1$ and $G_8 = 2$ (see Figure 2.12h). This is the same curve as that in Figure 2.12e, but with a much higher density of constructed points. Descartes' device allows us to geometrically carry out the calculational aims of Napier. Geometric sequences can be built as densely as one desires, and paired against any arithmetic sequence.

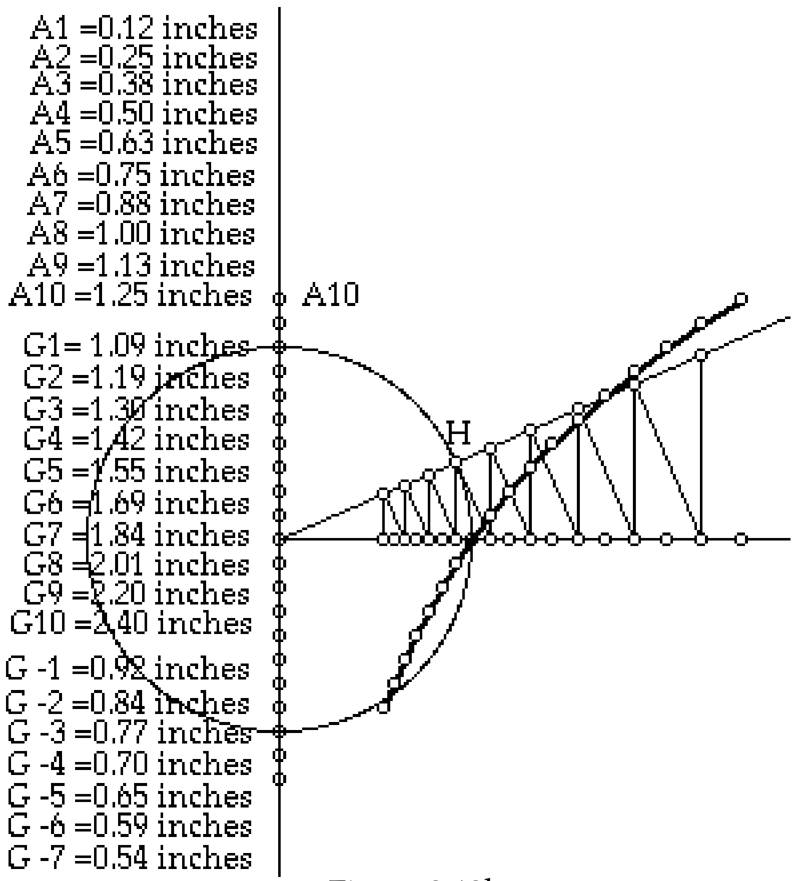


Figure 2.12h

After watching these log curves shift and bend dynamically, one can begin to look carefully at the slopes between points on the curves. Several interesting patterns come to light.[29] Suppose one wants to use the slopes between constructed points to approximate the tangent slope at a point, say for example at (1,0). It is visually apparent that using the point (1,0) in the calculation is not the best thing to do. The slope between the nearest points to the right and left gives a better approximation of the

[29] Thanks to David Henderson for his probing questions on this point.

tangent slope. This is true for most curves, not just the logarithm.[30] Here, at (1,0), I want to calculate the secant slope between G_{-1} and G_1. Letting r equal the common ratio of the geometric sequence, and d equal the common difference of the arithmetic sequence, one calculates the general slope at (1,0) as:

$$\text{slope at } (1{,}0) \;=\; \frac{2d}{r-\dfrac{1}{r}} = \frac{2rd}{r^2-1} = k.$$

Suppose one now approximates the slope at any other point on the constructed curve. This approximate slope at $(G_n, A_n) = (r^n, nd)$, is found by computing the secant slope between G_{n-1} and G_{n+1}. The calculation yields:

$$\text{slope at } (G_n, A_n) = \frac{2d}{r^{n+1}-r^{n-1}} = \frac{1}{r^n}\cdot\frac{2rd}{r^2-1} = \frac{k}{r^n}.$$

Here one has the approximate tangent slope at a point on a logarithm written as the inverse function times a constant. The constant k is the slope of the curve at (1,0). Of course these slopes are all approximations, but once the slope at (1,0) is approximated it can be divided by the x-coordinate at any other point to get the corresponding slope approximation at that point. By making the constructed points on the curve denser the approximations all improve together at the same rate. Thus the essential derivative property of logarithms is revealed without recourse to the usual formalisms of calculus. In fact, even more is being displayed here than the usual derivative of a logarithm. One sees that the all the slope approximations converge uniformly as the density of the constructed points is increased.

This constant k can be seen geometrically in another way. If we view these curves and tangent constructions using the vertical axis (i.e. as exponential curves), then

---

[30] It is strange that when the derivative is developed in calculus classes, it is defined using secant slopes from the point in question, rather than around the point. It would seem that nobody is directly interested in secant slope approximations, except as an algebraic device from which to define a limit. The practical geometry of secant slopes is ignored.

we find that the subtangent (see Section 2.3) is constant for all points, and is always equal to k, the slope at (1,0). This can be established algebraically from the previous discussion, but it is nice to see it geometrically on the curve and verify it using the meters on *Geometer's Sketchpad*. This is done in Figure 2.12i, for two different points on the log base 3 curve. The tangent lines and slope are approximated by using the points adjacent to the one under consideration, and the accuracy is quite good (a calculator gives $k = .910$).

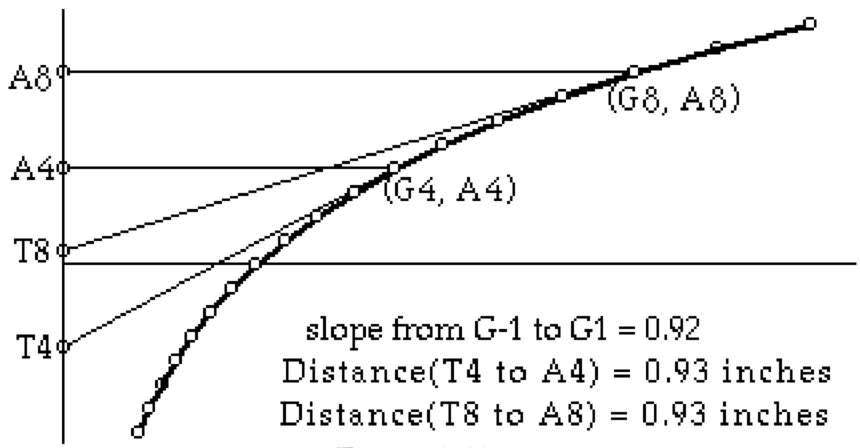


Figure 2.12i

This constant subtangent property was at the heart of Descartes' discussion of De Beaune's curve. The constant subtangent was the hallmark by which logarithmic and exponential curves were recognized during the seventeenth century (Lenoir, 1979; Arnol'd 1990). One way to think of this property is to imagine using Newton's method to search for a root an exponential curve. The method will march off to infinity at a constant arithmetical rate, where the size of the steps will be the constant k.

Compare this constant subtangent property of exponential curves to the properties of parabolas discussed in Section 2.4. For the parabola, the subnormal remained constant while the subtangents got very large. For the log curves just the opposite is the case. In standard modern functional language, the comparison would be between exponential functions and multiples of the square root function.

In order to construct the natural logarithm, one wants the slope at (1,0) to be 1. This is the property from which Euler first derived the number "e" (Euler, 1988). Returning to the construction, with a meter which monitors the approximate secant slope at (1,0), I now rotate H until the slope meter reads as close to 1 as possible. This constructs a close approximation to the graph of the natural logarithm. The approximate slope at any point on the curve is the inverse of its x-coordinate. Note that since $A_5 = 1$, the value of G_5 is approximately the number "e". See Figure 2.12j.

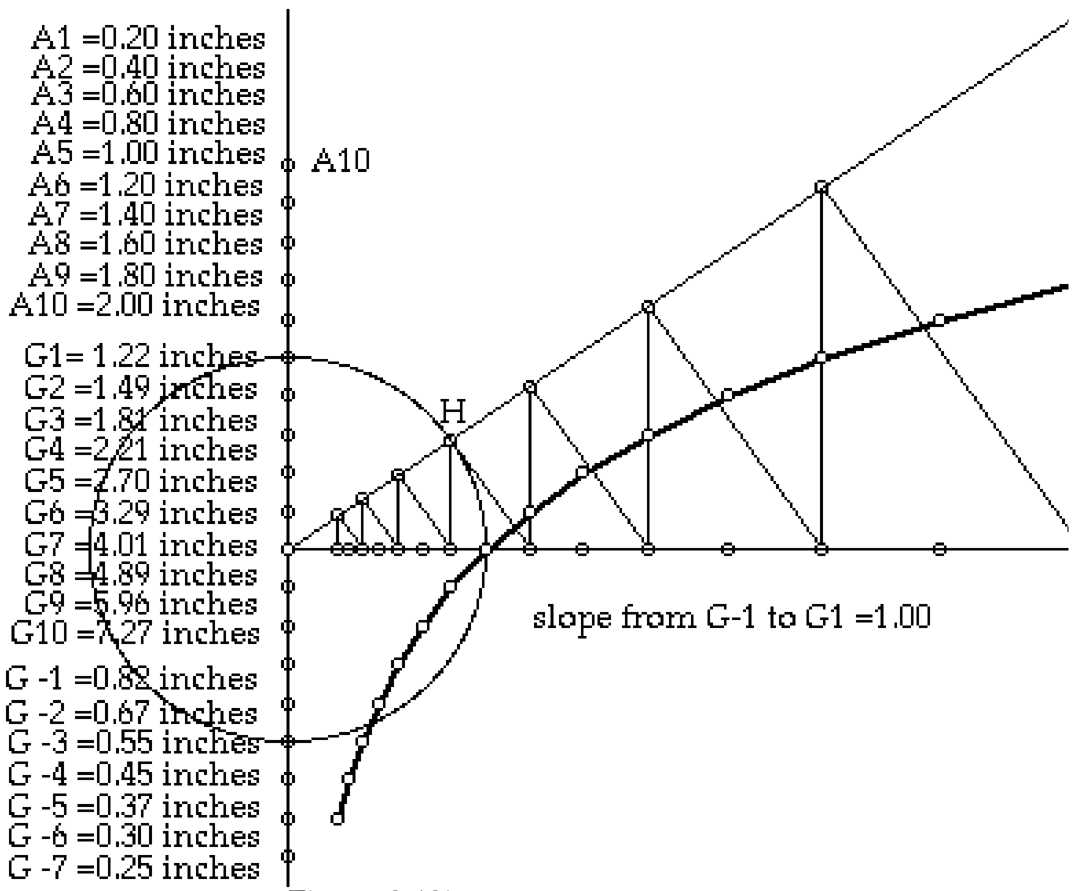


Figure 2.12j

This geometric construction of points on log curves achieves the goals set out by Napier. It allows one to construct logarithms (and also exponents) as densely as one desires. Of course Napier achieved these goals algebraically (Edwards, 1979), and throughout the seventeenth century increasingly subtle table calculations were developed, e.g. those of Wallis (Dennis & Confrey, 1993). The story of these calculational techniques is a very important one leading eventually to Newton's development of binomial expansions for fractional powers (Newton, 1967). Euler routinely used Newton's techniques to calculate log tables to over 20 decimal places (Euler, 1988). Theoretically, the geometric construction has unlimited accuracy, but directly using the meters in the software, *Geometer's Sketchpad*, one is limited to at most three decimal places.[31] Nevertheless, in order to discuss logarithms from the standpoint of an epistemology of multiple representations (Confrey & Smith 1995), a dynamic geometrical construction of the curves is very illuminating.

I ask the reader to compare this investigation of logarithms with the approaches more frequently taken in classrooms. Many students are introduced to logarithms in a formal algebraic way, with no references to geometry or to table construction. Such students often have no method for geometrically or numerically constructing, even a square root. Such an approach leads, at best, only to a superficial understanding of the grammar of logarithmic notation. There is no dialogue at all.

Another approach that is taken is to see logarithms as the accumulated area under a hyperbola (usually $y = 1/x$). This approach can provide many fascinating insights that connect logarithms to both geometry and to the numerical construction of tables. The study of hyperbolic area accumulation was fundamental in the early work of Newton, as he extended the table interpolations of John Wallis and created his first infinite binomial expansions (Dennis & Confrey, 1993; Edwards, 1979). Although this approach can create a fascinating and balanced dialogue, it is not usually taken with students until they are already involved with calculus. The fundamental theorem of calculus, for example, is usually invoked to show that the hyperbolic area function must have a derivative of $1/x$.

---

[31] By changing the scale in the meters by division unlimited accuracy can be achieved geometric ally.

The approach that I have described here is strictly pre-calculus. It involves only a systematic use of similar triangles, in a hands-on setting that is both visual, physical and geometric. It provides a specific form of grounded activity that allows to students to manipulate, extend, and interpolate both logarithms and continuous exponents. Rather than using calculus to create a balanced dialogue, this approach uses the dialogue to achieve some of the results of calculus in a very simple setting. It highlights the power of iterated geometric similarity (Confrey, 1994a), and provides another independent experience which can be used to validate the linguistic construction of calculus.

Reading this section can not truly convey the feeling one gets while physically manipulating the curves. The investigation of the slopes of log curves depends logically only on the properties of a table which maps a geometric sequence against an arithmetic sequence, but I did not notice this piece of algebra until many fluctuating examples of log curves had appeared on the screen. The geometry can heighten the intuition so that the appropriate question emerges. The power of suggestion should not be underestimated. The association of rotation around the unit circle with the building of logarithms is a wonderful foreshadowing of the connections between these functions and the trigonometric functions when extended to the complex numbers (Euler, 1988).

2.13 Roberval: Cycloids and Sine Curves

The circle is the curve with which we all have the most experience. It is an ancient symbol and a cultural icon in most human societies. It is also the one curve whose area, tangents, and arclengths are discussed in our mathematics curriculum without the use of calculus, and indeed long before students approach calculus. This discussion can take place, because most people have a lot of experience with circles, and know several ways to generate them. Pascal thought that, second only to the circle, the curve that he saw most in daily life was the cycloid (Bishop, 1936). Perhaps the large and slowly moving carriage wheels of the seventeenth century were more easily observed than those of our modern automobile, but the cycloid is still a curve that is readily generated and one in which many students of all ages easily take an interest. In a variety of settings, when I have mentioned, for example, the path of an ant riding on the side of a bicycle tire, some immediate interest has been sparked (see Figure 2.13a).

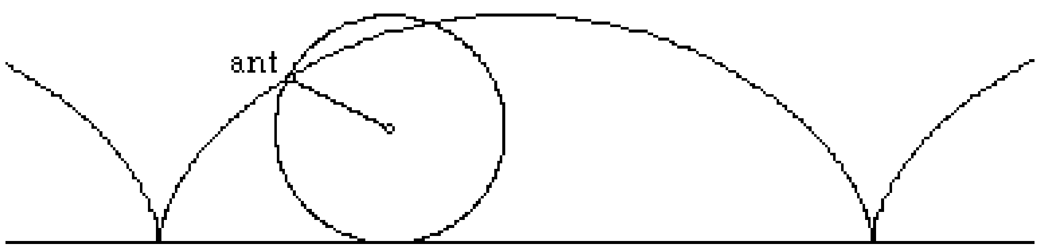


Figure 2.13a

The cycloid played an important role in the thinking of the seventeenth century. It was used in architecture and engineering (e.g. Wren's arches, and Huygens' clocks). As analytic methods were developed, their language was always tested against known curves, and the cycloid was the preeminent example for such testing (Whitman, 1946). Galileo, Descartes, Pascal, Fermat, Roberval, Newton, Leibniz and the Bernoullis, as well as the architect, Christopher Wren, all wrote on various aspects of the cycloid.

Besides the fact that it can be easily drawn, what makes this curve an excellent example for this discussion is that its areas, tangents, and arc-lengths were all known, from the geometry of its generation, many years before Leibniz first wrote an equation for the curve in 1686 (Whitman, 1946).

Some early observers thought that perhaps the cycloid was another circle of a larger radius than the wheel which generated it. Some careful observation will dispel this belief; for example at the cusps where the traced point touches the ground the tangents are already vertical, but this section of the curve is clearly not a half circle.

Galileo used the curve as a design for the arches of bridges. For this reason he sought to determine the area under one arch of a cycloid. He approached the problem empirically by cutting the shape out of a uniform sheet of material and weighing it. He found that the shape weighed the same as three circular plates of the same material cut with the radius of the wheel used to draw the curve. Galileo tried this experiment repeatedly and with care, and found again that the ratio of the area of the cycloidal arch to that of the wheel which drew it was three to one. He suspected however that the ratio must be incommensurable, probably involving π, and abandoned further attempts to more accurately determine the ratio (3:1 is correct as we shall see). Galileo gave the name "cycloid" to the curve, although it has also been known as a "roulette" and a "trochoid" (Struik, 1969; Whitman, 1946).

A French mathematician, Gilles Personne de Roberval (1602 - 1675), wrote a tract in 1634 that included both the area and tangent properties of the cycloid (Struik, 1969). This work was done just before the publication of Descartes' *Geometry*, and several important issues are raised by Roberval's mechanical methods which involve no algebra. He began by imagining a point P on a wheel drawing a cycloid, and at the same time observing a second point Q drawing a second curve which he called the "companion of the cycloid." This second point Q has, at all times, the same elevation off the ground as P, but always rides on a vertical diameter of the wheel. Q can be thought of as the projection of P onto the vertical diameter of the wheel. See Figure 2.13b which shows both curves traced by *Geometer's Sketchpad*. Q will move forward at a constant speed while monitoring the height of P. The path of Q is, therefore, what is now know as a sine or cosine curve.

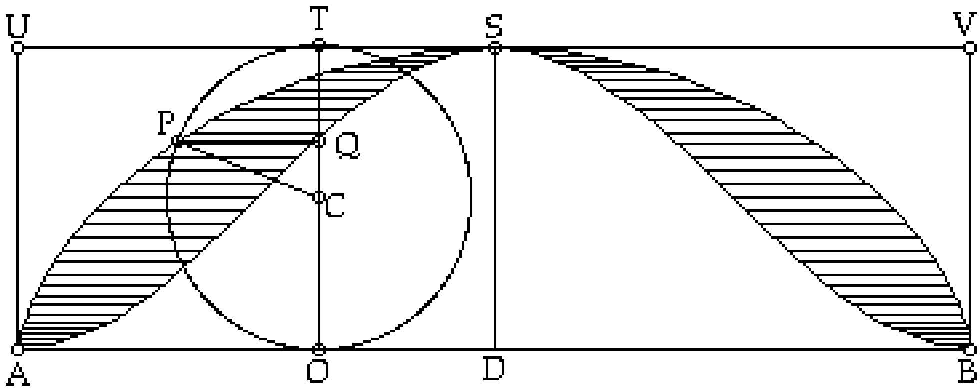


Figure 2.13b

Points P and Q start together at A, and come together again at S. In between the distance between P and Q takes on all the different horizontal segments that occur in half of the circle (i.e. all of the horizontal line segments $\overline{PQ}$ that form the shading). Thinking of the shaded area between the curves from A to S as a deck of cards, if one pushes them against a vertical line, they will form a half circle. Hence the entire shaded area in figure 2.13b is equal to the area of the circle. This reasoning employs what is known as the method of Cavalieri, also known as the method of indivisibles.

Looking at the symmetry of the companion curve traced by Q between A and S told Roberval that the area under that curve is one half the area of the entire rectangle $ABVU$. The entire rectangle has dimensions equal to the diameter and the circumference of the wheel, and is therefore equal to four times the area of the wheel (i.e. $(2\pi\, r)(2r) = 4(\pi\, r^2)$). The area under the cycloid is the shaded area plus the area under the companion curve, and therefore equals three times the area of the wheel that generated the curves, just as Galileo's weighing experiments had indicated.

Another way of stating this result is to say that the area of the cycloidal arch is always $3/4$ of the rectangle that contains it. Other mathematicians of the time (e.g. Wallis and Newton) would have called three quarters the characteristic ratio of the curve (Dennis & Confrey, 1993). This tradition goes back to ancient mathematics, like the result of Archimedes that says that if the curve under consideration was any downward parabola then the area under the curve would be $2/3$ of the rectangle containing it.

Roberval obtained tangents to the cycloid by thinking of the motion of point P as two separate motions, one rotational and the other forward (Struik, 1969). Since the wheel is rolling smoothly without slipping the rotational speed of the wheel must equal its forward speed (see Figure 2.13c). One can then construct the tangent as the sum of these two equal velocities. Thus Roberval constructed the tangent at P by considering a tangent to the circle at P (PH=rotational velocity), and a horizontal of the same length (PQ=forward velocity), and then forming the parallelogram on these two segments, and then drawing the diagonal $\overline{PV}$. Since $PH = PQ$, $\overline{PV}$ will bisect the angle $\angle HPQ$.[32]

[32] Roberval applied this same method of finding tangents by components to the parabola and the ellipse. For example a point on a parabola is increasing (or decreasing) its distance from the focus at the same rate as it is increasing (or decreasing) its distance from the directrix. Bisecting the angle, or drawing the diagonal between appropriate equal segments will yield the tangent. This is nearly the same tangent construction as van Schooten's (see Section 2.4).

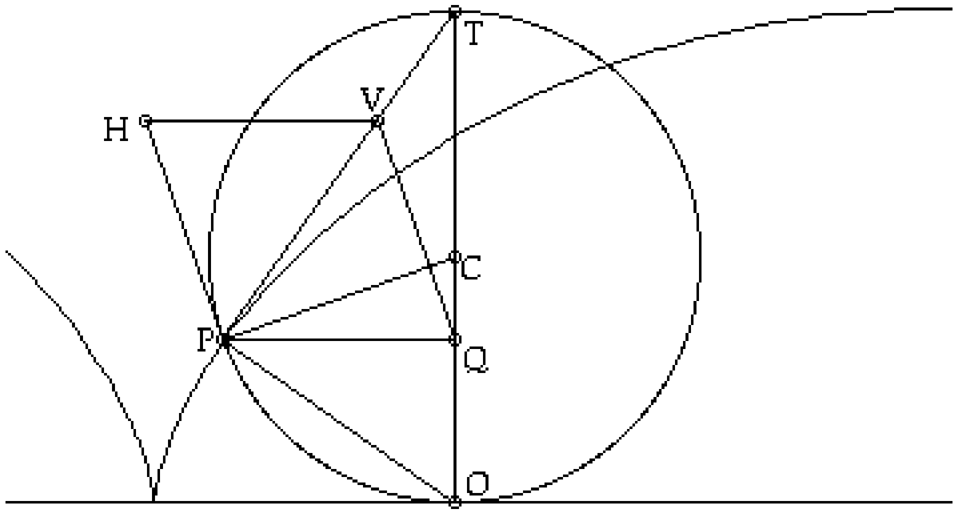


Figure 2.13c

Since $\overline{PH}$ is perpendicular to the radius $\overline{CP}$, and $\angle CPT = \angle CTP$ (isosceles triangle), and $\angle TPQ = 90°$, then $\angle HPV = 90° - \angle CPT = \angle TPQ$. Hence the bisector $\overrightarrow{PV}$ of angle $\angle HPQ$ lies along the line $\overleftrightarrow{PT}$. One can deduce from this geometry that this tangent $\overleftrightarrow{PV}$ to the cycloid at P always points at the top of the rolling circle T. Look back at Figure 2.13b to see the tangent $\overleftrightarrow{PT}$ in another position. Thus the ant on the bicycle wheel is always moving directly towards or away from the top of the wheel.

One can also deduce from Figure 2.13c that the tangent to the cycloid is always perpendicular to the line $\overleftrightarrow{PO}$ which connects P to the point of contact of the wheel with the ground. In 1638, Descartes saw this directly by approaching the tangent problem in a different way. Instead of a circular wheel, he started by imagining a rolling convex polygon (e.g. a square wheel). Such a figure pivots on one vertex until a side comes down flat on the ground and then it shifts to pivot on the next vertex. Thus any point P, moving on a rolling polygon, will have as its path a series of circular arcs of different radii. While the polygon is pivoting on any one vertex, the path of that point P will be a circle centered at that vertex, and thus its tangent will be perpendicular to the line connecting P to that vertex (i.e. the point of contact with the ground). Descartes then imagined regular polygons with an increasing number of sides, becoming closer and closer to a circle. From this he deduced that the tangents at each point P on a cycloid must always be perpendicular to the segment $\overline{PO}$ which connects that point P with the point of contact O of the wheel with the ground (Whitman, 1946).

The approaches of Roberval and Descartes to this problem display their different conceptions of mathematics. Roberval thought in terms of engineering and mechanics. He saw the cycloid as two combined motions and resolved them using the parallelogram law, in the manner of Galileo. Descartes' approach is more geometrical, and involves seeing a circle as a limit of polygons (a ancient view taken, for example, by Archimedes). Descartes called the cycloid one of the "mechanical" curves that he refused to admit to his *Geometry*, because the regulation of its motion was not "clear and distinct" (i.e. it involved matched simultaneous rotation and forward motion).

If a wheel rolls at a constant rate, both of these approaches will yield not only the tangent to the path of motion at each point (i.e. the direction of velocity), but also the magnitude of the velocity vector as well. With Roberval's construction, if the wheel is rolling at a constant rate, then the horizontal velocity has constant magnitude, and by adding it to a vector tangent to the circle, and of the same magnitude as the horizontal velocity; one can, at all points, construct the cycloidal velocity vector. Using Descartes' conception of polygonal rolling motion, and thinking of the rotational rate at each pivotal contact point as constant, one can see that the magnitude of the velocity vector is proportional to the distance of the moving point from the contact point. This will remain true as the polygons approach the circle.

One sees, in either case, that the velocity is zero at the cusp of the cycloid when the point P touches the ground, and twice the forward velocity of the wheel when P is at the top of the wheel. Using Roberval's conception, this can be nicely animated using *Geometer's Sketchpad* (see Figure 2.13d). At the point of contact with the ground the two motions (rotational and forward) cancel each other, and the velocity vector is zero. At the top they are both in same direction and the velocity is at its maximum of double the constant forward speed of the wheel.

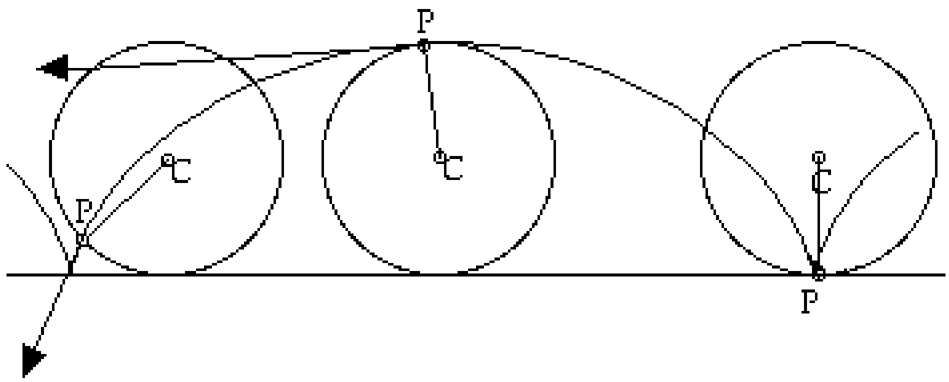


Figure 2.13d

As the seventeenth century progressed interest in the cycloid intensified and a variety of mathematical, physical, and engineering questions were investigated by Pascal, Huygens, Leibniz, Bernoulli and others (Arnol'd, 1990; Whitman, 1946; Whiteside, 1961; Smith, 1959). I will present one more investigation of the cycloid that gives the arclength of any portion of the curve in a simple geometric form. I first found this "rectification" (i.e. the finding of a straight segment equal to a given arclength) in the early notebooks of Newton from 1668 (1968, p. 193), but it also appears in a tract by John Wallis of 1659 (1972, p. 536). It was attributed by Newton to the famous London architect Sir Christopher Wren from a tract written in 1658 (Newton, 1968; Whitman, 1946). Like Galileo, Wren saw the cycloidal arch as well suited for architecture.

We have already seen that for any point on the rim of a rolling wheel, the segment that connects the point with the top of the wheel is tangent to its path of motion. Wren showed that the length of this segment is always exactly one half of the arclength between the point and the top of the cycloidal arch on which it is moving. That is to say in Figure 2.13e, the length of segment $\overline{QT}$ is exactly on half the arclength between P and T.

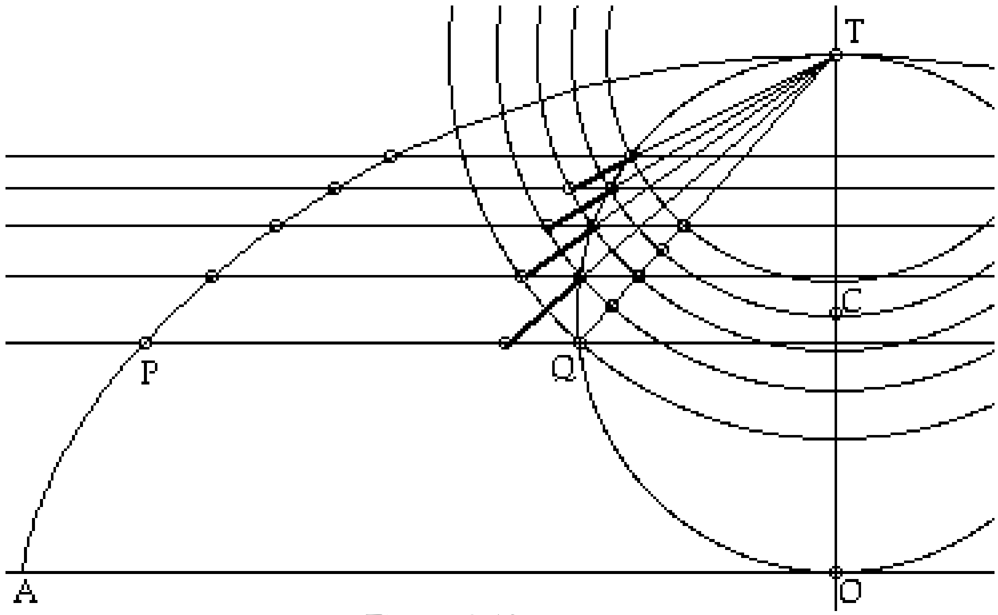


Figure 2.13e

$\overline{QT}$ is parallel to the tangent at P. Wren, like others in his time, imagined a curve to be made up of small line segments. Wren then imagined a series of points along the curve. Lines parallel to their tangents are shown radiating from T, and a series of circles, centered at T, pass through the intersection of these lines with the circle TQO. Each of the small darkened line segments is equal to a small tangent segment to the curve. Figure 2.13e then shows that the segment $\overline{QT}$ is the sum of pieces each of which are half of one of these tangential pieces of arclength. Thus twice QT must equal the entire arclength from P to T.

I find this theorem startling in its simplicity, especially after having calculated arclengths using the integral formulas from a calculus book. Wren presented his argument in the turgid formal Greek style known as the "method of exhaustion," but Newton provided only slightly more than what I have already said (1968, p. 193). The use of such methods was becoming quite natural to Newton (and also to Leibniz as we shall see in the next section).

This arclength property implies that the length of one entire cycloidal arch is exactly four times the diameter of the wheel which generated the curve. The circumference of the wheel is π ($\approx$ 3.14) times the diameter. For a point to traverse one cycloidal arch the wheel must revolve once. The extra distance that is added by the forward motion stretches the path of motion from <u>π diameters</u> to <u>4 diameters</u>. It is interesting to think back to the ant on the rim of a wheel. On the upswing, her motion is always headed straight for the top of the wheel, but the length of her cycloidal path to the top will always be twice her distance from the top at any given moment.

I would ask the reader to reflect here on the things which can be known about curves solely from considering the actions which produce them. An equation for the cycloid was not written down until after all of the above discussions. When I think of how, in the past, I have presented this curve in my calculus classes using the standard parametric equations, I feel that both I and my students learned very little. In the secondary curriculum, cycloids are rarely mentioned, because their equations are considered too difficult.

What is governing our choice of curriculum? It would seem to be regulated by algebraic convenience. Students are asked to consider many curves that I have never seen in daily life, simply because their equations are tractable. Analytic methods are powerful tools, but letting the tools govern the subjects of our thoughts can only lead to tedious and unnatural formalism. As Leibniz labored to create the language and notation that we call calculus, he had to test this language to see that it was consistent with what was known about areas, tangents, and arclengths. Curves such as the cycloid were used as critical experiments to test the validity of linguistic constructions. Leibniz first wrote an equation for the cycloid in 1686, and then used it to test his evolving notations (Whitman, 1946).

Leibniz wanted to create a universal language which was capable of expressing all known results about areas, tangents, arclengths, and other quantities. Newton accused Leibniz of plagiarism, because he never came up with any previously unknown answers to questions about areas, volumes, tangents, or arclengths. Newton misunderstood the intention of Leibniz. He was not a plagiarist; he was a linguist. He largely succeeded in his quest for a universal language capable of expressing all of the known results from the geometry of his day.

It is this sense of language as a human construction, evolving from experiences, and fitted to certain purposes, that I want suggest should be brought into the mathematics classroom. By eliminating the discussion of curves like the cycloid, the grounded activity which justifies language construction is taken away from students. They then have no critical experiments upon which to test the consistency and validity of the formalisms they learn. They learn only about the nature of representations that refer to themselves in an endless hall of mirrors. Students often see mathematics as perfect and unquestionable, because they have only experienced it within a self referential frame. Right at the very beginning of the scientific revolution, Pascal objected to this general linguistic trend in modern thought. Speaking theologically he said, "Nature possesses forms of perfection in order to show that it is an image of God; and faults to show that it is only an image" (Pascal, 1962, #262). I will take the liberty to paraphrase him and say that: Mathematics possesses forms of perfection in order to show that it is an image of Nature; and faults to show that it is only an image.

2.14 Pascal and Leibniz: Sines, Circles, and Transmutations

The title of the last section promised some discussion of sine curves. In Roberval's construction a sine curve was drawn and used to find the area of the cycloidal arch (see Figure 2.13b), but Roberval called this curve the "companion of the cycloid." He did not see this curve as a graph of a sine or cosine, and neither did any of his contemporaries. They did discuss sines and cosines, however, and it is important to understand the conceptual point of view that was then standard. Before I present my last example of curve drawing from Leibniz, I must also return to some of the issues raised in section 2.3. Leibniz's use of the characteristic triangle (Figure 2.3b) was directly inspired by the work of Pascal concerning sines.

