Spring 1979

Mathematical And Physical Aspects Of Heat Conduction Between A Rod And A Sleeve Of Similar Material And Infinite Length

Thomas Dillon Jr.
Carroll College

Follow this and additional works at: https://scholars.carroll.edu/mathengcompsci_theses

Part of the Applied Mathematics Commons, and the Physics Commons

Recommended Citation
https://scholars.carroll.edu/mathengcompsci_theses/111

This Thesis is brought to you for free and open access by the Mathematics, Engineering and Computer Science at Carroll Scholars. It has been accepted for inclusion in Mathematics, Engineering and Computer Science Undergraduate Theses by an authorized administrator of Carroll Scholars. For more information, please contact tkratz@carroll.edu.
MATHEMATICAL AND PHYSICAL ASPECTS OF HEAT CONDUCTION BETWEEN A ROD AND A SLEEVE OF SIMILAR MATERIAL AND INFINITE LENGTH

By

Thomas J. Dillon Jr.
This thesis for honors recognition has been approved for the Department of Mathematics.

Dr. Ronald N. Knoshaug

Mr. Alfred J. Murray

Dr. John E. Semmens

March 30, 1979
Date
Acknowledgements

I would like to thank my parents who have worked so hard to put me through college. Their unending love and support has helped me tremendously in completing this thesis and the B.A. degree.

I am grateful to Mr. Alfred J. Murray and Dr. John E. Semmens, who have read this thesis and made helpful criticisms. And most of all, I sincerely appreciate Dr. Ronald N. Knoshaug, my thesis director, for his unlimited help and encouragement throughout the writing of this thesis.

There are many others I would like to thank who have helped me complete this work. To all my friends, especially Frank Clinch and Glenn Garrison, Thank You!
Introduction

The modern world has experienced an increase in the applications of mathematics to physical activity. Many physical conditions can be described in terms of partial differential equations, which result in industrial applications. One such application deals with the problem of imbedding metal bolts into a metal frame, achieving a secure fit. The problem is in determining the factors which affect the security of the seal.

There are two ways by which the bolt is secured. The first involves heating the frame and inserting the bolt snugly into it. When the frame cools it contracts tightly around the bolt. The second involves cooling the bolt and fitting it snugly into the frame. Then a tight seal forms when the bolt warms up to the temperature of the frame. The problem that the engineer faces is in determining how hot or how cold the bolt should be in relation to the temperature of the frame. This problem is generally referred to as the problem of shrunken-fittings. Both methods can be described mathematically with the use of the heat equation.

The intent of this thesis is to indicate how this problem can be solved mathematically. Specifically it is to create a mathematical model of the process of heat conduction in the cylindrical coordinate system. Included in the study will be several basic requirements for the mathematical model. A description of the physical aspect of...
heat conduction will be introduced to derive the heat equation. The heat equation, in the Cartesian coordinate system, will be translated to the cylindrical coordinate system. A look at the method of separation-of-variables will show how it applies to the heat equation. A compact review of Bessel function theory and its use in the study of the Sturm-Liouville problem will be included in this work.

This body of knowledge will be used to solve a particular problem in heat conduction. The problem will involve a rod of infinite length, made up of isotropic material, and initially at a constant temperature A. Slipped over the rod will be a sleeve of the same material and length, but initially with a different constant temperature B. Also, at the outside boundary of the sleeve the temperature will always be B. What I intend to do is describe the heat distribution of this system for anytime forward of time zero. To conclude the thesis I will discuss briefly the limitations of this mathematical model.
Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>v</td>
</tr>
<tr>
<td>1) Physical Theory of Heat Conduction</td>
<td>1</td>
</tr>
<tr>
<td>a) Physical Description of Heat Conduction</td>
<td></td>
</tr>
<tr>
<td>b) Derivation of the Heat Equation</td>
<td></td>
</tr>
<tr>
<td>c) Transformation of the Heat Equation into the Cylindrical Coordinate System</td>
<td></td>
</tr>
<tr>
<td>2) Method of Separation of Variables</td>
<td>9</td>
</tr>
<tr>
<td>3) Bessel Function Theory</td>
<td>12</td>
</tr>
<tr>
<td>a) Derivation of the Bessel Function</td>
<td></td>
</tr>
<tr>
<td>b) Derivative and Integral Formula of the Bessel Function</td>
<td></td>
</tr>
<tr>
<td>c) The Sturm-Liouville Problem</td>
<td></td>
</tr>
<tr>
<td>d) Orthogonal Sets of Functions</td>
<td></td>
</tr>
<tr>
<td>e) Fourier Series</td>
<td></td>
</tr>
<tr>
<td>f) Orthonormal Functions and Fourier-Bessel Series</td>
<td></td>
</tr>
<tr>
<td>4) Solution of the Selected Problem</td>
<td>28</td>
</tr>
<tr>
<td>a) Definition of the Problem</td>
<td></td>
</tr>
<tr>
<td>b) General Solution</td>
<td></td>
</tr>
<tr>
<td>5) Extensions of the Model</td>
<td>34</td>
</tr>
<tr>
<td>References</td>
<td>37</td>
</tr>
</tbody>
</table>
1) Physical Theory of Heat Conduction
   a) Physical Description of Heat Conduction

The terms "hot" and "cold" are frequently used to describe an object. They allude to an object's temperature in relation to the temperature of other objects. When one object, which differs from another object in hotness or coldness, is placed in contact with that other object a change takes place. If we refer to one of the objects as "hot" and the other as "cold," the change involved is the transfer of heat from the hot object to the cold object. This process is called heat conduction. It should be noted that we can also talk about heat conduction in a single object when its environment changes in temperature. That is, when we hold a metal rod at one end and heat the other end, the entire rod will warm up due to the conduction of heat from the hot end to the cold end.

The physical reason for this process of heat conduction is as follows. When something is hot, like the end of a rod, the molecules that make up that something are very active. Their action brings them in contact with other slower moving molecules. When a collision of molecules takes place, the slow moving particle picks up speed and therefore warms up. This process continues through the entire object until an equilibrium is reached. Equilibrium can mean that the same temperature is throughout the system containing the object(s) or that there is some distribution of temperatures in the system.
Through experimentation it has been found that if the temperature difference between two objects or between an object and its environment is not large, the rate of cooling or warming that occurs is proportional to the temperature difference between the two objects or the object and its environment. That is to say: \( \frac{d\Delta T}{dt} \propto \Delta T \) or \( \frac{d\Delta T}{dt} = -k\Delta T \). The letter \( k \) is the constant of proportionality. The minus sign describes the cooling effect on an object that has a higher temperature than its environment or the object in contact with it. This, of course, also applies in reverse order for the warming effect. The name for the equation above is Newton's Law of Cooling.

Heat in a system is the result of the temperature changes in the system. The term "heat" is applied to the energy dispersed by something when it cools off. The heat given off is dissipated into the surroundings and therefore the surroundings warm up. To understand this process, we can imagine a slab of material that has a particular area and thickness. During the process of heat transfer there is a certain amount of heat that passes through that slab from the hot face to the cold face. Experiments have shown that heat flow can be measured perpendicular to the slab face and analyzed if we use small temperature differences and small slab thicknesses. The experiments tell us that for a given area and length of time, the flow of heat is proportional to the change in temperature over a certain thickness of material. In terms of an equation, this means that \( \frac{\Delta Q}{\Delta t} \propto \frac{\Delta T}{\Delta x} \), where \( \Delta Q \) is the heat. When we take
the limit of the equation using a slab of thickness \(dx\) and a temperature difference \(dT\), we obtain \(\frac{dQ}{dt} = -kA\frac{dT}{dx}\), where \(k\) is the proportionality constant. This equation is the fundamental law of heat conduction, in which \(\frac{dT}{dx}\) is called the temperature gradient and \(k\) is called the thermal conductivity of the material.\(^2\) We considered a small temperature difference because if it is small, \(k\) is indeed a constant. But in cases where work is done over a large range of temperatures, \(k\) becomes a function of several variables, including temperature.

