VISUALIZATION IN MATHEMATICS

by

JENNIFER ANN BEGGS

A THESIS

Presented to the Department of Mathematics
and the Honors College of the University of Oregon
in partial fulfillment of the requirements
for the degree of
Bachelor of Arts

June 1992

APPROVED: [redacted]

Prof. Richard Koch

An Abstract of the Thesis of

Jennifer Ann Beggs for the degree of Bachelor of Arts

in the Department of Mathematics to be taken June 1992

Title: VISUALIZATION IN MATHEMATICS

Approved: [redacted]

Prof. Richard Koch

Traditional methods of teaching mathematics have been aimed at one type of learner, the non-visual thinker. Because not everybody learns in this way, the old teaching method creates a problem for a good portion of students. Visualization in mathematics is a relatively new method of teaching mathematics which incorporates visual techniques into the classroom. This thesis investigates the visualization technique through articles, interviews and observations at schools that emphasize visualization in their mathematics classrooms. It is observed that students taught using the new method are generally more enthusiastic about mathematics and are better able to retain information than students taught in the traditional manner.

ACKNOWLEDGMENTS

I would like to thank Professor Richard Koch for assisting me in this project and for always being supportive and also to Professor Brad Shelton for his valuable input. Thanks also to: John Bruning for his love and the use of his computer, Lisa Millegan for her friendship and input, and to my parents, Mary Ann and Larry Beggs for always being there for me.

TABLE OF CONTENTS

FIGURES

INTRODUCTION

This thesis is a study of a method of teaching mathematics which uses the students' ability to visualize and to learn in ways other than symbolically. In order to meet the increasing need for other methods to teach students mathematics, visualization is a rapidly growing alternative to mainstream teaching methods.

The following chapter is a review of some the research done on this topic. The third chapter is an analysis of the "Math in the Mind's Eye" program in Portland, Oregon which provides teaching materials for visualization in the teaching of mathematics. And in the sixth chapter I use what I have learned about visualization to show how a small section of linear algebra could be taught visually. This last section is of the most interest because it shows that topics that are normally presented analytically can be presented visually.

The section of linear algebra I attempt to visualize is on matrices and properties of rotations and reflections. The fourth chapter is an analytical presentation of that same information, given much in the way the material is presented in a college level linear algebra course. And the fifth chapter is a historical look at the originator of the theory of matrices. Arthur Cayley first wrote about matrices in 1858, and was the first to do so.

These two chapters are important to this thesis for several reasons. The analytical presentation of linear algebra is important because it allows the reader to be able to compare how different the two methods of presentation are, but also that they reach the same goals. It is an example of how many analytical mathematical studies have their roots in geometric proofs. I studied the history of matrices, even though it is unusual to do so, in order to discover how the originators of this theory visualized the topic, if at all. Any indication of how Cayley visualized this topic would be usefull to my study on how to present this material visually.

The main goal of this paper is to present an alternate method of teaching mathematics, and to show that visualization is beneficial to all students and not restricted to the lower grades of K-12, because even highly analytical topics can be visualized.

VISUALIZATION

People learn in many different ways. Some prefer to listen to information, others need to write the information down and formulate it in their own words, while others still need visual pictures and the opportunity to work with their hands. Most people learn best with a combination of these learning methods and others. People who learn through reading and writing are thought to be more analytical and theoretical and those people who need pictures and the hands on experience are more visual and tactile. Mathematics is usually taught in an extremely analytical manner. For analytical learners, mathematics presented in a symbolic method is an acceptable way to learn. But for the visual person, the non-analytical learner, this method is unacceptable. Symbolic representation of mathematics does not provide enough visual information to a visual learner, hence there is little to which the student can relate. These visual learners need to "experience" what they are studying. The typical mathematics teacher needs to be able to teach to both types of learners in order to meet the needs of a greater range of students, and using the visualization technique might just be the way to do this. By integrating visualization into a mathematical lesson both analytical and visual peoples' needs can be satisfied.

Visualization is the act of creating a picture in your mind. If you are speaking about a dog you might have a picture of a dog in your mind. Instead of just thinking the word "dog" you are actually seeing one. Psychologists noticed their subjects were using this technique in memory exercises where the subject was required to remember a list of paired words. The subjects would visualize the two objects in conjunction with each other, and when the cue word was given the subject would recall the picture and then report the desired word. These people were using their visualization skills to improve their memory. For my purpose, visualization will also include actions that require drawing, cutting out pictures and looking at figures. It would be impossible to complete this study if only internal, or spatial visualization was included because spatial visualization is the act of creating a picture in the mind. I want to include drawing, cutting out pictures, etc. because the idea of visualization is for the student to become more involved with what they are studying instead of being a passive observer. The visualization technique allows the student to be more physically involved in their work which in turn gives them a greater understanding of the material. It is hoped that by using this technique, students will become more active and begin to understand mathematics on a regular basis, because they were given the chance to experience the mathematics. The more they are able to understand, the more excited they will be about what they are studying. And there is nothing more beneficial to the learning process than an enthusiastic student.

Visual thinking is a phenomenon that is integrated into all human activities [**16**, 8]. For example, people visualize where they left their car in a parking lot, plot out a route of travel before leaving for a destination, day dream and so on. Hence, visualization is a skill that we are born with and it does not have to be taught. The ability might have to be cultivated to be used effectively in the mathematics classroom, but educators certainly do not have to start at ground zero. Because visualization is a tool most people use on a daily basis, it seems reasonable that educators would try to use this skill to their advantage. According to Robert H. McKim in his book Experiences in Visual Thinking, people visualize in three ways. First they see, second they imagine, like a dream, and third they draw. McKim explains that the best type of visual thinker is one that can integrate all three types of visual imagery[**16**, 8]. Educators need to integrate these elements into their lessons, including also manipulation of materials such as blocks and paper shapes. By integrating these elements, the teacher will be taking advantage of an ability that we all have, therefore creating another realm in which the student can study the material.

Integrating visualization into the classroom is no easy task. It requires the instructor to be very creative, and students must be willing to work hard and try new ideas. Converting an analytical study of numbers into visual and tactile resources is very difficult. But by putting effort into this teaching method the results can be very positive. The technique allows students who are visual learners to grasp onto something that they can understand. For the analytical thinker the added information given through the visualization exercises may not be necessary for their comprehension, but it will certainly not hurt them. As long as the option is there for the students who need the visual information, the purpose has been served. "Visual illustrations can furnish insights into number relationships that are often not obtained by algebraic proofs."[3, 268] The idea of visual illustrations is very important in terms of allowing students to explore mathematics. They can discover things on their own instead of learning everything from the teacher. For students who have always felt frustrated in mathematics, the visualization technique may be a way to unlock the door. As McKim says,"Visual thinking complements abstract-language thinking by its power to concretize"[16, 26]. In other words, by using visualization we can turn abstract languages, such as the language of mathematical symbols, into concrete ideas that can be internalized.

I will now present an example of how this method can be used. The example comes from materials obtained from the Portland project, "Math in the Minds' Eye". This particular lesson is used to teach how to add and subtract negative numbers to fourth graders. The visual tools used are small paper squares that are colored red on one side and black on the other. First a small handful, 5-10, of the squares are thrown down somewhere that everybody can see. The students are taught that every pair (one red and one black) can be taken away; whatever is left is called the "yield". So typical answers are 3 red or 2 black. The students are then allowed time to work with this on their own or in groups. Next, a negative or positive value is assigned to each color, so black might be positive and hence red is negative. The same activity as before is preformed but now when there are three black squares showing after all of the matched pairs have been removed, the outcome is named +3, and if three reds are showing then the outcome would be called -3. In this way the statement 5-2=3 can be visually explained with five black squares and two red squares. Once the two pairs are taken away, the three black squares left over would yield the desired result. Such a method can be expanded to explain why subtracting negative numbers is the same as adding and also how to multiply and divide positive and negative numbers. The example would be very beneficial to students who have a hard time grasping what function negative numbers have and how to deal with them. For the student who does not have difficulty understanding negative numbers, the activity is just fun and interesting. While this activity might not have been necessary for the student who understood negative numbers already, it can still be beneficial for such a student because it puts a level of fun into mathematics and it gives the student one more item to base their knowledge on. In other words, it "reaffirms their knowledge".

Earlier I referred to McKim's idea that visualization consists of seeing, imagining and drawing. I would now like to explain a little about how this affects the mathematics classroom. A teacher using the visualization technique uses many pictures and graphical representations, and has the students draw and cut out figures. Students are encouraged to become involved in their work so that they can "own" it. By "owning" the information, I mean that the student has a firm grasp on the concept and is usually able to explain the concept to someone else. For deep comprehension of mathematics it is very important that students have a firm understanding of what is happening and its relation to matters other than numbers on a page. If teachers are able to create links between numbers on a page and real life situations, students are going to have a better understanding of what is happening on the page. "Teachers should strive to create learning environments which facilitate the development and reinforcement of such links" [9, 27]. Clements poses we do this by giving the students' the freedom to be creative in their study, and not to censor them if they don't do what the teacher wants them to do. He explains that there is too much passive listening in the classroom and it is too restrictive for students to learn in such a way, but visualization allows the students to be creative and to not be censored. McKim confirms this idea by saying "What we need is to be educated in how not to interfere with the inherent capacity of the human mind to think"[16,28]. The visualization method is designed to not censor the students ability to explore and to try things on their own, therefore it does not interfere with the students capacity to think on their own.

Another aspect of visualization is spatial visualization. As Michael Battista claims in his study, the ability to mentally visualize a figure can determine how accurate a picture the student will draw. Battista demonstrated that students who had better spatial ability were more successful in geometry then those who had trouble with spatial visualization. In the study the subjects were given a test over four areas, spatial visualization, logical reasoning, knowledge of geometry, and geometric problem solving/strategies. They were assessed on the following areas: "spatial visualization, logical reasoning, geometry achievement, geometric problem solving, use of drawing strategy, use of visualization without drawing, use of a non-spatial strategy, correct drawing made, and discrepancy between a student's spatial score and logical reasoning score"[1, 49]. Those with a higher spatial ability were able to create more accurate pictures and therefore their answers were more accurate. He explains that the non-spatial thinkers were probably trying to take an analytical approach to the problems, which was why they were running into problems[1, 56]. In his study he shows that the split between spatial and non-spatial thinkers is usually along male/ female lines, but as McKim asserts, almost everybody can learn to visualize. McKim states that it takes a lot of work to develop visualization skills if you are not used to it but it is an ability we all possess[16, 29]. In a research survey done by Eisenberg and Dreyfus, it was found that there was a bias against visual thinking in gifted math classrooms. But,"On the other hand, this reluctance to visualize may well be a result of the way mathematics is presented and communicated by most teachers and research mathematicians"[10, 2].

In the studies done over the last 20 years, many have examined differences between males and females in terms of their spatial abilities. In the study by Battista, it was hypothesized that there are two modes of thought, one verbal/logical and the other visual/pictorial. These two modes of thought cause "mathematical casts of mind", which can determine mathematical ability. What he found was that males were superior in spatial tasks while females were better in verbal tasks. This, he claimed, was because male's modes of thought tended to be more specialized while female's modes of thought were more balanced between the two types. This led him to conclude that high spatial performance is dependent on specialization. That is why, he believed, males tended to preform better in geometry classes than females. "Spatial visualization and logical reasoning were important factors in geometry achievement and geometric problems solving for both males and females"[1, 56]. While this may be true, he did not examine how instructional differences can have a large effect. A study done by Ben-Chaim, Lappan and Houary came to interesting conclusions about educational practices. Again, as in the previous study, they were assessing differences between males and females in terms of their spatial abilities. But in this study, the subjects were first taught using the spatial visualization methods. "The most important result of this investigation was that after the instruction intervention, middle school students, regardless of sex, gained significantly from the training program in spatial visualization tasks"[2, 66]. I think that this is the most important thing that can be found from one of these studies. No matter what the individual differences might be, everybody can benefit from the new teaching practices.

What I have shown in this chapter is that visualization techniques offer an alternative to the standard teaching methods and has the capability to be beneficial to most students. By integrating programs such as "Math in the Mind's Eye" into a mathematics curriculum, fewer students will be falling through the cracks of the system.

MATH IN THE MIND'S EYE

"I don't think there are any non-visual people...everybody dreams. Dreaming is really a form of visual thinking." says Eugene Maier, one of the originators of the "Math in the Mind's Eye" program. To learn more about the "Math in the Mind's Eye", I went to Portland, Oregon, to interview Eugene Maier about the program. A transcript of that interview is supplied in Appendix A. I also had the opportunity to observe three classes at Portsmouth Middle School in Portland and two classes at Roosevelt Middle School in Eugene, all of which use the "Math in the Mind's Eye" program. "Math in the Mind's Eye" is a program from the Math Learning Center in Portland, having offices in the Curriculum and Development department of Portland State University. Their main goal is stated clearly in their catalog;

> "Math in the Mind's Eye teaching materials are designed to draw upon the student's capacity for visual thinking. Through the use of manipulatives, models, sketches, and diagrams, students develop an understanding of mathematical concepts and processes, and create meaningful mental images that help them retain and recall this information."

The program started in 1982 with a group of mathematicians feeling frustrated with the current mainstream style of teaching mathematics which devalued imagery, and "knowing that it's almost totally absent from school after the first or second grade"[15]. The first developers were Maier, Ted Nelson and Don Rassmussen. Originally they set out to bring teachers' attention to the topic by some inservice presentations. The intent was to encourage teachers to incorporate visual thinking into the classroom and to suggest ways to go about doing that. But the teachers wanted more. Because the teachers had not had any experience in visual imagery, they needed more materials. Therefore, the developers began to produce materials to keep up with the demand. The project has grown and now it is supported by the sale of materials and by revenue from workshops.

The group secured a grant from the National Science Foundation in 1984, and still gets some funding from NSF. Originally the grant was given jointly to the Math Learning center and to PSU, but the Math Learning Center is now its own non-profit corporation. The two agencies still work together; the Math Learning Center does the publishing of the materials and organizes the workshops and the educators who attend the workshops are able to receive credit from PSU. Maier explained that any profit made from the workshops or the sale of their materials is put back into the program to fund further research. When the grant runs out in the next couple of years they believe they will be able to fully fund themselves from profits gained from materials and workshops.

The workshops are a chance for the teachers to learn about the program. There are generally three workshops, "Math in the Mind's Eye" Workshops I and II and Implementing "Math in the Mind's Eye". These workshops are held all over the United States all year round. Maier estimates there will be 71 workshops this summer. The workshops provide about 30 hours of contact with the program; in the first two workshops the teachers are taken through the material much as they would teach it to their students. The implementation workshop deals with how to structure the class and how to evaluate the students. Maier explained that a workshop on evaluation is needed "because what happens with teachers is that in addition to a different mind set about mathematics...the way they assess students changes. The single answer test and computation and single manipulation doesn't test what you have been doing"[15]. In a discussion with Heather Nelson, a math instructor at Portsmouth Middle School, where all of the math teachers use the "Math in the Mind's Eye" program, she explained that she assesses her students on the basis of short quizzes and a portfolio of their work. She feels that large examinations are unnecessary in this type of situation. Nelson is also a leader for some of the workshops given in the Portland area, and her experience in the workshops leads her to believe that the "Math in the Mind's Eye" program is very enlightening for the educators as well as the students. Both Maier and Nelson expressed a need for the teacher to really want to change their teaching methods for the program to work, as it requires a lot of time and effort to implement the program into a school. Nelson said that you have to be really frustrated with the old way. Audrey Manning at Roosevelt Middle School explained that she discovered the program when she was about ready to get out of education altogether because she had been so frustrated with the results. She said that it took a few years to implement the program, but she is much more satisfied with the results than she was with the old system. Given that the program is still relatively new, most of the educators using it have been teaching for many years; in time new teachers will begin teaching in this way and not feel the same frustration.

The materials, the Visual Math Course Guides and the units to supplement the Course Guides, are all available at several different levels, but most of the materials are geared toward middle school students. This is mainly because when the money became available from NSF, it was earmarked for middle school curriculum development, and hence the project started there. But even in the beginning there were clear indications that the program would work in higher levels. Linda Foreman was a high school teacher who got involved with the developmental stage. She had been very frustrated with the classes she was teaching, so she chose to throw out everything she was doing and teach using visual ideas. She had such success with visualization she was eventually encouraged to develop her course into what is now the Visual Math Course Guides.

We discussed the benefits of having a computer program in a higher level math class, one that was able to generate these figures, but Maier feels you should build a model first yourself.

> "You can put that on a computer screen and people can look at it, but there seems to be something about giving people pieces of graph paper and some scissors and have them cut out these things and tape them in. It's totally different. I think basically you just develop a more different sensory perception...Then afterwards it would be nice, if you wanted to look at others to have the computer program...But I don't think you can skip the hands on part of it"[15].

Maier said that while it is more difficult for secondary math teachers than for elementary math teachers and for people who have studied math for a long time to switch to this method, it is not impossible, especially since, as Maier believes, most mathematicians are visual people who don't want to admit it. Maier told a story of a mathematics professor who had become stuck while trying to do a proof for a class. The professor went to the side of the board and drew a diagram to figure out the problem. He erased the diagram and explained to the class the diagram was his cheat sheet. Maier asserts that it would be better for the students to know the diagram and where the proof came from than to learn the result alone. If the students had that kind of information, they would be able to recreate the proof for themselves if they got stuck instead of relying fully on notes and memory. The "Math in the Mind's Eye" program strives to give students that kind of background so that they can go back and recreate what they need to know instead of memorizing formulas and tricks.