Trigonometric functions were not defined using ratios, or the unit circle, until the textbooks of Euler were published in 1748 (Euler, 1988). Both Ptolemy and Arabic astronomers made detailed tables of the lengths of chords subtended by circular arcs (Katz, 1993). That is, given two points, A and B, on a circle, to find the length of the line segment $\overline{AB}$. Such tables were usually made for a circle of a given (large) radius, and then scaled for use in other settings. It was also found useful by Arabic astronomers to have tables of half chords, and such tables became known in Latin as tables of "sines." [33] Since the perpendicular bisector of any chord passes through the center of the circle, half chords on the unit circle are the same as our sines, but I want to stress that trigonometric quantities were seen as lengths. Tangents were seen as lengths marked on a tangent line. Secants were the lengths from the center of the circle to the tangent line. Etc. In Figure 2.14a, PB is a sine, CD is a tangent, and AD is a secant, regardless of what parameter is used to index these lengths.

[33] The Arabic word for "half chord" closely resembled the Arabic word for a bay of water. Early Latin translators confused the two and translated the Arabic "half chord" into the Latin word "sinus" meaning bay or cavity.

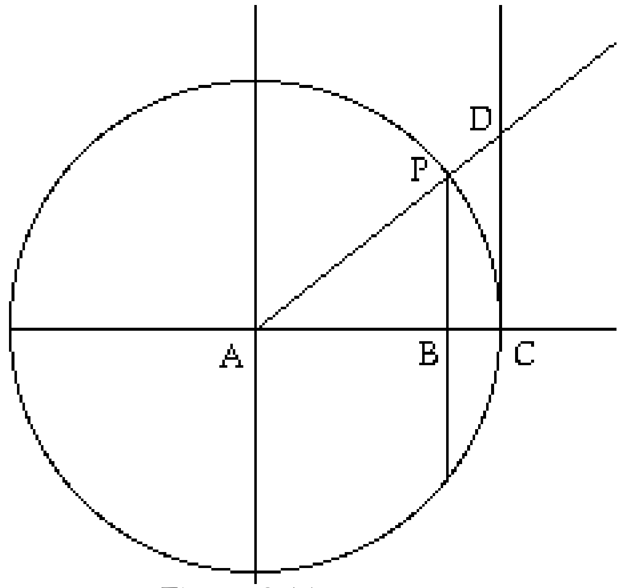


Figure 2.14a

The parameter from which these trigonometric lengths were indexed was implied by usage, but not always stated explicitly. Some form of systematic equal division was used to make any particular table. As analytic geometry evolved, parameters had to be stated more clearly. For example, in Figure 2.13a, is PB a "sine" or an "ordinate" to the circle with respect to the axis $\overleftrightarrow{AC}$? Pascal and Roberval were quite clear on this point. If the line $\overleftrightarrow{AC}$ is divided into equal increments, and then perpendiculars are erected to the circle, then those segments are "ordinates." On the other hand, if the circle is divided into equal pieces of arclength, and then perpendicular segments are dropped to the axis, then those segments are "sines" (Struik, 1969). See Figure 2.14b, where on the left the diameter is divided into sixteen equal segments, and on the right the half circle is divided into sixteen equal arclengths. Pascal did not restrict this terminology to the circle, but used it for any curve .

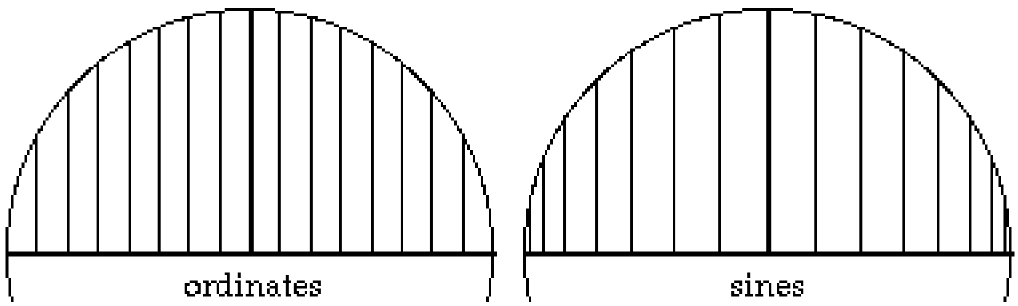


Figure 2.14b

In any given situation, the size of the increment was determined by the nature of the problem at hand, and the level of accuracy that was desired or possible. The smallest increment being used was most often taken as the unit of measurement for a problem. In the early work of Leibniz, $dx = 1$ represents these small increments, and there is no distinction between dy and dy/dx. Leibniz did not distinguish between dx and Δx; he used only dx, and it meant simply the difference of the quantity x. The ratio dy/dx entered into his later work, so that he could discuss rates using units of measurement other than the smallest increment dx (Child, 1920).

In 1659 Pascal published his work "On the sines of a quadrant of a circle" in which he established a series of propositions which are algebraically (but not conceptually) equivalent to integrating all of the integer powers of the sine function (Struik, 1969). I will describe only his first example. Figure 2.14c shows an increment of circular arclength DF, together with a segment $\overline{QR}$ tangent at the midpoint P. $\overline{CP}$ and $\overline{CA}$ are radii of the circle, and $\overline{QK}$, $\overline{PB}$, and $\overline{RL}$ are all perpendicular to $\overline{CA}$. Since $\overline{QR}$ is perpendicular to $\overline{PC}$, triangles ΔERQ and ΔBPC are similar. Hence $PB \cdot QR = ER \cdot CP$.

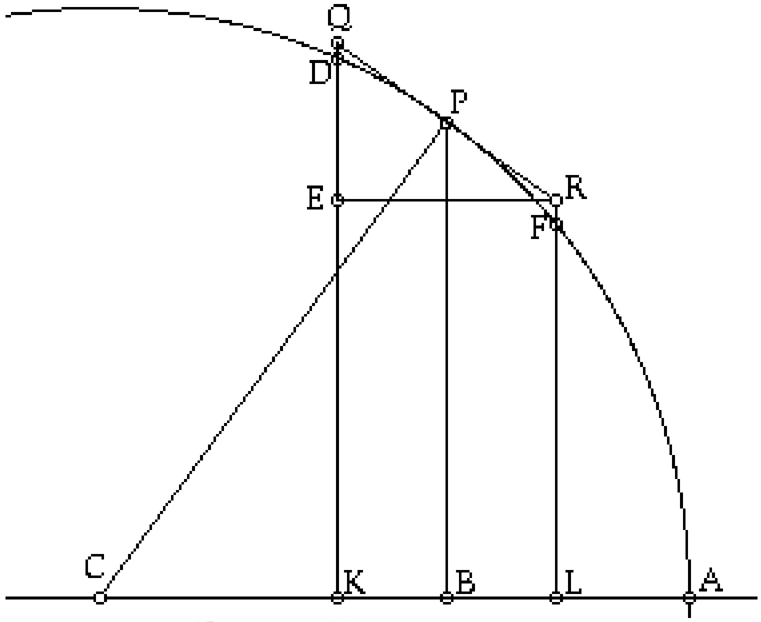


Figure 2.14c

Imagine any arc (≤ 90°) of a circle divided up into equal pieces of arclength such that it is impossible to distinguish those arclength increments from the corresponding tangent line segments. Even in Figure 2.14c they are very close ($QR = 1.40$ in.; arc $DF = 1.44$ in.). Using his arclength increment QR as his unit, Pascal then concluded, from the similarity statement above, that if one sums up the sines of the increments ($PB = PB \cdot QR$), then the sum is always equal to the portion of the horizontal axis between the first and last of the sines, multiplied by the radius of the circle; i.e. the sum of the changes ER in the cosine, times the radius CP (Struik, 1969).

One could transform this statement in the later notation of Leibniz (i.e. modern calculus) by a change of unit. Using the radius of the circle CP as a unit (rather than than the arclength increment QR), and using θ as an arclength parameter (i.e. radians), then $QR = d\theta$, $ER = KL = d(\cos\theta)$ and the similarity statement $PB \cdot QR = ER \cdot CP$ becomes $\sin\theta \cdot d\theta = d(\cos\theta)$. Summing this one obtains: $\int_a^b \sin\theta \cdot d\theta = \cos a - \cos b$. Leibniz created the modern integral symbol as an "s" for summation. Just as he did not separate the concepts of dx and Δx, he did not separate integration ($\int$) from summation ($\sum$). Thinking geometrically, the issue of minus signs does not arise, because ER (the change in cosine) is a geometric line segment. The arc can be measured either clockwise or counterclockwise.

Pascal made a whole series of increasingly complex arguments like the one above. Transformed into Leibnizian notation, they amount to a series of integrations for integer powers of the sine, solved by changes of variable. In the work of Pascal, however, all of these changes of variable were done geometrically, using similarity and projection. Two themes from this work deeply affected Leibniz, and became central in his thinking (Child, 1920). First, the small characteristic triangles along a curve can be analyzed by finding large ones which are similar to them (see Section 2.3). In the case just mentioned, triangle ΔERQ is the characteristic triangle which is similar to the large triangle ΔBPC. Second, there are useful connections between tangents and areas that can be exploited through the finding of such similar triangles.

A few years before his death in 1716, Leibniz wrote "The history of the origins of differential calculus" (Child, 1920). This essay centered on two themes. First, that his notation of differences and summations was developed from his study of tables of numbers, and the patterns that he found there. Second, that this notation from tables could be consistently applied to geometry, and was capable of yielding all known results concerning areas, volumes, tangents, and arclengths. Leibniz described his original insights into the consistency between geometry, and his new algebraic notation, by focusing on what came to be know as the "transmutation of curves" which involved a particular example of a large triangle which is similar to the characteristic one. He used this triangle as a way to draw new curves from existing ones. This method of curve drawing produces, from the original curve, a new "transmuted curve" which bounds areas that are closely related to the areas bounded by the original curve. As we shall see, this transmutation is closely akin to integration by parts. This technique was first investigated by Leibniz early in his career, in 1673, and was described in his letters to Newton (Turnbull, 1960; Child, 1920; Edwards, 1979).

In order to understand the curves and derivations of Leibniz, I have constructed the following three point rectilinear example, as an introduction. This example is not from the writings of Leibniz, but it I think it could help students understand his conceptual approach. Imagine a piece-wise linear "curve" passing through the three points A, B, and C, where A is the origin (see Figure 2.14d). If all three points are collinear then the area under this curve is the area of triangle ΔACG, which is equal to $\frac{1}{2}x'y'$, where x' and y' are the coordinates of C.[34] If B is moved up off the line $\overleftrightarrow{AC}$ then the area under the curve ABC equals $\frac{1}{2}x'y' + \text{Area}(\Delta ABC)$. Leibniz's transmutation technique shows us how to find the area of ΔABC, by looking at a new curve which is drawn by monitoring where the tangents ($\overleftrightarrow{AB}$ and $\overleftrightarrow{BC}$ in this case) to the original curve intersect the y-axis.

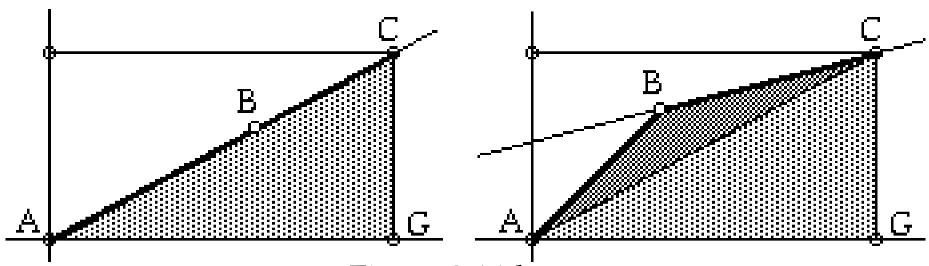


Figure 2.14d

If we let the point U be the intersection of line $\overleftrightarrow{BC}$ (tangent) with the y-axis, then Figure

shows how to construct a right triangle ΔDFG, which is equal in area to triangle ΔABC, by making $FG = AU$, and $DG = BH$. To see this area equality, first construct $\overline{AN}$ perpendicular to $\overline{BC}$. Now triangles ΔBHC and ΔANU are similar. This triangle ΔANU is an example of a triangle which is similar to the characteristic triangle at the point B. Letting $AU = z$ and $AN = p$, from the similarity one sees that $z \cdot dx = p \cdot ds$.

[34] Throughout this section prime notation such as x' and y' will be used to denote fixed endpoint values of variables, and has no relation to derivatives. Any mention of derivatives will use strictly Leibniz notation.

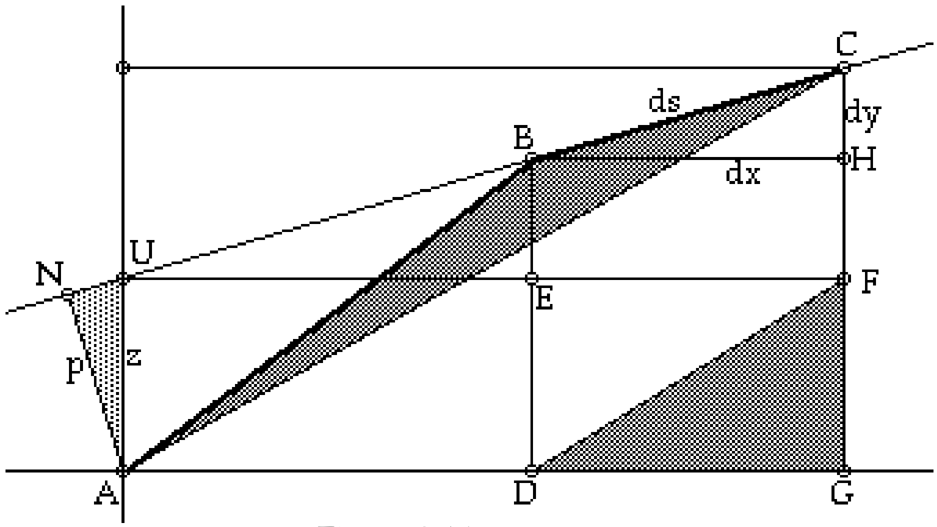


Figure 2.14e

Since triangle ΔABC can be seen as having a base of ds and height of p, its area is $\frac{1}{2}p \cdot ds$, while triangle ΔDFG has a base of dx and a height of z and hence an area of $\frac{1}{2}z \cdot dx$. Hence the similarity tells us that the two darkly shaded triangular areas are equal. This construction can be seen to be helpful in the sense that triangle ΔABC has been "transmuted" into the right triangle ΔDGF, which sits nicely in the coordinate system. If $C = (x', y')$, then the area under the curve ABC equals $\text{Area}(\Delta AGC) + \text{Area}(\Delta ABC) = \frac{1}{2}x'y' + \text{Area}(\Delta DGF)$. Using *Geometer's Sketchpad*, one can drag point B to different positions and watch the changes in triangle ΔDFG, while monitoring the areas with meters.

In order to generalize this construction to non linear curves, Leibniz defined a new "transmuted curve" constructed from any given curve, relative to an axis of ordinates, as follows:

Definition: Given any curve and a system of perpendicular abscissas and ordinates, for each point (x, y) on the given curve define a point (x,z) on the transmutation curve as the one which has the same abscissa x, and has, as its ordinate, the length z, where z is equal to the y-intercept of the tangent line to the original curve at the point (x, y).

In Figure 2.14e, starting with the curve ABC, this new transmuted curve would be the piecewise linear path through $ADEF$ (i.e. a step function in the modern sense). Leibniz then expressed the area under the original curve from A to C as: $\frac{1}{2}x'y' + \frac{1}{2}(\text{area under the transmuted curve})$, where $A = (0,0)$ and $C = (x', y')$. The first term is the area of triangle ΔACG, and second term is equal to the "extra area" under the curve contributed by triangle ΔABC.

This general curve drawing technique can applied to any curve where one can construct the tangent lines at all points and thus monitor their intersections with the axis of ordinates. The area under the new transmuted curve, drawn from the original, will then be used to find the area under the original curve. This geometrical construction yielded, for Leibniz, the area formula above, which is algebraically (but <u>not</u> conceptually) equivalent to the technique now known as "integration by parts." Leibniz, however, developed his transmutation of curves prior to his algebraic notations, such as the product rule. He developed a general algebraic language (i.e. the calculus) only after he had investigated many examples of his geometric transmutation, and had seen the generality of the technique. The extension of his language and notation to geometry grew from his experience with curve generation and transmutation.

I will next apply this curve drawing technique to the circle, because it was Leibniz's favorite example (Child, 1920). In doing so, I will also derive again the transmutation area formula in a more general setting using the mature notation of Leibniz, although this derivation is essentially the same as the one from Figure 2.14e. This general transmutation formula (2.14-1) which I will derive is then valid for any curve with tangent lines.

I start by letting the point P rotate around a circle while dragging its tangent line $\overleftrightarrow{PT}$ with it (see Figure 2.14f). Letting the diameter $\overline{AB}$ be the axis of abscissas measured from A (i.e. $A=(0,0)$), at each position of P, let $\overleftrightarrow{AN}$ be a perpendicular from A to the tangent line. The triangle ΔAUP (shaded) is then similar to the small characteristic triangle along the curve at P. Construct a new curve by tracing the locus of the point R, where R and P always have the same abscissa AS, and the new ordinate RS is always equal to AU (the y-intercept of the tangent line at P). This new curve is drawn from the circle in Figure 2.14f, using an animation in *Geometer's Sketchpad*. See Figures 2.14i and 2.14j to see this same technique applied to the cycloid and the hyperbola .

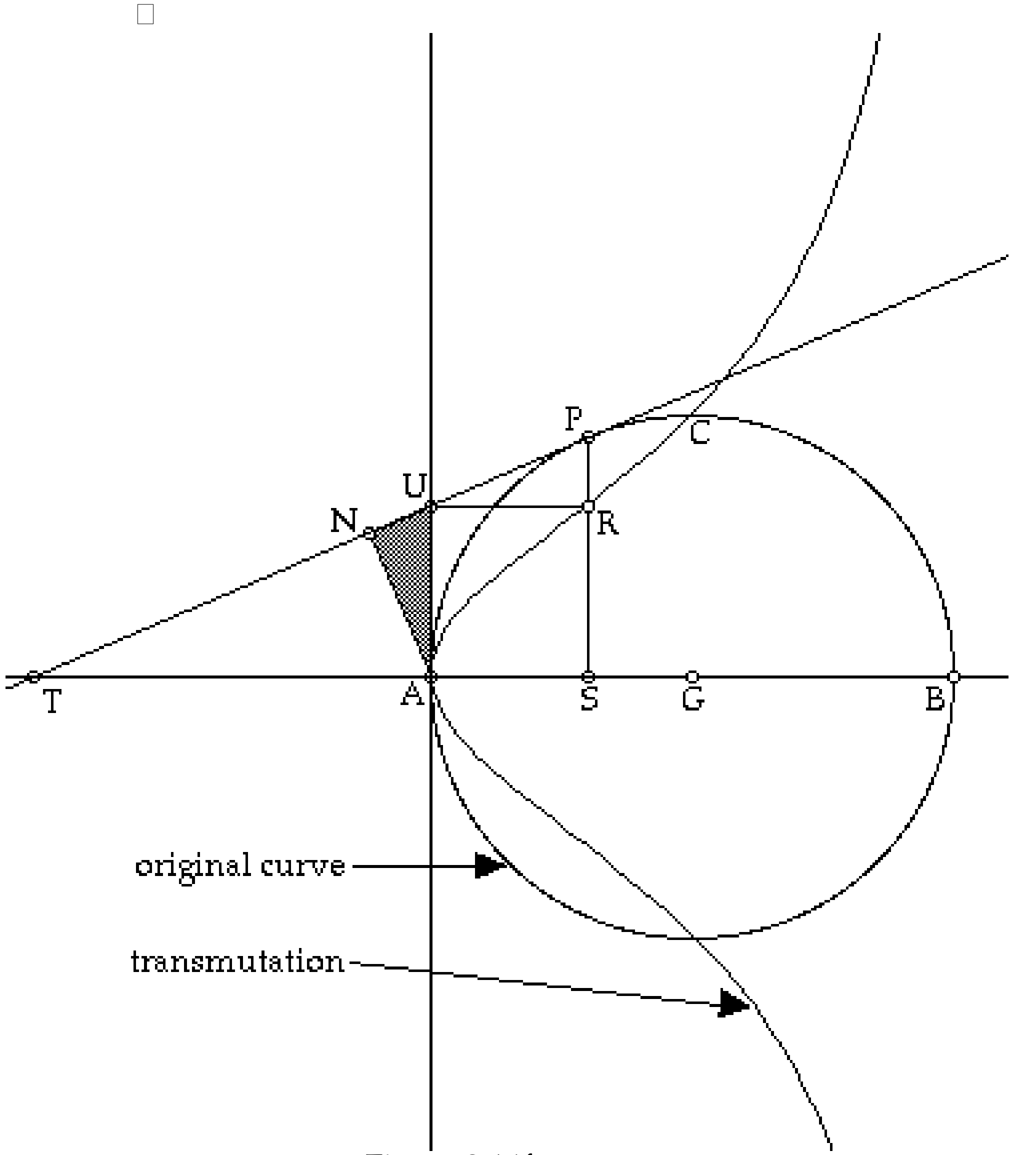


Figure 2.14f

Using the mature notation of Leibniz, his general transmutation area formula can be derived as follows. As before, let $x = AS$ and $y = SP$ be the abscissas and ordinates of the original curve, and let $z = AU$ be the ordinates of the transmuted curve (i.e. $P = (x,y)$ and $R = (x,z)$). Since z is the intercept of the tangent line to the original curve, $z = y - x \cdot \frac{dy}{dx}$. Now let $p = AN$. Since triangle ΔANU is similar to the characteristic triangle at the point P (sides dx, dy, and ds) we have: $\frac{dx}{ds} = \frac{p}{z}$, where s is arclength along the original curve. Hence $p \cdot ds = z \cdot dx$.

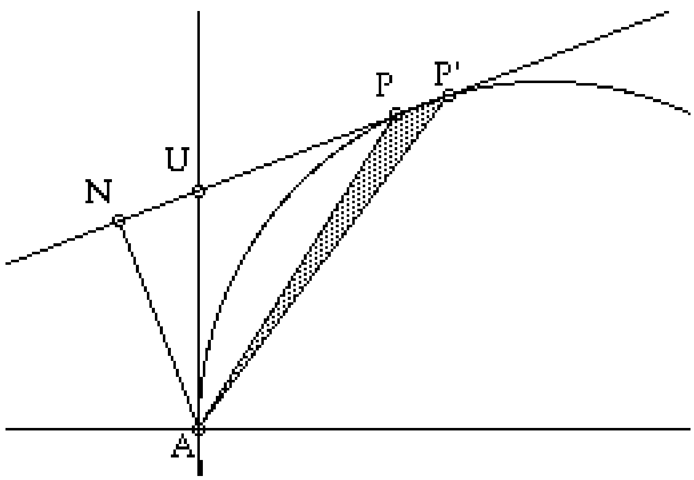


Figure 2.14g

Now imagine two points, P and P', on the original curve that are so close together that the line $\overleftrightarrow{PP'}$ is essentially the tangent line at P (see Figure 2.14g). Now imagine the slender triangle $\Delta APP'$ (shaded). Thinking of $PP' = ds$ as the base this triangle, it has a height of $AN = p$, and so its area is
$\frac{1}{2} \cdot p \cdot ds = \frac{1}{2} \cdot z \cdot dx$ (from the above similarity argument).

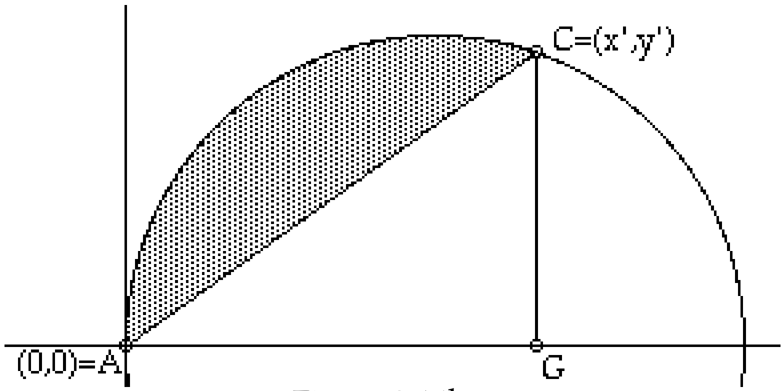


Figure 2.14h

Now choose any fixed point $C=(x',y')$ on the original curve (see Figure 2.14h). The area of the shaded sector AC is the sum of the slender triangular sectors in Figure 2.14g, which equals $\frac{1}{2}\int_0^{x'} p\cdot ds=\frac{1}{2}\int_0^{x'} z\cdot dx$. The area under the curve between A and C is equal to the area of triangle ΔAGC plus the area of the shaded sector, which is to say :

(2.14-1) $$\int_0^{x'} y\cdot dx=\frac{1}{2}(x'\cdot y')+\frac{1}{2}\int_0^{x'} z\cdot dx$$

This is Leibniz's transmutation formula. As I showed in the three point example, it says that the area under any curve is the triangle ΔAGC plus half the area under the transmutation curve. If one substitutes for z from the original definition $z=y-x\cdot\frac{dy}{dx}$, and then solves for the integral of y, then the statement becomes the usual integration by parts formula written for definite integrals: $\int_0^{x'} y\cdot dx=[xy]_{(0,0)}^{(x',y')}-\int_0^{y'} x\cdot dy$.

I make this last statement only for modern readers to see the connection with the usual algebraic approach to "integration by parts" which is taught as an application of the product rule written backwards. The formula which Leibniz applied with great success, in a large number of examples, is the geometric statement (2.14-1) which expresses the area under the original curve in terms of the triangle ΔAGC plus half the area under the new transmutation curve. This transmutation curve can be drawn by a simple linkage, provided that the method which drew the original curve included a construction of tangent lines. This chapter has already shown that that can be done for a great variety of curves. This technique of Leibniz is a response to experience with the tools of the day. The tools mediated the knowledge.

Leibniz first stated his formula in a purely geometric language. For him, it was a major breakthrough in seeing how tangents could be used to determine areas. He developed this method just before he constructed his first notations for calculus, which would later include such statements as the "product rule," the "quotient rule," and the "chain rule." Integration by parts is usually introduced to students as a purely algebraic manipulation of the product rule, but for Leibniz, such a derivation was an algebraic confirmation of the geometric experiments that he carried out using his transmutation technique.

This transmutation of curves can be used in a variety of ways to conduct the kind of critical experiment that tests the validity of a new language against independently established results. Here is how Leibniz applied the transmutation formula to obtain an expression for the area under any part of circle, and then tested his expression by using it to compute the area of a quarter of a circle of radius one. In Figure 2.14f, let G be the center of the circle of radius one: $AG = 1$, and $A = (0,0)$. The equation of the upper half of the circle is $y = \sqrt{2x - x^2}$, and since the tangent is always perpendicular to the radius, $\frac{dy}{dx} = \frac{1-x}{y}$. Hence the equation of the newly drawn, bell shaped, transmutation curve is:

$$z = y - x\frac{1-x}{y} = \sqrt{\frac{x}{2-x}} \quad \text{or} \quad x = \frac{2z^2}{1+z^2}$$

By synthetic division, $x = 2\{ z^2 - z^4 + z^6 - z^8 + \cdots\}$

Now the area under the circle between $A = (0,0)$ and any fixed point (x', y') is given by (2.14-1) as: $\int_0^{x'} y \cdot dx = \frac{1}{2}(x' \cdot y') + \frac{1}{2}\int_0^{x'} z \cdot dx$

Since z, in terms of x, also involves a square root, Leibniz rewrote the area under the transmutation curve by subtracting its complement from the rectangle containing it. Hence, as an integral in x subtracted from the circumscribed rectangle with area $x' \cdot z'$ (where z' is the value of z at $x = x'$), Leibniz obtained:

$$\int_0^{x'} z \cdot dx = (x' \cdot z') - \int_0^{z'} x \cdot dz$$

This new integral can now be expressed using the synthetic division above and then integrated term by term.[35] Hence the area under the circle is:

$$\begin{aligned}\int_0^{x'} y \cdot dx &= \frac{1}{2}(x' \cdot y') + \frac{1}{2}(x' \cdot z') - \frac{1}{2}\int_0^{z'} 2(z^2 - z^4 + z^6 - z^8 + \ldots)dz \\ &= \frac{1}{2}(x' \cdot y') + \frac{1}{2}(x' \cdot z') - \left[\frac{1}{3}z^3 - \frac{1}{5}z^5 + \frac{1}{7}z^7 - \frac{1}{9}z^9 + \cdots\right]_0^{z'} \\ &= \frac{1}{2}(x' \cdot y') + \frac{1}{2}(x' \cdot z') - \left(\frac{1}{3}z'^3 - \frac{1}{5}z'^5 + \frac{1}{7}z'^7 - \frac{1}{9}z'^9 + \cdots\right)\end{aligned}$$

In order to test the validity of this expression against known results in geometry, Leibniz checked it on the area of a quarter circle where $x' = y' = z' = 1$. The expression above then asserts that:

$$\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \frac{1}{9} - \frac{1}{11} + \cdots$$

Checking this series empirically one finds that it does indeed converge to $\frac{\pi}{4}$, although very slowly.[36] This alternating odd harmonic series provided a new way to compute π,

---

[35] The calculation of areas and volumes that involve only the summation (integration) of polynomials had been known since the tenth century from the work of the Arabic mathematician, al-Haitham. In the early seventeenth century the Arabic work on polynomial summations had been extensively elaborated by Cavalieri, Fermat, Pascal and Wallis (Struik, 1969; Dennis & Confrey 1993).

[36] Many terms of this series are needed to obtain even a minimal level of accuracy. As a practical calculation of π, this method is abysmal. More effective methods for evaluating π had been known since ancient times (Archimedes, 1952), such as the doubling on the number of sides in regular polygons, which gives π as a nested series of square roots. This was not a new practical calculation; it was a confirming experiment for the validity of a new language.

A modern calculus book would evaluate the area under a section of the circle by making a trigonometric substitution into an integral. Such an approach looks simple, but is only useful if one already has a means to evaluate trig functions. Such means usually involve an infinite series.

and it was the first known expression for π as a sum of rational numbers. Far more important for Leibniz was the fact that this calculation was a confirming experiment that showed that his new language was indeed a tool which was consistent with the long established traditions of geometry. This showed him that the notation that he developed from tables could be consistently applied to geometry.

The confirmation of a new mathematical approach using as a check a new computation of π was a favorite procedure during the seventeenth century. In 1655, John Wallis had used exactly this method to confirm his assertions about the consistency of algebra and geometry in his *Arithmetica Infinitorum* (1972). He had justified his extensive use of table interpolations by showing that they implied a value for π, given as an infinite product, which was consistent with other computations from geometric constructions, such as the iterated square root procedure of Archimedes (Dennis & Confrey, 1993). This work of Wallis had a profound effect on the work of both Leibniz and Newton (Newton,1967; Child, 1920).

The use of synthetic division to create an infinite series was a technique pioneered by Nicolaus Mercator (1620 -1687), and used extensively by Leibniz, Newton and other mathematicians of the period. Leibniz saw his transmutation technique as a general method for finding rational expressions that could replace those integrands which contained root extractions, thereby creating series expressions for those areas using, first, synthetic division, followed by term by term polynomial integration (i.e. what Leibniz called "Mercator's method") (Turnbull, 1960; Boyer, 1968). In the example above, for the circle, it should be noted that the transmutation curve for the circle is a cubic curve, since it contains an xz^2 term. In general the transmutation curves for conic

---

The series usually used for computing trig functions were first derived from series which expressed areas, like the one discussed here. (See Newton, 1967; Dennis & Confrey, 1993; Edwards, 1979).

sections are of third or fourth degree, or what Descartes called "curves of the second class."

Leibniz drew remarkably few figures in his descriptions of his work. He tended to prefer elegant tabular displays of the coefficients that occurred in his series expansions (see for example his August, 1676 letter to Newton in Turnbull, 1960, pp. 65 -71). These tabular displays reveal some remarkable connections between trigonometric series like the one above and logarithmic series which come from hyperbolic areas via transmutation (see figure 2.14j). Despite the paucity of figures in the original discussions of Leibniz, *Geometer's Sketchpad* allows one to create a remarkable set of curves based on the transmutation construction. It is these sets of curves with their corresponding equivalent areas which I feel could provide fertile ground for student investigations.

A simpler and interesting special case comes from transmuting a parabola using the line tangent to the vertex as the axis of ordinates (this example does not appear in Leibniz). I invite the reader to make his/her own figure (see Section 2.4). One could then choose a coordinate system so that the parabola has the equation: $y = \sqrt{x}$. From the geometric properties of parabolic tangents discussed in Section 2.4 (i.e. the subtangent is always twice the abscissa), it can be seen that the transmutation curve of this parabola is another parabola whose ordinates are half those of the original curve, i.e. $z = \frac{\sqrt{x}}{2}$. If one considers the area under the curve from (0,0) to (x', y') transmutation tells us that: $\int_0^{x'} \sqrt{x}\,dx = \frac{x'y'}{2} + \frac{1}{2}\int_0^{x'} \frac{\sqrt{x}}{2}\,dx$.

Since the same integral appears on both sides one can solve for it to obtain: $\int_0^{x'} \sqrt{x}\,dx = \frac{2}{3}x'y' = \frac{2}{3}y'^3$. By looking at the complement of the rectangle with sides x' and y', this is equivalent to the usual integration of the parabola written as $\int_0^{y'} y^2 dy = \frac{1}{3}y'^3$. Note that this does not involve any use of "anti-derivatives," but is instead a linguistic coding of purely geometric properties that built the parabola.

I wish to return to the cycloid and discuss another of Leibniz's application of the transmutation technique (Edwards, 1979). Roberval's argument (Section 2.13) showed that the area of an entire cycloidal arch was three times the area of the circle that generated it, but this argument depended on the symmetry of the companion (sine) curve, and thus did not yield the values of arbitrary sections of cycloidal area. Using the constructed tangents one can draw the transmutation curve, and arrive at this general result. Figure 2.14i shows one half of a cycloidal arch drawn sideways (i.e. the wheel is rolling along the vertical line $\overleftrightarrow{FD}$). It turns out that the transmutation curve traced by the point R is exactly the same as Roberval's companion curve, i.e. a sine-shaped curve. This follows from the fact that the tangent at P is parallel to $\overline{AB}$, and so $PB = AU$, and hence by subtraction $PR = SB$.