This section has covered a short description of the physical part of heat conduction. Our next step is to derive the heat equation for a solid so that we may look at heat conduction more closely.

b) Derivation of the Heat Equation\(^3\)

Consider a solid of any shape, with a certain temperature distribution at time \(t\), made up of a material that is isotropic: that is, the structure and properties at any point in the solid are the same in any direction. It is also necessary that we assume that the properties involved are continuous throughout the solid. This requirement makes sure that we do not have anomalies that would upset our analysis.
Let us consider a point \( P \) in the solid with a volume \( V \) and a temperature distribution \( U(x,y,z,t) \), where \( t \) denotes time. There is a small face of area, \( ds \), centered about the point \( P \). It is through this face that heat flows in a direction perpendicular to \( ds \), namely in the direction of \( \hat{n} \), the outward normal. We will call this flow of heat \( f_n \), the flux of heat, which can be derived according to the formula \( f_n = -\kappa \frac{\partial U}{\partial n} \). Of course, \( \frac{\partial U}{\partial n} \) denotes the change of heat distribution in the direction of the outward normal, while \( K(x,y,z,u) \), with \( u \) signifying heat, is the thermal conductivity of the solid. To standardize this notion, we can rely on the Cartesian coordinate system to give us

\[
\begin{align*}
  f_x &= -\kappa \frac{\partial U}{\partial x}, \\
  f_y &= -\kappa \frac{\partial U}{\partial y}, \\
  f_z &= -\kappa \frac{\partial U}{\partial z}.
\end{align*}
\]

In order to go any further in our analysis, we will assume that we have a fixed volume \( V \). Next we know that there are three things which will contribute to the heat distribution in the solid. First we know there is a certain amount of heat stored in the solid. If we let \( p(x,y,z,t,u) \) be the density of the material in the solid and \( c(x,y,z,t,u) \) be the heat per unit mass of the solid, or the specific heat, we can integrate over the volume and get the heat contained in the solid. That is

\[
\int_V p c u \, dx \, dy \, dz = (\text{heat stored in } V).
\]

Secondly, there is the amount of heat which flows into the solid. We can integrate the flux of heat over the surface of the solid as follows:

\[
- \int_S f_n \, ds = (\text{heat flow into } V),
\]

where the minus sign signifies that the flow is into the solid.
Thirdly, it is possible that heat is generated in the solid. If we let $A(x,y,z,t,u)$ be the rate of heat production at any point, integration over the entire volume will give us the total heat which is produced within the solid. Therefore we have $\iiint_V A \, dx \, dy \, dz =$ (heat produced in $V$).

According to the laws of energy conservation, we know that the change of heat stored in the solid must equal the amount that flows into the solid plus the amount that is produced in the solid. In other words,

$$\frac{\partial}{\partial t} \left( \iiint_V \rho c u \, dx \, dy \, dz \right) = -\iiint_S f_n \, ds + \iiint_V A \, dx \, dy \, dz.$$

This form is hard to work with, so we will try to make some observations that will simplify the equation. First of all, we stated above that we were using a fixed volume. Therefore the heat stored in the solid will only vary with time and so we can say that

$$\frac{\partial}{\partial t} \left( \iiint_V \rho c u \, dx \, dy \, dz \right) = \iiint_V \frac{\partial}{\partial t} (\rho c u) \, dx \, dy \, dz.$$

In the Cartesian coordinate system, we can define the flux of heat at any point in terms of the unit vectors $\hat{i}, \hat{j}, \hat{k}$, such that $\mathbf{f} = f_x \hat{i} + f_y \hat{j} + f_z \hat{k}$. To obtain the flux of heat in the direction of the normal $\hat{n}$, all we must do is take the dot product of the flux with the unit normal $f_n = \mathbf{f} \cdot \hat{n} \, ds$. Therefore our integral for heat flow into the solid becomes $\iiint_S f_n \, ds = \iiint_S \mathbf{f} \cdot \hat{n} \, ds$. Using Gauss' integral theorem, we can change the surface integral above to a volume integral.

Gauss' Integral Theorem: Let $\mathcal{S}$ be a piecewise-smooth bounding surface of a bounded, closed volume $V$ in $D$ the domain.
Then \( \int\int\int V_1 dS = \int\int\int d\text{inv}(\bar{v}) \, dV \) where
\[
V_1 = \bar{v} \cdot \hat{n} \quad \text{and} \quad d\text{inv}(\bar{v}) = \nabla \cdot \bar{v} = \frac{\partial V}{\partial x} + \frac{\partial V}{\partial y} + \frac{\partial V}{\partial z}.
\]

Since our solid follows these conditions, then
\[
\int\int\int F_n dS = \int\int\int d\text{inv}(\bar{f}) \, dV = \int\int\int \left( \frac{\delta F_1}{\delta x} + \frac{\delta F_1}{\delta y} + \frac{\delta F_1}{\delta z} \right) dx \, dy \, dz.
\]

Next we can show that \( \frac{\partial}{\partial t} \left( \int\int\int pcu \, dx \, dy \, dz \right) = -\int\int\int A \, dx \, dy \, dz \), leads to \( \int\int\int A \, dx \, dy \, dz \), so we have \( \int\int\int \left( \frac{\partial}{\partial t} (pcu) + \nabla \cdot \bar{f} - A \right) dx \, dy \, dz = 0. \)

This is true for any volume because we used an arbitrary volume at the start of our analysis. If the integrand was positive, or negative for that matter, the integration would yield a volume of some magnitude other than zero. This contradicts our equation. Therefore the integrand must equal zero. That is to say, \( \frac{\partial}{\partial t} (pcu) + \nabla \cdot \bar{f} - A = 0 \)

or \( \frac{\partial}{\partial t} (pcu) + \frac{\partial F_1}{\partial x} + \frac{\partial F_1}{\partial y} + \frac{\partial F_1}{\partial z} = A. \) Using the definition of \( f_n = -\kappa \frac{\partial u}{\partial n} \) and adding the \( \nabla \cdot \bar{f} \) to the right side, we get the following

\[
\frac{\partial}{\partial t} (pcu) - \frac{\partial}{\partial x} (k \frac{\partial u}{\partial x}) + \frac{\partial}{\partial y} (k \frac{\partial u}{\partial y}) + \frac{\partial}{\partial z} (k \frac{\partial u}{\partial z}) + A.
\]

This is a form of the equation that we can work with and it applies to any volume. There are some changes we can make in the equation that will simplify it further. First of all, for many applications of the equation, \( p \) and \( c \) are not functions of \( t \). Therefore \( \frac{\partial}{\partial t} (pcu) = pc \frac{\partial u}{\partial t} \)

Also \( K \) and \( A \) are generally functions of position only. So we can say that

\[
\frac{\partial}{\partial x} (k \frac{\partial u}{\partial x}), \frac{\partial}{\partial y} (k \frac{\partial u}{\partial y}), \frac{\partial}{\partial z} (k \frac{\partial u}{\partial z}) = K \nabla^2 u + \frac{\partial K}{\partial x} \frac{\partial u}{\partial x} + \frac{\partial K}{\partial y} \frac{\partial u}{\partial y} + \frac{\partial K}{\partial z} \frac{\partial u}{\partial z}
\]

by using the definition of the derivative, where
\[ \nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}. \]

Then because of the chain rule of differentiation we know that \( \frac{\partial K}{\partial n} = \frac{\partial K}{\partial u} \frac{\partial u}{\partial n} \). Joining these facts with the equation we derived above, we obtain the form

\[
\rho c \frac{\partial u}{\partial t} = K \nabla^2 u + A + \frac{\partial K}{\partial u} \left( \frac{\partial u}{\partial x} \right)^2 + \frac{\partial u}{\partial y} \left( \frac{\partial u}{\partial y} \right)^2 + \left( \frac{\partial u}{\partial z} \right)^2 \right). 
\]

The term \( K \) was defined to be the thermal conductivity and when there are small temperature gradients, such as in our case, it is a constant. Another change is found in \( A \). Since the scope of our problem does not include internal production of heat we can eliminate \( A \). These changes yield the following equation:

\[
\frac{\partial u}{\partial t} = k \nabla^2 u = k \left( \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} \right), \]

where \( k = K/(pc) \) which is called the thermal diffusivity. This equation describes the distribution of temperature throughout a solid in terms of the Cartesian coordinate system.

To complete the model of the physical heat equation, we must know what the heat distribution was at the initial time \( t=0 \). This initial condition gives us a starting place and it is denoted as follows \( u(t=0) = U_0 \). Along with the initial condition we need to know what is happening at the boundaries at any time forward. These conditions, called boundary conditions, can be in the form of heat flow through the boundary or no heat flow through the boundary, which is called insulation. Heat flow at the boundary can be described as a fixed temperature \( T \) for \( t>0 \) or as some function of \( \frac{\partial u}{\partial n} \), like convection. Insulation at a boundary is written as \( \frac{\partial u}{\partial n} = 0 \) for \( t>0 \).
This model that we have derived has been used to describe the physical problem of heat conduction with much success. Therefore we are going to use it with the assurance that it will indeed work.

c) Transformation of the Heat Equation into the Cylindrical Coordinate System

The heat equation derived in the last section is in the Cartesian coordinate system. To help facilitate the solution of our problem, which is cylindrical in nature, we will make a change of variables to the cylindrical system. The variables used in this system are \( r, \theta, \) and \( z, \) such that \( x = r \cos \theta, \) \( y = r \sin \theta, \) and \( z = z, \) which is defined for \( r \) greater than zero.

