SQUARE ROOTS OF 2x2 MATRICES [1]

Sam Northshield
SUNY-Plattsburgh

1. INTRODUCTION

What is the square root of a matrix such as $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$? It is not, in general, $\begin{pmatrix} \sqrt{A} & \sqrt{B} \\ \sqrt{C} & \sqrt{D} \end{pmatrix}$. This is easy to see since the upper left entry of its square is $A + \sqrt{BC}$ and not A. The square of the matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is $\begin{pmatrix} a^2+bc & b(a+d) \\ c(a+d) & d^2+bc \end{pmatrix}$ and, if this is to equal $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$, then the following system of equations must be solved:

$a^2+bc = A$, $b(a+d) = B$, $c(a+d) = C$, and $d^2+bc = D$.

We may return to solve this later. However, let's first look at some examples.

Example 1. The matrix $\begin{pmatrix} 4 & 0 \\ 0 & 9 \end{pmatrix}$ has four square roots:

$$\begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}, \begin{pmatrix} 2 & 0 \\ 0 & -3 \end{pmatrix}, \begin{pmatrix} -2 & 0 \\ 0 & 3 \end{pmatrix}, \begin{pmatrix} -2 & 0 \\ 0 & -3 \end{pmatrix}.$$

Example 2. $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ has two square roots:

$$\begin{pmatrix} 1 & \frac{1}{2} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} -1 & -\frac{1}{2} \\ 0 & -1 \end{pmatrix}.$$

Matrices which have just two square roots can often be recognized as geometric transformations which can be "halved" in an obvious way. For example, *shear* matrices $\begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}$ transform the plane to itself by sliding horizontal lines to the right by a times the y-intercept of the line (so its square root is $\begin{pmatrix} 1 & \frac{a}{2} \\ 0 & 1 \end{pmatrix}$). Rotation matrices $\begin{pmatrix} t & s \\ -s & t \end{pmatrix}$,$s^2+t^2=1$, rotate the plane around the origin by θ where $\cos\theta = t$ and $\sin\theta = s$ (so its square roots are the rotation matrices corresponding to rotation by $\frac{\theta}{2}$ and $\pi + \frac{\theta}{2}$.

Example 3. $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ has no square roots.

[1] AMS 2000 subject classification: Primary 15A24, Secondary 15A36, 11C20

To see this, suppose to the contrary that

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

Then $a^2+bc = d^2+bc = 0$ which implies $a = \pm d$. Since $b(a+d) = 1$, $a+d \neq 0$ and so $a = d \neq 0$. Finally, since $c(a+d) = 0$ it follows that $c = 0$ and thus $a = 0$ - a contradiction!

Example 4. $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ has infinitely many square roots. To see this, recall the Cayley-Hamilton Theorem which states that A satisfies its characteristic equation:

$$A^2 = \tau A - \delta I$$

where τ is the trace of A and δ is the determinant of A. Hence, if A has trace 0 and determinant -1, for example

$$A = \begin{pmatrix} a & b \\ \frac{a^2-1}{b} & -a \end{pmatrix},$$

then $A^2 = I$.

Here are some square roots for what we'll call *Jordan* matrices (matrices with lower left entry 0– also known as upper triangular matrices or the Jordan canonical form of a matrix).

$$\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}^{\frac{1}{2}} = \begin{pmatrix} \pm\sqrt{a} & 0 \\ 0 & \pm\sqrt{b} \end{pmatrix}.$$

Note that this covers all four square roots when $a \neq b$.

$$\begin{pmatrix} a & 1 \\ 0 & a \end{pmatrix}^{\frac{1}{2}} = \pm \begin{pmatrix} \sqrt{a} & \frac{1}{2\sqrt{a}} \\ 0 & \sqrt{a} \end{pmatrix}.$$

Note that this covers both square roots.

Most generally,

$$\begin{pmatrix} a & b \\ 0 & c \end{pmatrix}^{\frac{1}{2}} = \begin{pmatrix} \pm\sqrt{a} & \frac{b}{\pm\sqrt{a}\pm\sqrt{c}} \\ 0 & \pm\sqrt{c} \end{pmatrix}.$$

These are all the square roots; you see that there are four for the first type and two for the second. The last – the most general case– includes the other two. Notice that if $a = c$ then two of the possible square roots are undefined.

What follows are five methods for computing the square roots of arbitrary two-by-two matrices. I include a number of applications and examples. My assignment of names to each method is informal and has no historical significance as far as I know.

2. SIMILARITY METHOD

Although not every matrix is a Jordan matrix, every matrix A is similar to a Jordan matrix:

$$\forall A : \exists M : (M^{-1}AM)_{21} = 0.$$

If $M^{-1}AM = J$ and $J^{\frac{1}{2}}$ is a square-root of J, then

$$(MJ^{\frac{1}{2}}M^{-1})^2 = MJM^{-1} = A$$

and so $MJ^{\frac{1}{2}}M^{-1}$ is a square root of A.

It is well known, and easy to see, that if the columns of M are linearly independent eigenvectors for A, then $M^{-1}AM$ is diagonal. Hence, finding a diagonalizing matrix M is no harder than finding the eigenvectors of A. For our purposes, we simply want to find a matrix M which, upon conjugating A, gives a Jordan matrix. It turns out that (almost) all matrices are similar, via a rotation matrix, to a Jordan matrix. We get this algebraically. A matrix similar to $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ via a rotation looks like:

$$\begin{pmatrix} t & s \\ -s & t \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} t & -s \\ s & t \end{pmatrix} = \begin{pmatrix} ? & ? \\ X & ? \end{pmatrix}$$

where

$$X = ct^2 + (d-a)st - bs^2.$$

To be a Jordan matrix, $X = 0$ and so we want s, t that satisfy

- i) $s^2 + t^2 = 1$, and
- ii) $ct^2 + (d-a)st - bs^2 = 0$. As long as $X \neq s^2 + t^2$ (equivalently, A is not a scalar multiple of a rotation matrix), it is possible.

