

# Shapewise Equivalence and the Helmholtz Equation

David S. Spiegel

Submitted to: *The Journal of Mathematical Physics*, 8 January 2001

## Abstract

Separation of variables is by far the most widely used method of solving linear partial differential equations. The Helmholtz equation ( $\frac{\partial^2 \Psi}{\partial x^2} + \frac{\partial^2 \Psi}{\partial y^2} + \omega^2 \Psi = 0$ ) is one of the simplest equations that can be solved by this method. This equation is widely used in physics, and so it has been studied for many years. It is well-known that there are four orthogonal coordinate systems (rectangular, parabolic, polar, and elliptic) up to a certain kind of equivalence in which this equation separates. Both [Miller 1977] and [Morse and Feshbach 1953], among others, derive its four orthogonal separable coordinate systems. These references, however, never define what it means for a partial differential equation to separate in a coordinate system, and are not specific about in what way the other coordinate systems in which the Helmholtz equation is separable are equivalent to the four derived. In this paper, we provide a definition of separation of variables that is applicable in a small but important set of types of equations (a set that, of course, includes the Helmholtz equation in orthogonal coordinates), and we introduce *shapewise equivalence*, the equivalence relation holding between any coordinate system in which the Helmholtz equation is separable and one of the four coordinate systems commonly cited as the Helmholtz equation's separable coordinate systems.

# Shapewise Equivalence and the Helmholtz Equation

## 1 Separation of Variables in Orthogonal Coordinate Systems

When a solution to a partial differential equation is the product of several functions, each of which is a function of just one of the independent variables, most people say that the solution *separates*, and we may find this solution to the differential equation through the method of *separation of variables*. We will call this the naive notion of separation of variables, and we will revise this description later in the paper. One of the simplest equations that is solvable via this method is the Helmholtz equation:

$$\mathcal{H}\Psi(x, y) = (\Delta_2 + \omega^2)\Psi(x, y) = 0, \quad (1)$$

where  $\mathcal{H} = \Delta_2 + \omega^2$ , with  $\omega \in \mathbf{R} - \{0\}$ , is the so-called *Helmholtz operator*. Note: for the remainder of this paper,  $\mathcal{F}$  will denote the set of all  $\mathbf{C}$ -valued real-analytic functions on  $\Omega$ , a non-empty open, connected subset of  $\mathbf{R}^2$ .

The Helmholtz equation separates in rectangular coordinates. It is natural to inquire in what other coordinate systems the Helmholtz equation separates. We will examine only orthogonal curvilinear coordinates. Consider a new set of orthogonal coordinates  $\{u, v\}$  so that (i) there are real-analytic functions  $u$  and  $v : \Omega \rightarrow \mathbf{R}$ , where  $\Omega$  is a non-empty, open, connected subset of  $\mathbf{R}^2$ , such that there is a bijection between pairs  $(x, y)$  and pairs  $(u(x, y), v(x, y))$ , and (ii) the Jacobian determinant  $v_x u_y - u_x v_y$  never vanishes.

In order to write the Helmholtz equation in another coordinate system  $\{u, v\}$ , we must transform the Helmholtz operator into the new coordinate system. If  $\psi$  is a real-analytic function on the plane ( $\psi \in \mathcal{F}$ ) then the chain rule gives us

$$\partial_{xx}\psi = (u_x^2 \partial_{uu} + u_{xx} \partial_u + 2u_x v_x \partial_{uv} + v_x^2 \partial_{vv} + v_{xx} \partial_v) \psi.$$

The expression for  $\partial_{yy}\psi$  is similar, and so we have that in the  $\{u, v\}$  system,

$$\Delta_2 = (u_x^2 + u_y^2) \partial_{uu} + (u_{xx} + u_{yy}) \partial_u + 2(u_x v_x + u_y v_y) \partial_{uv} + (v_x^2 + v_y^2) \partial_{vv} + (v_{xx} + v_{yy}) \partial_v. \quad (2)$$

The condition that the system is orthogonal means that the unit vectors in the  $u$  and  $v$  directions ( $\hat{\mathbf{u}}$  and  $\hat{\mathbf{v}}$ , respectively) are perpendicular at every point, which means that  $\hat{\mathbf{u}}(x, y) \cdot \hat{\mathbf{v}}(x, y) \equiv 0$ . Observe that  $\hat{\mathbf{u}}$  is a unit vector in the direction of maximum increase of  $u$  — i.e., in the direction of  $\nabla(u) = (u_x, u_y)$ ; and similarly  $\hat{\mathbf{v}}$  is a unit vector in the direction of  $\nabla(v) = (v_x, v_y)$ . Since the dot-product of  $\hat{\mathbf{u}}$  and  $\hat{\mathbf{v}}$  is 0, the dot-product  $\nabla(u) \cdot \nabla(v) = u_x v_x + u_y v_y$  must equal 0 as well. As a result, in an orthogonal coordinate system, the coefficient of the mixed second partial derivative vanishes:

$$\Delta_2 = (u_x^2 + u_y^2) \partial_{uu} + (u_{xx} + u_{yy}) \partial_u + (v_x^2 + v_y^2) \partial_{vv} + (v_{xx} + v_{yy}) \partial_v. \quad (3)$$