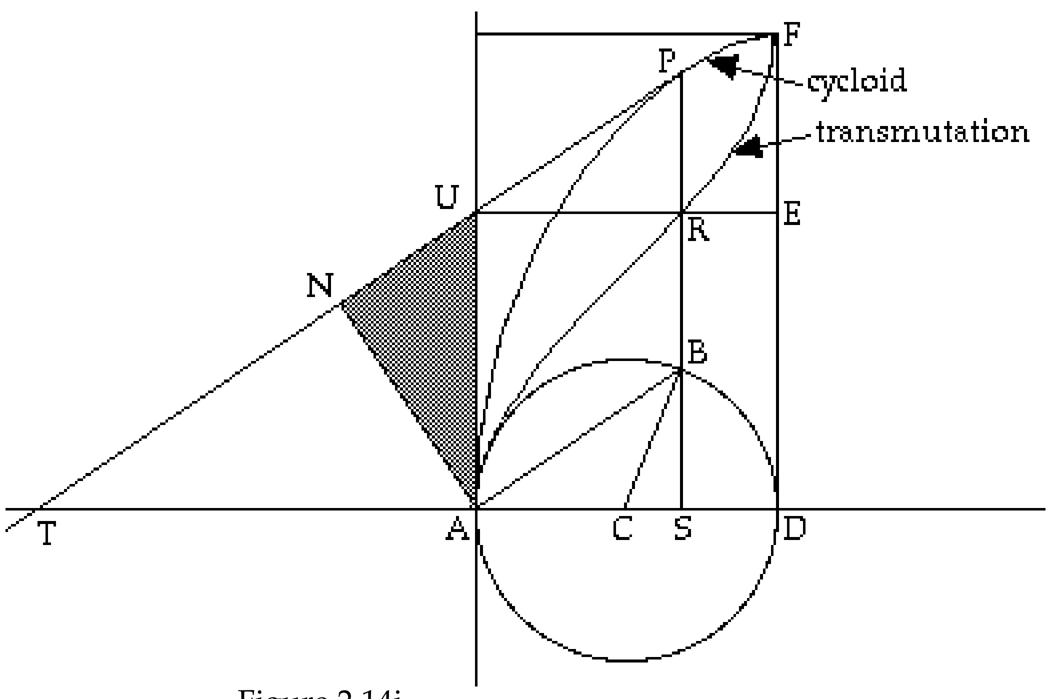


Figure 2.14i

Leibniz determined the area under the cycloid over any portion of the axis $\overleftrightarrow{AD}$ as follows. Introducing variables as before; let the abscissas x be measured from A along $\overleftrightarrow{AD}$, and let y, z, and w be the ordinates, respectively, of the three curves shown in the figure, those being the cycloid, sine, and circle. That is to say let: $P=(x,y)$, $R=(x,z)$, and $B=(x,w)$. Since $w=SB=PR$, then $y=z+w$. From the motion that produced the cycloid one knows that the circular arclength from B to D will equal FE, while the circular arclength from A to B will equal $ED=AU=PB=RS=z$. This says that the locus of R can be seen as a graph of arclength against height on the circle, which makes the curve sine shaped.

Let the constant radius of the generating circle be $r=AC$. Suppose one seeks the area under the cycloid between 0 and some x'. Let y', w', and z' be the corresponding endpoint values. The general transmutation formula

(2.14-1) says that:

$$\int_0^{x'} y\cdot dx=\frac{1}{2}\left(x'\cdot y'+\int_0^{x'} z\cdot dx\right)$$

$$\int_0^{x'} (z+w)\cdot dx=\frac{1}{2}\left(x'\cdot y'+\int_0^{x'} z\cdot dx\right)$$

Solving this for $\int_0^{x'} z\cdot dx$ yields: $\int_0^{x'} z\cdot dx = x'\cdot y'-2\int_0^{x'} w\cdot dx$

Since w represents the ordinates along the circle, $\int_0^{x'} w\cdot dx$ is the area under the circle from 0 to x', which can be expressed as the area of the circular sector ABC plus (or minus) the area of the triangle ΔCBS (see figure 2.14i). In my figure $x'>r$, but if $x'<r$ then one subtracts the triangle from the sector, in which case $x'-r$ will be negative. Using $z'=(\text{arclength from } A \text{ to } B)$ to express the area of the sector one obtains:

$$\int_0^{x'} w\cdot dx=\frac{1}{2}rz'+\frac{1}{2}w'(x'-r).$$

Now

$$\int_0^{x'} z\cdot dx=x'y'-2\left(\frac{1}{2}r\,z'+\frac{1}{2}w'(x'-r)\right)=x'(z'+w')-rz'-w'(x'-r)$$

$$=rw'-z'(r-x')$$

Substituting this back into the original transmutation formula yields:

(2.14-2) $$\int_0^{x'} y\cdot dx=\frac{1}{2}x'y'+\frac{1}{2}(rw'-z'(r-x'))$$

which gives the area under a cycloidal section in terms of the radius of the generating circle r, the endpoint $P=(x',y')$, the arclength of rotation z' and the ordinate to the circle w'. These are the constants which are natural to the action of drawing the curve. If one attempts to use the transmutation formula in a strictly algebraic sense then $\int_0^{x'} z \cdot dx$ calls for an integral of the arcsine function. This is not the conceptual approach taken here. In order to apply formula (2.14-2) one does need to know the arclength z' but the action which drew the cycloid produced line segments which were equal to this arclength. Although this approach is quite analytic one should note that the parametric equations of the cycloid were never written down or directly used. Only later did Leibniz write down the parametric equations and then use them to further test his calculus notations.

If one applies Formula 2.14-2 to one half of the cycloidal arch then $x' = 2r,\ y' = \pi r,\ w' = 0,\ z' = \pi r$. Using these values the formula yields $\frac{3}{2}\pi r^2$. This is in accordance with Galileo's experiments, and Roberval's geometry. Pascal had earlier given a general geometric solution for the area of a cycloidal section against which Leibniz's formula (2.14-2) could also be checked. It is this checking back and forth between geometry and the new language of calculus that gave people faith in the new linguistic constructions of Leibniz.

Leibniz applied his transmutation technique to the hyperbola and obtained an infinite series for the area between a hyperbola and its axis of symmetry, in much the same way that he found the circular areas. This series intrigued him because of its relations to other series, found by Newton and Mercator, for the calculation of areas between hyperbolas and their asymptotes (called by Mercator "natural logarithms") (Turnbull, 1960; Edwards, 1979). I will not describe all of these calculations, but I will draw the transmutation curve because most historical accounts provide either no figures or highly distorted ones. Using the envelope construction from Section 2.6 to draw the hyperbola with its tangents, one can trace the locus of the transmutation curve using one vertex A as the origin. Figure 2.14j shows all three branches of this curve for the hyperbola with vertices A and B. This curve has a cubic equation that is very similar in form (a single sign change) to the one for the circle, but the appearance of the curve is quite different. In this case, the equation of the hyperbola is $y = \sqrt{2x + x^2}$, and the equation of the transmutation curve is $x = \frac{2z^2}{1 - z^2}$. One then proceeds to construct an infinite series, as in the circular case, via Mercator's method and the transmutation area formula (Turnbull, 1960; Edwards, 1979).

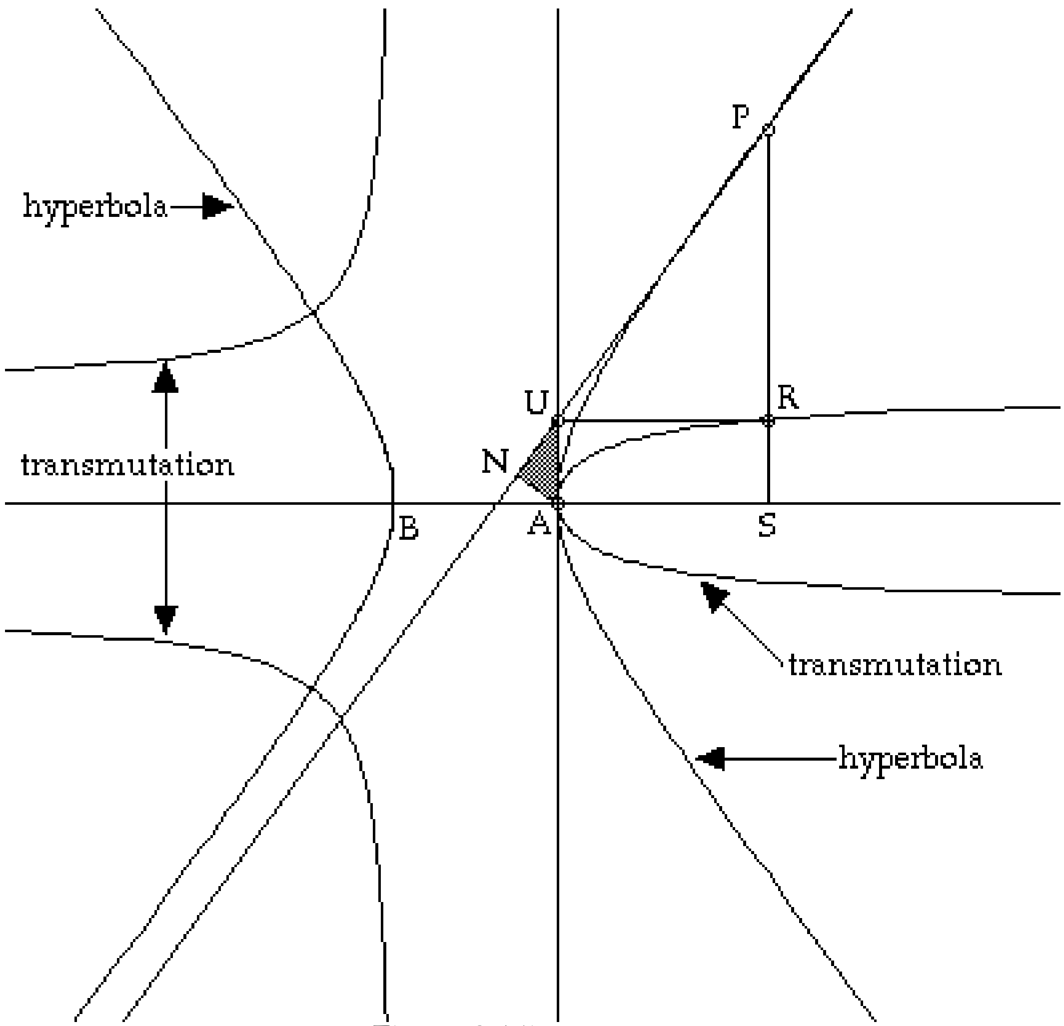


Figure 2.14j

What I find educationally valuable is the experience of accurately drawing these curves without any reference to equations or even the establishment of a scale. Using the previous conic drawing methods, one can draw a large variety of transmutation curves, depending on where one chooses to place the line on which the tangent intercepts z are to be observed. The examples chosen by Leibniz had a specific algebraic purpose which was to rewrite areas, in a rational form, so that synthetic division and term by term integration would yield an infinite series, but I find these curves interesting in their own right. They provide an iterative way to generate curves of increasing degree whose areas have a special relation to the original curve form which they were drawn.

Choosing different lines on which to intersect the tangents gives a fascinating geometric view of all possible different ways to integrate by parts. Not all of these choices will yield an algebraic simplification. For example if one constructs a transmutation curve for the hyperbola $y = 1/x$ by using the y-axis (i.e. an asymptote) for the construction then the transmutation curve is another hyperbola with the same asymptotes which is just a multiple of the original curve. This is equivalent to an application of integration by parts which is algebraically circular.

Experimenting along these lines can produce beautiful and fascinating new curves. Figure 2.14k shows a another hyperbolic transmutation curve drawn with respect to a line at a skewed angle to the axis of symmetry of the hyperbola. This time the transmutation curve has two branches instead of three. Leibniz's area relation still holds. That is, the difference between the area under the hyperbola (traced by P) and the triangle ΔAPG is always one half of the area under the curve traced by R (from A to any given x). Looking at this figure and the previous one, a variety of graphic observations could emerge for discussion. For example, what does it mean when the original curve and the transmutation curve cross each other? When the original curve becomes quite straight near its asymptotes, then the transmutation curve becomes quite flat and nearly constant. What does this say about related area accumulation? How does one interpret the places where the transmutation curve crosses the x-axis in terms of area accumulation on the original curve?

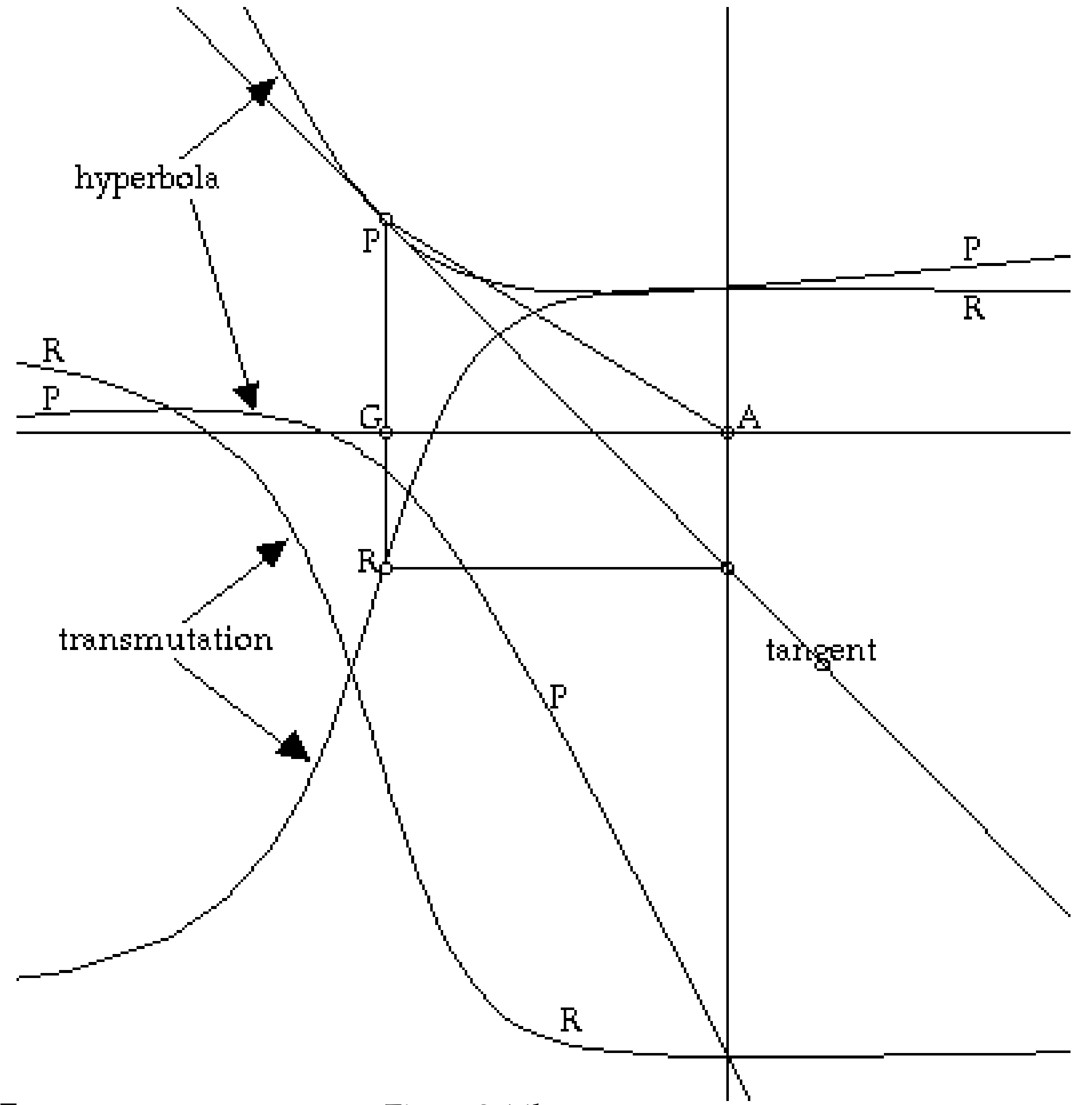


Figure2.14k

2.15 Summary Remarks

All of the figures in this chapter were accurately drawn with all ratios of lengths and areas shown in proper proportion, and yet never once was an equation plotted or entered into a computer as a means of generation. Geometrically defined motions, all of which can be mechanically produced, created the curves. The algebra and tables, and even the numerical scales, presented were all created after the fact as a form of syntax through which one conducts a systematic inquiry into the geometric actions that created the curves. If I had presented only one or two good examples of investigation based on curve drawing (e.g. the parabola in Section 2.4) a possible response might have been to think that they were beautiful special cases, but not a fully general, pervasive and independent mathematical approach. I hope that I have presented enough material to eliminate that response.

This chapter began with the notion of diameters and axes and their corresponding ordinate directions which were the main tool of Apollonius for the investigation of conic sections. This notion has many functional aspects in that it leads towards the view that a curve is a sequence of ordinate segments which are ordered by an axis, and that specific ratio statements can be made about such orderings. Combining this idea with the mechanical devices of the seventeenth century and with the emerging algebraic symbolism, led to the development of analytic geometry, not as a formal language, but as a generalized tool for the investigation of curves and motion. The examples presented give a sense of how successfully these generalized tools were used to investigate a wide variety of curves and motions.

I have tried to sketch this development right up to the brink of calculus in order to show how language is tailored to fit experience. Although education is not a reliving of history, one aim of this thesis is to advocate for how important it is for students to see mathematical language as a tool forged within a set of experiences. Students must first be involved in some set of experiential activities before this understanding can take place. Sometimes examples taken directly from history can best provide such experiences (e.g. Sections 2.4, 2.7 or 2.13), and at other times historical examples can be modified for educational purposes (e.g. Section 2.12).

Much more could have been said about the relation of curve drawing to other mathematical representations, for example, the relations of the transmutation technique to the development of tables and summations (Child, 1920), but my intention here is not to provide fully developed curriculum materials. This chapter contains enough material to make the argument that valuable conceptual diversity has been lost from mathematics by the representational constraints of our current curriculum. The existence of software like *Geometer's Sketchpad* makes it impossible to justify these constraints by simply saying that geometric curve drawing is too complicated and unwieldy for classroom use. The combination of available software and physical tools with the diverse historical representations of curves indicate specific directions for curricular reform.

More geometry is needed in our curriculum, but if that geometry is to provide the experiential base which leads to an understanding of calculus then it must involve motion and curves. The traditional static geometry with two column proofs, based loosely on Euclid's Books 1-4, fails to provide the appropriate experiences. As early as possible geometry must be connected to mechanics, motion, and curves. This chapter has indicated how such a curriculum could begin to be constructed. When curves are investigated dynamically, their tangents, areas, and arclengths can all be investigated, empirically at first, and then with increasing analytic precision. Reflect, for example, on the area of a cycloidal arch first measured empirically by Galileo, and then determined by Roberval using mechanical geometry (Section 2.13), and finally analyzed by Leibniz using an algebraic expression that emerged from a more general view of the mechanics of curve generation (Section 2.14). A precise analysis like Roberval's need not be algebraic, and could be conducted by students before they study analytic geometry.

The material in this chapter establishes the first and third claims of this thesis as stated at the end of Section 1.1 (p. 9). The first claim ,that curve drawing devices played a fundamental historic and conceptual role in the development of analytic geometry and calculus, could be established by simply looking at Descartes' *Geometry*, and then realizing the impact that that book had on seventeenth century mathematical thought. I have gone beyond that limited argument in an attempt to show how a larger technological experience with tools was shaping the mathematical thought of at least two generations of European mathematicians. I have also made some attempt to show how this development was rooted in the work of Apollonius and Arabic mathematicians.

My third claim , concerning discussion of tangents, areas, and arclengths, was repeatedly established for a wide variety of curves. In particular, the tangent properties of conic sections could be easily investigated informally by students with little or no algebraic background, by using the devices described in the earlier sections of this chapter. Similar investigations involving trigonometric and logarithmic curves could occur well in advance of calculus, as part of a reformed analytic geometry curriculum. Historical discussions of the social and technological history of the scientific revolution would connect such mathematical investigations directly with larger cultural issues, but most importantly these investigations would provide students with more appropriate, dynamic, geometric experience. Such experience would lend far greater meaning to the syntactical structures of calculus.

The role of functions as conceptual tools for the analysis of curve drawing actions reverses the usual epistemic role that they play in current curriculum where functions are used to create curves. This reversed role could allow for a meaningful introduction of functions in an empirical geometric setting quite early in the curriculum. Functional notions in this sense could be introduced before analytic geometry, and could provide an important way to build early connections between geometry and algebra. How early this material could be introduced would have to be the subject of further study, but I suspect that the middle schools years would be the appropriate place to begin such research.

An issue raised by the historical material in this chapter concerns the use of parameters and parametric functions. This issue comes up again in the student interviews of Chapter 3. In many cases a given geometrical method for generating a curve implies that the points on the curve are being defined from a particular parameter space. For example, at the end of Section 2.5 points on the ellipse are defined from the set of points on the hodographic circle. The circle forms an explicit parameter space in a very tactile way, since in Geometer's Sketchpad one drags a point around this circle in order to generate the ellipse. Newton's method for drawing curves described in Section 2.9 always creates new curves from a given "directrix curve." In the case of conic sections, this directrix is always a line and there is a one to one correspondence between points on the "directrix" and points on the curves. Again one finds here a very specific parameter from which the curve is being defined. This suggests that perhaps parametric equations as a form of analytic representation of curves could be introduced earlier in the curriculum and that this form of equation writing might be very natural in the setting of mechanically generated curves. This thesis will not make any specific educational claims regarding this issue, but since it arises both here and in the student interviews, it would seem to be a fruitful direction for further research.

Another more general philosophical question that arises here: what is a representation of what? Are curves a visual representation of functions? or are functions syntactical tools for the representation of curves? The first view is the one usually taken by modern curriculum, although the concept of a curve is more general than the notion of a "graph of a function." The second view is implied by Leibniz's original concept of function as described in Section 2.3, and by the parametric issue just discussed. I propose that this question can provide enlightening reflection, but is best left unanswered. Both views provide important insights. Neither concept can be completely contained within the other, just as the equator of the earth divides its surface into two hemispheres neither of which can be said to definitively contain the other, although each is surrounded by the other. Any attempt at hierarchical structure will diminish the creative diversity of mathematics. The inside is the outside is the inside is the outside . . . . . . in a perpetual feedback loop.

Chapter Three: Interviews with Students Exploring Elliptic Devices

3.1 The Content and Purpose of the Interviews

This chapter will present the results of interviews with two high school students as they explored three different devices for drawing ellipses. My intention here was to see how the physical devices themselves mediate knowledge. In particular how such tools can create a situation where curves have a primary existence and algebraic quantification, and symbolism play a secondary, facilitating role. The students were given no prior instruction in the historical, cultural or mathematical significance of the devices with which they worked. With only the most basic hands on instruction as to how to operate the curve drawing devices they were then asked questions about what kinds of curves each device could draw, the possible situations where different devices might or might not draw the same curve, and how the action of each device might give rise to an algebraic representation. They were asked to justify their answers in any way that seemed appropriate to them.

The three devices chosen have been described and analyzed in Sections 2.5 and 2.9. They are the loop of string over two tacks (Figure 2.5a), the right angled trammel device (Figures 2.9a, 2.9d), and the folding arm device (Figure 2.9g). All three devices were carefully built to be accurate and easily adjustable. They all worked on the fairly large scale of a three foot by four foot drawing area.

The string device involved a 3ft. x 4ft., paper covered sheet of soft plywood into which tacks could be inserted. An adjustable loop of string could then be placed over the tacks and then drawn taut with a colored pencil. The string was tough braided nylon which would not stretch and the length of the loop could be quickly and easily adjusted by a spring locked slide such as those found on the drawstrings of coats.

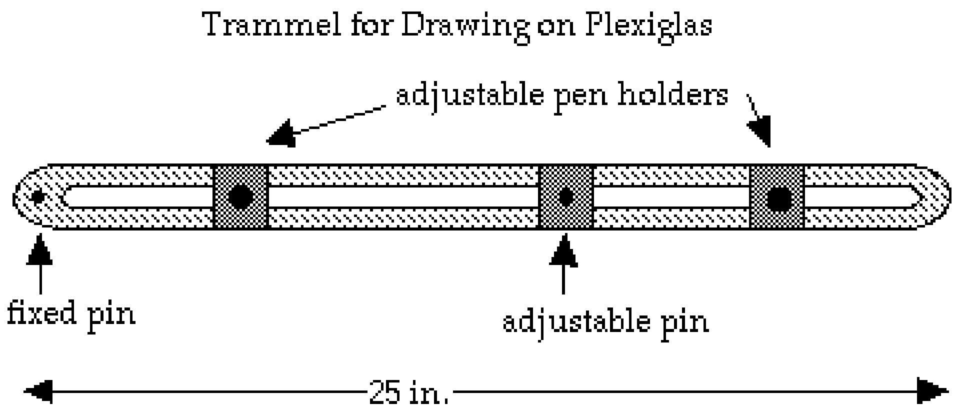


Figure 3.1a

The trammel device worked on a 3ft. x 4ft. sheet of Plexiglas which had two narrow grooves (1/8 in.) carved into it at right angles to each other and bisecting both dimensions of the sheet. The trammel itself (see Figure 3.1a) was made from a 25in., slotted wooden stick with a fixed pin at one end which could slide in one of the grooves in the Plexiglas. An identical pin which could slide in the other groove protruded from a small aluminum holder which fitted into the slotted stick and could be locked with a thumbscrew at any chosen distance from the fixed pin (range =2in. - 25in.) thus creating a trammel of adjustable length. The path of motion of any point on the trammel could be traced on the Plexiglas by a pen fitted in an adjustable penholder. In the slot of the trammel stick were fitted aluminum pieces which were drilled as a penholders for dry-erase, colored marking pens. These penholders could be locked at any position on the trammel with a thumb screw. The pen fit tightly into the penholder and did not need to be held by hand during the drawing motion of the trammel. One penholder was placed between the two pins to form a trammel device as in Figure 2.9a, and the other was beyond the second pin for drawing curves as in Figure 2.9d.

The folding arm device used the same sheet of Plexiglas, pens and penholders as the trammel device, but this time two 15in. slotted sticks, similar to the trammel, were hinged together at their ends (i.e. where the fixed pin is marked in Figure 3.1a). One stick (arm) was fitted with an adjustable long pin which rotated in a hole drilled in the center of the Plexiglas. The distance of this pin from the hinge could be adjusted and locked with a thumbscrew (2 in. - 15 in). The other stick (arm) had a short pin which could slide in one of the grooves and whose distance from the hinge could also be adjusted and locked with a thumbscrew. As with the trammel adjustable penholders were placed on the this second arm on either side of the short sliding pin, hence both of the drawing configurations shown in Figure 2.9g were possible. As was shown in Section 2.9, this device will draw ellipses as long as the rotation pin and the sliding pin are equidistant from the hinge between the two arms. If these distances are not equal then this device will draw curves as in Figure 2.9h. While the loop of string and trammel devices could draw an entire ellipse in one smooth motion, this folding arm device could draw only half of an ellipse before the arms collided. To complete the entire ellipse the second arm had to be lifted or flipped over to the other side of the rotating pin.

I chose to build these three particular devices for use in student interviews for the following reasons. Ellipses are common in visual experience, science and art, aesthetically pleasing, and although the curve is not the graph of a function in a narrow sense, its equation is fairly simple and familiar to the students. These three devices are all fairly simple to build, demonstrate and experiment with. The actions involved in all three can be felt directly and intuitively, and although they are each capable of drawing the entire family of ellipses, they all feel quite different and the adjustments for changing the elliptic parameters work in different ways and hence any algebraic relations that emerge directly from the actions will, at first, have different forms. There is an immediate physical element of surprise when three such different actions all produce curves which look the same. As the interviews will show, this first impression provided a strong motivation for the students to search for a reconciliation of their physical experience with mathematical representations.

The high school students had heard about the loop of string device, and had seen their teacher derive an elliptic equation from the constant sum of the two focal distances implied by action of the device. For the purpose of interviews, this device provided an initial sense of familiarity for the students, although they had not personally experimented with such a device with their own hands, and hence had little practical instinct for exactly how variations in the length of the loop and the distance between the tacks would affect the curve. They had seen various videos and computer animations of conic sections, but their experience had been passive and visual, not active and tactile.

The students had never seen a trammel device used to draw ellipses, although the action involved in the device is perhaps the most simple of the three (e.g. a ladder sliding down the wall). From a practical standpoint (e.g. in carpentry) it is the most useful since it draws an ellipse directly from its most apparent visual dimensions (i.e. major and minor axes). Making connections between the loop of string and the trammel provided a rich problematic experience grounded in an immediately tangible situation. No matter how it was approached, it involved some kind of profound geometrical or algebraic transformation, since the former device starts by establishing the foci whereas the latter gives no immediate indication as to focal position. Algebraic representation of the loop of string device depends on the distance formula while the other two depend on similarity and proportion.

The folding arm device was included because, although its action feels quite different from the trammel, it has close geometric relations with the trammel device. For example, once one sees that the midpoint of the trammel draws a circle one can directly relate the trammel and the folding arm device (with equal arms), and see that they draw the same curves (see Section 2.9). I wondered to what extent high school students might try to use some such direct physical or geometric argument verses a more algebraic one. The folding arm device is also the only one of the three that can be adjusted to draw curves other than ellipses, but as it turned out, in the space of time that I spent with each student (two - approx. 2 hr. interviews), examining this was beyond the scope of the investigation. After an initial interview, I restricted my questions to the case of equal length folding arms.

The purpose of the investigation was to create a set of physical tools, and then to use that environment to ask a series of questions in a setting where direct physical experiments with curves could shape a student's initial beliefs. This investigation focused solely on the second main point of this thesis, i.e. that a rich mathematical experience results from giving geometrically generated curves a primary epistemic role. Chapter 2 touched on a broad range of issues connecting curve drawing actions to tangents, areas, rates of motion, and computer simulations. The investigation with two students, the subject of Chapter 3, was designed to examine the plausibility of using curve drawing devices as a touchstone for exploring and contrasting the mathematical pursuits of students. To this end, three elliptical devices were chosen, and the students here were asked only about those three devices, and about how (if possible) each one could be set up to reproduce curves drawn by another of the three, and how they could be sure that the curves were the same. They were also asked how the action of each device might give rise to an equation of the curve. At no time, however, did I suggest the use of any particular coordinate system, origin, axes, or unit of measure.

They were asked to justify their assertions in any way that seemed convincing to them and with as much detail as possible. From their own hypotheses, formed directly from their experience with primary curve drawing actions, they moved, in different ways, to represent geometric actions with algebraic relations. For both of the students that I will discuss, the ways in which they described their sense of the geometric actions strongly shaped the kind of algebraic language that they employed. This was especially apparent in the case of the folding arm device. Although neither of the students attempted a solely geometric approach of the sort described in Section 2.9, nevertheless the students employed physical geometry in quite different ways.

3.2 The Structure and Method of the Interviews

The subjects for these interviews were chosen from a single class of New York State Regents' Course Four Mathematics at Ithaca High School in Ithaca, NY in the Spring of 1994. Other than the fact that they were in that class, they were not chosen for any special background or ability. I simply took the first three volunteers who were willing to participate in an after school project (one as a pilot and two for analysis). A class presentation had been previously required by their teacher, and they were told that after participating in the project, they would be free to use anything they learned as possible material for such a presentation. They were also paid $5.00 per hour for their time.

Course Four Mathematics in New York State is a high school precalculus course that is taken by seniors as part of the regular Regents sequence beginning in ninth grade with Course One. Some of the students in Course Four are juniors who began an honors mathematics sequence in eighth grade. Most of the juniors in Course Four would subsequently go on to take some form of Advanced Placement Calculus course in their senior year. The class from which my volunteers came was roughly half seniors and half juniors. Among all students at Ithaca High School roughly half will take Course Four at some time during their secondary education.

In order to minimize bias in my expectations of these students, I did not inquire about the backgrounds, records, and teacher opinions of the students that I interviewed until after the completion of the project. I videotaped two individual interviews with each student for approximately two hours per interview. In each case, the second interview occurred one week after the first one. The students were asked not to discuss the project with others until after the completion of all of the interviews. During the week between their two interviews, although they did not have the devices, they were free to work by themselves on any unresolved questions, and consult any mathematical source material that they thought might be helpful. I did not, however, provide the students with any references or background material until after the completion of the interviews. The interviews were all conducted in the same quiet room at Cornell University with the curve drawing devices on one big table. Besides myself and one student subject, the only other person present each time was a video camera-person who was initially introduced to the student, but then remained silent throughout the interview.

After conducting two pilot interviews with one student, I structured the interviews in the following way. I first showed the student the paper covered board and gave them two tacks, an adjustable loop of string, and several colored pencils. I explained how the device worked asked them to draw a curve with this device. I then asked them if they had ever seen such a device (they both said yes), and asked them what kind of curve this device produced (they both answered "an ellipse"). I then encouraged them to experiment with various positions of the tacks and loop lengths. I asked them how they knew that such a device produced ellipses. Depending on their answers I tried to probe in various ways to see what the word "ellipse" meant to them intuitively, geometrically or algebraically. If words such as "focus" or "eccentricity" came up, I asked them what they meant by such terms. I did not introduce such terms until they were used by the students. If they felt unsure as to what such terms meant, I abandoned them and tried to use terms that related directly to the physical devices (e.g. using "tack point" instead of "focus").

I next introduced them to the trammel device which laid on the table beside the loop of string device. I showed them how to use the pen holder and how to adjust and lock the positions of both the pen holders and the moveable pin. I asked them if they had ever seen such a motion before and if so in what context. I then encouraged them to experiment by drawing curves with the device in various adjustments. After some experiments I asked them what kind of curves they thought the device was capable of drawing.

After some more experimentation, I eventually asked them if they thought this trammel device could produce any curves that exactly matched those drawn by the loop of string device and vise versa. I encouraged them to make guesses and then try to verify or refute their ideas with experiments. If they thought a pair of devices were capable of drawing some curves that were the same, I then asked them if this was always the case for any curve drawn by one of the devices. In particular, I asked if there were any curves that one device could draw that the other one could not. As their experiments progressed I asked them to be as specific as possible about how they would go about setting up one device so as to reproduce a curve that they had previously drawn with the other device. I told them to feel free to use any other tools that they thought might be appropriate, such as string, a ruler, a protractor, a square, or a calculator.

Depending on the direction of their responses, I tried to ask questions to find out what convinced them that two curves were same. I tried to discover what aspects, parameters, or forms of analysis were important to them. In particular, I tried distinguish to what extent their beliefs depended on physical appearance, empirical measurements, geometric properties (such as distances, ratios, similarity and angles), or algebraic descriptions (such as equations). I asked them if it was possible to find an algebraic equation for these curves directly from the action of the devices. Following the direction of their inquires, I tried to get them to articulate how they coordinated the results of their experiments in these different areas, and what representations were most important to them in order to feel convinced about their assertions.

After about an hour and half into the first interview the students had all made, and partially justified, some series of conjectures, although none of them were completely satisfied that the trammel and the loop of string devices were drawing exactly the same set of curves. At that point I told them that we would return to these questions, but for now we would look at the operation of a third curve drawing device. I then showed them how to operate the folding arm device on the same sheet of Plexiglas on which the trammel had just operated. I encouraged them to experiment with device, and then asked them the same general questions as before, both about this new device itself and its possible relations with the other two devices. After some initial experimentation and conjectures, I ended the first interview and told them that in one week we return to all of these same questions.