To find the square root of A in the applicable cases:

- 1) Find roots of
$$cx^2 + (d-a)x - b = 0.$$
- 2) Find s, t which satisfy $s^2 + t^2 = 1$ and
$$ct^2 + (d-a)st - bs^2 = 0$$
and form $M = \begin{pmatrix} t & -s \\ s & t \end{pmatrix}$.
- 3) Calculate $J = M^{-1}AM$.
- 4) Find $J^{\frac{1}{2}}$.
- 5) Calculate $MJ^{\frac{1}{2}}M^{-1}$.

Example 5. Let $A = \begin{pmatrix} 8 & -2 \\ 6 & 1 \end{pmatrix}$.

- 1) Solving $6x^2 - 7x + 2 = 0$, we find $x = \frac{1}{2}$ or $x = \frac{2}{3}$.
- 2) Choosing the root ($\frac{1}{2}$), we next find s, t so that $\frac{t}{s} = \frac{1}{2}$ and $s^2 + t^2 = 1$. Namely, $s = \frac{2}{\sqrt{5}}$, $t = \frac{1}{\sqrt{5}}$. Then

$$M = \begin{pmatrix} \frac{1}{\sqrt{5}} & \frac{-2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{pmatrix}.$$

- 3,4) Calculating $J = M^{-1}AM = \begin{pmatrix} 4 & -8 \\ 0 & 5 \end{pmatrix}$, gives four square roots of J:

$$\begin{pmatrix} \pm 2 & \frac{-8}{\pm 2 \pm \sqrt{5}} \\ 0 & \pm\sqrt{5} \end{pmatrix}.$$

  Choosing one, say

$$J^{\frac{1}{2}} = \begin{pmatrix} 2 & 16 - 8\sqrt{5} \\ 0 & \sqrt{5} \end{pmatrix},$$

- 5) We get

$$MJ^{\frac{1}{2}}M^{-1} = \begin{pmatrix} -6 + 4\sqrt{5} & 4 - 2\sqrt{5} \\ -12 + 6\sqrt{5} & 8 - 3\sqrt{5} \end{pmatrix}$$

  which, indeed, is a square root of A.

We note that the choice of the other root ($\frac{2}{3}$) in step 2 will still give the same set of square roots of A.

Example 6. Although dealing with real numbers is desirable, it is not essential. Let $A = \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix}$.

- 1) $x^2 - x + 1 = 0$ has two solutions: $\frac{1}{2}(1 \pm \sqrt{3})$.
- 2) We take $s = \frac{1}{2}(\sqrt{3} + i)$ and $t = \frac{1}{2}(\sqrt{3} - i)$. Then

$$M = \frac{1}{2}\begin{pmatrix} \sqrt{3} - i & -\sqrt{3} - i \\ \sqrt{3} + i & \sqrt{3} - i \end{pmatrix}.$$

- 3) $J = M^{-1}AM = \dfrac{1}{2}\begin{pmatrix} 1 - \sqrt{3}i & -2 \\ 0 & 1 + \sqrt{3}i \end{pmatrix}$.
- 4) $J^{\frac{1}{2}} = \dfrac{1}{2}\begin{pmatrix} \sqrt{3} - i & -4/\sqrt{3} \\ 0 & \sqrt{3} + i \end{pmatrix}$.
- 5) $A^{\frac{1}{2}} = MJ^{\frac{1}{2}}M^{-1} = \dfrac{1}{\sqrt{3}}\begin{pmatrix} 2 & -1 \\ 1 & 1 \end{pmatrix}$.

3. ABEL-MÖBIUS METHOD

The equation $cx^2 + (d-a)x - b = 0$ actually has a geometric significance. We note the following chain of equivalences:

$$cx^2 + (d-a)x - b = 0$$

$$\frac{ax+b}{cx+d} = x$$

$$\exists\lambda : \begin{pmatrix} \lambda x \\ \lambda \end{pmatrix} = \begin{pmatrix} ax+b \\ cx+d \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ 1 \end{pmatrix}$$

$\begin{pmatrix} x \\ 1 \end{pmatrix}$ is an eigenvector.

Another interesting connection between the equation $cx^2 + (d-a)x - b = 0$ and the square roots of A is via Abel's functional equation.

Theorem 1. *Let* $p(x) = cx^2 + (d-a)x - b$. *Then*

$$F(x) = \int \frac{dx}{p(x)}$$

satisfies Abel's functional equation:

$$F(\frac{ax+b}{cx+d}) = F(x) + k.$$

This can be used to find a closed formula for powers of A (in particular, the $\frac{1}{2}$ power). To see this, given a matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, define $\Phi_A(x) = \dfrac{ax+b}{cx+d}$. It is easy to see that

$$\Phi_A \circ \Phi_B = \Phi_{AB}.$$

Hence, if $F(\Phi_A(x)) = F(x) + k$, then

$$\Phi_{A^n}(x) = F^{-1}(F(x) + nk).$$

Example 7. Let $A = \begin{pmatrix} 8 & -2 \\ 6 & 1 \end{pmatrix}$. Then $p(x) = 6x^2 - 7x + 2$ and

$$F(x) = \int \frac{dx}{6x^2 - 7x + 2} = \ln(\frac{3x-2}{2x-1}).$$

Then,

$$F(\frac{8x-2}{6x+1}) = \ln(\frac{12x-8}{10x-5}) = \ln(\frac{4}{5}) + \ln(\frac{3x-2}{2x-1}) = F(x) + \ln(\frac{4}{5}).$$

Since

$$F^{-1}(x) = \frac{e^x - 2}{2e^x - 3},$$

it works out that

$$\Phi_{A^n}(x) = F^{-1}(F(x) + n\ln(\frac{4}{5})) = \frac{(4\cdot 5^n - 3\cdot 4^n)x + (2\cdot 4^n - 2\cdot 5^n)}{(6\cdot 5^n - 6\cdot 4^n)x + (4\cdot 4^n - 3\cdot 5^n)}.$$

Coming full circle, this shows (with a little more work)

$$A^n = \begin{pmatrix} 4\cdot 5^n - 3\cdot 4^n & 2\cdot 4^n - 2\cdot 5^n \\ 6\cdot 5^n - 6\cdot 4^n & 4\cdot 4^n - 3\cdot 5^n \end{pmatrix}. \quad (1)$$

Letting $n = \frac{1}{2}$, we find

$$A^{\frac{1}{2}} = \begin{pmatrix} 4\sqrt{5} - 6 & 4 - 2\sqrt{5} \\ 6\sqrt{5} - 12 & 8 - 3\sqrt{5} \end{pmatrix}.$$

The form of equation (1) is not surprising. A consequence of the Cayley-Hamilton is that

$$A^{n+1} = \tau A^n - \delta A^{n-1}$$

and so the ij-th entry of A^n satisfies a second order recurrence (like the Fibonacci numbers) and so satisfies a Binet-type formula (like the Fibonacci numbers).