And so, since  $\omega^2$  is unchanged under the coordinate transformation, (1) is now written

$$\mathcal{H}\Psi = [(u_x^2 + u_y^2) \partial_{uu} + (u_{xx} + u_{yy}) \partial_u + (v_x^2 + v_y^2) \partial_{vv} + (v_{xx} + v_{yy}) \partial_v + \omega^2] \Psi = 0. \quad (4)$$

For convenience in viewing, we will write (4) as follows:

$$\mathcal{H}\Psi = (A_{11} \partial_{uu} + A_1 \partial_u + A_{22} \partial_{vv} + A_2 \partial_v + \omega^2) \Psi = 0, \quad (5)$$

where the functions  $A_{11}, A_1, A_{22}$ , and  $A_2$  are equal to the corresponding coefficients in (4).

There are certain well-known “facts” concerning the form that the coefficient functions must assume if an equation of the form (5) is to be separable. For instance, [Miller 1977, p. 14] states that, assuming  $\omega^2 \neq 0$ , (5) is separable only if there exist functions  $\mathcal{U}, \mathcal{V}, \mathcal{U}_1$ , and  $\mathcal{V}_1$  such that

$$(a) \quad A_{11} = \frac{\mathcal{U}(u)}{\mathcal{U}_1(u) + \mathcal{V}_1(v)} \quad \text{and} \quad (b) \quad A_{22} = \frac{\mathcal{V}(v)}{\mathcal{U}_1(u) + \mathcal{V}_1(v)}. \quad (6)$$

Also, [Robertson 1928] asserts that if (5) is separable, then there exist functions  $B(u)$  and  $C(v)$  such that

$$(c) \quad \frac{A_1(u, v)}{A_{11}(u, v)} = B(u) \quad \text{and} \quad (d) \quad \frac{A_2(u, v)}{A_{22}(u, v)} = C(v). \quad (7)$$

These “facts” prove very useful when characterizing the separable coordinate systems of (5), and they hold true in all separable coordinate systems that are commonly used in practice. But, if one adopts the naive notion of separability that we present at the begining of this paper, (6) and (7) are not always true.<sup>1</sup> The following counterexample demonstrates this.

### Example 1

Let  $\Psi_0(u, v) = e^u e^v$ , and let the coefficient functions of the operator  $(A_{11}\partial_{uu} + A_1\partial_u + A_{22}\partial_{vv} + A_2\partial_v + \omega^2)$  be defined by

$$\begin{aligned} A_{11} &= u - v & A_1 &= v \\ A_{22} &= -u - v & A_2 &= -\omega^2 + v. \end{aligned}$$

One can easily verify that  $\Psi_0$  is a solution. Since  $\Psi_0$  is a product of a function only of  $u$  and a function only of  $v$ , it is a separated solution according to the naive definition. Yet, it is clear that none of (a) — (d) in (6) and (7) hold for the above coefficients.

Since (6) and (7) are so useful, we wish to define separation so that they do hold.

**Definition 1** *Consider a second-order partial differential equation that involves two independent variables and has no mixed partial derivatives. That is, our partial differential equation is of the form*

$$(A_{11}\partial_{uu} + A_1\partial_u + A_{22}\partial_{vv} + A_2\partial_v + A)\Psi = 0,$$

*where  $A_{11}, A_1, A_{22}, A_2$ , and  $A$  are real-analytic functions of  $u$  and  $v$ . We say that this equation is weakly solvable by separation of variables or weakly separable if there exists a solution  $\Psi_0(u, v)$  that can be written  $U_0(u)V_0(v)$ , for some functions  $U_0$  and  $V_0$ , and any such solution is said to be weakly separated. We say that this equation is strongly solvable by separation of variables or strongly separable if  $\frac{A_1}{A_{11}}$  is a function only of  $u$ ,  $\frac{A_2}{A_{22}}$  is a function only of  $v$ , and the equation has at least five linearly independent weakly separated solutions.*

It is not immediately clear that the above definition implies (6) in the case of the Helmholtz equation, but we will show that indeed it does.

**Lemma 1** *Let a partial differential equation of the form*

$$(A_{11}\partial_{uu} + A_1\partial_u + A_{22}\partial_{vv} + A_2\partial_v + A)\Psi = 0 \quad (8)$$

*be strongly solvable by separation of variables, and assume that  $A_{11}$  and  $A_{22}$  are always positive. Then*

$$\frac{A_{11}(u, v)}{A_{22}(u, v)} = \frac{\mathcal{U}(u)}{\mathcal{V}(v)},$$

*for some  $\mathcal{U} > 0$  that is a function only of  $u$  and some  $\mathcal{V} > 0$  that is a function only of  $v$ .*

<sup>1</sup>In [Hildebrand 1976, p. 505], we encounter another characterization of the coefficients in a separable equation that is useful but not true in general if one adopts the naive definition of separability.

