

## MUIR'S RULES FOR MATRICES: ANOTHER LOOK AT STABILITY

*David Brown*

Much attention has been given to the concept of stability in recent SEP volumes. This is certainly understandable, since for computational purposes, the term "instability" is closely synonymous with "total uselessness." A well-known example occurs when deriving one-way wave equations: By expanding the square root operator in a Taylor series we obtain a sequence of one-way equations with dispersion relations of arbitrarily high accuracy. Yet all of these higher-order equations beyond second order are unstable, and therefore simply cannot be used for the extrapolation of wavefields in processes such as migration. The solution to this problem has been known for quite some time: the square root operator can be expanded using the rational fraction expansions suggested by Francis Muir and first reported by Jon Claerbout and Bjorn Engquist in SEP-8. In that report, Engquist showed that for the constant velocity case, square root approximations derived in this manner always yield stable one-way wave equations. Furthermore, the "15-degree" and "45-degree" equations turn out to be the first two equations given by this expansion. The introduction of "causal positive real" or "impedance" functions in SEP-15 and -16 (articles by Brown and Claerbout) gave more insight into the stability question for these operators. Muir's rational fraction expansion fit the definition for causal positive real (CPR) operators, and, furthermore, the expansion could be formed making use only of three simple rules for the combination of CPR operators (Muir's rules).

Successful as it was for finding stable extrapolation equations for constant velocity or stratified media, Muir's theory seemed to run into problems for earth models containing strong lateral velocity variation (see examples by Bloxson and Kjartansson in SEP-15). Some quick work reported by Godfrey, Muir, Claerbout and Jacobs in SEP-16 showed, however, that at least for the 45-degree equation, stable extrapolation, or "bullet-proofing," was indeed possible. The object of the present paper is to show that Muir's rules of combination can be generalized in a simple way to include the laterally varying case, and hence to show that all higher-order rational expansions can be bullet-proofed. Since the introduction of lateral velocity variation results in equations that cannot be Fourier-transformed over the lateral coordinate, we must consider matrix systems of equations rather than scalar equations; and hence the generalization to lateral variation will yield "Muir's Rules for Matrices." In addition, we will discuss briefly the questions of side boundary condition stability and internal boundary conditions since these questions are closely linked to the problem of bullet-proofing.

### *The Scalar Theory*

Let us begin by reviewing the theory of causal positive real operators. A passable definition for a CPR operator is that, when Fourier-transformed over all space variables and Laplace-transformed over time, such an operator will have a positive real part if the time derivatives are realized in a causal manner. More specifically, let  $s$  denote the Laplace transform of  $\partial_t$ .<sup>1</sup> Then causal realization of  $\partial_t$  is equivalent to requiring that  $\text{Re}(s) > 0$ . To test the transformed operator we set  $\text{Re}(s) > 0$  everywhere and check to see if the result has a positive real part. If it does, it is a CPR operator.

As an example consider the 15-degree operator:

---

<sup>1</sup> In this paper, subscripting of  $\partial$  will denote partial differentiation with respect to the subscript, i.e.  $\partial_t = \partial/\partial t$ . Superscripting denotes the inverse of subscripting, i.e. integration:  $\partial^q = 1/(\partial/\partial q)$ .

$$L_{15} = \frac{1}{v} \partial_t - \frac{v}{2} \partial_{xx}^2 \quad (1)$$

Laplace-transform over  $t$  and Fourier-transform over  $x$  to obtain

$$\hat{L}_{15} = \frac{s}{v} + \frac{v}{2s} k_x^2 \quad (2)$$

Since  $k_x$  is a purely real quantity, and by assuming causality  $\text{Re}(s) > 0$ , it is clear that  $\text{Re}(\hat{L}_{15}) > 0$ , and hence the 15-degree operator is CPR. The usefulness of CPR operators is easy to see, also. Let us solve the problem

$$\frac{\partial p}{\partial z} = -L_{15} p \quad (3)$$

with initial conditions

$$p(x, t, z=0) = p_0(x, t)$$

Transforming over  $t$  and  $x$ , we get

$$\frac{\partial \hat{p}}{\partial z} = -\hat{L}_{15} \hat{p} \quad ; \quad \hat{p}(z=0) = \hat{p}_0 \quad (4)$$