The equation in the Cartesian coordinate system is

\[
\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} = \frac{1}{k} \frac{\partial u}{\partial t}
\]

If we use the chain rule for derivatives and the fact that the mixed second derivatives \( \frac{\partial^2 u}{\partial r \partial \theta} \) and \( \frac{\partial^2 u}{\partial \theta \partial r} \) are equal, we obtain the general format of the heat equation in cylindrical coordinates. That is to say:

\[
\frac{1}{k} \frac{\partial u}{\partial t} = \frac{1}{r} \frac{\partial}{\partial r} \left( r \frac{\partial u}{\partial r} \right) + \frac{1}{r^2} \frac{\partial^2 u}{\partial \theta^2} + \frac{\partial^2 u}{\partial z^2},
\]

which is the equation we will use in our problem.
2) Method of Separation of Variables

The separation of variables technique has been developed by mathematical physicists to obtain solutions to partial differential equations, one of which is the heat equation. The technique involves the use of an assumption that the solution can be written in terms of the product of functions of a single variable each. For example, with \( U(x,t) \) the technique would assume that it could be written \( U(x,t) = X(x)T(t) \). The reason for the use of this method is that many times it yields several ordinary differential equations, which are easier to solve. In fact, for the heat distribution function in \( n \) variables, the separation-of-variables technique may yield \( n \) ordinary differential equations from the one partial differential equation.

The equation we obtained in the last chapter was as follows:
\[
\frac{1}{k} \frac{\partial U}{\partial t} = \frac{\partial^2 U}{\partial r^2} + \frac{1}{r} \frac{\partial U}{\partial r} + \frac{1}{r^2} \frac{\partial^2 U}{\partial \theta^2} + \frac{\partial^2 U}{\partial z^2}
\]
and it is in terms of four variables, \( r, \theta, z, \) and \( t \). We assume that we can write \( U(r,\theta,z,t) = R(r)\Theta(\theta)Z(z)T(t) \) and then try to separate the partial differential equation of \( U \) into four ordinary differential equations.

Let \( U(r,\theta,z,t) = R(r)\Theta(\theta)Z(z)T(t) \) or \( U = R\Theta Z T \) for short. Then we know that \( \frac{\partial U}{\partial t} = R\Theta Z T' \) where \( T' = \frac{dT}{dt} \) because of properties of partial differentiation. Also \( \frac{\partial U}{\partial \theta} = R\Theta Z \), \( \frac{\partial U}{\partial z} = R\Theta Z \), \( \frac{\partial U}{\partial r} = R\Theta Z \) and \( \frac{\partial U}{\partial \theta} = R\Theta Z \) are forms that we can use. Putting these new forms into the partial differential equation, we arrive at the relation.
\[(1/k)R\theta ZT' = R''\theta ZT + (1/r)R'\theta ZT + (1/r^2)R\theta''ZT + R\theta''ZT.\]

The next step is to divide the equation by \(R\theta ZT\) to get
\[(1/k)(T'/T) = (1/R\theta Z)(R''\theta Z + (1/r)R'\theta Z + (1/r^2)R\theta''Z).\]

We can see that the left side of the equation is in terms of \(T\) only and the right side is in terms of \(R, \theta,\) and \(Z\) only. Therefore if we took the derivative of both sides with respect to \(T\) we would get \(\frac{1}{k} \frac{dT'}{dT} = C\) because the right side of the equation was free of \(T\). Then we can say that \(\frac{T'}{kT}\) equals some constant. This constant, \(c_1\), is called the separation constant such that \(\frac{T'}{kT} = c_1\).

Next we know that from the division of \(R\theta ZT\) and the separation constant \(c_1\) that we arrive at a new form of the equation. That is \(c_1 = R''/R + R'/rR + \theta''/r^2\theta + Z''/Z\). Now by subtraction we can obtain \(Z''/Z = c_1 - (R''/R + R'/rR + \theta''/r^2\theta)\). Since \(Z\) is isolated, we may assign another separation constant, \(c_2\), such that \(Z''/Z = c_2\).

We have finally broken the original equation down to an equation of two variables, namely \(c_2 = c_1 - R''/R - R'/rR - \theta''/r^2\theta\), an equation of just \(R\) and \(\theta\). The next step is to multiply by \((-r^2)\) and then subtract terms to isolate \(\theta\). We obtain \(\theta''/\theta = -r^2R''/R - rR'/rR + r^2(c_2 - c_1)\). Therefore there is one more separation constant, \(c_3\), because the left side is a function of \(\theta\) only and the right side is a function of \(R\) only. This result gives us \(\theta''/\theta = c_3\) and \(r^2(c_2 - c_1) - r^2R''/R - rR'/R = c_3\).

Through the technique of the method of separation of variables, we have broken down the heat equation into four ordinary differential equations. They are as follows
Now that the equations are ordinary differential equations, we can use some of the techniques available to solve these first and second order linear differential equations.

\[ T' - c_1 k T = 0, \quad Z'' - c_2 Z = 0, \quad \theta'' - c_3 \theta = 0 \quad \text{and} \quad r R'' + r R' - R (r^2 (c_1 - c_2) - c_3) = 0. \]
3) Bessel Function Theory

a) Derivation of the Bessel Function

When we used the separation of variables technique in the last chapter, we obtained four ordinary differential equations. One of these equations, namely the equation in terms of \( R \), was of the form

\[ R'' + \frac{1}{r}R' + \left( \lambda^2 - \mu^2/r^2 \right) R = 0. \]

With a few changes we can translate this equation into one more useful to the purpose of this study. Let \( x = \lambda r \), then

\[ R(r) = R(x/\lambda) = y(x). \]

Using the chain rule for derivatives, the equation above is transformed into

\[ \lambda^2 y'' + \left( \lambda^2/x \right) y' + \left( \lambda^2 - \mu^2 \lambda^2/x \right) y = 0. \]

When we divide by \( \lambda^2 \) we get the form

\[ y'' + \left( 1/x \right) y' + \left( 1 - \mu^2/x^2 \right) y = 0, \]

which is called a Bessel differential equation of order \( \mu \).

A typical solution used for the Bessel differential equation is a power series solution of the form

\[ y(x) = \sum_{n=0}^{\infty} a_n x^n. \]

There are certain rules governing this type of solution. A theorem designated for this process reads as follows.

**Theorem:** Let \( y'' + p_1(x)y' + p_0(x)y = 0 \) have coefficients \( p_1(x) \) and \( p_0(x) \) which can be expanded in a power series for \( |x| < r \). Then every solution, \( y \), can be expanded in a power series for \( |x| < r \).

Therefore, if we can expand \( p_0(x) \) and \( p_1(x) \) in a power series for \( |x| < r \), where \( r \) is some radius greater than zero, we can get every solution of the equation in terms of a power series. A problem arises when we look at \( p_1(x) = 1/x \) and \( p_0(x) = 1 - \mu^2/x \). As we can see, a power series cannot be expanded about either \( p_0(x) \) or \( p_1(x) \) for any
radius of r, because |x|<r includes zero, which is an undefined quantity in both functions.

The problem encountered above may suggest that we cannot find a power series solution to the Bessel differential equation. However this is not the case because there is a theorem called Frobenius' Theorem, which will help us. It states the following.

**Theorem I**: Let \(xp_1(x)\) and \(x^2p_0(x)\) have power series expansions valid for \(|x|<r\). Then the equation \(y''+p_1(x)y'+p_0(x)y=0\) has a solution of the form

\[
y(x) = \sum_{n=0}^{\infty} a_n x^n + \sum_{n=0}^{\infty} a_n x^{n+s}, \quad a_0 \neq 0
\]

also valid for \(|x|<r\), where \(s\) is a constant.

Using this theorem and our \(p_0(x)\) and \(p_1(x)\), we get \(xp_1(x)=1\) and \(x^2p_0(x)=x^2-M^2\). We can see that the power series can be found valid for any \(x\). Therefore we can proceed with the assumption that \(y(x)=\sum_{n=0}^{\infty} a_n x^{s+n}\) is a form of the solution.

The equation we are working with, once again, is

\[
x^2 \frac{d^2 y}{dx^2} + x \frac{dy}{dx} + (x^2 - \lambda^2) y = 0.
\]

Since the solution is of the form \(y(x)=\sum_{n=0}^{\infty} a_n x^{s+n}\), then

\[
\frac{dy}{dx} = \sum_{n=0}^{\infty} a_n (s+n) x^{s+n-1} \quad \text{and} \quad \frac{d^2 y}{dx^2} = \sum_{n=0}^{\infty} a_n (s+n)(s+n-1)x^{s+n-2}.
\]

By placing \(y\) and its first two derivatives into the equation, we get

\[
x^2 \sum_{n=0}^{\infty} a_n (s+n)(s+n-1)x^{s+n-2} + x \sum_{n=0}^{\infty} a_n (s+n)x^{s+n-1} + (x^2 - \lambda^2) \sum_{n=0}^{\infty} a_n x^{s+n} = 0.
\]

To consolidate the summation of terms within one summation sign, we must find the first two terms of each of the first three sums. In doing so, we obtain the following results

\[
a_o (s(s-1)+s-\lambda^2)x^s + a_1 ((s+1)s+(s+1)-\lambda^2)x^{s+1} + \sum_{n=1}^{\infty} (((s+n)(s+n-1)+(s+n)-\lambda^2)a_n + a_{n-1})x^{s+n} = 0.
\]

Therefore we have the relations

\[
a_o (s(s-1)+s-\lambda^2)x^s = 0, \quad a_1 ((s+1)s+(s+1)-\lambda^2)x^{s+1} = 0 \quad \text{and} \quad (((s+n)(s+n-1)+(s+n)-\lambda^2)a_n + a_{n-1})x^{s+n} = 0.
\]
For this to be true, we must say that $x^{n+1}$, for $n=1, 2, \ldots$ cannot equal zero or it would be a trivial solution. We also know, from Frobenius' theorem, that $a_0 \neq 0$.