*Proof:* Let us write our differential equation as  $(A_{11}\partial_{uu} + A_1\partial_u + A_{22}\partial_{vv} + A_2\partial_v + A)UV = 0$ , where  $\Psi = UV$  represents a weakly separated, nonzero solution. This implies that

$$A_{11}U''V + A_1U'V + A_{22}UV'' + A_2UV' + AUV = 0. \quad (9)$$

We divide through by  $A_{11}V$  and rearrange terms to arrive at

$$U'' + \frac{A_1}{A_{11}}U' + \left( \frac{A_{22}}{A_{11}} \frac{V''}{V} + \frac{A_2}{A_{11}} \frac{V'}{V} + \frac{A}{A_{11}} \right) U = 0. \quad (10)$$

We also re-group terms to arrive at

$$V'' + \frac{A_2}{A_{22}}V' + \left( \frac{A_{11}}{A_{22}} \frac{U''}{U} + \frac{A_1}{A_{22}} \frac{V'}{V} + \frac{A}{A_{22}} \right) V = 0. \quad (11)$$

The crucial aspect of (10) is that, since  $U''$ ,  $\frac{A_1}{A_{11}}U'$ , and  $U$  are functions only of  $u$ , so must be  $\frac{A_{22}}{A_{11}} \frac{V''}{V} + \frac{A_2}{A_{11}} \frac{V'}{V} + \frac{A}{A_{11}}$ . Write

$$\mathcal{U}_0(u) = \frac{A_{22}}{A_{11}} \frac{V''}{V} + \frac{A_2}{A_{11}} \frac{V'}{V} + \frac{A}{A_{11}}, \quad (12)$$

for some  $\mathcal{U}_0$  that is a function only of  $u$ . Similarly, from (11), we see that there exists some  $\mathcal{V}_0$  that is a function only of  $v$  such that

$$\mathcal{V}_0(v) = \frac{A_{11}}{A_{22}} \frac{U''}{U} + \frac{A_1}{A_{22}} \frac{V'}{V} + \frac{A}{A_{22}}. \quad (13)$$

We multiply (10) by  $A_{11}V$  and (11) by  $A_{22}U$ , and we add the results:

$$A_{11}U''V + A_1U'V + A_{22}UV'' + A_2UV' + (A_{11}\mathcal{U}_0 + A_{22}\mathcal{V}_0)UV = 0. \quad (14)$$

Comparing this to (9), we have  $(A - (A_{11}\mathcal{U}_0 + A_{22}\mathcal{V}_0))UV = 0$ , and so, since  $\mathcal{F}$  is an integral domain and  $UV \not\equiv 0$ , we see that  $A$  must equal  $A_{11}\mathcal{U}_0 + A_{22}\mathcal{V}_0$ :

$$A \equiv A_{11}\mathcal{U}_0 + A_{22}\mathcal{V}_0. \quad (15)$$

Since, for a particular choice of  $\mathcal{U}_0$  and  $\mathcal{V}_0$ , both (10) and (11) are linear, homogeneous, second-order ordinary differential equations, each has at most two linearly independent solutions. We will label these  $U_{01}, U_{02}$ , and  $V_{01}, V_{02}$ . Thus, for a particular choice of  $\mathcal{U}_0$  and  $\mathcal{V}_0$ , there are at most four linearly independent solutions to (9):  $U_{01}V_{01}, U_{01}V_{02}, U_{02}V_{01}$ , and  $U_{02}V_{02}$ . Since (8) is strongly separable, Definition 1 requires that there be at least five linearly independent weakly separated solutions. By the preceding argument, the existence of a fifth linearly independent weakly separated solution necessitates the existence of a pair of solutions  $(\mathcal{U}_1, \mathcal{V}_1) \neq (\mathcal{U}_0, \mathcal{V}_0)$  that satisfies (12) and (13). Suppose, without loss of generality, that  $\mathcal{V}_1 \neq \mathcal{V}_0$ . We claim that  $\mathcal{U}_1 \neq \mathcal{U}_0$ .