The solution is

$$\hat{p}(z) = e^{-\hat{L}_{15} z} \hat{p}_0 \quad (5)$$

$$= e^{-\text{Re}(\hat{L}_{15})z} e^{-i\text{Im}(\hat{L}_{15})z} \hat{p}_0$$

Since  $\text{Re}(\hat{L}_{15}) > 0$ , it is clear that the solution does not blow up, so the differential equation (3) is stable.



increasing function of  $z$ , or equivalently that

$$\frac{\partial}{\partial z} \|q\|^2 \leq 0. \quad (10)$$

Substituting (8) into (10), we see that this is equivalent to the requirement that

$$q^*(\tilde{M} + \tilde{M}^*)q \geq 0 \quad \text{for all possible choices of } q, \quad (11)$$

i.e.  $\tilde{M} + \tilde{M}^*$  must be a positive semi-definite matrix. If  $M$  happens to have a full set of orthonormal eigenvectors, this is equivalent to the requirement that the real part of the eigenvalues of  $M$  be non-negative. This is very much like the condition for a CPR operator. In fact we only need add the condition that  $\text{Re}(s) \geq 0$  in equation (8) to assure that (6) is a stable equation for extrapolation in  $z$  and  $t$ .

So far we have shown that the conditions for stability carry over to the matrix case. Next we would like to find out what the rules of combination are for the matrix case. Recall that for the scalar case Muir's rules are (see SEP-16, p. 143)

i)	Multiplication by a positive scalar	$R' = a R$	if $a > 0$
ii)	Inversion	$R' = 1/R$	
iii)	Addition	$R' = R_1 + R_2$	

where we mean that  $R$ ,  $R_1$  and  $R_2$  being CPR operators implies that  $R'$  is a CPR operator.

For the matrix case, we want the following to be true: if  $M$  is a matrix CPR operator, then by using one of the allowed rules of combination we want to get another CPR operator  $M'$ . This is equivalent to saying that if  $M + M^* \geq 0$ ,<sup>1</sup> then  $M' + M'^* \geq 0$ . As we will demonstrate

---

<sup>1</sup>  $M + M^* \geq 0$  is shorthand for " $x^*(M + M^*)x \geq 0$  for all  $x$ ,"  
i.e.  $M + M^*$  is a positive semi-definite matrix.

later, the following are rules for the combination of matrix causal positive real (MCPR) operators. They are somewhat more restrictive than the rules for the scalar case.

*Muir's Rules for Matrices*

- |    |   |                  |   |
|----|---|------------------|---|
| 1) | Multiplication<br>by a positive<br>definite<br>matrix | $M' = T M$       | if $T = T^* > 0$ and $M$ and $T$ can be diagonalized by the same similarity transformation (this implies as well that $M$ must be a normal matrix).   |
| 2) | Inversion   | $M' = M^{-1}$    | if and only if $MM^* = M^*M$ , i.e. $M$ is a normal matrix. Note that, equivalently, $M$ must be diagonalizable by a unitary transformation. This will be possible if $M$ has a full set of orthonormal eigenvectors. |
| 3) | Addition  | $M' = M_1 + M_2$ | (no restrictions on this rule)  |

The first rule says that an MCPR operator  $M$  can be multiplied by a positive definite Hermitian matrix  $T$  and yield an MCPR operator if  $M$  and  $T$  share the same eigenvectors. The requirement that  $M$  and  $T$  have the same eigenvectors is rather restrictive, although it is probably not really necessary. The example given below falls into this category, however, so for our purposes it is not too severe a restriction. It is also possible to get around the requirement that  $T$  be a Hermitian matrix. The appropriate tricks to do this are also discussed in the example below.

Since  $T$  is Hermitian, if  $T$  and  $M$  are to share the same eigenvectors,  $M$  must be a *normal matrix*, i.e. it must be diagonalizable by a unitary transformation. (A unitary transformation is a similarity transformation where the transformation matrix is unitary.) An equivalent condition is that  $MM^* = M^*M$ . The requirement that  $M$  be a normal matrix also comes up in the inversion rule, so this again is not really an additional restriction.

