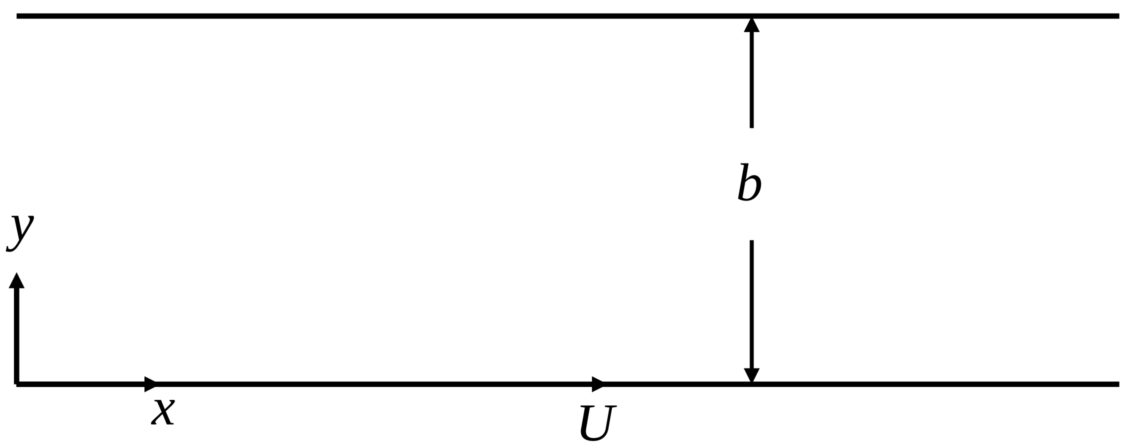
Solution of Partial Differential Equations
Separation of Variables

R. Shankar Subramanian
Department of Chemical and Biomolecular Engineering
Clarkson University

Introduction and Problem Statement

We encounter partial differential equations routinely in transport phenomena. Some examples are unsteady flow in a channel, steady heat transfer to a fluid flowing through a pipe, and mass transport to a falling liquid film. Here, we shall learn a powerful method for solving many of these partial differential equations. We shall also learn when the method can be used. The model problem we consider is the motion induced in fluid contained between two long and wide parallel plates placed with a distance b between them as shown in the sketch below.



The fluid is initially assumed to be at rest. Motion is initiated by suddenly moving the bottom plate at a constant velocity of magnitude U in the $x-$direction. The velocity of the bottom plate is maintained at that value for all future values of time t while the top plate is held fixed in place. There is no applied pressure gradient, with motion being caused strictly by the movement of the bottom plate.

We shall assume the flow to be incompressible with a constant density ρ and Newtonian with a constant viscosity μ. We neglect edge effects in the $z-$direction so that we can set $v_z = 0$ and $\dfrac{\partial \boldsymbol{v}}{\partial z} = 0$, and assume fully developed flow, implying $\dfrac{\partial \boldsymbol{v}}{\partial x} = 0$. Here, $\boldsymbol{v}$ stands for the velocity vector, and the subscripts denote components.

It can be established from the continuity equation and the kinematic condition at one of the walls that $v_y = 0$. Therefore, the only non-zero velocity component is $v_x(t, y)$, which can be shown to satisfy the following partial differential equation.

$$\frac{\partial v_x}{\partial t} = \nu \frac{\partial^2 v_x}{\partial y^2} \qquad (1)$$

In Equation (1), t represents time, and ν is the kinematic viscosity. The initial condition is

$$v_x(0, y) = 0 \qquad (2)$$

and the boundary conditions are

$$v_x(t, 0) = U \qquad (3)$$

and

$$v_x(t, b) = 0 \qquad (4)$$

It is convenient to work with scaled variables. Scaling minimizes the number of parameters in the problem, and helps us identify the true (dimensionless) parameters so that we can perform asymptotic analyses where desired. Whenever possible, we use a reference quantity, termed a "scale" for the variable involved, that will normalize that variable, meaning that the range of values assumed by the dimensionless variable will be from 0 to 1. The natural variables that normalize the velocity and the $y-$coordinate in this problem are U and b, respectively. Therefore, we define a scaled velocity $V = \frac{v_x}{U}$ and a scaled distance variable $Y = \frac{y}{b}$. Introducing these definitions into the differential equation, we obtain

$$\frac{\partial V}{\partial t} = \frac{\nu}{b^2} \frac{\partial^2 V}{\partial Y^2} \qquad (5)$$

which suggests that we might choose the time scale as b^2 / ν, defining a scaled time $T = \frac{\nu t}{b^2}$. Thus, Equation (5) can be rewritten as