What I saw during my observations at Portsmouth Middle School and at Roosevelt Middle School were sixth, seventh and eighth graders ready and willing to work on the mathematics put before them. The teachers ran the classes by either posing problems or going over the homework from the night before and having a student go to the overhead projector and work out the problem for the class. Not only did the students solve the problem, but they had to explain their reasoning. Since one of the goals of "Math in the Mind's Eye" is to not inhibit creativity and to allow the students to find their own way of doing the work, all answers were given consideration and praise.

The first class I observed was a class of eighth graders being taught by Paul Griffith, and the students were working on proportions. They were given two numbers and asked to find the proportion between them by using a line of that many units. The students were then given a third number and their task was to find a fourth number so that the proportion between the third and fourth number was the same proportion as that between the first two numbers. The first problem given had 6 and 10 as the first two numbers and 12 and ? as the second two numbers. At least two different ways to reach the same answer were presented. The first student drew two lines 6 and 10 units long respectively. He divided both lines in half and saw that the first line was divided into two equal lengths of three and the second was divided into two equal lengths of five.

```
 6  ---/---
10  -----/-----          Figure 1
```

He concluded that the proportion was three to five. Next the student drew the third line 12 units long and divided it into three equal parts of length four. Since he now had the three part of the equation, he figured that if he found five segments of length four he would have his answer.

```
12  ----/----/----
 ?  ----/----/----/----/----      Figure 2
```

From the line constructed he was able by counting to discover that the proper ratio was 12 to 20.

Another method presented was to divide each of the first two lines into equal lengths of two, such as:

6 --/--/--

10 --/--/--/--/-- Figure 3

Seeing that the first line was divided into three parts prompted the student to divide the third line of 12 units into three parts. From that he got that the fourth line needed to be five sections of four units.

12 ----/----/----

? ----/----/----/----/---- Figure 4

Again the correct answer of 12 to 20 was reached.

I found this an innovative way to talk about ratios and I felt that the students had a firm understanding of how they reached the answers that they did. I also felt that the students were willing and even eager to present their method to the class. The liberal doling of praise and explanations by Griffith gave the classroom a very comfortable atmosphere.

The second class I observed was a class of eighth graders being taught by Heather Nelson. They were using areas to talk about percentages. Students in the third class, taught by Dorothy Geary, were seventh graders. They were using the integer pieces that I described before to add, subtract and multiply negative and positive numbers. I found it impressive that the students were learning about multiplication of negative and positive number without ever being told that was what they were doing.

At Roosevelt I observed two of Audrey Manning's classes. The first were eighth graders and the second were sixth graders, but both classes were doing the same work, and the sixth graders seemed to have a better understanding of the material. This is an illustration of a comment that Nelson made. She said the earlier the students begin to learn in this way, the better they are at mathematics.

Manning's classes were using areas of rectangles to solve division problems. If given the problem 4.8/1.2=?, the students would solve this by knowing that the area of the rectangle is 4.8 and one of the dimensions is 1.2. Their task was to find the second dimension of the rectangle. They were given grid paper to work on, where one small grid was a .1 by .1 square. I liked the fact that this exercise was a direct result of a multiplication exercise where the answer was the area of the rectangle when the two dimensions were given. Students were encouraged to use what they already knew to solve the problem. Manning explained to the class that the purpose of exercise was to "build models to make sense out of concepts." Her goal was for them to be able to visualize the rectangle so at a later date they could decide if an answer they got from their calculator was reasonable or not.

I was very impressed by what I saw in my observations and with my discussion with Eugene Maier. All of the instructors expressed a sense of satisfaction with what they were able to teach to their students. The good things that I saw in the program were: the positive and encouraging atmospheres of the classrooms. Since the method allows for more than one way of solving a problem, the students were not intimidated by the fact that someone else did the problem in the different way. I liked the fact that some problems were so open ended that there was not one right answer. This forced the student to think about what was reasonable rather than what the ultimate right answer was. I was impressed with how willing students were to work on the material and to share their method with the rest of the class. I think the constant positive feedback encourages the students to be brave and not afraid to try. I also think the program allows the teacher to incorporate the material into their classroom as they see fit. It is not an all or nothing situation. Overall the program in action is a very exciting thing to watch.

Despite the many promising attributes of this program, there are several elements that cause some concern. One being, is it possible to emphasize visualization too much, in that the student cannot understand a symbolic representation? In the real world mathematics is not presented visually, so can the students adjust to that after only working with visual mathematics? To avoid such a problem, I believe that it is important for the teacher to create a balance between symbolic representation and visual representation. The students need to see the connection between the two different types of representations, and need to be able to move from one to the other. This could also help alleviate problems for a student that moves on to a mathematics class that does not use visualization.

Another point of concern involves the question whether or not the connection between real life situations and mathematics is valid. Word problems are real life situations and yet most students find word problems to be the most difficult type of problems. How does the program deal with the problem of students just wanting straightforward questions with straightforward answers? This issue did not seem to be addressed in the curriculum, which is unfortunate because its premise is to get students more involved with the mathematics and to not be robots. But how do you convince the students that this is the better way to learn when they don't want to get involved? Here, again, a balance needs to be found between the symbolic presentation and the visual presentation.

Some other questions that should be addressed in regards to this program are: whether or not all material is appropriate to visualization? Is it really better that there is not always one right answer? Doesn't it frustrate the students that sometimes there is not a right answer? Are there alternate ways to do things if the student is physically unable to manipulate the learning tools? What if the student is blind? These and many other questions should be addressed when it comes to the development of any new method in the teaching of mathematics.

Of course, the program is not a cure-all. Some students struggled with the information and some said that it was boring. No matter what system you use, there will always be behavior problems to contend with and students who do not seem motivated in the least bit. But, on that point, Heather Nelson made an encouraging comment. She said that she and the other teachers at Portsmouth feel that the students who appear to be tuned out during class are still picking up more than they had under the mainstream analytical method.

LINEAR ALGEBRA

Linear algebra is a very broad topic that can be intertwined into every section of mathematics. I am going to concentrate on linear transformations and matrices. Matrices are a topic that is taught in high-school, but without any theory. Students learn how to add, subtract and multiply matrices but they never learn why the operations work as they do or what the matrices are capable of doing. At the college level, all prospective math majors take a linear algebra course, but here, once again, the students do not get a real feeling about why things are done as they are. Rarely are diagrams given to describe what is happening with the linear transformations which the matrices are representing. Both in high school and in college, students do not have any idea of the geometry behind the above; this can be detrimental to the student's ability to comprehend the material.

I will be reviewing an introduction to the theory of linear transformations and matrices. This chapter will be a brief overview of the theory of matrices and will be done much in the way that the material is presented in a linear algebra class. The following chapter will be a peek at how the originator of this material, Arthur Cayley, looked at matrices and how he imagined the matrices worked. Then, in the last chapter I will suggest a way to introduce the linear algebra material to a class in a visual and tactile manner so that the students have a clear sense of how the matrices work. For the mathematician, the first two sections will be familiar material; nonmathematicians, by glancing through the following sections can get a feel for how theoretically the material can be presented. It is my hope that even the non-mathematician will be able to follow my introductory presentation of visual linear algebra.

First I will begin with a definition of a linear transformation in two dimensions. A linear transformation is a map from R^2 to R^2, where R^2 is a two dimensional plane, which associates a vector L(v) with each vector v. The linear transformation is required to have the following properties: if v and w are vectors and r is a real number then L(v+w)=L(v)+L(w) and L(rv)=rL(v).

A theorem following directly from this definition states that such an L can be represented by a matrix. In the two dimensional case, this result can be proved as follows:

Proof: Suppose $L\begin{pmatrix}1\\0\end{pmatrix}=\begin{pmatrix}a\\c\end{pmatrix}$, and $L\begin{pmatrix}0\\1\end{pmatrix}=\begin{pmatrix}b\\d\end{pmatrix}$

Then
$$L\begin{pmatrix}x\\y\end{pmatrix}= L(x\begin{pmatrix}1\\0\end{pmatrix}+ y\begin{pmatrix}0\\1\end{pmatrix})$$
$$=xL\begin{pmatrix}1\\0\end{pmatrix}+yL\begin{pmatrix}0\\1\end{pmatrix}$$
$$=x\begin{pmatrix}a\\c\end{pmatrix}+y\begin{pmatrix}b\\d\end{pmatrix}$$
$$=\begin{pmatrix}ax\\cx\end{pmatrix}+\begin{pmatrix}by\\dy\end{pmatrix}$$
$$=\begin{pmatrix}ax+by\\cx+dy\end{pmatrix}$$
$$=\begin{pmatrix}a & b\\c & d\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}$$

This is what we hoped to find, so the theorem is proved.

The matrix that corresponds with $x_1=ax_0+by_0$ and $y_1=cx_0+dy_0$ is

$\begin{pmatrix}a & b\\c & d\end{pmatrix}$ because $\begin{pmatrix}a & b\\c & d\end{pmatrix}\begin{pmatrix}x_0\\y_0\end{pmatrix}=\begin{pmatrix}x_1\\y_1\end{pmatrix}$.

Multiplication of matrices corresponds to composition of linear transformations. So if $L=\begin{pmatrix} a & b \\ c & d \end{pmatrix}$, and $M=\begin{pmatrix} e & f \\ g & h \end{pmatrix}$ then (LM)(v) is equivalent to L(M(v)) and this is given by:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} e & f \\ g & h \end{pmatrix}=\begin{pmatrix} ae+bg & af+bh \\ ce+dg & cf+dh \end{pmatrix}$$

This is the manner in which matrices are multiplied.

These equations take a point (x_0, y_0) and send it to (x_1, y_1). What happens if I want the transformation to be a rotation? What is the matrix that corresponds to a rotation? Assume that we are rotating by angle α. First we need to find what the rotation does to the vector (1,0). By using trigonometry we can easily see that rotation of (1,0) by α yields $(\cos\alpha, \sin\alpha)$. Hence we have:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix}=\begin{pmatrix} \cos\alpha \\ \sin\alpha \end{pmatrix}$$

Hence, $a(1)+b(0)=\cos\alpha$ so $a=\cos\alpha$ and

$c(1)+d(0)=\sin\alpha$ so $c=\sin\alpha$.

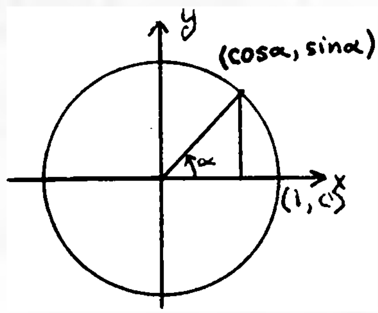

Figure 5

At this point we know a and c. Now to find b and d we need to consider the image of (0,1) under the rotation by angle α . By considering the figure below and trigonometry, we know that (0,1) goes to $(-\sin\alpha, \cos\alpha)$ under the rotation by α . Hence we have:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix}=\begin{pmatrix} -\sin\alpha \\ \cos\alpha \end{pmatrix}$$

Hence, $a(0)+b(1)= -\sin\alpha$ so $b= -\sin\alpha$ and

$c(0)+d(1)= \cos\alpha$ so $d= \cos\alpha$.

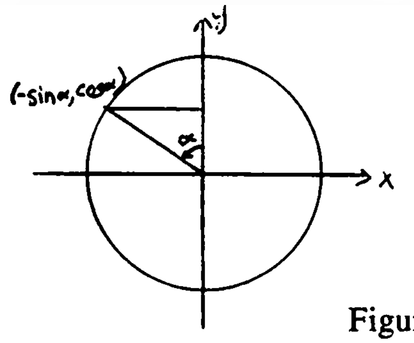

Figure 6

We have our rotation matrix: $\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$.

One of the amazing consequences of these rotation matrices is that many trigonometric identities can be derived with these matrices. I am going to derive the formulas for $\cos(\alpha+\theta)$ and $\sin(\alpha+\theta)$. To do this I want to find the matrix for rotation by $(\alpha+\theta)$. From the definition of matrices I know that rotating by $(\alpha+\theta)$ is equivalent to rotating by α, then rotating by θ.

So

$$\begin{pmatrix} \cos(\alpha+\theta) & -\sin(\alpha+\theta) \\ \sin(\alpha+\theta) & \cos(\alpha+\theta) \end{pmatrix} = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

$$= \begin{pmatrix} \cos\alpha\cos\theta - \sin\alpha\sin\theta & -(\cos\alpha\sin\theta + \sin\alpha\cos\theta) \\ \sin\alpha\cos\theta + \cos\alpha\sin\theta & -\sin\alpha\sin\theta + \cos\alpha\cos\theta \end{pmatrix}$$

From these matrices we can simply read off the identities.

$\cos(\alpha+\theta)=\cos\alpha\cos\theta- \sin\alpha\sin\theta$

$\sin(\alpha+\theta)=\sin\alpha\cos\theta+ \cos\alpha\sin\theta$

This has been a brief introduction to matrices and linear transformations. The presentation has been done much in the same way that the material would be presented in a university level linear algebra course. As it is, there is very little appeal to the senses. One might have the feeling that matrices are very powerful tools, but an understanding of the full implications of the theory is lacking. Students simply memorize the formulas without knowing what the matrices are doing. I will propose a visual method to introduce this material, so that students can see and understand the power of matrices.

ARTHUR CAYLEY

At this point in my thesis I am going to deviate a little from the main idea and discuss the original paper on matrices, written by Arthur Cayley. While it is unusual for a mathematician to discuss the history of the topic in a paper such as this one, I feel that the history is important and interesting. By looking at the original papers on a topic, one might find some insights into how the originators visualized their topics, and hence give the instructors something to base their attempts to teach the subject in a visual manner.

A brief biography of Cayley follows; the information was obtained from a biography in Volume 8 of his Collected works, written by A.R. Forsyth in June of 1895 soon after Cayley's death. Arthur Cayley was born August 16, 1821 to Henry Cayley and Maria Antonia Doughty. At an early age Arthur Cayley showed great mathematical and scientific ability. Despite his father's wishes that Arthur would take over the family business, Arthur was sent to Cambridge at the age of 17 to nurture his abilities. He entered Trinity College in 1838 and was coached by William Hopkins of Peterhouse. Here he excelled and eventually became the Senior Wrangler in 1842. The Senior Wrangler was the student who scored the highest on a very difficult examination called the Tripos Exam. In 1840 he was admitted as a scholar and in 1842 he was elected a Fellow of Trinity. At that time he was the youngest student ever to be elected to this position. As an undergraduate he began publishing papers, as he would do at a incredible rate for the rest of his life. His first paper was published in 1841 in the second volume of the Cambridge Mathematical Journal. Over the years the topics of his papers included every aspect of pure mathematics as well as theoretical dynamics and physical astronomy.

In 1846, at the age of 25, he was forced to make a decision about his future. In order to stay at Cambridge he either had to obtain a permanent teaching post, which was not available at that time, or take holy orders. Since he was unwilling to take the holy orders he chose to enter into a career of law. He left Cambridge to study under Mr. Christie at Lincoln's Inn. In 1849 he was "called to the Bar" [8, xiv] and became a very successful lawyer. But his interest in mathematics kept him from becoming part of a large firm. He chose to keep his clientele to a minimum to allow himself the time to do research in mathematics. In the 14 years he spent as a lawyer he still produced two to three hundred mathematical papers.

In 1860 Cambridge created a new position called the Sadlerian Professorship of Pure Mathematics. The position was named after Lady Mary Sadleir, who had set up an endowment to the university in her will of 1701. Originally the endowment was meant to be used to fund "lectureships in algebra at nine of the colleges in Cambridge"[8, xv]. Due to advancements in analysis, these classes became inadequate to encourage advancement in mathematics. Also, the lectureships were being poorly attended by the undergraduates for whom they had been intended. To put the money to better use, Cambridge created this new position. The description of the post was "to explain and teach the principals of pure mathematics and to apply himself to the advancement of that science"[8, xvi]. Cayley was elected to this position in June of 1863 and he remained in this post for the rest of his life.

He was considered very influential as a lecturer and an investigator. It was reported that many times in his lectures he would report on discoveries he had made since the last session. Students also found him very willing to work with them individually and to give advice on their work. Despite the teaching demands on him he continued publishing papers of his own work. And in 1889 he began editing his works for the Cambridge University Press. Unfortunately he died before the task was completed and the remaining volumes of his works had to edited without his input.

During the last few years of his life, he was confined mainly to his house at Cambridge. He suffered a great deal from a "painful internal malady" [8, xxii] and he died January 26, 1895. He left behind his wife Susan, whom he had married in 1863, and their two children Mary and Henry.

While he was a very serious mathematician, he did have other interests in his life. He enjoyed traveling and spent summers during his undergraduate years in Switzerland. He had a love of mountain climbing and was a member of the Alpine Club. Hiking was an activity he pursued as long as his health permitted. He also had a love for paintings, architecture and for reading novels. This last activity was considered a "popular weakness"[8, xi]. Among his favorite authors were Scott, Jane Austin and George Eliot. He read their books several times, and while he was sick, he had them read to him. While he was considered a great scholar with incredible analytical skill, he also had a well rounded personality with the ability to discuss everything from book plots to art.

Cayley is considered the creator of the theory of matrices, although, unknown to Cayley, a few isolated results had been obtained by Hamilton in 1852. Cayley was the first to speak of letting a matrix be represented by a single symbol and to do the calculations with this symbol instead of returning to the equations at each stage. He created a new class of quantity. His work on this topic was validated by work done by other mathematicians stemming from Cayley's original paper. Work was done by Sylvester, Tait, Peirce, Clifford and Buchheim, to name a few. There is really no way to measure of influence that Cayley had on the mathematical world.