The first relation in the paragraph above is called the indicial equation and in the reduced form of $a(s^2-s+s-\mu^2) = a\{s^2-\mu^2\}$ it yields the roots $s=\mu$ and $s=-\mu$. The third relation can be reduced to $(s^n+s-n+s-n+\mu^2)a_n=-a_{n-2}$ or $a_n = \frac{-a_{n-2}}{(s+n+\mu)(s+n-\mu)}$ for $n=2, 3, \ldots$ which is called the recursion formula. This formula can be used to determine all of the coefficients, $a_n$, of the solution.

If we first consider the root $s=\mu$, the recursion formula gives us $a_n = \frac{-1}{n(2\mu+n)} a_{n-2}$. The even coefficients are, therefore, $a_2 = \frac{-1}{2(2\mu+2)} a_0$, $a_4 = \frac{-1}{4(2\mu+4)} a_2$ or in general, $a_{2k} = \frac{-1}{2^k k(\mu+k)} a_{2k-2}$. The odd coefficients are as follows $a_3 = \frac{-1}{(2\mu+3)3} a_1$, but we must get $a_1$ from the second equation $a_1((s+1)s+(s+1)-\mu^2)=0$. In reduced form we obtain $a_1((s+1)^2-\mu^2)=0$ and since $s=\mu$ we get $a_1(2\mu+1)=0$. For this to be possible, either $\mu=-1/2$ or $a_1=0$. We are going to need Bessel functions of order $\mu$ not equal to $-1/2$; therefore we will ignore that case. Then with $a_1=0$, all odd coefficients, which depend upon $a_1$, go to zero.

We found that $a_{2k} = \frac{-1}{2^k k(\mu+k)} a_{2k-2}$, which leads to $a_{2k-2} = \frac{-1}{2^k (\mu+k-1)(k-1)} a_{2k-4}$ and $a_{2k-4} = \frac{-1}{2^k (k-2)(\mu+k-2)} a_{2k-6}$. If this process was continued, we would obtain the form

$$a_{2k} = \frac{(-1)^k c_0}{(2^k)^k (k-1)! (\mu+k)^{(k+\mu-1)} \cdots (\mu+1)} \cdot \frac{(-1)^k c_0}{(2^k)^k (k+\mu)^{(k+\mu-1)} \cdots (\mu+1)}$$

For the sake of convenience, it is a common practice to standardize the solutions with $a_0 = \frac{1}{2^\mu / (1+\mu)}$. Therefore
we obtain a form of the coefficients for \( s = \mu \). That is to say, 
\[
\alpha_{2k} = \frac{(-1)^k}{2^{\mu+2k} k! \Gamma(\mu+1) \Gamma(\mu+2) \cdots \Gamma(\mu+k+1)}.
\]

The Gamma function \( \Gamma(n) \) is written as follows:
\[
\Gamma(n) = \int_0^\infty y^{n-1} e^{-y} \, dy.
\]
It can be integrated by parts to yield the formula
\[
\Gamma(n) = (n-1) \Gamma(n-1)
\]
and if \( n \) is an integer \( \Gamma(n) = (n-1)! \). Then we know that
\[
(\mu+1) \Gamma(\mu+1) = \Gamma(\mu+2) \quad \text{and} \quad (\mu+2) \Gamma(\mu+2) = (\mu+3).
\]
If we continue this operation several times we will obtain a new gamma function which, when placed into the coefficient formula will yield
\[
\alpha_{2k} = \frac{(-1)^k}{2^{\mu+2k} k! \Gamma(\mu+k+1)}.
\]

Now we have a solution \( y(x) \) for the root \( s = \mu \). It is written as follows using the new coefficient formula and \( n = 2k \) to gather just the even coefficients:
\[
y(x) = \sum_{k=0}^{\infty} \frac{(-1)^k x^{\mu+2k}}{2^{\mu+2k} k! \Gamma(\mu+k+1)}, \quad k = 0, 1, 2, \ldots.
\]
This solution \( y(x) \) is usually written in the form \( y(x) = J_\mu(x) \) which is called the Bessel function of order \( \mu \).

For the Bessel function to work as a solution to the Bessel differential equation it is required that we test its convergence. To do so we will use the ratio test. That is, 
\[
P = \lim_{k \to \infty} \left| \frac{U_{k+1}}{U_k} \right|
\]
is tested with \( U_k \) being the function in question. If \( P \) is greater than one, the function diverges and if it is less than one, it converges. When \( P \) is equal to one we must try another test. Let 
\[
U_k = \frac{(-1)^k x^{\mu+2k}}{k! \Gamma(\mu+k+1) 2^{\mu+2k}}
\]
and 
\[
U_{k+1} = \frac{(-1)^{k+1} x^{\mu+2k+2}}{(k+1)! \Gamma(\mu+k+2) 2^{\mu+2k+2}}.
\]
Then 
\[
P = \lim_{k \to \infty} \left| \frac{U_{k+1}}{U_k} \right| = \lim_{k \to \infty} \frac{|(-1)^{k+1} x^{\mu+2k+2} | (k+1)! \Gamma(\mu+k+2) 2^{\mu+2k+2}}{|(-1)^k x^{\mu+2k}| k! \Gamma(\mu+k+1) 2^{\mu+2k}}.
\]
After some cancellation we obtain 
\[
P = \lim_{k \to \infty} \frac{|x^2 \Gamma(\mu+k+1)|}{|x^2 \Gamma(\mu+k+2)|}.
\]
\[(n-1)!\Gamma(n-1)\), and so \(P = \lim_{k \to \infty} \frac{x^2}{4(n+1)(n+K+1)} \). We can factor out \(\frac{x^2}{4}\) and then \(P = \frac{x^2}{4} \lim_{k \to \infty} \frac{1}{(n+1)(n+K+1)} = 0\), which means that \(J_\mu(x)\) converges everywhere regardless of its order.

Considering the second root of the indicial equation, \(s=-\lambda\), we must look at the recursion formula again. Since \(s=-\lambda\), then \(a_n = \frac{-1}{n(n-2\lambda)} a_{n-2}\) where \(n \geq 2\), and \(a_1((s+1)^2 - \lambda^2) = 0\), which becomes \(a_1(1-2\lambda) = 0\). The recursion formula runs into difficulties when \(\mu\) is of the form of an integer divided by two. We must note here that an even integer divided by two is an integer and an odd integer divided by two is called a half-integer. The problem, obviously, is that \(a_n\) is undefined for this kind of \(\mu\). Therefore we will first consider the case where \(\mu\) is not an integer or a half-integer. Because of this assumption, \(a_1 = 0\) and all the odd coefficients, which depend on \(a_1\), are zero. The only difference in the coefficient formula is that we use \(-\lambda\) instead of \(\lambda\). Hence, the solution is different only in the fact that we have \(-\lambda\). That is to say that the solution is \(J_{-\lambda} (x) = \sum_{k=0}^{\infty} \frac{(-1)^k x^{L_k-\lambda}}{k! \Gamma(k+\lambda+1) 2^{2k-\lambda}}\), and it can be shown to converge for all \(x\). The two solutions are linearly independent because neither one is a multiple of the other. Therefore the total solution of the Bessel differential equation for \(\mu\), assuming that \(\mu\) is not an integer or a half-integer, is \(y(x) = A_1 J_\mu(x) + A_2 J_{-\mu}(x)\).

There is, however, a theorem that can be used to help us analyze the problem where \(\mu\) is an integer or a half-integer. The theorem is as follows.
Theorem II: When the roots $p_1$ and $p_2$ of the indicial equation are distinct and do not differ by an integer, the method of Theorem I yields two linearly independent solutions. If the roots differ by an integer, a second solution can be found by assuming that

$$y_2(x) = C y_1(x) \log(x) + \sum_{n=0}^{\infty} a_n x^{n+1} p_2$$

where $y_1(x)$ is the solution given by Theorem I for the root $p=p_1$.

According to this theorem, we see that when $\mu$ is an integer or a half-integer, the difference between the roots $p_1=\mu$ and $p_2=-\mu$ is always going to be an integer. Therefore the second solution is going to be of the form

$$y_2(x) = C y_1(x) \log(x) + \sum_{n=0}^{\infty} a_n x^{n-\mu}.$$ 

We have come up with two solutions for the Bessel differential equation of the form $y_1(x)=A_1 J_{\mu}(x)+A_2 J_{-\mu}(x)$ and $y_2(x)=C y_1(x) \log(x) + \sum_{n=0}^{\infty} a_n x^{n-\mu}$.