It should be pointed out that the requirement that all MCPR matrices we consider be normal matrices implies that those matrices must be diagonalizable. Diagonalization of the operators amounts to reducing the matrix problem to a scalar problem to which we can then apply the scalar theory. So Muir's rules for matrices can be summarized by saying that if the problem can be diagonalized then Muir's (old) rules will apply. However, the formulation of the rules in terms of matrices can turn out to be more convenient to use since then we need only do simple checks of matrix properties rather than worry about actually diagonalizing the problem and checking eigenvalues.

### *Proof of Rule 1*

If  $T$  is a Hermitian matrix, then there is a unitary matrix  $Q$  such that

$$T = Q^* \Lambda_t Q \quad (12)$$

where  $\Lambda_t$  is a diagonal matrix containing the eigenvalues of  $T$ . If  $M$  can be diagonalized by this same unitary transformation then we can find  $\Lambda_m$ , another diagonal matrix, such that

$$M = Q^* \Lambda_m Q. \quad (13)$$

Now note that since  $T$  and  $M$  share the same eigenvectors, they commute with each other, and hence since  $T$  is Hermitian,  $TM + (TM)^* = T(M + M^*)$ . Using (12) and (13), we have therefore that

$$TM + (TM)^* = Q^* \Lambda_t (\Lambda_m + \Lambda_m^*). \quad (14)$$

Since  $M + M^* > 0$  and  $T = T^* > 0$ , it is clear that the right-hand side of (14) is a positive definite matrix, and hence

$$TM + (TM)^* > 0 \quad (15)$$

**Proof of Rule 2**

(I'll just show the "sufficient" part of the proof.) If  $M$  is a normal matrix, then we can find a unitary matrix  $Q$  so that (13) holds. Then

$$M^{-1} + M^{-1*} = Q^*(\Lambda_m^{-1} + \Lambda_m^{-1*})Q \quad (16)$$

Since  $\Lambda_m + \Lambda_m^* > 0$  by assumption,  $M^{-1} + M^{-1*} > 0$ .

The proof of Rule 3 is obvious ("left to the reader").

Now we will show that these three rules can be used to construct Muir's square root recursion formula for one-way wave equation operators. For the matrix case, Muir's recursion can be written as

$$S_{n+1} = sI + (sI + S_n)^{-1}T \quad (17)$$

where  $T = T^* > 0$  is a (Hermitian) positive definite matrix and  $s$  is again the Laplace transform of  $\partial_t$ .  $S_n$  is an approximation to  $S_\infty = (s^2I + T)^{\frac{1}{2}}$  and is essentially the operator on the right-hand side of a one-way wave equation. The usual "starting value" for the recursion is  $S_0 = sI$ . We will use a standard inductive proof: If we take  $S_0$  to be causal, i.e.  $\text{Re}(s) > 0$ , then  $\text{Re}(S_0) > 0$  and so  $S_0$  is an MCPR operator. The formula for  $S_1$  is

$$S_1 = sI + \frac{1}{2s}T \quad (18)$$

We can consider  $2s$  to be a diagonal MCPR matrix. By rule 2,  $1/2s$  is also MCPR. By rule 1,  $(1/2s)T$  is MCPR, because since  $1/2s$  is diagonal, we can choose any vectors we like for its eigenvectors, so we will use the eigenvectors of  $T$  (thus satisfying the requirement that  $T$  and  $1/2s$  have the same eigenvectors). By rule 3,  $S_1$  is therefore an MCPR operator. Now observe that  $S_1$  has the same eigenvectors as  $T$  since it is the sum of a diagonal matrix and a matrix with those eigenvectors. In fact, it is easy to see by inspection of equation (17) that all of





and then

$$\frac{\partial}{\partial z} \|q\|^2 = -q^* \Lambda^{-\frac{1}{2}} (S_n + S_n^*) \Lambda^{-\frac{1}{2}} q \quad (22)$$

The trick is the following: since  $x^* (S_n + S_n^*) x > 0$  for all choices of  $x \neq 0$  when  $\text{Re}(s) > 0$ , then obviously we can write  $x = \Lambda^{-\frac{1}{2}} q$ . And since  $\Lambda^{-\frac{1}{2}} > 0$ , the right-hand side of (22) must be negative for all nonzero choices of  $q$ .