Assume for the moment that  $\mathcal{U}_1 = \mathcal{U}_0$ . If this is so, then by following the same procedure that led to (14), we see that

$$A \equiv A_{11}\mathcal{U}_0 + A_{22}\mathcal{V}_1. \quad (16)$$

Taking the difference between (16) and (15) yields

$$0 = A_{11}(\mathcal{U}_0 - \mathcal{U}_0) + A_{22}(\mathcal{V}_1 - \mathcal{V}_0),$$

and hence

$$0 = A_{22}(\mathcal{V}_1 - \mathcal{V}_0).$$

Since, by supposition,  $\mathcal{V}_1 - \mathcal{V}_0 \not\equiv 0$ , this forces  $A_{22} \equiv 0$ .  $\implies \Leftarrow$ ! This contradicts the hypothesis that  $A_{22}$  is everywhere positive. Thus, we must have  $\mathcal{U}_1 \neq \mathcal{U}_0$ . We can therefore correctly rewrite (16) as  $A \equiv A_{11}\mathcal{U}_1 + A_{22}\mathcal{V}_1$ . Subtracting (15) from this gives us

$$0 = A_{11} (\mathcal{U}_1 - \mathcal{U}_0) + A_{22} (\mathcal{V}_1 - \mathcal{V}_0). \quad (17)$$

For no value of  $u$  does  $\mathcal{U}_1(u) = \mathcal{U}_0(u)$ ; and for no value of  $v$  does  $\mathcal{V}_1(v) = \mathcal{V}_0(v)$ . Why? Well, suppose for instance that for  $v = v_0$ , we have  $\mathcal{V}_1(v_0) = \mathcal{V}_0(v_0)$ . Then, we see from (17) that for all  $u$ ,  $A_{11}(u, v_0)(\mathcal{U}_1(u) - \mathcal{U}_0(u)) = 0$ , but then, since  $\mathcal{F}$  is an integral domain and  $A_{11} > 0$ , this forces  $\mathcal{U}_1 \equiv \mathcal{U}_0$ .  $\implies \Leftarrow$ ! Therefore, we may rearrange the terms in (17) to get

$$\frac{A_{11}}{A_{22}} = \frac{\mathcal{V}_1 - \mathcal{V}_0}{\mathcal{U}_0 - \mathcal{U}_1} = \frac{\frac{1}{\mathcal{U}_0 - \mathcal{U}_1}}{\frac{1}{\mathcal{V}_1 - \mathcal{V}_0}}, \quad (18)$$

which is a quotient of a function only of  $u$  and a function only of  $v$ , as required. And, note that (i) both  $\mathcal{U}(u) \equiv \frac{1}{\mathcal{U}_0(u) - \mathcal{U}_1(u)}$  and  $\mathcal{V}(v) \equiv \frac{1}{\mathcal{V}_1(v) - \mathcal{V}_0(v)}$  are defined everywhere, for their denominators never vanish; and (ii) neither  $\mathcal{U}$  nor  $\mathcal{V}$  is ever 0, since neither  $A_{11}$  nor  $A_{22}$  ever vanishes. Since  $A_{11}$  and  $A_{22}$  are both everywhere positive, their quotient is too. Thus, if it happens that  $\mathcal{U} < 0$ , we will have that  $\mathcal{V} < 0$  too. In this case, replace both  $\mathcal{U}$  and  $\mathcal{V}$  with their negatives — this will make them both positive, and will preserve their quotient.  $\diamond$

**Lemma 2** *Let  $A_{11}, A_1, A_{22}$ , and  $A_2$  be real-analytic functions of both  $u$  and  $v$  that satisfy the following conditions:*

- (i)  $\frac{A_1(u, v)}{A_{11}(u, v)} = B(u)$  for some  $B$  that is a function only of  $u$ ,
- (ii)  $\frac{A_2(u, v)}{A_{22}(u, v)} = C(v)$  for some  $C$  that is a function only of  $v$ ,
- and
- (iii)  $\frac{A_{11}(u, v)}{A_{22}(u, v)} = \frac{\mathcal{U}(u)}{\mathcal{V}(v)}$  for some never-zero  $\mathcal{U}$  that is a function only of  $u$  and some never-zero  $\mathcal{V}$  that is a function only of  $v$ .

Consider a differential equation that has the following form:

$$A_{11} \frac{U''}{U} + A_1 \frac{U'}{U} + A_{22} \frac{V''}{V} + A_2 \frac{V'}{V} + \omega^2 = 0,$$

where  $U$  is a real-analytic function only of  $u$  and  $V$  is a real-analytic function only of  $v$ . Then, if  $\omega^2 \neq 0$ , there exist functions  $\mathcal{U}_1$  of  $u$  alone and  $\mathcal{V}_1$  of  $v$  alone such that

$$(a) \quad A_{11} = \frac{\mathcal{U}(u)}{\mathcal{U}_1(u) + \mathcal{V}_1(v)} \quad \text{and} \quad (b) \quad A_{22} = \frac{\mathcal{V}(v)}{\mathcal{U}_1(u) + \mathcal{V}_1(v)}.$$

*Proof:* From hypotheses (i) and (ii) of the lemma,  $A_1 = A_{11}B$  and  $A_2 = A_{22}C$ . Thus, we may rewrite the equation as follows:

$$A_{11} \left( \frac{U''}{U} + B \frac{U'}{U} \right) + A_{22} \left( \frac{V''}{V} + C \frac{V'}{V} \right) + \omega^2 = 0.$$