The structure of the second interview depended on the direction of the thoughts of the student. All three devices were available to them and I stuck to the same set of questions that had been raised during the first interview, allowing the students to work with the devices in any order or fashion that they chose. For all three devices they were asked to determine:

(1) Are any of the devices capable of drawing the same curves?

(2) Is there any curve which one device can draw which another one cannot?

(3) How exactly do you go about setting up one device so as to reproduce a curve drawn by another device?

(4) Is there any way to find an equation of a curve directly from the actions involved in the device used to draw that curve?

(5) What convinces you of your claims and how would you go about justifying them?

3.3 A General Comparison of Two Student Approaches

The two students that I will discuss both resolved all of the questions for the three devices, but only after considerable effort. One student, whom I will call Tom, came to the second interview with some specific notes that he had made during the week which satisfied him concerning the first two devices. After an hour and half into the second interview, he was able to satisfy himself about the third device. The other student, whom I will call Jim, brought to the second interview only a few general notes about the loop of string device. Jim picked up from where he had left off the previous week, and then spent two and one half hours, eventually satisfying himself about all three devices. Much more than Tom, Jim seemed to need the physical devices in hand in order to think about them. He paid very close attention to "things that move in proportion," and his physical perceptions were astute. The relative uses by the two students of physical empiricism, geometry, and algebra were strikingly different as were their conceptions of motion, especially concerning the folding arm device.

As I stated earlier, I did not inquire about the backgrounds of Tom and Jim until after I had conducted the interviews, however, a few general comments beforehand may help the reader to understand the interviews and the conclusions that I wish to draw from them. Some of these general observations are drawn directly from what I saw during the interviews, while others come from my informal conversations with Jim and Tom, and with their teacher after the interviews.

Tom was an eleventh grade student who was taking Course Four Mathematics as part of an honors program. He would go on in his senior year to take the most difficult Advanced Placement Calculus course that was offered at Ithaca High School. He told me that he intended to pursue a career in mathematics or science. Tom was described by his teacher as a very good student with exceptional algebraic skills. I witnessed these skills during the second interview when Tom tried to find an equation from the action of the folding arm device. With no hesitation he plunged into a whole series of tedious and complicated algebraic manipulations never losing track of any of his variables or his overall intentions. Throughout his work he made only one algebraic mistake which he eventually corrected himself. Tom, however, was more hesitant when translating his algebraic results back into conclusions about geometry or the physical context of motion.

Tom felt very comfortable with algebra and clearly preferred it to experimentation or geometric language. He did carry out some very systematic and telling physical experiments upon which he based his original algebraic statements, but once he had some equations with which to work he preferred to stay with them. Tom was a quiet and soft spoken person who considered his actions and words carefully. He never made wild guesses and his curve drawing experiments were very premeditated. His answers to questions strongly favored algebraic over geometric language. For example, when dealing with the trammel device, both Tom and Jim sketched in the same set of similar right triangles, but Tom did not use the word "similar," nor did he directly write down any ratio statements, nor did he mention the "Pythagorean theorem." Instead he noted that two angles were the same and then wrote down two trigonometric statements which he then combined using a trigonometric identity, thereby minimizing direct mention of physical geometric concepts.

Tom seemed to regard his time in front of the video camera as a kind of test, despite my contrary urgings. He wanted to impress me with his skills, and indeed he succeeded. At times it was all I could do to keep up with his rapid algebraic derivations. When I was driving him home after the interviews he spoke a bit more openly about his feelings and beliefs. He said that he never believed anything in mathematics until he worked through some set of equations to verify it. He told me that he liked most of his high school mathematics program except the sections of Course Two that dealt with classical geometry. This he found very boring and unsatisfying, and he was very glad when the focus of his mathematics classes returned to algebra. His teacher confirmed this, and added the observation that when the Course Four class had been studying curves whose equations were very simple in polar coordinates, Tom had insisted on finding, on his own, Cartesian equations of these curves despite their enormous complexity. It was as if Tom saw polar equations as some sort of geometric trick, and would not be satisfied that the curves had been properly defined until he saw a Cartesian equation.

Jim was a twelfth grade student taking Course Four Mathematics. Several times during the interviews Jim described himself laughingly as a "terrible student," and said that if his teacher "saw these videotapes he would probably be horribly embarrassed." His teacher described Jim as a fair to average student who had to struggle hard to keep up with his work. Jim's teacher also found him to be very helpful and cooperative in class. He was a very open, friendly and talkative person which made it easy to interview him. He talked almost constantly about what he was doing and thinking with little or no prodding. He seemed to have no inhibitions about being videotaped.

Jim expressed a strong preference for geometry over algebra, and most of all he liked physical experimentation. He said that he really enjoyed "fooling around with stuff." He said that he was fascinated by engineering and that he might eventually study it (I have since found out that Jim chose not to take up Engineering in college). He told me that he wished that there could have been more geometry in Course Four, and that he had really enjoyed the geometry sections of Course Two. Jim obviously enjoyed experimenting with the curve drawing devices, and rapidly generated and rejected a whole series of conjectures about how they might relate to each other. Although many of his guesses seemed, at first, a bit wild and random, his overall pattern of refining his experiments showed an astute sense of geometric proportion and invariance. He voiced many guesses based on things which were visually and physically suggestive to him and in the case of the trammel he immediately saw that it operated in much the same way as a desktop toy that he owned known as a "B. S. grinder."

Jim openly admitted that he easily got lost in algebra, and that he found it very boring. He said that he wished that his algebra skills were better, and he thought that this was something he would "have to work on." During his second interview, Jim eventually expressed algebraically, proportions that he had found from the geometry of the trammel device and the folding arm device, but both times when I asked him if these equations were equivalent to the one that came from the loop of string device he paled at the thought of having to attempt an algebraic reduction. His usual cheerful demeanor seemed to darken abruptly. I told him that he did not have to do this, and I reminded him that he was free to end the interview whenever he wished, but he said that it would it give him some real satisfaction to see the algebra "come out." He asked me to watch his algebra carefully, because he knew that he would make mistakes. Sometimes he even predicted in advance exactly what type of algebra mistakes he was prone to make, and then proceeded several minutes later to confirm his predictions.

I asked him how important it was for him to see the algebra "come out" in order for him to believe that the devices were drawing the same curves. Jim said that he had made a big jump in his belief based upon his procedures for reproducing the curves visually, and that the algebraic confirmation was just one more little step. He gestured geometrically with his hands showing the big jump and the little step. He then estimated the proportions in his gesture at around 8 to 2 and laughed. Jim had very little confidence in his own algebraic skills, and this seemed to transfer over somewhat to his confidence about algebra in general, yet he still wanted to see the algebra confirm what he had learned from his experiments. When he got frustrated, he directly asked me for some algebraic advice, and I offered him a few procedural hints. Once he had corrected and completed his algebra he had no trouble interpreting the results back in terms of the physical reality of the curve drawing devices.

3.4 Tom's First Interview

I began by showing Tom a large paper covered board into which I put two tacks and then placed over them an adjustable loop of string and then demonstrated how to draw a curve by keeping the string taut with a pencil. I did not actually draw a curve, but only indicated the action on a small section keeping the pencil slightly above the paper. I told him to feel free to experiment with this device in any way that he wished, and that I was interested in what kinds of curves could be drawn with this device. I asked him if he had ever seen such a device and he nodded yes. Before he drew any curves, I asked him what he knew about this device and he answered:

T: You can draw ellipses with that.

D: Is that all that it draws?

T: That's all that I've seen it used for.

D: What's your definition of an ellipse?

T: An ellipse is all the set of points ... ummm...... There's the two Foci here (points at the tacks) and these two (points at the string lengths out to the tacks) always have to add up to the same thing..... and all the set of points that qualify that, that's the ellipse.

D: And that's pretty much your definition of an ellipse?

T: yeah.....pretty much

D: Do you know any other way to describe an ellipse?

T: Oh yeah.... something to do with eccentricity less than one, I think.

D: What's eccentricity?

T: I don't remember......but it has to be between zero and one.

D: Do you have any vague or intuitive memories of what eccentricity is?

T: It has something to do with.....like..... how spread out the sides are of a type of.....uhhhh.... conic section (gestures with his hands).

D: How spread out?

T: Like the hyperbolas would have a high eccentricity so they would spread out very far (gestures to indicate a very flat curve)........and then the parabola would be exactly one........

D: And you said that ellipses would be between zero and one?

T: yeah

D: OK so it's some kind of measure of spread-outness?

T: yeah........ a ratio I think.

D: a ratio between what?

T: I don't know

Despite my encouragements, after five minutes Tom was still hesitant to actually draw an ellipse with the loop of string, so I reminded him that I was interested in having him explore the actual physical curves that could that could be drawn with this particular apparatus given control over where to put the tacks and how to adjust the string. I asked him if the particular device on the table could draw all possible ellipses or just some special type of ellipse.

T: It would be able to draw all of the ellipses.... as far as the string is long.

As Tom still made no move to actually draw an ellipse with the loop of string, I next showed him how to operate the trammel curve drawing device, but I did not draw any curves. This time he seemed quite ready to try the device, but before he did I asked him if he had ever seen a device like this one before, or if he had any idea what kind of curve it might draw. He said he had no idea, so I let him experiment with it. He drew one curve using the pen holder between the pins closer to the horizontal one. He then moved the pen holder closer to the vertically moving pin and drew another curve with the pins in the same position. These curves are labeled 1 and 2 in Figure 3.4a

T: Looks elliptical to me.

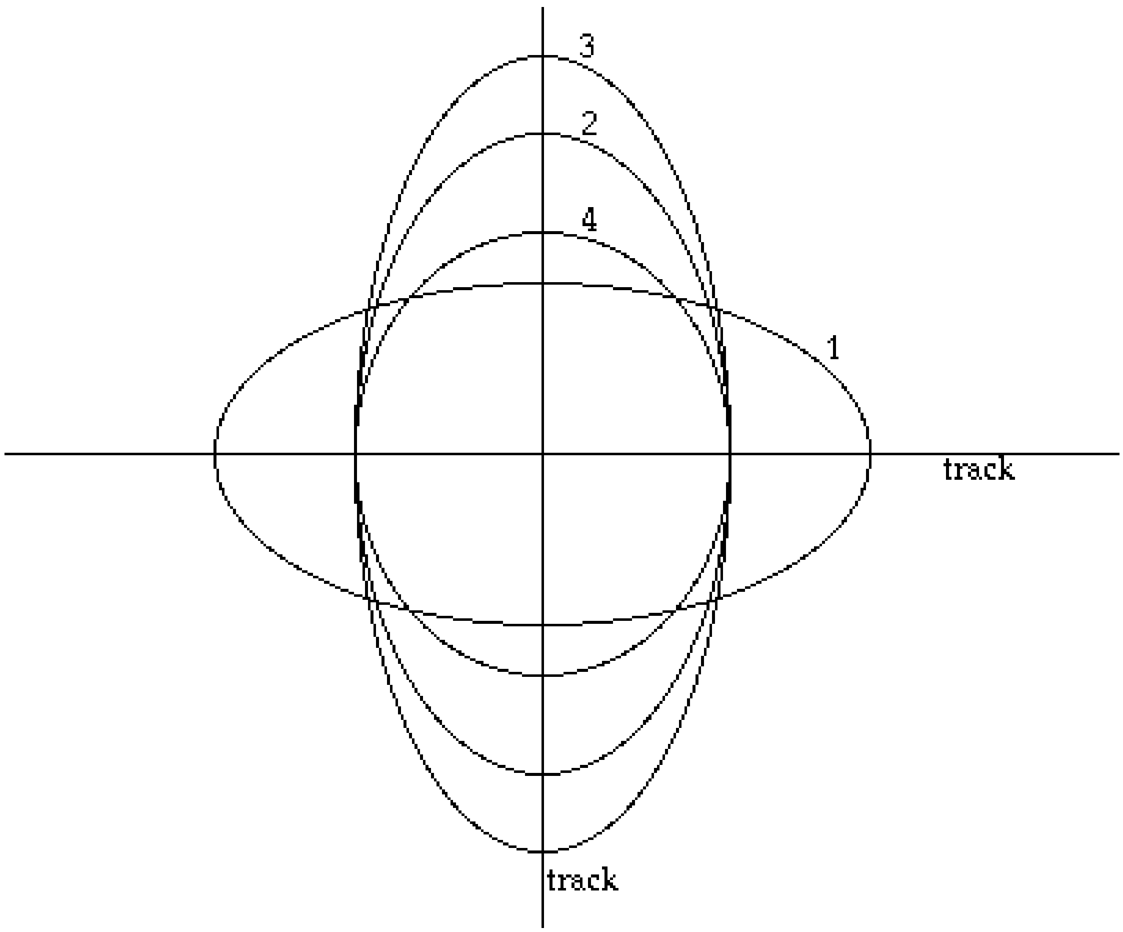


Figure 3.4a

Tom then increased the distance between the pins leaving the pen holder fixed and drew another curve (labeled 3 in Figure 3.4a).

D: What about those last two..... what changed?

(long pause as Tom studies the figure)

T: It sort of dilated...... this way (gestures vertically)...... I stretched it out.

Tom then moved the adjustable pin so as to decrease the distance between the pins, again leaving the pen holder in the same position and drew a fourth curve (labeled 4 in Figure 3.4a)

T: The closer this pin is to that one the more compressed it is.

Tom then asked me if I remembered the order in which he had drawn the curves, and since they were all the same color I was not exactly sure. I offered him another colored pen and an eraser. He then erased these curves, and drew two new ones which confirmed his idea that shortening the trammel while leaving the pen holder fixed between the pins fixed would compress the shape of the curve. These new curves looked much like 3 and 4 in Figure 3.4a. I asked him if these curves still looked elliptical to him and he looked at them for a while.

T: Well I could try to draw them with that (indicating the loop of string device).

D: Do you think you could copy one of these with that device?

T: Well..... if this is an ellipse......

I encouraged him to experiment and he eventually took the loop of string and the tacks, and put them on the Plexiglas. Tom then said there was no way to stick the tacks in the Plexiglas, so I offered to hold the tacks in any position that he wanted while he used the string to draw a curve on the Plexiglas. He told me to hold the tacks in one of the pin tracks at equal distances from the intersection of the tracks. He experimented several times telling me where to hold the tacks as he adjusted the loop of string. Eventually he succeeded by trial and error in tracing with the loop of string one of the ellipses that he had previously drawn with the trammel.

(The interview has now gone on for 17 min.)

T: That looks pretty close. I think if we made the foci exactly where they should be, it would be exactly the same shape.

I then asked him if could come up with any systematic way to reproduce a curve drawn by the trammel with the loop of string, or vise versa how to reproduce with the trammel a curve drawn by the loop of string. I said for example that if someone told him exactly how the trammel device was set up, could he figure out how to place the tacks, and adjust the string so as to draw the same curve. Vise versa, I asked him if someone told him how far apart they had put the tacks and how long a loop of string they had used could he tell me how to set up the trammel, if possible, so as to draw the same curve. I was trying to see if Tom could explore or systematize the physical experiment which he had just carried out, but after thinking for a minute he offered some general guesses as to how he might go about analyzing the two curve drawing devices.

T: I remember that for this one (indicates loop of string) when we did an analysis of how this worked in class..... it was.... like the distance formula between the foci and the point and then you do all kinds of algebra with it and then it come out into a neat equation.... x^2/a plus b^2 over whatever equals uhh.....one. Yeah......and so ....ummmm.... I don't know.... maybe there's some type of analysis to do on that (indicates trammel device)...... It seems like........ what we're doing in class right now is working with parametric equations... like x and y independent according to time.......This seems kind of similar. This seems like the x is moving and the y is moving (gestures towards the trammel's pins moving in the tracks) and if there was some way of getting t out of the equation and going back to the non-parametric equation.... just a function of y.

D: OK so this thing (trammel device) somehow looks parametric to you?

T: Something like that, yeah.

D: It looks like that might be a way to approach analyzing this device? parametric?

T: Yeah.

Note here that Tom seems to have essentially reinterpreted my questions about how to map direct physical settings of the devices in order to reproduce curves. He interpreted them as calling for an algebraic map between the devices. I asked Tom for more details concerning his analysis of the loop of string device. In particular, I asked him what a and b in his equation represented, and how they might relate to actual curve drawing. After some struggling he eventually told me that they represented the lengths of the axes or perhaps half of the axes he wasn't sure which. I then asked him if he could draw with either device a curve with specific length axes; say, for example, a semi-major axis of 12 in. and a semi-minor axis of 5 in. (I did not use these terms I pointed at curves to indicate what I was asking). Tom said that he had no idea how to use the trammel, but that he could do it with the loop of string.

He explained to me how the two lengths in his equation (a and b) and the distance from the center to the foci would form a right triangle, and that he could double this to determine the distance between the two tacks. This seemed to come from his memory, and he did not immediately offer an explanation. Tom started out by writing the equation: $c^2 + b^2 = a^2$, and then putting in the numbers that I gave him ($b = 5$ and $a = 12$), and solving the equation to find that $c = \sqrt{119} = 10.9$ in. He then placed two tacks in the board each 10.9 in. from a marked center along a line. He then placed the loop of string over the tacks to form an isosceles triangle, and adjusted the size of the loop empirically until this triangle had a height of 5 in. He then drew the ellipse and checked that the vertices along the horizontal were 12 in. from the center. He seemed pleasantly surprised that the algebraic relations that he had remembered had worked out in this physical context. Tom was not entirely sure that they would work, since he had never before tested this knowledge in a physical setting.

I then asked Tom to clarify what he had said previously about knowing how to find an equation for this ellipse. He then told me that the general equation of such a curve was $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$, and that in this case the equation was $\frac{x^2}{144}+\frac{y^2}{25}=1$. He said that he had learned this in his math class.

I next asked Tom to try to draw something as close as possible to this same curve using the trammel device. He did not experiment at first, but instead began talking about parametric equations, and what he had learned by experimenting with a graphic calculator in its parametric mode. Throughout this first interview, Tom felt that the action of the trammel device had some direct connection with parametric equations. Eventually he began experimenting with the actual device. After looking at the positions of the trammel when the pen was crossing the tracks, he then said that he could reproduce the curve by setting it up so that the pen was five inches from the fixed pin, and then placing the other pin twelve inches from the pen. (See Section 2.9.) Tom then drew the curve and said that it looked like the one that he had drawn with the loop of string. He always used the pen holder that was between the two pins.

(The interview had now lasted for 36 min.)

I then asked Tom to try to provide some sort of argument for why these two curves coming from two different actions might be the same. Since he had already written an equation for the loop of string device, I asked him how that equation was related to the action which produced the curve. He said that he thought he could explain that. He chose an "arbitrary point on the curve, (x, y)" and then labeled the midpoint between the two foci as (0,0). He labeled the foci as $(c,0)$ and $(-c,0)$. He went on to explain that since the loop of string had a constant length, he could use the distance formula to write down the distances from the foci. He wrote that:

$$\sqrt{(y-0)^2+(x-c)^2}+\sqrt{(y-0)^2+(x+c)^2}=2a, \text{ or } 4a$$

When I asked him a question about right triangles in his figure Tom gave me a puzzled look. He did not see these radicals as coming from the Pythagorean theorem. He viewed it solely as a "distance formula."

Tom showed, from the action of the device, that the sum of these two distances must be equal but he was not quite sure what that constant was. He said you had to write it as $2a$ or $4a$ "because it's convenient to do that, and it simplifies nicely later." Tom said that you had to "do lots of algebra," and that eventually the equation would turn into: $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$. When I asked about the meaning of the letters a, b, and c, he said that "you have to introduce a and b somewhere along the line." He then said the a's in both equations were the same and that $b^2=a^2-c^2$, from his previous relation, and that somewhere in the algebraic derivation he would make this substitution. He said that he thought he could go through the algebra, and figure out the details, and that would tell him whether he wanted $2a$ or $4a$ in the original distance formula equation. Since he seemed to be remembering something that he had seen in class I did not press him for further details. Tom showed no inclination to check the value of the constant sum of the two distances geometrically from the string in the device. The value of the constant on the right side of the distance formula equation was solely a matter of algebraic convenience for Tom.

I then asked Tom to return to the trammel device and try to see whether he could do something like what he had just done for the loop of string device. I said that he had shown me how his equation involving the distance formula had described the action of the loop of string device. I asked him if he could see any way to find an equation of the curves drawn by the trammel device "from the nature of the device." Tom then began to slowly and carefully move the trammel over the curve that he had previously drawn. He reiterated what he knew about setting up the lengths a and b on the trammel, and then engaged in some more slow and careful motion of the trammel along the same curve. He did not draw any new curves but seemed very thoughtful about the action of the device, and how it traced the curve. He mumbled to himself a few vague thoughts about the relative rates of movement of different part of the trammel, and then lit up and began speaking once again about parametric equations.

(The interview had now lasted 49 min.)

T: Oh yeah, I remember.... the parametric equation of the circle was.... it works like this........... It (his calculator) would ask you for graph of x, y... like this ... and for x you want the cosine of t, and for y you want the sine of t.

(Tom writes: Graph $(x, y) = (\cos t,\ \sin t)$)

T: This would be a circle so to make this (the curve drawn by the trammel) I guess you would want..... if this is x and this is y.... 5 inches cosine of t, and 12 inches sine of t.

(Tom writes: Graph $(x, y) = (5\cos t,\ 12\sin t)$)

D: So you think that that might be the appropriate parametric equation of this?

T: It seems that's how it's working, yeah.

D: So what justification could you give me?

Tom explained to me why using just sine and cosine would give a circle, and then turned to his second parametric equation .

T: All I'm doing right here is dilating the circle five out here, and twelve out here, and the ellipse looks sort of like a dilated circle, because the circle's equation is $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ where the things under here are the same, whereas in the ellipse the things under x and y are different.

D: I'm going to be a skeptic, I can see that this might go with the 5 and the 12 (I point to the places where his trammel drawn curve crosses the axes), but the points in between..... how do we know that this (his parametric equation) gives points on this curve?

Tom then began plugging in some values of t into the parametric equation. After checking the intercepts, he chose to check $t = 45^\circ$, because he knew that the sine and cosine were both equal to $\sqrt{2}/2$. Using his proposed parametric equation, he then calculated that there should be a point on the ellipse at roughly (3.5, 8.5). When he went to check for it on the curve, I realized that he also believed that this point would be 45˚ off the x axis. Using a ruler he found that the point (3.5, 8.5) was approximately on the curve but that the angle was more like 65˚.

T: Well on the circle it would be 45 degrees but maybe on the ellipse it's slightly different... how I'm going to account for that I don't know..... It makes sense that it would be slightly different in an ellipse because its stretched out... I don't know exactly how.

D: So the angles don't seem to be checking but at least in terms of the numbers.....?

T: Basically this ellipse is just the dilated coordinates of the circle, and so the angle measurement will work for the circle, but once I dilate it this angle will be off. (Gestures with his hands to show the how the point is stretched away from the 45˚.)

I briefly summarized Tom's claims thus far, and then asked him again how his parametric equations related to the action of the trammel device in the sense that the distance formula equation related to the action of the loop of string.

T: I think it will be easier if we just draw a circle.

D: Can this device draw a circle?

T: It should be able to.... if I just make a and b equal.

Tom then erased the Plexiglas and adjusted the trammel so that the pins were 12 in. apart and the penholder was halfway between them. He predicted that he would get a circle of radius 6 in. He then drew the curve, and said that it looked circular. Looking at where the pen crossed the axes he showed that the radius must be 6 in. I asked him what his definition of a circle was and he said "all points equidistant from a center point." Once again I challenged him by saying that I could see that when the pen crossed the axes it was 6 in. from the center, but how could we be sure about the points in between. He then began slowly moving the trammel and watching the pen trace a quarter of his curve. He began measuring the sides of the right triangles formed by the pins and the tracks with a ruler. He eventually said that these triangles all have the same hypotenuse of 12 in., and that the pen was always on the midpoint of that hypotenuse. He decided that it must have "something to do with the trigonometry of these triangles." He looked at various angles that the trammel made with the horizontal axis.

Eventually he told me that the y-coordinate of the pen would always be half the distance of the vertical pin from the center and the x-coordinate of the pen would always be half the distance of the horizontal pin from the center. He said that the sum of the squares of the pin distances would always have to be equal to $12^2 = 144$. He said that that was "pretty much the equation of a circle," and so the pen was just being "compressed by half."

When I asked him why the coordinates of the pen were always half of the pin distances he wasn't sure what to say. His gestures made it apparent that he had a clear insight into the action of the device, but he could not (or would not) express himself in geometric terms. I expected him to say that the triangles that he was looking at were similar or proportional as his gestures indicated. Throughout both interviews Tom never used the words "similar triangles" or "proportional." This made it much more difficult for him to express certain relationships like the one here. Tom was extremely fluent algebraically (as one can see in the second interview), but when he had to express geometric proportions his gestures were clear, but he did not know what to say. After asking him about this three times I felt that he was getting flustered, and so I said that I understood what he meant and let it go.

Tom then said that if he let x and y be the distances of the pins from the center that since the trammel was 12 in. long then $x^2 + y^2 = 144$. The coordinates of the pen would then be $\left(\frac{x}{2}, \frac{y}{2}\right)$ and then $\frac{x^2}{4} + \frac{y^2}{4} = c^2$, and then $c = 6$ in., so it would be a circle of radius 6 in.

Our time was growing short, so I said that we would return to this device one week later during the second interview. I next showed Tom how to operate the folding arm device. After some experiments Tom had some good ideas about how to adjust the device so that the axes of the curves would be certain lengths. He said that he thought that this device was more related to the loop of string device than to the trammel device. He then said of the folding arm device, "This has something to do with eccentricity."

I told him that I was particularly interested in the case when the lengths of the two folding arms were equal. I said that I would like to ask him in more detail about that case, and also about the first two devices when he came back in one week. In particular I told him that I was interested in how he would know when the curves drawn by different devices were the same, and also how the action of these three devices might generate equations.

I then asked him if had any other specific thoughts about any of the three devices. For the remainder of the interview he talked about the concept of eccentricity, and how he had seen it displayed in computer animations. When I again asked him to define what he meant by eccentricity, he showed me a method for generating discrete sets of points on conic sections of any pre-specified rational eccentricity. The method involved using equally spaced concentric circles, and parallel lines, and was closely related to the curve drawing methods of Ibn-Sina described in Section 2.2. Tom was a little unsure of the definition of eccentricity, and asked if he could consult his notebook. He showed me a worksheet from his math class that plotted a finite number of points on various conic sections with rational eccentricities. This involved using a directrix line. Tom said if you knew the eccentricity of the curve you could use the distance formula to find its equation.

After he talked about eccentricity for a few minutes I asked him what was the eccentricity of the curve that he had drawn with the loop of string where $a = 12$ in. and $b = 5$ in. Tom tried in various ways to relate his homework sheet to this curve, but he never could come up with a clear answer. Finally he said that he could not figure this out because he did not know where the directrix was. He worked very hard for about 15 min. to figure this out. He tried to work backwards from the homework sheet to figure out the relationship between eccentricity and the lengths of the axes and focal distance. At one point he said that eccentricity might just be c/a but he then rejected that idea. In the context of the homework sheet Tom could generate discrete points on curves with a given eccentricity, but he could not transfer that idea to curves that had been generated by another method even for the loop of string which was familiar to him. Tom was familiar with generating curves from given algebraic parameters, but he could not easily reverse his thinking when those parameters had to be determined from preexisting curves. When Tom returned a week later he abandoned his attempts to use eccentricity as tool for understanding these curve drawing devices.

I think Tom's teacher went further than most in his attempt to give his students a visual and geometric idea of eccentricity. Many teachers simply define it as c/a and leave it at that. Perhaps such a straightforward algebraic statement would have been clearer for Tom given his algebraic preferences. His instincts about eccentricity being a measure of how spread out a curve is were very well described, and I found it surprising that he could not measure it in a broader context. I think that this was related to his general reluctance to use geometric language, and to connect such language to exact statements about ratio and proportion. When Tom returned a week later he resolved all of the questions that I asked about the curve drawing devices, but his uses of geometric ratios are all indirect and cloaked in algebra and trigonometric identities which make his derivations more complicated. He never returned to his discussion of eccentricity.

3.5 Tom's Second Interview

When Tom returned a week later for a second interview he brought with him his own written notes (see Figure 3.5a and 3.5b) on why he thought the trammel device drew ellipses. Even before I could turn on the video camera he cheerfully announced that he had figured out how the trammel worked. When the camera was on he began explaining his notes to me.

T: I figured out why this (picks up the trammel) would make the ellipse.

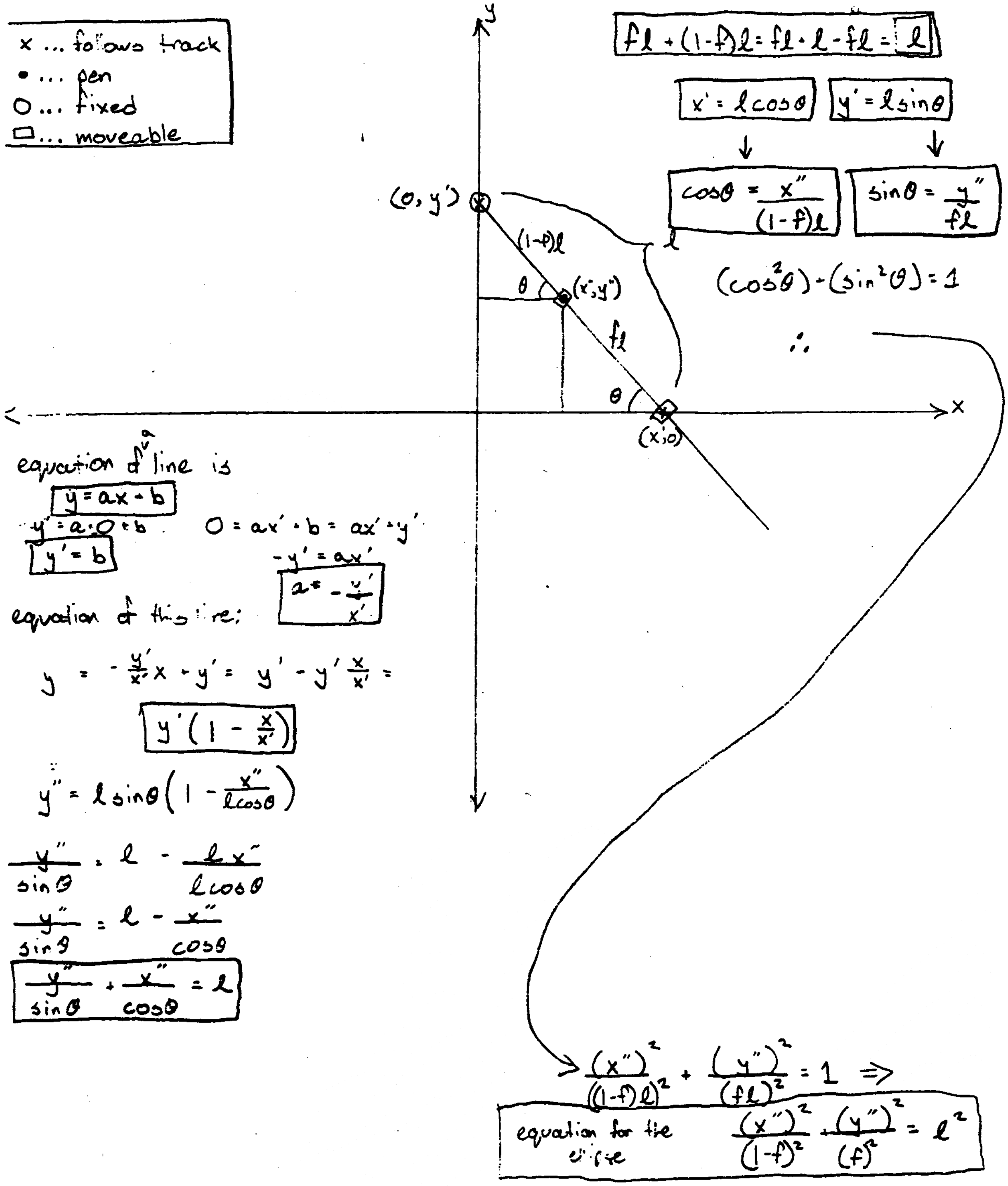
D: Tell me what you thought about.

T: I did this algebraically.

Tom picked up the trammel and put it in the tracks and began labeling things on the Plexiglas using the notation from his notes (Figure 3.5a). He reconstructed for me the figure and the equations that appear on the right side of his sheet (Figure 3.5a). He was not simply copying things from his notes, but truly reconstructing them since he hardly looked at his notes, and on the Plexiglas, he reversed the roles of $f - l$ and $f(1-l)$, from how they appear in his notes. His equations on the Plexiglas stayed consistent with his labels there.

He seemed excited to show me his work, but he showed no inclination to draw a curve, and avoided as much as possible using geometrical language. I had to stop him, and ask him to actually draw the curve that he was analyzing. In his own notes there is a figure of the trammel, but no curve appears. Note that he designated the position of the pen on the trammel by using an algebraic notation for proportion (i.e. $f - l$ and $(1-f)l$) instead of using geometric lengths. Note also the abundance of variables in Tom's set up especially his use of x', x'', y', and y''. A major difference between Tom and Jim is Jim's reluctance to introduce intermediate variables, while Tom never hesitated to introduce algebraic variables.