We shall now prove Theorem 1 in two ways; the first utilizing the fact that the roots of $p(x)$ are slopes of eigenvectors, the second related to the system of differential equations defined by A. We assume that $p(x)$ has distinct real roots (which, since the discriminant of $p(x)$ is the same as that of the characteristic polynomial of A, is equivalent to A having distinct real eigenvalues).

Proof 1. Let x_1 and x_2 be the roots of $p(x)$ and define

$$M = \begin{pmatrix} x_1 & x_2 \\ 1 & 1 \end{pmatrix}.$$

As was noted above, the columns of M are eigenvectors and so $D = M^{-1}AM$ is diagonal. Then $\Phi_D(x) = kx$ for some k and

$$\Phi_{M^{-1}}(\Phi_A(x)) = \Phi_D(\Phi_{M^{-1}}(x)) = k\Phi_{M^{-1}}(x).$$

If $F(x) = \displaystyle\int \frac{dx}{F(x)}$, then, by partial fractions,

$$F(x) = c\ln\left|\frac{x - x_1}{x - x_2}\right| = c\ln|\Phi_{M^{-1}}(x)|$$

and thus

$$F(\frac{ax+b}{cx+d}) = c\ln|k| + F(x).$$

QED

Proof 2. Let $x = x(t)$ and $y = y(t)$ be the solutions to the system of differential equations

$$\begin{cases} x' = ax + by \\ y' = cx + dy \end{cases}.$$

By the quotient rule,

$$\left(\frac{x}{y}\right)' = -p\left(\frac{x}{y}\right)$$

and, similarly,

$$\left(\frac{x'}{y'}\right)' = -p\left(\frac{x'}{y'}\right).$$

If $F(x) = \displaystyle\int \frac{dx}{p(x)}$, then $[F(x/y)]' = -1 = [F(x'/y')]'$ and so

$$F(\frac{ax+by}{cx+dy}) = F(\frac{x'}{y'}) = F(\frac{x}{y}) + k$$

for some k and therefore, for all z in the range of x/y,

$$F(\Phi_A(z)) = F(z) + k.$$

QED

Although the proofs of Theorem 1 require $p(x)$ to have real roots, it still works to some extent for other matrices.

Example 8. Let $A = \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix}$. Then $p(x) = x^2 - x + 1$ and

$$F(x) = \int \frac{dx}{x^2 - x + 1} = \frac{2}{\sqrt{3}} \arctan\left(\frac{2x-1}{\sqrt{3}}\right).$$

Then

$$F(1 - \frac{1}{x}) = \frac{2}{\sqrt{3}} \arctan\left(\frac{x-2}{\sqrt{3}x}\right) = \frac{2}{\sqrt{3}} \arctan\left(\frac{2x-1}{\sqrt{3}}\right) - \frac{2}{\sqrt{3}} \arctan(\sqrt{3}).$$

Since $F^{-1}(x) = \frac{1}{2} + \frac{\sqrt{3}}{2}\tan(\frac{\sqrt{3}}{2}x)$,

$$\Phi_{A^n}(x) = \frac{1}{2} + \frac{\sqrt{3}}{2} \tan\left(\arctan\left(\frac{2x-1}{\sqrt{3}}\right) - n \cdot \arctan(\sqrt{3})\right).$$

It is a challenging exercise to use the addition formula for arctangents to show

$$\Phi_{A^{\frac{1}{2}}}(x) = \frac{1}{2} + \frac{3}{2}\left(\frac{x-1}{x+1}\right) = \frac{2x-1}{x+1}$$

and therefore

$$A^{\frac{1}{2}} = \pm\frac{1}{\sqrt{3}}\begin{pmatrix} 2 & -1 \\ 1 & 1 \end{pmatrix}.$$

4. NEWTON'S METHOD

Newton's method is a way of approximating roots of a given function. It works as follows. Given a function $f(x)$ and an initial value x_0, define

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}.$$

The sequence often converges to a root of the function $f(x)$.

Its effectiveness varies according to the type of function and initial guess. Define

$$x_{n+1} = \frac{x_n^2 + a}{2x_n} = \frac{1}{2}(x_n + \frac{a}{x_n}).$$

If $x_0 > 0$, then $x_n \to \sqrt{a}$ and if $x_0 < 0$, then $x_n \to -\sqrt{a}$. This follows from the following easily proved formula:

$$\frac{x_{n+1} - \sqrt{a}}{x_{n+1} + \sqrt{a}} = (\frac{x_n - \sqrt{a}}{x_n + \sqrt{a}})^2.$$

We now attempt Newton's method for matrices. That is, given a starting guess X_0, define

$$X_{n+1} = \frac{1}{2}(X_n + AX_n^{-1}).$$

Example 9. Let $A = \begin{pmatrix} -1 & -2 \\ 4 & -1 \end{pmatrix}$, and $X_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. Then

$X_1 = \begin{pmatrix} 0 & -1 \\ 2 & 0 \end{pmatrix}$,

$X_2 = \begin{pmatrix} 1 & -.75 \\ 1.5 & 1 \end{pmatrix}$,

$X_3 = \begin{pmatrix} .9706 & -1.022 \\ 2.0441 & .9706 \end{pmatrix}$,

$X_4 = \begin{pmatrix} .9995 & -.9998 \\ 1.9996 & .9995 \end{pmatrix}$, and

$X_5 = \begin{pmatrix} 1.000 & -1.000 \\ 2.000 & 1.000 \end{pmatrix}$.

Hence X_n rapidly converges to a square root of A.