Notice that  $\frac{U''}{U} + B \frac{U'}{U}$  is a function only of  $u$ , so we may write it  $b(u)$ ; and  $\frac{V''}{V} + C \frac{V'}{V}$  is a function only of  $v$ , so we may write it  $c(v)$ . From hypothesis (iii) above,  $A_{22} = A_{11} \frac{\mathcal{V}}{\mathcal{U}}$ . Thus, we again rewrite the equation as follows:

$$A_{11}b + A_{11} \frac{\mathcal{V}}{\mathcal{U}} c + \omega^2 = 0,$$

which implies that

$$A_{11} = \frac{-\omega^2}{b + c\frac{\mathcal{V}}{\mathcal{U}}} = \frac{-\omega^2\mathcal{U}}{b\mathcal{U} + c\mathcal{V}}.$$

Now, since  $\omega^2 \neq 0$ , we may divide numerator and denominator by  $-\omega^2$ , to arrive at

$$A_{11} = \frac{\mathcal{V}}{\left(-\frac{b\mathcal{U}}{\omega^2}\right) + \left(-\frac{c\mathcal{V}}{\omega^2}\right)}, \quad (19)$$

indicating that our choices of  $\mathcal{U}_1$  and  $\mathcal{V}$  ought to be  $\mathcal{U}_1 = -\frac{b\mathcal{U}}{\omega^2}$  and  $\mathcal{V}_1 = -\frac{c\mathcal{V}}{\omega^2}$ . Note that  $\mathcal{U}_1$  and  $\mathcal{V}_1$  are clearly functions only of  $u$  and only of  $v$ , respectively, as required. This verifies result (a) of the lemma, and it is clear that (19) in conjunction with hypothesis (iii) now implies result (b) of the lemma.  $\diamond$

## 2 Orthogonal Coordinate Systems In Which the Helmholtz Equation is Strongly Separable

We are seeking orthogonal coordinate systems in which (1) is strongly separable, so we will assert (4) to be strongly separable and try to characterize  $u$  and  $v$ . First, we note that the condition that the Jacobian determinant never vanishes implies that  $u_x^2 + u_y^2 = A_{11}$  and  $v_x^2 + v_y^2 = A_{22}$  never vanish either, and hence are everywhere positive. Now, we see immediately that Lemma 1 implies that  $\frac{A_{11}}{A_{22}}$  is of the form required by hypothesis (iii) of Lemma 2, and, since (5) is strongly separable, Definition 1 implies that hypotheses (i) and (ii) of Lemma 2 are satisfied as well. Thus, Lemma 2 implies that

$$u_x^2 + u_y^2 = \frac{\mathcal{U}(u)}{\mathcal{U}_1(u) + \mathcal{V}_1(v)}, \quad \text{and} \quad v_x^2 + v_y^2 = \frac{\mathcal{V}(v)}{\mathcal{U}_1(u) + \mathcal{V}_1(v)}. \quad (20)$$

Note: in the above equation we replaced  $A_{11}$  and  $A_{22}$  in accordance with (4).

Recall that we are working with orthogonal coordinate systems, and so  $u_x v_x + u_y v_y = 0$ . As a result,  $-\frac{v_x}{u_y} = \frac{v_y}{u_x}$ . Define the function  $\mathcal{R} \equiv -\frac{v_x}{u_y} = \frac{v_y}{u_x}$ . Then,

$$v_y = \mathcal{R}u_x \quad (21)$$

and

$$v_x = -\mathcal{R}u_y. \quad (22)$$

Now, since  $v_x^2 + v_y^2 = \mathcal{R}^2(u_x^2 + u_y^2)$ , (20) implies that

$$\mathcal{R}^2 = \frac{\mathcal{V}}{\mathcal{U}}. \quad (23)$$

In order to proceed, we will need to define some new terminology.

**Definition 2** Let  $\{u, v\}$  be a coordinate system on  $\Omega \subseteq \mathbf{R}^2$ . We will say that the image of a curve is a coordinate-curve if it is defined by an equation of the form  $u(x, y) = c$  or  $v(x, y) = c$ , for a number  $c \in \mathbf{R}$ . It is a coordinate-curve of  $u$  if it is defined by  $u = c$  for some  $c \in \mathbf{R}$ , and a coordinate-curve of  $v$  if it is defined by  $v = c$  for some  $c \in \mathbf{R}$ .

**Definition 3** A matrix  $M \in GL(n, \mathbf{R})$  is said to be orthogonal if the linear transformation that it represents preserves lengths and angles.<sup>2</sup> The set of all orthogonal, real  $n \times n$  matrices is denoted  $\mathcal{O}(n)$ .