Throughout his description he never used the term "similar triangles." He did say that the two marked angles were the same because of parallel lines and he quoted the standard trigonometric definitions, e.g. "cosine equals adjacent over hypotenuse." When he formed his equation by substituting into $\sin^2\theta + \cos^2 = 1$, he said that he was "using the Pythagorean theorem as a trig. equation." Other than that, he used no geometric language. Tom is not really using the angle θ as a parameter in his analysis. The angle along with the trigonometric functions serve as means to avoid mentioning similar triangles, and yet still arrive at the equivalent ratio statements. The trigonometric functions are mentioned as codes for ratios, and then immediately disappear by substitution. Later in this interview one can see how Tom used this same approach as he worked with the folding arm device.



x ... follows track
• ... pen
O ... fixed
□ ... moveable
$f\ell + (1-f)\ell = f\ell + \ell - f\ell = \ell$
$x' = \ell\cos\theta$
$y' = \ell\sin\theta$
$\cos\theta = \frac{x''}{(1-f)\ell}$
$\sin\theta = \frac{y''}{f\ell}$
$(\cos^2\theta) + (\sin^2\theta) = 1$
∴
(0, y')
(1-f)ℓ
(x'', y'')
fℓ
(x', 0)
equation of line is
$y = ax + b$
$y' = a \cdot 0 + b$
$y' = b$
$0 = ax' + b = ax' + y'$
$-y' = ax'$
$a = -\frac{y'}{x'}$
equation of this line:
$y = -\frac{y'}{x'}x + y' = y' - y'\frac{x}{x'} =$
$y'\left(1 - \frac{x}{x'}\right)$
$y'' = \ell\sin\theta\left(1 - \frac{x''}{\ell\cos\theta}\right)$
$\frac{y''}{\sin\theta} = \ell - \frac{\ell x''}{\ell\cos\theta}$
$\frac{y''}{\sin\theta} = \ell - \frac{x''}{\cos\theta}$
$\frac{y''}{\sin\theta} + \frac{x''}{\cos\theta} = \ell$
$\frac{(x'')^2}{((1-f)\ell)^2} + \frac{(y'')^2}{(f\ell)^2} = 1 \Rightarrow$
equation for the ellipse
$\frac{(x'')^2}{(1-f)^2} + \frac{(y'')^2}{(f)^2} = \ell^2$

Figure 3.5a

I tried to get him to explain how he came up with his idea, and he told me a little bit about his original attempts to use the equation of the line of the trammel. He told me that he struggled with various algebraic forms for the equation of this line and that he eventually began writing it using the $\sin\theta$ and the $\cos\theta$. This got him thinking in terms of trigonometry and then he found the elliptic equation without using the equation of the line. On the left side of his notes (Figure 3.5a) he wrote down various forms of these linear equations for the trammel line that he had looked at before he arrived at his method for finding an equation of the path of a point on the trammel.

D: Could you tell me,.... or could you reconstruct for me how you came up with using the trig. identity and these triangles to see it?

T: This is like my final version, my rough draft was much messier....... Well first I tried to figure out what the equation of the line would be (indicates the line of the trammel)...... And I got all confused in that, and I came up with the equation which was $y'' = l \cdot \sin\theta \cdot \left(1 - \frac{x''}{l \cdot \cos\theta}\right)$ (see Figure 3.5a) and that didn't help me at all.

I asked him if he had used the concept of eccentricity in his work on this problem

D: We had a long discussion last time about eccentricity...

Y: (interrupts) I couldn't find anything anywhere.

D: So eccentricity seemed <u>not</u> to be a helpful idea here?

T: No it didn't help me, I couldn't find anything anywhere.

D: So you never worked on that further?

T: No.

I asked him if his equations fit with the results of his experiments from last time, and told me he had tested a few examples. He then described what appears on second sheet of the notes that he brought (Figure 3.5b). He said that his equation fit perfectly with what he knew about the device in terms of how the lengths of the axes corresponded to the lengths on the trammel. His idea of a test was a mental visualization, and did not involve any actual curve drawing. To be sure of what he was saying, I did ask him to demonstrate his example for me by drawing with the trammel the curve that appears in his notes where $a = 11$ in. and $b = 6$ in. He did this for me but it seemed unimportant to him.

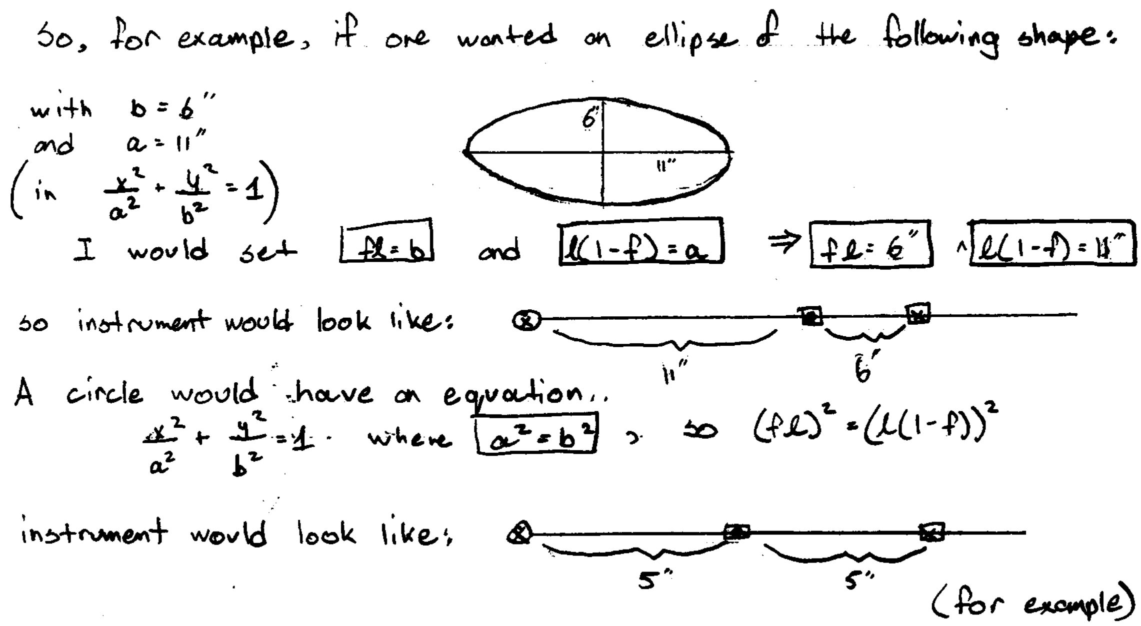


Figure 3.5b

Finally I asked him about what really convinced him that the first two devices were drawing the same curves, and he indicated his equation.

D: So that (his equation of the ellipse) convinced you that the loop of string and the two tacks draws the same curves that this (the trammel) draws.

T: Yeah, you just replace this and this (points to the denominators in his equation from Figure 3.5a) with a^2 and b^2.

D: Is this equation a thoroughly convincing thing to you?

T: Yeah (nods with an air of finality).

(The interview had now lasted 15 min.)

I next asked Tom to consider again the folding arm device with the lengths of the two arms set equal to each other. I asked him if could determine what kind of curve this device could draw. In particular I asked him to compare the curves that it drew to those drawn by the other two devices. Tom erased the Plexiglas and began experimenting with the device leaving the length of the two folding arms fixed, and moving the pen closer and closer to the hinge. He drew four curves as shown in Figure 3.5c. Near the vertical track the two arms collide it is only possible to draw the entire curve by lifting one arm over the other and continuing on the other side. Only for the last curve did Tom eventually draw the other half. Tom studied these and then slowly moved the arm back and forth carefully watching the action as it traced back and forth over a piece of the curve.

Tom then began looking at the lengths of the axes in the last curve that he drew. He placed the arms so that they were both laying flat on the horizontal track and then folded them up as close as they would go to being vertical and on top of each other. He then folded them back flat onto the horizontal track, studying where the curve reached its horizontal and vertical extremes.

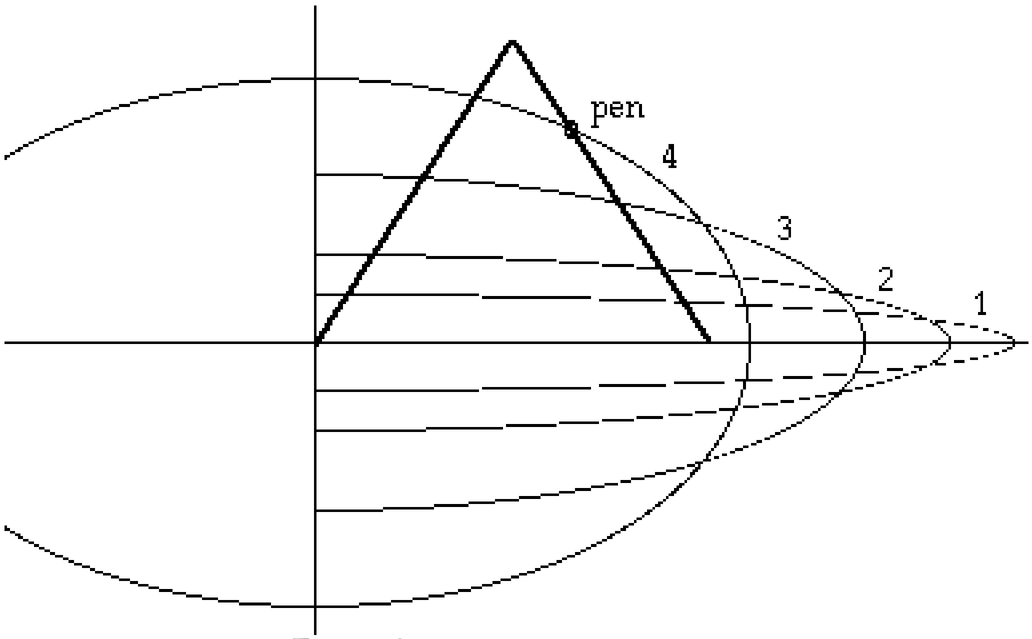


Figure 3.5c

T: Well if this was an ellipse...... then the length between here and here would be the "a" (indicates the semi-major axis)....... and.......... this length (semi-major axis)...... minus twice this length (indicates the distance from the hinge to the pen) would be the "b."

D: Well, is there some way we could be sure what kind of curves these are?

T: (moves the stick back and forth slowly over the curve and begins speaking very softly to himself) What's changing?...... hmmmm....... This distance is changing (indicates the distance from the center to the pin in horizontal track)....... The fact that this bends (points to the hinge) makes it different....

D: Different?

T: Yes different.... More difficult for me to visualize........ Well... maybe I could try to do the same type of analysis.....

Tom then erased all four curves and placed the folding arms back on the clean Plexiglas and contemplated its motion in the first quadrant again. He then stopped the arms in an intermediate position and began writing labels on the Plexiglas. He began as before by labeling the position of the pin moving in the horizontal track as $(x',0)$. He next labeled the position of the hinge between the two arms as $(x'/2, y')$. When I asked him how he knew that that hinge would always have half the x-coordinate of the moving pin Tom indicated the two equal length arms. I pressed him for more justification in an attempt to elicit some geometric language from him. He gestured that this was obvious, and when I asked him again he drew a vertical line from the hinge to the track and said that the two triangles were the same. I then asked him why they were the same.

T: Side angle side..... So there!

Tom 's tone seemed to imply, "Why are you bothering me with these silly geometric details, can't you see I'm trying to work on this!" so for a while I stopped asking about such details and let him proceed.

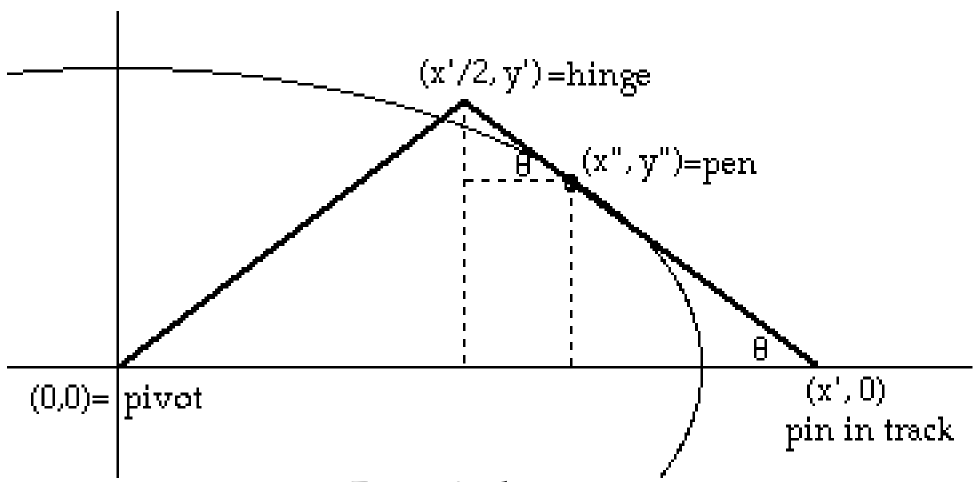


Figure 3.5d

As before he labeled the point on curve (i.e. the pen holder) as (x'', y''), and let the lengths of the two equal arms both be l. He then began to consider the circular path of the hinge pin at $(x'/2, y')$. He then began thinking about his experiments with moving the pen holder. Tom pointed out that if the pen were on the hinge it would draw a circle, while if it were moved all the way down to the pin it would just trace a line segment.

T: The degenerate thing for a circle as a conic section was a point.... but for ellipses..... let's see......What's an ellipse?..... an ellipse is when you cut it at an angle....... it would still just be a point. I don't see how it would be a line.... I don't know...... How would I make an ellipse with a's and b's and have it form a line? I don't see it...... I guess I could have $b = 0$ and $a =$...... whatever.....

Tom drew a crude sketch of an elliptic slice of a cone and thought about how to slice it so as to get very long and thin ellipses. He eventually thought about slicing it right through the apex and along one side to get a line.

T: The moment that I get right here (indicates his sketch of cone being sliced through the apex along one edge).... then it would be infinitely long .... but this has endpoints (indicates degenerate case on the folding arm device when the pen is at the pin).

D: So that bothers you?

T: yeah.

D: It looks somehow different?

T: um hmm.

(The interview had now lasted 33 min.)

Tom decided to return to his analysis of the folding arm device, not sure at all what kind of equation it might produce. He had some serious doubts about whether these curves were ellipses. Several times during his algebraic analysis, Tom was quite willing to believe that he had proved that these curves are not ellipses.

T: I guess I could try the same trick as last time.

Tom finished labeling his figure and then began talking to himself sometimes inaudibly.

T: So what is the difference between this and the previous one?......In this one the entire line moves over...... in addition to going like that (indicates a trammel type of motion by gesture)..... you see?....... In that one it was just the line part going like this (picks up the trammel and demonstrates it's motion) and changing the slope..... In this one it does the same thing but also moves at the same time (gestures to indicate that the outer arm of the folding arm device with the pen moves like a trammel, but with an additional horizontal motion). So... ummm..... wait.... but I know that the uhhh...... the distance between the beginning of this line and the origin will always be one half of the total.

D: Beginning of which line?

T: This one right here (indicates vertical line dropped from hinge).

(long pause)

T: Well OK... so I can assume that without the movement, this is just going to be the formula for the same one.

At this point Tom began writing a sequence of algebraic statements based on his vision of the folding arm device as a combination of two movements. The first being the trammel like action of the outer arm on which the pen is riding, while the central pivoting arm adds a second horizontal motion. He began by copying down his equation for the trammel and then trying to figure out how to account for the second horizontal motion caused by the pivoting arm. Tom wrote:

$$\left(\frac{x''}{f \cdot l}\right)^2 + \left(\frac{y''}{(1-f) \cdot l}\right)^2 = 1$$

T: But then I think I'll have to add something else here (points to the x in the equation).

D: So now your trying to see this stick as the same as the other stick but with...

T: With some type of translation......

Tom gestures to indicate a trammel on a moving wall.

D: So now you're going to take...

T: The equation for this point (the pen) with respect to that line (a vertical line from the moving hinge point).... and how that moves (gestures again) .... is just like that (points to his equation).... but also (gestures for the forward motion).....

D: But add on a translation? ..... OK ...... well what translation would you try to add on?

T: The translation would always be..... well OK...... it would always be..... I don't want to put it in terms of x'. I'll put it in terms of x''.... and I know that... hmmm.... How do I do that? (long pause)

D: Are these (the variables in his equation) really the same now as what you had in this device? (the trammel).

T: It seems to me.

D: Tell me why you think these are the same.

T: It's basically the same triangles... these are the coordinates...and.... the same angle (points at right half of Figure 3.5d, and gestures as if to move the whole right side of the diagram over to the origin).

I questioned him about some of the details and what his variables represented, and he insisted that the same derivation that he had used on the trammel was valid here if he could find the correct translation.

D: Are these really the same? The stuff that you showed me before would work on these two triangles?

T: It works the same way..... But what would I move it by?..... hmmm

At this point I did not really understand where Tom was going with this approach. I could not see any way to express the kind of translation that he was thinking of. I asked a few questions trying to return the discussion to the physical geometric action involved in the device. These questions were entirely irrelevant to what Tom was trying to do. He patiently answered them, but gave me a look that said "So What! why are you asking me this?" I knew of at least four different analytic approaches to this device, and my own preconceptions temporarily blocked my understanding of what Tom was attempting. After two or three minutes Tom ignored me, and abruptly returned to what he was trying to do, and I let him proceed. Soon, I was to learn another way to analyze the folding arm device.

Tom decided to begin by solving for y'' in the equation that he had taken from the trammel device. He began by letting $a = f \cdot l$ and $b = (1-f) \cdot l$.

T: I'm going to work this into a function because its hard for me to see this.

Tom wrote: $\frac{y''^2}{b^2} = 1 - \frac{x''^2}{a^2}$

$$y''^2 = b^2 - \frac{b^2 x''^2}{a^2}$$

$$y'' = \pm\sqrt{b^2 - \frac{b^2 x''^2}{a^2}}$$

T: I also want to translate the x'' by $x'/2$.

D: Why did you want solve this for y?

T: Because it helped me to see it.

Tom proceeded to put in the translation straightforwardly by replacing x'' with $x'' - x'/2$. He said that he wanted the translation to be to the right so it had to be a minus $x'/2$. He explained this algebraically. He did not point to the length in the small upper triangle (see Figure 3.5d) that equaled $x'' - x'/2$. For him to think about this translation algebraically, he had to have the equation solved for y'' i.e. written as a function.

He next wrote: $y'' = \pm\sqrt{b^2 - \frac{b^2\left(x'' - x'/2\right)^2}{a^2}}$

D: These are two different x's ?

T: Yeah I want to find x' in terms of x'', that's what I'm trying to think of..... of course it would be different for every $\left(x'', y''\right)$ ... Let's see, how would that work?

D: Which one are you trying to find from which?

T: I want to find $x' =$ so and so, and so and so x'' (writes: $x' = \text{..........} x''$)

D: So you want to compute x' as a function of x'' ?

T: Yeah.

D: Which variables do you want in your equation at the end?

T: I just want y'', which is this one, x'', which is that, b and a. That's it. Then I can see if it's an ellipse.

(The interview had now lasted 50 min.)

Tom looked back at his notes (Figure 3.5a), and at his figure on the Plexiglas, Figure 3.5d. He thought for a minute and then began writing down trigonometric statements using the triangles Figure 3.5d. Each time that he wrote down a statement, I asked him which triangle he was looking at. Recall that here Tom refers to the distance from the hinge to the pen as "a" and the distance from the pen to the moving pin as "b" and the length of the arms as "l" (hence: $l = a + b$). Tom wrote:

$$\frac{x'}{2} = l\cos\theta \quad \text{(from the triangle with the arm as hypotenuse)}$$

$$\frac{x' - \frac{x'}{2}}{a} = \cos\theta \quad \text{(from the small upper triangle)}$$

$$y' = l\sin\theta \quad \text{(from the triangle with the arm as hypotenuse)}$$

$$\sin\theta = \frac{y''}{b} \quad \text{(from the small lower triangle)}$$

Tom next began working with the last two statements and got y' as a function of y'', but then realized that that was not what he was looking for. He next worked with first two statements by getting two expressions for $\cos\theta$, and then setting them equal to each other. Again one sees Tom using his previous approach. He never mentioned similarity; instead using trig. functions to code ratios, and then eliminating them by substitution. He wrote down as many algebraic statements as he could, and then abandoned the figure and began experimenting with various algebraic substitutions and simplifications. His algebraic manipulations were rapid, and he seemed to enjoy this. He had finished substituting the last two equations with the y's before he ever thought about what it was that he was looking for.

In order to relate x' and x'', Tom wrote:

$$\frac{x' - \frac{x'}{2}}{a} = \frac{x'}{2 \cdot l} \qquad \text{(both equal to } \cos\theta\text{)}$$

$$\frac{2x'' - x'}{2a} = \frac{x'}{2 \cdot l}$$

$$2x'' \cdot l - x' \cdot l = x' \cdot a$$

$$2x'' \cdot l = x' \cdot (l + a)$$

$$x' = \frac{2x'' \cdot l}{(l + a)}$$

T: (Looking at the curve's equation into which he intends to substitute this) But really I'm looking for $x'/2$, so I can just get rid of that 2.

Tom writes: $\frac{x'}{2} = \frac{x'' \cdot l}{(l + a)}$

T: And so my finalized equation is......

At this point Tom was writing so fast that I could hardly follow him . I stopped him for moment to get him to say what he was doing, but he was anxious to make the substitution and see the curve's equation. He proceeded to make the substitution and write: $y'' = \pm\sqrt{b^2 - \frac{b^2\left(x'' - \frac{x''\cdot l}{1+a}\right)^2}{a^2}}$

T: So I'm going to try to rearrange that in this form right here (indicates the standard elliptic equation).

D: To see if it looks like an ellipse?

T: Yeah.[37]

Tom's first attempt at rewriting his new equation led him to a fourth degree equation. He then concluded that the curve was not an ellipse. He worked so quickly, that I could not follow what he had done. I asked him to review what he had done step by step, and long before I saw what was going on, Tom spotted a mistake in his derivation. He had read the minus sign after x'' in $b^2\left(x'' - \frac{x''\cdot l}{l+a}\right)^2$ as multiplication. He quickly erased what he had written before I could even look at it, and began again. This time he decided to replace the l's in his equation with $(a+b)$'s since they were equivalent. He also decided to drop the prime notation since all occurrences of x and y in the equation were now double primes. After some work he wrote:

[37] At this point Tom eagerly dove into another round of algebraic manipulations. Several times during the interview I had asked him to write on paper so that I could save his work, but he much preferred to write on the large Plexiglas sheet with the dry-erase pens. He wrote large and kept the rag handy for easy erasures. In the next set of interviews, Jim also loved writing on the Plexiglas. I was reminded of an historical description of how L. Euler worked with his students on a large slate table using chalk. Tom did not always write sequentially. He loved to jump around checking, comparing, substituting and rearranging terms in his equations. Having everything spread out on the table seemed to help him.

$$\frac{y^2}{b^2} + \frac{(x(2a+b) - x(a+b))^2}{a^2(2a+b)^2} = 1$$

T: That's it.

D: Does that tell us whether it's an ellipse or not?

Tom began scrutinizing his equation in this form and his opinion began shifting back and forth as to what it meant. He saw that it was of second degree, but he thought for a while that it was the equation of an ellipse with its center on the x-axis, but not at the origin. He was seeing the subtraction in the numerator of the fraction involving x as some sort of horizontal translation. I asked him when he drew the curves where did the center appear to be? He told me that the center seemed to always be at the origin. He became perplexed. I asked him to remember what he had said in the beginning about the lengths of the axes of the curve in terms of how the device was set up. After looking at the device he decided that the axes, when expressed in terms of a and b, were always of length b on the y-axis and $2a+b$ on the x-axis. He then went back to work on his equation and simplified it further and wrote:

$$\frac{y^2}{b^2} + \frac{x^2}{(2a+b)^2} = 1$$

I was very surprised at his next comment.

T: I don't think that's an ellipse.

D: Oh......you don't think that's an ellipse?

T: Cause an ellipse would not have this b here (points at the b under the x term)

D: Hmmm......So you think that that tells us that we have the equation of some different curve?

T: I think so..... I'm not sure....... well it might just be that a would be bigger than b, by ahhh....... b......

I once again pointed back to the device and asked him how his equation related to it. He looked again at how the device crossed the axes.

T: So it's like..... a restricted ellipse...... it's like an ellipse that can only be certain values....maybe...........No.....it's an ellipse (smiles broadly).

D: It is an ellipse?........ You're changing your mind?

T: Yeah.

D: Why is it an ellipse?

T: Well..... I was confused because there was a $2a+b$ under the same term.... I wasn't sure if that was.... ahh..... good enough....... but now I see how that's just another quantity.... so..............You could do the same thing that we did with the $f \cdot l$ and $(1-f) \cdot l$, and with the a and b.

D: OK

T: And so this looks elliptical to me, and I could draw any ellipse with it.

(The interview had now lasted 1 hr. and 15 min.)

At this point Tom was completely convinced that the folding arm device with equal length arms was a general elliptic device. The constants a and b were not playing the same roles as in the standard equation that he was used to seeing, but he did see how they could be adjusted in this new situation to obtain any lengths as major and minor axes.

What comes out in Tom's last few doubts about this device is the dominance in his thinking of algebraic format. If he could not see the new equation in exactly the same form as his previous elliptic equations then he was quite ready to reject its being an ellipse without reference to a geometric representation or to his physical experience. With the exception of his thoughts about comparing degenerate conic sections to the extreme cases in the folding arm device, his doubts and beliefs were based on algebra with which he was quite comfortable. I think it was especially telling when he looked at the equation: $\frac{y^2}{b^2}+\frac{(x(2a+b)-x(a+b))^2}{a^2(2a+b)^2}=1$, and thought that it must be a translated ellipse with a center on the x-axis away from the origin. He seemed quite ready to believe this based on a hasty inspection of the general form of the equation, even though it contradicted all of his physical experience with the device. I had to specifically ask him to look at the device and the curve again. He then saw that the pivot point was always the center, and that he was always using this point as the origin of his coordinate system. This spurred him to reconsider the form of his equation, but I do not believe that he would readily have gone back and forth between representations on his own. This was certainly not his habit.

This situation illustrates one of the main points of this thesis; that we must create situations were different mathematical representations are used to generate knowledge cyclically. Tom was not used to dealing with a situation where moving geometry generates algebra. His natural inclinations were to see the algebra as controlling the shape of graphs. His graphing calculator worked entirely from this epistemic view. By having the physical tools right in front him, once he was reminded to look at them, he could see that he had to reconsider his algebraic form. By directing Tom's attentions to the physical actions of the tools themselves, he had to look at knowledge flowing in a different direction.

Even though this was not his usual habit of thought and action, Tom was not frustrated or alienated by this situation. Quite the contrary, Tom found this new situation challenging and interesting. He came back (without pay) after the interviews were over, to talk to me about other curve drawing devices, their history, and their computer simulations. He said that he thought there should be more "engineering type stuff" in high school mathematics. Towards the end of the first interview when it had lasted nearly two hours, it was late in the day, and I was tired and had to go home for dinner; I asked Tom, "So, are you about burned out for the day?" He answered in his usual quiet monotone, "I'm still having fun."

3.6 Jim's First Interview

As with Tom, I began the interview by showing Jim the adjustable loop of string, the two tacks, and the paper covered board, and demonstrating the action without actually drawing a curve.

D: Have you ever seen this device before?

J: Something similar, we've sort of covered ellipses in class.

D: So you already have a name for what this draws (Jim nods). Would you like to try drawing with this just to see?

J: Sure, I'm expecting an ellipse at least. (Jim draws one complete curve.)

D: Did you get what you expected?

J: Looks like it, yeah.

D: That looks like an ellipse to you?

J: Umm hmm, of course as these (the tacks) become further and further apart it's going to become more and more eccentric; closer this way (gestures vertically along minor axis) and wider out this way (gestures horizontally along the major axis).

D: OK, would you like to try it one more time?

J: Sure.

D: Go ahead, move it any way you want.

J: (Leaves one tack alone and moves the other one further away. Draws a new curve using the same size loop of string.) That's the basic idea.

D: You say that you moved the tacks further apart and it became more eccentric?

J: I think so, yeah.

D: Is that what you expected?

J: Yeah.

D: What does "eccentric" mean to you?

J: Well..... further away from a circle. I mean if you just had one tack here, say right in the middle (Remove both tacks and places one in the middle and places the loop over the one tack) I would expect, I think most people would,.... just to draw the simple circle. (Draws a circle) That's what it looks like to me. As... with two tacks, of course I could have used two right next to each other and it would have still looked like a circle even though it wouldn't have been probably because there's still a slight distance between them ...... what we call focal points in our class (indicates a large loop of about 20 in. over two tacks about 1 in. apart).

D: OK, call them whatever you feel comfortable with. So if they were close together you say it would still look like a circle but might not actually be one?

J: Yeah, that would be my guess.

D: So "eccentric" to you is something that's farther away from a circle?

J: Yeah, I hope that's not too far from the real meaning but that's what it means to me.

D: Is eccentricity just a word or is it associated with anything else? Are there any other thoughts you have about what eccentricity is?

J: Well, from the definition of the word, when you say someone is eccentric, they're a bit off.... away from the norm, which sort of makes sense cause if something is eccentric (gestures towards ellipse), it's further away from a circle, which might be a norm because its relatively stable. But as far as this thing goes, there's a definition that my teacher gave us, and I'm accustomed to using it.

D: Do you remember what that definition is?

J: We had a certain point scale. As eccentricity increased, an object went from a point, to a circle, to an ellipse, to a hyperbola at, I think, eccentricity one.

Jim described a computer animation of conic sections that he seen and how he remembered the numeric scale of eccentricity in terms of the visual pictures he had seen, where one focal point went to infinity, and the remaining visible piece of an ellipse resembled a parabola.

D: So eccentricity can be a number?

J: Yeah

D: Is there any way that you could estimate or compute the eccentricity of one the curves that you've drawn here?

J: Probably, yeah if I assigned... (begins putting in the tacks to draw a new curve).... I'm a terrible student, I don't remember the actual number system, but what we did was ahh..... it was a basic relation between this length here, and this length here, and this length here (indicates the sides of a right triangle formed by the loop of string with its right angle at one of the tacks).

Jim used the loop of string to create a right triangle with the right angle at one of the tacks, and said that he would use that tack as the origin of a coordinate system for an equation of the curve, letting the other tack be on the x-axis.

J: I like to keep things as simple as possible and right triangles are relatively simple. You can use the Pythagorean theorem.

Jim studied the motion of the string carefully along the curve and began describing how the distance between the two tacks caused the point on the curve to be pulled in towards "the absolute center" as the point traveled away from the end of the major axis. I asked him what he meant by the center and he then decided to mark the midpoint between the two tacks and told me that it was the center because the curve was "symmetrical through that point." Jim then showed me how the distance from the center varied from the end of the major axis to the point where the string formed a right triangle at one tack (not the end of the minor axis). He told me that eccentricity had something to do with the difference in the distances from the center of these two points on the curve.

J: As an object becomes more eccentric there's going to be a greater difference between these two lines (the segments from the curve points to the center).

Jim began trying to tell me about the eccentricity of ellipses as compared with hyperbolas. I asked him if he thought this device could draw hyperbolas, and he began experimenting. He began pulling at the loop while he traced a curve. He said that he thought maybe he could do it if he could used a third tack. I asked him for the time being to stick with only two tacks and a fixed loop of string.

J: I don't really see how you could draw a hyperbola from this arrangement. Maybe you can..... I'm just probably not looking...... I don't see it.

(The interview had now lasted 11 min.)

I next asked Jim if he knew any way to find an equation of these curves from the action of the loop of string device.

J: Well we were supposed to know this... ahhh......

D: What might you try?

Jim explained how he would first measure the distance between the two tacks, and then measure the loop of string which he looked at when it lay in two equal pieces along the line of the two tacks. He then repositioned the loop at another point on the curve, and said of the length of the loop that it was "of course the perimeter of this triangle."

J: Then at any point, of course, the perimeter is going to be fixed. Soooo.... oh how would I do this?.......(long pause as Jim examines the action of the device along a quarter of the curve)..... As far as writing a direct equation, I'm drawing a blank here, but I would definitely look at these two lengths here (again places the loop so that it collapses onto the line through the two tacks, and indicates half the loop length and the distance between the two tacks).

D: Would those two measurements be enough to determine an equation or would you need more information?

J: It seems to me that that should be enough, because all that we're using are these two things.

Jim explained with many gestures and motions of the device along the curve how those two measurements were enough to completely determine the curve.

J: By varying these two distances we can vary the shape of these drawings, so as far as writing an equation, I would think that these two distances would be the only pertinent information, because certainly the angles aren't fixed as we slide it around the angles are changing (demonstrates).

Jim played with the device some more considering various positions but had no idea how to translate his observations into an equation.

D: You don't have to do it right now, but you think that you could do it maybe?

J: Yeah, if I thought about it, probably, yeah.

D: You're pretty convinced that that would be enough to get an equation?

J: I hope so! Because that's all I see right now.... yeah, I'm pretty certain, because it seems those are the only two things that are interacting on this system right now (indicates the two distances, although he seems to have switched from the entire string length to half of it as seen in the collapsed position).

D: Well let me show you something different. This is another device that draws a curve.

I pointed towards the trammel and the grooved Plexiglas. Before I had even moved the trammel in the tracks, Jim said:

J: Oh yeah, I have one of these. It slides along the track right.

Jim described a desktop toy that he had been given known as a "B.S. grinder" which moved in the same way as the trammel with a crank attached to a point on the trammel extended beyond the two pins in the tracks. His toy could not be adjusted, and moved in only one curve. I explained how to use the pens and adjust the pins and the pen holders (one between the pins and one outside them). Jim was anxious to get his hands on the device and draw some curves.

D: Before you draw anything, what do you think that pen's going to draw? Do you have any guesses?
J: Well in the outer one my guess is that it would be an ellipse right away. I'm relatively sure of that because that's what the one at home looks like. It looks like it draws an ellipse....... Now looking at the inner one (the pen holder between the two pins).... I like to try to picture things..... ummm.... if you..... I'm just trying think what would happen. Now when I look at this, I think about when one of these points is going to be relatively fixed over here say...

Without having moved the device Jim then gave a very detailed description of what he thought would happen when either one of the pins was near the intersection of the two tracks. He explained how when the pin in the vertical track was near the junction, that both that pen and the other pin experienced almost no horizontal motion and vies versa. He seemed anxious to demonstrate what he meant, and so I told him to go ahead and draw a curve. With the pen between the pins, closer to the one in the horizontal track, and the pins about 15 in. apart, Jim drew a curve and explained his idea.

He then added the observation that when the vertical pin is near the junction, the pen is moving vertically "at a fraction" of the pin's motion. His sense of a geometrically determined proportion seemed very astute.
J: Say it's moving about an inch in either direction here (vertical pin near junction), this (the horizontal pin) is actually moving in and out a very very small amount, so the pen doesn't move forward and backward very much at all, while it moves up and down..... oh........ what is this? (indicates motion of vertical pin) say I have about two inches here..... It (the pen) is moving at a fraction of that distance here (gestures to indicate a shrinking proportion of vertical motion along the line of the trammel).

Jim also observed that when one of the pins was near the junction the motion of the pen was "essentially a line," either horizontal or vertical depending on which pin was near the junction. He expressed these ideas quite clearly using both hands to gesture about relative rates of motion and making "V" gestures with both hands to indicate the proportionality between the motion of the pen and a pin in the track.

Jim next wanted to verify his guess that when the pen was outside of the two pins that he would get a ellipse. He wanted to make sure that the trammel device worked in the same way as his toy, the "B.S. grinder." He drew a curve using the outside pen holder, and was satisfied that his guess was confirmed.

D: Were you surprised at all by the curve that you got with the pen on the inside?

J: A little bit. I thought it was going to draw a shape more like.....uhhh.... (begins to sketch on paper)..... I sort of expected..... or I was hoping that it would do something like this (see Figure 3.6a). I'm trying to think now if it's possible to get it to do something like that.