An equivalent way of looking at this is to begin by defining a new norm. A vector norm is essentially just any positive definite quantity which can be defined as a measure of the size of a vector. Any book on matrix theory will give all the specifics on this. Up until now we have used the definition given in equation (9) for a norm. An equally good definition, however, is the following:

$$\|q\|_{\Lambda}^2 \equiv q^* \Lambda q \quad (23)$$

Since velocity is always a positive quantity,  $\|q\|_{\Lambda}^2$  will always be a non-negative quantity, which is essentially the reason that this definition can be used. Now by substituting in from equation (19), we have that

$$\frac{\partial}{\partial z} \|q\|_{\Lambda}^2 = -q^* (S_n + S_n^*) q \quad (24)$$

from which it follows immediately that

$$\frac{\partial}{\partial z} \|q\|_{\Lambda}^2 \leq 0 \quad (25)$$

and so the differential equation is stable.

### Internal Boundary Conditions

When we solve an equation like (19) for a medium where the velocity function has discontinuities in the lateral direction, the difference approximation must be implicitly satisfying some "internal boundary conditions" at the interface in the medium. As an example, let's consider solving a downward-continuation problem in a medium with a vertical interface at  $x = 0$  using the 15-degree equation  $p_z = -\Delta^{-1} S_1 p$  (see figure 1). If we were to do this using a pure differential equation, then the 15-degree equation would be

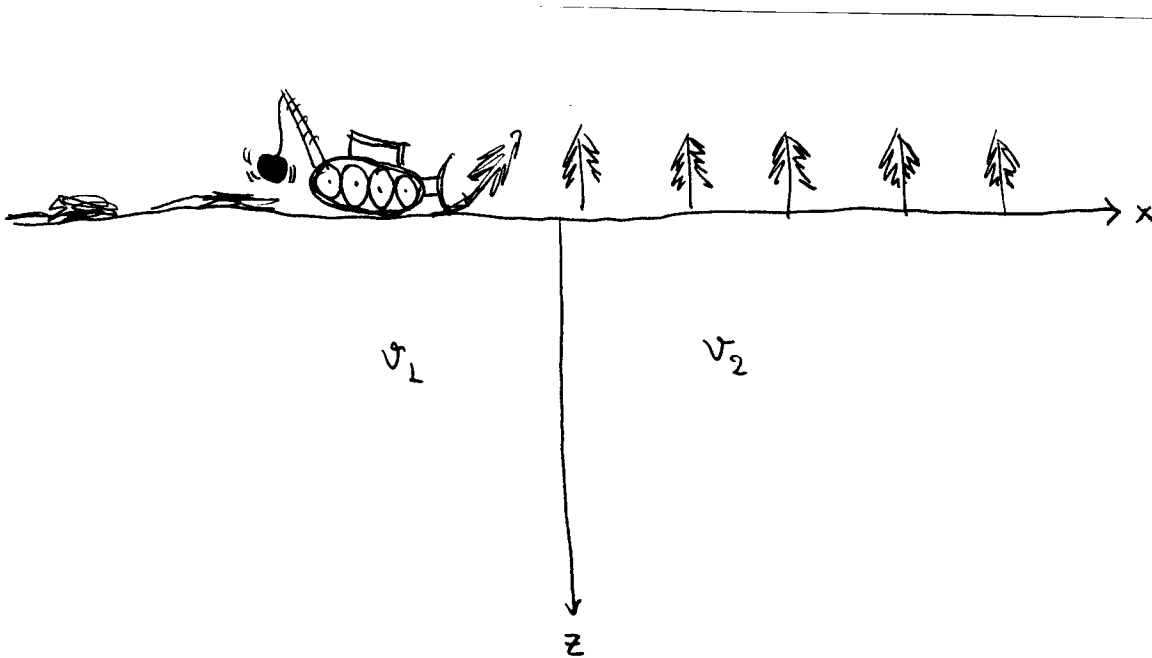


FIG. 1. Earth model with a vertical interface at  $x = 0$ .

$$p_z = \frac{-s}{v_1} p + \frac{v_1}{2s} p_{xx} \quad (26a)$$

in the medium on the left and

$$p_z = \frac{-s}{v_2} p + \frac{v_2}{2s} p_{xx} \quad (26b)$$

in the medium on the right. If, for instance, we were using an acoustic earth model where the compressibility was assumed constant across the interface, and so the velocity discontinuity was due only to a change of density across the interface, then the appropriate interface conditions would be

$$[p(0)] = 0 \quad (27a)$$

and

$$[v^2 p_x(0)] = 0 \quad (27b)$$

where  $[q(x)] = q(x+) - q(x-)$  is the jump in  $q$  at  $x$ . The analytic solution of (26) cannot be determined unless we use these extra conditions (27). When we discretize the problem in  $x$ , however, the resulting system of equations can be solved without specifying any special conditions at the interface. This must mean that when solving the discretized system, we are implicitly satisfying some conditions at the interface. It might be nice, for instance, if those implicit conditions were some approximation to (27).