We say that a matrix is *positive* if it has positive eigenvalues. We then reserve the notation $\sqrt{A}$ to denote the positive square root of A (there is indeed only one such square root; the other(s) having spectrum with at least one negative element). For example

$$\sqrt{\begin{pmatrix} -1 & -2 \\ 4 & -1 \end{pmatrix}} = \begin{pmatrix} 1 & -1 \\ 2 & 1 \end{pmatrix}.$$

It turns out that the convergence in Example 9 is true in general.

Theorem 2. *Let A and X_0 be positive. If A and X_0 can be simultaneously upper triangularized (for example, if A and X_0 commute), then $X_n \to \sqrt{A}$.*

Proof. By hypothesis, there exists M such that for some a, b, c, x, y and z, $MAM^{-1} = \begin{pmatrix} a & b \\ 0 & c \end{pmatrix}$ and $MX_0M^{-1} = \begin{pmatrix} x & y \\ 0 & z \end{pmatrix}$. It follows that the matrix $(X_0 - \sqrt{A})(X_0 + \sqrt{A})^{-1}$ has spectrum

$$\{\frac{x-\sqrt{a}}{x+\sqrt{a}}, \frac{z-\sqrt{c}}{z+\sqrt{c}}\} \subset (-1, 1).$$

Let $B_n = (X_n - \sqrt{A})(X_n + \sqrt{A})^{-1}$. It is easy to verify that $B_{n+1} = B_n^2$ and therefore

$$MB_nM^{-1} = \begin{pmatrix} a_n & b_n \\ 0 & c_n \end{pmatrix}$$

where $a_n, c_n \to 0$. Since $b_{n+1} = b_n(a_n + c_n)$, $b_n \to 0$ and therefore $B_n \to 0$. Since

$$X_n = [2(I - B_n)^{-1} - I]\sqrt{A},$$

it follows that $X_n \to \sqrt{A}$. QED

A more general version of this theorem has been done by Higham [4].

Interestingly, the choice of X_0 is important if X_0 and A do not commute. For example, consider $A = \begin{pmatrix} 1 & 4 \\ 0 & 1 \end{pmatrix}$ for which $\sqrt{A} = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$. If $X_0 = \begin{pmatrix} 1 & 0 \\ c & 2 \end{pmatrix}$ then, apparently, $X_n \to \sqrt{A}$ if c is less than but near $\frac{5}{6}$ but X_n diverges if c is larger than but near $\frac{5}{6}$. The number $\frac{5}{6}$ is critical in the sense that if $X_0 = \begin{pmatrix} 1 & 0 \\ \frac{5}{6} & 2 \end{pmatrix}$, then X_1 is not invertible. In general, there are infinitely many matrices $\begin{pmatrix} 1 & 0 \\ c & 2 \end{pmatrix}$ such that some X_n is not invertible and so one might expect that the set of matrices X_0 for which Newton's method converges is quite complicated. This is indeed borne out by computer experimentation.

Let $\mathcal{S}$ be the set of all matrices X_0 for which X_n converges. $\mathcal{S}$ is a subset of the four dimensional space of two-by-two matrices. By Theorem 2, $\mathcal{S}$ contains the plane $\{sA + tI : s, t \in \mathbf{R}\}$ but computer experiments indicate that $\mathcal{S}$ is a self-similar fractal. Following are examples of slices through $\mathcal{S}$; $\begin{pmatrix} s & t \\ t & 1 \end{pmatrix}$ in Figure 1, for example, indicates the plane $\{\begin{pmatrix} s & t \\ t & 1 \end{pmatrix} : s, t \in [-100, 100]^2\}$ and the black pixels represent matrices $X_0 = \begin{pmatrix} s & t \\ t & 1 \end{pmatrix}$ such that X_n (apparently) converges.