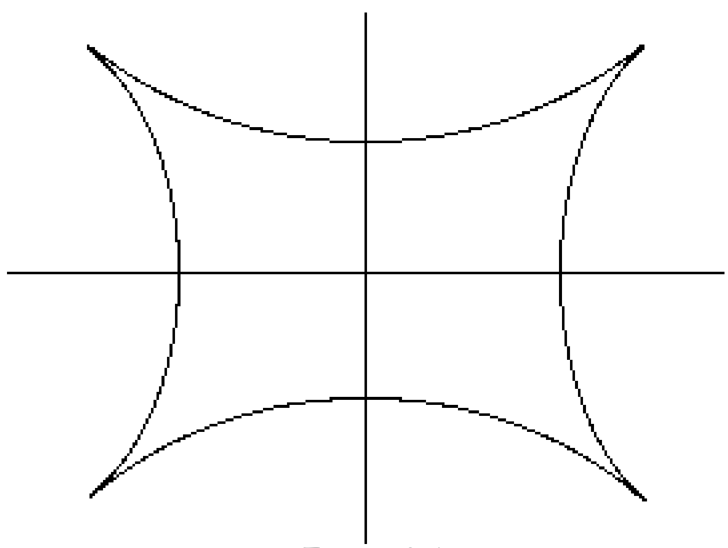


Figure 3.6a

D: Why did you think that? could you explain?

J: Well just immediately looking at this, I was trying to think what would happen when uhhh....... well you see I hadn't seen the relation that this had with that (begins moving the trammel and pointing to the motion of the inside pen with respect to the pins). I was imagining.....I just didn't think....but of course this has to stay in that same position, but I was imagining that it would come down somehow (indicates a possible cusp as in Figure 3.6a)... I don't know but I was hoping somehow that it would look like that. It's obvious now that it's not going to do that.

Jim explained that he now thought the trammel would draw elliptic shapes in almost any position. This is an excellent example of the powerful and immediate impact that physical tools can have on one's conceptions as well as a classic example of Piaget's notion of assimilation and accommodation. Jim's vision of the vertical motion of the pen when crossing the horizontal track and its horizontal motion when crossing the vertical track led to envision a cusp in the middle where the motion changed. While essentially correct, Jim's theory had to assimilate the experience of the inner pen holder having drawn an ellipse. This led him to a new theory which accommodates this experience.

D: Is there any other different shape that you think you could get by putting these (pins and pen holders) in different positions?

J: Well I would expect that I could get a circle if I were to...... bring these (the pins) very close together..... or get something that looks sort of like a circle.

Jim then placed the pins as close together as possible (about 2 in.) and put the outside pen holder near the other end of the trammel (about 18 in. away). He began to trace the motion of outside pen holder.

J: As it (the pen holder) gets further and further out it's going to look less and less like an ellipse, and more and more like a circle (draws large curve and stands up to get an overview)....... It does look a great deal more circular, although it does look a little bit longer in this direction than it is in this direction, which is what I'd expected.

D: So if we could put these (pins) closer together and the pen further out there.... that's a way to get something closer to a circle?

J: Yeah, yeah, sure it would look more circular...... I wonder if it would be more circular? It certainly would look more circular.

Jim pondered this point and decided to draw some curves by just changing the position of the outside pen holder and keeping the pins as close together as the device would allow.

J: What I wanted to see was..... I wonder if this is any more circular than this is or if it's just the way we see it (compares large and small "circular" curves by measuring the distance between the curves on the horizontal and on the vertical finding them to be the same). What I was wondering was... if this is just simply a blown-up version of this, or if we somehow changed..... the uhhh equation of this. I wonder if these equations are same, or if it's just the way we see it. For example if you were to stand two stories up.... if this (large curve) would look the same as this (small curve).

Jim thought about this question and decided that just measuring the distance between the curves was not enough to establish whether one was a "blow-up" of the other. He then measured out from the center to both curves on the vertical and again on the horizontal. He got out his calculator and then computed the ratios of these pairs of numbers. He got the ratio of vertical distance out to the curves as: $\frac{15.5 \text{ in.}}{4.25 \text{ in.}} = 3.64$, and the ratio of horizontal distances as: $\frac{17.75 \text{ in.}}{6.25 \text{ in.}} = 2.84$. He then concluded that these two ratios were not equal, and so the large curve was not "blow-up" of the smaller one. I asked him which curve was more circular, and he said that the larger one "looked more circular", and that he "wanted to have some way of supporting that observation."

J: I hope that using a ratio like that is the right way of doing that.

D: Is there any other way to get something more circular with this device?

J: I think we've tried all the different possibilities.

(The interview had now lasted 36 min.)

Jim expressed here a very clear conception of geometric similarity here, although he did not use the word "similar." Later when looking at triangles Jim used the word "similar" with precision and comfort. He was very interested in the similarity of these conics, but it seems the word "similar" is used in math classes only for rectilinear figures, and so Jim's use of the word was restricted, despite his clear conception. Apollonius wrote a whole series of propositions for determining when conic sections are similar, and I think Jim would have been fascinated by such ideas. In my teaching I have found that most secondary and college students have no idea what is meant by statements like "all parabolas are similar to each other." As his use of ratios demonstrates, Jim was thinking very clearly about the concept but had never experienced in mathematics a more general use of the word "similar."

Since Jim had said that the curves drawn by the trammel "looked like ellipses," I next asked him whether he thought he could duplicate curves drawn by one device with the other.

J: Probably...... well given this thing's limitations (trammel).... You can only get the pins about that far apart (2 in.)..... You can draw something closer to a perfect circle with that (loop of string) because you can get those (tacks) very close together.... but if this were able to slide in further I think that they would draw essentially the same things.... given any combination (indicates adjustments in the devices).

I next asked Jim to try to duplicate with the trammel any specific curve drawn with the loop of string. Jim began studying the motion of the trammel in different positions. Jim decided to use specific even numbers of inches. He used a 26 in. loop of string over two tacks 6 in. apart. He thought of this as an ellipse based on a 6 in. tack distance and what he called an "L" of 13 in. He was measuring the distance from the far tack (focus) to the end of the major axis or half the total loop of string. Because of his thoughts about circles, he equated the tack distance with the distance between the pins on the trammel; and so he began by setting the pins 6 in. apart. He then set the outside pen holder 13 in. from the far pin to match the set up of the string device. He then drew the curve.

Just by looking at the curves he decided that they were not the same. He then noticed that the trammel was drawing a curve with 13 inches between the center and the end of one axis, while the loop was drawing a curve with 13 inches between a focus and the end of the major axis. He then decided that he had to measured from the center (halfway between the tacks) which gave him 10 inches. He then reset the pen holder on the trammel at 10 in. from the far pin leaving the pins 6 in. apart. Just by looking at the motion of the newly adjusted trammel, he saw that it was still not going to draw the same curve, because it's minor axis was much smaller than the curve drawn by the loop of string. Jim did not measure the minor axes with the ruler but simply used his eyes. The semi-minor axis of the curve drawn by the loop of string was about 9.5 in. while the trammel curve's semi-minor axis was 4 in. Using the outside pen, Jim was still conceptually committed to matching the tack (focal) distance with the pin distance on the trammel, and so he next abandoned using the outside pen, without considering a readjustment of the pin distance, which could have achieved the stretch he said he needed (see below).

D: Did anything improve? Did it get any closer to that curve?

J: We got closer because the distance from this center (intersection of the tracks) to here (end of curve's major axis) should be the same, because I measured 10 inches...... but I basically need to stretch.... if I could grab a hold of it and stretch it out in this direction (indicates widening the minor axis of the trammel drawing)....... Well it looks like it's not going to happen using this outside one (pen holder) so.... I'll abandon the outside one, and use the inside one (pen holder between the pins)........ Now rather than going with something completely arbitrary I'll try ten (sets the pins 10 in. apart)..... and the six from here to here (sets the inside pen holder 6 in. from the fixed pin and draws a new curve). I don't know if I got closer. Obviously it's going to be a lot smaller, so..... but it least it less stretched out this way (gestures that his new curve is rounder than the one previously drawn with the trammel—new curve has half axes of 6 in. and 4 in.).

Jim explained that he now had a curve which appeared to his eye to be roughly the same shape as the one drawn by the loop of string, but smaller. They both appeared to him as "smooth and round." He then thought about how to increase the size of the one drawn by trammel.

D: What might you do?

J: Well I was thinking of just.... uhhh.... basically..... dilating from this point here (points to fixed pin in the end of the trammel) both these things outward (indicates pen holder and adjustable pin moving farther out the trammel).

I asked Jim what he meant by the word "dilating," and he explained by giving an example where all the distances on the trammel were doubled. He said that doubling was an "easy thing to dilate by," and that it looked like it would get him closer what he wanted to get the trammel to draw. He then set the pins 20 in. apart and the pen holder 12 in. from the fixed pin. He drew a new curve and decided just by eye that it was not the same curve (half axes of 8 in. and 12 in.). Jim then took the ruler and the trammel and began considering new possible settings for the pin and the inside pen holder. Jim began trying to use the numbers 6, 10, or 13 inches in some way to set up the trammel, because those were the distances that seemed to important when setting up the loop of string device. He considered that maybe instead of dilating he should use a "square function," which he described as doubling one distance while quadrupling another. He believed that the loop of string involved the Pythagorean theorem, which involves squares so this might make some sense.

J: You see I was trying to use numbers to be sure that I was getting close to the right thing. .... It might make sense to try something with squares but before I do that.......ummm.... I guess I'm overlooking the obvious, I can get......... I mean before, I did make something that sort of looked sort of like a circle. It was a much less eccentric looking object, simply by getting the two focal points real close together (Jim is talking about the pins on the trammel and is calling them "focal points"). So maybe by doing the same thing I can get something close. (Jim returns to considering the outside pen holder of the trammel with the two pins close together). I didn't really give this a chance.

Jim played with this idea by eye, and drew a curve that was fairly close. He told me that he thought by doing this he could get "something that was a blow-up of that one" but that he could never get it exactly because the pins on the trammel could not be moved close enough together "to get the same scale." He was indeed correct in this observation since the closest setting of the pins was 2 in. and in order to reproduce his curves he would need to have the pins 1.5 in apart. Jim did not, however, express this limitation quantitatively.

Jim decided to start again, and he erased the Plexiglas and decided to draw a new ellipse with loop of string as well. He left the tacks 6 in. apart but shortened the loop of string to 24 in. or what he called an "L" of 12 in. He wanted to have "L" equal to twice the distance between the tacks. He thought that might help him to see how to reproduce the curve. He then decided to try placing the pins 12 in. apart, and the inside pen holder 6 in. from the fixed pin to match the lengths that he saw on the loop of string. He then placed the trammel on the tracks, and before he started drawing the curve he said:

J: This is going to be six out there and six out there..... Oh!... This is obviously going to be like a circle. I should have seen this before (draws the curve and gets what he expects).

(The interview had now lasted for 1 hr. and 3 min.)

Jim then explained how the distances of the pins from the pen holder determined where the curve would cross the horizontal and vertical tracks which he now called the x and y axes. In this case those two distances were both equal, and he said that was "a characteristic of a circle."

J: This is pretty much as close as we're going to get to a perfect circle. That's my prediction.

D: Do you think that this is a perfect circle? Or as close as you can get with device?

J: Theoretically, yeah it probably is a perfect circle, because this distance here and this distance here (indicates the half axes) are supposed to be exactly the same...... It looks circular to me.

I told Jim that I could see that the curve crossed the axes at equal distances, but I asked him why the curve remained equidistant from the center at the points in between. He studied how the trammel moved through one quarter of the curve, and then went back to describing qualitatively the relative rates of horizontal and vertical motion as they varied between the axis crossing points, just as he had in the beginning of the interview. He said that these rates behaved just like the sine and cosine function, and that that was evidence for why the curve was circular. He said that at 45° the rate at which the vertical was increasing was equal to the rate at which the horizontal was decreasing. He could not be more specific. He summed it up by saying:

J: The whole reason I think it's a circle is that it's behaving like I would expect a circle to behave.

I tried to get Jim to be more specific, and he studied the trammel device some more, and then said that the whole device depended on looking at a series of right triangles that all had the same constant hypotenuse, i.e. the trammel. In this case the pins were 12 in. apart, and so he said he was looking at the Pythagorean relation $A^2 + B^2 = C^2$, where C remained constant at 12 in. and A and B were the distances of the pins from the center. He then tried to think about the relative rates of change of A and B, because they would be would be related to the x and y coordinates of the pen moving on the trammel. He thought it might have something to do with the graphs of the curves $y = 1/x$ and $y = 1/x^2$. He got out his graphing calculator and looked at the graphs of these curves and decided these would not help him.

(The interview had now lasted 1 hr. and 15 min.)

Jim next calculated some values for A and B using his Pythagorean relation. He showed me that as B increased from 10 in. to 11 in. , A would decrease from about 7 in. to about 5 in.; or roughly twice as much change as B since B was nearing the top at 12 in. I said that that was a good demonstration of his idea about the relative rates, but that in order to convince me that this trammel setting was really drawing a circle he would have to show me that the pen holder stayed 6 in. from the center at all points along the curve. Jim looked a little flustered, and then simply got the ruler and laid it on the curve and showed me empirically that all points were 6 in. from the center. Jim had a triumphant smile on his face, and I felt foolish and pedantic.

D: OK, you got me on that one.

I then asked him to return to the task of copying the curve drawn by the loop of string with the trammel. Jim first tried just moving the pen slightly off center leaving the pins 12 in. apart. He thought this was close in shape but smaller than the other curve. He played with the device some more, and he eventually saw that the distances of the pen from the pins would have to match the half axes on the other curve. He could see that the semi-major axis of the loop-drawn curve was $12 - 3 = 9$ in. He then looked for the length of the semi-minor axis on the loop drawn curve. As in the beginning, Jim pulled the string, so that if made a right triangle with the right angle at one of the tacks (instead of an isosceles one). See the dotted triangle in Figure 3.6b. He measured the leg of this triangle towards the curve as an axis getting 8 in. (instead of 8.5 in. if he had measured out from the center to the apex of the isosceles triangle).

He set up the trammel with the pen 9 in. from the fixed pin and then moved the other pin 8 in. form the pen. He drew the curve, and said that it looked about right, but that he had "no way of really knowing." He then looked back at the loop of string, and saw that he had measured the minor axis wrong, and that he should have measured from the center out to the apex of the isosceles triangle. He remeasured, and found the semi-minor axis to be 8.5 in. He made the adjustment and drew a new curve.

J: Looks reasonably close.

D: Do you have a system at this point for copying any curve over there (loop) with this thing here (trammel).

J: I should be able to.

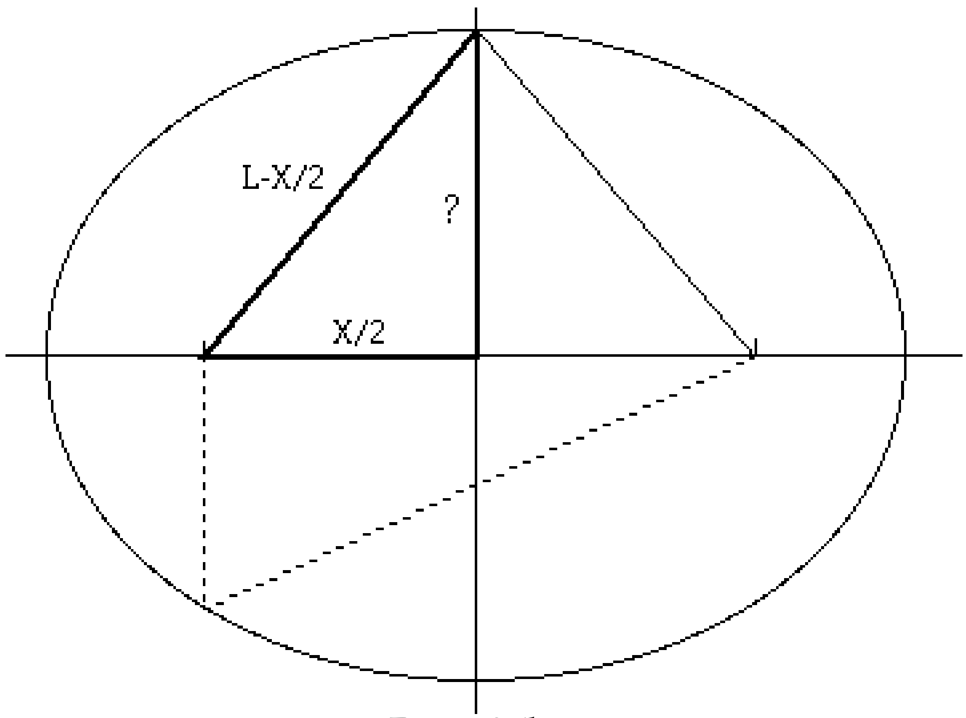


Figure 3.6b

Jim explained how he would use the midpoint between the tacks as a "center" or "origin" and that he would measure the half axes, and then set up the trammel accordingly. I then asked him if he could calculate these distances from the tack distance and the length of the loop of string. He told me that the semi-major axis was "$L-\frac{1}{2}X$" (where L=half the loop, and X=dist. between the tacks). Jim then explained that by dividing the isosceles triangle formed by the loop when the pen was at the end of the minor axis, he would get two equal right triangles each having a hypotenuse of

$L - \frac{1}{2}X$, and a leg of $\frac{1}{2}X$. See the bold triangle in Figure 3.6b. Using the Pythagorean theorem he could then find the semi-minor axis that he needed to set up the trammel.

(The interview had now lasted 1 hr. and 30 min.)

I told Jim that we would return to our discussion of these devices in the next interview. For the remainder of the time that we had (about 15 min.) I let Jim experiment with the folding arm device with equal arms. He first drew a curve which was very eccentric. Looking at a piece of this curve, he thought it might be a parabola. He then move the pen closer to the hinge and drew another curves which was much rounder. I told him that although the arms of the device collided at some point that he could lift them or flip them over to continue the curves. He then flipped the device over and continued his curves on the other side, and decided that they might both be ellipses. He then took the folding arms off the Plexiglas, and began moving them slowly in his hands. He said that the rates of motion in the device were very much same as in the trammel device. I told him that when we returned, I would ask him the same sort of questions about this device, and its possible relations to the other two. I also asked him to think about how the action of each of the devices might give rise to equations for the curves that they drew.

3.7 Jim's Second Interview

When Jim returned a week later for his second interview, he began by telling me that he had looked over some of his notes on conic sections, and that he had thought about what was "important" in the loop on string device. He put two tacks in the board, and said the distance between them was "important." He then used the loop of string to draw an ellipse and chose a point on the curve and labeled it (x, y). I asked him how he was measuring x and y.

J: Oh... Ok.... yeah...... a coordinate system..... let's see..... I think I'll have the center be..... (sketches in the axes of the ellipse)..... You know,... so we're going through here with our y and x axes..... and these (the tacks) are on the x-axis like that.

D: By "center" you mean half way between the tacks?

J: Yeah, yeah.... and I'm assuming that that's going to be the center of whatever I'm drawing. It's more or less half way between. That's what it looks like. I think it's easier that way, because then you have the symmetry to deal with rather than having one tack centered on the ahhhh..... origin.

Jim explained how he was changing from his original idea of the previous week of using one tack as the origin. He again mentioned that he was motivated by the symmetry of the ellipse, so I asked him about that.

D: What are the symmetries of the ellipse?

J: You've got the definite x and y axis symmetries (indicates his sketch) where you can reflect it over either way.... you can flip it over.

D: And those are the lines you want to use as axes?

J: Yeah.... It's also got point reflection (indicates center), but that's kind of irrelevant...... for right now at least. So starting out with a point like this (labeled (x, y)) I was trying to think how you could relate this to..... whatever it was you were drawing..... you know, given this, trying to make an equation for it..... And so what I thought about was the distance from the origin to these things here (the tacks) and the distance here (marks the distance from the tack out to the end of the major axis, see Figure 3.7a).

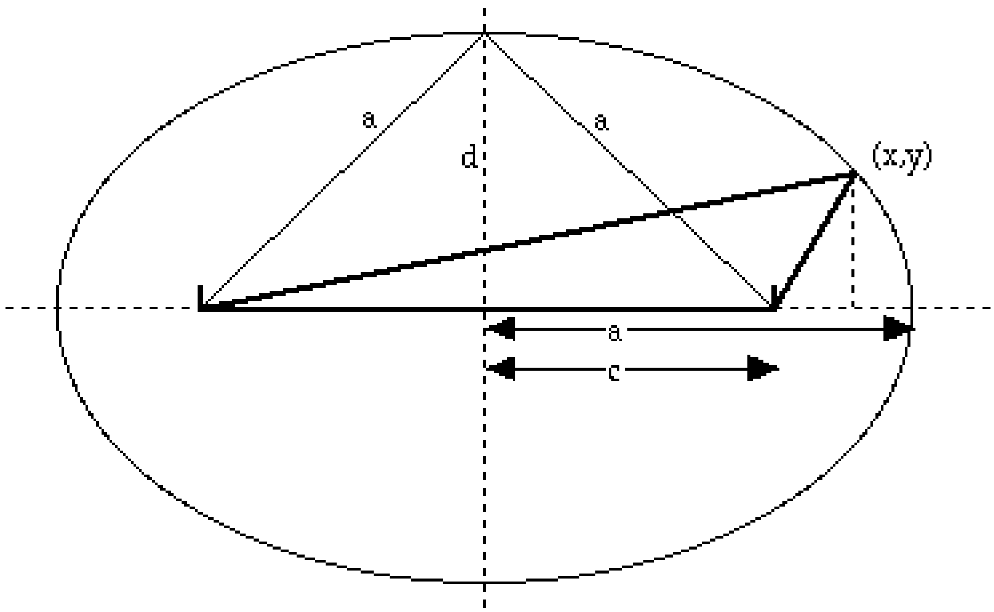


Figure 3.7a

Jim started looking at the x-coordinate of his labeled point. He drew in the perpendicular from the point to the major axis. He looked at the figure, and then decided that he wanted to change his previous marked constant to be the "total distance" from the center out to the end of the major axis. He decided to call this "a" and then he labeled the distance from the center to the tacks as "c." He then began to label the x-coordinate of his point as "b."

D: Where are x and y in the picture? Can you show me geometrically?

Jim showed me where they were, and then got rid of "b" since he saw that it was "better to call that x."

D: What do you want to do with this?

Jim thought for a while and then consulted his notes. He said that he "had this all thought out before," but that now he had forgotten some of it.

J: I decided last time that the basic governing principle...... the basic things that you needed to have.... to say, write an equation for this were the length of the loop.... this here (pulls string to lay along the x-axis), and the distance between the two focal points (indicates tacks). I was trying to remember how to write an equation using x and y to create any.... uhh..... ellipse (gestures along a piece of the curve).

I asked Jim about the relationship between the "a" and "c" that he had labeled in his picture, and those "two basic things" used to draw the curve i.e. the loop's length and the tack distance. He gestured to show me that c was half the distance between the tacks. He then put the string back on his general labeled point on the curve, and told me that he was "guessing" that the two string lengths that connected that point to the tacks "added together would equal $2a$." Jim kept moving the string back and forth between his labeled point and the end of the major axis where the string triangle collapsed onto the x-axis. I asked him to explain that to me. He said the reason that he thought that it was always equal to $2a$ is that when the string all lay along the x-axis he imagined grabbing the string at the nearest tack and sliding it along like a conveyer belt until the he had moved the that point to the center. This would amount to a slide of length c. When this was done the point on the string that started at the other tack would also have moved to the center and the piece of string which was the sum of the distances from the tacks to the curve would now go from the center to end of the major axis and back again and thus would equal to $2a$. Jim expressed this physically by actually marking the string and demonstrating the sliding motion. It was both simple and convincing and avoided the algebraic subtraction that is usually used to show this (i.e. $(a - c) + (a + c) = 2a$).

Jim said that he "didn't need to worry" about the piece of string between the tacks (of length $2c$) because the important part was the two pieces that went out from the tacks that always added up to $2a$.

J: Now where am I going with this.... uhhh... (consults notes and then returns the string to his labeled point).... OK, what I trying to do is to get the distance of this one here and this one here (indicates the two string lengths between his point and the tacks), and the easiest way to do that is to use right triangles.

Jim then set up right triangles with these two string lengths as hypotenuses, and calculated their distances using the Pythagorean theorem (see Figure 3.7a). He clearly indicated the bases of the these triangles as $x+c$ and $x-c$, and their heights as y and then wrote:

$$\sqrt{(x+c)^2+y^2}+\sqrt{(x-c)^2+y^2} = 2a.$$

I invite the reader compare Jim's explanation of this equation with Tom's (Sec. 3.4). Although Tom eventually remembered this equation during his first interview, without reference to any notes, he explained it as an application of the distance formula and did not refer to triangles. Tom knew from looking at the string that the sum of these two distances was constant, but he was not quite sure whether it should equal $2a$ or $4a$. Tom said that the choice had to be made so as to make the "algebra come out nicely." Jim's sliding string argument was based on his own physical experience with the device and expressed a clear geometrical reason for the constant $2a$.

J: That (the equation) is a really nasty, but meaningful expression for this creation here, whatever we have.....uhhh.... given that.... uhh..... what we said was that..... since the string is determining where it (the point on the curve) is going to be, and this (the equation) is telling you the lengths of the string. You're given these points here (tacks). It (the equation) relates all of the pertinent information together..... into a nasty equation which .... you know.... given higher algebraic skills, I'm sure I could simplify, but I don't really want to.... unless you're asking me to...... I know what it is.....

D: You've been through it before?

J: Yeah, it should be..(writes: $\frac{x^2}{a^2}+\frac{y^2}{a^2-c^2}=1$). I saw my teacher do it.

(The interview had now lasted 18 min.)

Jim said that he felt that he might be able to show that the two equations were algebraically equivalent, but that his algebra skills were not good, and that he was "prone to making mistakes when it comes to following rules." I told Jim that his explanation of the first equation "made perfect sense to me," and that we would just assume for now that the two equations were equivalent without going through the derivation. He said that that "pretty much nailed down" the loop of string device, and so he turned his attention to the trammel device.

J: Since these two are equivalent devices, as far of being able to draw the same objects......

D: You think they are?

J: I think these are equivalent devices. Experimentally we've been able to draw the same things..... and..... given a distance here (loop) I can sort of relate..... and get a similar looking object over here on this thing (trammel).

Jim then reviewed his method from the previous interview for copying the curves from the loop of string device with the trammel device. He now chose to call the semi-minor axis "d" and showed me, by looking at an isosceles triangle formed by the loop of string, that $d^2+c^2=a^2$ (from half of the isosceles triangle in Figure 3.7a). He then showed me that using a and d, he could set up the trammel to copy a curve. As an example, he copied his string drawn ellipse with the trammel.

I next asked him if he could go the other way, and copy with the loop of string, a curve first drawn by the trammel. Jim was instantly sure that he could, and set about showing me. He changed the trammel setting arbitrarily and drew a new curve with the pen between the pins. He took the trammel and used it like a compass to mark off the semi-major axis, a, on the string board from a center. Jim then thought about where to place the tacks.

J: Well I could work backwards (he marked the half axes on the trammel curve as a and d). (long pause as Jim studies his figures, and marks the length d from the trammel onto the board). Well I could just do it by trial and error but I'd rather not....... I know that it's got to max-out up at this point here (indicates top of minor axis)..... So since these are going to be an equal distance out (tacks from center)...... hmmm

D: So what's the piece of information that you need here?

J: I need to know c.

D: OK, you know a and you know d, and you've got to find c?

J: ummmmmm.

Jim then measured a and d precisely in inches from his trammel curve as $a = 8.25$ in. and $d = 3.25$ in. He then took his calculator and began figuring.

D: What are you doing?

J: I'm using this equation right here to find c (indicates: $d^2 + c^2 = a^2$)....... I got an answer of 7.58 in.

Jim then used the ruler to position the two tacks on a horizontal line each 7.58 in. from a marked center. He then adjusted the size of the loop of string, so that when it was placed over the tacks it reached out to a point on the horizontal axis 8.25 in. from the center. The string loop then went less than an inch past the tacks on the x-axis. This looked a little tight to Jim.

J: I don't know if it's going to make it. We'll find out..... But.... I mean it's close (examines the loops action with his finger). I don't know if its a problem with my logic, or if it's something with just the mechanical limitations of the measurements and stuff like that.

D: Well draw it and see how close it looks. Give it a try. (Jim draws the curve with some difficulty because of string being so tight on the tacks.) Does it look reasonably close?

J: It looks reasonably close, yeah. I think...... I mean given the inaccuracies of the measurements. Hopefully.... hopefully it's not just a coincidence. I don't think it is.... because unless I'm not seeing something, the logic follows that it would be the same.

I told Jim that his logic convinced me that the lengths of the axes on both curves would be the same, and that this would guarantee that the four points on the axes would match up. I asked him how he could be sure that the other points along the curve were really the same "since the actions that produced the curves were different." I pointed out that he had showed me in detail how any point on the loop of string device always had the sum of its distances to the tacks equal to $2a$. I asked Jim if he could give any kind of argument; physical, geometric or algebraic; to show that the points on the two curves were "really the same or perhaps different." Jim began readjusting the trammel so that the pen was close to the midpoint between the pins.

J: Yeah, I was trying to think about something sort of along the same lines. You asked me last time how I could know that something was a circle (points at trammel). I drew this.... (indicates the first trammel drawn curve); it looks semi circular. And I was trying to think of some argument for that, and I found myself having a difficult time (studies the motion of the trammel along a large close-to-circular curve).

D: Last time you found a way to draw a circle. Do you remember what that was?

J: Yeah, I just have these two of equal lengths (distances from pen to pins). (Jim uses the ruler to get the two distances the same at 6 in. each. He then draws a new curve which appears circular.) Now the last time that I was working on this I remember trying to talk about the sine and the cosine and the unit circle..... things like that...... trying to show the way things were increasing and decreasing at varying rates. And I thought about this for a while, and I couldn't really think of a really conclusive argument to show that it was a circle beyond saying that it's got the same radius here and here (indicates where the curve crosses the tracks). And sort of looking at it and saying well it's not doing anything special like moving back and forth so there's sort of a fixed ratio between here and here (moves the trammel through on quarter of the curve), and so I'm guessing that it's going to be sort of a constant relation along there, but ahh.... beyond that I was trying to think of how I could make a good convincing argument...... Well when we transferred from here to here (trammel to loop device) we said that distance from the first pin to the pen was equivalent to a. Right?

D: unnn hun.

(The interview had now lasted 40 min.)

J: And over here when we were drawing this thing (with the loop) we said that $2a$ was constant throughout. The length of the string being $2a$ doesn't change (moves loop of string to demonstrate). So if this distance here is "a" (on trammel), and this distance here is $2a$ (length between the pins on the trammel).... I just trying to get some corresponding pieces from these two different apparatuses. Because they're both doing the same thing in the end, and they have the same sort of measurements, so I might as well call it the same thing. I going to draw myself a little diagram.

Jim traced a picture of the trammel and labeled the length from one pin to the pen as "a." Since he had the trammel set up to draw a circle he then labeled the distance from the pen to the other pin also with an "a." I then asked him to review again how in general he transferred curves from the trammel to the loop. He then labeled the second distance on the trammel as "d," and said that in this special case $a = d$. My question was somewhat leading for Jim because although in a physical sense he had discovered quite clearly how to set up a and d on the trammel, he did not always label things consistently.

J: Although a and d, in this case, are equal, I should call one "a" and one "d" for the purposes of keeping my mind straight..... because again maybe.... uhh.... I wouldn't have thought of that..... Now over here (looks at loop of string)..... I don't know how to put all this together. I'm seeing a lot of different things here. When you label them appropriately things start to correspond. Obviously when d becomes equal to a over here (trammel diagram).... this right triangle (on loop, see Figure 3.7a) goes down to nothing. The two lines (strings) overlap over each other and the two focal points (tacks) have to come together to a point, and that's when you get a circle in this case. Then again over here (trammel), when you're getting a circle is when a and d are equal to each other.

Jim explained in detail how the right triangle in his loop picture with sides a, d and c would "collapse" if $a=d$ forcing $c=0$, which is how he first thought of a circle, i.e. as a curve drawn with the loop over one tack. Jim then took the trammel and left "a" the same and made "d" larger and drew another curve which touched tangent to his circle on the x-axis. He looked unhappy.

J: I should have made a longer than d for the purposes of keeping everything the same, because now my focal points are on the y-axis (points to where he thinks the foci would be on the Plexiglas, and then erases the new curve keeping the circle). What I want to do is find c on this somewhere, because if I can do that it will show me where the focal points are..... It's hard to do that with a circle because there is no c.

Jim readjusted the trammel so that half the vertical axis, d, was equal to the radius of his circle but half the horizontal axis, a, was "a little bit bigger." He then drew a new curve with the trammel that touched tangent to his circle at the two points on the y-axis.

J: Now I want to find c on this..... and since c is the base of the triangle formed by d and a...... and since I said that a is this distance here (Jim lays the trammel on the horizontal axis so that half the axis of the curve and the length on the trammel match up)...... if I want this like this, then I get the point there... which should be one of the focal points (Jim takes the trammel and uses it as a compass to draw an arc of radius a from the top of his curve's minor axis intersecting the horizontal axis at his proposed focal points. He then traces the a, d, c triangle on both sides, see Figure 3.7b).

I was amazed at his ingenious use of the trammel as a compass to find the focal points in this simple and exact physical way. The tool worked quite well as a compass, since the pen distance was already set, and all he had to do was hold the pin fixed in the track at one point and rotate the trammel stick. After Jim marked the foci on the Plexiglas, I offered to hold the tacks at these points while Jim traced over the trammel drawn curve using the loop of string. The tracing seemed very accurate to Jim (and to me).

(The interview had now lasted 50 min.)