It turns out that these implicit interface conditions are determined by the choice for the tridiagonal matrix  $T$  in equation (19). One of the choices for  $T$  given by Godfrey, Muir, and Claerbout (SEP-16, p.86) corresponds to discretizing the equation

$$p_z = -\frac{1}{v(x)} \left( s - \frac{1}{2s} \partial_x (v^2(x) \partial_x) \right) p \quad (28)$$

in the  $x$ -direction, giving

$$\frac{\partial p(x)}{\partial z} = -\frac{1}{v(x)} \left\{ s p(x) - \frac{1}{2s \Delta x^2} (v^2(x + \frac{\Delta x}{2})) p(x + \Delta x) \right. \quad (29)$$

$$- \left. \left( v^2 \left( x + \frac{\Delta x}{2} \right) + v^2 \left( x - \frac{\Delta x}{2} \right) \right) p(x) + v^2 \left( x - \frac{\Delta x}{2} \right) p(x - \Delta x) \right\}$$

To find out what the implicit interface conditions are, look at equation (29) for  $x = 0$  and  $x = \Delta x$ . Using the velocity model given above, we get, by taking the first term on the right-hand side over to the left-hand side and multiplying through by  $2s\Delta x^2$ ,

$$v_2^2 p(\Delta x) - (v_1^2 + v_2^2) p(0) + v_1^2 p(-\Delta x) = 0(\Delta x^2) \quad (30a)$$

and

$$v_2^2 p(2\Delta x) - 2v_2^2 p(\Delta x) + v_2^2 p(0) = 0(\Delta x^2) \quad (30b)$$

If we add these two equations together, we get

$$v_2^2 (p(2\Delta x) - p(\Delta x)) = v_1^2 (p(0) - p(-\Delta x)) + 0(\Delta x^2) \quad (31)$$

which is at least a first-order approximation to (27b). Subtracting (30b) from (30a) yields

$$v_2^2 p(2\Delta x) - 3v_2^2 p(\Delta x) + (v_1^2 + 2v_2^2) p(0) - v_1^2 p(-\Delta x) = 0(\Delta x^2) \quad (32)$$

We can expand  $p(2\Delta x)$  and  $p(\Delta x)$  in Taylor series in terms of  $p(0+)$  and  $p(-\Delta x)$  in terms of  $p(0-)$ . The term  $p(0)$  can be written as  $p(0) = \frac{1}{2}(p(0+) + p(0-))$ . Then (32) becomes

$$\left[ \frac{v_1^2}{2} - v_2^2 \right] p(0+) + \left[ v_2^2 - \frac{v_1^2}{2} \right] p(0-) = 0(\Delta x) \quad (33)$$

which approximates equation (27a).

Other choices for the tridiagonal matrix will result in different interface conditions. For example, any straightforward difference approximation to

$$p_z = -\frac{1}{v(x)} \left( s - \frac{1}{2s} v^2(x) \partial_{xx} \right) p \quad (34)$$

will implicitly specify the conditions

$$[p] = 0 \quad \text{and} \quad [p_x] = 0 \quad (35)$$

wherever discontinuities of the velocity function occur.

The consideration of internal boundary conditions can offer another point of view for bullet-proofing. By "bullet-proofing" we have come to mean that we are stabilizing a differential equation which, although stable in the constant velocity case, becomes unstable when lateral velocity gradients become too large. In particular, we are concerned with destabilizing effects that vertical discontinuities in the velocity function have on the difference approximations. Since the introduction of interfaces will result in reflection and transmission effects at those interfaces, it is possible to consider bullet-proofing as a process which assures that the transmissions and reflections at these interfaces do not become too large. In other words, if we make sure that the implicit interface conditions are associated with reasonable reflection and transmission coefficients, then the difference approximation can be considered "bullet-proof." Thus, we expect (29) to be a bullet-proof approximation because the interface conditions associated with it are the same as those conditions for the true physical problem. Here at Stanford we have found the internal boundary condition point of view for stability to be a useful one when dealing with one-way elastic equations (see article(s) by Clayton and Brown, this report).