When we first started work on the Bessel differential equation, we made the substitution $x=\lambda r$. Therefore the solutions to the Bessel differential equation are $R_1(r)=A_1 J_{\mu}(\lambda r)+A_2 J_{-\mu}(\lambda r)$ and $R_2(r)=C R_1(r) \log(\lambda r) + \sum_{n=0}^{\infty} c_n (\lambda r)^{n-\mu}$.

b) Derivative and Integration Formula of the Bessel Function

In the last section we derived the Bessel function of order $\mu$ such that

$$J_{\mu}(r) = \sum_{k=0}^{\infty} \frac{(-1)^k}{2^{\mu+2k} k! (\mu+k+1)} r^{\mu+2k}.$$ 

For $\mu=n$, where $n$ is an integer, we will develop the derivative and the integration formula. We will find that these two developments will assist us in the analysis of Bessel functions.
First of all, from previous knowledge of the Gamma function, we can say \( \Gamma(n) = (n-1)! \) when \( n \) is an integer.

This leads to a modified version of the Bessel function:
\[
J_n(x) = \sum_{k=0}^{\infty} \frac{(-1)^k}{k!(n+k)!} \left( \frac{x}{2} \right)^{n+2k}.
\]

Because the sum is in terms of \( k \), we can factor out the term \( \left( \frac{x}{2} \right)^n \) and multiply the result with \( x^{-n} \) to yield
\[
J_n(x) = \frac{1}{2^n} \sum_{k=0}^{\infty} \frac{(-1)^k}{k!(n+k)!} \left( \frac{x}{2} \right)^{2k}.
\]

The reason we multiplied the Bessel function by \( x^{-n} \) is that it gives us an easier form to differentiate. In doing so the derivative form is
\[
\frac{d}{dx} \left[ x^{-n} J_n(x) \right] = \frac{1}{2^n} \sum_{k=0}^{\infty} \frac{(-1)^k}{k!(n+k)!} \left( \frac{x}{2} \right)^{2k},
\]

This, of course, looks familiar. So with the use of our original definition for a Bessel function, the derivative form is
\[
\frac{d}{dx} \left[ x^{-n} J_n(x) \right] = -x^{-n} J_{n+1}(x).
\]

A similar method can be used to get the other derivative form
\[
\frac{d}{dx} \left[ x^n J_n(x) \right] = x^n J_{n-1}(x).
\]

Of course, if we integrate this expression from zero to some radius \( r \), we obtain the integration formula, namely:
\[
\int_0^r \frac{x^n J_n(x)}{r^n J_n(r)} dx = \int_0^r x^n J_n(x) dx = \int_0^r y^n J_n(y) dy.
\]

Use will be made of this formula later.
c) The Sturm-Liouville Problem

Many times, when using the separation-of-variables technique, we come across a homogeneous equation of the form $X''(x) + R(x)X'(x) + (Q(x) + \lambda P(x))X(x) = 0$ with the boundary conditions $a_1X(a) + a_2X'(a) = 0$ and $b_1X(b) + b_2X'(b) = 0$. This is called a Sturm-Liouville problem over the interval from $a$ to $b$.\(^\text{16}\)

Multiplying the Sturm-Liouville equation by the integrating factor $r(x) = e^{\int P(x) \, dx}$ yields $d\left[\frac{d}{dx} r(x) X(x)\right] + \left[q(x) + \lambda r(x)\right] X(x) = 0$. It has been found that there exist many solutions of that equation that depend on certain values of the variable $\lambda$. These values, namely $\lambda_i$ for $i=1,2,\ldots,$ are called eigenvalues and each $\lambda_i$ has a solution $X_i(x)$ identified with it, which is called an eigenfunction. Each eigenfunction, by way of definition, is a solution to the linear homogeneous differential equation used in the Sturm-Liouville problem. The principle of superposition states that if the functions $X_1, X_2, \ldots, X_n$ are solutions of that homogeneous equation, the sum $X = c_1X_1 + c_2X_2 + \cdots + c_nX_n$ is a solution.\(^\text{17}\) Also if each $X_i$ is a solution of some linear homogeneous boundary condition, the sum of $X_i$'s is a solution of that boundary condition. This property will allow us to include all the solutions in our final analysis.

Eigenvalues and eigenfunctions are important in the study of Bessel functions because they lead to the important concept of orthogonality, which we will develop here.
There are several things we must require in order to develop orthogonality. The functions \( p(x) \), \( q(x) \), \( r(x) \) and \( r'(x) \) must be real valued and continuous on the closed interval from \( a \) to \( b \). We must have \( p(x) \) and \( r(x) \) greater than zero over the open interval, as well as the condition that the constants \( a, b, a_2, b_2 \) are real valued and independent of \( \lambda \).

If we have two eigenvalues, \( \lambda_n \) and \( \lambda_m \), and their corresponding eigenfunctions, \( X_n \) and \( X_m \), of the Sturm-Liouville equation, we know the following: \((rX_n')' + (q + \lambda_n p)X_n = 0\) and \((rX_m')' + (q + \lambda_m p)X = 0\). When we multiply the first equation by \( X_m \) and the second by \( X_n \), and take the difference of the two, we obtain \((rX_n')X_m + \lambda_n pX_n X_m - (rX_m')X_n - \lambda_m pX_m X_n = 0\). By removing the function \( p \) to the right side of the equation, the form \((rX_n')X_m - (rX_m')X_n = (\lambda_m - \lambda_n) p X_n X_m \) is obtained. If we reverse the chain rule of the derivative, we can show that

\[
\frac{d}{dx} \left[ (rX_n')X_m - (rX_m')X_n \right] = (rX_n')(X_m' X_n) - (rX_m')(X_n' X_n) = (\lambda_m - \lambda_n) p X_n X_m.
\]

If we were to integrate this expression in terms of \( dx \), over the interval from \( a \) to \( b \), we have

\[
\int_{a}^{b} (\lambda_m - \lambda_n) p X_m X_n = (\lambda_m - \lambda_n) \int_{a}^{b} p X_m X_n.
\]

Applying the first boundary condition to the eigenfunctions \( X_m \) and \( X_n \), we obtain a system of equations \( a_1 X_n(a) + a_2 X_n'(a) = 0 \) and \( a_1 X_m(a) + a_2 X_m'(a) = 0 \). In order for there to be a non-trivial solution for \( a_1 \) and \( a_2 \), the coefficient determinant must be zero. That is to say \((X_n(a)X_m(a) - X_n'(a)X_m(a)) = 0\).

We can also show that this is true for the boundary condition at \( b \). Therefore when we evaluate the left hand side of the equation above, we obtain \( r(b)(X_n(b)X_m(b) - X_n'(b)X_m(b)) \)
\(-r(a)(X_n'(a)X_n(a) - X_m'(a)X_m(a))\), which, of course, equals zero. This leaves the form \((\lambda_m - \lambda_n)\int_a^b \rho X_n X_m d\chi = 0\). Since we assumed that \(\lambda_m\) and \(\lambda_n\) were two different eigenvalues, we must have \(\int_a^b \rho X_n X_m d\chi = 0\). This describes the orthogonality of two eigenfunctions with respect to what is called the weight function \(\rho\). There are other cases where the orthogonality holds. When \(r(a)\) or \(r(b)\) equal zero, the corresponding boundary conditions are dropped, yielding a single-bounded Sturm-Liouville problem. Also when \(r(a)=r(b)\) and there are periodic boundaries \(X(a)=X(b)\) and \(X'(a)=X'(b)\), then \(r(a)\ det\ (a)=r(b)\ det\ (b)\), and so orthogonality holds. As we will see later, the orthogonality of eigenfunctions is an important concept in the analysis of Bessel functions.

d) Orthogonal Sets of Functions

It is often useful to describe functions as vectors. This enables applying the properties we know about vectors to functions. Assuming we know what a vector space is, we can describe anything in the space in terms of a family of vectors that span the space. It is customary to work with a group of perpendicular vectors because they can describe any other vectors in the space. This property of vectors being perpendicular is called the orthogonality of vectors. If two vectors \(\vec{V}_1(\chi)\) and \(\vec{V}_2(\chi)\) are orthogonal, they satisfy the inner product relation \(\left( \vec{V}_1(\chi), \vec{V}_2(\chi) \right) = \int_a^b \vec{V}_1(\chi) \vec{V}_2(\chi) d\chi = 0\). It is important to note that the
units of measurement on the axes may be different. Therefore we must introduce a function \( p(x) \geq 0 \), called the weight function, that will insure orthogonality. The resulting relation \( \int_a^b p(x) \overrightarrow{v_1}(x) \overrightarrow{v_2}(x) \, dx = 0 \) is the orthogonality of \( \overrightarrow{v_1} \) and \( \overrightarrow{v_2} \) with respect to the weight function \( p(x) \), which can be a constant, as it is above.