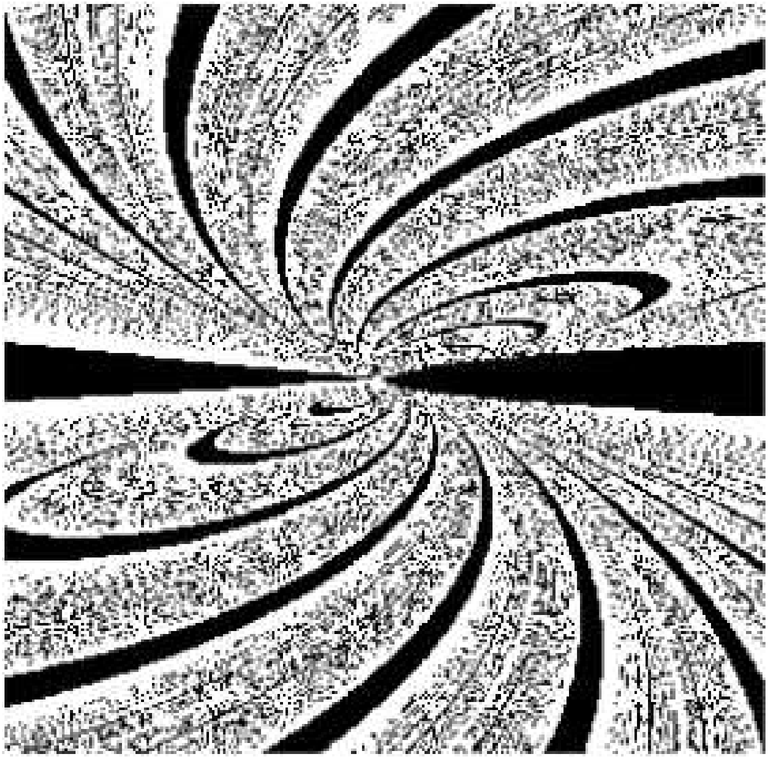

Fig. 1 $\begin{pmatrix} s & t \\ t & 1 \end{pmatrix}$

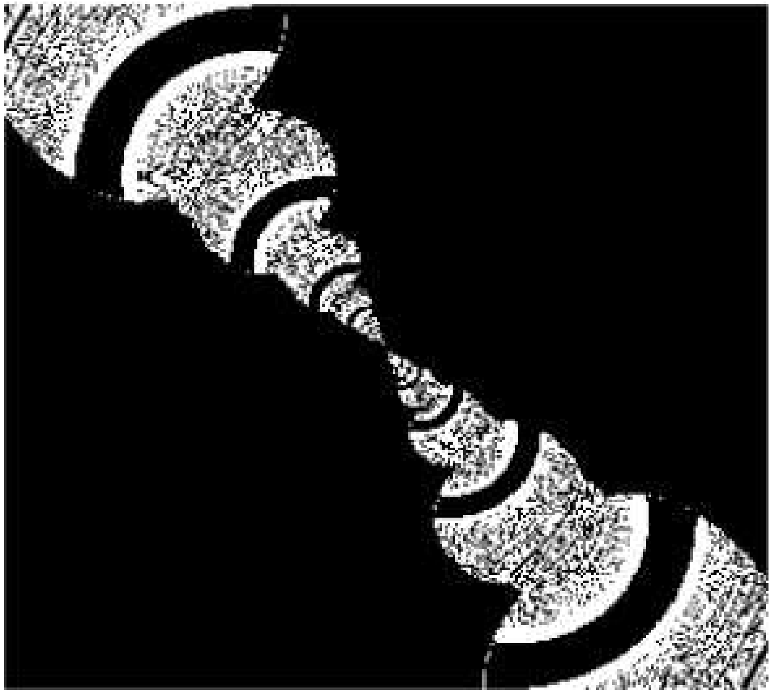

Fig. 2 $\begin{pmatrix} s & t \\ -t & s \end{pmatrix}$

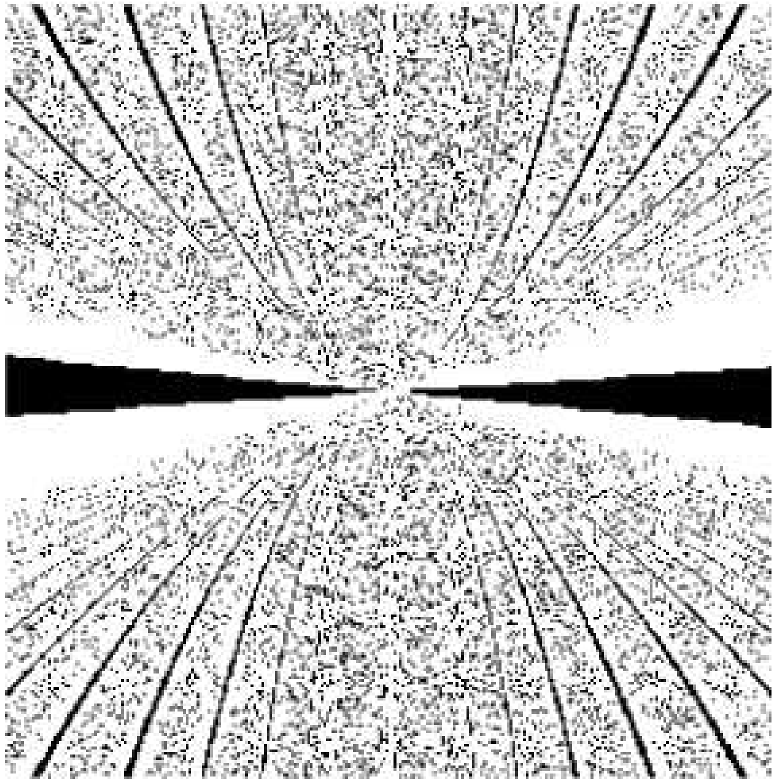

Fig. 3 $\begin{pmatrix} s & 0 \\ t & -s \end{pmatrix}$

Figure 3 represents part of the plane which is the orthogonal complement to the plane of matrices which commute with $\begin{pmatrix} 1 & 4 \\ 0 & 1 \end{pmatrix}$.

Some facts/questions jump out at one upon seeing these pictures. It is easy to see that if $X_0 \in \mathcal{S}$ then $-X_0 \in \mathcal{S}$. Apparently, $\mathcal{S}$ is invariant under multiplication by 2; does $X_0 \in \mathcal{S}$ imply $2X_0 \in \mathcal{S}$? As far as I know, this is an open question. Is $\mathcal{S}$ a true fractal? That is, is the Hausdorff dimension of a two-dimensional slice of $\mathcal{S}$ ever less than 2? What is the Hausdorff dimension of $\mathcal{S}$?

5. EXTENSION METHOD

We now consider functions of matrices. That is, if a function $f(x)$ is given, is there a way to define $f(A)$? There is extensive literature on this; see for example, Rinehart [6] and Uhlig [7]. This, of course, is of interest when $f(x) = \sqrt{x}$.

Consider first the general Jordan matrix $\begin{pmatrix} a & b \\ 0 & c \end{pmatrix}$. Then

$$\begin{pmatrix} a & b \\ 0 & c \end{pmatrix}^2 = \begin{pmatrix} a^2 & b(a+c) \\ 0 & c^2 \end{pmatrix},$$

$$\begin{pmatrix} a & b \\ 0 & c \end{pmatrix}^3 = \begin{pmatrix} a^3 & b(a^2+ac+c^2) \\ 0 & c^3 \end{pmatrix},$$

$$\begin{pmatrix} a & b \\ 0 & c \end{pmatrix}^4 = \begin{pmatrix} a^4 & b(a^3+a^2c+ac^2+c^3) \\ 0 & c^4 \end{pmatrix},$$

and, in general, $\begin{pmatrix} a & b \\ 0 & c \end{pmatrix}^n$ is either $\begin{pmatrix} a^n & \frac{b}{a-c}(a^n-c^n) \\ 0 & c^n \end{pmatrix}$ or $\begin{pmatrix} a^n & bna^{n-1} \\ 0 & a^n \end{pmatrix}$ according to whether $a \neq c$ or $a = c$ respectively. Hence for any polynomial $p(x)$,

$$p\left(\begin{pmatrix} a & b \\ 0 & c \end{pmatrix}\right) = \begin{pmatrix} p(a) & \frac{b}{a-c}(p(a)-p(c)) \\ 0 & p(c) \end{pmatrix}$$

or

$$p\left(\begin{pmatrix} a & b \\ 0 & c \end{pmatrix}\right) = \begin{pmatrix} p(a) & bp'(a) \\ 0 & p(a) \end{pmatrix}$$

according to whether $a \neq c$ or $a = c$ respectively.