### *Side Boundary Conditions*

An assumption we have made in the arguments above for the stability of matrix systems is that there are no side boundary conditions. Although this is a common simplifying assumption in any work of this kind, it would be nice if, for instance, side boundary conditions could be incorporated into Muir's Rules for Matrices. As of yet, this has not been done, but some work for the 15-degree equation, given below, may

prove to be a useful first step.

Let us consider the problem of downward continuation in a semi-infinite medium  $x \geq 0$  using the 15-degree equation,

$$\Delta \frac{\partial p}{\partial z} = -s_1 p. \quad (36)$$

Here,

$$s_1 = s\tilde{I} + \frac{1}{2s}T \quad (37)$$

where now,  $T$  is not a Hermitian matrix, but consists of a Hermitian part plus a non-Hermitian part which contains nearly all zeros except in the upper left-hand corner:

$$\Delta x^2 T = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ -v_1^2 & v_1^2 + v_2^2 & -v_2^2 & 0 & 0 \\ 0 & -v_2^2 & v_2^2 + v_3^2 & -v_3^2 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \text{O} & \text{O} & \text{O} & \text{O} & \text{O} \end{pmatrix} = \Delta x^2 (\tilde{T} + A) \quad (38)$$

where  $\tilde{T} = \tilde{T}^*$ , and hence

$$\Delta x^2 A = \begin{pmatrix} -2v_1^2 & v_1^2 & \text{O} \\ 0 & 0 & \text{O} \\ \text{O} & \text{O} & \text{O} \\ \text{O} & \text{O} & \text{O} \end{pmatrix} \quad (39)$$

The semi-infinite vector  $p$  is given by  $p = (p_0, p_1, p_2, \dots)^T$ ,





$$\begin{pmatrix} \bar{p}_0 & \bar{p}_1 \end{pmatrix} \begin{pmatrix} \frac{2 \operatorname{Re}(s)}{|s|^2} & \frac{\bar{s}}{2|s|^2} \\ \frac{s}{2|s|^2} & 0 \end{pmatrix} \begin{pmatrix} p_0 \\ p_1 \end{pmatrix} \\ = \frac{1}{2|s|^2} (\bar{s} \bar{p}_0 p_1 + s p_0 \bar{p}_1) - \frac{\operatorname{Re}(s)}{|s|^2} \bar{p}_0 p_0 \geq 0 \quad (44)$$

An oft-used absorbing boundary condition is

$$\frac{p_1 - p_0}{\Delta x} = a p_0 \quad (45)$$

Substituting (45) into (44) to eliminate  $p_1$ , we get

$$a \bar{p}_0 p_0 \geq 0 \quad (46)$$

which means that if  $a \geq 0$ , (45) is an absorbing boundary condition and (42) is satisfied.

In summary, it seems that when side boundaries are present, in order to get a stability estimate like (42), the side boundary conditions must be used to take care of the slightly non-CPR parts of the difference operator. The side boundary problem deserves some further work.

## REFERENCES

- Brown, D., 1978, Causal positive real operators: SEP-15, p. 203.
- Bloxsom, H. and Kjartansson, E., 1978, Steeply faulted regions, artifacts, and stability in the 45-degree  $\omega$ -finite difference algorithm: SEP-15, p. 21.
- Claerbout, J. F., 1976, Slanted waves in slanted frames: SEP-8, p.17.
- Claerbout, J. F., 1979, Impedance, reflectance and transference functions: SEP-16, p.141.
- Clayton, R. and Brown, D., 1979, Choice of variables in elastic wave extrapolation: this report.
- Engquist, B., 1976, Well-posedness of one-way wave equations: SEP-8, p. 48.
- Godfrey, B., Muir, F. and Claerbout, J. F., 1979, Stable extrapolation: SEP-16, p. 83.
- Jacobs, B., Godfrey, B., Muir, F. and Claerbout, J. F., 1979, Bullet-proofing the code for the 45-degree equation: SEP-16, p. 89.

Broom Hilda / Russell Myers