The paper I am considering appears in Volume 2 of The Collected Mathematical Papers of Arthur Cayley, Sc.D., F.R.S., pages 475-496. The title of the paper is "A Memoir on the Theory of Matrices", which was first published in 1858 in the "Philosophical Transactions of the Royal Society of London", volume CXLVIII. He began his paper by explaining that matrices are an alternate way to look at linear transformations and linear equations. In his introduction he explains that many things we can do with numbers can also be done with matrices, such as addition, multiplication, division, and finding roots of equations. He explains that his memoir will attempt to prove that "every rational and integral function (or indeed every rational function) of a matrix may be considered as a rational and integral function, the degree of which is at most equal to that of the matrix, less unity"[8, 476]. I will discuss this more later, but first I will briefly give a summary of the properties of matrices that he discusses. Interested readers will find a copy of Cayley's paper as Appendix B.

Cayley first explains that, for his purposes, he will be dealing mainly with 3x3 matrices but it can be assumed that all of the material presented can be applied to matrices of any size. After explaining how the matrix is a representation of a linear function, he describes the zero matrix and the unity matrix. The zero matrix is a matrix where all of the entries of the matrix are zero. The zero matrix has the same properties as the real number zero. Any matrix multiplied by the zero matrix results in the zero matrix. The zero matrix can be represented by 0. The unity matrix is a matrix where the integer 1 appears in the left to right diagonal, and all other entries are 0: $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

The unity matrix can be represented by 1 and it has the same properties as the real number 1 does: another matrix multiplied by the unity matrix results in the original matrix.

Cayley next explores addition of matrices and the properties of addition for matrices. To add matrices, terms in like entry placements are added together. This is represented in this manner:

$$\begin{pmatrix} a & b & c \\ a' & b' & c' \\ a'' & b'' & c'' \end{pmatrix} + \begin{pmatrix} x & y & z \\ x' & y' & z' \\ x'' & y'' & z'' \end{pmatrix} = \begin{pmatrix} a+x & b+y & c+z \\ a'+x' & b'+y' & c'+z' \\ a''+x'' & b''+y'' & c''+z'' \end{pmatrix}$$

I will quickly list the properties of addition for matrices. A matrix will not be changed in any way if added or subtracted together with the zero matrix. If L=M, where L and M are both matrices, then it follows that L-M=0. Similarly, if L=-M, then L+M=0. In this last case the matrices are opposite of each other. Commutativity and associativity hold for addition of matrices. In other words, when L,M, and N are matrices then L+M=M+L and (L+M)+N=L+(M+N)=L+M+N.

The rule for multiplying a matrix by a single quantity is defined by:

$$m \begin{pmatrix} a & b & c \\ a' & b' & c' \\ a'' & b'' & c'' \end{pmatrix} = \begin{pmatrix} ma & mb & mc \\ ma' & mb' & mc' \\ ma'' & mb'' & mc'' \end{pmatrix}$$

This operation satisfies the commutativity and the distributive law, hence mL=Lm and m (L+M)=m L+m M. We notice that :

$$m \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} m & 0 & 0 \\ 0 & m & 0 \\ 0 & 0 & m \end{pmatrix}$$

Because the unity matrix can be written as 1, we can rewrite the previous statement as:

$$m = \begin{pmatrix} m & 0 & 0 \\ 0 & m & 0 \\ 0 & 0 & m \end{pmatrix}$$

This "is said to be the single quantity m considered as involving the matrix unity" [8, 478].

Next, Cayley looks at multiplication of matrices, which is a difficult operation to explain, but unfortunately Cayley gives no motivation to why multiplication is done in such a way. Let $L = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ and $M = \begin{pmatrix} p & q \\ r & s \end{pmatrix}$

then LM is defined by $LM = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} p & q \\ r & s \end{pmatrix}$

$$= \begin{pmatrix} ap+br & aq+bs \\ cp+dr & cq+ds \end{pmatrix}$$

This operation is rarely motivated, and it is generally given as a rule to be memorized. A few properties follow: as stated before, a matrix multiplied by the zero matrix results in the zero matrix and a matrix multiplied by the unity matrix is the original matrix. Example:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} a(1)+b(0) & a(0)+b(1) \\ c(1)+d(0) & c(0)+d(1) \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

One interesting fact about multiplication of matrices is that the operation is associative, but it is not commutative. The following is an illustration of the non-commutative law.

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} 5 & 2 \\ 1 & 3 \end{pmatrix} = \begin{pmatrix} 5+2 & 2+6 \\ 15+4 & 6+12 \end{pmatrix} = \begin{pmatrix} 7 & 8 \\ 19 & 18 \end{pmatrix}$$

$$\begin{pmatrix} 5 & 2 \\ 1 & 3 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 5+6 & 10+8 \\ 1+9 & 2+12 \end{pmatrix} = \begin{pmatrix} 11 & 18 \\ 10 & 14 \end{pmatrix}$$

Cayley next considers powers of matrices. If L is a matrix then L^p is a power of that matrix. And given that p and q are positive integers then (as with powers of real numbers) $L^p L^q = L^{p+q}$. Notice that $L^0 = 1$ so $L^p L^0 = L^p$. This leads directly to a discussion of inverses of matrices because if p = 1 and q = -1 or vis versa then $L^1 L^{-1} = L^{-1} L^1 = L^0 = 1$. We call L^{-1} the inverse of L. This is what would have to be true in order that $\begin{pmatrix} p & q \\ r & s \end{pmatrix}$ be an inverse of $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$.

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} p & q \\ r & s \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

and $$\begin{pmatrix} p & q \\ r & s \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

This leads to a discussion of inverses and determinants. It is interesting to note that determinants had been known in relation to linear algebra for at least 100 years before the invention of matrices. In the 2x2 case of matrices, the determinate for the matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is defined by det(A)=ad-bc.

The determinate is used to find if the inverse of a 2x2 matrix exists or not. If det(A)=ad-bc=0 then the inverse of A does not exist. To prove this we must look at how to find the inverse of a 2x2 matrix in the general case. If $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ and $B = \begin{pmatrix} p & q \\ r & s \end{pmatrix}$, and B is the inverse of A, we see from above that AB=1=BA. So

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} p & q \\ r & s \end{pmatrix} = \begin{pmatrix} ap+br & aq+bs \\ cp+dr & cq+ds \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

We can equate the entries to get

$$ap+br = 1 \qquad aq+bs = 0$$
$$cp+dr = 0 \qquad cq+ds = 1.$$

We want to get p, q, r, and s in terms of a, b, c, and d. By solving this system of equations we get that p=d/(ad-bc), q=-b/(ad-bc), r=-c/(ad-bc) and s=a/(ad-bc). Now, if we remember that det(A)=ad-bc we can see that:

$$\begin{pmatrix} p & q \\ r & s \end{pmatrix} = \begin{pmatrix} d/\det(A) & -b/\det(A) \\ -c/\det(A) & a/\det(A) \end{pmatrix}$$

If det(A)=0, then each of the terms of the inverse would be divided by 0, which is impossible. Hence if det(A)=ad-bc=0 then the A inverse does not exist. If det(A)=t (for some t in the real numbers not equal to zero), then A inverse exists and is in the form given above.

The next topic that Cayley covers in his memoir is what is now called the "Cayley-Hamilton Theorem". This is a theorem that is now considered very difficult and is rarely proved in an introduction course to linear algebra. Here Cayley is showing "the determinant, having for its matrix a given matrix less the same matrix considered as a single quantity involving the matrix unity, is equal to zero"[8, 482]. The modern version of this theorem says: Let A be a 2x2 matrix and form the polynomial

$$P(m)=\det(A-mI) = \det\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix} - \begin{pmatrix} m & 0 \\ 0 & m \end{pmatrix}\right) = \det\begin{pmatrix} a-m & b \\ c & d-m \end{pmatrix}$$

$$=(a-m)(d-m)-bc.$$

Then A is of root of this polynomial, i.e. P(A)=0. The following is how Cayley proves this theorem for the 2x2 case by calculating the result. It should be noted that it is very difficult to prove this for the general case using Cayley's method, so a different proof is given today. To see why Cayley emphasizes this point, please consult the memoir in Appendix B.

We let $M=\begin{pmatrix} a & b \\ c & d \end{pmatrix}$.

Here we notice that the values of M^2, M^1 and M^0 are:

$$M^2 = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a^2+bc & b(a+d) \\ c(a+d) & d^2+bc \end{pmatrix}$$

$$M^1 = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

$$M = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Now we can substitute these values into our equation for the determinant to get:

$$P(M) = \begin{pmatrix} a^2+bc & b(a+d) \\ c(a+d) & d^2+bc \end{pmatrix} - (a+d)\begin{pmatrix} a & b \\ c & d \end{pmatrix} + (ad-bc)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} a^2+bc & b(a+d) \\ c(a+d) & d^2+bc \end{pmatrix} - \begin{pmatrix} a^2+ad & b(a+d) \\ c(a+d) & ad+d^2 \end{pmatrix} + \begin{pmatrix} ad-cb & 0 \\ 0 & ad-cb \end{pmatrix}$$

$$= \begin{pmatrix} (a^2+bc)-(a^2+ad)+(ab-cb) & b(a+d)-b(a+d)+0 \\ c(a+d)-c(a+d)+0 & (a^2+bc)-(ad+d^2)+(ad-cb) \end{pmatrix}$$

$$= \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

Hence, we can conclude that if $P(m) = \det \begin{pmatrix} a-m & b \\ c & d-m \end{pmatrix}$ then $P(M)=0$.

Here Cayley explains that he has verified this theorem for a 3x3 matrix "but I have not thought it necessary to undertake the labour of a formal proof of the theorem in the general case of a matrix of any degree"[8, 483]. You can see this process becomes laborious very quickly as the expanded determinant for a 3x3 matrix is as follows:

$M^3 - (a+e+i)M^2 + (ei+ia+ae-fh-cg-bd)M - (aei+bfg+cdh-afh-bdi-ceg) = 0.$

This theorem is known now as the "Cayley-Hamilton Theorem". As I said before, what he gives in his memoir is a demonstration that the theorem works for the two dimensional case, but he has not proved the theorem in the general case.

While Cayley does go on to discuss other elements of matrices, such as square roots of matrices, it is beyond the range of this paper. A reader notices when reading Cayley is that his ideas are not cut and dry and he is experimenting. Some of his ideas, such as his proof of the characteristic polynomial, do not seem complete and there is still work to be done on the area. His memoir does show how the topics in matrices flow together and how one idea comes from another. This presents a good way to look at matrices when introducing the topic to a class.

By simply flipping through Appendix B containing this memoir, it is easy to see that all of Cayley's work is done algebraically. There are no graphs or figures to illustrate what he is investigating and nowhere does even his English suggest a geometric interpretation. Cayley probably did not visualize this topic, but it is important to note that since Cayley's time we have discovered the geometric implications and interpretations of matrices. But also note that this information is very difficult to visualize and to illustrate, especially once you pass two-dimensions. I believe that three-dimensions would be the limit to how far you could attempt to display this material in a visual manner, and to do that you would have to have access to a powerful computer.

The hope when first reading this memoir was to get an idea of how Cayley visualized his topic. The reality of the situation is Cayley did not develop the subject geometrically or even use geometry as an aid in the discovery process. Unfortunately, we will never know how if Cayley visualized the topic, but very important information comes from reading the original papers. For an educator the most important information would be the order in which the material was presented. Cayley's memoir is written in such a way to show that he was experimenting. This presents a logical order to present the material to a class, for if the students are able to use the previous knowledge and to experiment with the material in order to learn the new information, they are given a clearer sense of where the material comes from.

VISUAL LINEAR ALGEBRA

After reading our earlier treatment of linear algebra, one can see how extremely algebraic the material is. It is my assertion that it does not have to be that way. When this material is presented to a class, it is usually only presented in the general case and the instructor assumes that the student can transfer that into a specific case on their own if needed. I believe the starting point should be the specific case, with generalizations to follow later. Also, it is important to look at non-trivial cases. Many times if only a trivial or simple case of a problem is presented, many of the steps are glossed over because they are easy, while these steps are not so easy in a non-trivial case. I believe it is important to look at the difficult parts with the students instead of assuming they will be able to figure them out on their own. It is dangerous to assume what a student might or might not know.

You will notice that I will use very little of the technical language used in the previous chapter. Many times the meaning of a theorem gets lost in the language. The technical language can be found in the books and studied there but should not be the emphasis in the classroom.

My approach to visualizing the short section of linear algebra I presented is to look at specific cases and allow the student the time to play around with the equations before presenting the theory. Once the students are familiar with how the equations act and what they can do, a foundation is present on which the theory can be built. I can imagine that if the students were allowed to work with the material first with a little direction provided by the teacher, they could create some of the theory on their own, without it being hand fed to them by the teacher.

The goal of this section is to present the relationship between reflections and rotations in two dimensions using matrix theory. It is hoped the theory will develop as a consequence of the specific cases and not the other way around. I feel that the three dimensional case has a lot of potential for being presented visually and that it deserves investigation, but that is beyond the scope of this paper.

The students need to have certain tools to work with. The most important tools are graph paper, colored pens and pencils. Linear algebra in two dimensions is about shapes and lines within the coordinate plane, and therefore the work should be done on the coordinate plane. Using colored pens and pencils allows the figure to be drawn clearly. By giving certain elements specific colors, the student can personalize their work so that it makes sense to them. Students should be allowed to work this out on their own, but a few suggestions would be to have all the original vectors or shapes be of one color, and their images be of another color. Another possibility is to color each separate vector and it's image in the same way, but with different colors for other vectors and their images. Another valuable tool is to create what I call a Vector Board. This is a piece of cardboard with graph paper attached to it. There is a hole punched through the paper and cardboard at the origin. Some form of string is threaded through the hole to represent vectors. Embroidery thread or curling ribbon are possibilities. It is preferable that the string be of at least two colors, for the reason described above. There should be several strands of each color, and all of them should be tied together at one end, to allow you to thread through only the strands needed. The hole should be big enough to allow this, unless a needle is handy to thread the string through the hole. Also needed are pins or tape to hold the string down onto the paper.

Using the Vector Board will give the students the opportunity to physically move vectors around and to get a real feel for what a vector is.

For these exercises I will be working in two dimensions only. By using the Vector Board, one can easily show that the plane is a vector space and the Vector Board is a physical representation of the vector space. A vector is represented by threading in one piece of string and placing the tip at some point, say (3,4); the string is the vector that corresponds to the point (3,4). It is very important at this stage to make sure that there is no confusion about vector spaces and vectors. If the student does not understand what a vector is then there is no way the theory is going to make sense. Therefore, students should be allowed the time to draw vectors or work with the Vector Board to become comfortable with the concept of a vector, even though it might seem trivial.

The next step is to talk about reflections. A reflection is done by drawing a line from the given point, say X, to the line of reflection, ℓ, at a 90 degree angle and then continuing on through by an equal length to the reflection point Y. In other words, line ℓ is the perpendicular bisector of XY.

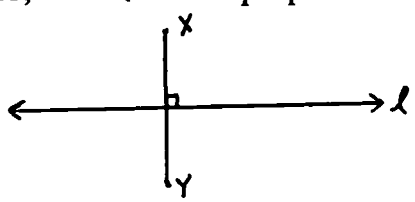

Figure 7

The exercise is to reflect points and vectors over the x-axis, and to try to find the coordinate relationship between the point and its image. The students should find that the x coordinate stays the same and the sign of the y coordinate is reversed, or $x=x_0$, and $y=-y_0$.

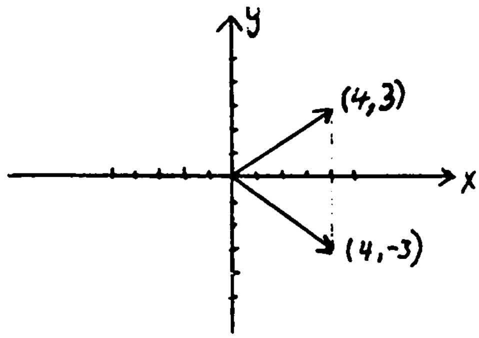

Figure 8

Next, the students are to do the same exercise with the y-axis. They should find that $x=-x_0$, and $y=y_0$.

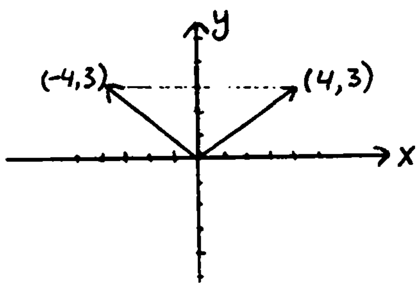

Figure 9

Now what happens it you were to reflect over the 45 degree angle line that goes through the origin? Because this exercise is not as trivial as the previous ones it is important that the students draw and estimate carefully. First they should consider points along the x or y-axis such as (1,0), (5,0) or (0,3) and then try arbitrary points, such as (6,3). What happens to points such as (1,1) and (3,3)? The students should be able to find the relationship between the points and their images, that being $x=y_0$, and $y=x_0$.

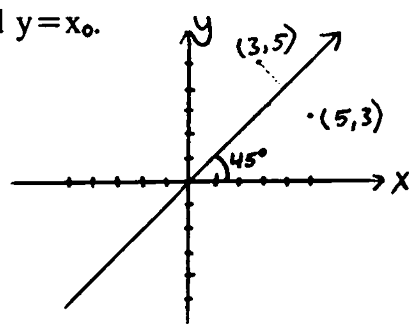

Figure 10

And finally, consider the case of reflecting over the 22.5 degree line. This is getting more difficult because the answer now depends on both x and y. The best way to find the coordinate relationship is to consider what happens with the vectors (1,0) and (0,1).