$$\frac{\partial V}{\partial T} = \frac{\partial^2 V}{\partial Y^2} \qquad (6)$$

and the initial and boundary conditions become

$$V(0,Y)=0 \tag{7}$$

$$V(T,0)=1 \tag{8}$$

$$V(T,1)=0 \tag{9}$$

Now, we are ready to learn the mathematical technique of "Separation of Variables." The usual way to solve a partial differential equation is to find a technique to convert it to a system of ordinary differential equations. Then, we can use methods available for solving ordinary differential equations. One important requirement for separation of variables to work is that the governing partial differential equation and initial and boundary conditions be linear. Another is that for the class of partial differential equation represented by Equation (6), the boundary conditions in the $Y-$coordinate be homogeneous. This means that any constant times the dependent variable should satisfy the same boundary condition. Also, the differential equation itself should be homogeneous. A condition in which the variable or a linear combination of the variable and its spatial or time derivative is set equal to 0 can be seen to be a homogeneous condition.

We see that Equation (6) is homogeneous because a constant times V will satisfy the same equation. Equation (9) is homogeneous as well, but Equation (8) is not. Therefore, we must first define a new problem in which homogeneous boundary conditions can be written. The approach we follow is based on the physical aspects of the problem. Consider the same fluid mechanical problem at steady state, wherein we set the time derivative of the velocity equal to zero. This means that we can no longer expect to satisfy the initial condition, but the boundary conditions still hold. The resulting steady velocity field $V_s(Y)$ can be seen from Equations (6), (8), and (9) to satisfy

$$\frac{d^2V_s}{dY^2}=0 \tag{10}$$

$$V_s(0)=1 \tag{11}$$

$$V_s(1)=0 \tag{12}$$

The solution is seen to be

$$V_s(Y)=1-Y \tag{13}$$

Now, write the solution of Equations (6) - (9) as the sum of the above steady solution and a transient contribution that we expect will decay to zero as $T\to\infty$.

$$V(T,Y)=V_s(Y)+V_t(T,Y) \tag{14}$$

Equation (14) is substituted into Equations (6) - (9), and use is made of Equations (10) - (12). This yields the governing equation and the initial and boundary conditions for the transient field $V_t\left(T,Y\right)$.

$$\frac{\partial V_t}{\partial T} = \frac{\partial^2 V_t}{\partial Y^2} \tag{15}$$

$$V_t\left(0,Y\right) = -V_s\left(Y\right) \tag{16}$$

$$V_t\left(T,0\right) = 0 \tag{17}$$

$$V_t\left(T,1\right) = 0 \tag{18}$$

It is seen that the inhomogeneity in the boundary condition for $V\left(T,0\right)$ has been taken up by $V_s\left(0\right)$, leaving us with a homogeneous boundary condition for $V_t\left(T,0\right)$. If the governing differential equation had a time-independent inhomogeneity, we can expect the same will happen. That inhomogeneity will be included in the governing equation for the steady field, leaving the governing equation for the transient field homogeneous.

Product Class Solution

Now, we attempt a solution of Equation (15) in the form of a product

$$V_t\left(T,Y\right) = G\left(T\right)\phi\left(Y\right) \tag{19}$$

This is not to suggest that the final solution will be exactly like this. It is a trial solution, just like the trial solution e^{mx} that is used in the case of a linear ordinary differential equation with constant coefficients. The approach will be to substitute this trial solution in the governing equation and the initial and boundary conditions to see if it might possibly satisfy them. First inserting it into Equation (15) yields

$$G'\phi = G\phi'' \tag{20}$$

where we have used primes to denote differentiation with respect to the argument of the function. Thus, G' stands for dG/dT whereas ϕ'' connotes $d^2\phi/dY^2$. Divide both sides of Equation (20) by $G\phi$. This yields

$$\frac{G'}{G} = \frac{\phi''}{\phi} \tag{21}$$

But, the left side of the above equation can depend only on T, whereas the right side can depend only on Y. How can it be possible for a function of only T to be equal to a function of only Y? The answer is: Never, unless we force both functions to be a constant that is independent of T and Y. For reasons that will become clear later, we require this "constant of separation" to be negative. So, we set it equal to $-\lambda^2$ where λ is a real number.