<sup>2</sup>One can prove that  $M$  is orthogonal iff  $M^t = M^{-1}$ .

**Definition 4** A map  $\mathcal{L} : \mathbf{R}^n \rightarrow \mathbf{R}^n$  is said to be affine if there exist a matrix  $M \in GL(n, \mathbf{R})$  and a vector  $\mathbf{v}_0 \in \mathbf{R}^n$  such that for every vector  $\mathbf{v} \in \mathbf{R}^n$ ,  $\mathcal{L}(\mathbf{v}) = \mathbf{v}_0 + M\mathbf{v}$ . An affine transformation  $\mathcal{L}(\mathbf{v}) = \mathbf{v}_0 + M\mathbf{v}$  is said to be an affine similitude if  $M = cM'$ , where  $c \in \mathbf{R}$  and  $M' \in \mathcal{O}(n)$ .

It turns out that an affine similitude is the composition of a dilation, a translation, a rotation, and possibly a reflection.

**Definition 5** We will say that the coordinate system  $\{u, v\}$  is shapewise equivalent to the coordinate system  $\{\tilde{u}, \tilde{v}\}$  if there exist invertible real-analytic functions  $\eta(u)$  and  $\phi(v)$ , where  $\eta'(u)$  and  $\phi'(v)$  never vanish, such that  $\tilde{u}(x, y) = \eta(u(\mathcal{L}(x, y)))$  and  $\tilde{v}(x, y) = \phi(v(\mathcal{L}(x, y)))$ , where  $\mathcal{L}$  is an affine similitude.

We summarize the relationship that holds among shapewise equivalent coordinate systems in the schematic diagram below.

$$\begin{array}{ccc}
 \mathbf{R}_{\{x, y\}}^2 & \xleftarrow{\mathcal{L}} & \mathbf{R}_{\{x, y\}}^2 \\
 \downarrow & & \downarrow \\
 \mathbf{R}_{\{u, v\}}^2 & \xrightarrow{\eta, \phi} & \mathbf{R}_{\{\tilde{u}, \tilde{v}\}}^2
 \end{array} \tag{24}$$

The vertical maps are given by  $(x, y) \mapsto (u(x, y), v(x, y))$  on the left, and  $(x, y) \mapsto (\tilde{u}(x, y), \tilde{v}(x, y))$  on the right. We say that the diagram is schematic because in order to be rigorous, we would need to replace each  $\mathbf{R}^2$  with an appropriate open set. Still, the diagram makes it clear that one gets the same result going from  $\{x, y\}$  coordinates to  $\{\tilde{u}, \tilde{v}\}$  regardless of whether one uses the direct functions  $\tilde{u}(x, y)$  and  $\tilde{v}(x, y)$  or one follows the path around the other way, using the affine similitude  $\mathcal{L}$ , the functions  $u(x, y)$  and  $v(x, y)$ , and the reparametrizations  $\eta(u)$  and  $\phi(v)$ . The diagram is suggestive of the fact (which we will not prove) that shapewise equivalence is an equivalence relation.

The definition of shapewise equivalence has to do with the analytical, not the geometrical, properties of shapewise equivalent coordinate systems. We define shapewise equivalence this way because the analytical properties are what are useful for proving that a change of coordinates that preserves shapewise equivalence also preserves strong separability of the Helmholtz equation (i.e., if (1) is strongly separable in one orthogonal coordinate system, it is strongly separable in any shapewise equivalent coordinate system). It is the geometrical similarities among coordinate systems that are shapewise equivalent, however, that lead to the term “shapewise equivalent”. We will therefore prove a lemma about the geometrical relationships among shapewise equivalent coordinate systems.

**Lemma 3** Let two coordinate systems  $\{u, v\}$  and  $\{\tilde{u}, \tilde{v}\}$  be shapewise equivalent. Let  $\mathcal{L}$  be the affine similitude that relates  $\{\tilde{u}, \tilde{v}\}$  to  $\{u, v\}$ , as in Definition 5. Then, every non-empty coordinate-curve  $C_u$  of  $u$  is equal to  $\mathcal{L}(C_{\tilde{u}})$  for some non-empty coordinate-curve  $C_{\tilde{u}}$  of  $\tilde{u}$ .