It is customary to use unit vectors because all vectors can be described as a scalar multiple of some unit vector. Since a vector has magnitude and direction, we can obtain a unit vector in any direction by dividing the vector by its magnitude. The magnitude of a vector is called the norm of the vector and it is described as follows. The norm of a vector \( \overrightarrow{V} \) is written \( \| \overrightarrow{V} \| \), and it is defined by the relation \( \| \overrightarrow{V} \| = \left( \sum_{i=1}^{n} V_i^2 \right)^{\frac{1}{2}} \). Therefore the unit vector \( \hat{V} \) is defined in a form such that \( \hat{V} = \overrightarrow{V}/\| \overrightarrow{V} \| \). It is important to note that the inner product of an orthogonal unit vector with itself is one; that is, \( (\hat{V}, \hat{V}) = \int_a^b (\hat{V})^2 \, dx = 1 \).

When we group these orthogonal unit vectors together, we obtain a set which is called an orthonormal set, \( \{ \overrightarrow{V}_n \} \). It is with this set that we can describe functions in our vector space.

e) Fourier Series

If we have an orthonormal set \( \{ \phi_n \} \) on some interval \((a,b)\), every vector in the space defined by that set can be expressed as a linear combination of the vectors \( \phi_n \). Therefore if we have some function \( f(x) \), there is a chance
that it can be expressed as a linear combination of the vectors $\phi_n$. This form of the function, \( f(x) = c_1\phi_1 + \cdots + c_n\phi_n \), can frequently be written as an infinite series. If this series representation of \( f(x) \) converges, we may use some properties of vectors to obtain the constants \( c_n \).

By multiplying the function, \( f(x) \), by some orthogonal vector, \( \phi_k \), from the orthonormal set, we obtain \( \phi_k f(x) = c_1\phi_k + c_2\phi_k + \cdots \), which, when integrated over the interval \((a, b)\), yields
\[
\int_a^b f(x) \phi_k \, dx = \sum_{m=1}^{\infty} c_m \int_a^b \phi_m \phi_k \, dx.
\]
This relation will give us the constants because it is a representation of the orthogonality property
\[
\int_a^b \phi_m \phi_k \, dx = \delta_{m,k} \quad \text{and} \quad \delta_{m,K} = \sum_{m=1}^{\infty} c_m \phi_k \int_a^b \phi_k \phi_m \, dx.
\]
Therefore \( \int_a^b \phi_k f(x) \, dx = c_k \), which indeed are the required constants, and which are called the Fourier constants for the function \( f(x) \). Hence the function \( f(x) \) can be written
\[
\sum_{k=1}^{\infty} c_k \phi_k = \sum_{k=1}^{\infty} \phi_k \int_a^b f(y) \phi_k (y) \, dy.
\]
This, of course, is provided that the orthonormal set is appropriate for the function, \( f(x) \). That is to say, we do not want the \( f(x) \) to be orthogonal to each of the vectors from the orthonormal set. Also, we want the function and the orthonormal set to describe the same space.

A very important consideration in describing a function in terms of an orthonormal set is the activity of the function itself. So far we have assumed that the function is continuous on the interval \([a, b]\). There is another family of functions, though, which can be described as a linear combination of some orthonormal set. That family is the group of sectionally continuous functions. A function on an interval \([a, b]\), such that it has a finite num-
ber of jump-discontinuities, $x_1, x_2, \ldots, x_n$, can be broken into smaller intervals of the form $(a, x_1), (x_1, x_2), \ldots, (x_{n-1}, x_n), (x_n, b)$. Then the function is continuous on each interval $(x_j, x_{j+1})$. The end points of each interval, $x_j$ where $i=1, 2, \ldots, n$, are not defined by the function. But as long as there is a limit to the function, when approaching the end points from within the interval, the function is called a sectionally continuous function. If the function and its first derivative are sectionally continuous over an interval, we can describe it in terms of the orthonormal set, with the value of the representation converging to $(1/2)(f(x_1^-0)+f(x_1^+0))$ at jump-discontinuities, $x_j$.

f) Orthonormal Families and Fourier-Bessel Series

In considering the single-bounded Sturm-Liouville equation, on the interval $(0, c)$ with $\lambda^j$ as the eigenvalue parameter instead of $\lambda$, $p(x)=r(x)=x$ and $q(x)=-n^2/x$, we get the relation $(xx')'+(\lambda^2 x-n^2/x)x=0$. According to the conditions placed on the equation, this resembles the Bessel differential equation with $J_n(\lambda x)$ as a solution. If we multiply by the integrating factor $2xx'$, we get the equation $2xx'(xx')+2(\lambda^2 x^2-n^2)xX'=0$, which can also be written in the form $\frac{d}{dx}(\lambda x^2) + (\lambda^2 x^2-n^2) \frac{d}{dx}(x^2) = 0$.

Multiplying the equation by the differential $dx$, we can integrate the result over the interval $(0, c)$: $\int_0^c d[(\lambda x^2)] = \int_0^c (n^2-\lambda^2 x^2) d(x^2)$. This process can be completed by ordinary procedures and by using integration by parts where
it is needed. When we have completed the integration, we will get the relation 
\[ \left[ (\chi x')^2 + (\lambda^2 x^2 - n^2) X^2 \right]_{x=0}^{x=c} = 2 \lambda^2 \int_0^c x^2 \, dx. \]

It can be seen that the left hand side of the equation evaluated at zero is zero. The solution of the differential equation is \( J_n(\lambda x) \) where \( \frac{d}{dx} (J_n(\lambda x)) = \lambda J_n'(\lambda x). \) Therefore we have \( (c\lambda J_n'(c\lambda))^2 + (\lambda^2 c^2 - n^2)(J_n(\lambda c))^2 = 2 \lambda^2 \int_0^c \left[ J_n(\lambda x) \right]^2 \, dx \) or \( \int_0^c \left[ J_n(\lambda x) \right]^2 \, dx = \frac{c^2}{2} \left[ J_n'(\lambda c) \right]^2 + \left( \frac{c^2}{2} - \frac{n^2}{2\lambda^2} \right) \left[ J_n(\lambda c) \right]^2. \)

Looking back at the definition of the norm, we can see that the left side of the equation is the square of the formula for the norm of \( J_n(\lambda x) \). Since we know \( \lambda \) has many eigenvalues, \( \lambda_j \) where \( j=1,2,\ldots \), then \( J_n(\lambda_j x) \) are the eigenfunctions of the Sturm-Liouville problem.

The boundary condition for this simple-bounded Sturm-Liouville problem is \( a_1 X(c) + a_2 X'(c) = 0 \). We said that \( X(x) = J_n(\lambda x) \), which leads to the new condition \( a_1 J_n(\lambda c) + a_2 (J_n(\lambda c))' = 0 \) or \( a_1 J_n(\lambda c) + a_2 \lambda J_n'(\lambda c) = 0. \) Of course we could have \( a_1 = 0, a_2 = 0 \) or \( a_1, a_2 \) not equal to zero. If \( a_1 \) and \( a_2 \) are not zero then the eigenvalues, \( \lambda_j \), are the positive roots of the equation \( a_1 J_n(\lambda c) + a_2 \lambda J_n'(\lambda c) = 0. \) For the case where \( a_2 = 0 \), the condition is \( J_n(\lambda c) = 0 \), where the eigenvalues are the positive roots of the equation. When \( a_1 = 0 \), we have the condition that \( J_n'(\lambda c) = 0. \) Then the eigenvalues, \( \lambda_j \), are the positive roots of the equation \( J_n'(\lambda c) = 0. \) We must note that the negative roots in all the cases yield the eigenfunction corresponding to the positive root times a constant.\(^2\)

The first case above gives us the relation \( a_1 J_n(\lambda c) + a_2 \lambda J_n'(\lambda c) = 0. \) Hence the corresponding norm formula is as follows: \[ \| J_n(\lambda_j c) \|^2 = \left( \frac{\lambda_j^2 c^2 - n^2 + a_2^2}{2 \lambda_j} \right) \left[ J_n'(\lambda_j c) \right]^2. \]
In the second case we know that the eigenvalues are the positive roots of the equation $J_n(\lambda c) = 0$. The formula we obtained for the square of the norm here is $||J_n(\lambda_j c)||^2 = \frac{c^2}{2} \left[ J'_n(\lambda_j c) \right]^2 + \left( \frac{n^2 - \lambda_j^2}{2m_a} \right) \left[ J_n(\lambda_j c) \right]^2$ or $||J_n(\lambda_j c)||^2 = \frac{c^2}{2} \left[ J'_n(\lambda_j c) \right]^2$.

In a previous section we determined the derivative formula $\frac{d}{dx} \left( x^n J_n(x) \right) = x^n J_{n-1}(x)$, which can be reduced to the relation $J'_n(x) = \left( \frac{n^2 - \lambda_j^2}{2m_a} \right) J_n(x)$. Since $J_n(\lambda_j c) = 0$, the relation gives us a new way to determine the norm. That is to say, the norm is of the form $||J_n(\lambda_j c)||^2 = \frac{c^2}{2} \left[ J'_{n+1}(\lambda_j c) \right]^2$. The third case is similar to the first case because it has the same formula for the norm as the first case, except that $\lambda = 0$.