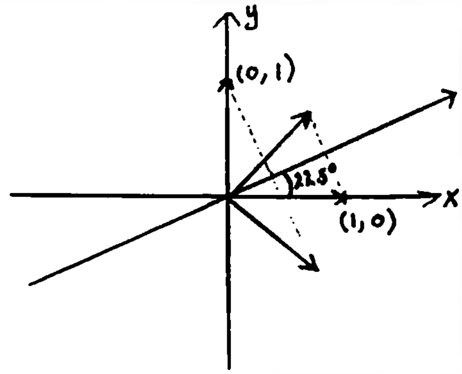

Figure 11

By using graph paper, and by estimating carefully, it can be seen that the image of (1,0) is on the 45 degree line and the image of (0,1) is on the -45 degree line. Because we know that reflections preserve distance we know these reflections remain on the unit circle and by using trigonometry we can deduce that the image of (1,0) is $(\sqrt{2}/2, \sqrt{2}/2)$ and the image of (0,1) is $(\sqrt{2}/2, -\sqrt{2}/2)$. We can see from this that the coordinate relationship is $x = (\sqrt{2}/2)x_0 + (\sqrt{2}/2)y_0$, and $y = (\sqrt{2}/2)x_0 - (\sqrt{2}/2)y_0$. Now we can use these equations to calculate several reflections over the 22.5 degree line.

At this point I would introduce matrices, explaining that we can calculate the points we found geometrically with a matrix. In the two dimensional case, the linear equations $x = ax_0 + by_0$ and $y = cx_0 + dy_0$, where (x_0, y_0) is the original point, (x,y) is the image and a,b,c,d are all real numbers, corresponds to the matrix:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}$$

Now that matrices have been introduced, the students should go back to the equations they found with the reflections and find the matrices that correspond to the reflections over the x-axis, the y-axis, the 45 degree line, and the 22.5 degree line. The answers should be $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, $\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$, $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$

and $\begin{pmatrix} \sqrt{2}/2 & \sqrt{2}/2 \\ \sqrt{2}/2 & -\sqrt{2}/2 \end{pmatrix}$ respectively.

Use these matrices to find out what happens when you reflect a vector over one line and then another. Graph the results using (1,0) and (0,1) and then several different vectors. Using colored pens and pencils or the Vector Board here would be very beneficial in being able to keep track of the answers. Compare the final answer to the original vector in order to find a single coordinate relationship that will send the original vector to that image. If you reflect over the x-axis or $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, then over the y-axis or $\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$, the reflection would look like this for point (2,3):

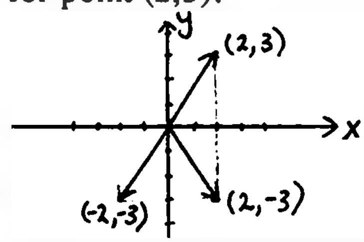

Figure 12

We can calculate with the matrices or by observing the graph that the coordinate relationship is $x = -x_0$, and $y = -y_0$. This corresponds to the matrix $\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$. We see from this that the composition of two matrices is another matrix of the same degree. Using the same method, try to find the single matrix that will preform the same transformation as the composition of any two of the above matrices.

Here would be a good place to discuss matrix multiplication. The composition of two reflections done with matrices is matrix multiplication. Some of the results from the exercise above should look like: $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ with $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ is $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$,

and $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ with $\begin{pmatrix} \sqrt{2}/2 & \sqrt{2}/2 \\ \sqrt{2}/2 & -\sqrt{2}/2 \end{pmatrix}$ is $\begin{pmatrix} \sqrt{2}/2 & \sqrt{2}/2 \\ -\sqrt{2}/2 & \sqrt{2}/2 \end{pmatrix}$.

Now if we did the same exercise with the general case of $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ and

$\begin{pmatrix} e & f \\ g & h \end{pmatrix}$, we can find the general formula for matrix multiplication.

We know we can compose the first matrix with the point (x,y) to get $\begin{pmatrix} ax+by \\ cx+dy \end{pmatrix}$, and now we want to compose this with the second matrix. We get:

$$\begin{pmatrix} e & f \\ g & h \end{pmatrix}\begin{pmatrix} ax+by \\ cx+dy \end{pmatrix} = \begin{pmatrix} aex+bey+cfx+fdy \\ agx+gby+chx+dhy \end{pmatrix}$$

$$= \begin{pmatrix} x(ac+fc)+y(be+df) \\ x(ag+ch)+y(bg+dh) \end{pmatrix}$$

$$= \begin{pmatrix} ae+fc & be+df \\ ag+ch & bg+dh \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$$

This gives us the general formula for the composition of two 2x2 matrices.

Now, let's switch to rotations. On the Vector Board draw a circle around the origin with a compass. Thread in several different strands and place the vectors on the circle. Rotate each of those vectors by 90 degrees, and mark where they go. Next, try rotating by other angles, say 45 degrees, 60 degrees, 180 degrees or anything else. What is the result if you rotate a vector first by 30 degrees and then by 50 degrees. Is there another way to get this result, or is there a single rotation that preforms this function? You could have rotated by 80 degrees to begin with. Make a conjecture about two rotations. A reasonable conjecture would be: Two rotations composed together is the same as a single rotation by the sum of the two rotations.

As with the reflections, there are matrices that correspond to rotations. To find the general rotation matrix let us consider the unit circle and remember trigonometry. This section can be introduced at this point much in the same way it was described in the previous chapter. The main difference would be to describe how the answers were found with trigonometry in more detail and to allow the students to investigate the answer on their own first before going through the steps. Once the rotation matrix is found, use the rotation matrix to calculate and graph several vectors and their images under the rotation. Use either graph paper or the Vector Board.

Do: $\begin{pmatrix} \cos 60 & -\sin 60 \\ \sin 60 & \cos 60 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix}$

and $\begin{pmatrix} \cos 60 & -\sin 60 \\ \sin 60 & \cos 60 \end{pmatrix}\begin{pmatrix} 3 \\ 4 \end{pmatrix}$

Answers: $\begin{pmatrix} \cos 60 \\ \sin 60 \end{pmatrix} = \begin{pmatrix} 1/2 \\ \sqrt{3}/2 \end{pmatrix}$ which corresponds to $(1/2, \sqrt{3}/2)$.

and $\begin{pmatrix} 3\cos 60 - 4\sin 60 \\ 3\sin 60 + 4\cos 60 \end{pmatrix} = \begin{pmatrix} 3(1/2) - 4(\sqrt{3}/2) \\ 3(\sqrt{3}/2) + 4(1/2) \end{pmatrix}$

$$= \begin{pmatrix} (3-4\sqrt{3})/2 \\ (3\sqrt{3}+4)/2 \end{pmatrix}$$

which corresponds to $((3-4\sqrt{3})/2, (3\sqrt{3}+4)/2)$.

Create and compute a couple rotations on your own using the rotation matrix.

Go back to our conjecture about the product of two rotations being another rotation. Can we show this with matrices? Try and find out. Calculate using matrix multiplication what happens when you rotate by one angle and then another, as we did with the reflections. Equate your answer with a rotation matrix that has the angles already added together. If you used specific angles, the answers should be the same. If you have used arbitrary angles, then you might recognize that the answers you have obtained are the same as the trigonometric identities found in a trigonometry course. This was done in the Linear Algebra chapter.

We now want to find a relationship between reflections and rotations. Look back at the composition of two reflections that we did earlier. With the compositions that we calculated, graph a few vectors and their images under the new transformations. Look for a relationship between the different points and their images, looking specifically at the angle that the two vectors form at the origin. You should see that the angle is consistent no matter which vector you choose. I will consider the case of reflecting over the 22.5 degree line, and then the 45 degree line. By matrix multiplication one can see that the composition of the two matrices corresponding to the two reflections is $\begin{pmatrix} \sqrt{2}/2 & -\sqrt{2}/2 \\ \sqrt{2}/2 & \sqrt{2}/2 \end{pmatrix}$.

The following figure is a graph of (1,0) and (0,1) and their images under this transformation.

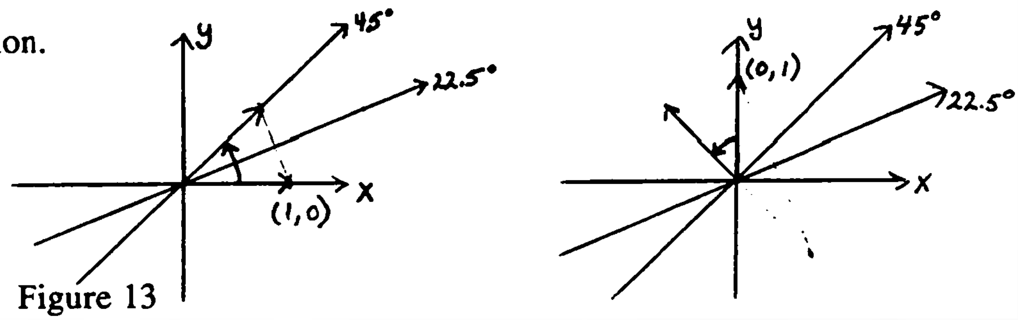

Figure 13

Notice that the angle between the original vector and its' image is 45 degrees in both cases. How is that related to the angle between the two lines of reflection? Notice that the angle between the two lines of reflection is 22.5 degrees.

What does this tell us? A reasonable conjecture would be that the product of reflections over two intersecting lines is a rotation by twice the angle between the lines. Using several of the compositions from before, confirm for yourself that this is true. If you remember the trigonometric forms, you can probably recognize that $\begin{pmatrix} \sqrt{2}/2 & -\sqrt{2}/2 \\ \sqrt{2}/2 & \sqrt{2}/2 \end{pmatrix} = \begin{pmatrix} \cos 45 & -\sin 45 \\ \sin 45 & \cos 45 \end{pmatrix}$.

We now have one very interesting conjecture: The composition of reflections over intersecting lines is a rotation by twice the angle of intersection. I want to look at one more question: What happens when you compose a reflection with a rotation? I am going to look at a reflection over the 45 degree angle line composed with a reflection by 60 degrees. One important point to notice at this point is that order is important, because if you reverse the order of the matrices the answer will be different. For my purposes I will only be considering one case. Notice $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} \cos 60 & -\sin 60 \\ \sin 60 & \cos 60 \end{pmatrix} = \begin{pmatrix} \sin 60 & \cos 60 \\ \cos 60 & -\sin 60 \end{pmatrix}$.

Is this a rotation or a new reflection? To check if this is a rotation I will graph a few points to see if the angle is consistent. Again I will consider (1,0) and (0,1)

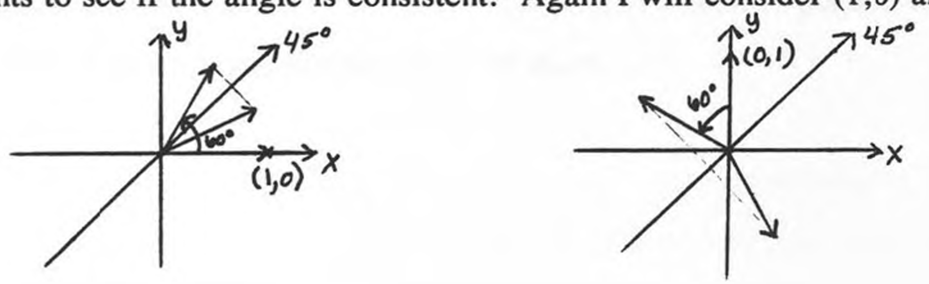

Figure 14

It is obvious that the angle is not consistent, so this must be a reflection. Can you predict the line of reflection geometrically? Because we know that a line of reflection is the perpendicular bisector of the segment connecting the original point and its' image, we can easily construct the line of reflection, as follows:

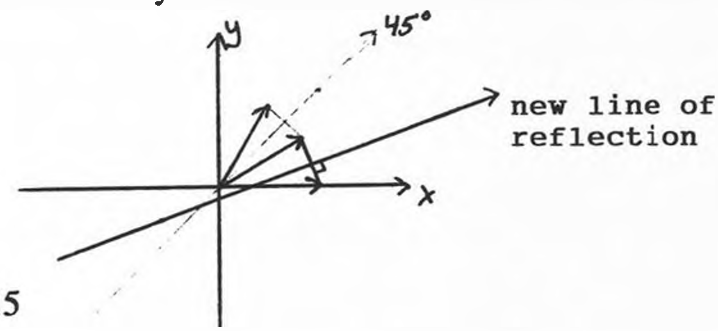

Figure 15

Hence we have shown that the composition of a rotation with a reflection is another reflection over a new line.

This section of my paper has shown that typically algebraic material can be presented geometrically and visually, with a little effort on the part of the instructor and the student. It is my assertion that mathematics should be presented geometrically whenever possible in order to introduce the visual aspect into the classroom. As I explained previously, drawing graphs and creating physical models of the mathematics greatly increases the amount of comprehension of the material.

A very valid concern might be whether or not it would be possible to cover the required amount of material using this method. It's true that teaching in this method is slower than teaching purely analytically, but I would argue that it is more important to have a firm understanding of a smaller portion of material than to have a shaky understanding of a lot more.

CONCLUSION

I have shown in the course of this paper that visualization is a viable option for a method of teaching mathematics. Because it has been found that many people need visual cues and hands-on experience to learn new material, visualization teaching techniques provide a reasonable alternative to teaching in a purely symbolic manner.

I believe that programs such as "Math in the Mind's Eye" are an evolutionary result of educational trends. Educators are always trying to find better ways to teach their students, to make the material more tangible for them. Visualization is one way to do this and "Math in the Mind's Eye" is a direct result of this kind of direction in the educational field.

With more and more computers in the classroom and increasing attention put on the issue of alternative teaching techniques, I believe that visualization methods will become a prominant technique used in the mathematics classroom. I believe the method can be integrated into all levels of education as an integral part of the learning process. As more information becomes available on the benefits and drawbacks of using visualization methods, even more people will become interested in using the programs. Integrating visualization into a mathematics classroom will become a trend for future generations of teachers.

APPENDIX A

Interview: Eugene Maier, April 15, 1992

Jennifer Beggs: How did this project get started and when?
Eugene Maier: Well, I don't know. I'll have to think back. Sometime around '82 or '83. I think the actual funding from the National Science Foundation probably came sometime in 1984-85. I just recently gave a talk down in Eugene and I was curious about when they first used the title "Math in the Mind's Eye" and I discovered it was twelve years ago. Twelve years ago this fall actually. And the project finally came about, well it has a number of origins. Some is from reading people like Arnheim and McKim. The book by Robert Somners, The Mind's Eye, I was just browsing in a book store and I picked this book up because it looked curious, then I noticed it had a chapter in it called "The New Math". And one of the points he was making was that the reason the New Math was so unsuccessful in most settings, was that it devalued imagery. And there wasn't any imagery, it was all simple. I also have an interest in visual thinking because I am interested in dreams. Done a fair amount of things that involve dreams. Things of that nature. So I also have that kind of background. Another part of this comes from working with teachers of Native Americans and that was back in the early 80's. And when you got other parties involved, you go to teachers of Native Americans and say "What kind of things do you think would be useful to your students in the classrooms?" and we had a project we were bringing to them at a Northwest Math conference. They'd go to sessions and look at things. Part of the things that were always picked out were highly sensory, highly visual. My sense is that Native Americans are probably highly visual people. As little as I know about that it seems to be that way. So from that and my own interest in visual thinking and dreams and things of that nature and having read things like... I don't know how I got into it, visual thinking, experiences in visual thinking. I first started doing a lot of reading back in the 70's about split brain things and people like, what's his name? Bob Samples and Betty Edwards and Drawing from the Right Side of the Brain, things of that nature. Putting it together with the visual thinking and that having read things like Hadamard book, which is an old book The Psychology of Invention in the Mathematical Field. It, gee whiz, the Dover reprint, when was this done. The Dover reprint was 1954, the book itself is copy written in 1945, so it has been around for a long time. In that book he basically concludes, what he was trying

to do was to determine how mathematicians thought. He sent a bunch of questionnaires out to a body of mathematicians. And he decided after reading their responses that most mathematicians were highly visual. And the most famous response comes from Einstein, and his reply is accredited in its entirety in the index and it said something about word and symbols are not really a part of his thought processes, and he'd do the math later. So the roles of those kind of things and my own experience is in mathematics where way back , probably two or three life times ago, I did do some mathematics. I knew that the stuff I did, well some of it was called geometry of numbers. The work that I contributed to those papers was all done off of sketches and diagrams and I know that's how I thought about this stuff and if you look at the papers you see that it is all algebraic symbols. It's my own inclination to think in highly visual terms and then also knowing that a lot of mathematicians did. And also when I was going to school, in the 50's and before, it was denigrated. The idea that diagrams and sketches could mislead you and lead you astray. I finally figured out that I could lead somebody astray with some kind of a logical argument just as easily as with a diagram or something. So I decided that one of the things that could happen instead of saying, well because you made a mistake, not to devalue that imagery and more intuitive part of things but to try to nurture and to inform you so you don't make any mistakes anymore, just like when you make a mistake in logic. People don't tell you to quit using this logical argument but they try to inform you or bring to the surface what errors you were making. So there is that aspect. Also the way that it is one of those things that mathematicians did but didn't talk about in a sense and I've told a story a couple of times. Did you hear my talk in Eugene?

J.B.: I didn't. Unfortunately I saw it after the fact but it was right in the middle of one of my finals. I wish I had.