We may extend further to analytic functions or even to *any* function $f(x)$: if $MAM^{-1} = \begin{pmatrix} a & b \\ 0 & c \end{pmatrix}$, then define

$$f(A) = M^{-1} \begin{pmatrix} f(a) & \frac{b}{a-c}(f(a)-f(c)) \\ 0 & f(c) \end{pmatrix} M$$

if A has distinct eigenvalues a, c and define, for the 'confluent' case when A has equal eigenvalues:

$$f(A) = M^{-1} \begin{pmatrix} f(a) & bf'(a) \\ 0 & f(a) \end{pmatrix} M.$$

Note that, of course, if f is not differentiable everywhere, then there exist matrices for which $f(A)$ is undefined. The fact that this definition is well-defined (i.e., the result is independent of the choice of M) is left to the reader.

In general, it is clear that A and $f(A)$ are simultaneously upper triangularizable and thus $f(A) = xA + yI$ for some x and y (possibly depending on both f and A). This is a classical formula appearing, for example, in Horn and Johnson [5]. If $MAM^{-1} = \begin{pmatrix} a & b \\ 0 & c \end{pmatrix}$, then the trace of $\sqrt{A}$ is $\sqrt{a} + \sqrt{c}$, the determinant of $\sqrt{A}$ is $\sqrt{ac}$ and therefore, since $\sqrt{A}$ satisfies its characteristic equation,

$$\sqrt{A} = \frac{1}{\sqrt{a} + \sqrt{c}}(A + \sqrt{ac}I). \tag{2}$$

Hence we have a formula for the square root of A depending only on A and its eigenvalues.

Example 10. Let $A = \begin{pmatrix} 8 & -2 \\ 6 & 1 \end{pmatrix}$. Then $\tau = 9$, $\delta = 20$ and so A has characteristic equation

$$x^2 - 9x + 20 = 0$$

and the eigenvalues are 4 and 5. By (2),

$$\sqrt{A} = \frac{1}{2 + \sqrt{5}}(A + 2\sqrt{5}I) = \begin{pmatrix} 4\sqrt{5} - 6 & 4 - 2\sqrt{5} \\ 6\sqrt{5} - 12 & 8 - 3\sqrt{5} \end{pmatrix}.$$

We may also apply this method to matrices without real eigenvalues.

Example 11. As in Example 8, let $A = \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix}$. Then A has eigenvalues $a, c = \frac{1}{2}(1 \pm i\sqrt{3})$. Since $ac = \delta = 1$ and $a + c = \tau = 1$,

$$(\sqrt{a} + \sqrt{c})^2 = a + c + 2\sqrt{ac} = 3$$

and so, by (2),

$$A^{\frac{1}{2}} = \frac{1}{\sqrt{3}}(A + I) = \frac{1}{\sqrt{3}} \begin{pmatrix} 2 & -1 \\ 1 & 1 \end{pmatrix}.$$

More generally, if $f(A) = xA + yI$ and A has distinct eigenvalues, then conjugation gives

$$\begin{pmatrix} f(a) & \frac{b}{a-c}(f(a) - f(c)) \\ 0 & f(c) \end{pmatrix} = f(\begin{pmatrix} a & b \\ 0 & c \end{pmatrix}) = x \begin{pmatrix} a & b \\ 0 & c \end{pmatrix} + y \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

and so $x = (f(a) - f(c))/(a - c)$, and $y = (af(c) - cf(a))/(a - c)$. When A has equal eigenvalues, then $x = f'(a)$ and $y = f(a) - af'(a)$. Therefore, if A has distinct eigenvalues a and c then

$$f(A) = \frac{f(a) - f(c)}{a - c}A + \frac{af(c) - cf(a)}{a - c}I \tag{3a}$$

while if A has eigenvalue a of multiplicity 2, then

$$f(A) = f'(a)A + (f(a) - af'(a))I. \tag{3b}$$

As an application, we consider continued fractions of square roots of matrices. Recall $\sqrt{2}$ can be written as an infinite continued fraction:

$$\sqrt{2} = 1 + \frac{1}{2 + \frac{1}{2+\frac{1}{2+\dots}}}.$$

We shall use the standard notation:

$$\sqrt{2} = [1, 2, 2, 2, \ldots].$$

In general, every irrational number x has an infinite continued fraction expansion:

$$x = a_0 + \frac{1}{a_1 + \frac{1}{a_2+\frac{1}{a_3+\dots}}} = [a_0, a_1, a_2, a_3, \ldots]$$

but 'quadratic surds' (i.e., irrational numbers of the form $r + \sqrt{s}$ where r and s are rational or, equivalently, irrational roots of quadratic polynomials with integer coefficients) are special in that they are precisely the numbers with eventually repeating continued fractions. For example,

$$\sqrt{2} = [1, \overline{2}]$$

and

$$\frac{3+\sqrt{7}}{5} = [1, 7, \overline{1, 2, 1, 8, 13, 8}].$$

This is a standard result in the theory of continued fractions; see, for example, [1] or [2].

Does the square-root of an integral matrix A satisfy

$$\sqrt{A} = A_0 + (A_1 + (A_2 + \ldots)^{-1})^{-1}$$

where A_k are integral and eventually repeat?

A natural attempt to answer this question is to extend the floor function to matrices. For example, given a matrix A, if $MAM^{-1} = \begin{pmatrix} s & 0 \\ 0 & t \end{pmatrix}$ then

$$[A] = M^{-1} \begin{pmatrix} [s] & 0 \\ 0 & [t] \end{pmatrix} M.$$

It is worth pointing out that if A is integral, $[A]$ need not be. For example if $A = \begin{pmatrix} 1 & 2 \\ 1 & 3 \end{pmatrix}$, then $[A] = \frac{1}{2}\begin{pmatrix} 3-\sqrt{3} & 2\sqrt{3} \\ \sqrt{3} & 3+\sqrt{3} \end{pmatrix}$. The reason for the discrepancy is that the eigenvalues of A are not rational. If a matrix A is integral with integral eigenvalues however, then $[A] = A$. This is more in line with what we would expect of integral matrices; we henceforth call such matrices *strongly integral*.