*Proof:* We can prove this most efficiently with the following chain of equivalences.

$$\begin{aligned}
 (x, y) \in \mathcal{L}(C_{\tilde{u}}) & \iff (x, y) = \mathcal{L}(\tilde{x}, \tilde{y}) \quad \text{and} \quad (\tilde{x}, \tilde{y}) \in C_{\tilde{u}} \\
 & \iff (x, y) = \mathcal{L}(\tilde{x}, \tilde{y}) \quad \text{and} \quad \tilde{u}(\tilde{x}, \tilde{y}) = \tilde{c}, \text{ for some } \tilde{c} \in \mathbf{R} \\
 & \iff (x, y) = \mathcal{L}(\tilde{x}, \tilde{y}) \quad \text{and} \quad \eta(u(\mathcal{L}(\tilde{x}, \tilde{y}))) = \tilde{c} \\
 & \iff \eta(u(x, y)) = \tilde{c} \\
 & \iff u(x, y) = c, \text{ where } \eta(c) = \tilde{c} \\
 & \iff (x, y) \in C_u \quad \diamond
 \end{aligned}$$

## Example 2

Frequently, when people graph a data set that includes both some very small and some very large values, they use so-called log graph-paper. Log graph-paper is based on a coordinate system  $\{\tilde{x} = x, \tilde{y} = \log y\}$  that is, by the above argument, shapewise equivalent to rectangular coordinates  $\{x, y\}$ . The reason it is said to be shapewise equivalent should be clear: a rectangle that is parallel to the coordinate axes in ordinary rectangular coordinates is a rectangle on log graph-paper, as well.

**Theorem 1** *Let two coordinate systems  $\{u, v\}$  and  $\{\tilde{u}, \tilde{v}\}$  be shapewise equivalent. Then if the Helmholtz equation is strongly separable in  $\{u, v\}$ , it is also strongly separable in  $\{\tilde{u}, \tilde{v}\}$ , although the transformation of  $\mathcal{H}$  from  $\{u, v\}$  into  $\{\tilde{u}, \tilde{v}\}$  might change  $\omega^2$  into  $\tilde{\omega}^2$ , a different positive quantity.*

*Proof:* Suppose, without loss of generality, that the Helmholtz equation is strongly separable in  $\{u, v\}$ . We will first verify that a reparametrization of  $u$  and  $v$  leaves the equation strongly separable; then, we will verify that under an affine similitude the equation remains strongly separable (although possibly changing  $\omega^2$ ). It will therefore follow that the composition of a reparametrization and an affine similitude leaves the equation strongly separable.

Let there be reparametrizations  $\tilde{u} = \eta(u)$  and  $\tilde{v} = \phi(v)$ . Now,  $A_{11}(u, v) = u_x^2 + u_y^2$ . Note that

$$\tilde{u}_x = \frac{\partial \tilde{u}}{\partial x} = \frac{d\tilde{u}}{du} \frac{\partial u}{\partial x} = \eta'(u) u_x.$$

Similarly,  $\tilde{u}_y = \eta'(u) u_y$ . Thus, we have that the corresponding coefficient

$$\tilde{A}_{11}(\tilde{u}, \tilde{v}) = \tilde{u}_x^2 + \tilde{u}_y^2 = (\eta'(u))^2 (u_x^2 + u_y^2) = (\eta'(u))^2 A_{11}(u, v).$$

Note that  $\eta'(u)$  can be expressed as a function of  $\tilde{u}$  using the inverse function of  $\eta$ . Similarly, one can determine that, if  $A_1(u, v) = u_{xx} + u_{yy}$ , then

$$\tilde{A}_1(\tilde{u}, \tilde{v}) = \eta''(u)(u_{xx} + u_{yy}) = \eta''(u) A_1(u, v).$$

Note, as above, that  $\eta''(u)$  can be expressed as a function of  $\tilde{u}$  using the inverse function of  $\eta$ . Thus, the ratio  $\frac{\tilde{A}_1}{\tilde{A}_{11}}$  equals a function of  $\tilde{u}$  multiplied by  $\frac{A_1}{A_{11}}$ . The latter is a function only of  $u$ . Again, we can use the inverse function of  $\eta$  to express this as a function of  $\tilde{u}$ , thus showing that the ratio  $\frac{\tilde{A}_1}{\tilde{A}_{11}}$  is a function only of  $\tilde{u}$ , as required. Similarly,  $\frac{\tilde{A}_2}{\tilde{A}_{22}}$  is a function only of  $v$ . We must also show that there are five linearly independent solutions. If  $\Psi(u, v) = U(u)V(v)$  is a solution in  $\{u, v\}$  coordinates, then a straightforward calculation reveals that  $\tilde{\Psi}(\tilde{u}, \tilde{v}) = \Psi(\eta^{-1}(\tilde{u}), \phi^{-1}(\tilde{v})) = U(\eta^{-1}(\tilde{u}))V(\phi^{-1}(\tilde{v}))$  is a solution in  $\{\tilde{u}, \tilde{v}\}$  coordinates, where  $\eta^{-1}$  and  $\phi^{-1}$  are the inverse functions of  $\eta$  and  $\phi$ , respectively. And since each  $\tilde{\Psi}$  takes on the same value at a given  $(x, y)$  point as the corresponding  $\Psi$ , and thus the functions  $\tilde{\Psi}$  are essentially the same as the functions  $\Psi$ , it is clear that the transformations of the five linearly independent solutions form a linearly independent set of solutions.