E.M.: There was a story about a professor, who was here in Portland at a school and I remember one of the students telling me one time he was trying to prove this theorem in an analysis class and he got stuck and so he goes off to the side of the board and draws this diagram. And then what he does is the interesting part. He draws a diagram and then he goes back, erases it, goes back and finishes his theorem. And then he says to the class, "That's my cheat sheet." Which is kind of an interesting way to refer to that. He was carrying all this information in the diagram. What he was doing was reconstructing the proof from a symbolic proof onto the diagram instead. It always struck me that it would be more important to say to the students, it would be more useful to say that's my thinking pad and here's how I'm carrying that information into the problem like that. And then knowing what we call visual thinking, now we call visual thinking sort of how McKim uses the words. Also knowing that's almost totally absent from school after the first or second grade. So we thought, well, it would be interesting to try to do some things around visual thinking in the schools programs. So that's how it started and NSF had some money available in their theory development. It could have started, I suppose, in a number of different ways. Since the money was primary available for middle schools, so we said we will do it there. So that is where we started.

J.B.: Who originally funded the group?

E.M.:The Math in the Mind's Eye project was funded by the National Science Foundation. Still gets some funding from the National Science Foundation.

J.B.: Now is it self supportive at all, by sale of the materials?

E.M. Yeah, to large extent. Our grant is running out but...Well the project is a project of Portland State University and the Math Learning Center. Now the Math Learning Center is a separate corporation and it's a private corporation. It's at not for profit, tax exempt. When I left the University of Oregon, what had happened was someone else had formed the organization. I did not think at that time, I did not expect to be here at Portland State University. I just moved to Portland because that's where the population was in Oregon, and we were planning on working with teachers. And, so then over the years, the Math Learning Center developed this relationship with the university. I'm actually an employee of the Math Learning Center and not the university. I am retired out of the university system. So the grant was jointly to the Math Learning Center and to Portland State University. And the Math Learning Center publishes the materials and organizes the workshops. The workshops give credit through the University and it has become fairly popular so we do make some money. I don't know if we make that much off the Mind's Eye, or the long haul of the materials of the Math Learning Center. So we earn some money off of that and we plough that money back into it. So we are matching the NSF money or maybe more. I think this summer we've got 71 workshops going on around in different places.

J.B.: The workshops are for the teachers, right?

E.M.: Right.

J.B.: So the project itself doesn't actually hold classes for students, its for training of the teachers.

E.M.: Yeah, the materials go along with it. At first we weren't going to develop materials as much as we did. Basically what we thought we would do was conduct some inservice courses, and other things that brought teachers attention to that this is a mode of thinking that has some value in mathematics, and it would be nice to incorporate it into the schools, and here are a few ideas for doing that. Most teachers had not had any experience in that, so they were wanting examples. And so we started putting out these units and from that came the course guides. Now it's heavily into materials.

J.B.: What kind of training does it involve for the teachers themselves? Teachers who have been teaching in an analytical layout their whole lives and now they suddenly want to change.
E.M.: First thing they have to want to change. I think it varies from teacher to teacher. Some teachers...most teachers I would have to say... most teachers start by taking what we call...we have several workshops and we have what we call Math in the Mind's Eye Workshop I and Math in the Mind's Eye Workshop II, Implementing Math in the Mind's Eye. Each workshop is about 30 contact hours and what we do in those workshops is basically go through the materials and learn about some of the stuff that's in the Mind's Eye units. We simply have the teachers do it much as the kids would do it in class. We do that or there is a second class available or we have a what called a implementation courses. Because what happens with teachers is that in addition to a different mind set about mathematics, and different approaches, there is something else that happens. The way they assess students changes. The single answer test and computation and single manipulation doesn't test what you have been doing. And that is one change. Another change, and this is almost serendipity, is we find that when we ask people to do things like create images and cross sketches and ask them how they think about things in those terms, why there is a lot of variety and a lot of different ways a students approach things. For example how they might see a pattern. They might see that in a variety of different ways. And we begin to talk about that. And we've discover that when first in workshops with teachers that there are a variety of different ways that a person might see something and that they would be interested in telling about it and other people interested in hearing about it. In the course of doing that more people got and generated more ideas. It seems to lend itself well to that.
J.B.: What would you tell someone about the program that didn't know anything about it.
E.M.: That's one reason we made a video about it. Hand it to them and say "go stick this in their VCR." Or a statement that is similar to what is in the catalog. Math in the Mind's Eye draws upon the students capacity for visual thinking. They use models, sketches and diagrams to create images that they can use and call upon to carry mathematical information forward. Might give them an example or something, have some blocks.
J.B.: Just carrying them around.
E.M.: Yeah, I used to go around and throw some blocks out to folks and say, well, you know, do something with these blocks that demonstrates or a representation of that(2exp3). You know what is interesting? Took a long time before I found anybody who would build a 2x2x2 cube. People would, like my accounting son would put two cubes like this and then stick three up here for an exponent. I remember the first two people who did this. One was my son who

is a science professor and the other one's a medical researcher up on the hill. But his daughter in the 8th grade did that..[same as accounting son]. So we did a lot of that kind of thing. Ask people what images they carry in their head of 3x4. And they say 3x4 and we try to point out to them what that might look like. The image of it.

J.B.: You said that Portsmouth uses this progam. Is it throughout the whole school or just the math department?

E.M.: The whole math department.

J.B.: Are there any other schools in the area, or could you estimate how many programs use this.

E.M. No, not really. I think teachers use it to varying degrees. I think we've probably sold six to eight thousand copies of the Math in the Mind's Eye materials. All you need is one set if you are a teacher.

J.B.: Do you find that it is hard for people, for older students to suddenly to start learning in this way? I have found that some people think it would be difficult to suddenly change the way you were thinking about something.

E.M.: It's hardest for people who have...It's harder for secondary math teachers than for elementary math teachers. The people who have studied more mathematics, perhaps. In terms of teachers. Well, all of the people we work with are adults and don't find it...For most people it works out. Every now and then we find a person it doesn't work for. But by and large they seem to pick it up. Then also, here at Portland State you use a lot of those kinds of methods in the math teacher's course. There is a whole program for middle school teachers here. Seems to be fine. I did an experimental course with some upper division first year graduate students doing some number theoretic topics. It was a little hard for some folks because they're used to doing the symbol manipulation kind of thing and looking at things algebraically and some people struggled with it.

J.B.: In several of the studies that I read they talked about that in geometry classes... because I thought that real visual people were having problems in their math classes and these studies were saying in geometry classes it was the visual people who were doing better than the non-visual people. And it took me a while to think about that to kind of rationalize it to myself.

E.M.: I suppose it kind of depends on the kind of geometry course. If it is a geometry course that is kind of informal, and doing a lot of diagram type things. I don't think there are any non-visual people.

J.B..: You don't think there are any?

E.M.: No, everybody dreams. Dreaming is really a form of visual thinking. You are taking and expressing feelings and that sort of things. There is meaning there.

J.B..: I talk a little bit about that in my thesis, that everybody is born with the ability to visualize.

E.M.: I think everybody probably innately, has abilities for almost any kind of thinking. Symbolic thinking and other kinds of thinking and there are all these things that try to test students for their preferred learning style. I think once you do that you should teach to the one that is least preferred because that's the one that needs developing, instead of doing the other. Well, now there are some folk, like if you read McKim he would say that your visual thinking is the more fundamental mode of thinking. I just think that it is like a lot of other things. You know everybody can play basketball but some people happen to be better than others at it. But basically I think everybody can sort of do it to the point where they have some enjoyment. Music is another thing. I think most people can sing, I don't know how well they might play the violin or the piano. I think those things are things that everybody has the chance to do them well enough so that they can get some type of enjoyment out of it.

J.B..: Since it does take a lot of work to implement this program, what do you recommend for students of lower skills or less enthusiastic? Is the program beneficial for them or does it take even more effort on the part of the teacher to get them involved?

E.M.: I think probably less, but you are better off asking the teachers tomorrow. But actually where the visual math course guides came from, now Linda is not here. Another school district that has been into a lot of this stuff is Madres, several schools over there. Linda Foreman is over there. Actually, in addition to the Mind's Eye units there is the Visual Math Course Guides. The Math in the Mind's Eye units are sort of supplementary, some idea pieces for people. Actually you can take any book and if you take a topic out of it and you can develop some ways to incorporate some visual thinking into that. Pretty much, just with some experience. So you can probably use any book to do some visual thinking in your classroom. What happened back in the beginning of the project, we worked with a couple of school districts and then Linda Foreman, who was a high school teacher and she wasn't really supposed to be involved since it was for middle school teachers. But she asked to be involved and we said fine. She was the chairman of the math department of the school district. And what she did was, she said, well she wasn't having a terrible a lot of success, taught the whole range including what you might call pre-general math class, last chance math class for high school students. So what she decided to do was just toss out everything she was doing and use visual things. She had so much success at that, that she actually developed a course that later became the Visual Math Course Guides. Other people wanted to know about it, so we asked her if she was available if we found some funding to get her develop it and she did. So it seems to work with those kids, and sometimes even more with those than with some kids are having, experiencing success in mathematics in school. The ones who know how to get the right answer and to manipulate the symbols because they have a lot of rewards for doing that. All of a sudden you are asking them to do something different, when they have already had success. But I think a lot of people have success in mathematics courses that are symbol manipulation. Basically they have learned how to swindle their way because you can go through a math class, I've done it, maybe you've done it too, and you know how to get good grades and you get all done and you think, well what in the world was that all about? You just don't have any feeling about it. Some people claim that's always going to happen, the feeling comes later.

J.B..: I've been very frustrated in classes where I kind of had a vague idea of what to do with the symbols but I had no idea why. Especially in a Linear Algebra courses I took.

E.M.: Linear Algebra can be that way and linear algebra to a large extent well all the matrix series gets it origins in geometric transformations. It's a highly geometric topic, but that gets hidden.

J.B.: Last term I had a class in geometric transformations, by Schlomo Libeskind. So it was interesting because it was what I was looking at in linear algebra and I went, wow this is what this comes from.

E.M. I remember my own experiences in that. After going through a linear algebra course I took this statistics course and I had a little bit more of a sense for it. What a multivariable distribution looked like and they had this, I was supposed to normalize this multivariable. So I did some kind of transformation and this stuff and reduced it down to some diagram. And I remember the instructor saying, "but you didn't preserve any distances." It wasn't orthagonal and until then I had taken this thing and going through it, moving it about so it was sitting on the coordinate axis. You know all the way through a linear algebra course I didn't realize what was going on.

J.B.: Do you feel this method can be used in high school or in college, extensively or maybe just in parts? Because the higher you get in math the more difficult it would become.

E.M.: Yes and no. A lot of the origins can be done this way. It would also be nice if mathematicians would talk about whatever kind of images they use. Well you get the topology stuff, Jim Bembesco, you know the knot theory guy. He stands there and talks about, and moving his hands around while he talks about knots doing what-not. Course that is all been written down in a bunch of algebraic symbols. There is something there that somebody has visualized. There's a [he produces a model] model for teaching a calculus course, the derivative of a product. That's one way that you can thing about the product function. There are other ways to think about the product function. You see there is one function here and another function here, so at any point x, why this times this is this cross section. So you can think about the product of these two functions as represented by the area of these cross sections. So then looking at the derivative of a product what you are really looking for what the rate of change is for the cross section where the variables are. You can prepare a couple of them and you get a nice expression for the difference quotient. This is great too as a model and then later on when you are trying to do this integration stuff, you cut things up into slices. People who have built these things don't have any particular problem. As a matter of fact, when you say "now can you draw a sketch or can you can build a model?" and "what does it look like?" You can ask them long before they know about integration, ask them to approximate the volume and many people do it by slabs and they approximate the slabs. So later on when you talk about an integral, which is just taking the sum of a bunch of slabs like that and adding them up and taking more and more of them, why it isn't so abstract. There's a difference between looking at a picture in a book and building a model. I don't know exactly what it is. You can put that on a computer screen and people can look at it, but there seems to be something about giving people pieces of graph paper and some scissors and have them cut out these things and tape them in. It's totally different. I think basically you just develop a different sensory perception.

J.B.: I think the more senses you are able to involve in something the more you own it.

E.M.: Yeah, that's right. You're thinking all the time you are doing that. I mean you're thinking in different ways about what is happening. Pretty hard to look at a computer screen and get the same kind of interaction.

J.B.: I know that there have been some programs done on linear algebra and visualization. I've got an article on it and I haven't had a chance to read it. What I want to do is to figure a way, without a computer screen. It will mainly consist of having people draw a lot of pictures of what is happening, and also thinking of, I'm not really sure yet, but for three dimensions bringing in the axis, because it is really difficult to look at three dimensions on a two dimensional page.

E.M.:Yeah, you could construct it with pieces of paper, like the one in the corner. And you can locate points with string. I think if you just had the person build one model like that and then you had a computer program that could generate what that would look like for different functions, I think you would be OK. It doesn't take much. I don't think you have to do a lot but I think it is important to do it. Then afterwards it would be nice, you know if you wanted to look at others, to have the computer program. But then you would know what was going on. But I don't think you can skip the hands on part of it.

APPENDIX B

152.

A MEMOIR ON THE THEORY OF MATRICES.

[From the *Philosophical Transactions of the Royal Society of London*, vol. CXLVIII. *for the year*, 1858, pp. 17—37. Received December 10, 1857,—Read January 14, 1858.]

THE term matrix might be used in a more general sense, but in the present memoir I consider only square and rectangular matrices, and the term matrix used without qualification is to be understood as meaning a square matrix; in this restricted sense, a set of quantities arranged in the form of a square, e.g.

$$\begin{array}{|ccc|} \hline a\,, & b\,, & c \\ a'\,, & b'\,, & c' \\ a''\,, & b''\,, & c'' \end{array}$$

is said to be a matrix. The notion of such a matrix arises naturally from an abbreviated notation for a set of linear equations, viz. the equations

$$\begin{aligned} X &= ax + by + cz\,, \\ Y &= a'x + b'y + c'z\,, \\ Z &= a''x + b''y + c''z\,, \end{aligned}$$

may be more simply represented by

$$(X,\ Y,\ Z) = \left(\begin{array}{ccc} a\,, & b\,, & c \\ a'\,, & b'\,, & c' \\ a''\,, & b''\,, & c'' \end{array}\right)(x,\ y,\ z),$$

and the consideration of such a system of equations leads to most of the fundamental notions in the theory of matrices. It will be seen that matrices (attending only to those of the same order) comport themselves as single quantities; they may be added, multiplied or compounded together, &c.: the law of the addition of matrices is precisely similar to that for the addition of ordinary algebraical quantities; as regards their multiplication (or composition), there is the peculiarity that matrices are not in general convertible; it is nevertheless possible to form the powers (positive or negative, integral or fractional) of a matrix, and thence to arrive at the notion of a rational and integral function, or generally of any algebraical function, of a matrix. I obtain the remarkable theorem that any matrix whatever satisfies an algebraical equation of its own order, the coefficient of the highest power being unity, and those of the other powers functions of the terms of the matrix, the last coefficient being in fact the determinant; the rule for the formation of this equation may be stated in the following condensed form, which will be intelligible after a perusal of the memoir, viz. the determinant, formed out of the matrix diminished by the matrix considered as a single quantity involving the matrix unity, will be equal to zero. The theorem shows that every rational and integral function (or indeed every rational function) of a matrix may be considered as a rational and integral function, the degree of which is at most equal to that of the matrix, less unity; it even shows that in a sense the same is true with respect to any algebraical function whatever of a matrix. One of the applications of the theorem is the finding of the general expression of the matrices which are convertible with a given matrix. The theory of rectangular matrices appears much less important than that of square matrices, and I have not entered into it further than by showing how some of the notions applicable to these may be extended to rectangular matrices.

1. For conciseness, the matrices written down at full length will in general be of the order 3, but it is to be understood that the definitions, reasonings, and conclusions apply to matrices of any degree whatever. And when two or more matrices are spoken of in connexion with each other, it is always implied (unless the contrary is expressed) that the matrices are of the same order.

2. The notation

$$\left(\begin{array}{ccc} a, & b, & c \\ a', & b', & c' \\ a'', & b'', & c'' \end{array}\right. \between x, y, z)$$

represents the set of linear functions

$$((a, b, c \between x, y, z),\ (a', b', c' \between x, y, z),\ (a'', b'', c'' \between x, y, z)),$$

so that calling these (X, Y, Z), we have

$$(X, Y, Z) = \left(\begin{array}{ccc} a, & b, & c \\ a', & b', & c' \\ a'', & b'', & c'' \end{array}\right. \between x, y, z)$$

and, as remarked above, this formula leads to most of the fundamental notions in the theory.

3. The quantities (X, Y, Z) will be identically zero, if all the terms of the matrix are zero, and we may say that

$$\left(\begin{array}{ccc} 0, & 0, & 0 \\ 0, & 0, & 0 \\ 0, & 0, & 0 \end{array}\right)$$

is the matrix zero.

Again, (X, Y, Z) will be identically equal to (x, y, z), if the matrix is

$$\left(\begin{array}{ccc} 1, & 0, & 0 \\ 0, & 1, & 0 \\ 0, & 0, & 1 \end{array}\right)$$

and this is said to be the matrix unity. We may of course, when for distinctness it is required, say, the matrix zero, or (as the case may be) the matrix unity *of such an order*. The matrix zero may for the most part be represented simply by 0, and the matrix unity by 1.