$$\frac{G'}{G} = \frac{\phi''}{\phi} = -\lambda^2 \tag{22}$$

So we see that we have made a lot of progress. We now have two ordinary differential equations in place of the partial differential equation. They are

$$G' + \lambda^2 G = 0 \tag{23}$$

and

$$\phi'' + \lambda^2 \phi = 0 \tag{24}$$

The solution of Equation (23) can be written as

$$G(T) = \alpha \exp\left[-\lambda^2 T\right] \tag{25}$$

where α is a constant of integration. Notice that Equation (25) implies that as $T \to \infty$, $G \to 0$, which is consistent with our idea that the transient solution will decay as $T \to \infty$. The general solution of Equation (24) for $\phi(Y)$ can be written as

$$\phi(Y) = c_1 \sin \lambda Y + c_2 \cos \lambda Y \tag{26}$$

where c_1 and c_2 are constants of integration. Because the boundary conditions on V_t at $Y = 0$ and $Y = 1$ that are given in Equations (17) and (18), respectively, are both homogeneous, they can be satisfied by setting

$$\phi(0) = 0 \tag{27}$$

$$\phi(1) = 0 \tag{28}$$

Application of these boundary conditions yields the following results.

$$c_2 = 0 \tag{29}$$

$$c_1 \sin \lambda = 0 \tag{30}$$

If we try to satisfy Equation (30) with the choice $c_1 = 0$, we obtain the result that $\phi(Y) \equiv 0$. This yields the trivial solution $V_t = 0$. This is incorrect because it does not satisfy the initial condition on V_t given in Equation (16). Therefore, we must choose the alternative

$$\sin \lambda = 0 \tag{31}$$

This equation has an infinite number of roots that occur in pairs.

$$\lambda = \lambda_n = \pm n\pi\,, \qquad n = 0,1,2,\ldots \tag{32}$$

First, we note that the case $n=0$ can be discarded because it again leads to the trivial solution that is unacceptable. Second, the negative roots do not yield an independent solution because $\sin(-n\pi Y) = -\sin(n\pi Y)$. Therefore, we can write the solution for $\phi(Y)$ as

$$\phi(Y) = \phi_n(Y) = c_n \sin \lambda_n Y \tag{33}$$

with

$$\lambda_n = n\pi, \qquad n = 1, 2, 3, \ldots \tag{34}$$

Note that we have replaced the single constant c_1 with a subscripted constant c_n to underscore the fact that each of these acceptable solutions can be multiplied by a different arbitrary constant.

The net result of the exercise has been to produce an infinite set of product class solutions for V_t. By representing the product of the arbitrary constants α and c_n using a new constant A_n, we can write the n'th solution as $A_n e^{-\lambda_n^2 T} \sin \lambda_n Y$. At $T=0$, this becomes $A_n \sin \lambda_n Y$. This is a periodic function and does not at all look like the function $-V_s(Y)$, which happens to be a straight line in the interval $[0,1]$. Fortunately, because the governing equation and boundary conditions are linear and homogeneous, we can add all of these solutions and try to see if the sum can be used to satisfy the initial condition by judicious choice of the constants A_n. Therefore, we write

$$V_t(T,Y) = \sum_{n=1}^{\infty} A_n \; e^{-\lambda_n^2 T} \sin \lambda_n Y \tag{35}$$

Notice that there is no problem when we add a finite number of solutions, but when the upper limit of summation is infinity, we need to concerned with the issue of whether the right side converges. Such mathematical issues are considered in detail in Weinberger [1]. Here, we assume that the sum uniformly converges for all values of scaled time T and all values of Y in the interval $[0,1]$. By applying the initial condition given in Equation (16), we obtain

$$-V_s(Y) = \sum_{n=1}^{\infty} A_n \sin \lambda_n Y \tag{36}$$

Equation (36) represents the expansion of an arbitrary function ($-V_s(Y)$) in a Fourier series, named after the scientist Fourier who studied such expansions a long time ago. Fourier series do not necessarily have to be expansions in trigonometric functions, and you can learn more about them from Weinberger [1]. The most important aspect of such an expansion is that the set of functions $\{\sin \lambda_n Y\}$ is orthogonal in the interval $[0,1]$. That is

$$\int_0^1 \sin\lambda_m Y \sin\lambda_n Y \, dY = 0, \qquad m \neq n \tag{37}$$

Of course, when $m = n$, the integral is not zero, but is given by

$$\int_0^1 \sin^2\lambda_n Y \, dY = \frac{1}{2} \tag{38}$$

Therefore, we can use the following recipe for calculating the expansion coefficients A_n. Multiply both sides of Equation (36) by $\sin\lambda_m Y$ where m is a specific integer, and integrate from $Y = 0$ to 1. Every term in the infinite series will reduce to zero because of Equation (37), with the exception of the term that involves an integral that is of the form of Equation (38) with the index n replaced by m. As a result, we obtain

$$\int_0^1 \left(-V_s\left(Y\right)\right)\sin\lambda_m Y \, dY = A_m \int_0^1 \sin^2\lambda_m Y \, dY = \frac{A_m}{2} \tag{39}$$

so that we can write

$$A_n = 2\int_0^1 \left(-V_s\left(Y\right)\right)\sin\lambda_n Y \, dY \tag{40}$$

where we have replaced the index m, which is just a placeholder, with the index n. When the result for $V_s\left(Y\right)$ given in Equation (13) is used and the integration is performed, we ultimately find

$$A_n = -\frac{2}{n\pi} \tag{41}$$

The final result for $V\left(T,Y\right)$ can be written as follows.