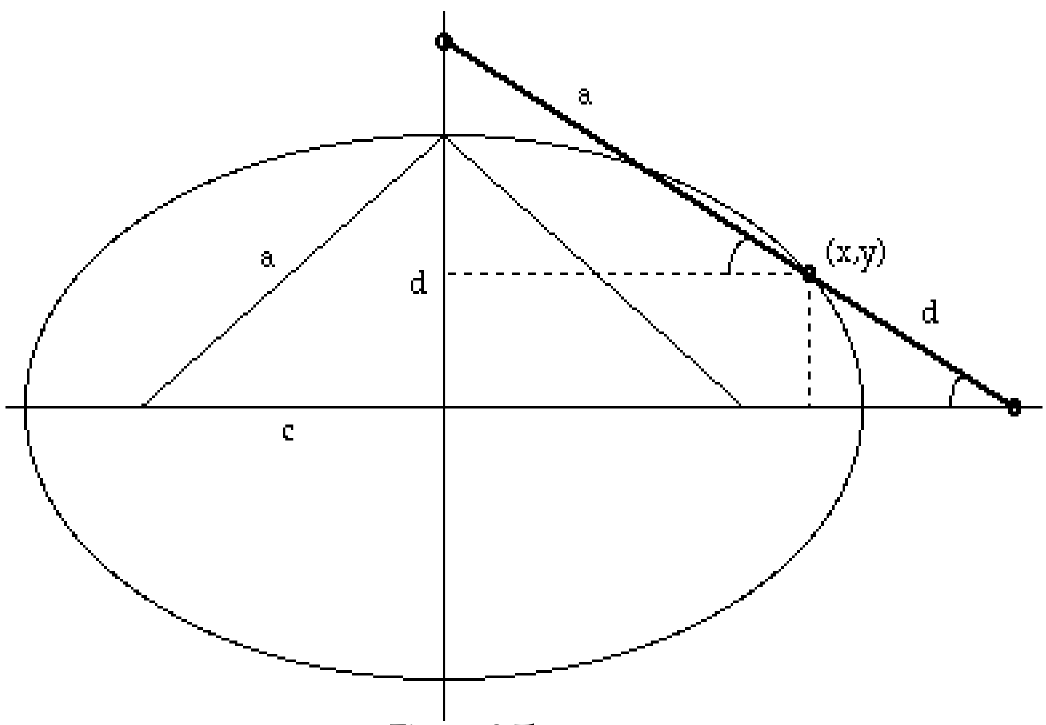


Figure 3.7b

D: Is there any other kind of argument that could really nail this down?

J: Well I'm guessing that the equation is going to be the same for both, since we have the equivalent pieces.

D: Is there a way to get an equation out of this device (the trammel) that talks about the geometry of this device?

J: I'll do the same thing that I did before. I'll take a point here, (x, y) ... (labels an arbitrary point on the curve in the first quadrant, see Figure 3.7b)..... And then look at what this thing doing while it's drawing that point there (places the trammel so that the pen is on his marked point). I said before that the important piece of information with this machine was the distance from here to here (pin to pin), and then these distances here (pen to each pin) call them d and a (labels Plexiglas as in Figure 3.7b).... hmmm..... Then I started looking at the right triangles....... Now what describes that action?.... hmmm... let's see..... given a certain x.... (draws a vertical line from his point and marks the x-coordinate on the horizontal axis)......... (long pause as Jim studies both the trammel and the loop of string figure. He then draws a horizontal line from his point to the vertical axis, dotted in Figure 3.7b)..... What I'm doing is I'm looking at these two similar triangles (points to small upper and lower right triangles with hypotenuses a and d on the trammel, see Figure 3.7b)... I think that they're similar.

D: Why do you think they're similar?

J: Well they share...ummm..... (Moves the trammel to watch its action. Seems to be checking to see whether what he is about to say is invariant along the curve.)..... first of all they share the same angle (marks angles, see Figure 3.7b)..... Since it's (base of upper triangle) parallel to the x-axis .... and their hypotenuses are in a constant ratio (points to a and d on the trammel)...... And I'm guessing that these sides here, and these sides here (other pairs of sides in the upper and lower triangles).... are also in constant ratios.... that's what you mean by similar..... How to make that clear?.... unnnn..... (long pause)

D: Could you write down some of these ratios you're talking about so I could see an example?

J: As far as numbers.... numerical....

D: Any way.... just so we have it written down.... which things are... uhh

J: Well I'll have to label these so I can talk about them (indicates similar triangles)..... well I said that this was d and this was a (hypotenuses)..... and let's see........ information.........

D: You're measuring your coordinates from where?

J: I meant this point here (labeled point (x, y)).

D: That's the point on the curve, but in terms of lengths, just to be real clear, what's x and what's y?

J: I meant this distance on the x-axis and this distance on the y-axis.

D: So from the center here that's x and that's y. OK.

Jim then decided to clean up the figure by erasing the a, d, c triangle that he had drawn to show the position of the curve's foci. He said that those lines were "distracting him." Originally he was trying to copy the method of generating an equation that he had done previously with the loop of string, and so he thought it was important to know the foci, but now that he was looking at the similar triangles he found this focal triangle distracting. I then asked Jim to review what he had told me about the similar triangles that he had mentioned. He pointed out again the pairs of sides that he thought "would be in constant ratio."

J: I'm forgetting my geometry here, but is Side, Angle, Angle enough for triangle similarity?.... I can't remember... (pause)

D: I'll believe that those triangles are similar.

J: I'm trying to convince myself.....

D: Well.... you've got right angles.

J: yeah

D: And then you told me that these two angles are equal (marked in Figure 3.7b).

J: Should be. Right.

D: Now if they've got two angles the same, what about the third angle?

J: Of course it's going to be the same. Right. That is a similar triangle. Angle, angle, angle.

D: So, I believe your statement about the constant ratio. I'm just wondering what that has to do with the curve?

This interchange shows the disparity between Jim's confidence in his own precise and accurate observations, and his confidence in his ability of apply rules learned in mathematics classes. Even when it comes to geometry, there is a gap. As we shall see later this gap is much greater for Jim, when it comes to algebraic thinking.

Jim began pacing around and looking at the figure and the curve from various perspectives. He took off his glasses and appeared deep in thought.

(The interview had now lasted 1 hr. and 3 min.)

J: Often when I'm looking at something I like to move around..... Sometimes I'm looking at something for such a long time that I kind of forget about...... you know... I miss something obvious..... (long break. I get Jim a coke and he paces around thinking)...... Yeah, my problem is that I'm getting stuck in the same.... uhh.... because it worked so nicely I think with that setup (loop of string), and I'm trying to think about what the... uhhh.... (pause).

D: Well what do these similar triangles say? You were telling me something's in a to d?

Jim pointed at the lengths in the triangles that he knew were proportional, but he floundered around when it came to giving any of these sides names other that "this distance," or "the base of that triangle." He had now expressed several times with gestures and pointing the proportions in the triangles, but he would not label or name any of the sides other than the hypotenuses a and d. He tried to express the base of the lower triangle as "something minus x." He had previously showed me where x and y were in the picture so I asked him a review question.

D: Are x and y the sides of any of these triangles?

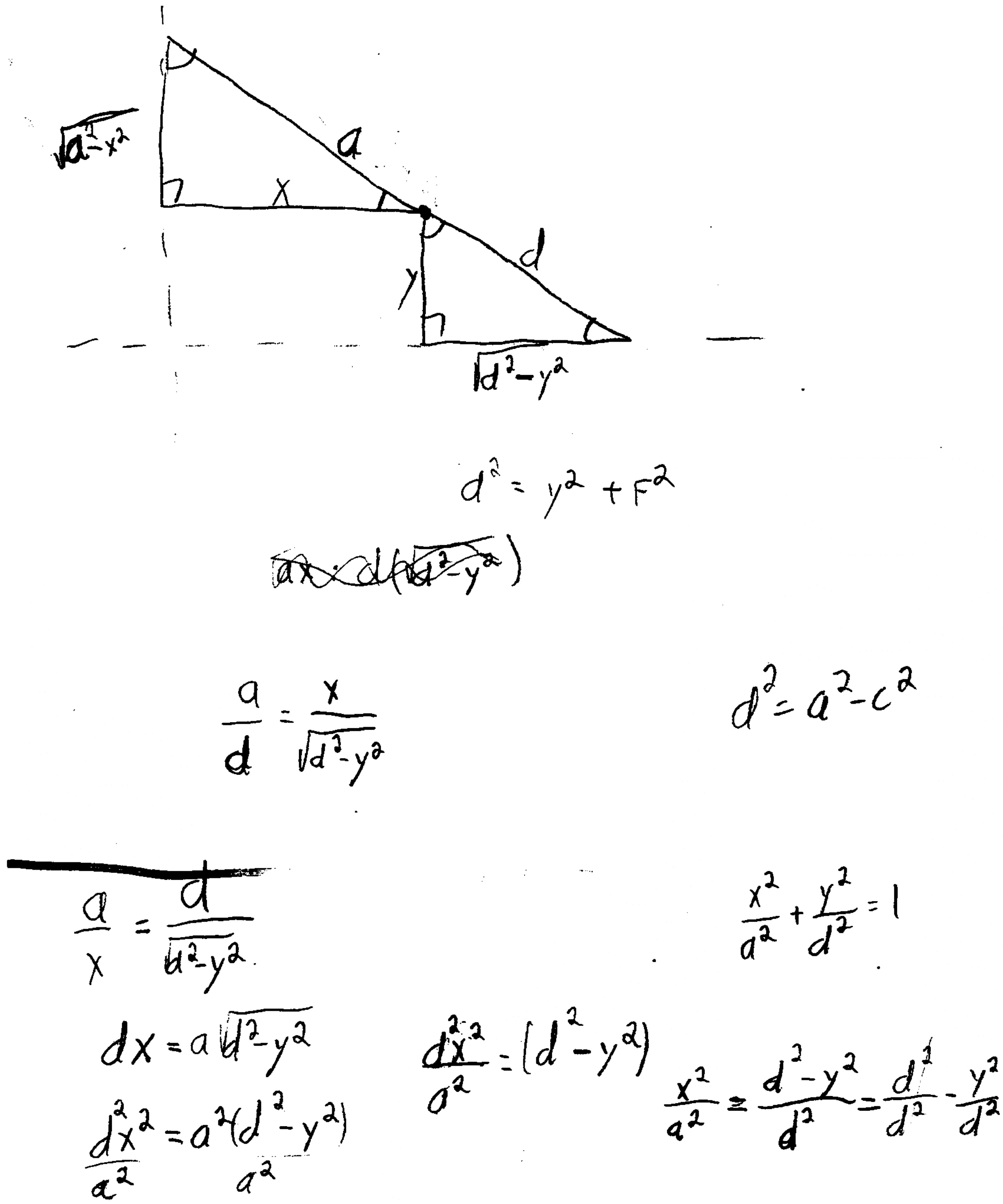
J: x is the base of this one here (the upper triangle). y will be the side of that one (the lower triangle).

D: Can we write down anything using what we know?

Jim took a sheet of paper and made a copy of his figure and labels a, d, x, and y (see Figure 3.7c). He then said that he wanted to "declare something new." At first he said he might want to give a name to the distance of the horizontal pin from the center, so that he could then subtract x from it and get a name for the base of the lower triangle. He never declared such a name, but he told me what he wanted to do.

J: Yeah I'm finding the best way to express that length.... and then once I get that I can express this here (height of upper triangle), and this here (base of lower triangle).... these lengths in these triangles so I can get these triangles all pinned down. I need to get names for all the sides.

I encouraged Jim to work on this, and I specifically encouraged him to introduce a new variable if he needed one. I said "why don't you give some of these things names, and maybe we'll find out what they are later," but Jim was very hesitant to add any new algebraic variables to his picture even though he became a little flustered using "this length" and "that distance" all the time. Jim was physically convinced that all you needed to know to set up the device and draw a curve were a and d, so he wanted to get an equation using only what he saw as relevant. Unlike Tom, Jim was extremely uncomfortable with the idea of introducing any intermediate or superfluous variables. Algebraic convenience did not suit Jim's purposes since he had no faith in his algebraic abilities. He finally turned to the lower triangle in Figure 3.7c.



a
x
d
y
d² = y² + F²
a/d = x/√(d²−y²)
d² = a² − c²
a/x = d/√(d²−y²)
x²/a² + y²/d² = 1
dx = a√(d²−y²)
d²x²/a² = (d² − y²)
x²/a² = (d² − y²)/d² = d²/d² − y²/d²

Figure 3.7c

J: Well I can express these sides using.... just saying.... since d^2 equals y^2 plus ...... whatever..... plus..... uhhh..... I don't really know what to call it....
F maybe?... uhhh ....squared.. (indicates base of lower triangle).

D: OK so call it F, the base the of the little triangle?

J: Yeah...... oh...... it's got to be $\sqrt{d^2 - y^2}$.

At this point Jim was off and running. He had immediately eliminated the variable, F, and it was never mentioned again. It only appeared once on his work sheet (Figure 3.7c). Right away Jim saw that he could use the Pythagorean theorem on the upper triangle to find it's height as $\sqrt{a^2 - x^2}$, without introducing another variable. He told me again that things were in constant proportion, and so I asked him what things. He told me that a and d were in the same proportion as "x and that expression there" (i.e. $\sqrt{d^2 - y^2}$). I asked him to write it down.

J: You see I don't really know how to express the ratio (long pause).

D: Well how do you usually write ratios? Do you have any notation or way of writing ratios?

J: Well you could say something is in a one to two ratio like (writes 1:2)

D: You like to write them with colons?

J: Well I mean you could say something is in an "a times x" to "d times $\sqrt{d^2 - y^2}$" (writes: $ax : d\sqrt{d^2 - y^2}$, scratched out in Figure 3.7c).... You see I don't know if I'm going in the right direction here with the ratios. Sure they're similar but...........

D: Don't ratios give equations in some way, shape, or form?

J: Yeah....... It's a lot easier to say that a is to x, as d is to $\sqrt{d^2 - y^2}$.
(Jim writes: $\frac{a}{d} = \frac{x}{\sqrt{d^2 - y^2}}$, see Figure 3.7c)

D: Is that an equation for this curve?

J: I don't know.

D: Looks like an equation.

J: It's an equation that's for sure (laughs)....... but what's it saying...... It's giving you.....uhhhh...... I don't see why not. I mean it's giving you this distance x and y, given an "a" and a "d" which we can get from those things (points at trammel).

D: OK, so it's an equation that talks about this curve. Is it the same equation that we got over there with that string device? Is this equation equivalent to those two over there, or is it different?

Jim looked very glum at the thought of having to do any algebra.

J: It's got the same look to it as far as the ratios go.... things like that .... you know.... the relation of the ..... but the thing is that there are no squares besides down here (indicates y^2 under the radical but no square on x, or a). Whereas on the other side over there (indicates: $\frac{x^2}{a^2}+\frac{y^2}{a^2-c^2}=1$, from loop of string) there are no square roots there are just squared numbers.

D: Well play around with it; maybe it's different?

Jim was hesitant to believe that this could be an equation for this curve, because it looked very different from the reduced elliptic equation that he knew, and because it had been too easy to obtain (he says this later). He was expecting some complicated use of the distance formula as he had seen in class for the loop of string. Using similarity seemed too easy to him. Jim was also very hesitant to perform any kind of algebraic manipulation. He said he was very "bad at algebra," and the thought of having to do it made him very anxious. He muttered to himself with a foreboding tone "here come the rules." He stared at his new equation for while trying to think what to do.

Jim's every algebraic move was made with trepidation. He repeatedly asked for help. Before each step he would ask me "is it equivalent to say.....?" or "is it legal to.....?" He tended to get lost in his notation for several reasons. He liked using the eraser to make changes an algebraic expression, rather than writing a new modified equation.

When simplifying an expression, he would also tend to run the writing together without putting in an equal sign (e.g. in the lower right hand corner of Figure 3.7c, he did not at first have the second "="). I cautioned Jim against such practices. In spite of all this, he did not make any gross algebraic errors, although he did not know what to do with the radical in his equation. When he asked me I suggested that he would have to "square both sides of his equation."

J: I need to work on my rules. I need to get down and do some of this stuff. But I just hate doing it so much that I have neglected it.

Jim did know what he was trying to do. Once the radical was gone he immediately tried to obtain the term x^2/a^2, because it appeared in the other equation. Once he had that, he continued on and was pleased when the equation eventually appeared as: $\frac{x^2}{a^2}+\frac{y^2}{d^2}=1$. He looked over at the loop of string equation (i.e. $\frac{x^2}{a^2}+\frac{y^2}{a^2-c^2}=1$) and smiled. I asked Jim about the difference between the two equations, and he knew that d^2 and a^2-c^2 were the same. That, after all, was the geometric relation which he had demonstrated so well when he used the trammel as a compass.

J: I'm happy.

D: Does this convince you that the curves are same?

J: (with resignation) Well, if they have the same equation, I guess I <u>should</u> be convinced.

D: But equations, deep down, don't seem to convince you very much. Is that what you're trying to tell me? Do I detect a skeptical note?

J: No, I'm happy. I mean seeing the equation the same makes me happy, but I was more convinced the first time I saw the similar..... uhh..... graph... or drawing... or whatever you want to call it....... Well I can't say I was more convinced..... I was quite certain...... I mean I took a large step when I saw the relationship between the drawing tool we had here (trammel) and the string over here, and getting those two to draw the same thing; I immediately thought OK there doing the same operation. They're making the same kind of picture. Therefore they're doing the same thing. They're operating in the same way. And they probably do have a similar equation. And getting that equation to work out, you know, confirms it..... but it's not like it's a great shock........ It's something I already knew..... You know, I kind of assumed that it was like that.

D: So the physical experience was really a more convincing experience to you than an algebraic experience?

J: Well, not to belittle the power of the algebra to show you, without a doubt that it's like that, but I mean I was relatively certain........ If you can look at my steps of certainty..... I took a large step from here to here (gestures about one foot on the table) when I first saw it drawn out and I could get it to do the same thing.... and from here to here (gestures about 2 in.) when I saw it (the algebra)...well yeah OK...... This being my total amount of certainty.

D: (laughing) I see... if you had to put it on a one to ten scale? I'm looking at a ratio on your fingers there of ahhh...... looks like maybe...

J: Eight to two.

D: Eighty percent confident with the experimentation, and the algebra give you another twenty percent on top of that? Something like that?

J: (laughs and nods) yeah..... Some people really like the algebra..... I need to get more familiar with it...... But it's.... (shrugs).

D: Well, just looking at this algebra..... We arrived at this equation (Jim's trammel equation), and here you worked it all out.

J: Right.

D: Over there (two loop of string equations) we skipped some big horrible step that you said is in some book, or that you saw your teacher do. If you had to derive an equation of an ellipse which method would you rather do?

J: I'd definitely rather do that (his trammel equation).

D: You like the similar triangles better?

J: Yeah

D: That was all distance formula, although we had some Pythagoration in here too.

J: Actually it surprised me that I was able to get it so easily. I thought I was going to have to go with something like finding this distance here and this distance here (indicates the distances of the trammel pins from the center), and then subtracting the x's and getting an idea of what x was. But it worked out nicely. I guess doing it with similar triangles was a good idea, I mean it looked right.

D: That was what jumped out to your eye: these two triangles?

J: Yeah, I mean I saw them. When I approach something I try to draw in everything that I can, so I can get an overall sense of what it's going to look like, and then look at each piece of it with the greatest amount of ...uhhh ... greatest degree of.... uhhh.... I want to have all the detail, including it. So then I can look at the overall thing, and then look at each piece and how it relates to the overall drawing, rather than getting caught up in algebra (voice drops). Algebra for me, it helps to make something certain and to give it a great deal of shape, but the actual thought of how something's going to work out happens in the geometry.

D: I see. Geometry is somehow more deeply convincing to you?

J: (nods) Also much easier to understand the way that things interact with each other. Watching this piece move along like this (moves trammel along curve), and watching this decrease as this decreases (indicates two horizontal distances, one from the pen to the vertical-axis and the other from the horizontally moving pin to the center), I can see that those are in a fixed ratio from watching this thing move.

D: Which of these devices do you most enjoy drawing with?

J: The string is more convenient. There's less to worry about physically speaking. On a basic level the tacks hold the string nicely while this (trammel) has to slide through slots and things like that. But it's also.... it's kind of mystical, the way this slides around and draws it like that (makes trammel action gesture), whereas with the string you can definitely see, because there's definitely something holding back the pen. Moving around, you can see the thing moving around in a prescribed ellipse. Whereas with this (the trammel), you're not directly controlling where this thing (the pen) is; you're controlling where it's sliding, and you sort of watch it moving around (shows that when he draws with the trammel his hand is on the pins rather than on the pen). I think it's initially a little bit more difficult to understand, but it's more interesting...... As I said the first time that I saw it I expected it to do something completely different. I expected it to make a sort of a star, you know something with points (see Figure 3.6a). It's definitely not as intuitive as the string and the tacks. Not just because I'd seen the string and the tacks work before, but because you can definitely see how it's limiting the distance the pen is going to go.

(The interview had now lasted 1 hr. and 33 min.)

After Jim's long hard work and algebraic frustrations I was not sure whether either of us had the energy to fully discuss the folding arm device. Jim's philosophizing, however, seemed to lift his spirits and so I brought out the device, and asked him if he wanted to experiment with it. This return to a physical and geometric setting enlivened Jim even more. He stayed for more than another hour and worked with the device until he fully understood its action. As I shall describe, he found similarity relations in it's action, and eventually used these to write an equation of the curves that it drew. Reducing this equation algebraically to his previous elliptic equation was more difficult than for the trammel device, and Jim felt lost and frustrated, although he very much wanted to see an algebraic confirmation of his experiments. He pleaded for advice and I eventually did come to his aid in a purely formal algebraic sense, offering him a series of strategic hints, which enabled him to simplify the equation. Jim was very satisfied to see his geometric thoughts confirmed.

My description thus far gives a good picture of Jim's conceptions and procedures. His work with the folding arm device was of the same nature, and so I will not give such a fully detailed account of the remaining 1 hr. and 10 min. of this interview. I will instead summarize the main points in the order of Jim's thinking. This summary will be important, however, when comparing Jim's conceptions with those of Tom. Jim will eventually draw the same figure that Tom drew, and work with same pair of triangles, but his approach and his conception of motion and especially his use of algebraic variables are all very different from Tom's.

Jim began by reconsidering one of the curves that he drawn the previous week with the pen set fairly close to the sliding pin. This curve was very eccentric and looking at only one half of the curve Jim thought that it strongly resembled a parabola. Jim began thinking about how one could empirically distinguish half of a highly elongated ellipse from a parabola. He told me about the focal properties of the parabola, and its uses in such things as solar heaters and car headlights. He then considered the focal properties of an ellipse in terms of bouncing light, and thought about how the light would converge to the other focus with an ellipse, and how this might provide an experimental way to distinguish ellipses from parabolas. He described an experiment that he had seen with marbles bouncing inside of an elliptic tray.

Jim then discussed how he would set up the first two devices to copy curves drawn with the folding arm device and vise versa. He looked immediately at where the curves crossed the tracks, and saw in the folding arm device how he could control these lengths (a and d). He then said that these could be immediately transferred to the trammel, and from these he could calculate the focal distance (c) as before for setting up the string device. Jim was very confident about getting at least the lengths of the axes to match up.

Since Jim had explicitly stated his visual preferences in approaching these problems, I next asked him what he "saw in the action of the folding arm device." He studied the motion of the arms along a curve that he had drawn. He then marked a point on the curve in the middle of the first quadrant and drew a vertical line from it to the horizontal axis. Jim first saw the two arms making an isosceles triangle with the horizontal track. He then drew in a vertical from the hinge (see Figure 3.7d), and said that this divided that triangle into "two congruent pieces by Side Angle Side."[38] He then tried to see how the x-coordinate of the pen related to this.

Jim looked at the device in silence for several minutes moving it slowly. He thought of several triangles that he might consider, and decided to first label the base and height of the large isosceles triangle that he first saw formed by the folding arms. Figure 3.7d shows what Jim drew on the Plexiglas. He let M be its height (i.e. y-coordinate of the hinge), and N be its base (i.e. x-coordinate of the pin). He then said that when the arm was fully extended ($M = 0$), and the pen was crossing the x-axis, that N would be equal to $a + d$ (the sum of the half axes of the curve).

[38] I did not challenge Jim on his reasoning here. Tom gave the same reason at this point, and I accepted it with him as well. Looking more carefully at the situation, one sees that "Side, Angle, Side" is actually not the appropriate reason for the congruence, but rather "Side, Side, Angle," where the angle is a right angle. This congruence is sometimes more accurately referred to as "Right, Leg, Hypotenuse."

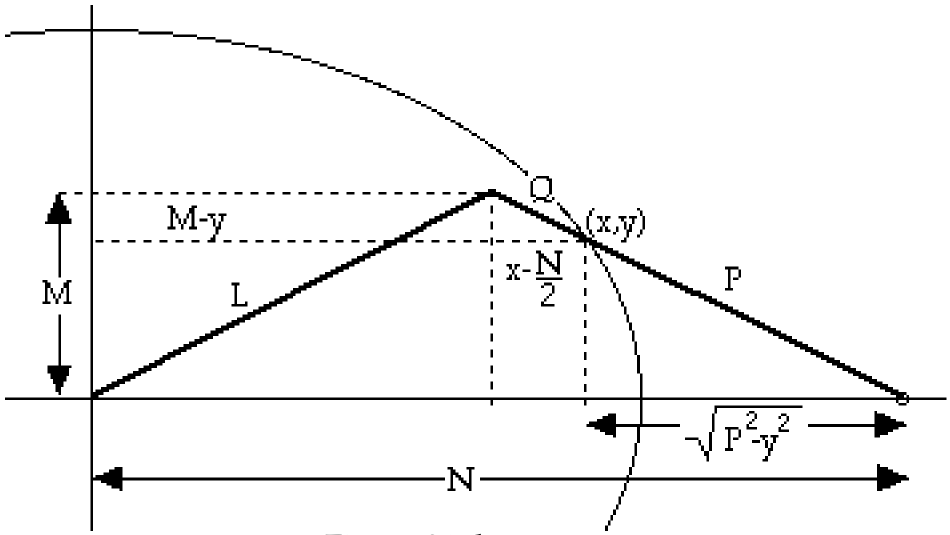


Figure 3.7d

J: I want get a hold of what his thing is doing. I mean, I can see what it's doing, but as far as mapping it out...... hmmmm.... Well this distance, y, (indicates y-coordinate of the marked point on the curve)..... is dependent on..... (draws in the horizontal line from the point to the altitude of the isosceles triangle formed by the arms)..... I think I may again do similar triangles..... I'm saying that this y distance can be found..... if you're given an N ....and you're given an x?.... well $x - \frac{1}{2}N$ would be this distance here (base of small upper triangle in Figure 3.7d). This is $M - y$ (side of the same triangle). M can be expressed on a more basic level. I think all you need to know are N.... at any time.... and this distance here (length of the arms)..... Which we'll call L.

(The interview had now lasted for 2 hr.)

Jim then wrote that: $\left(\frac{1}{2}N\right)^2 + M^2 = L^2$, although at first he neglected to write the parenthesis. He then solved for M and wrote: $M = \sqrt{L^2 - \frac{1}{4}N^2}$ and explained how this eliminated M. He still was not happy working with N because it was not constant.

J: It's a moving piece (N). It's constantly changing as x and y are changing whereas the d and the a in the previous things were fixed lengths. So I'm thinking about just talking about this length here, call it Q (from the hinge to the pen in Figure 3.7d)..... (long pause)..... I'm thinking about a similar triangles argument again (points to figure).

D: So what triangles do you want to look at here?

J: These two. It's the same as we had before. I'll talk about Q and P (labels length P in Figure 3.7d from the pen to the pin and then points to right triangles with hypotenuses Q and P). It seemed to work before so maybe it'll turn out OK now..... We know that P and Q are in a fixed ratio as before. And I can say that this side here is... ahhh... (points at base of lower triangle in Figure 3.7d with hypotenuse P, and then looks back at his trammel derivation worksheet, Figure 3.7c). I had uhh.... yeah, it's not as simple. What I'm trying to do is basically set up what I had before. The only problem is that, again this here is y (height of lower triangle), but this here is not x like it was before, it's another quantity (base of upper triangle). In this case, it's $x-\frac{1}{2}N$. And I don't like having that N in there because of...... uhh I was trying to think of putting it in terms of something else. But my other sides are $M-y$ (height of upper triangle) and M is also a measured number, and how can I get it. OK I know P, L and Q. Those are always constant..... (long pause)......

Mimicking his trammel derivation, Jim then labeled the base of the lower triangle as $\sqrt{P^2-y^2}$. Jim was then happy with his expressions for all three sides the lower triangle because they did not involve M or N. He then looked hard at the upper triangle and searched for some way to express it's sides without using N and M. Jim was stuck for a while on this problem. He eventually asked me for some sort of hint that might get him going again. He seemed intuitively convinced that there was some way to express these sides using only x, y, Q, and P. I eventually pointed to the base of the lower triangle which he had just labeled with $\sqrt{P^2-y^2}$.

D: You just solved for this. Right?

J: Right.

D: And what's this? (I now pointed to the x-coordinate of his point along the horizontal axis)

J: It's just N minus that (the base of the lower triangle).

D: But it's also x. Right?

J: Oh, OK, so x and that (points to: $\sqrt{P^2 - y^2}$) add up to N.

Jim then wrote that: $\sqrt{P^2 - y^2} + x = N$. Jim puzzled for a while over what to do with this statement. He wanted to relate it to the base of the upper triangle which he had written as: $x - \frac{1}{2}N$, but his anxiety over algebra began to emerge, and he got flustered. I told him to calm down, and just write down what made sense from his picture. He said that he had to take his expression for N and cut it in half and subtract it from x, but doing this algebraically seemed a daunting task to Jim. I asked him to just write down what he wanted to do without trying to simplify it first. He then wrote that the base of the upper triangle was equal to: $x - \frac{\sqrt{P^2 - y^2} + x}{2}$. He then said that he had two sides of the upper triangle, so he could use the Pythagorean theorem to find the third side of the upper triangle.

D: OK, where would you go with that?

J: Well I think I'd sort of end up in the same spot where I had... (looks at his trammel worksheet, Figure 3.7c).... You know, I was starting off with a ratio here, and then I'd say like.... uhh.... the same way I did here, you know a is to this side as d is to this side...... So here I'd say Q is to this thing up here, as P is to that one here.

Jim wrote: $\frac{Q}{P} = \frac{x - \frac{\sqrt{P^2 - y^2} + x}{2}}{\sqrt{P^2 - y^2}}$

D: Is that an equation with the right things in it? You were worried before about having some things that varied.

J: Yeah I think it will be OK, because the only things that I'm dealing with here are Q and P. I said that P was this length here, and Q is this length here and I didn't even use L (points to arm of device).

I next asked Jim what kind of equation he would expect of get if this curve really were an ellipse. He looked at the arms in the positions where they crossed the axes, and said that what he had previously called d, was now equal to P, and what he previously called a, was now equal to $L+Q=2Q+P$. He then said if this were really an ellipse then he would expect the equation to come out as: $\frac{x^2}{(2Q+P)^2}+\frac{y^2}{P^2}=1$.

D: I know algebra is not your strong point, but do you think there's any chance that this equation is the same as that?

J: Ohhh... (sighs)..... How would I go about changing that....... Well, I simplified it last time, and it's the same argument.

D: Do you have any intuitions here?

J: I had intuitions about geometry, but not about this (looks gloomy and tired).

(The interview had now lasted for 2 hr. and 30 min.)

D: Do you want to try to do this? (algebraic reduction).

J: Yeah. It would give me some satisfaction.

With a very clear idea about where he was going, Jim started in on trying to simplify his ratio equation : $\frac{Q}{P}=\frac{x-\frac{\sqrt{P^2-y^2}+x}{2}}{\sqrt{P^2-y^2}}$. He failed at first to distribute a minus sign, and I eventually pointed this out to him. He knew that like last time, at some point he would have square the equation to get rid of the radicals, but this time radicals occurred twice in the equation, and this presented a real problem for him. He wanted to square the equation when one side was still an expression minus the radical, which would have led to a real mess. He was tired and asked for some strategic advice, and so I explained to him that since both radicals were same he should put all the terms containing the radical on one side, and then factor it out before squaring. This was a very difficult maneuver for him to envision, since it involved treating the radical as a single variable entity. Although he had easily see the radical as a length in the figure, it was not easy for him to see it as an entity when performing algebra manipulations, and he certainly did not want to introduce an intermediate variable to stand for the radical. After understanding my suggestion Jim worked hard to simplify his algebra. Eventually was able to reduce his ratio equation to the standard elliptic form as: $\frac{x^2}{(2Q+P)^2}+\frac{y^2}{P^2}=1$.

When he arrived at this form he was immediately aware that the equation represented a general ellipse, and that it was consistent with his geometric experiments. As before Jim's personal confidence was not based on achieving this algebraic result, but this confirmation of his experiments in another representation enhanced his certainty. He very much wanted to see a clear confirmation of what he already believed. Far more than his beliefs about the curves being the same, Jim's confidence in the language of algebra was enhanced. I congratulated him on his fine work, and he beamed with satisfaction.

J: It makes me feel good to get that!

(The interview lasted for 2 hr. and 40 min.)

3.8 A Comparison of the Methods and Epistemology of Tom and Jim

Let me begin with an example of the kind of reductive argument that can be so misleading when trying to think about educational issues. From a strictly formal mathematical view, the work of Tom and Jim on both the trammel and the folding arm devices might be said to be the same. They both drew the same figures, and analyzed the same pair of triangles on both problems. They both worked with the same sets of variable quantities using only different letters. Consider their work on the folding arm device. Jim's N is the same as Tom's x'. They both set up the same expression for the base of the upper triangle, Jim's $x-\frac{1}{2}N$ and Tom's $x''-\frac{x'}{2}$. They both replaced the x-coordinate of a point on the trammel drawn curve, with these expressions to obtain an equation for the folding arm motion. They then both faced the problem of how to eliminate the extra variable (N or x'), so as to get an equation in the appropriate variables that they could compare to a standard elliptic equation. The immediate form of the equation into which they substituted these expressions appears different, but that was only because Jim wrote the equation directly as a ratio from similar triangles, while Tom went to a functional form and used trigonometry to make his substitutions.

If, after their experiences, Jim and Tom had both submitted the kind of terse, summary lab reports that are often encouraged in math and science classes, their teacher might well have said that they had both used the same method of analysis to solve the problem. Tom would undoubtedly have cleaned up and condensed his big messy strings of equations, while Jim would probably never have mentioned his long rounds of physical experiments. Tom might never have described his view of the outer folding arm as a trammel action being transported by the inner rotating arm. Jim might never have described his crucial moments of vision when he saw things that were "in constant ratio." A teacher might look at these hypothetical lab reports, and the only difference that would appear would be Tom's use of trigonometry to circumvent any mention of the similar triangles from which Jim directly constructed his original ratio statements. Looking at such reports this difference might very well seem trivial.

Even a cursory look at what actually went on in these interviews will show the superficiality of the preceding reductive argument. Tom and Jim were very different in what they saw, what they did, what they said, and what they <u>believed</u>. These differences are profound and have educational implications that are both epistemological and curricular. These differences and their implications provide a powerful justification for the use of "qualitative methods" such as these interviews. I can not imagine a more effective way to investigate these issues.

Tom's algebraic skill has become an epistemological stance that affects his behavior and beliefs. His stance is one that is constantly reinforced by both the content and method of most mathematics classes. Tom's skills are the ones that are most likely to be praised and rewarded by school mathematics. His habits of thought make him unlikely to engage in any, but the barest minimum of physical experimentation. I can not say whether Tom developed these habits because of the nature of mathematics classes and their importance in our school culture, or whether he already had such predilections and found a home for them in mathematics classes. It matters little either way.