4. The equations

$$(X, Y, Z) = \left(\begin{array}{ccc} a, & b, & c \\ a', & b', & c' \\ a'', & b'', & c'' \end{array}\right)(x, y, z), \quad (X', Y', Z') = \left(\begin{array}{ccc} \alpha, & \beta, & \gamma \\ \alpha', & \beta', & \gamma' \\ \alpha'', & \beta'', & \gamma'' \end{array}\right)(x, y, z)$$

give

$$(X+X', Y+Y', Z+Z') = \left(\begin{array}{ccc} a+\alpha, & b+\beta, & c+\gamma \\ a'+\alpha', & b'+\beta', & c'+\gamma' \\ a''+\alpha'', & b''+\beta'', & c''+\gamma'' \end{array}\right)(x, y, z)$$

and this leads to

$$\left(\begin{array}{ccc} a+\alpha, & b+\beta, & c+\gamma \\ a'+\alpha', & b'+\beta', & c'+\gamma' \\ a''+\alpha'', & b''+\beta'', & c''+\gamma'' \end{array}\right) = \left(\begin{array}{ccc} a, & b, & c \\ a', & b', & c' \\ a'', & b'', & c'' \end{array}\right) + \left(\begin{array}{ccc} \alpha, & \beta, & \gamma \\ \alpha', & \beta', & \gamma' \\ \alpha'', & \beta'', & \gamma'' \end{array}\right)$$

as a rule for the addition of matrices; that for their subtraction is of course similar to it.

5. A matrix is not altered by the addition or subtraction of the matrix zero, that is, we have $M \pm 0 = M$.

The equation $L = M$, which expresses that the matrices L, M are equal, may also be written in the form $L - M = 0$, i.e. the difference of two equal matrices is the matrix zero.

6. The equation $L = -M$, written in the form $L + M = 0$, expresses that the sum of the matrices L, M is equal to the matrix zero, the matrices so related are said to be *opposite* to each other; in other words, a matrix the terms of which are equal but opposite in sign to the terms of a given matrix, is said to be opposite to the given matrix.

7. It is clear that we have $L+M=M+L$, that is, the operation of addition is commutative, and moreover that $(L+M)+N=L+(M+N)=L+M+N$, that is, the operation of addition is also associative.

8. The equation

$$(X,\ Y,\ Z)=\left(\begin{array}{ccc} a\ , & b\ , & c \\ a'\ , & b'\ , & c' \\ a'', & b'', & c'' \end{array}\right)(mx,\ my,\ mz)$$

written under the forms

$$(X,\ Y,\ Z)=m\left(\begin{array}{ccc} a\ , & b\ , & c \\ a'\ , & b'\ , & c' \\ a'', & b'', & c'' \end{array}\right)(x,\ y,\ z)=\left(\begin{array}{ccc} ma\ , & mb\ , & mc \\ ma'\ , & mb'\ , & mc' \\ ma'', & mb'', & mc'' \end{array}\right)(x,\ y,\ z)$$

gives

$$m\left(\begin{array}{ccc} a\ , & b\ , & c \\ a'\ , & b'\ , & c' \\ a'', & b'', & c'' \end{array}\right)=\left(\begin{array}{ccc} ma\ , & mb\ , & mc \\ ma'\ , & mb'\ , & mc' \\ ma'', & mb'', & mc'' \end{array}\right)$$

as the rule for the multiplication of a matrix by a single quantity. The multiplier m may be written either before or after the matrix, and the operation is therefore commutative. We have it is clear $m(L+M)=mL+mM$, or the operation is distributive.

9. The matrices L and mL may be said to be similar to each other: in particular, if $m=1$, they are equal, and if $m=-1$, they are opposite.

10. We have, in particular,

$$m\left(\begin{array}{ccc} 1, & 0, & 0 \\ 0, & 1, & 0 \\ 0, & 0, & 1 \end{array}\right)=\left(\begin{array}{ccc} m, & 0, & 0 \\ 0, & m, & 0 \\ 0, & 0, & m \end{array}\right),$$

or replacing the matrix on the left-hand side by unity, we may write

$$m=\left(\begin{array}{ccc} m, & 0, & 0 \\ 0, & m, & 0 \\ 0, & 0, & m \end{array}\right);$$

the matrix on the right-hand side is said to be the single quantity m considered as *involving the matrix unity.*

11. The equations

$$(X,\ Y,\ Z)=\left(\begin{array}{ccc} a\ , & b\ , & c \\ a'\ , & b'\ , & c' \\ a'', & b'', & c'' \end{array}\right)(x,\ y,\ z),\quad (x,\ y,\ z)=\left(\begin{array}{ccc} \alpha\ , & \beta\ , & \gamma \\ \alpha'\ , & \beta'\ , & \gamma' \\ \alpha'', & \beta'', & \gamma'' \end{array}\right)(\xi,\ \eta,\ \zeta),$$

give

$$(X,\ Y,\ Z) = \begin{pmatrix} A\ , & B\ , & C \\ A'\ , & B'\ , & C' \\ A''\ , & B''\ , & C'' \end{pmatrix} \between (\xi,\ \eta,\ \zeta) = \begin{pmatrix} a\ , & b\ , & c \\ a'\ , & b'\ , & c' \\ a''\ , & b''\ , & c'' \end{pmatrix} \between \begin{pmatrix} \alpha\ , & \beta\ , & \gamma \\ \alpha'\ , & \beta'\ , & \gamma' \\ \alpha''\ , & \beta''\ , & \gamma'' \end{pmatrix} \between (\xi,\ \eta,\ \zeta),$$

and thence, substituting for the matrix

$$\begin{pmatrix} A\ , & B\ , & C \\ A'\ , & B'\ , & C' \\ A''\ , & B''\ , & C'' \end{pmatrix}$$

its value, we obtain

$$\begin{pmatrix} (a\ ,b\ ,c \between \alpha,\alpha',\alpha''), & (a\ ,b\ ,c \between \beta,\beta',\beta''), & (a\ ,b\ ,c \between \gamma,\gamma',\gamma'') \\ (a',b',c' \between \alpha,\alpha',\alpha''), & (a',b',c' \between \beta,\beta',\beta''), & (a',b',c' \between \gamma,\gamma',\gamma'') \\ (a'',b'',c'' \between \alpha,\alpha',\alpha''), & (a'',b'',c'' \between \beta,\beta',\beta''), & (a'',b'',c'' \between \gamma,\gamma',\gamma'') \end{pmatrix} = \begin{pmatrix} a\ ,b\ ,c \\ a',b',c' \\ a'',b'',c'' \end{pmatrix} \between \begin{pmatrix} \alpha\ ,\beta\ ,\gamma \\ \alpha',\beta',\gamma' \\ \alpha'',\beta'',\gamma'' \end{pmatrix}$$

as the rule for the multiplication or composition of two matrices. It is to be observed, that the operation is not a commutative one; the component matrices may be distinguished as the first or further component matrix, and the second or nearer component matrix, and the rule of composition is as follows, viz. any *line* of the compound matrix is obtained by combining the corresponding *line* of the first or further component matrix successively with the several *columns* of the second or nearer compound matrix.

[We may conveniently write

$$\begin{array}{c|ccc} & (\alpha,\ \alpha',\ \alpha''), & (\beta,\ \beta',\ \beta''), & (\gamma,\ \gamma',\ \gamma'') \\ \hline (a\ ,\ \ b\ ,\ \ c\ ) & ,, & ,, & ,, \\ (a',\ \ b',\ \ c') & ,, & ,, & ,, \\ (a'',\ \ b'',\ \ c'') & ,, & ,, & ,, \end{array}$$

to denote the left-hand side of the last preceding equation.]

12. A matrix compounded, either as first or second component matrix, with the matrix zero, gives the matrix zero. The case where any of the terms of the given matrix are infinite is of course excluded.

13. A matrix is not altered by its composition, either as first or second component matrix, with the matrix unity. It is compounded either as first or second component matrix, with the single quantity m considered as involving the matrix unity, by multiplication of all its terms by the quantity m: this is in fact the before-mentioned rule for the multiplication of a matrix by a single quantity, which rule is thus seen to be a particular case of that for the multiplication of two matrices.

14. We may in like manner multiply or compound together three or more matrices: the order of arrangement of the factors is of course material, and we may

distinguish them as the first or furthest, second, third, &c., and last or nearest component matrices: any two consecutive factors may be compounded together and replaced by a single matrix, and so on until all the matrices are compounded together, the result being independent of the particular mode in which the composition is effected; that is, we have $L\,.\,MN = LM\,.\,N = LMN$, $LM\,.\,NP = L\,.\,MN\,.\,P$, &c., or the operation of multiplication, although, as already remarked, not commutative, is associative.

15. We thus arrive at the notion of a positive and integer power L^p of a matrix L, and it is to be observed that the different powers of the same matrix are convertible. It is clear also that p and q being positive integers, we have $L^p\,.\,L^q = L^{p+q}$, which is the theorem of indices for positive integer powers of a matrix.

16. The last-mentioned equation, $L^p\,.\,L^q = L^{p+q}$, assumed to be true for all values whatever of the indices p and q, leads to the notion of the powers of a matrix for any form whatever of the index. In particular, $L^p\,.\,L^0 = L^p$ or $L^0 = 1$, that is, the 0th power of a matrix is the matrix unity. And then putting $p=1$, $q=-1$, or $p=-1$, $q=1$, we have $L\,.\,L^{-1} = L^{-1}\,.\,L = 1$; that is, L^{-1}, or as it may be termed the inverse or reciprocal matrix, is a matrix which, compounded either as first or second component matrix with the original matrix, gives the matrix unity.

17. We may arrive at the notion of the inverse or reciprocal matrix, directly from the equation

$$(X,\ Y,\ Z) = \left(\begin{array}{ccc} a, & b, & c \\ a', & b', & c' \\ a'', & b'', & c'' \end{array}\right)(x,\ y,\ z),$$

in fact this equation gives

$$(x,\ y,\ z) = \left(\begin{array}{ccc} A, & A', & A'' \\ B, & B', & B'' \\ C, & C', & C'' \end{array}\right)(X,\ Y,\ Z) = \left(\left(\begin{array}{ccc} a, & b, & c \\ a', & b', & c' \\ a'', & b'', & c'' \end{array}\right)^{-1}\right)(X,\ Y,\ Z),$$

and we have, for the determination of the coefficients of the inverse or reciprocal matrix, the equations

$$\left(\begin{array}{ccc} A, & A', & A'' \\ B, & B', & B'' \\ C, & C', & C'' \end{array}\right)\left(\begin{array}{ccc} a, & b, & c \\ a', & b', & c' \\ a'', & b'', & c'' \end{array}\right) = \left(\begin{array}{ccc} 1, & 0, & 0 \\ 0, & 1, & 0 \\ 0, & 0, & 1 \end{array}\right),$$

$$\left(\begin{array}{ccc} a, & b, & c \\ a', & b', & c' \\ a'', & b'', & c'' \end{array}\right)\left(\begin{array}{ccc} A, & A', & A'' \\ B, & B', & B'' \\ C, & C', & C'' \end{array}\right) = \left(\begin{array}{ccc} 1, & 0, & 0 \\ 0, & 1, & 0 \\ 0, & 0, & 1 \end{array}\right),$$

which are equivalent to each other, and either of them is by itself sufficient for the complete determination of the inverse or reciprocal matrix. It is well known that if ∇ denote the determinant, that is, if

$$\nabla = \begin{vmatrix} a, & b, & c \\ a', & b', & c' \\ a'', & b'', & c'' \end{vmatrix}$$

then the terms of the inverse or reciprocal matrix are given by the equations

$$A = \frac{1}{\nabla}\begin{vmatrix} 1, & 0, & 0 \\ 0, & b', & c' \\ 0, & b'', & c'' \end{vmatrix}, \qquad B = \frac{1}{\nabla}\begin{vmatrix} 0, & 1, & 0 \\ a', & 0, & c' \\ a'', & 0, & c'' \end{vmatrix}, \text{ \&c.}$$

or what is the same thing, the inverse or reciprocal matrix is given by the equation

$$\begin{pmatrix} a, & b, & c \\ a', & b', & c' \\ a'', & b'', & c'' \end{pmatrix}^{-1} = \frac{1}{\nabla}\begin{pmatrix} \partial_a\nabla, & \partial_{a'}\nabla, & \partial_{a''}\nabla \\ \partial_b\nabla, & \partial_{b'}\nabla, & \partial_{b''}\nabla \\ \partial_c\nabla, & \partial_{c'}\nabla, & \partial_{c''}\nabla \end{pmatrix}$$

where of course the differentiations must in every case be performed as if the terms a, b, &c. were all of them independent arbitrary quantities.

18. The formula shows, what is indeed clear *à priori*, that the notion of the inverse or reciprocal matrix fails altogether when the determinant vanishes: the matrix is in this case said to be indeterminate, and it must be understood that in the absence of express mention, the particular case in question is frequently excluded from consideration. It may be added that the matrix zero is indeterminate; and that the product of two matrices may be zero, without either of the factors being zero, if only the matrices are one or both of them indeterminate.

19. The notion of the inverse or reciprocal matrix once established, the other negative integer powers of the original matrix are positive integer powers of the inverse or reciprocal matrix, and the theory of such negative integer powers may be taken to be known. The theory of the fractional powers of a matrix will be further discussed in the sequel.

20. The positive integer power L^m of the matrix L may of course be multiplied by any matrix of the same degree: such multiplier, however, is not in general convertible with L; and to preserve as far as possible the analogy with ordinary algebraical functions, we may restrict the attention to the case where the multiplier is a single quantity, and such convertibility consequently exists. We have in this manner a matrix cL^m, and by the addition of any number of such terms we obtain a rational and integral function of the matrix L.

21. The general theorem before referred to will be best understood by a complete development of a particular case. Imagine a matrix

$$M = \begin{pmatrix} a, & b \\ c, & d \end{pmatrix},$$

and form the determinant

$$\begin{vmatrix} a - M, & b \\ c & , d - M \end{vmatrix},$$

the developed expression of this determinant is

$$M^2 - (a+d) M^1 + (ad - bc) M^0 ;$$

the values of M^2, M^1, M^0 are

$$\begin{pmatrix} a^2 + bc & , b(a+d) \\ c(a+d), & d^2 + bc \end{pmatrix}, \quad \begin{pmatrix} a, & b \\ c, & d \end{pmatrix}, \quad \begin{pmatrix} 1, & 0 \\ 0, & 1 \end{pmatrix},$$

and substituting these values the determinant becomes equal to the matrix zero, viz. we have

$$\begin{vmatrix} a - M, & b \\ c & , d - M \end{vmatrix} = \begin{pmatrix} a^2 + bc & , b(a+d) \\ c(a+d), & d^2 + bc \end{pmatrix} - (a+d) \begin{pmatrix} a, & b \\ c, & d \end{pmatrix} + (ad - bc) \begin{pmatrix} 1, & 0 \\ 0, & 1 \end{pmatrix}$$

$$= \begin{pmatrix} (a^2 + bc) - (a+d)a + (ad - bc), & b(a+d) - (a+d)b \\ c(a+d) - (a+d)c & , d^2 + bc - (a+d)d + ad - bc \end{pmatrix} = \begin{pmatrix} 0, & 0 \\ 0, & 0 \end{pmatrix};$$

that is

$$\begin{vmatrix} a - M, & b \\ c & , d - M \end{vmatrix} = 0,$$

where the matrix of the determinant is

$$\begin{pmatrix} a, & b \\ c, & d \end{pmatrix} - M \begin{pmatrix} 1, & 0 \\ 0, & 1 \end{pmatrix},$$

that is, it is the original matrix, diminished by the same matrix considered as a single quantity involving the matrix unity. And this is the general theorem, viz. the determinant, having for its matrix a given matrix less the same matrix considered as a single quantity involving the matrix unity, is equal to zero.

22. The following symbolical representation of the theorem is, I think, worth noticing: let the matrix M, considered as a single quantity, be represented by $\bar{M}$, then writing 1 to denote the matrix unity, $\bar{M}.1$ will represent the matrix M, considered as a single quantity involving the matrix unity. Upon the like principles of notation, $\bar{1}.M$ will represent, or may be considered as representing, simply the matrix M, and the theorem is

$$\text{Det.}\ (\bar{1}.M - \bar{M}.1) = 0.$$

23. I have verified the theorem, in the next simplest case of a matrix of the order 3, viz. if M be such a matrix, suppose

$$M = \begin{array}{c} (\; a, \;\; b, \;\; c \;), \\ \left|\; d, \;\; e, \;\; f \;\right| \\ \left|\; g, \;\; h, \;\; i \;\right| \end{array}$$

then the derived determinant vanishes, or we have

$$\begin{vmatrix} a-M, & b & , c \\ d & , e-M, & f \\ g & , h & , i-M \end{vmatrix} = 0,$$

or expanding

$$M^3 - (a+e+i)\, M^2 + (ei + ia + ae - fh - cg - bd)\, M - (aei + bfg + cdh - afh - bdi - ceg) = 0\,;$$

but I have not thought it necessary to undertake the labour of a formal proof of the theorem in the general case of a matrix of any degree.

24. If we attend only to the general form of the result, we see that any matrix whatever satisfies an algebraical equation of its own order, which is in many cases the material part of the theorem.

25. It follows at once that every rational and integral function, or indeed every rational function of a matrix, can be expressed as a rational and integral function of an order at most equal to that of the matrix, less unity. But it is important to consider how far or in what sense the like theorem is true with respect to irrational functions of a matrix. If we had only the equation satisfied by the matrix itself, such extension could not be made; but we have besides the equation of the same order satisfied by the irrational function of the matrix, and by means of these two equations, and the equation by which the irrational function of the matrix is determined, we may express the irrational function as a rational and integral function of the matrix, of an order equal at most to that of the matrix, less unity; such expression will however involve *the coefficients of the equation satisfied by the irrational function*, which are functions (in number equal to the order of the matrix) of the terms, assumed to be unknown, of the irrational function itself. The transformation is nevertheless an important one, as reducing the number of unknown quantities from n^2 (if n be the order of the matrix) down to n. To complete the solution, it is necessary to compare the value obtained as above, with the assumed value of the irrational function, which will lead to equations for the determination of the n unknown quantities.