Consider now the continued fraction expansion of a matrix A. Let $X_0 = A$ and define, recursively, $A_n = [X_n]$ and $X_{n+1} = (X_n - A_n)^{-1}$.

The following theorem answers the question above (partially).

Theorem 3. *If A is strongly integral with distinct, positive, eigenvalues neither of which is a perfect square, then $\sqrt{A} = A_0 + (A_1 + (A_2 + \ldots)^{-1})^{-1}$ for a sequence of rational matrices (A_n) and, furthermore, the sequence is eventually periodic.*

Proof. Let $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ and suppose s, t are the eigenvalues of A. Since the discriminant of

$$cr^2 + (d-a)r - b = 0 \tag{4}$$

is the same as that of the characteristic equation and since the eigenvalues of A are integral, the solutions of (4) are rational. That is, there exist integers x, u, y, v such that $\frac{x}{u}$ and $\frac{y}{v}$ satisfy (4). Recall, from section 2, that this implies that $M^{-1}AM$ is diagonal where $M = \begin{pmatrix} x & y \\ u & v \end{pmatrix}$.

Since s, t are not both perfect squares, $\sqrt{s} = [s_0, s_1, ...]$ and $\sqrt{t} = [t_0, t_1, ...]$ for integers s_i and t_i. Furthermore, it's easy to see that

$$A_n = M \begin{pmatrix} s_n & 0 \\ 0 & t_n \end{pmatrix} M^{-1}.$$

Since M is integral, each A_n is rational. Since $\sqrt{s}$ and $\sqrt{t}$ are quadratic surds, it follows that the pairs (s_i, t_i) eventually repeat and therefore the matrices A_n eventually repeat. QED

It is too much to hope that A_i are integral. For example, if $A = \begin{pmatrix} 1 & 3 \\ -4 & 9 \end{pmatrix}$, then A satisfies the hypotheses of Theorem 3 but

$$[\sqrt{A}] = \frac{1}{4} \begin{pmatrix} 2 & 3 \\ -4 & 10 \end{pmatrix}.$$

An interesting fact is the following:

Proposition 1. *If A is a rational matrix, then the sequence A_i eventually repeats or is eventually undefined.*

Proof. An eigenvalue of A is either a rational number or a quadratic surd. QED

6. CAYLEY-HAMILTON METHOD

For what A is $A^{\frac{1}{2}}$ integral? To answer this question, we apply the Cayley-Hamilton Theorem to $A^{\frac{1}{2}}$ to get perhaps our simplest method. Note

$$A = \tau\sqrt{A} - \delta I \tag{5}$$

where τ is the trace of $\sqrt{A}$ and δ is the determinant of $\sqrt{A}$. Suppose A has trace T and determinant Δ and is not a multiple of I. By (5), $\tau \neq 0$ and we have

$$\sqrt{A} = \frac{1}{\tau}(A + \delta I). \tag{6}$$

Furthermore, $\delta^2 = \Delta$ or $\delta = \pm\sqrt{\Delta}$. Using (5) and (6),

$$TA - \Delta I = A^2 = (\tau\sqrt{A} - \delta I)^2 = \tau^2 A - 2\tau\delta\sqrt{A} + \delta^2 I = (\tau^2 - 2\delta)A - \Delta I$$

and so $T = \tau^2 - 2\delta$. Hence $\tau = \pm\sqrt{T + 2\delta}$ and, finally

$$\sqrt{A} = \frac{\pm 1}{\sqrt{T + 2\delta}}(A + \delta I), \quad \delta = \pm\sqrt{\Delta} \qquad (7).$$

Example 12. $A = \begin{pmatrix} 8 & -2 \\ 6 & 1 \end{pmatrix}$ has no integral square roots. Since $T = 9, \Delta = 20$, we have $\delta = \pm 2\sqrt{5}$. Hence

$$\sqrt{T + 2\delta} = \sqrt{9 + 4\sqrt{5}} = 2 \pm \sqrt{5}$$

and therefore

$$A = \frac{\pm 1}{2 \pm \sqrt{5}}\left[\begin{pmatrix} 8 & -2 \\ 6 & 1 \end{pmatrix} \pm 2\sqrt{5}I\right] = \begin{pmatrix} \dfrac{8 \pm 2\sqrt{5}}{2 \pm \sqrt{5}} & \dfrac{-2}{2 \pm \sqrt{5}} \\ \dfrac{6}{2 \pm \sqrt{5}} & \dfrac{1 \pm 2\sqrt{5}}{2 \pm \sqrt{5}} \end{pmatrix}.$$

Obviously, none of the four square roots of A are integral.

Example 13. $A = \begin{pmatrix} 2 & 7 \\ 7 & 25 \end{pmatrix}$ has two rational square roots but no integral ones. Since $T = 27$ and $\Delta = 1$, when $\delta = -1$ we get $A^{\frac{1}{2}} = \frac{1}{5}\begin{pmatrix} 1 & 7 \\ 7 & 24 \end{pmatrix}$ and when $\delta = 1$ we get $A^{\frac{1}{2}} = \frac{1}{\sqrt{29}}\begin{pmatrix} 3 & 7 \\ 7 & 26 \end{pmatrix}$

Example 14. $A = \begin{pmatrix} 7 & 10 \\ 15 & 22 \end{pmatrix}$ has two integral square roots and two irrational square roots. Since $T = 29, \Delta = 4$, and $\delta = \pm 2$, we have

$$A^{\frac{1}{2}} = \frac{\pm 1}{\sqrt{29 \pm 4}}\begin{pmatrix} 7 \pm 2 & 10 \\ 15 & 22 \pm 2 \end{pmatrix}$$

and so the square roots of A are $\pm\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ and $\pm\dfrac{1}{\sqrt{33}}\begin{pmatrix} 9 & 10 \\ 15 & 24 \end{pmatrix}$.