Since  $\mathcal{H}$  is invariant under the one-parameter subgroups of its first-order symmetry operators (proof omitted — see, for instance [Spiegel 1999, pp. 41-43]),  $\mathcal{H}$  is invariant under rotations and translations. We first observe that dilations serve only to multiply  $\omega^2$  by the square of the dilation factor. For reflections, we may consider only reflections over the  $y$ -axis, for all other reflections are the composition of reflections over the  $y$ -axis and translations and rotations. If we reflect our space over the  $y$ -axis,  $\partial_x$  turns into  $-\partial_{\tilde{x}}$ , but then taking a second derivative yields  $\partial_{xx} = -(-\partial_{\tilde{x}\tilde{x}}) = \partial_{\tilde{x}\tilde{x}}$ . That the  $\omega^2$  term will remain unchanged follows from the fact that the  $\partial_{xx}$  and  $\partial_{yy}$  terms maintain the same form (only  $\partial_{xx}$  is re-labeled as  $\partial_{\tilde{x}\tilde{x}}$ ).  $\diamond$



We now return to the Helmholtz equation's separable coordinate systems. If it is not already the case that  $\mathcal{U} \equiv 1$  and  $\mathcal{V} \equiv 1$  in (20), then there exist real-analytic functions  $\tilde{u}(u)$  and  $\tilde{v}(v)$  such that

$$\tilde{u}'(u) = \mathcal{U}^{-\frac{1}{2}} \quad \text{and} \quad \tilde{v}'(v) = \mathcal{V}^{-\frac{1}{2}}.$$

Since  $\mathcal{U}$  and  $\mathcal{V}$  are both always positive (from Lemma 1),  $\tilde{u}'(u)$  and  $\tilde{v}'(v)$  are everywhere defined. It is clear that  $\tilde{u}'(u)$  and  $\tilde{v}'(v)$  are never 0, for there are no real values of  $\mathcal{U}$  and  $\mathcal{V}$  that would lead to this. As a result, these functions describe a coordinate system that is shapewise equivalent to  $\{u, v\}$ . We now wish to come up with expressions for  $\tilde{u}_x^2 + \tilde{u}_y^2$  and for  $\tilde{v}_x^2 + \tilde{v}_y^2$ . Note that

$$\frac{\partial \tilde{u}}{\partial x} = \frac{d\tilde{u}}{du} \frac{\partial u}{\partial x} = \mathcal{U}^{-\frac{1}{2}} u_x$$

and so

$$\tilde{u}_x^2 = \frac{1}{\mathcal{U}} u_x^2.$$

Expressions for  $\tilde{u}_y^2, \tilde{v}_x^2$  and  $\tilde{v}_y^2$  are analogous, and so, (20) now implies that

$$\tilde{u}_x^2 + \tilde{u}_y^2 = \frac{1}{\mathcal{U}} (u_x^2 + u_y^2) = \frac{1}{\mathcal{U}_1 + \mathcal{V}_1} \quad \text{and} \quad \tilde{v}_x^2 + \tilde{v}_y^2 = \frac{1}{\mathcal{V}} (v_x^2 + v_y^2) = \frac{1}{\mathcal{U}_1 + \mathcal{V}_1}. \quad (25)$$

At this point it is important to make sure that we understand what we have just derived. Our discovery is that if  $\{u, v\}$  is a coordinate system in which (5) is strongly separable, and thus in which (20) holds, then there is a shapewise equivalent coordinate system  $\{\tilde{u}, \tilde{v}\}$  in which (5) is strongly separable such that (25) holds. At this point, it is a straightforward (but tedious) matter to show that every coordinate system (choices of  $\{u, v\}$ ) in which (5) is strongly separable is shapewise equivalent to rectangular coordinates, to parabolic coordinates, to polar coordinates, or to elliptic coordinates. This analysis is done in [Miller 1977] and [Morse and Feshbach 1953], albeit without the language of shapewise equivalence, and in [Spiegel 1999].

## Acknowledgements

I thank Professor David A. Cox of Amherst College, whose efforts and support in this research were outstanding.

## References

- [Hildebrand 1976] Hildebrand, F. *Advanced Calculus for Applications: SECOND EDITION*. Prentice-Hall, Inc., Englewood Cliffs, NJ, 1976
- [Miller 1977] Miller, W. *Symmetry and Separation of Variables*. Addison-Wesley Publishing Company, Reading, MA, 1977
- [Morse and Feshbach 1953] Morse, P. M. and Feshbach, H. *Methods of Theoretical Physics*. McGraw-Hill Book Company, New York, NY, 1953
- [Robertson 1928] Robertson, H. P. "Bemerkung uber separierbare Systeme in der Wellenmechanik". *Math. Ann.* 98(1928), pp. 749-752
- [Spiegel 1999] Spiegel, D. S. *Symmetries of the Helmholtz Equation And Its Separable Coordinate Systems*. Senior Thesis, Amherst College, 1999