26. As an illustration, consider the given matrix

$$M = \begin{array}{c} (\; a, \;\; b \;), \\ \left|\; c, \;\; d \;\right| \end{array}$$

and let it be required to find the matrix $L = \sqrt{M}$. In this case M satisfies the equation,

$$M^2 - (a+d)M + ad - bc = 0;$$

and in like manner if

$$L = \begin{pmatrix} \alpha, & \beta \\ \gamma, & \delta \end{pmatrix}$$

then L satisfies the equation

$$L^2 - (\alpha + \delta) L + \alpha\delta - \beta\gamma = 0;$$

and from these two equations, and the rationalized equation $L^2 = M$, it should be possible to express L in the form of a linear function of M: in fact, putting in the last equation for L^2 its value $(= M)$, we find at once

$$L = \frac{1}{\alpha + \delta}[M + (\alpha\delta - \beta\gamma)],$$

which is the required expression, involving as it should do the coefficients $\alpha + \delta$, $\alpha\delta - \beta\gamma$ of the equation in L. There is no difficulty in completing the solution; write for shortness $\alpha + \delta = X$, $\alpha\delta - \beta\gamma = Y$, then we have

$$L = \begin{pmatrix} \alpha, & \beta \\ \gamma, & \delta \end{pmatrix} = \begin{pmatrix} \frac{a+Y}{X}, & \frac{b}{X} \\ \frac{c}{X}, & \frac{d+Y}{X} \end{pmatrix},$$

and consequently forming the values of $\alpha + \delta$ and $\alpha\delta - \beta\gamma$,

$$X = \frac{a + d + 2Y}{X},$$

$$Y = \frac{(a+Y)(d+Y) - bc}{X^2},$$

and putting also $a + d = P$, $ad - bc = Q$, we find without difficulty

$$X = \sqrt{P + 2\sqrt{Q}},$$

$$Y = \sqrt{Q},$$

and the values of α, β, γ, δ are consequently known. The sign of $\sqrt{Q}$ is the same in both formulæ, and there are consequently in all four solutions, that is, the radical $\sqrt{M}$ has four values.

27. To illustrate this further, suppose that instead of M we have the matrix

$$M^2 = \begin{pmatrix} a, & b \\ c, & d \end{pmatrix}^2 = \begin{pmatrix} a^2 + bc, & b(a+d) \\ c(a+d), & d^2 + bc \end{pmatrix},$$

so that $L^2 = M^2$, we find

$$P = (a+d)^2 - 2(ad - bc),$$
$$Q = (ad - bc)^2,$$

and thence $\sqrt{Q} = \pm (ad - bc)$. Taking the positive sign, we have

$$Y = \quad ad - bc,$$
$$X = \pm (a + d),$$

and these values give simply

$$L = \pm \begin{pmatrix} a, & b \\ c, & d \end{pmatrix} = \pm M.$$

But taking the negative sign,

$$Y = -ad + bc,$$
$$X = \pm \sqrt{(a-d)^2 + 4bc},$$

and retaining X to denote this radical, we find

$$L = \begin{pmatrix} \dfrac{a^2 - ad + 2bc}{X}, & \dfrac{b(a+d)}{X} \\ \dfrac{c(a+d)}{X}, & \dfrac{d^2 - ad + 2bc}{X} \end{pmatrix},$$

which may also be written

$$L = \frac{a+d}{X} \begin{pmatrix} a, & b \\ c, & d \end{pmatrix} - \frac{2(ad - bc)}{X} \begin{pmatrix} 1, & 0 \\ 0, & 1 \end{pmatrix},$$

or, what is the same thing,

$$L = \frac{a+d}{X} M - \frac{2(ad-bc)}{X};$$

and it is easy to verify *à posteriori* that this value in fact gives $L^2 = M^2$. It may be remarked that if

$$M^2 = \begin{pmatrix} 1, & 0 \\ 0, & 1 \end{pmatrix}^2 = 1,$$

the last-mentioned formula fails, for we have $X = 0$; it will be seen presently that the equation $L^2 = 1$ admits of other solutions besides $L = \pm 1$. The example shows how the values of the fractional powers of a matrix are to be investigated.

28. There is an apparent difficulty connected with the equation satisfied by a matrix, which it is proper to explain. Suppose, as before,

$$M = \begin{pmatrix} a, & b \\ c, & d \end{pmatrix},$$

so that M satisfies the equation

$$\begin{vmatrix} a-M, & b \\ c \quad , & d-M \end{vmatrix} = 0,$$

or

$$M^2-(a+d)M+ad-bc=0,$$

and let $X_{,}$, $X_{,,}$ be the single quantities, roots of the equation

$$\begin{vmatrix} a-X, & b \\ c \quad , & d-X \end{vmatrix} = 0$$

or

$$X^2-(a+d)X+ad-bc=0.$$

The equation satisfied by the matrix may be written

$$(M-X_{,})(M-X_{,,})=0,$$

in which $X_{,}$, $X_{,,}$ are to be considered as respectively involving the matrix unity, and it would at first sight seem that we ought to have one of the simple factors equal to zero; this is obviously not the case, for such equation would signify that the perfectly indeterminate matrix M was equal to a single quantity, considered as involving the matrix unity. The explanation is that each of the simple factors is an indeterminate matrix, in fact $M-X_{,}$ stands for the matrix

$$\begin{pmatrix} a-X_{,}, & b \\ c \quad , & d-X_{,} \end{pmatrix},$$

and the determinant of this matrix is equal to zero. The product of the two factors is thus equal to zero without either of the factors being equal to zero.

29. A matrix satisfies, we have seen, an equation of its own order, involving the coefficients of the matrix; assume that the matrix is to be determined to satisfy some other equation, the coefficients of which are given single quantities. It would at first sight appear that we might eliminate the matrix between the two equations, and thus obtain an equation which would be the only condition to be satisfied by the terms of the matrix; this is obviously wrong, for more conditions must be requisite, and we see that if we were then to proceed to complete the solution by finding the value of the matrix common to the two equations, we should find the matrix equal in every case to a single quantity considered as involving the matrix unity, which it is clear ought not to be the case. The explanation is similar to that of the difficulty before adverted to; the equations may contain one, and only one, common factor, and may be both of them satisfied, and yet the common factor may not vanish. The necessary condition seems to be, that the one equation should be a factor of the other; in the case where the assumed equation is of an order equal or superior to the matrix, then if the equation contain as a factor the equation which is always satisfied by the matrix, the assumed equation will be satisfied identically, and the condition is sufficient as well as necessary: in the other case, where the assumed equation is of an order inferior to that of the matrix, the condition is necessary, but it is not sufficient.

30. The equation satisfied by the matrix may be of the form $M^n = 1$; the matrix is in this case said to be periodic of the nth order. The preceding considerations apply to the theory of periodic matrices; thus, for instance, suppose it is required to find a matrix of the order 2, which is periodic of the second order. Writing

$$M = \begin{pmatrix} a, & b \\ c, & d \end{pmatrix},$$

we have

$$M^2 - (a+d)M + ad - bc = 0,$$

and the assumed equation is

$$M^2 - 1 = 0.$$

These equations will be identical if

$$a + d = 0, \quad ad - bc = -1,$$

that is, these conditions being satisfied, the equation $M^2 - 1 = 0$ required to be satisfied, will be identical with the equation which is always satisfied, and will therefore itself be satisfied. And in like manner the matrix M of the order 2 will satisfy the condition $M^3 - 1 = 0$, or will be periodic of the third order, if only $M^3 - 1$ contains as a factor

$$M^2 - (a+d)M + ad - bc,$$

and so on.

31. But suppose it is required to find a matrix of the order 3,

$$M = \begin{pmatrix} a, & b, & c \\ d, & e, & f \\ g, & h, & i \end{pmatrix}$$

which shall be periodic of the second order. Writing for shortness

$$\begin{vmatrix} a-M, & b &, c \\ d &, e-M, & f \\ g &, h &, i-M \end{vmatrix} = -(M^3 - AM^2 + BM - C),$$

the matrix here satisfies

$$M^3 - AM^2 + BM - C = 0,$$

and, as before, the assumed equation is $M^2 - 1 = 0$. Here, if we have $1 + B = 0$, $A + C = 0$, the left-hand side will contain the factor $(M^2 - 1)$, and the equation will take the form $(M^2 - 1)(M + C) = 0$, and we should have then $M^2 - 1 = 0$, provided $M + C$ were not an indeterminate matrix. But $M + C$ denotes the matrix

$$\begin{pmatrix} a+C, & b &, c \\ d &, e+C, & f \\ g &, h &, i+C \end{pmatrix}$$

the determinant of which is C^3+AC^2+BC+C, which is equal to zero in virtue of the equations $1+B=0$, $A+C=0$, and we cannot, therefore, from the equation $(M^2-1)(M+C)=0$, deduce the equation $M^2-1=0$. This is as it should be, for the two conditions are not sufficient, in fact the equation

$$M^2 = \left(\begin{array}{lll} a^2+bd+cg, & ab+be+ch, & ac+bf+ci \\ da+ed+fg, & db+e^2+fh, & dc+ef+fi \\ ga+hd+ig, & gb+he+ih, & gc+hf+i^2 \end{array}\right) = 1$$

gives nine equations, which are however satisfied by the following values, involving in reality four arbitrary coefficients; viz. the value of the matrix is

$$\left(\begin{array}{lll} \dfrac{\alpha}{\alpha+\beta+\gamma}, & \dfrac{-(\beta+\gamma)\,\nu\mu^{-1}}{\alpha+\beta+\gamma}, & \dfrac{-(\beta+\gamma)\,\nu\mu^{-1}}{\alpha+\beta+\gamma} \\ \dfrac{-(\gamma+\alpha)\,\mu\nu^{-1}}{\alpha+\beta+\gamma}, & \dfrac{\beta}{\alpha+\beta+\gamma}, & \dfrac{-(\gamma+\alpha)\,\lambda\mu^{-1}}{\alpha+\beta+\gamma} \\ \dfrac{-(\alpha+\beta)\,\mu\nu^{-1}}{\alpha+\beta+\gamma}, & \dfrac{-(\alpha+\beta)\,\nu\lambda^{-1}}{\alpha+\beta+\gamma}, & \dfrac{\gamma}{\alpha+\beta+\gamma} \end{array}\right)$$

so that there are in all five relations (and not only two) between the coefficients of the matrix.

32. Instead of the equation $M^n-1=0$, which belongs to a periodic matrix, it is in many cases more convenient, and it is much the same thing to consider the equation $M^n-k=0$, where k is a single quantity. The matrix may in this case be said to be periodic to a factor *près*.

33. Two matrices L, M are convertible when $LM=ML$. If the matrix M is given, this equality affords a set of linear equations between the coefficients of L equal in number to these coefficients, but these equations cannot be all independent, for it is clear that if L be any rational and integral function of M (the coefficients being single quantities), then L will be convertible with M; or what is apparently (but only apparently) more general, if L be any algebraical function whatever of M (the coefficients being always single quantities), then L will be convertible with M. But whatever the form of the function is, it may be reduced to a rational and integral function of an order equal to that of M, less unity, and we have thus the general expression for the matrices convertible with a given matrix, viz. any such matrix is a rational and integral function (the coefficients being single quantities) of the given matrix, the order being that of the given matrix, less unity. In particular, the general form of the matrix L convertible with a given matrix M of the order 2, is $L=\alpha M+\beta$, or what is the same thing, the matrices

$$\left(\begin{array}{ll} a, & b \\ c, & d \end{array}\right), \quad \left(\begin{array}{ll} a', & b' \\ c', & d' \end{array}\right)$$

will be convertible if $a'-d' : b' : c' = a-d : b : c$.

34. Two matrices L, M are skew convertible when $LM = -ML$; this is a relation much less important than ordinary convertibility, for it is to be noticed that we cannot in general find a matrix L skew convertible with a given matrix M. In fact, considering M as given, the equality affords a set of linear equations between the coefficients of L equal in number to these coefficients; and in this case the equations are independent, and we may eliminate all the coefficients of L, and we thus arrive at a relation which must be satisfied by the coefficients of the given matrix M. Thus, suppose the matrices

$$\begin{pmatrix} a, & b \\ c, & d \end{pmatrix}, \quad \begin{pmatrix} a', & b' \\ c', & d' \end{pmatrix}$$

are skew convertible, we have

$$\begin{pmatrix} a, & b \\ c, & d \end{pmatrix}\begin{pmatrix} a', & b' \\ c', & d' \end{pmatrix} = \begin{pmatrix} aa' + bc', & ab' + bd' \\ ca' + dc', & cb' + dd' \end{pmatrix},$$

$$\begin{pmatrix} a', & b' \\ c', & d' \end{pmatrix}\begin{pmatrix} a, & b \\ c, & d \end{pmatrix} = \begin{pmatrix} aa' + b'c, & a'b + b'd \\ c'a + d'c, & c'b + d'd \end{pmatrix},$$

and the conditions of skew convertibility are

$$\begin{aligned} 2aa' + bc' + b'c &= 0, \\ b'(a+d) + b(a'+d') &= 0, \\ c'(a+d) + c(a'+d') &= 0, \\ 2dd' + bc' + b'c &= 0. \end{aligned}$$

Eliminating a', b', c', d', the relation between a, b, c, d is

$$\begin{vmatrix} 2a, & c, & b, & . \\ b, & a+d, & ., & b \\ c, & ., & a+d, & c \\ ., & c, & b, & 2d \end{vmatrix} = 0,$$

which is

$$(a+d)^2 (ad - bc) = 0.$$

Excluding from consideration the case $ad - bc = 0$, which would imply that the matrix was indeterminate, we have $a + d = 0$. The resulting system of conditions then is

$$a + d = 0, \quad a' + d' = 0, \quad aa' + bc' + b'c + dd' = 0,$$

the first two of which imply that the matrices are respectively periodic of the second order to a factor *près*.

35. It may be noticed that if the compound matrices LM and ML are similar, they are either equal or else opposite; that is, the matrices L, M are either convertible or skew convertible.

36. Two matrices such as

$$\begin{array}{|cc|} (\,a, & b\,), \\ c, & d \end{array} \qquad \begin{array}{|cc|} (\,a, & c\,), \\ b, & d \end{array}$$

are said to be formed one from the other by transposition, and this may be denoted by the symbol tr.; thus we may write

$$\begin{array}{|cc|} (\,a, & c\,) \\ b, & d \end{array} = \text{tr.} \begin{array}{|cc|} (\,a, & b\,). \\ c, & d \end{array}$$

The effect of two successive transpositions is of course to reproduce the original matrix.

37. It is easy to see that if M be any matrix, then

$$(\text{tr. } M)^p = \text{tr. } (M^p),$$

and in particular,

$$(\text{tr. } M)^{-1} = \text{tr. } (M^{-1}).$$

38. If L, M be any two matrices,

$$\text{tr. } (LM) = \text{tr. } M. \ \text{tr. } L,$$

and similarly for three or more matrices, L, M, N, &c.,

$$\text{tr. } (LMN) = \text{tr. } N. \ \text{tr. } M. \ \text{tr. } L, \ \&\text{c.}$$

40. A matrix such as

$$\begin{array}{|ccc|} (\,a, & h, & g\,) \\ h, & b, & f \\ g, & f, & c \end{array}$$

which is not altered by transposition, is said to be symmetrical.

41. A matrix such as

$$\begin{array}{|ccc|} (\quad 0, & \nu, & -\mu\,) \\ -\nu, & 0, & \lambda \\ \mu, & -\lambda, & 0 \end{array}$$

which by transposition is changed into its opposite, is said to be skew symmetrical.

42. It is easy to see that any matrix whatever may be expressed as the sum of a symmetrical matrix, and a skew symmetrical matrix; thus the form

$$\begin{array}{|ccc|} (\quad a\;, & h+\nu, & g-\mu\,) \\ h-\nu, & b\;, & f+\lambda \\ g+\mu, & f-\lambda, & c \end{array}$$

which may obviously represent any matrix whatever of the order 3, is the sum of the two matrices last before mentioned.