$$V\left(T,Y\right) = 1 - Y - \frac{2}{\pi}\sum_{n=1}^{\infty} \frac{\exp\left[-n^2\pi^2 T\right]}{n} \sin n\pi Y \tag{42}$$

We can infer how long it will take to achieve steady state. Of course, the correct answer is infinite time because the exponential functions in the infinite series are never quite zero approaching that value only when $T \to \infty$. But as a practical matter, we can see that when $T = 1$, which corresponds to physical time $t = b^2 / \nu$, the contribution from the infinite series will be negligible, except in a situation where we wish to be extremely precise. We refer to this "time scale" b^2 / ν as the time it takes for momentum to diffuse a distance b. Analogous time scales can be defined for diffusion of thermal energy or diffusion of species by replacing the kinematic viscosity by the thermal diffusivity or the mass diffusivity, respectively.

Summary

Here is a brief summary of the method of “Separation of Variables.” It may be used to find solutions of linear partial differential equations. After identifying the governing partial differential equation and the initial and boundary conditions for our physical system, we

1. scaled the problem by using suitable reference quantities;

2. found a solution of the steady-state problem;

3. expressed the solution of the original problem as the sum of the steady-state solution and a transient contribution, in that process formulating a partial differential equation and the initial and boundary conditions for the transient contribution;

4. found a solution of the transient problem by assuming a product form for that solution;

5. invoked the principle of superposition to express the general solution of the transient problem as an infinite sum;

6. used the orthogonality of the basis functions (sines in our problem) to obtain the coefficients that appear in the general transient solution;

7. wrote the complete solution as the sum of the steady and transient solutions.

The method of “Separation of Variables” also can be used to find the solution of other linear problems such as steady-state multi-dimensional conduction or diffusion problems. In such a case, we would not have an initial condition, but there would be more boundary conditions.

Concluding Remarks

If the process for finding the expansion coefficients A_n reminds you of the process we use for expanding spatial vectors in an orthogonal basis set, the resemblance is not superficial. The idea of geometrical orthogonality, which we can visualize in three-dimensional space, is extended to an infinite-dimensional “function space” in developing a basis set for expanding functions. The dot product that we use with spatial vectors is generalized to the “inner product” which is defined as the integral over the interval that we used, for example, in Equation (37). Just as the eigenvectors of a real symmetric tensor can be used to generate an orthogonal set of basis vectors, a certain type of differential operator, called a self-adjoint operator, is used to generate a basis set of “orthogonal eigenfunctions” in the context of expanding arbitrary functions. You can learn more about such ideas from Greenberg [2].

It is worthy of note that the problem of unsteady heat conduction in a solid slab (or a quiescent liquid layer) of thickness b when the temperature at the surface $y = b$ is maintained at the same value that it is initially everywhere in the slab, while the temperature at the surface $y = 0$ is changed to a new value, is described by the same governing equations and boundary conditions in scaled form. The assumptions are that there are no sources or sinks, heat transport occurs only by conduction with a constant thermal conductivity, the density and specific heat of the material are constant, and that the slab is very long and very wide so that end effects and edge effects can be neglected. By analogy, it can be seen that the same equations also describe unsteady diffusion in a similar situation. Other boundary conditions are possible in these problems. For example, one can prescribe the heat or mass flux at a boundary instead of prescribing the temperature, or write the flux at a boundary as being proportional to the temperature difference between the surface and a constant ambient temperature. All of these cases can continue to be handled by the same solution method, which gives you some idea about the versatility of the mathematical technique in the case of this type of partial differential equation and boundary conditions. As noted in the summary, the method also can be used with other types of linear partial differential equations such as the Laplace equation or the convective diffusion equation that arise in heat or mass transport.

Acknowledgment

I am indebted to Professor John B. McLaughlin of Clarkson University who read a draft of this module and provided suggestions for improvement.

References

1. H.F. Weinberger, *A First Course in Partial Differential Equations*, Xerox College Publishing, Lexington, 1965.

2. M.D. Greenberg, *Foundations of Applied Mathematics*, Prentice-Hall, Englewood Cliffs, 1978.