Now that we have the formulas for the norms, we can create an orthonormal set of the Bessel functions. This is necessary so that we can define some function in terms of the Bessel functions. It is common to signify normalized eigenfunctions with the symbol $\Phi$. Therefore with what we know about normalization of a vector, the set of normalized eigenfunctions is $\Phi_{n_j}(x) = J_n(\lambda_j x) / ||J_n(\lambda_j c)||$ where $j = 1, 2, ..., n$ is the order of the Bessel function.

According to the information from the section on the Sturm-Liouville problem, the orthonormal set, $\{ \Phi_{n_j}(x) \}$ obeys the relation $\int_0^c \chi \Phi_{n_j}(x) \Phi_{n_k}(x) \ d\chi = \sum_{i \neq k} \delta_{ij} \delta_{jk}$. We know that since $\Phi_{n_j}$, as they are defined, are orthogonal, they can represent a properly characterized function on the interval $(0, c)$. From our knowledge of Fourier series, we know that the correct function, $f(x)$, can be represented by a linear combination of the members of the
orthonormal set. That is, $f(x) = \sum_{j=1}^{\infty} c_{nj} \phi_{nj}(x)$
where the constants are of the form $c_{nj} = \int_{o}^{c} \phi_{nj}(x) f(x) dx$.

The orthonormal functions have been defined as $\phi_{nj}(x) = J_n(\lambda_j x)\|J_n(\lambda_j x)\|$, leading to a new form of the constants
$c_{nj} = \int_{o}^{c} \frac{J_n(\lambda_j x)f(x)}{\|J_n(\lambda_j x)\|^2} dx$ or $c_{nj} = \frac{1}{\|J_n(\lambda_j x)\|^2} \int_{o}^{c} \frac{J_n(\lambda_j x)f(x)}{J_n(\lambda_j x)} dx$ because the norm yields a constant. Therefore the function can be written as
$f(x) = \sum_{j=1}^{\infty} \frac{J_n(\lambda_j x)}{\|J_n(\lambda_j x)\|^2} \int_{o}^{c} J_n(\lambda_j x)f(x) dx$.

The first and third case of functions were those in which the eigenvalues are the positive roots of the equation $a_1 J_n(\lambda c) + a_2 \lambda J'_n(\lambda c) = 0$. After applying the formula of the norm, the constants have the form
$c_{nj} = \frac{2 \lambda_j}{(\lambda_j^2 - \eta^2 + a_2^2) [J_n(\lambda_j c)]^2} \int_{o}^{c} \frac{J_n(\lambda_j x)f(x)}{J_n(\lambda_j c)} dx$.

Of course in case three, $a_1 = 0$, and the constant formula is changed accordingly. The second case of functions we have are those in which the eigenvalues are the positive roots of the equation $J_n(\lambda c) = 0$. For this, then, the constants are of the form
$c_{nj} = \frac{2}{\lambda_j^2 [J_m(\lambda_j c)]^2} \int_{o}^{c} \frac{J_n(\lambda_j x)f(x)}{J_n(\lambda_j c)} dx$.

It has been found that the representation of a function, which conforms to specifications, will indeed converge. This form of solution, called a Fourier-Bessel series representation of the function $f(x)$, will be used to solve a problem in heat conduction.
4) Solution of the Selected Problem
   
   a) Definition of the Problem

   The problem I have chosen is a heat distribution problem in the cylindrical system. The problem consists of a solid rod of infinite length made up of some isotropic material. Slipped over this rod is a sleeve of infinite length made up of the same material. The rod has a radius of one, while the sleeve has an inner radius of one and an outer radius of two. The initial temperature of the rod is $A$. $B$ is the initial temperature of the sleeve and also the temperature at the outer boundary of the sleeve. $A$ and $B$, in this problem, are constants.

   The heat equation for the cylindrical system is
   \[ \frac{1}{k} \frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial r^2} + \frac{1}{r} \frac{\partial u}{\partial r} + \frac{1}{r^2} \frac{\partial^2 u}{\partial \theta^2} + \frac{\partial^2 u}{\partial z^2}. \]

   Since I have chosen to do a problem with cylinders of infinite length, I can eliminate $z$ from the heat distribution. The parameter $\theta$ can also be eliminated because of symmetry, yielding a distribution, $U(r,t)$. Therefore the heat equation is reduced to
   \[ \frac{1}{k} \frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial r^2} + \frac{1}{r} \frac{\partial u}{\partial r}. \]

   b) General Solution

   The problem I have solved has the following equation and conditions:
   \[ \frac{1}{k} \frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial r^2} + \frac{1}{r} \frac{\partial u}{\partial r} \quad 0 < r < 2, \quad t > 0 \]
   \[ U(2,t) = B \quad t > 0 \]
   \[ U(r,0) = \begin{cases} A, & r < 1 \\ B, & 1 \leq r \leq 2 \end{cases} \]
In order to use the methods that I have described in previous sections, I have to have a homogeneous boundary condition. Therefore I will use a technique that is commonly used to change a non-homogeneous boundary condition to a homogeneous one. The heat distribution \( U(r,t) \) can be described as the sum of a transient solution, \( V(r,t) \), and a steady state solution, \( W(r) \). That is, \( U(r,t) = V(r,t) + W(r) \), which means that \( U \) can be described in terms of \( V \) and \( W \) in the heat equation. First of all \( \frac{\partial^2 U}{\partial r^2} + \frac{1}{r} \frac{\partial U}{\partial r} = \frac{\partial^2 (V + W)}{\partial r^2} + \frac{1}{r} \frac{\partial (V + W)}{\partial r} = \frac{1}{K} \left( \frac{\partial W}{\partial t} + \frac{\partial V}{\partial t} \right) \). The relation thus becomes \( \frac{\partial^2 V}{\partial r^2} + \frac{1}{r} \frac{\partial V}{\partial r} = \frac{1}{K} \frac{\partial W}{\partial t} \) and \( \frac{\partial^2 W}{\partial r^2} + \frac{1}{r} \frac{\partial W}{\partial r} = 0 \).

Looking at the steady state solution, \( W \), the heat equation is \( \frac{\partial^2 W}{\partial r^2} + \frac{1}{r} \frac{\partial W}{\partial r} = 0 \) or \( W'' + \left( 1/r \right) W' = 0 \). I can solve this by using a reduction of order technique. That is, I will let \( S = W' \). Then the equation becomes \( S' + \left( 1/r \right) S = 0 \), which has the solution \( S = (c_1/r) \), where \( c_1 \) is a constant. With the substitution replaced, I know that \( W'(r) = (c_1/r) \) and therefore \( W(r) = c_1 \ln(r) + c_2 \), where \( c_1 \) and \( c_2 \) are constants. The object of solving this problem is to get a solution which is valid at all places in the cylinder. Since the logarithm function is undefined at zero, \( c_1 \) must be zero, leaving \( W(r) = c_2 \). I also know that the function, \( W \), must conform to the boundary condition; therefore, \( W(2) = B \) or \( c_2 = B \).

The transient solution, \( V(r,t) \), has the heat equation \( \frac{\partial^2 V}{\partial r^2} + \frac{1}{r} \frac{\partial V}{\partial r} = \frac{1}{K} \frac{\partial V}{\partial t} \). The function \( V \) can
be described in terms of \( U \) and \( W \), such that \( V(r,t) = U(r,t) - W(r) \). Therefore the initial condition of \( V \) becomes \( V(r,0) = U(r,0) - W(r) = f(r) - W(r) \), where \( f(r) = \sum_{i \leq 1} A_i \), \( i \leq r \leq 2 \). The boundary condition of \( V \) becomes \( V(2,t) = U(2,t) - W(2) = B - B = 0 \), which is a homogeneous boundary condition. To solve this equation I will use the method of separation of variables. I can let \( V(r,t) = R(r)T(t) \), which leads to \( R''(r) + (1/r)R'(r) - sR(r) = 0 \) and \( T'(t) - k\lambda^2 T(t) = 0 \).

To solve the equation of \( T \), there are three cases, \( s = 0 \), \( s > 0 \) and \( s < 0 \). The first case \( s = 0 \) gives the equation \( T'(t) = 0 \), which means \( T(t) = c \), where \( c \) is a constant. For \( s > 0 \) the equation becomes \( T'(t) - k\lambda^2 T(t) = 0 \), where \( s = \lambda^2 \). This equation has the solution \( T(t) = c_3 \exp(k\lambda^2 t) \), where \( k \) and \( c_3 \) are constants. The positive exponential function is not bounded above; therefore \( s \) cannot be greater than zero. The final case is for \( s < 0 \), which can be written as \( s = -\lambda^2 \), which yields the equation \( T'(t) + k\lambda^2 T(t) = 0 \). This equation has the solution \( T(t) = c_4 \exp(-k\lambda^2 t) \).