43. The following formulæ, although little more than examples of the composition of transposed matrices, may be noticed, viz.

$$\left(\begin{array}{cc} a, & b \\ c, & d \end{array}\right) \between \left(\begin{array}{cc} a, & c \\ d, & b \end{array}\right) = \left(\begin{array}{cc} a^2+b^2, & ac+bd \\ ac+bd, & c^2+d^2 \end{array}\right)$$

which shows that a matrix compounded with the transposed matrix gives rise to a symmetrical matrix. It does not however follow, nor is it the fact, that the matrix and transposed matrix are convertible. And also

$$\left(\begin{array}{cc} a, & c \\ b, & d \end{array}\right) \between \left(\begin{array}{cc} a, & b \\ c, & d \end{array}\right) \between \left(\begin{array}{cc} a, & c \\ b, & d \end{array}\right) = \left(\begin{array}{cc} a^3+bcd+a(b^2+c^2), & c^3+abd+c(a^2+d^2) \\ b^3+acd+b(a^2+d^2), & d^3+abc+d(b^2+c^2) \end{array}\right)$$

which is a remarkably symmetrical form. It is needless to proceed further, since it is clear that

$$\left(\begin{array}{cc} a, & c \\ b, & d \end{array}\right) \between \left(\begin{array}{cc} a, & b \\ c, & d \end{array}\right) \between \left(\begin{array}{cc} a, & c \\ b, & d \end{array}\right) \between \left(\begin{array}{cc} a, & b \\ c, & d \end{array}\right) = \left(\left(\begin{array}{cc} a, & c \\ b, & d \end{array}\right) \between \left(\begin{array}{cc} a, & b \\ c, & d \end{array}\right)\right)^2.$$

44. In all that precedes, the matrix of the order 2 has frequently been considered, but chiefly by way of illustration of the general theory; but it is worth while to develope more particularly the theory of such matrix. I call to mind the fundamental properties which have been obtained, viz. it was shown that the matrix

$$M = \left(\begin{array}{cc} a, & b \\ c, & d \end{array}\right),$$

satisfies the equation

$$M^2 - (a+d)M + ad - bc = 0,$$

and that the two matrices

$$\left(\begin{array}{cc} a, & b \\ c, & d \end{array}\right), \quad \left(\begin{array}{cc} a', & b' \\ c', & d' \end{array}\right),$$

will be convertible if

$$a' - d' : b' : c' = a - d : b : c,$$

and that they will be skew convertible if

$$a + d = 0, \quad a' + d' = 0, \quad aa' + bc' + b'c + dd' = 0,$$

the first two of these equations being the conditions in order that the two matrices may be respectively periodic of the second order to a factor *près*.

45. It may be noticed in passing, that if L, M are skew convertible matrices of the order 2, and if these matrices are also such that $L^2 = -1$, $M^2 = -1$, then putting $N = LM = -ML$, we obtain

$$L^2 = -1, \qquad M^2 = -1, \qquad N^2 = -1,$$
$$L = MN = -NM, \quad M = NL = -NL, \quad N = LM = -ML,$$

which is a system of relations precisely similar to that in the theory of quaternions.

46. The integer powers of the matrix

$$M = \begin{pmatrix} a, & b \\ c, & d \end{pmatrix},$$

are obtained with great facility from the quadratic equation; thus we have, attending first to the positive powers,

$$\begin{aligned} M^2 &= (a+d)\,M - (ad-bc), \\ M^3 &= [(a+d)^2 - (ad-bc)]\,M - (a+d)(ad-bc), \\ &\text{\&c.}, \end{aligned}$$

whence also the conditions in order that the matrix may be to a factor *près* periodic of the orders 2, 3, &c. are

$$\begin{aligned} a + d \qquad\qquad &= 0, \\ (a+d)^2 - (ad-bc) &= 0, \\ &\text{\&c.}; \end{aligned}$$

and for the negative powers we have

$$(ad-bc)\,M^{-1} = -\,M + (a+d),$$

which is equivalent to the ordinary form

$$(ad-bc)\,M^{-1} = \begin{pmatrix} d, & -b \\ -c, & a \end{pmatrix};$$

and the other negative powers of M can then be obtained by successive multiplication with M^{-1}.

47. The expression for the nth power is however most readily obtained by means of a particular algorithm for matrices of the order 2.

Let h, b, c, J, q be any quantities, and write for shortness $R = -h^2 - 4bc$; suppose also that h', b', c', J', q' are any other quantities, such nevertheless that $h' : b' : c' = h : b : c$, and write in like manner $R' = -h'^2 - 4b'c'$. Then observing that $\dfrac{h}{\sqrt{R}}, \dfrac{b}{\sqrt{R}}, \dfrac{c}{\sqrt{R}}$ are respectively equal to $\dfrac{h'}{\sqrt{R'}}, \dfrac{b'}{\sqrt{R'}}, \dfrac{c'}{\sqrt{R'}}$, the matrix

$$\begin{pmatrix} J\left(\cot q - \dfrac{h}{\sqrt{R}}\right), & \dfrac{2bJ}{\sqrt{R}} \\ \dfrac{2cJ}{\sqrt{R}} & , \; J\left(\cot q + \dfrac{h}{\sqrt{R}}\right) \end{pmatrix}$$

contains only the quantities J, q, which are not the same in both systems; and we may therefore represent this matrix by (J, q), and the corresponding matrix with

h', b', c', J', q' by (J', q'). The two matrices are at once seen to be convertible (the assumed relations $h' : b' : c' = h : b : c$ correspond in fact to the conditions,

$$a' - d' : b' : c' = a - d : b : c,$$

of convertibility for the ordinary form), and the compound matrix is found to be

$$\left(\frac{\sin(q+q')}{\sin q \sin q'} JJ',\ q + q'\right);$$

and in like manner the several convertible matrices (J, q), (J', q'), (J'', q'') &c. give the compound matrix

$$\left(\frac{\sin(q+q'+q''\ldots)}{\sin q \sin q' \sin q''\ldots} JJ'J''\ldots,\ q + q' + q''\ldots\right).$$

48. The convertible matrices may be given in the first instance in the ordinary form, or we may take these matrices to be

$$\begin{pmatrix} a, & b \\ c, & d \end{pmatrix},\quad \begin{pmatrix} a', & b' \\ c', & d' \end{pmatrix},\quad \begin{pmatrix} a'', & b'' \\ c'', & d'' \end{pmatrix},\ \text{\&c.}$$

where of course $d - a : b : c = d' - a' : b' : c' = d'' - a'' : b'' : c'' =$ &c. Here writing $h = d - a$, and consequently $R = -(d-a)^2 - 4bc$, and assuming also $J = \frac{1}{2}\sqrt{R}$ and $\cot q = \frac{d+a}{\sqrt{R}}$, and in like manner for the accented letters, the several matrices are respectively

$$(\tfrac{1}{2}\sqrt{R},\ q)\,(\tfrac{1}{2}\sqrt{R'},\ q'),\ (\tfrac{1}{2}\sqrt{R''},\ q''),\ \text{\&c.},$$

and the compound matrix is

$$\left(\frac{\sin(q+q'+q''\ldots)}{\sin q \sin q' \sin q''\ldots} (\tfrac{1}{2}\sqrt{R})(\tfrac{1}{2}\sqrt{R'})(\tfrac{1}{2}\sqrt{R''})\ldots,\ \ q + q' + q'' + \ldots\right).$$

49. When the several matrices are each of them equal to

$$\begin{pmatrix} a, & b \\ c, & d \end{pmatrix},$$

we have of course $q = q' = q''\ldots$, $R = R' = R''\ldots$, and we find

$$\begin{pmatrix} a, & b \\ c, & d \end{pmatrix}^n = \left(\frac{\sin nq}{\sin^n q}(\tfrac{1}{2}\sqrt{R})^n,\ nq\right);$$

or substituting for the right-hand side, the matrix represented by this notation, and writing for greater simplicity

$$\frac{\sin nq}{\sin^n q}(\tfrac{1}{2}\sqrt{R})^n = (\tfrac{1}{2}\sqrt{R})\,L, \text{ or } L = \frac{\sin nq}{\sin^n q}(\tfrac{1}{2}\sqrt{R})^{n-1},$$

we find

$$\begin{pmatrix} a, & b \\ c, & d \end{pmatrix}^n = \begin{pmatrix} \tfrac{1}{2} L\,(\sqrt{R}\cot nq - (d-a)), & Lb \\ Lc & , \tfrac{1}{2} L\,(\sqrt{R}\cot nq + (d-a)) \end{pmatrix}$$

where it will be remembered that

$$R = -(d-a)^2 - 4bc \text{ and } \cot q = \frac{d+a}{\sqrt{R}},$$

the last of which equations may be replaced by

$$\cos q + \sqrt{-1}\sin q = \frac{d+a+\sqrt{-R}}{2\sqrt{ad-bc}}.$$

The formula in fact extends to negative or fractional values of the index n, and when n is a fraction, we must, as usual, in order to exhibit the formula in its proper generality, write $q + 2m\pi$ instead of q. In the particular case $n = \frac{1}{2}$, it would be easy to show the identity of the value of the square root of the matrix with that before obtained by a different process.

50. The matrix will be to a factor *près*, periodic of the nth order if only $\sin nq = 0$, that is, if $q = \frac{m\pi}{n}$ (m must be prime to n, for if it were not, the order of periodicity would be not n itself, but a submultiple of n); but $\cos q = \frac{d+a}{2\sqrt{ad-bc}}$, and the condition is therefore

$$(d+a)^2 - 4\,(ad-bc)\cos^2\frac{m\pi}{n} = 0,$$

or as this may also be written,

$$d^2 + a^2 - 2ad\cos\frac{2m\pi}{n} + 4bc\cos^2\frac{m\pi}{n} = 0,$$

a result which agrees with those before obtained for the particular values 2 and 3 of the index of periodicity.

51. I may remark that the last preceding investigations are intimately connected with the investigations of Babbage and others in relation to the function $\phi x = \frac{ax+b}{cx+d}$.

I conclude with some remarks upon rectangular matrices.

52. A matrix such as

$$\begin{pmatrix} a, & b, & c \\ a', & b', & c' \end{pmatrix}$$

where the number of columns exceeds the number of lines, is said to be a broad matrix; a matrix such as

$$\begin{pmatrix} a, & b \\ a', & b' \\ a'', & b'' \end{pmatrix}$$

where the number of lines exceeds the number of columns, is said to be a deep matrix

53. The matrix zero subsists in the present theory, but not the matrix unity. Matrices may be added or subtracted when the number of the lines and the number of the columns of the one matrix are respectively equal to the number of the lines and the number of the columns of the other matrix, and under the like condition any number of matrices may be added together. Two matrices may be equal or opposite, the one to the other. A matrix may be multiplied by a single quantity, giving rise to a matrix of the same form; two matrices so related are similar to each other.

54. The notion of composition applies to rectangular matrices, but it is necessary that the number of lines in the second or nearer component matrix should be equal to the number of columns in the first or further component matrix; the compound matrix will then have as many lines as the first or further component matrix, and as many columns as the second or nearer component matrix.

55. As examples of the composition of rectangular matrices, we have

$$\left(\begin{array}{l} a, b, c \\ d, e, f \end{array}\right) \between \left(\begin{array}{l} a', b', c', d' \\ e', f', g', h' \\ i', j', k', l' \end{array}\right) = \left(\begin{array}{l} (a, b, c \between a', e', i'),\ (a, b, c \between b', f', j')\ (a, b, c \between c', g', k'),\ (a, b, c \between d', h'\ l') \\ (d, e, f \between a', e', i'),\ (d, e, f \between b', f', j')\ (d, e, f \between c', g', k'),\ (d, e, f \between d', h'\ l') \end{array}\right),$$

and

$$\left(\begin{array}{l} a, d \\ b, e \\ c, f \end{array}\right) \between \left(\begin{array}{l} a', b', c', d' \\ e', f', g', h' \end{array}\right) = \left(\begin{array}{l} (a, d \between a', e'),\ (a, d \between b', f'),\ (a, d \between c', g'),\ (a, d \between d', h') \\ (b, e \between a', e'),\ (b, e \between b', f'),\ (b, e \between c', g'),\ (b, e \between d', h') \\ (c, f \between a', e'),\ (c, f \between b', f'),\ (c, f \between c', g'),\ (c, f \between d', h') \end{array}\right).$$

56. In the particular case where the lines and columns of the one component matrix are respectively equal in number to the columns and lines of the other component matrix, the compound matrix is square, thus we have

$$\left(\begin{array}{l} a, b, c \\ d, e, f \end{array}\right) \between \left(\begin{array}{l} a', d' \\ b', e' \\ c', f' \end{array}\right) = \left(\begin{array}{l} (a, b, c \between a', b', c'),\ (a, b, c \between d', e', f') \\ (d, e, f \between a', b', c'),\ (d, e, f \between d', e', f') \end{array}\right)$$

and

$$\left(\begin{array}{l} a', d' \\ b', e' \\ c', f' \end{array}\right) \between \left(\begin{array}{l} a, b, c \\ d, e, f \end{array}\right) = \left(\begin{array}{l} (a', d' \between a, d),\ (a', d' \between b, e),\ (a', d' \between c, f) \\ (b', e' \between a, d),\ (b', e' \between b, e),\ (b', e' \between c, f) \\ (c', f' \between a, d),\ (c', f' \between b, e),\ (c', f' \between c, f) \end{array}\right)$$

The two matrices in the case last considered admit of composition in the two different orders of arrangement, but as the resulting square matrices are not of the same order, the notion of the convertibility of two matrices does not apply even to the case in question.

57. Since a rectangular matrix cannot be compounded with itself, the notions of the inverse or reciprocal matrix and of the powers of the matrix and the whole resulting theory of the functions of a matrix, do not apply to rectangular matrices.

58. The notion of transposition and the symbol tr. apply to rectangular matrices, the effect of a transposition being to convert a broad matrix into a deep one and reciprocally. It may be noticed that the symbol tr. may be used for the purpose of expressing the law of composition of square or rectangular matrices. Thus treating (a, b, c) as a rectangular matrix, or representing it by $\begin{pmatrix} a, & b, & c \end{pmatrix}$, we have

$$\text{tr.} \begin{pmatrix} a', & b', & c' \end{pmatrix} = \begin{pmatrix} a' \\ b' \\ c' \end{pmatrix},$$

and thence

$$\begin{pmatrix} a, & b, & c \end{pmatrix} \text{tr.} \begin{pmatrix} a', & b', & c' \end{pmatrix} = (a, b, c \between \begin{pmatrix} a' \\ b' \\ c' \end{pmatrix} = (a, b, c \between a', b', c'),$$

so that the symbol

$$(a, b, c \between a', b', c')$$

would upon principle be replaced by

$$\begin{pmatrix} a, & b, & c \end{pmatrix} \text{tr.} \begin{pmatrix} a', & b', & c' \end{pmatrix}:$$

it is however more convenient to retain the symbol

$$(a, b, c \between a', b', c').$$

Hence introducing the symbol tr. only on the left-hand sides, we have

$$\begin{pmatrix} a, & b, & c \\ d, & e, & f \end{pmatrix} \text{tr.} \begin{pmatrix} a', & b', & c' \\ d', & e', & f' \end{pmatrix} = \begin{pmatrix} (a, b, c \between a', b', c'), & (a, b, c \between d', e', f') \\ (d, e, f \between a', b', c'), & (d, e, f \between d', e', f') \end{pmatrix}.$$

or to take an example involving square matrices,

$$\begin{pmatrix} a, & b \\ d, & e \end{pmatrix} \text{tr.} \begin{pmatrix} a', & b' \\ d', & e' \end{pmatrix} = \begin{pmatrix} (a, b \between a', b'), & (a, b \between d', e') \\ (d, e \between a', b'), & (d, e \between d', e') \end{pmatrix};$$

it thus appears that in the composition of matrices (square or rectangular), when the second or nearer component matrix is expressed as a matrix preceded by the symbol tr., any *line* of the compound matrix is obtained by compounding the corresponding *line* of the first or further component matrix successively with the several *lines* of the matrix which preceded by tr. gives the second or nearer component matrix. It is to be noticed that the terms 'symmetrical' and 'skew symmetrical' do not apply to rectangular matrices.

BIBLIOGRAPHY

1. Battista, Michail T. Spatial Visualization and Gender Differences in High School Geometry. Journal for Research in Mathematics Education, 1990, Vol. 21, No. 1, 47-60.

2. Ben-Chaim, David & Lappan, Glenda & Housang. The Effect of Instruction on Spatial Visualization Skills of Middle School Boys and Girls. American Educational Research Journal, 1988, Vol. 25, No. 1, 51-71.

3. Bennett. Visual Thinking and Number Relationships. Unknown.

4. Bishop, Alan J. Review of Research on Visualization in Mathematics Education. Focus on Learning Problems in Mathematics, 1989, Vol. 11, No. 1, 7-16.

5. Bishop, Alan J. Spatial Abilities and Mathematics Education - A Review. Educational Studies in Mathematics, 1980, Vol 11, 257-269.

6. Brown, Stephen &Walter, Marion. The Art of Problem Posing, Lawrence Erlbaum Associates, Hillsdale, New Jersey, 1983.

7. Burke, Maurice J. The Big Picture. Mathematics Teacher, April 1990, 258-262.

8. Cayley, Arthur. The Collected Mathematical Papers of Arthur Cayley, Cambridge University Press, London. 1889, 1895, Vol. 2 & 8.

9. Clements, M.A. & Del Campo, Gina. Linking Verbal Knowledge, Visual Images, and Episodes for Mathematical Learning. Focus on Learning Problems in Mathematics, 1989, Vol 11, No. 1, 25-33.

10. Eisenberg, Theodore & Dreyfus, Tommy. Spatial Visualization in the Mathematics Curriculum. Focus on Learning Problems in Mathematics, 1989, Vol. 11, No. 1, 1-5.

11. Harel, Guershon. Learning and Teaching Linear Algebra: Difficulties and an Alternative Approach to Visualizing Concepts and Processes. Focus on Learning Problems in Mathematics, 1989, Vol 11, No. 1, 139-148.

12. Kaput, James J. Supporting Concrete Visual Thinking in Multiplicative Reasoning: Difficulties an Opportunities. Focus on Learning Problems in Mathematics, 1989, Vol. 11, No. 1, 35-47.

13. Larecombe, Tony. Mathematical Learning Difficulties in the Secondary School Pupil Needs and Tearcher Roles, Open University Press, Milton Keynes, England, 1985.

14. Maier, Eugene. Mathematics and Visual Thinking. The Oregon Mathematics Teacher, Sept, 1985, 21-23.

15. Maier, Eugene. Interview, April 15, 1992. Portland, OR.

16. McKim, Robert. Experiences in Visual Thinking, PWS Publishers, Boston, Mass, 1980.