Jim's skills and habits of observation and investigation are much less likely to be engaged by our traditional mathematics curriculum. His ability to play and tinker and hypothesize in a physical setting are not often called for in mathematics classes. Even his refined visual sense of ratio helps him on only few occasions, due to the paucity of geometry in our curriculum. What passes for "context" in classrooms is most often sets of "word problems" that may describe some situation but rarely involve designing or physically experiencing that "context." Most "contextual problems" are decontextualized.

For example, the trammel involves the same action as a ladder sliding down a wall, a common rate problem in calculus; yet few teachers of mathematics know that the motion of any point on that sliding ladder is elliptical. I have asked many experienced calculus teachers this question and they were all familiar with the rate problem, but they were all very surprised that the motion of points on ladder was elliptical. The most common first guess was hyperbolic (i.e. like $y = 1/x$), followed by some kind of cusped curve like Jim's star. Calculus teachers (myself included) have taught this "contextual rate problem" for years, but have never physically examined the action involved, and therefore tend to have very poor instincts for motion. A student like Jim is far more creative and inspired when given a physical action to control and observe. He can see rates and constant ratios long before he can express them in algebra.

Tom is a very successful mathematics student, but placing him in the kind of physical problematic situation that these curve drawing devices provided, challenged him in a way that can only add depth, flexibility, and perhaps an entirely new dimension to his well developed talents. During his first interview, Tom wanted so much to use his recently developed skill with parametric equations to analyze the trammel. His skill with this algebraic form had been largely developed through long sessions with his graphing calculator, but this tool was entirely embedded in the usual approach where equations have the primary epistemic role. The graphing calculator proved to be a frustrating and useless tool here, because of its built-in epistemic hierarchy which is the reverse of the problem at hand. It took Tom a while to realize that in order to analyze motion, he could not start with algebraic equations no matter how good his "graphic visualization" skills were. The algebra had to emerge from the motion. The recent educational reform emphasis on "visualization" is still locked into the same epistemic hierarchy where equations create curves. Graphs are still secondary facilitators that help one visualize an equation (see, for example, any of the many articles in Romberg et. al. , 1993). Although such reform efforts contribute many important educational insights, they do not give truly independent status to different representations, and the analytic geometry of Descartes is entirely absent. They fail to complete a cognitive feedback loop.

It is interesting to note how Tom's thinking is tied to the narrow traditional notion of a function even when this complicates the algebra. His fundamental insight into the folding arm device was to see it as a trammel being transported, and hence his algebraic approach was to write the trammel equation, and then modify it with a non-constant horizontal translation. He clearly saw the need to write his translation ($x'/2$) as a function of his x-coordinate (x''), and this was important. In order to make the translational substitution, he felt much more comfortable having the elliptic equation in a functional form (i.e. solved for y'). Why didn't he just make his translation substitution directly into the standard elliptic form? This would have made his algebraic reduction very simple. He said at the time, "it helps me to see it," although he did not mean this in a visual or geometric sense, but referred rather to his sense of how to keep his algebra straight.

Jim, on the contrary, showed no strong inclinations towards traditional functional notation. He was much happier using statements about changing rates and ratio equality that directly expressed his geometric vision. His saw that the folding arm configuration contained the same similar triangles as the trammel, but he did not associate this directly with his previous equation. Jim instead preferred to see the ratios inherent in the new system, and study their operation. He too faced the problem of how to write the horizontal distance out to the hinge ($N/2$) in terms of the variables that he wanted to see in his final equation, but he saw this problem in terms of the physical geometry. N was not constant, and was not something one needed to know to set up the device; hence it was not a determining physical parameter of the curve, and therefore did not belong in the final equation of the curve. Jim was very clear on this point. Although he never used functional language his notion of physical parameters that one controls in the device was astute, and his sense of how an equation "talked about" what was happening in the motion of the device was well expressed.

I find it sad that a student as talented as Tom would so completely avoid geometric language, especially the phrase of "similar triangles." His experience with classical geometry in Course Two must have indeed been negative, as he told me after the interviews. The brief piece of Euclidean formality that passes as geometry in our schools was, for Tom, an isolated departure from his path of mathematical development as he saw it, and therefore best forgotten. Perhaps even sadder is Jim's longing to return to geometry, as the piece of mathematics that he most loved. I think that some experience earlier in their curriculum with curve drawing and dynamic geometry could have helped to inspire both of these students. Tom might then have seen geometry as connected far more directly with the branches of mathematics that he loves. More importantly, Jim might have found a way to engage more profoundly his gift for seeing ratios. This might have gone a long way towards changing Jim's attitude about algebra and mathematics in general.

Jim's and Tom's teacher certainly tries harder than many to bring more visualization and geometry into his classroom. They both had strong intuitive and visual notions of eccentricity. It was interesting that neither of them could apply that notion to the curves that they drew. Perhaps this was because their visual experience of eccentricity from computer animations was passive. It was not a concept that they had ever applied in the physical sense of measuring it empirically from a pre-existing curve. Although the very creative worksheet that Tom showed me from his class, where he constructed points on a curve with a given eccentricity, was neither passive nor entirely algebraic, it still created a curve from a pre-specified algebraic property. The curves still did not have a primary epistemic role, and so Tom could not reverse the process.

The most pronounced difference between Jim and Tom involved the process of how they came to believe that differently drawn curves were the same shape. Up until his last algebraic revision Tom, for example, was quite willing to believe that the curves drawn by the folding arm device were not ellipses. Algebraic equations formed the foundation of all of his mathematical beliefs. Jim's beliefs were formed mainly by physical and geometric experience as he so directly expressed (80% according to him). Both students found ways to reproduce curves from one device to the other, but its impact on their beliefs was not the same. Perhaps one might say that Tom had a healthy skepticism, while Jim was too willing to believe what he saw. This view, however, is oversimplified. Tom let his algebra run away with him, and convince him temporarily of things that were clearly contradicted by his physical experience. Jim would never have allowed this to happen to him. To be meaningful to Jim, algebra had to be a slow, careful and precise confirmation of what he had physically experienced.

For both students the experience of connecting and confirming geometric experience with algebraic expression was both engaging and satisfying. All of these interviews point up the need to bring about a more balanced dialogue both between geometry and algebra, and between physical experience and theoretical language. Tom and Jim could both have benefited greatly from experiences with curve drawing long before they reached Course Four. Curve drawing could have been introduced in middle school long before the equations of conic sections were studied. It could have been connected to many other activities; sundials for example. Having a base of such grounded activity would have been beneficial in many ways. It could have given Tom a more flexible sense of geometric expression with deeper connections to algebra.

More importantly, it might have given Jim an entirely different feeling about algebra. If Jim saw algebra as a systematic language developed to allow for the expression of his physical and mechanical visions, it is doubtful that he would have come to see it as boring and fearsome. Even after having developed these debilitating attitudes, he was still able to work clearly and precisely within a problematic situation where his visual skills were clearly valuable and connected to the problematic situation. He did not want to avoid algebra at all costs. He wanted to see how it could express what he saw, and validate what he experienced. By reversing the usual epistemic hierarchy, the curve drawing devices gave him the stamina to work on a difficult problem for over two and half hours, even when he seemed exhausted. His physical certainty as to what would "come out" gave him the determination to finish.

If mathematical language is to become comprehensible to a broader audience, it must display early on its capacity for expressing a wide variety of situations. Most often in our curriculum, the linguistic form of mathematics (usually algebra) dictates in advance both the forms of classroom discourse and the allowable span of activities. That is to say that physical activities and "contextual problems" are introduced as examples or applications of pre-established linguistic skills and concepts. The language and symbolisms are not being generated in response to student activity, but vise versa. Since semiotics usually dictates in advance the content of mathematics curriculum, students are only allowed to discuss activities that fit those forms, and often even simple "activities" are only discussed hypothetically, and never materially explored.

Such a situation severely disadvantages a student like Jim. His skills, thoughts, and epistemic inventions remain largely unengaged. Jim does not really hate algebra; what he hates is the way that linguistic rules have come to dominate the content of his mathematics courses. When language flowed from physical experience, Jim was quite ready to push very hard to coordinate and reconcile language with experience. As he said "the thinking happens in geometry." Jim had a vision of what he expected of geometry, but that vision remained out of touch with school mathematics. Jim's vision was largely a seventeenth century geometric vision, like that of Descartes and Pascal, that involved architecture, civil engineering, and mechanical devices. Jim told me about his continued thoughts concerning the operation of a mechanical apparatus that reset pins in a bowling alley. He was disappointed that the geometry that he learned in high school never helped him to even begin to analyze what he saw there.

The material in this chapter is more than sufficient to establish the second major claim of this thesis (end of Section 1.1). Both Tom and Jim clearly benefited from their experience with these three curve drawing devices. Tom's firmly established algebraic epistemology was gently threatened, and he subsequently gained a broader and more balanced view of the possible dialogue between physical curves and algebraic language. Jim's engagement with the curve drawing devices was even more profound because they satisfied in him a longing for what he saw as the geometry of the world. I learned a great deal from watching and listening to Jim. The phrase "these move in a fixed ratio" combined with certain hand gestures will always remain with me. They have already become part of my thinking about the learning and teaching of dynamic geometry.

If mathematics is allowed to confront the uncertainties and ambiguities of the physical world; if its language, symbols and notations are allowed to grow directly from experiences, then a fully circular feedback loop will evolve an epistemological model. The algebra of equations and functions would then be more than just "the rules." Only then will more students be able to genuinely say, as Jim did at the end of his algebraic derivation, "It makes me feel good to get that."

Chapter Four: Summary and Conclusions The main goal of this thesis, as stated in Section 1.1, was to establish the following three claims:

1). To establish the fundamental historic and conceptual importance of curve drawing devices in the development of analytic geometry, algebraic symbolism, calculus and the notion of functions.

2). To show how two secondary students of mathematics benefited from their experiences with physical curve drawing devices, and that both the geometric and algebraic analysis of these devices raised, for them, crucial epistemic issues, the consideration of which led them to engage in a more balanced dialogue between the physical world and symbolic languages.

3). To show that a discussion of the tangents, areas, and arclengths associated with many curves need not be deferred until calculus, and that, quite the contrary, an understanding of the semiotic importance of calculus depends upon being able to correlate its symbolisms with independently verifiable geometric experience. Such experience can be readily gained from the use of physical curve drawing devices, and from simulations of such devices using available dynamic geometry computer applications.

All three of these claims have been clearly and firmly established by the material presented, the first and third claims in Chapter 2, and the second claim in Chapter 3. What I wish to focus on here is interconnections between the three assertions, and their larger implications. Several questions and directions for future research will arise.

The first two claims formed the initial concept that shaped this research; forged, as it was, within the theoretical perspective of Jere Confrey's research program as described in Section 1.3. Initially, the main theoretical components guiding the research were: genetic epistemology, epistemology of multiple representations, and listening for the epistemic invention of students. The first two theoretical components created a stance from which I conducted historical investigations which, in turn, led me to create a set of tools and problematic situations which allowed the third component of student voice to emerge. The third claim of this thesis came later as the work was in progress, and was not tested in terms of student voice, but it emerged so strongly from the historical research that it deserved the status of being listed as a basic claim.

As this research progressed, I began to interpret it in light of the increasingly unified theories of Jere Confrey (1993a, 1994e), whose evolution and discussion I was privileged to witness first hand. One element of these newly revised theories is the increasing stress that they place on tools and their effect on student voice, in moving towards the creation of a balanced dialogue between grounded activity and systematic inquiry. This viewpoint helped to solidify the historical stance taken in Chapter 2, and to clarify the analysis of the activities of the students in Chapter 3.

The mathematical history of seventeenth century Europe shows a profound and balanced dialogue between physical tools and the linguistic development of analytic geometry and calculus. The historical tools involved activities that we would now classify largely as mechanical or civil engineering, and these have mostly disappeared from mathematics curriculum. A reading of original mathematical sources from the period of the genesis of analytic geometry and calculus, reveals the formation of a fluid and active feedback loop between physically generated, geometrical curves, and algebraic language. A person trained in modern mathematics is immediately struck by how closely intertwined these tools were with the evolving language. This is a very different perspective; the otherness of history. I have interpreted this history of evolving language as a process of verification through cross checks within an epistemology of multiple representations, with physical tools forming an important part of the feedback loop.

Taking this view of history helped me to create the physical and intellectual structure of the student interviews. The students were not being taught a direct historical lesson, and neither were they being given a problem stated solely within modern mathematical language. The tools and problematic situations presented to the students were meant to enable them to enter into a more balanced dialogue between the physical and the linguistic, in an open way that would allow for both freedom of invention and personal expression. The students were not reliving history. This was not possible because they had already been formally exposed to the language and notations that evolved during the seventeenth century. The interviews were structured to allow them to make contact with a part of a feedback loop which is now rarely mentioned in mathematics classrooms, although it historically preceded the form of analytic geometry taught in their classes.

The invention and expression of the students was clearly influenced by the beliefs and linguistic tools that they brought with them. Tom believed in algebraic notation, and it was already a powerful and effective tool for him. The graphing calculator, with which he was so comfortable, embodied (and probably helped to create) his epistemological hierarchy. The curve drawing devices perplexed him at first because, by reversing this hierarchy, they did not immediately yield to his favorite linguistic approaches. His notions of representation were challenged, and he was compelled to respond in new ways in order to reconcile his physical experience with algebraic analysis. His basic belief in algebraic language was not strongly altered, but it was broadened to include a new class of interesting problems that he referred to as "engineering type stuff."

Jim's epistemology was at odds with formal algebraic analysis, which he dismayingly referred to as "the rules." The curve drawing devices placed him in a setting that was much closer to his personal belief structure. He was delighted to carry out extensive physical experiments, and the coordination of a systematic set of physical actions was fundamentally convincing to him without algebraic analysis. He showed great inventiveness in his use of the physical tools (e.g. his location of the foci of a trammel drawn curve). After successfully conducting an algebraic analysis, he was pleased by his ability to check things in an independent representation, but this linguistic setting was not the primary seat of his belief. Recall his words when he first derived an elliptic equation from the action of the trammel device:

D: Does this convince you that the curves are the same?

J: (with resignation) Well, if they have the same equation, I guess I should be convinced.

In large part, Jim was already convinced by his systematic physical experiments (as he eloquently and explicitly stated in Section 3.7).

Both Tom and Jim responded creatively to the challenge to coordinate physical and algebraic representations of curves, but while Tom's voice is rewarded and encouraged in the classroom, Jim's voice is repressed. Current mathematics curriculum gives Tom an inflated view of the power of his algebraic skills, while Jim's beliefs and inventiveness are rarely offered an arena for expression. Analytic geometry as a feedback loop, promoting a balanced dialogue between coordinated physical actions and algebraic language, would provide diversified expressive options, as well as more flexible and profound mathematical content. Tom would reap significant benefits. Jim would be given a voice.

The third claim of the thesis (concerning tangents, etc.) raises the issue of what kind of geometry curriculum might best provide the kind of experience from which the language of calculus could emerge. If language is a tool which codifies, and correlates with experience, then perhaps the static shadow of ancient Greek geometry that is usually taught is not the most appropriate. Greek geometry was certainly not designed for dynamic or mechanical purposes, but rather for logical and philosophical ones (Klein, 1968). These purposes are still valid ones for some students, but Jim's vision of what he expected geometry to be was connected with engineering and architecture. Jim's vision is not an isolated one; many students have expressed to me similar geometric dreams. These dreams are seventeenth century dreams, and they are the genesis of calculus. They deserve a place in the secondary curriculum.

Chapter 2 repeatedly showed how one might construct situations for the investigation of tangents, areas, and arclengths, through the use of dynamic curve drawing devices. I suspect that one would not have to be overly concerned with a precise definition of, for example, tangency if one is dealing with mechanically generated curves. Language like that of Apollonius might well serve introductory educational purposes. A line either "touches" a curve, or "cuts" a curve in the case conic sections (no cusps or inflection points). An exploration of these ideas, through the use student interviews, might prove very fruitful. On a larger scale, it would be very interesting to see how the teaching of calculus could change if students approached the subject with a systematic physical, and geometric experience of tangents, areas, and arclengths. Calculus would undergo an epistemic transformation when seen by students as a language built from, and confirmed by prior experience.

Another direction for further research would be to introduce curve drawing activities to students prior to algebraic notation. In this way one might allow for inventive physical expression like Jim's at an early stage. Curves could be studied systematically through their geometric methods of generation before the advent of graphing equations. One could then explore how the teaching of analytic geometry might change if students approached the subject with a body of systematic experience with curves generated in non-algebraic ways. Once again, the students would see language built from and confirmed by prior experience. Such a journey through genetic epistemology might radically transform curriculum, and require an extended set of educational experiments, but seems to me entirely feasible.

In terms of student activities, another question left unanswered by this thesis is the appropriate role of dynamic geometry computer software. My own struggles to understand the historical material presented in Chapter 2 involved the making of many physical models and extensive experiments with *Geometer's Sketchpad*. How does one find the most meaningful mix of mechanical and electronic tools? In my own experience, it was not easy to get *Geometer's Sketchpad* to simulate linkages. While the program has nice graphic features, it's linear elements do not want to behave like rigid bodies. The definitional aspects of *Geometer's Sketchpad* are based on Euclidean notions of lines and circles, and sometimes to simulate the simplest physical linkage several levels of hidden elements are required. On the other hand, once one is accustomed to these awkward constraints, the software eliminates the problems of friction and collisions that can seriously bog down a physical device. My own investigations of seventeenth century mechanical devices was greatly facilitated by the software, but it required a considerable investment of time and experience. Perhaps new and simpler software will appear that will more directly simulate curve drawing linkages.

The affect on students of mixing physical and simulated devices remains to be investigated. My only conclusion here is that it would be very misguided to approach the subject using only computer simulations. Many of the devices described in Chapter 2 can be easily built and explored. Jim's statement that things that moved "in a fixed ratio" was as much a tactile as a visual observation. He used his hands with care and accuracy throughout his explorations. A computer simulation would never have had the same effect, and would not have given full range to Jim's inventiveness. My own personal experiences while conducting the historical research confirmed this as well. Simulating curve drawing devices would affect both student voice, and epistemology, and investigating that would require a more involved study over a longer period of time. Such a study might lead to important insights into how modern epistemology is evolving in response to new tools which can both revive and supplant the devices of the seventeenth century.

Before embarking on this series of investigations, I began with a large set of mathematical experiences and beliefs coming from years of formal training, a body unorganized historical knowledge, as well as a vague set of pragmatic educational beliefs based upon classroom experience. After coming into contact with the educational research program of Jere Confrey and her associates, all of these were transformed and unified by educational and epistemological theories. I witnessed detailed theoretical discussions and saw how these theories shaped and guided work with very real schoolchildren, "inspiring and inspired at both ends."[39] Without a vision of educational theory and its active implications for shaping student investigations, this work would not have been possible.

Like the feedback model of analytic geometry which I propose, so too educational theory and practice form a feedback loop whose epistemic shape is transformed by tools. My reading of seventeenth century documents was shaped by both educational philosophy, and by the computer simulations that I built. The student interviews were constructed within a theoretical framework using reconstructions of seventeenth tools, but my interpretations of those interviews were transformed by the use of the video camera. I spent many hours as a camera man for Jere Confrey before setting up my own student investigations.

The larger question that remains is how my historical and student investigations will transform my educational theories. As I continue to investigate any of the questions mentioned above, there must be a cyclical revision of the theories that shape the experiments. This preceding research was situated largely within the model of radical constructivism, for the methods employed were genetic epistemology, and

---

[39] This phrase is taken from my opening quote in Section 1.1, p.5, where Courant spoke about Gauss.

videotaped clinical interviews. The study, however, focused on a set of physical tools and their capacity to transform knowledge and language. This brings one face to face with the issues raised by Vygotsky. Putting radical constructivism together with the theories of Vygotsky is not a clear and simple task (Confrey, 1995; 1994e). Important issues of society and its relations to tools must be considered. In the near future both computers and video equipment will greatly expand their capacities. Their impact on both the conduct, content and theory of educational research is difficult to predict, but the feedback loops of knowledge coming from historical and student perspectives, can only be studied within the larger feedback loop which simultaneously transforms educational theory and methods.

References

Arnol'd, V. I. (1990). *Hygens & Barrow, Newton & Hooke* . Boston: Birkhäuser Verlag

Anton, H. (1988). *Calculus with Analytic Geometry* . New York: Wiley

Apollonius, of Perga. (1952). *On Conic Sections.* In Vol. 11 of *The Great Books of the Western World.* Chicago: Encyclopaedia Britannica Inc.

Archimedes. (1952). *Measurement of a Circle*. In Vol. 11 of *The Great Books of the Western World.* Chicago: Encyclopaedia Britannica Inc.

Apostol, T. (1969). *Calculus, Vol. I.* Waltham, Mass.: Xerox College Pub.

Artobolevskii, I. I. (1964). *Mechanisms for the Generation of Plane Curves* . New York: Macmillan Co.

Berggren, J. L. (1986). *Episodes in the Mathematics of Medieval Islam* . New York: Springer- Verlag.

Boyer, C.B. (1956). *History of Analytic Geometry* . New York: Scripta Mathematica Chapters III-V.

Boyer, C.B. (1968). *A History of Mathematics.* New York: Wiley.

Cajori, F. (1913). *History of Exponential and Logarithmic Concepts.* Am. Math. Mon. 20.

Cajori, F. (1929). Controversies on Mathematics Between Wallis, Hobbes, and Barrow. *The Mathematics Teacher*, Vol.XXII Num. 3 p.146 - 151.

Child, J. M. (1920). *The Early Mathematical Manuscripts of Leibniz*.. Chicago: Open Court.

Confrey, J. (1988). *Multiplication and Splitting: their Role in Understanding Exponential Functions.* Proc. of the Tenth Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education (PME-NA), Dekalb, Ill.

Confrey, J. (1992). Using Computers to Promote Students' Inventions on the Function Concept. In Malcom, Roberts, and Sheingold (Eds.), *This Year in School Science 1991*, Wasington D.C. : American Assoc. for the Advancement of Science, p. 131 - 161.

Confrey, J. (1993a). The role of technology in reconceptualizing functions and algebra. In Joanne Rossi Becker and Barbara J. Pence (eds.) *Proceedings of the Fifteenth Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education , Pacific Grove, CA, October 17-20*. Vol. 1 pp. 47-74. San José, CA: The Center for Mathematics and Computer Science Education at San José State University

Confrey, J. (1993b). Learning to See Children's Mathematics: Crucial Challenges in Constructivist Reform. In Tobin, K. (Ed.), *Constructivist Perspectives in Science and Mathematics*.. Washington D.C. : American Assoc. for the Advancement of Science, pp. 299 - 321.

Confrey, J. (1993c). Diversity, Tools, and new Approaches to Teaching Functions. A paper presented for the China-Japan-United States Mathematics Education Conference on Problem Solving, October, 1993. Avaiable from the author.

Confrey, J. (1994a). Splitting , Similarity, and Rate of Change: New Approaches to Multiplication and Exponential Functions. In Harel, G. and Confrey, J. (Eds.), *The Development of Multiplicative Reasoning in the Learning of Mathematics.* Albany N.Y. : State University of New York Press, pp. 293 - 332.

Confrey, J. (1994b). Six approaches to transformation of functions using multirepresentational software. In J. Pedro da Ponte & J. Filipe Matos (eds.) *Proceedings of the Eighteenth International Conference for the Psychology of Mathematics Education , Lisbon, Portugal August 1994.* Vol. 2 pp. 217-224. Lisbon, Portugal: Program Committee of the 18th PME Conference.

Confrey, J. (1994c). *Learning About Functions Through Problem-Solving*. Santa Barbara, CA: Intellimation.

Confrey, J. (1994d). *Function Probe*. [Computer Program]. Santa Barbara, CA: Intellimation.

Confrey, J. (1994e). A Theory of Intellectual Development. *For the Learning of Mathematics*. Appearing in three parts in consecutive issues: **14**, 3, pp.2-8; **15**, 1, pp. 38-48; **15**, 2. Vancouver, Canada: FLM Publishing Association.

Confrey, J. (1995). How Compatible are Radical Constructivism, Sociocultural Approaches, and Social Constructivism? In Steffe, L. and Gale, J. Constructivism in Education. Hove, NJ: Lawrence Erlbaum Associates. Confrey, J. & Smith, E. (1991). A framework for functions: Prototypes, multiple representations, and transformations. In Robert Underhill and Catherine Brown (eds.), Proceedings of the 13th Annual Meeting of PME-NA, Blacksburg, VA, Oct. 16-19, 1991. (pp. 57-63).

Confrey, J & Smith, E. (1995). Splitting, covariation, and their role in the development of exponential functions. *Journal for Research in Mathematics Education*. Vol. 26, No. 1

Confrey, J. & Smith, E. (in press). Applying an Epistemology of Multiple Representations to Historical Analysis: A Review of "Democratizing the Access to Calculus: New Routes to Old Roots" by Kaput, J. To appear in Schoenfeld, A. *Mathematical Thinking and Problem Solving* , Hillsdale N.J.: Lawrence Erlbaum Assoc. Inc.

Coolidge J. L. (1968). *A History of Conic Sections and Quadratic Surfaces* . New York: Dover Publications Inc.

Courant, R. (1984). Gauss and the present situation of the exact sciences. In Campbell, D. & Higgins, J. (Eds.), *Mathematics: People, Problems, Results.* Vol. I, pp. 125 - 133. Belmont CA: Wadsworth International.

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

A. Coxford, A.; Usiskin, Z. & Hirschhorn, D. (1993). *University of Chicago School Mathemtaics Project: Geometry*. Chicago: Scott Foresman.

Dennis, D. & Confrey, J. (1993). The Creation of Binomial Series: A Study of the Methods and Epistemology of Wallis, Newton, and Euler. Presented at the Joint Mathematics Meetings (AMS - CMS - MAA) Vancouver, August 1993. Manuscript submitted for publication, available from the authors.

Descartes, R. (1989). *Discourse on Method*. LaSalle, Ill.: Open Court.

Descartes, R. (1952). *The Geometry.* translated by D.E. Smith and M.L. Latham. LaSalle, Ill. : Open Court.

Edwards, C. H. (1979). *The Historical Development of Calculus.* New York: Springer-Verlag.

Euler, L. (1988). *Introduction to Analysis of the Infinite Book One.* translated by J. D. Blanton, New York: Springer- Verlag.

Euler, L. (1990). *Introduction to Analysis of the Infinite Book Two.* translated by J. D. Blanton, New York: Springer- Verlag.

Gardner, M. (1989). *Penrose Tiles to Trapdoor Ciphers.* W.H. Freeman and Co.

Hall, L. (1992). Trochoids, roses and thorns - beyond the Spirograph. *The College Mathematics Journal* vol. 23. p. 20-35

Heath, T. L. (1961). *Apollonius of Perga: Treatise on Conic Sections.* New York: Barnes & Noble Inc.

Henderson, D. (in press). *Experiencing Geometry on Plane and Sphere.* Engelwood Cliffs, NJ: Prentice Hall

Hilbert, S. et al. (1994). *Calculus: An Active Approach with Projects*. New York: John Wiley & Sons, Inc.

Jackiw, Nicholas (1994). *Geometer's Sketchpad* ™ (version 2.1), [Computer Program]. Berkeley, CA.: Key Curriculum Press.

Jahnke, H. N. (1994). The historical dimension of mathematical understanding - Objectifying the subjective. In J. Pedro da Ponte & J. Filipe Matos (eds.) *Proceedings of the Eighteenth International Conference for the Psychology of Mathematics Education , Lisbon, Portugal August 1994.* Vol. 1 pp. 139-156. Lisbon, Portugal: Program Committee of the 18th PME Conference

Joseph, G. G. (1991). *The Crest of the Peacock, Non-Euopean Roots of Mathematics*. New York: Penguin Books.

Kaput, J. (1993). The Urgent Need for Proleptic Research in the Representation of Quntitative Relationships. In T. Romberg et al. (Eds.), *Integrating Research on the Graphical Representation of Functions.* Hillsdale, N.J.: Lawrence Erlbaum Associates.

Kaput, J. (in press) Democratatizing acccess to calculus: New routes to old roots. In A. Schoenfeld (Ed.), *Mathematical Thinking and Problem Slolving* . Hillsdale, N.J.: Lawrence Erlbaum Associates.

Katz, V. J. (1993). *A History of Mathematics: An Introduction*. New York: Harper Collins

Klein, J. (1968). *Greek Mathematical Thought and the Origin of Algebra*.. Cambridge, MA.: M.I.T. Press.

Lakatos, I. (1976). *Proofs and Refutations, The Logic of Mathematical Discovery* New York: Cambridge University Press.

Lenard, A. (1994). Kepler Orbits, *More Geometrico*.. *The College Mathematics Journal* . Vol. 25, No. 2, March 1994. Washington D. C.: Mathematical Assoc. of America.

Lenoir, T. (1979). Descartes and the geometrization of thought: The methodological background of Descartes' geometry. *Historia Mathematica*, 6, pp. 355-379.

Laubenbacher, R. & Pengelley, D. (1994). Eisenstein's misunderstood geometric proof of the quadratic reciprocity theorem. *The College Mathematics Journal* . Vol. 25, No. 1, January, 1994. Washington D. C.: Mathematical Assoc. of America.

Mahoney, M. S. (1973). *The Mathematical Career of Pierre de Fermat*, Princeton N.J.: Princeton University Press.

Milroy, W. (1990). *An Ethnographic Study of the Mathematical Ideas of a Group of Carpenters*. (Doctoral Dissertation, Cornell University). Dissertation Abstracts International, 57/9, 3007A, 9106270.

National Council of Teachers of Mathematics (NCTM). (1991). *Professional Standards for Teaching Mathematics*. Reston, VA: Author.

Newton, I. (1967). *The Mathematical Papers of Issac Newton., Vol. I: 1664 - 1666* Cambridge: Cambridge University Press.

Newton, I. (1968). *The Mathematical Papers of Issac Newton., Vol. I I: 1667 - 1670*. Cambridge: Cambridge University Press.

Pedoe, D. (1976). *Geometry and the Visual Arts.* New York: Dover Pub.

Richards, J. (1988). *Mathematical Visions, The Pursuit of Geometry in Victorian England*. Boston: Academic Press Inc.

Rizzuti, J. (1991). *Students Conceptualizations of Mathematical Functions: The Effects of a Pedagogical Approach Involving Multiple Representations*. Unpublished Doctoral Dissertation, Cornell University, Ithaca, New York.

Romberg, T. & Fennema, E. & Carpenter, T. (Eds.) (1993). *Integrating Research on the Graphical Representation of Functions.* Hillsdale, N.J.: Lawrence Erlbaum Associates.

Row, T. S. (1966). *Geometric Exercises in Paper Folding*. New York: Dover Pub.

Rubin, A. & Nemirovsky, N. (1991). Cars, Computers, and Airpumps: Thoughts on the Roles of Physical and Computer Models in Learning the Central Concepts of Calculus. In Robert Underhill and Catherine Brown (eds.), Proceedings of the 13th Annual Meeting of PME-NA, Blacksburg, VA, Oct. 16-19, 1991. (Vol. 2, pp. 168-174).

Schooten, Franz van. (1657). *Exercitationum Mathematicorum , Liber IV, Organica Coniccarum Sectionum in Plano Descriptione.* Leiden. (original edition in the rare books collection of Cornell University, Ithaca New York).

Sfard, A. (1992). Operational origins of mathematical objects and the quandary of reification - The case of function. In G. Harel & E. Dubinsky (Eds.) *The Concept of Function, Aspects of Epistemology and Pedagogy* . Washington D. C. : Mathematical Association of America.

Sierpinska, A. (1992). On understanding the notion of function. In G. Harel & E. Dubinsky (Eds.) *The Concept of Function, Aspects of Epistemology and Pedagogy* . Washington D. C. : Mathematical Association of America.

Smith, D. A. (1959). *A Source Book in Mathematics* . New York: Dover Pub.

Smith, E. & Confrey, J. (1994). *Multiplicative Structures and the Development of Logarithms: What was Lost by the Invention of Functions?.* In G. Harel & J. Confrey (eds.), *The Development of Multiplicative Reasoning in the Learning of Mathematics.* Albany N.Y. : State University of New York Press, pp. 331 - 360.

Smith, E. & Dennis, D. & Confrey, J. (1992). Rethinking Functions, Cartesian Constructions. In *The History and Philosophy of Science in Science Education , Proceedings of the Second International Conference on the History and Philosophy of Science and Science Education,* vol. 2 pp. 449 - 466, S. Hills (Ed.) Kingston, Ontario: The Mathematics, Science, Technology and Teacher Education Group; Queens University.

Struik, D. J. (1969). *A Source Book in Mathematics, 1200-1800.* Cambridge Mass. : Harvard University Press.

Turnbull, H. W. (ed.). (1960). *The Correspondence of Isaac Newton, Volume II, 1676 - 1687*. Cambridge: Cambridge University Press

van Hiele-Geldof, D. (1984). The Didactics of Geometry in the Lowest Class of Secondary School. In Fuys, Geddes, and Tischler (Eds.), *English Translation of Selected Writings of Dina van Hiele-Geldof and P.M. van Hiele,* p. 1-214. Brooklyn: Brooklyn College.

Van Maanen, J. (1992). Seventeenth century instruments for drawing conic sections. In *The Mathematical Gazette* , vol. 76, n. 476, pp. 222-230 New York: Macmillan Co.

von Glasersfeld, E. (1978). Cybernetics, experience and the concept of self. In M. Oser (Ed.) *Towards a More Human Use of Human Beings* . Boulder, CO: Westview Press, pp. 109 - 122.

von Glasersfeld, E. (1982). An interpretation of Piaget's constructivism. *Review Internationale de Philosophy* , No. 142-3. Diffussion: Presses Universitaires de France, pp. 612 - 635.

von Glasersfeld, E. (1984). An introduction to radical constructivism. In P. Watzlawick (Ed.), *The Invented Reality* . New York: W. W. Norton, pp. 17-40.

von Glasersfeld, E. (1990). An exposition of constructivism: Why some like it radical. *Journal for Reasearch in Mathematics Education* , 4. pp. 19-31.

von Neumann J. (1984). The Mathematician. In D. Campbell & J. Higgins (Eds.), *Mathematics: People, Problems, Results* . Belmont CA: Wadsworth.

Vygotsky, L. (1962). *Thought and Language* . Cambridge, Mass.: M.I.T. Press & Wiley.

Wallis, J. (1972). *Opera Mathematica*, Vol. I. New York: Georg Olms Verlag.

Whiteside, D. T. (1961). Patterns of mathematical thought in the later 17th century. *Archive for History of Exact Sciences,* **1**, p. 179-388.

Whitman, E. A. (1943). Some Historical Notes on the Cycloid. *American Mathematical Monthly* . Vol. 50, p. 309 - 315.