Example 15. $A = \begin{pmatrix} -11 & 6 \\ -30 & 16 \end{pmatrix}$ has four integral square roots: $\pm\begin{pmatrix} -3 & 2 \\ -10 & 6 \end{pmatrix}$ and $\pm\begin{pmatrix} -13 & 6 \\ -30 & 14 \end{pmatrix}$.

Based on (7), a matrix A with trace T and determinant Δ has integral square roots if and only if $\sqrt{T \pm 2\sqrt{\Delta}}$ is an integer which divides each entry of $A \pm \sqrt{\Delta} I$.

Suppose a and b are relatively prime. When does $A = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}$ have an integral square root? Answer: when $\sqrt{2a + 2\sqrt{a^2 + b^2}}$ is an integer or, equivalently, when there exists an integer c such that $a^2 + b^2 = c^2$ and $2(a+c)$ is a square. Assuming this, since a and b are relatively prime, either a or b is odd but, since $a + c$ must be even, a is odd. Conversely, if a is odd, b is even, and there is some c such that $a^2 + b^2 = c^2$, then

$$\frac{c-a}{2}\frac{c+a}{2} = (\frac{b}{2})^2.$$

Since the two factors on the left are relatively prime, $2(a+c)$ is a square which divides both $(a+c)^2$ and b^2. Therefore, $\sqrt{2a+2c}$ divides both $a+c$ and b. But this is exactly the condition for A to have an integral square root. Therefore, A has an integral square root if and only if a is odd, b is even and $a^2 + b^2 = c^2$ for some c.

This leads easily to the standard parametrization of Pythagorean triples. Suppose a, b and c are relatively prime and $a^2 + b^2 = c^2$ with a odd and b even. Then

$$\begin{pmatrix} x & y \\ u & v \end{pmatrix}^2 = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}$$

for some integers x, y, u, v. It's not hard to verify that $u = -y$ and $v = x$ from which it follows that

$$\begin{pmatrix} x & y \\ -y & x \end{pmatrix}^2 = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}$$

and therefore $a = x^2 - y^2$, $b = 2xy$, and $c = x^2 + y^2$.

It is worth noting that the set of matrices of the form $\begin{pmatrix} x & y \\ -y & x \end{pmatrix}$ with real entries forms a field isomorphic to the field of complex numbers via the map

$$\begin{pmatrix} x & y \\ -y & x \end{pmatrix} \to x + iy.$$

When x and y are integers, the resulting subring is isomorphic to the ring of “Gaussian integers”.

More generally, consider the quadratic field $Q(\sqrt{d}) = \{x + y\sqrt{d} : x, y \in Q\}$ where d is an integer. It is easy to see that the set of matrices of the form $\begin{pmatrix} a & b \\ db & a \end{pmatrix}$ (a, b rational) is a field isomorphic to $Q(\sqrt{d})$ via the mapping

$$\begin{pmatrix} a & b \\ db & a \end{pmatrix} \to a + b\sqrt{d}.$$

We may devise a test for when an element of $Q(\sqrt{d})$ is the square of another element in $Q(\sqrt{d})$: $a + b\sqrt{d}$ has square root in $Q(\sqrt{d})$ if and only if $\begin{pmatrix} a & b \\ db & a \end{pmatrix}$ has a rational square root if and only if $\sqrt{2a \pm \sqrt{a^2 - db^2}}$ is rational.

Example 16 Is $\frac{3+\sqrt{5}}{2}$ the square of a number of the form $x + y\sqrt{5}, x, y$ rational? Let $a = \frac{3}{2}$ and $b = \frac{1}{2}$. Then $a^2 - 5b^2 = 1$ and $\sqrt{2a + \sqrt{a^2 - 5b^2}} = 2$ and so the answer is yes. Computing the square root,

$$\begin{pmatrix} 3/2 & 1/2 \\ 5/2 & 3/2 \end{pmatrix}^{\frac{1}{2}} = \begin{pmatrix} 1/2 & 1/2 \\ 5/2 & 1/2 \end{pmatrix}$$

and thus

$$\sqrt{\frac{3+\sqrt{5}}{2}} = \frac{1+\sqrt{5}}{2}.$$

The question of when A has an integral square root is also related to the theory of quadratic forms. A binary quadratic form is a polynomial

$$Q(x, y) = ax^2 + 2bxy + cy^2.$$

Such a form is related to the matrix $\begin{pmatrix} a & b \\ b & c \end{pmatrix}$ by the equation

$$Q(x, y) = \begin{pmatrix} x \\ y \end{pmatrix}^t A \begin{pmatrix} x \\ y \end{pmatrix}$$

and therefore

$$Q(x, y) = \left| A^{\frac{1}{2}} \begin{pmatrix} x \\ y \end{pmatrix} \right|^2.$$

If A has an integral square root, then the corresponding quadratic form is the sum of squares of two linear forms. This is not the only case where this happens however. A theorem of Mordell [3] gives sufficient conditions for a quadratic form to be the sum of squares of two linear forms: the gcd of a,b,and c is a sum of two square, the determinant of A is a square, and Q is non-negative. The matrix $13 \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$ satisfies these conditions but does not have an integral square root, for example.

References.

[1] G.H. Hardy and E.M. Wright, *An Introduction to the Theory of Numbers, 4th ed.*, Oxford Univ. Press, New York, 1971.

[2] A.M. Rockett, P. Szusz, *Continued Fractions*, World Scientific, Singapore, 1992.

[3] Mordell, L.J., *On the representation of a binary quadratic form as a sum of squares of linear forms*, Math. Z. **35** (1932), 1-15.

[4] Higham, N.J., *Newton's Method for the Matrix Square Root*, Math. of Computation, **46** (1986) 537-549.

[5] Horn and Johnson, *Topics in Matrix Analysis*, Cambridge University Press, 1991.

[6] Rinehart, R.F., *The Equivalence of Definitions of a Matrix Function*, American Math. Monthly **62** (1955) 395-413.

[7] Uhlig, F., *Explicit Polar Decomposition...*, Linear Algebra Appl. **38** (1981) 239-249.

Sam Northshield
Dept. of Mathematics
Plattsburgh State Univ.
Plattsburgh, NY 12901
e-mail: samuel.northshield@plattsburgh.edu