The equation in terms of \( R \) is as follows: \( R''(r) + (1/r)R'(r) - sR(r) = 0 \). There are two possibilities for \( s \), due to what I found when solving the \( T \) equation: namely, \( s = 0 \) and \( s < 0 \). When \( s = 0 \) the equation becomes \( R''(r) + (1/r)R'(r) = 0 \), with the solution \( R(r) = b_1 \ln(r) + b_2 \). The behavior of the logarithm function dictates the form \( R(r) = b_2 \). The second case, \( s < 0 \), gives me the equation \( R''(r) + (1/r)R'(r) + \lambda^2 R(r) = 0 \).
From Bessel function theory I know that this equation has the solution $J_0(\lambda r)$. The boundary condition on the function $V$, is such that $V(2,t)=R(2)T(t)=0$. Since $T$ is not zero for all $t$, the eigenvalues for $R$ must be the solutions of the equation $R(2)$. That is to say, the eigenvalues, $\lambda_j$, are the positive roots of the equation $J_0(2\lambda)=0$. The eigenfunctions for $R(r)$ and $T(t)$ are $J_0(\lambda_j r)$ and $\exp(-k\lambda_j^2 t)$ respectively.

For the case $s=0$, I obtained the solutions $T(t)=c$ and $R(r)=b_2$, which leads to $V(r,t)=cb_2$. This case can be dropped because it is the steady state solution, which is already taken care of in the function $W(r)$. When $s<0$ the solutions are $T(t)=c_4\exp(-k\lambda_j^2 t)$ and $R(r)=J_0(\lambda_j r)$ or $V(r,t)=R(r)T(t)=J_0(\lambda_j r)\exp(-k\lambda_j^2 t)$. The solution $V$ is a linear combination of the eigenfunctions, therefore each eigenfunction, $V_j$, has a constant, $c_j$, identified with it. That is, $V(r,t)=\sum_{j=1}^{\infty} c_j J_0(\lambda_j r) \exp(-k\lambda_j^2 t)$.

From the theory of Fourier-Bessel series, I know that the constants are of the form $c_j = \frac{1}{2[\beta_j(2\lambda_j)]^2} \int_0^2 r J_0(\lambda_j r) f_1(r) \, dr$. The function $f_1(r)$, which is the initial condition of $V$, was described to be $f(r)-W(r)$, where $f(r)=\begin{cases} A, & r \leq 1 \\ B, & 1 < r \leq 2 \end{cases}$ and $W(r)=B$. Therefore the integral can be expressed as $\int_0^2 r J_0(\lambda_j r) [f(r)-B] \, dr = \int_0^1 r J_0(\lambda_j r)(A-B) \, dr + \int_1^2 r J_0(\lambda_j r)(B-B) \, dr$. The original integral thus becomes $(A-B)\int_0^1 J_0(\lambda_j r) \, dr$.

The integration formula for the Bessel function gives me the value $(A-B)\frac{J_1(\lambda_j)}{\lambda_j}$. This leads to the new formula for the constants: namely, $c_j = \frac{(A-B)}{2\lambda_j} \frac{J_1(\lambda_j)}{[\beta_j(2\lambda_j)]^2}$, which means the solution is of the form $V(r,t)=\sum_{j=1}^{\infty} c_j J_0(\lambda_j r) \exp(-k\lambda_j^2 t)$.
I described the heat distribution over the cylinder of radius two to be \( U(r,t) = W(r) + V(r,t) \). Therefore the heat distribution of my problem is as follows:

\[
U(r,t) = B + \sum_{j=1}^{\infty} \left( \frac{(A-B)}{2\lambda_j} \frac{J_1(\lambda_j r)}{(J_1(2\lambda_j))^2} \right) J_0(\lambda_j r) e^{\exp(-k\lambda_j^2 t)}
\]

where the eigenvalues, \( \lambda_j \), are the positive roots of the equation \( J_0(2\lambda) = 0 \). This type of representation has been used often in mathematical physics. Since it is convergent over the interval I have used, I am assured that it describes my problem well.

The zeros, \( \alpha_j \) where \( j = 1, 2, \ldots \), of the Bessel function of order zero are defined to be such that \( J_0(\alpha_j) = 0 \). These zeros have been obtained by numerical methods of approximation on a computer. The following is a list of the first twenty zeros of the Bessel function of order zero:

<table>
<thead>
<tr>
<th>( \alpha_j )</th>
<th>( J_0(\alpha_j) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.40482 ( 55577 )</td>
</tr>
<tr>
<td>2</td>
<td>5.52007 ( 81103 )</td>
</tr>
<tr>
<td>3</td>
<td>8.65372 ( 79129 )</td>
</tr>
<tr>
<td>4</td>
<td>11.79153 ( 44391 )</td>
</tr>
<tr>
<td>5</td>
<td>14.93091 ( 77086 )</td>
</tr>
<tr>
<td>6</td>
<td>18.07106 ( 39679 )</td>
</tr>
<tr>
<td>7</td>
<td>21.21163 ( 66299 )</td>
</tr>
<tr>
<td>8</td>
<td>24.35247 ( 15308 )</td>
</tr>
<tr>
<td>9</td>
<td>27.49347 ( 91320 )</td>
</tr>
<tr>
<td>10</td>
<td>30.63460 ( 64684 )</td>
</tr>
<tr>
<td>11</td>
<td>33.77582 ( 02136 )</td>
</tr>
<tr>
<td>12</td>
<td>36.91709 ( 83537 )</td>
</tr>
<tr>
<td>13</td>
<td>40.05842 ( 57646 )</td>
</tr>
<tr>
<td>14</td>
<td>43.19979 ( 17132 )</td>
</tr>
<tr>
<td>15</td>
<td>46.34118 ( 83717 )</td>
</tr>
<tr>
<td>16</td>
<td>49.48260 ( 98974 )</td>
</tr>
<tr>
<td>17</td>
<td>52.62405 ( 18411 )</td>
</tr>
<tr>
<td>18</td>
<td>55.76551 ( 07550 )</td>
</tr>
<tr>
<td>19</td>
<td>58.90698 ( 39261 )</td>
</tr>
<tr>
<td>20</td>
<td>62.04846 ( 91902 )</td>
</tr>
</tbody>
</table>

I know that the eigenvalues of my problem can be obtained from the table because \( 2\lambda_j = \alpha_j \).
The solution I obtained is sufficiently general to enable describing many problems of a similar nature. To solve a certain problem, only the material of the rod and sleeve and the temperatures A and B need to be defined. In this way, the effect of an unstable heat distribution in a cylindrical solid of a particular material can be studied.
5) Extensions of the Model

The model I used to solve my problem is normally considered a representation of the problem of shrunken-fittings. That is, generally the rod signifies a rivet, and the sleeve represents the surrounding piece of apparatus which will receive the rivet. In this procedure the temperature of the rod is much lower than the sleeve. Then the cooled rod is slipped into the sleeve with very little room to spare. When the rod warms up it creates a tight seal with the sleeve, one which cannot be loosened. Below are illustrated some of the assumptions used to form the mathematical model along with changes which can be made to extend and refine the application of this model.

An assumption was made involving the use of an infinite cylindrical system. This supposition led to a solution with two variables, r and t. Indeed, most of the time the physical problem is in a finite system. Therefore I would have to include θ and z in my work. This yields four ordinary differential equations, as described in an earlier section. The solution to this type of problem would be much more involved, but it does give a better approximation to the situation.

The second assumption used referred to the thermal conductivity, k, and it was in two parts. The first part required that the temperature difference between the rod and the sleeve be small. This may not always be the case. Therefore a more general solution would involve k as a
function of different values over the range of temperatures in the problem. The second part of this assumption involved the use of similar material in the rod and sleeve. Therefore, the solution of the problem of materials of different k's would involve separate solutions to each with a joining condition used at the common boundary.

The state of the materials was a third assumption used. It was necessary to solve a problem involving solids, but the physical problem might involve a heat gradient that would melt one of the solids. This would create many problems that couldn't be handled by the methods used. Therefore the analysis would have to be extended to include this possibility.

A final assumption made was to solve only the problem of heat conduction. There are two other modes of heat transfer that can be studied: namely, convection and radiation. Convection occurs when air or liquid is passed over an object and heat is acquired by either the object or the liquid. Radiation of heat is a process of heat emanating from a hot body into the surroundings. Therefore to form a better solution of the physical problem, these factors, in the form of boundary conditions, might make a contribution to this analysis.

In conclusion, this study is not complete because it doesn't include all possible cases. The background material needed to explore the mathematical physics problem chosen is complete. Many of the references used contain
material on the other types of problems mentioned above. Therefore, anyone interested in these problems should consult this list of books on the subject of mathematical physics.
References


2. Ibid., p. 552.


4. Ibid., p. 415.


10. Ibid., p. 236.


21. Ibid., pp. 54-5.

22. Ibid., p. 92.

23. Ibid., p. 183.


25. Ibid., pp. 455-6.